IBM Z

Processor Resource/Systems Manager Planning Guide

Level 01a

IBM

Note:

Before you use this information and the product it supports, read the information in "Safety" on page xiii, Appendix C, "Notices," on page 197 , and *IBM Systems Environmental Notices and User Guide*, Z125-5823.

Contents

Figures

Tables

Safety

Safety notices

Safety notices may be printed throughout this guide. **DANGER** notices warn you of conditions or procedures that can result in death or severe personal injury. **CAUTION** notices warn you of conditions or procedures that can cause personal injury that is neither lethal nor extremely hazardous. **Attention** notices warn you of conditions or procedures that can cause damage to machines, equipment, or programs.

World trade safety information

Several countries require the safety information contained in product publications to be presented in their translation. If this requirement applies to your country, a safety information booklet is included in the publications package shipped with the product. The booklet contains the translated safety information with references to the US English source. Before using a US English publication to install, operate, or service this product, you must first become familiar with the related safety information in the *Systems Safety Notices*, G229-9054. You should also refer to the booklet any time you do not clearly understand any safety information in the US English publications.

Laser safety information

All IBM Z® (Z) and IBM® LinuxONE (LinuxONE) models can use I/O cards such as FICON®, Open Systems Adapter (OSA), InterSystem Channel-3 (ISC-3), RoCE Express, Integrated Coupling Adapter (ICA SR), zHyperLink Express, or other I/O features which are fiber optic based and utilize lasers (short wavelength or long wavelength lasers).

Laser compliance

All lasers are certified in the US to conform to the requirements of DHHS 21 CFR Subchapter J for Class 1 or Class 1M laser products. Outside the US, they are certified to be in compliance with IEC 60825 as a Class 1 or Class 1M laser product. Consult the label on each part for laser certification numbers and approval information.

Laser Notice: U.S. FDA CDRH NOTICE if low power lasers are utilized, integrated, or offered with end product systems as applicable. Complies with 21 CFR 1040.10 and 1040.11 except for conformance with IEC 60825-1 Ed. 3., as described in Laser Notice No. 56, dated May 8, 2019.

CAUTION: Data processing environments can contain equipment transmitting on system links with laser modules that operate at greater than Class 1 power levels. For this reason, never look into the end of an optical fiber cable or open receptacle. (C027)

CAUTION: This product contains a Class 1M laser. Do not view directly with optical instruments. (C028)

About this publication

This information is intended for system planners, installation managers, and other technical support personnel who need to plan for operating in logically partitioned mode (LPAR mode) on the, IBM z15™ (z15™) / IBM LinuxONE III

This publication assumes previous knowledge of the characteristics and functions of the installed central processor complex (CPC).

To improve readability, we refer to the different CPCs using the following terminology whenever possible:

Note: Reference to IBM Z are also applicable to IBM LinuxONE.

Table 1. Terminology used in this publication	
Terminology	**Central Processor Complex (CPC)**
z15 / LinuxONE III	Models T02 and LT2
z15 / LinuxONE III	Models T01 and LT1

Some features, windows, and functions are model-dependent, engineering change (EC) level-dependent, machine change level-dependent (MCL-dependent), or control program-dependent. For this reason, not all of the functions discussed in this publication are necessarily available on every CPC.

Some illustrations and examples in this publication describe operation with as few as 2 logical partitions (LPs), although up to 85 LPs for models T01 and LT1 and up to 40 LPs for models T02 and LT2 can be defined on a z15 machine.

Figures included in this document illustrate concepts and are not necessarily accurate in content, appearance, or specific behavior.

Sample tasks and panels explained in this publication reference tasks and windows available from the Support Element console. Hardware Management Console (HMC) and Support Element (SE) information can be found on the console help system. Also on http://www.ibm.com/servers/resourcelink. Select **Library** and then select your system.

Control program operators should refer to the appropriate control program publication for information on control program commands.

However, for the most current coupling facility control code information for all models, use this publication.

What is included in this publication

The information presented in this publication is organized as follows:

- Chapter 1, "Introduction to logical partitions," on page 1 describes the prerequisites for establishing and using PR/SM, the general characteristics and some potential applications for LPs.
- Chapter 2, "Planning considerations," on page 35 presents considerations and guidelines for I/O configuration planning and coupling facility planning.
- Chapter 3, "The characteristics of logical partitions," on page 91 includes a list of the panels, provides guidelines for determining the CPC resources, and describes the operator tasks used to define the characteristics of LPs.
- Chapter 4, "Monitoring the activities of logical partitions," on page 163 describes the panels and operator tasks used to monitor LP activity.
- Appendix A, "Coupling facility control code support," on page 183 lists and explains the support provided at different levels of coupling facility control code Licensed Internal Code (LIC).

    xv

- Appendix B, "Developing, building, and delivering a certified system," on page 185 provides guidance in setting up, operating, and managing a secure consolidated environment using PR/SM.
- Appendix C, "Notices," on page 197 contains electronic emission notices, legal notices, and trademarks.

Related publications

The following publications provide information about the functions and characteristics of the different CPCs and the related operating systems that run on them.

z/Architecture

- *z/Architecture Principles of Operation*, SA22-7832

Hardware

z15

- Hardware Management Console (HMC) and Support Element (SE) information can be found on the console help system.
- *Secure Service Container User's Guide*, SC28-6978
- *IOCP User's Guide for ICP*, SB10-7172
- *Stand-Alone IOCP User's Guide*, SB10-7173

z14

- Hardware Management Console (HMC) and Support Element (SE) information can be found on the console help system.
- *Secure Service Container User's Guide*, SC28-6978
- *IOCP User's Guide for ICP*, SB10-7172
- *Stand-Alone IOCP User's Guide*, SB10-7173

z13s

- Hardware Management Console (HMC) and Support Element (SE) information can be found on the console help system.
- *Secure Service Container User's Guide*, SC28-6971
- *Input/Output Configuration Program User's Guide for ICP IOCP*, SB10-7163
- *Stand-Alone IOCP User's Guide*, SB10-7166

z13

- Hardware Management Console (HMC) and Support Element (SE) information can be found on the console help system.
- *IBM z Systems Secure Service Container User's Guide*, SC28-6971
- *IBM z Systems Advanced Workload Analysis Report (IBM zAware) Guide*, SC27-2632
- *IBM z Systems Input/Output Configuration Program User's Guide for ICP IOCP*, SB10-7163
- *IBM z Systems Stand-Alone IOCP User's Guide*, SB10-7166

zEnterprise® BC12

- The content from the following publications is now incorporated into the Hardware Management Console (HMC) and Support Element (SE) (Version 2.12.1) help system:
  - *System z Hardware Management Console Operations Guide*

   – *zEnterprise System Support Element Operations Guide*

- *System z Advanced Workload Analysis Report (IBM zAware) Guide*, SC27-2623
- *zEnterprise BC12 System Overview*, SA22-1089
- *System z Input/Output Configuration Program User's Guide for ICP IOCP,* SB10-7037
- *System z® Stand-alone IOCP User's Guide*, SB10-7152

zEnterprise EC12

- *System z Advanced Workload Analysis Report (IBM zAware) Guide*, SC27-2623
- *System Overview*, SA22-1088
- *System z Input/Output Configuration Program User's Guide for ICP IOCP,* SB10-7037
- *System z Stand-alone IOCP User's Guide,* SB10-7152
- *System z Hardware Management Console Operations Guide*, SC28-6919 (Version 2.12.0)
- *Support Element Operations Guide*, SC28-6920 (Version 2.12.0)
- The content from the following publications is now incorporated into the Hardware Management Console (HMC) and Support Element (SE) (Version 2.12.1) help system:

   – *System z Hardware Management Console Operations Guide*

   – *zEnterprise System Support Element Operations Guide*

FICON

- *ESCON and FICON Channel-to-Channel Reference*, SB10-7034

Crypto features

The following publications provide additional information on the Crypto features:

- *System z Crypto and TKE Update*, SG24-7848
- Hardware Management Console (HMC) and Support Element (SE) information can be found on the console help system.
- *User Defined Extensions Reference and Guide,* website: http://www.ibm.com/security/cryptocards (Select a crypto card, and then click **Library**)

Software

z/OS

zSeries Parallel Sysplex

The following publications provide additional information about the z/OS® Parallel Sysplex® environment:

- *z/OS Parallel Sysplex Overview*, SA22-7661
- *z/OS Parallel Sysplex Application Migration,* SA22-7662
- *z/OS MVS Setting Up a Sysplex*, SA23-1399
- *z/OS MVS Programming: Sysplex Services Guide*, SA23-1400
- *z/OS MVS Programming: Sysplex Services Reference*, SA38-0658

Multiple Image Facility

The following publications provide additional information about Multiple Image Facility in the z/OS environment:

- *z/OS Hardware Configuration Definition User's Guide,* SC34-2669

Dynamic I/O Configuration

The following publication provides information about dynamic I/O configuration in the z/OS environment:

- *z/OS Hardware Configuration Definition Planning*, GA32-0907

Dynamic Storage Reconfiguration

The following publications provide additional information on the commands, functions, and capabilities of dynamic storage reconfiguration in the z/OS environment:

- *z/OS MVS Initialization and Tuning Reference*, SA23-1380
- *z/OS MVS Recovery and Reconfiguration Guide*, SA22-7623
- *z/OS MVS System Commands*, SA38-0666

Crypto features

The following publications provide additional information on the Crypto features:

- *z/OS ICSF Administrator's Guide,* SC14-7506
- *z/OS ICSF System Programmer's Guide*, SA14-7507

Sysplex Failure Manager

The following publication provides an overview of SFM and practical information for implementing and using SFM in the z/OS environment:

- *z/OS MVS Setting Up a Sysplex*, SA23-1399

LPAR Management Time

The following publication provides information about the RMF Partition Data Report that includes LPAR Management Time reporting in a z/OS environment:

- *z/OS Resource Measurement Facility User's Guide*, SC34-2664

Intelligent Resource Director (IRD)

The following publication provides information about Intelligent Resource Director in a z/OS environment:

- *z/OS MVS Planning Workload Management*, SC34-2662

z/VM

Hardware Configuration Definition (HCD)

The following publication provides information about the Hardware Configuration Definition (HCD):

- *z/VM I/O Configuration*, SC24-6291

Hardware Configuration Manager

The following publication provides information about the Hardware Configuration Manager:

- *z/OS and z/VM Hardware Configuration Manager User's Guide*, SC34-2670

Dynamic I/O Configuration

The following publication provides information about dynamic I/O configuration:

- *z/VM CP Planning and Administration*, SC24-6271
- *z/VM I/O Configuration*, SC24-6291

Guest Operating Systems

The following publication provides information about running **guest** operating systems:

- *z/VM Running Guest Operating Systems*, SC24-6321

z/VSE

The following publication provides information about the z/VSE® environment:

- *z/VSE Planning V6.2*, SC34-2681-01.

Revisions

A technical change from the previous edition of this document is indicated by a thick vertical line to the left of the change.

Accessibility

Accessible publications for this product are offered in EPUB format and can be downloaded from Resource Link® at http://www.ibm.com/servers/resourcelink.

If you experience any difficulty with the accessibility of any IBM Z and IBM LinuxONE information, go to Resource Link at http://www.ibm.com/servers/resourcelink and click **Feedback** from the navigation bar on the left. In the **Comments** input area, state your question or comment, the publication title and number, choose **General comment** as the category and click **Submit**. You can also send an email to reslink@us.ibm.com providing the same information.

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

Accessibility features

The following list includes the major accessibility features in IBM Z and IBM LinuxONE documentation, and on the Hardware Management Console and Support Element console:

- Keyboard-only operation
- Interfaces that are commonly used by screen readers
- Customizable display attributes such as color, contrast, and font size
- Communication of information independent of color
- Interfaces commonly used by screen magnifiers
- Interfaces that are free of flashing lights that could induce seizures due to photo-sensitivity.

Keyboard navigation

This product uses standard Microsoft Windows navigation keys.

Consult assistive technologies

Assistive technology products such as screen readers function with our publications, the Hardware Management Console, and the Support Element console. Consult the product information for the specific assistive technology product that is used to access the EPUB format publication or console.

IBM and accessibility

See http://www.ibm.com/able for more information about the commitment that IBM has to accessibility.

How to send your comments

Your feedback is important in helping to provide the most accurate and high-quality information. Send your comments by using Resource Link at http://www.ibm.com/servers/resourcelink. Click **Feedback** on the Navigation bar on the left. You can also send an email to reslink@us.ibm.com. Be sure to include the name of the book, the form number of the book, the version of the book, if applicable, and the specific location of the text you are commenting on (for example, a page number, table number, or a heading).

Chapter 1. Introduction to logical partitions

This chapter introduces the characteristics of logical partitioning and migration and compatibility considerations. Processor Resource/Systems Manager (PR/SM) is standard on all z15 models.

Prerequisites for operation

The prerequisites for operation are:

- Programming compatibility
- Programming support
  - Control program support
  - Input/Output Configuration Program (IOCP) support
- Hardware support
- Operator training

PR/SM

PR/SM enables logical partitioning of the central processor complex (CPC).

Logical partitioning

PR/SM enables the logical partitioning function of the CPC. The operator defines the resources that are to be allocated to each logical partition (LP). Most resources can be reconfigured without requiring a power-on reset. After a **General** , **z/VM®**, or **Linux®-Only** LP is defined and activated, you can load a supported control program into that LP. If a coupling facility logical partition is defined and activated, the coupling facility control code is automatically loaded into the LP. If a Secure Service Container partition is defined and activated, a supported software appliance feature can be installed and initialized.

Central storage

Central storage is defined to LPs before LP activation. When an LP is activated, the storage resources are allocated in contiguous blocks. These allocations can be dynamically reconfigured. Sharing of allocated central storage among multiple LPs is not allowed.

All storage is defined as central storage. See . The sum total of all LP central storage cannot exceed the amount of customer storage.

Central processors

Central processors (CPs) can be dedicated to a single LP or shared among multiple LPs. CPs are allocated to an LP when the LP is activated. You can use operator tasks to limit and modify the use of CP resources shared between LPs while the LPs are active.

Virtual Flash Memory

The Virtual Flash Memory is designed to improve availability and handling of paging workload spikes when running z/OS V2.1 or higher. The Virtual Flash Memory support with z/OS is designed to help improve system availability and responsiveness using Virtual Flash Memory across transitional workload events, such as market openings and diagnostic data collection.

The Virtual Flash Memory can also be used in coupling facility images to provide extended availability for workloads making use of WebSphere MQ Shared Queues structures. Virtual Flash Memory can help availability by reducing latency from paging delays that can occur at the start of the workday or during transitional periods. Virtual Flash Memory is also designed to help eliminate delays that can occur when collecting diagnostic data during failures.

The initial Virtual Flash Memory represents the amount of Virtual Flash Memory allocated to an activated logical partition. The maximum Virtual Flash Memory amount represents the maximum Virtual Flash Memory the logical partition is allowed. This means, if the initial and maximum amounts are specified, the maximum amount minus the initial amount is the Virtual Flash Memory amount that the logical partition's operating system can dynamically configure.

Multiple Image Facility

The Multiple Image Facility (MIF) is available on all CPCs discussed in this publication. MIF allows channel sharing among LPs. For information about accessing devices on shared channel paths and defining shared channel paths, see "Defining shared channel paths" on page 122.

MCSS

Multiple Logical Channel Subsystems (CSS) are available on all CPCs discussed in this publication. Each CSS supports a definition of up to 256 channels.

Channel paths

Active LPs can share channels. Shared channels require that the channel subsystem create a logical path for each channel image corresponding to an active LP that has the channel configured online. CNC, CTC, OSC, OSD, OSE, OSM, OSN, OSX, CL5, CS5, ICP, FC, FCP, and IQD channel path types can be shared. CVC and CBY channel paths *cannot* be shared.

For information about accessing devices on shared channel paths and defining shared channel paths, see "Defining shared channel paths" on page 122.

Simultaneous multithreading (SMT)

Higher workload throughput may be achieved because the z15 processor chip offers intelligently implemented 2-way simultaneous multithreading. Simultaneous multithreading (SMT) allows two active instruction streams per core, each dynamically sharing the core's execution resources. SMT is in z15 for workloads running on SAPs, Integrated Facility for Linux (IFL), and z Integrated Information Processor (zIIP).

Each software Operating System has the ability to intelligently drive SMT in a way that is best for its unique requirements. z/OS SMT management consistently drives the cores to high thread density to reduce SMT variability and deliver repeatable performance across varying CPU utilization. This provides more predictable SMT capacity. The z/VM SMT management optimizes throughput by spreading a workload over the available cores until it demands the additional SMT capacity.

The operating system running in a logical partition will optionally enable that logical partition to use SMT. Some logical partitions may be running with SMT enabled while others are not. It is important to understand that logical partition management of processor resources is performed at the logical core level. The number of processors you define for use in the logical partition is the number of logical cores it has. When the partition is not enabled for SMT, this is equivalent to the number of CPUs in the logical partition. When a logical partition enables SMT, each logical core gets two CPUs (separate instruction streams) defined. The operating system then decides when one or both of the CPUs of a particular core are to be used. PR/SM then manages the dispatching of the logical cores to physical cores.

IBM Secure Service Container

The IBM Secure Service Container was known as the IBM z Appliance Container Infrastructure (zACI) which is documented in the *Systems Appliance Container Infrastructure (zACI) User's Guide*, SC28-6970. The IBM Secure Service Container is a container technology through which you can quickly and securely deploy platform and software appliances on Z and LinuxONE servers.

A Secure Service Container partition is a specialized container for installing and running specific firmware or software appliances. An appliance is an integration of operating system, middleware and software components that work autonomously and provide core services and infrastructures that focus on consumability and security.

The suggested practice is to use the latest available firmware for a Secure Service Container partition. The latest firmware for the HMC/SE uses Secure Service Container mode. For more information, see *Systems Secure Service Container User's Guide*, SC28-6978.

IBM System Recovery Boost

The IBM System Recovery Boost is a new function on a z15 that is available with z/OS environments and with z/VM and z/TPF when running on general purpose processors. It can help recover workloads substantially faster than on prior Z machines. The System Recovery Boost expedites system shutdown processing, system IPL (Initial Program Load), middleware/workload restart and recovery, and the client workload execution that follows. It provides higher processor capacity for a limited period of time, called the "Boost period."

IBM System Recovery Boost allows your system to return to work faster, not just from catastrophes, but after all kinds of disruptions, both expected and unexpected. It helps catch up for lost time and get back to production and recover lost business transactions faster than previously possible, reducing impacts to service level agreements (SLAs). The IBM System Recovery Boost provides you with the optional ability to allocate additional resources to get the system back and ready for work faster and catch up for lost time, without increasing monthly software costs. For planning considerations, see "IBM System Recovery Boost planning considerations" on page 81.

Crypto features

The CP Assist for Cryptographic Function (CPACF) enablement (#3863) feature must be installed on your system prior to using the Crypto features.

Crypto Express

Crypto Express7S is state-of-the-art, tamper-sensing, and tamper responding, programmable cryptographic cards. The cryptographic electronics and microprocessor, housed within a tamper-responding container, provide a secure cryptographic environment, designed to meet FIPS 140-2 Level 4 requirements. The Crypto Express7S feature provides a PCI Express (PCIe) interface to the host. The concurrent update for CCA firmware is supported.

The Crypto Express7S feature contains one adapter. Crypto Express7S feature can be configured as an accelerator, a Common Cryptographic Architecture (CCA) coprocessor, or an Enterprise PKCS #11 (EP11) coprocessor.

Key features of Crypto Express7S includes:

- Consolidation and simplification. Each crypto adapter can be defined as a coprocessor or accelerator
- For Crypto Express7S virtualization of the crypto allows up to 85 logical partitions for models T01 and LT1 and up to 40 logical partitions for models T02 and LT2.
- Improved Reliability, Availability & Serviceability (RAS)
- Dynamic power management to maximize RSA performance while keeping within temperature limits of the tamper-responding package
- User Defined Extensions (UDXs) which provide the ability to embed customized function in the coprocessor firmware
- Secure code loading that ensures the card will only accept firmware that has not been modified and which comes from an IBM-approved source
- Concurrent patch and driver update to allow updating card functionality while installed in application systems. Applications can continue to use the cards while firmware is being updated
- Lock-step checking of dual CPUs for enhanced error detection
- Dynamic addition and configuration of cryptographic features to logical partitions without an outage
- Updated cryptographic algorithms used in firmware loading and with the TKE workstation to keep up with current recommendations for cryptographic security
- Support for EMV smart card applications.

TKE migration wizard

A wizard is available on the TKE Workstation to allow a user to migrate the roles and authorities of a crypto card to another crypto card, and the domain controls and the master keys for a particular crypto domain to another crypto domain. The migration data is moved quickly and securely from one card to another. Additionally, the source and target Crypto Express7S coprocessors must be configured as CCA coprocessors.

To locate the migration wizard and to view the criteria for migrating from Crypto Express7S, log on to the TKE workstation, and click **What's New**, listed under **Additional Resources** on the **Welcome** page.

CP Assist for Cryptographic Functions

CP Assist for Cryptographic Functions (CPACF), supporting clear and protected key encryption, is activated using the no-charge enablement feature (#3863). It offers the following:

- For data privacy and confidentially: Data Encryption Standard (DES), Triple Data Encryption Standard (TDES), Advanced Encryption Standard (AES) for 128-bit, 192-bit and 256-bit keys.
- For data integrity: Secure Hash Algorithm-1 (SHA-1) 160-bit, and SHA-2 for 224-, 256-, 384-and 512-bit support. SHA-3 for 224-, 256- 384-, 512 bit support, and SHAKE for 128-, and 256 bit support. SHA-1, SHA-2 and SHA3 are shipped enabled and do not require the no-charge enablement feature.
- For Key Generation: Pseudo Random Number Generation (PRNG), Deterministic Random Number Generation (DRNG), and True Random number generation (TRNG).
- CPACF instructions for SHA-3 hashing, TRNG (True Random Number Generation), and Improved performance of AES GCM encryption.
- The Elliptical Curve Cryptography (ECC) key import functions to control the enablement of digital signatures (KDSA instruction).

Protected key CPACF

Protected key CPACF blends the speed of CPACF with the secure key cryptography offered by the Crypto Express7S coprocessor feature. This function ensures the privacy of cryptographic key material when used by the CPACF for high-performance data encryption. Protected key CPACF helps ensure that key material is not visible to applications or by the operating system, maintaining the security of the key material used for the encryption operation. This capability provides better performance than secure key operations (that execute on the Crypto Express card) while providing better security than clear key operations. This function can work with the Crypto Express card to protect secure keys when they are used on the CPACF.

Security Comparison

compares CPACF clear key, CPACF protected key, and Crypto Express5S/6S/7S security.

Table 2. Comparison between the security and Cryptos			
	CPACF clear key	**CPACF protected key**	**Crypto Express5S/6S/7S**
Security of keys	**Low**: Keys are not encrypted in storage and when sent to CPACF for use.	**Medium**: Keys are stored in the key repository as secure keys, encrypted under the appropriate master key. When used as a protected key, the key material is decrypted from under the master key and re-encrypted under a wrapping key.	**High**: Keys are protected by encryption at all times, and hardware has tamper sensors and zeroizes keys when tampering is detected.

Table 2. Comparison between the security and Cryptos (continued)

	CPACF clear key	CPACF protected key	Crypto Express5S/6S/7S
Functions provided	Basic symmetric key with additional chaining options, and hashing and MACs functions.	Basic symmetric key with additional chaining options, and hashing and MACs functions.	Full CCA and EP11 function set including symmetric and public-key cryptography and key management, hashing, special banking and finance functions, and others.

Parallel Sysplex support

Parallel sysplex uses a broad range of hardware and software products to process in parallel a transaction processing workload across multiple z/OS images running in a sysplex and sharing data in a coupling facility.

Parallel sysplex allows you to manage a transaction processing workload, balanced across multiple z/OS images running on multiple CPCs, as a single data management system. It also offers workload availability and workload growth advantages.

The parallel sysplex enhances the capability to continue workload processing across scheduled and unscheduled outages of individual CPCs participating in a sysplex using a coupling facility by making it possible to dynamically reapportion the workload across the remaining active sysplex participants. Additionally, you can dynamically add processing capacity (CPCs or LPs) during peak processing without disrupting ongoing workload processing.

CPC support enables you to:

- Install coupling facility channels
- Define, as an LP, a portion or all the CPC hardware resources (central processors, storage, coupling facility channels, and flash memory) for use as a coupling facility that connects to z/OS images for data sharing purposes
- Connect to a coupling facility to share data
- Define a z15 with only ICFs to serve as a stand-alone coupling facility, which might contain one or more coupling facility images, but which cannot run z/OS or any other operating system.
- Define a z15 with both ICFs and other types of processors, where the ICF engines can be used to serve one or more coupling facility images, and the other types of processors can be used to run z/OS or any other operating system

For more information about the coupling facility including z/OS and CPC support for coupling facility levels, see "Coupling facility planning considerations" on page 63.

Guest coupling simulation

Guest coupling simulation is available with z/VM. The z/VM guest coupling simulation allows you to simulate one or more complete parallel sysplexes within a single z/VM system image, providing a test environment for parallel sysplex installation. The simulated environment is not intended for production use since its single points of failure diminish the availability advantages of the parallel sysplex environment. There are no special hardware requirements (external coupling facility channels, external coupling facilities, and Sysplex Timers are not necessary or supported). Guest operating systems within a simulated sysplex can only be coupled (through simulated coupling facility channels) to virtual coupling facilities also running as guests of the same z/VM system. You can have up to 32 virtual machines running z/OS within a simulated sysplex, with each z/OS virtual machine coupled to up to eight virtual machines running as coupling facilities.

There is no system-imposed limit to the number of guest parallel sysplex environments that z/VM can simulate. However, practical limits on the number of guests that can be supported by a particular hardware configuration constrain the number of simulated parallel sysplex environments.

Control program support in a logical partition

Control programs require certain characteristics. Before planning or defining LP characteristics, call your installation management to determine which control programs are in use or planned for operation.

Notes:

1. Use the support system to view the appropriate PSP bucket subset ID for hardware and software maintenance information.

2. For more detailed information about support for coupling facility levels (including hardware EC, driver, and MCL numbers and software APAR numbers), see "Coupling facility level (CFLEVEL) considerations" on page 72.

z/OS

Support for z15	z/OS 2.1	z/OS 2.2	z/OS 2.3	z/OS 2.4
Support for z15	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
OSA-Express7S GbE LX and GbE SX				
CHPID type OSD	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
OSA-Express7S 10 GbE LR and GbE SR				
CHPID type OSD	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
OSA-Express7S 25 GbE SR				
CHPID type OSD	with applicable PTFs	with applicable PTFs	with applicable PTFs	Yes
OSA-Express7S 1000BASE-T Ethernet				
CHPID Type OSC supporting TN3270E and non SNA-DFT	Yes	Yes	Yes	Yes
CHPID type OSD with exploitation of two ports per CHPID	Yes	Yes	Yes	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes	Yes	Yes	Yes
OSA-Express6S 10 GbE LR and GbE SR				
CHPID type OSD	with applicable PTFs	with applicable PTFs	with applicable PTFs	Yes (with applicable PTFs)

Support for z15	z/OS 2.1	z/OS 2.2	z/OS 2.3	z/OS 2.4
CHPID type OSX for access control to the intra-ensemble data network (IEDN) from Unified Resource Manager functions	with applicable PTFs	with applicable PTFs	with applicable PTFs	Yes (with applicable PTFs)
OSA-Express5S 10 GbE LR and GbE SR				
CHPID type OSD and OSX. A PTF is required for z/OS 1.13 support for OSD	Yes	Yes	Yes	Yes
OSA-Express6S 10 GbE LX and GbE SX				
CHPID type OSD. A PTF is required for z/OS 1.13 support	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
OSA-Express5S 10 GbE LX and GbE SX				
CHPID type OSD	Yes	Yes	Yes	Yes
OSA-Express6S 1000BASE-T Ethernet				
CHPID type OSC supporting TN3270E and non-SNA DFT	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
CHPID type OSD with exploitation of two ports per CHPID	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
OSA-Express5S 1000BASE-T Ethernet				
CHPID type OSD	Yes	Yes	Yes	Yes
CHPID type OSX for access control to the intraensemble data network (IEDN)	Yes	Yes	Yes	Yes
CHPID type OSM	Yes	Yes	Yes	Yes
CHPID type OSD without maximum port exploitation	Yes	Yes	Yes	Yes
CHPID type OSN supporting OSA-Express for NCP	Yes	Yes	Yes	Yes
CP Assist for Cryptographic Function (CPACF)	Yes	Yes	Yes	Yes

Support for z15	z/OS 2.1	z/OS 2.2	z/OS 2.3	z/OS 2.4
Protected Key CP Assist for Cryptographic Function (CPACF)	Yes	Yes	Yes	Yes
Crypto Express7S Toleration (1 port)	Yes	Yes	Yes	with applicable PTFs
Crypto Express7S Exploitation	with enhanced cryptographic support	with enhanced cryptographic support	with enhanced cryptographic support	with enhanced cryptographic support
Crypto Express6S	Yes	Yes	Yes	Yes
Crypto Express6S Toleration	Yes	Yes	Yes	Yes
Crypto Express5S Toleration	Yes	Yes	Yes	Yes
Crypto Express7S, Crypto Express6S, and Crypto Express5S support of Visa Format Preserving Encryption	No	with applicable PTFs	with applicable PTFs	with applicable PTFs
Crypto Express7S, Express6S, and Crypto Express5S support for more than 16 domains	No	with applicable PTFs	with applicable PTFs	with applicable PTFs
Secure Key Advanced Encryption Standard (AES)	Yes	Yes	Yes	Yes
STP System Management with new z/OS Messaging	Yes	Yes	Yes	Yes
STP for CTN split and merge capability	No	with applicable PTFs	with applicable PTFs	with applicable PTFs
CHPID type OSD supporting Checksum offload for IPv6 packets	Yes	Yes	Yes	Yes
CHPID type OSD supporting Checksum offload for LPAR-to-LPAR traffic for IPv4 and IPv6 packets	Yes	Yes	Yes	Yes
CHPID type OSD supporting Large Send for IPv6 packets	Yes	Yes	Yes	Yes
CHPID type CS5 supporting Integrated Coupling Adapter (ICA SR)	Yes	Yes	Yes	Yes
CHPID type CL5 supporting Coupling Express LR (CE LR)	with applicable PTFs	Yes	Yes	Yes
CHPID type OSM supporting intranode management network (INMN)	No	with applicable PTFs	with applicable PTFs	Yes
CHPID type OSE supporting 4 or 2 ports per feature	with applicable PTFs	with applicable PTFs	with applicable PTFs	Yes
CHPID type FC using native FICON or Channel-To-Channel (CTC)	Yes	Yes	Yes	Yes

Support for z15	z/OS 2.1	z/OS 2.2	z/OS 2.3	z/OS 2.4
CHPID type FC for support of zHPF single-track operations	Yes	Yes	Yes	Yes
CHPID type FC for support of zHPF multi-track operations	Yes	Yes	Yes	Yes
RoCE Express 2.1 10 GbE for Shared Memory Communications - Remote Direct Memory Access (SMC-R)	No	with applicable PTFs	with applicable PTFs	Yes
RoCE Express 2.1 25 GbE for Shared Memory Communications - Remote Direct Memory Access (SMC-R)	No	with applicable PTFs	with applicable PTFs	Yes
RoCE Express2 for Shared Memory Communications - Remote Direct Memory Access (SMC-R)	No	with applicable PTFs	with applicable PTFs	Yes
RoCE Express2 25 GbE for Shared Memory Communications - Remote Direct Memory Access (SMC-R)	No	with applicable PTFs	with applicable PTFs	Yes
Simultaneous multithreading (SMT)	No	No	Yes	Yes
Support for 384 Coupling CHPIDs	with applicable PTFs	Yes	Yes	Yes
zHyperLink Express Writes support	No	with applicable PTFs	with applicable PTFs	Yes
zHyperLink Express	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
Coupling Express LR (CE LR)	with applicable PTFs	Yes	Yes	Yes
Integrated Coupling Adapter 2 Fanout (ICA SR)	with applicable PTFs	Yes	Yes	Yes
IPL from an alternate subchannel set	Yes	Yes	Yes	Yes
Support for dynamically increasing the number of CPs, ICFs, IFLs, or zIIPs, without an intervening IPL	Yes	Yes	Yes	Yes
Large Page Support (1 MB pages)	Yes	Yes	Yes	Yes
HiperDispatch	Yes	Yes	Yes	Yes
Server Time Protocol	Yes	Yes	Yes	Yes

Support for z15	z/OS 2.1	z/OS 2.2	z/OS 2.3	z/OS 2.4
Maximum number of CPs	256 cores when SMT is not enabled and 128 cores/256 threads when SMT is enabled	256 cores when SMT is not enabled and 128 cores/256 threads when SMT is enabled	256 cores when SMT is not enabled and 128 cores/256 threads when SMT is enabled	100
Maximum central storage on	4 TB	4 TB	4 TB	4 TB
Maximum number of channel paths	256	256	256	256
Support for multiple LCSSs	Yes	Yes	Yes	Yes
Maximum CFLEVEL supported	20 with applicable PTFs	24 with applicable PTFs	24 with applicable PTFs	24 with applicable PTFs
Asynchronous Cache XI capability	No	Yes with applicable PTFs	Yes with applicable PTFs	Yes with applicable PTFs
Asynchronous CF duplexing for lock structures	No	Yes with applicable PTFs	Yes with applicable PTFs	Yes with applicable PTFs
System-Managed CF Structure Duplexing support	Yes	Yes	Yes	Yes
Multiple subchannel sets (MSS) for FICON (FC)	Yes	Yes	Yes	Yes
XES Coupling Facility cache structure architecture extensions for batch write, castout, and cross-invalidate functions	Yes	Yes	Yes	Yes
z/Architecture® 64- bit addressing	Yes	Yes	Yes	Yes
Support for message time ordering facility (MTOF)	Yes	Yes	Yes	Yes
Support z ELC software pricing structure	Yes	Yes	Yes	Yes
Intelligent Resource Director	Yes	Yes	Yes	Yes
Workload pricing	Yes	Yes	Yes	Yes
Peer mode channels (ICP)	Yes	Yes	Yes	Yes
XES Coupling Facility List Structure architecture extensions for shared queues	Yes	Yes	Yes	Yes
Logical Partition time offset	Yes	Yes	Yes	Yes
Internal Coupling Facility channels (ICP)	Yes	Yes	Yes	Yes
System-Managed Rebuild	Yes	Yes	Yes	Yes
Transactional Memory	Yes	Yes	Yes	Yes

Support for z15	z/OS 2.1	z/OS 2.2	z/OS 2.3	z/OS 2.4
IBM zAware	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
Flash Express® (exploitation by CF)	with applicable PTFs and web deliverable	with applicable PTFs and web deliverable	with applicable PTFs and web deliverable	with applicable PTFs and web deliverable
CFCC Flash Express or Virtual Flash Memory Exploitation	Yes	Yes	Yes	Yes
IBM Virtual Flash Memory	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
XL C/C++ support of ARCH(11) and TUNE(11) parameters	No	with applicable PTFs	with applicable PTFs and web deliverable	with applicable PTFs and web deliverable
XL C/C++ support of ARCH(13) and TUNE(13) parameters	No	No	No	with applicable PTFs
2GB Large Pages	Yes	Yes	Yes	Yes
z/OS global resource serialization (GRS) support for FICON CTCs	Yes	Yes	Yes	Yes
Guarded Storage	No	Yes	with applicable PTFs	with applicable PTFs
Instruction Execution Protection Facility	No	with applicable PTFs	with applicable PTFs	with applicable PTFs
CPU Measurement Facility support for z15	No	with applicable PTFs	with applicable PTFs	with applicable PTFs
Compression/CPU Deflate and Asynchronous Compression	No	with applicable PTFs	with applicable PTFs	with applicable PTFs
Coupling Facility Scalability Enhancements	with applicable PTFs	with applicable PTFs	Yes	Yes
Coupling Facility List Notification Enhancements	with applicable PTFs	with applicable PTFs	Yes	Yes
Fair Latch Manager	with applicable PTFs	with applicable PTFs	with applicable PTFs	with applicable PTFs
Quantum Safe Support	No	with FMID HCR77D1	with FMID HCR77D1	with FMID HCR77D1
System Recovery Boost	No	No	with applicable PTFs	with applicable PTFs
Sysplex Time Synchronization	Yes	Yes	Yes	Yes
IBM Z Data Privacy for Diagnostics	No	No	with applicable PTFs	with applicable PTFs
Sort Acceleration support for DFSORT	No	with applicable PTFs	with applicable PTFs	with applicable PTFs

Support for z15	z/OS 2.1	z/OS 2.2	z/OS 2.3	z/OS 2.4
Note:				
1. z/OS V2.1 supports z15; however, z/OS V2.1 support was withdrawn September 30, 2018. After that date, the IBM Software Support Services for z/OS V2.1 is required for z15. Talk to your representative for details. Certain function and features of the z15 system require later releases of z/OS. For a complete list of software support, see the 2827DEVICE Preventive Planning (PSP) bucket. For more information on the IBM Software Support Services for V2.1, see Software Announcement 614-010, dated June 24, 2018.				

z/VM

Support for z15	z/VM 6.4	z/VM 7.1	z/VM 7.2
Support for z15	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
OSA-Express7S GbE LX and GbE SX			
CHPID type OSD	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
OSA-Express7S 10 GbE LR and GbE SR			
CHPID type OSD	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
OSA-Express7S 25 GbE SR			
CHPID type OSD	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
OSA-Express7S 1000BASE-T Ethernet			
CHPID Type OSC supporting TN3270E and non-SNA DFT	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
CHPID type OSD with exploitation of two ports per CHPID	Yes (with applicable PTFs)	Yes (with applicable PTFs)	
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes (with applicable PTFs)	Yes (with applicable PTFs)	Yes
OSA-Express6S 10 GbE LR and GbE SR			
CHPID type OSD	Yes	Yes	Yes
CHPID type OSX for access control to the intra-ensemble data network (IEDN) from Unified Resource Manager functions	Only to define, modify, and delete CHPID type OSX, when z/VM is the controlling LP for dynamic I/O	Only to define, modify, and delete CHPID type OSX, when z/VM is the controlling LP for dynamic I/O	Only to define, modify, and delete CHPID type OSX, when z/VM is the controlling LP for dynamic I/O

Support for z15	z/VM 6.4	z/VM 7.1	z/VM 7.2
OSA-Express5S 10 GbE LR and GbE SR			
CHPID type OSD	Yes	Yes	Yes
CHPID type OSX	Only to define, modify, and delete CHPID type OSX, when z/VM is the controlling LP for dynamic I/O.	Only to define, modify, and delete CHPID type OSX, when z/VM is the controlling LP for dynamic I/O.	Only to define, modify, and delete CHPID type OSX, when z/VM is the controlling LP for dynamic I/O.
OSA-Express6S 10 GbE LX and GbE SX			
CHPID type OSD	Yes	Yes	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes	Yes	Yes
OSA-Express5S 10 GbE LX and GbE SX			
CHPID type OSD	Yes	Yes	Yes
OSA-Express6S 1000BASE-T Ethernet			
CHPID type OSC supporting TN3270E and non-SNA DFT	Yes	Yes	Yes
CHPID type OSD with exploitation of two ports per CHPID	Yes	Yes	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes	Yes	Yes
OSA-Express5S 1000BASE-T Ethernet			
CHPID type OSD	Yes	Yes	Yes
CHPID type OSX for access control to the intraensemble data network (IEDN) from zBC12 and zEC12 to Unified Resource Manager functions	Only to define, modify, and delete OSX CHPID types when z/VM is the controlling LP for dynamic I/O	Only to define, modify, and delete OSX CHPID types when z/VM is the controlling LP for dynamic I/O	Only to define, modify, and delete OSX CHPID types when z/VM is the controlling LP for dynamic I/O
CHPID type OSM	Only to define, modify, and delete CHPID type OSM, when z/VM is the controlling LP for dynamic I/O.	Only to define, modify, and delete CHPID type OSM, when z/VM is the controlling LP for dynamic I/O.	Only to define, modify, and delete CHPID type OSM, when z/VM is the controlling LP for dynamic I/O.
CP Assist for Cryptographic Function (CPACF)	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation

Support for z15	z/VM 6.4	z/VM 7.1	z/VM 7.2
Protected Key CP Assist for Cryptographic Function (CPACF)	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
Regional Crypto Enablement (RCE)	Yes (for guest exploitation)	Yes (for guest exploitation)	Yes (for guest exploitation)
Crypto Express7S Toleration (1 port)	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
Crypto Express7S Toleration (2 ports)	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
Crypto Express7S (1 port) supporting VISA Format Preserving Encryption	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
Crypto Express7S (1 port) supporting greater than 16 domains	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation and exploitation within the z/VM TLS/SSL server)	for guest exploitation
Crypto Express7S (1 port) Exploitation	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation and exploitation within the z/VM TLS/SSL server)	for guest exploitation and exploitation within the z/VM TLS/SSL server
Crypto Express7S (2 port) Exploitation	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation and exploitation within the z/VM TLS/SSL server
Crypto Express6S	for guest exploitation	for guest exploitation	for guest exploitation
Crypto Express5S	Yes (for guest exploitation)	for guest exploitation	for guest exploitation
Crypto Express6S toleration	for guest exploitation	for guest exploitation	for guest exploitation
Crypto Express6S and Crypto Express5S support of Visa Format Preserving Encryption	for guest exploitation	for guest exploitation	for guest exploitation
Crypto Express6S and Crypto Express5S supporting more than 16 domains	for guest exploitation	for guest exploitation	for guest exploitation
FICON Express16SE (CHPID type FCP) support of hardware data router	for guest exploitation	for guest exploitation	for guest exploitation

Support for z15	z/VM 6.4	z/VM 7.1	z/VM 7.2
FICON Express16SA (CHPID type FC) using native FICON or Channel-To-Channel (CTC)	Yes	Yes	Yes
FICON Express16SA (CHPID type FCP) for support of SCSI disks	Yes	Yes	Yes
FICON Express16SA (CHPID type FCP) support of hardware data router	for guest exploitation	for guest exploitation	for guest exploitation
FICON Express16SA (CHPID type FC), for support of zHPF single-track operations	Yes	Yes	Yes
FICON Express16SA (CHPID type FC), for support of zHPF multitrack operations	Yes	Yes	Yes
RoCE Express2 25 GbE for Ethernet communications including SingleRoot I/O Virtualization	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2 25GbE for TCP/IP	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2 10 GbE for Shared Memory Communications - Remote Direct Memory Access (SMC-R)	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2 25 GbE for Shared Memory Communication - Remote Direct Memory Access (SMC-R)	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2 for Ethernet communications including Single Root IO Virtualization (SR-IOV)	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2.1 10 GbE for Shared Memory Communication - Remote Direct Memory Access (SMC-R)	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2.1 10 GbE for Ethernet communications including SingleRoot I/O Virtualization	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2.1 10 GbE for TCP/IP	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2.1 25 GbE for Shared Memory Communication - Remote Direct Memory Access (SMC-R)	for guest exploitation	for guest exploitation	for guest exploitation

Support for z15	z/VM 6.4	z/VM 7.1	z/VM 7.2
RoCE Express2.1 25 GbE for Ethernet communications including SingleRoot I/O Virtualization	for guest exploitation	for guest exploitation	for guest exploitation
RoCE Express2.1 25 GbE for TCP/IP	for guest exploitation	for guest exploitation	for guest exploitation
Integrated Coupling Adapter 2 Fanout (ICA SR-2)	for guest exploitation	for guest exploitation	for guest exploitation
Support for 256 Coupling CHPIDs	for guest exploitation	for guest exploitation	for guest exploitation
Coupling Express LR (CE LR)	Only to define, modify, and delete CHPID type CL5, when z/VM is the controlling LP for dynamic I/O.	Only to define, modify, and delete CHPID type CL5, when z/VM is the controlling LP for dynamic I/O.	Only to define, modify, and delete CHPID type CL5, when z/VM is the controlling LP for dynamic I/O.
Asynchronous CF duplexing for lock structures	for guest exploitation	for guest exploitation	for guest exploitation
Simultaneous multithreading (SMT)	Yes	Yes	Yes
12x InfiniBand coupling links (12x IB-SDR or 12x IB-DDR)	Only to define, modify, and delete an InfiniBand coupling link, CHPID type CIB, when z/VM is the controlling LP for dynamic I/O	Only to define, modify, and delete an InfiniBand coupling link, CHPID type CIB, when z/VM is the controlling LP for dynamic I/O	Only to define, modify, and delete an InfiniBand coupling link, CHPID type CIB, when z/VM is the controlling LP for dynamic I/O
Long reach 1x InfiniBand coupling links (1x IB-SDR or 1x IB-DDR)	Only to define, modify, and delete an InfiniBand coupling link, CHPID type CIB, when z/VM is the controlling LP for dynamic I/O	Only to define, modify, and delete an InfiniBand coupling link, CHPID type CIB, when z/VM is the controlling LP for dynamic I/O	Only to define, modify, and delete an InfiniBand coupling link, CHPID type CIB, when z/VM is the controlling LP for dynamic I/O
Secure Key Advanced Encryption Standard (AES)	for guest exploitation	for guest exploitation	for guest exploitation
Support for z/VM-mode partition	Yes	Yes	Yes
Support for dynamic add of Reserved Central Storage	Yes	Yes	Yes
Support for z/VM Systems Management from the HMC	No	No	No
Support for installing Linux from the HMC	Yes	Yes	Yes
Dedicated OSA port to an operating system	Yes	Yes	Yes

Support for z15	z/VM 6.4	z/VM 7.1	z/VM 7.2
z/VM integrated systems management	Yes	Yes	Yes
HiperDispatch	Yes	Yes	Yes
zIIP Simulation on CPs (only on processor that support the specialty engine type)	Yes	Yes	Yes
Maximum number of CPs	64 cores when SMT is not enabled and 32 cores/64 threads when SMT is enabled	80 cores when SMT is not enabled and 40 cores/80 threads when SMT is enabled with applicable PTFs	80 cores when SMT is not enabled and 40 cores/80 threads when SMT is enabled with applicable PTFs
Maximum central storage	2 TB	2 TB	2 TB
Maximum number of channel paths	256	256	256
Maximum CFLEVEL supported	level 24 with applicable PTFs	level 24 with applicable PTFs	level 24
Support for multiple LCSSs	Yes	Yes	Yes
IFLs for OpenSolaris workloads	Yes	Yes	Yes
IFLs for Linux workloads	Yes	Yes	Yes
System-managed Coupling Facility structure duplexing, for z/OS guests	Yes	Yes	Yes
CHPID type OSD supporting Checksum offload for IPv6 packets	for guest exploitation	for guest exploitation	for guest exploitation
CHPID type OSD supporting Checksum offload for LPAR-to-LPAR traffic for IPv4 and IPv6 packets	for guest exploitation	for guest exploitation	for guest exploitation
CHPID type OSD supporting Large Send for IPv6 packets	for guest exploitation	for guest exploitation	for guest exploitation
CHPID type OSE supporting 4 or 2 ports per feature	Yes	Yes	Yes
CHPID type OSA performance enhancements	Yes (for guest exploitation)	Yes (for guest exploitation)	for guest exploitation
CHPID type FCP performance enhancements	Yes	Yes	Yes
Hardware Decimal Floating Point facilities	Yes	Yes	Yes
z/Architecture 64- bit addressing	Yes	Yes	Yes
Guest Coupling simulation	Yes	Yes	Yes

Support for z15	z/VM 6.4	z/VM 7.1	z/VM 7.2
Dynamic I/O configuration support through the CP configurability function	Yes	Yes	Yes
Shared Memory Communications - RDMA (SMC-R)	Yes (for guest exploitation)	Yes (for guest exploitation)	for guest exploitation
Shared Memory communications - Direct Memory Access (SMC-D)	Yes (for guest exploitation)	Yes (for guest exploitation)	for guest exploitation
Single Instruction Multiple Data (SIMC)	Yes (for guest exploitation)	Yes (for guest exploitation)	for guest exploitation
Performance assist by pass-through of adapter I/O operations and interruptions for CHPID types FCP, IQD, and OSD	Yes	Yes	Yes
Multi-VSwitch Link Aggregation	Yes	Yes	Yes
Guarded Storage	with applicable PTFs (for guest exploitation)	for guest exploitation	for guest exploitation
Instruction Execution Protection Facility	with applicable PTFs (for guest exploitation)	for guest exploitation	for guest exploitation
NVMe	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
CF Scalability Enhancements	Yes (for guest exploitation)	Yes (for guest exploitation)	for guest exploitation
CF List Notification Enhancements	Yes (for guest exploitation)	Yes (for guest exploitation)	for guest exploitation
Fair Latch Manager	Yes (for guest exploitation)	Yes (for guest exploitation)	for guest exploitation
CPU Measurement Facility	with applicable PTFs	with applicable PTFs	Yes
Compression/CPU Deflate and Asynchronous Compression	with applicable PTFs	with applicable PTFs	Yes
System Recovery Boost	No	with applicable PTFs	Yes
Quantum Safe support	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
Greater than 16 CEX adapters	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
Format preserving encryption	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
ECC curve support	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation

Support for z15	z/VM 6.4	z/VM 7.1	z/VM 7.2
Post Quantum support	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation
Sort Acceleration support for DFSORT	with applicable PTFs (for guest exploitation)	with applicable PTFs (for guest exploitation)	for guest exploitation

z/VSE

	z/VSE 6.2
Support for z15	Yes
OSA-Express7S GbE LX and GbE SX	
CHPID type OSD	with applicable PTFs
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	with applicable PTFs
OSA-Express7S 10 GbE LR and GbE SR	
CHPID type OSD	with applicable PTFs
OSA-Express7S 25 GbE SR	
CHPID type OSD	with applicable PTFs
OSA-Express7S 1000BASE-T Ethernet	
CHPID Type OSC supporting TN3270E and non-SNA DFT	with applicable PTFs
CHPID type OSD with exploitation of two ports per CHPID	with applicable PTFs
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	with applicable PTFs
CHPID type OSE support for 4 or 2 ports per feature	with applicable PTFs
OSA-Express6S 10 GbE LR and GbE SR	
CHPID type OSD	Yes
CHPID type OSX for access control to the intra-ensemble data network (IEDN) from Unified Resource Manager functions	Yes
OSA-Express5S 10 GbE LR and GbE SR	
CHPID type OSD	Yes
CHPID type OSX	Yes
OSA-Express6S GbE LX and GbE SX	
CHPID type OSD	Yes
OSA-Express5S GbE LX and GbE SX	
CHPID type OSD	Yes

	z/VSE 6.2
OSA-Express6S 1000BASE-T Ethernet	
CHPID type OSC, OSD, and OSE	Yes
OSA-Express5S 1000BASE-T Ethernet	
CHPID type OSC, OSD, and OSE	Yes
CP Assist for Cryptographic Function (CPACF)	Yes
Crypto Express7S Toleration (1 port)	with applicable PTFs
Crypto Express7S Toleration (2 ports)	with applicable PTFs
Crypto Express7S, Crypto Express6S, and Crypto Express5S supporting more than 16 Domains	with applicable PTFs
Crypto Express6S Toleration	Yes
Crypto Express5S Toleration	Yes
FICON Express16SA (CHPID type FC) using native FICON or Channel-To-Channel (CTC)	with applicable PTFs
FICON Express16SA (CHPID type FCP) for support of SCSI disks	with applicable PTFs
FICON Express16SA (CHPID type FC), for support of zHPF single-track operations	with applicable PTFs
FICON Express16SA (CHPID type FC), for support of zHPF multitrack operations	with applicable PTFs
CHPID type FC using native FICON or Channel-To-Channel (CTC)	with applicable PTFs
CHPID type FCP for support of SCSI disks	with applicable PTFs
CHPID type FC for support of zHPF single-track operations	with applicable PTFs
CHPID type FC for support of zHPF multitrack operations	with applicable PTFs
z/Architecture mode only	Yes
31- bit addressing	Yes
64 - bit addressing	Yes
Maximum central storage	32 GB
Maximum number of CPs	10
Support for dynamically increasing the number of logical cores, without an intervening IPL	Yes

TPF (Transaction Processing Facility)

Support for z15	**z/TPF 1.1**
Support for z15	Yes
OSA-Express7S GbE LX and GbE SX	
CHPID type OSD	with applicable PTFs
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	with applicable PTFs

Support for z15	z/TPF 1.1
OSA-Express7S 10 GbE LR and GbE SR	
CHPID type OSD	with applicable PTFs
OSA-Express7S 25 GbE SR	
CHPID type OSD	with applicable PTFs
OSA-Express7S 1000BASE-T Ethernet	
CHPID Type OSC supporting TN3270E and non-SNA DFT	Yes
CHPID type OSD with exploitation of two ports per CHPID	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes
OSA-Express6S 10 GbE LR and GbE SR	
CHPID type OSD	Yes
CHPID type OSX for access control to the intra-ensemble data network (IEDN) from Unified Resource Manager functions	Yes
OSA-Express5S 10 GbE LR and GbE SR	
CHPID type OSD	Yes
CHPID type OSX	Yes
OSA-Express4S 10 GbE LR and 10 GbE SR require at a minimum:	
CHPID type OSD	Yes
CHPID type OSX for access control to the intraensemble data network (IEDN) from zEC12 to Unified Resource Manager functions	at PUT 4 with APARs
OSA-Express6S GbE LX and GbE SX require at minimum:	
CHPID type OSD	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes
OSA-Express5S GbE LX and GbE SX require at minimum:	
CHPID type OSD	Yes
OSA-Express4S GbE LX and GbE SX require at minimum:	
CHPID type OSD with exploitation of two ports per CHPID	at PUT 4 with APARs
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes
OSA-Express6S 1000BASE-T Ethernet	
CHPID Type FC for support of zHPF single-track operations	with applicable PTFs
CHPID Type FC for support of zHPF multi-track operations	with applicable PTFs

Support for z15	z/TPF 1.1
CPU Measurement Facility	with applicable PTFs
Crypto Express7S Toleration (1 port)	Yes
Crypto Express7S Toleration (2 ports)	Yes
CP Assist for Cryptographic function (CPACF)	Yes
CHPID type OSD with exploitation of two ports per CHPID	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes
CHPID type FC using native FICON or Channel-To-Channel (CTC)	Yes
Maximum number of CPs (either shared or dedicated LP)	64
Maximum central storage	1 TB
Maximum CFLEVEL supported	9 (with APAR support)
Maximum number of channel paths	256
System Recovery Boost	with applicable PTFs

Linux

Support for z15	Linux
Support for z15	Yes
OSA-Express7S GbE LX and GbE SX	
CHPID Type OSD	Yes
CHPID Type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes
OSA-Express7S 10 GbE LR and GbE SR	
CHPID Type OSD	Yes
OSA-Express7S 25 GbE SR	
CHPID Type OSD	Yes
OSA-Express7S 1000BASE-T Ethernet	
OSA-Express6S 10 GbE LR and 10 GbE SR require at a minimum:	
CHPID type OSD	Yes
CHPID type OSX for access control to the intra-ensemble data network (IEDN) from Unified Resource Manager functions	Yes
OSA-Express5S 10 GbE LR and 10 GbE SR require at a minimum:	
CHPID type OSD	Yes
CHPID type OSX	Yes
OSA-Express6S 10 GbE LX and 10 GbE SX require at a minimum:	
CHPID type OSD	Yes

Support for z15	Linux
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes
OSA-Express5S 10 GbE LX and 10 GbE SX require at a minimum:	
CHPID type OSD	Yes
OSA-Express6S 1000BASE-T Ethernet require at a minimum:	
CHPID type OSC supporting TN3270E and non-SNA DFT	Yes
CHPID type OSD with exploitation of two ports per CHPID	Yes
CHPID type OSD without maximum port exploitation (one port on the PCIe adapter is available for use)	Yes
OSA-Express5S 1000BASE-T Ethernet require at a minimum:	
CHPID type OSD	Yes
CHPID type OSX for access control to the intra-ensemble data network (IEDN) from Unified Resource Manager functions	Yes
CHPID type OSM	Yes
CP Assist for Cryptographic Function (CPACF)	Yes
Crypto Express7S Toleration (1 port)	Yes
Crypto Express7S Toleration (2 ports)	Yes
Crypto Express7S (1 port) supporting greater than 16 domains	Yes
Crypto Express7S (1 port) Exploitation	Yes
Crypto Express6S toleration	Yes
Crypto Express6S support of greater than 16 Domains	Yes
CHPID type OSD supporting Checksum offload for IPv6 packets	Yes
CHPID type OSD supporting Checksum offload for LPAR-to-LPAR traffic for IPv4 and IPv6 packets	Yes
CHPID type OSD supporting Large Send for IPv6 packets	Yes
CHPID type FC using native FICON or Channel-To-Channel (CTC)	Yes
CHPID type FCP for support of hardware data router	Yes
CHPIDD type FCP for support of SCSI disks	Yes
CHPID type FCP support of hardware data router	Yes
CHPID type FC for support of zHPF single-track operations	Yes
CHPID type FC for support of zHPF multi-track operations	Yes
RoCE Express2.1 for Shared Memory Communications - Remote Direct Memory Access (SMC-R)	Yes
RoCE Express2.1 for Ethernet communication including Single Root IO Virtualization (SR-IOV)	Yes
RoCE Express2.1 for TCP/IP	Yes

Support for z15	Linux
RoCE Express2 for Shared Memory Communications - Remote Direct Memory Access (SMC-R)	Yes
RoCE Express2 for Ethernet communications including Single Root IO Virtualization (SR-IOV)	Yes
RoCE Express2 for TCP/IP	Yes
Linux-Only mode	Yes
Maximum number of CPs	64
Maximum central storage	2048 MB
Maximum number of channel paths	256
PKCS #11 API support	Yes
WLM Management of shared logical processors	Yes
Performance assist via pass through of adapter interruptions for FCP, IQD, and OSD CHPID types	Yes
Support for SSL clear key RSA operations	Yes
Support for CHPID type OSM for intranode management network (INMN)	Yes
zHyperLink Express	Yes
IBM Virtual Flash Memory	Yes
Transactional memory	Yes
Guarded storage	Yes
Instruction Execution Protection Facility	Yes

Hardware Configuration Definition (HCD)

You can use HCD in z/OS or z/VM to define configuration information both to the CPC and to the operating system. In z/OS, you can use the interactive panels of HCD or Hardware Configuration Manager (HCM) for definition; in z/VM, use HCM.

HCD allows you to dynamically change the current I/O configuration of the CPC. HCD allows you to dynamically change the current I/O configuration of the operating system and to create an IOCDS and make it the active IOCDS.

In z/OS, HCD is required to define the I/O configuration to the operating system. HCD is also the recommended way to define hardware configurations. HCD must be used, if the I/O configuration is dynamically changed in z/OS.

In z/VM, HCD is optional for defining the I/O configuration. If HCD is used for dynamic changes, the hardware configuration must be defined with HCD, the I/O configuration for the operating system may be defined with HCD.

HCD allows you to define the hardware and software I/O configuration information necessary for a parallel sysplex solution environment, including the capability to define:

- peer-mode channel paths (CE LR, ICA SR coupling links, and ICP) to connect z/OS systems to coupling facility images, and
- peer-mode channel paths (CE LR, ICA SR coupling links, and ICP) to connect coupling facility images to one another, in support of System-Managed CF Structure Duplexing.

In addition to these two uses, the external coupling links (CE LR, ICA SR coupling links) also support STP timing signals.

Additionally, HCD in z/OS allows you to remotely write IOCDSs from one Support Element to another Support Element as long as both Support Elements are powered-on, LAN-attached, enabled for remote licensed internal code (LIC) update, and defined to the same Hardware Management Console.

Dynamic I/O configuration does **not** support:

- Adding or deleting LPs
- Changing MIF image ID numbers (the MIF image ID number is different from the LP identifier [ID])

When using HCD, you can define and control the configuration of the CPC affecting all LPs. Those LPs that run with HCD or z/VM can dynamically change their software configuration definitions. Other LPs might require an IPL in order to use the new configuration.

HCD allows you to discover and automatically define switched FICON attached storage control units and devices.

When you use HCD you must install, in the LP, the appropriate version of IOCP. Throughout the remainder of this publication, all the capabilities or restrictions documented regarding the IOCP program, also apply to definitions entered and controlled through HCD.

For more information about dynamic I/O configuration on z/OS, see:

- *z/OS Hardware Configuration Definition Planning*, GA32-0907
- *z/OS Hardware Configuration Definition User's Guide*, SC34-2669

For more information about dynamic I/O configuration on z/VM, see:

- *z/VM I/O Configuration*, SC24-6291

z/VM dynamic I/O configuration

You can dynamically change the current I/O configuration of the CPC. You can also change the current I/O configuration of the operating system and create an IOCDS and make it the active IOCDS.

Dynamic I/O configuration does **not** support:

- Adding or deleting LPs
- Changing MIF image ID numbers (the MIF image ID number is different from the LP identifier (ID)).

You can define and control the configuration of the CPC affecting all LPs. Those LPs that run z/VM can dynamically change their software configuration definitions.

Input/Output Configuration Program (IOCP) support

To perform a power-on reset you must use an LPAR IOCDS. To generate an LPAR IOCDS you need to use the ICP IOCP program.

PTFs for supported IOCP versions must be applied and can be obtained from the Software Support Center. For more information about ICP IOCP, see *Input/Output Configuration Program User's Guide for ICP,* SB10-7172.

Hardware support

LPs operate independently but can share access to I/O devices and CPC resources. Each active LP must have sufficient channel paths and storage to meet the particular requirements of that LP. Additional central storage, channel paths, consoles, and other I/O devices might be necessary for the planned configuration of LPs.

Operator training

A general knowledge of z/Architecture is useful and, in some cases, required of all technical support personnel, PR/SM planners or LP planners. Generally, the operator performs the following tasks:

Editing activation profiles

You can edit reset, image, and load profiles for configurations using the **Customize/Delete Activation Profiles** task.

Reset profiles

Use the reset profile to:

- Select an IOCDS
- Optionally specify an LP activation sequence
- Enable I/O Priority Queuing.

You can select an IOCDS using the General page of the reset profile. To specify an optional LP activation sequence, use the Partitions page of the reset profile.

Image profiles

Use the image profile to:

- Define LP characteristics
- Optionally specify automatic load settings.

To define LP characteristics, use the General, Processor, Security, Storage, Options, Crypto, and Time Offset pages of the image profile (see "Defining logical partitions" on page 130). To specify optional automatic load settings, use the Load page of the image profile (see "Load information" on page 146).

Load profiles

If you are not using the image profile to specify load options for an LP, use the load profile.

Activating a CPC

To activate a CPC, locate the CPC and open the **Activate** task. This task will activate the hardware system and, if LP activation sequence is enabled for LPs, will activate those LPs in the order specified in the reset profile for the CPC. You can also automatically load the LP'ss operating system as specified in the image profile for the LP.

Activating an LP

To activate an LP locate the partition and open the **Activate** task. This task will activate the LP and, if automatic load is specified for an LP, will automatically load the LP's operating system as specified in the image profile for the LP.

Performing a load on an LP or activating a load profile

Perform a load on an LP or activate a load profile for an LP by locating the LP for a previously activated LP and opening the **Customize/Delete Activation Profiles** task.

- Select or, if necessary, customize or create a load profile for the LP.
- Assign the load profile to the LP's activation profile, save your changes, and exit the **Customize/Delete Activation Profiles**.
- Open the **Activate** task available from the Daily Tasks list.

In recovery situations, you can locate the LP and open the **Load** task.

Deactivating a logical partition

To deactivate a logical partition locate the logical partition and open the **Deactivate** task. This task deactivates the logical partition and any operating system running in the logical partition.

Locking and unlocking a logical partition

Lock a logical partition by selecting the CPC image representing the logical partition. Select the **Image Details** task, then select the **Lockout disruptive tasks** radio button to **Yes** and click **Apply**.

You can use this same procedure to unlock a logical partition by selecting the **Lockout disruptive tasks** radio button to **No** and click **Apply**.

Locking a logical partition can prevent accidentally performing disruptive tasks on a logical partition.

Deactivating a CPC

To deactivate a CPC locate the CPC and open the **Deactivate** task. This task will deactivate the CPC and any activated LPs and their associated operating systems.

Logical partitions

This section provides an overview of LP characteristics. Some of the characteristics described in this section are model-dependent, EC-level dependent, MCL-dependent, LP mode dependent, or control-program dependent. For this reason, all the characteristics described here are not necessarily available on all CPCs.

The resources of a CPC can be distributed among multiple control programs that can run on the same CPC simultaneously. Each control program has the use of resources defined to the logical partition in which it runs.

You can define an LP to include:

- One or more CPs
- Central storage
- Channel paths
- Two or more optional cryptos (Crypto Express). A single crypto engine can be defined, for test purposes, but it is not recommended for production LPs.

An LP can be defined to include CPs, zIIPs, ICFs, and IFLs. The allowed combinations of defined processor types for an LP depends on the defined mode of the logical partition. Refer to Table 16 on page 136.

You can also define an LP to be a coupling facility running the coupling facility control code.

Characteristics

LPs can have the following characteristics. For more information or details about exceptions to any of these characteristics, see "Determining the characteristics" on page 92.

- The maximum number of LPs you can define on a z15 is 85 for models T01/LT1 and 40 for models T02/LT2.
- LPs can operate in **General** , **Linux-Only**, **z/VM**, **Coupling facility**, **SSC** mode.
- The storage for each LP is isolated. Central storage cannot be shared by LPs.
- Using dynamic storage reconfiguration, an LP can release storage or attach storage to its configuration that is released by another LP.
- All channel paths can be defined as reconfigurable. Channel paths are assigned to LPs. You can move reconfigurable channel paths between LPs using tasks available from either the Hardware Management Console or the Support Element console. If the control program running in the LP supports physical channel path reconfiguration, channel paths can be moved among LPs by control program commands without disruption to the activity of the control program.
- MIF allows channel paths to be shared by two or more LPs at the same time. All CHPID types **except** CVC and CBY can be shared.
- The total number of logical processors that can be defined is 190.
- CPs can be dedicated to LPs or shared by them. CPs that you define as dedicated to an LP are not available to perform work for other active LPs. The resources of shared CPs are allocated to active LPs as needed. You can cap (limit) shared CP resources, if required.

Figure 1 on page 28 shows some of the characteristics that can be defined for an LP. You can view each LP as a CPC operating within the physical CPC.

- You cannot define a mix of shared and dedicated CPs for a single LP. CPs for an LP are either all dedicated or all shared. However, you can define a mix of LPs with shared CPs and LPs with dedicated CPs and activate them concurrently.

- For security purposes, you can:
  - Reserve reconfigurable channel paths for the exclusive use of an LP (unless overridden by the operator)
  - Limit the authority of an LP to read or write any IOCDS in the configuration and limit the authority of an LP to change the I/O configuration dynamically
  - Limit the authority of an LP to retrieve global performance data for all LPs in the configuration
  - Limit the authority of an LP to issue certain control program instructions that affect other LPs

- A coupling facility LP has a maximum of 16 processors (ICFs or CPs), regardless of the model.
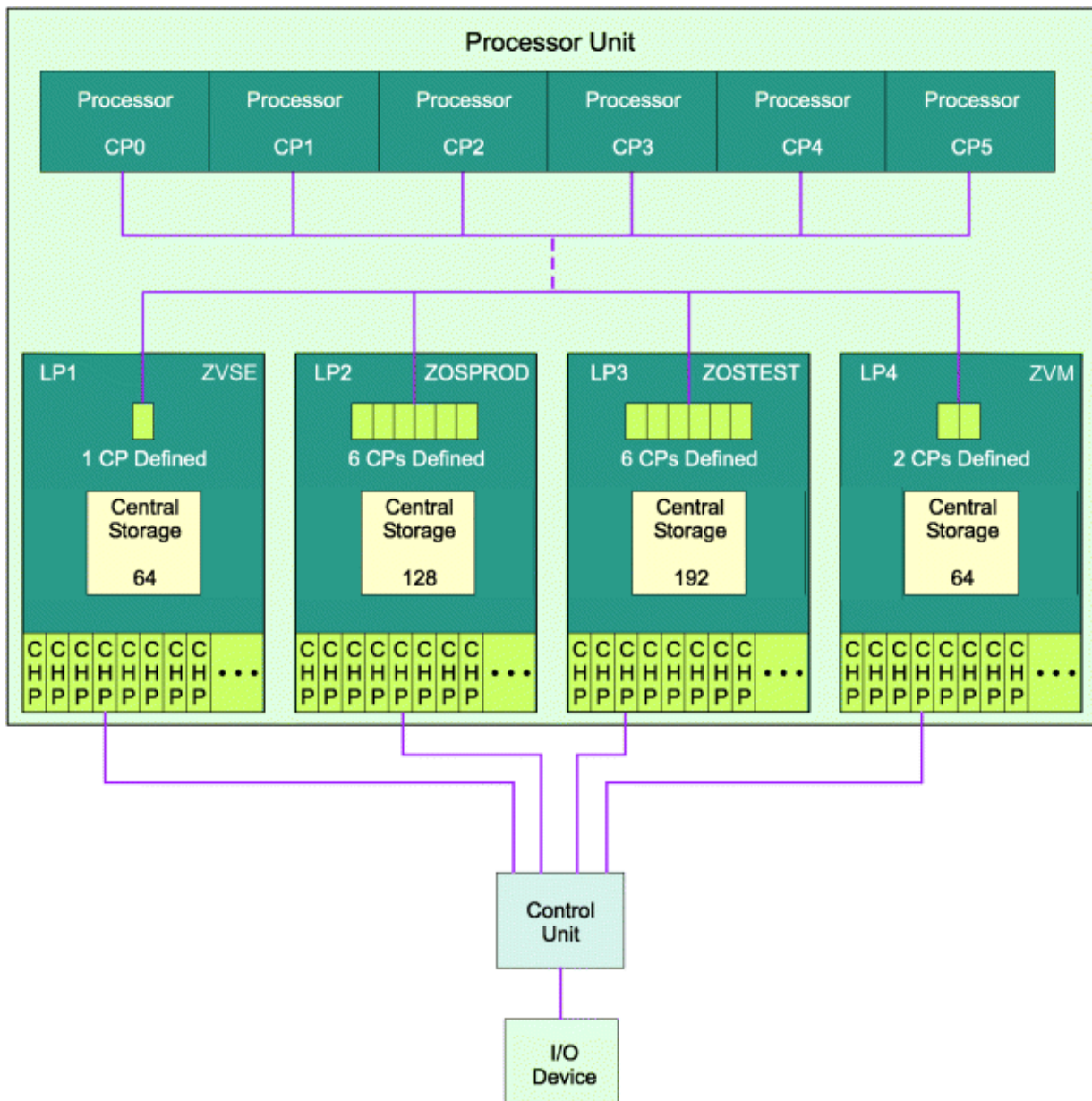

Figure 1. Characteristics of logical partitions

Potential applications

The use of LPs allows multiple systems, including the I/O for the systems, to be migrated to a single CPC while maintaining the I/O performance, recovery, and multi-pathing capability of each system, and with minimum impact to the system generation procedures.

LPs are suitable for consideration in the following environments:

Consolidation
Multiple production system images can be consolidated onto 1 CPC without having to merge them into one image.

Migration
Control programs or applications can be migrated by running the old and new systems or applications in independent LPs that are active on the same CPC at the same time.

Production and test
Multiple production and test systems can run on the same CPC at the same time.

Coupling facility
A coupling facility enables high performance, high integrity data sharing for those CPCs attached to it and configured in a sysplex.

Coupled systems
Multiple instances of the same workload can be run in multiple LPs on one or more CPCs as part of a sysplex configuration that takes advantage of the centrally accessible, high performance data sharing function provided by the coupling facility.

Extended Recovery Facility (XRF)
Primary and alternate XRF systems can run on 1 CPC. Multiple and alternate XRF systems can run on 1 CPC.

Communications Management Configuration (CMC)
The communications management configuration (CMC) machine, typically run on a separate CPC, can be run as an LP on the same CPC.

Departmental systems
Multiple applications can be isolated from one another by running each in a separate LP.

Constrained systems
Those systems that cannot fully use a large system because of storage constraints can alleviate the problem by using LPs to define multiple system images on the same CPC.

Diverse workloads
Interactive workloads such as the Customer Information Control System (CICS®) and time-sharing option (TSO) can be isolated by running each in a separate LP.

Compatibility and migration considerations

This section provides migration and compatibility information for a z15.

Device numbers

When multiple systems are migrated to a z15, the combination of systems could include different devices or shared devices with identical device numbers. Each system can operate in an LP without changing the device numbers as long as identical device numbers do not occur in the same LP. However, duplicate device numbers can exist in the same LP if these device numbers are in different subchannel sets.

Duplicate device number conflicts can occur when the I/O configuration is reconfigured. For example, if a reconfigurable channel path is reassigned to another LP and devices attached to the channel path have device numbers that are already assigned in the receiving LP to other online channel paths, a conflict results. When IOCP generates an LPAR IOCDS, the initial configuration contains no duplicate device number conflicts in an LP.

Device number conflicts are also detected when operator tasks change the I/O configuration (channel path tasks from the Hardware Management Console or Support Element console; or control program configuration command) or during LP activation.

Duplicate device number conflicts are also detected when a dynamic I/O configuration change is made.

Multiple Subchannel Sets (MSS)

The Multiple Subchannel Sets (MSS) structure allows increased device connectivity for Parallel Access Volumes (PAVs). Four subchannel sets per Logical Channel Subsystem (LCSS) are designed to enable a total of 63.75K subchannels in set-0 and the addition of 64K - 1 subchannels in set-1, set-2, and in new set-3. MSS is supported by FICON (CHPID type FC (both native FICON and zHPF paths) and z/OS.

Control programs

PTFs for supported control programs must be applied and can be obtained from the Software Support Center. A supported control program operates in an LP as it does in one of the basic modes, with the following exceptions:

z/OS

- Physical reconfiguration, either offline or online, of CPs is not supported on the system. Logical core reconfiguration, either offline or online, is supported in an LP. It does not affect the online/offline status of the physical cores. To reconfigure a logical core offline or online, use the following z/OS operator command:

```
CF CPU(x),<OFFLINE/ONLINE>
```

- Physical reconfiguration, either offline or online, of central storage is supported.

  To reconfigure a central storage element offline or online, use the following z/OS operator command:

```
CF STOR(E=1),<OFFLINE/ONLINE>
```

  Additionally you can use the following command to reconfigure smaller amounts of central storage online or offline:

```
CF STOR(nnM),<OFFLINE/ONLINE>
```

  Reconfigurable Storage Unit (RSU) Considerations: Set the RSU parameter to the same value that you specified in the central storage Reserved field divided by the storage granularity for your logical partition size (see Table 12 on page 94. See *z/OS MVS Initialization and Tuning Reference* for appropriate RSU parameter syntax.

- Reconfiguration, either offline or online, of channel paths by z/OS operator commands is supported on a Z mainframe. This capability also allows channel paths to be moved among LPs using z/OS operator commands.

- Preferred paths to a device are supported on a Z mainframe. If the preferred path parameter is specified in an LPAR IOCDS, it is accepted.

- Specifying SHAREDUP for devices is not recommended. If used, z/OS treats the device as a SHARED device.

- Each z/OS LP can run the Resource Measurement Facility (RMF). RMF enhancements for PR/SM allow a single LP to record system activities and report them in the Partition Data Report. To enable this function, use the **Change LPAR Security** task and select **Performance Data Control** for the LP.

- For z/OS, RMF reporting includes LPAR Management Time.

- RMF provides enhanced reporting for coupling facility configurations.

- RMF with APAR support identifies which logical and physical cores are of each type when any combination of general purpose IFL, zIIP, and ICF processors are present in the configuration on its partition data report. You can have up to 32 virtual machines running z/OS within a simulated sysplex, with each z/OS virtual machine coupled to up to eight virtual machines running as coupling facilities.

EREP

Each control program operating in an LP has its own environmental recording, editing, and printing (EREP) processing. EREP records for ICP channel paths go to the z/OS logs of the z/OS systems attached to a coupling facility LP.

CPU IDs and CPU addresses

Check application packages and software products that are licensed to run under specific CPU identification (CPU ID) information, because they might need to be updated.

CPU ID information is system-generated for each logical core in the LP during the LP activation. It consists of a version code for the CPC machine type, a CPU identification number that is unique for each logical partition, a model number for the CPC machine type, and a value of X'8000'.

The Store CPU ID (STIDP) instruction stores the CPU ID for each logical core in storage in the following format (Figure 2 on page 31):

0	7 8	31 32	47 48	63
Version Code	CPU Identification Number	Machine Type	8000	

Figure 2. CPU ID format

Figure 3 on page 31 shows the format of the CPU identification number (bits 8 through 31 of the CPU ID format).

PPnnnn
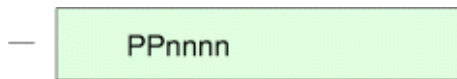
Legend :

P Logical partition identifier
n Digit derived from the serial number

Figure 3. CPU identification number format

CPU ID fields

The CPU identification number, with the version code and the machine type, permits a unique CPU ID for each logical partition.

- The **version code** for the system is always zero and is not affected by the operating mode.
- The **CPU identification number** for each logical core (see Figure 3 on page 31) consists of a two-digit LP identifier, and digits derived from the serial number of the CPC.
  - The **logical partition identifier** is specified using the Partition identifier field on the General page in either the reset or image profile used by the LP and must be unique for each active LP in the configuration.
- The following machine types (CPC model numbers) are returned as indicated:

Table 3. Machine types and models	
Machine type	**Models**
8562	z15 / LinuxONE III (T02 and LT2)
8561	z15 / LinuxONE III (T01 and LT1)

Note: STIDP is provided for purposes of compatibility with earlier version. It is recommended that you use the Store System Information instruction (STSI), rather than STIDP. STSI is the preferred means to

obtain all CPU information including machine serial number. When a unique logical CPU address is all that is required, use the Store CPU Address (STAP) instruction.

Examples of CPU ID information

The following examples show the format and contents of the CPU ID information stored by the STIDP instruction for logical cores in active LPs. Table 4 on page 32 shows the CPU ID information for a z15 with 3 active LPs.

Table 4. CPU IDs			
LP name	**LP identifier**	**Number of CPs defined**	**CPU ID returned by STIDP**
ZVSE	1	1	00 019999 8561 8000
ZOSTEST	2	1	00 029999 8561 8000
ZOSPROD	3	8	00 039999 8561 8000 00 039999 8561 8000 00 039999 8561 8000 00 039999 8561 8000 00 039999 8561 8000 00 039999 8561 8000 00 039999 8561 8000 00 039999 8561 8000

HSA allocation

The z15 Hardware System Area (HSA) has a fixed size of 256 GB for models T01 and LT1 and a fixed size of 160 GB for models T02 and LT2.

TOD clock processing

The CPC TOD clocks of all the CPs are automatically set during CPC activation. The time reference used depends on whether Server Time Protocol (STP) is enabled. When STP is enabled, a CPC can participate in a STP CTN. In this case the Current Time Server for the STP CTN provides the time information.

Server Time Protocol not enabled

During PR/SM initialization, the CPC TOD clocks for each CP are set to the TOD value of the Support Element. Each LP starts out with this CPC TOD value at the completion of LP activation. The operating system running in an LP can set a TOD value for itself and this is the only TOD reference it sees. Setting the TOD clock for one logical core in the LP sets the TOD clock for all logical cores in that LP, but does not affect the logical cores in any other LP. The TOD clock value is used for the duration of the LP activation, or until a subsequent Set Clock instruction is issued in the LP.

Server Time Protocol enabled

The enablement of STP is supported. Also, during PR/SM initialization, when STP is enabled, the CPC TOD clocks for each CP are set to the TOD value from STP.

The operating system in each LP can independently choose whether to synchronize to the current time source for STP, if present. Operating systems in LPs that do synchronize to STP run with identical TOD values. Operating systems in LPs that do not synchronize to STP do not need to be aware of the presence of STP and can set their TOD values independently of all other LPs.

z/OS does not allow you to change the value of the TOD setting when synchronized to STP (STPMODE=YES in the CLOCKxx parmlib member).

The z15 supports the specification of a logical partition time offset. When all members of a sysplex are in logical partitions on these supported models, the logical partition time offset can be used for:

- Different local time zone support in multiple sysplexes using the STP Coordinated Timing Network (CTN). Many sysplexes have the requirement to run with a LOCAL=UTC setting in a sysplex (STPMODE=YES) where the time returned from a store clock (STCK) instruction yields local time. To fulfill this requirement, the time initialized for the STP CTN must be local time. With logical partition time offset support, multiple sysplexes can each have their own local time reported to them from a STCK instruction if wanted. For instance, the STP CTN can be set to GMT, one set of sysplex partitions could specify a logical partition time offset of minus 5 hours, and a second set of sysplex partitions could specify a logical partition time offset of minus 6 hours.

External coupling links are also valid to pass time synchronization signals for Server Time Protocol (STP). Therefore the same coupling links can be used to exchange timekeeping information and Coupling Facility messages in a Parallel Sysplex.

Sysplex testing without Server Time Protocol enabled

You can do sysplex testing without Server Time Protocol enabled by setting up a test sysplex of several z/OS LPs in multiple LPs in the same PR/SM configuration. Use the SIMETRID keyword in the CLOCKxx parmlib member for z/OS to synchronize the members of the sysplex in the LPs.

Synchronized Time Source and the coupling facility

Improved processor and coupling facility link technologies inherent on the Z mainframe necessitate more rigorous time synchronization tolerance for members of a parallel sysplex hosted by those models. To help ensure that any exchanges of time-stamped information between members of a sysplex observe the correct time ordering, time stamps are now included in the message-transfer protocol between the systems and the coupling facility.

Consequently, a coupling facility hosted by any Z mainframe requires connectivity to the same synchronized time source as the other z/OS systems in its parallel sysplex. If a member of its parallel sysplex is on the same server as the coupling facility, required connectivity is already provided to the synchronized time source. However, when a coupling facility is a resident of a Z mainframe, which **does not** include a member of the coupling facilities parallel sysplex, connectivity attached to the synchronized time source must be implemented.

STP CTN Split and Merge

STP CTN split and merge is a new sysplex timing capability for availability that allows two distinct timing networks to be merged into one or allows you to split one timing network into two, nondisruptively. Previously, these timing network reconfigurations and transitions were disruptive to the running sysplex(es) operating within the CTN(s). This feature is especially helpful when working to combine or redistribute servers within a corporate structure.

Extended TOD-clock facility

The extended TOD-clock facility provides an extended form TOD clock and a TOD programmable register. The extended form TOD clock is a 128- bit value that extends the current basic form by appending 8 bits on the left and 56 bits on the right. The extended form TOD clock is returned by a problem-program instruction, STORE CLOCK EXTENDED (STCKE). The contents of the TOD programmable register are stored into the rightmost portion of the extended form TOD value when the TOD clock is inspected by STCKE. A TOD programmable register exists for each CPU and contains the TOD programmable field in bits 16-31. The TOD programmable register is set by a new privileged instruction, SET TOD PROGRAMMABLE FIELD (SCKPF). The leftmost byte of the extended form TOD clock is the TOD Epoch Index (TEX), and is stored as zeros in machines running General.

The extended TOD clock facility satisfies three main objectives:

- Relieve constraints that exist in the current 64- bit TOD clock
- Extend the TOD-clock architecture to multi-system configurations

- Help ensure sysplex-wide uniqueness of the STCKE TOD values

The TOD Programmable Field (TODPF) is a 16- bit quantity contained in bit positions 16-31 of the TOD programmable register. The contents of the register can be set by the privileged instruction SET TOD PROGRAMMABLE FIELD. The contents of the register can be stored by the instruction STORE CLOCK EXTENDED, which stores the TOD programmable field in the last 16 bits of the extended form TOD clock. The contents of the register are reset to a value of all zeros by an initial CPU reset.

Clock Comparator on Shared Processors

The clock comparator has the same format as bits 0-63 of the TOD clock. The clock comparator nominally consists of bits 0-47, which are compared with the corresponding bits of the TOD clock. On some models, higher resolution is obtained by providing more than 48 bits. In most cases, a logical processor running in a logical partition receives the model's resolution for the clock comparator.

However, when using shared logical processors in a logical partition, if the operating system running in a logical partition loads an enabled wait state with a clock comparator set on that logical processor, the PR/SM Hypervisor tracks that clock comparator value for the logical partition's processor at a less granular resolution. The granularity can be reduced to as little as bits 0-45 of the intended clock comparator value. This effect is not seen on dedicated logical processors nor is it seen on logical processors that are not in wait state.

Chapter 2. Planning considerations

This chapter describes planning considerations for I/O configuration and for coupling facility logical partitions.

Planning the I/O configuration

This section describes the planning considerations and guidelines for creating an IOCDS. It assumes that you understand the IOCP configuration and coding requirements described in the *Input/Output Configuration Program User's Guide for ICP*, SB10-7172.

Control program support

The maximum number of supported devices is limited by the control program. In planning an I/O configuration, determine the maximum number of devices supported by the control program run in each LP. See the documentation for the respective operating systems.

Hardware Configuration Definition (HCD) support

HCD supports definition of the I/O configuration for an entire installation. It is required for parallel sysplex and LPAR clusters. A single I/O data file is created for the installation and used for multiple machines and I/O configuration data sets.

HCD supports:

- Up to 85 logical partitions (LPs) for models T01/LT1 and up to 40 logical partitions for models T02/LT2 per central processing complex (CPC) on a z15.
- Coupling facility configurations
- Multiple Image Facility (MIF)
- Dynamic CHPID Management (DCM) channel paths.
- MCSS
- Assigning reserved logical partitions a meaningful name

Table 5. HCD function support		
HCD Function	**z/OS**	**z/VM**
Define 85 logical partitions?	Yes	Yes
Define shared channel paths?	Yes	Yes
Define coupling facility channel paths?	Yes	Yes (Note 2)
Define dynamically managed channel paths?	Yes	Yes (Note 2)
Write IOCDSs remotely?	Yes (Note 1)	No
Access I/O devices on shared channel paths?	Yes	Yes
Use software-only dynamic I/O?	Yes	Yes
Use hardware and software dynamic I/O?	Yes	Yes
Defined shared FICON CTC?	Yes	Yes
Use hardware-only dynamic I/O	Yes	No

Table 5. HCD function support (continued)		
HCD Function	**z/OS**	**z/VM**
Notes: 1. HCD, running on z/OS, allows you to remotely write IOCDSs from one CPC to another CPC that is powered on, LAN-attached, enabled for remote LIC update, and defined to the same Hardware Management Console. 2. HCD, running on z/VM, allows you to define coupling facility channel paths or dynamically managed channel paths for a z/OS LP but z/VM does not support coupling facility channel paths or dynamically managed channel paths for use by z/VM or guest operating systems. z/VM does virtualize the coupling facility for guest test purposes.		

For more information about using HCD with Multiple Image Facility, see

- *z/OS Hardware Configuration Definition: User's Guide*, SC34-2669
- *z/OS Hardware Configuration Definition Planning*, GA22-7525
- *z/VM I/O Configuration*, SC24-6291.

z/VM Dynamic I/O configuration support

This section described z/VM dynamic I/O configuration support for the Z mainframe.

z/VM support for the coupling facility

z/VM allows you to define configurations that use the coupling facility. However, z/VM does ***not*** support the coupling facility itself. (z/VM does virtualize the coupling facility for guest test purposes). Instead, the dynamic I/O configuration capability available on z/VM allows you to define resources that can be used by a z/OS system in another LP. For a summary of the support of dynamic I/O configuration on z/VM, see .

z/VM support for the Multiple Image Facility (MIF)

You can use z/VM to define shared channel paths. For a summary of z/VM support of dynamic I/O configuration, see .

Table 6. z/VM dynamic I/O support for MIF and the coupling facility	
	Release
z/VM Function	**6.4, 7.1, and 7.2**
Define shared channel paths?	Yes
Define coupling facility channel paths?	Yes (Note)
Write IOCDSs remotely?	No
Access I/O devices on shared channel paths?	Yes
Use software-only dynamic I/O?	Yes
Use hardware and software dynamic I/O?	Yes
Note: z/VM can define coupling facility channel paths for a z/OS LP but does ***not*** support real coupling facility channel paths for use by z/VM or guest operating systems. z/VM does virtualize the coupling facility for guest test purposes.	

Input/Output Configuration Program (IOCP) support

You can create up to four IOCDSs. ICP IOCP is the required supported version. You can define as many as 85 LPs for models T01 and LT1 is 85 and 40 LPs for models T02 and LT2. For more information about ICP IOCP, see *Input/Output Configuration Program User's Guide for ICP*, SB10-7172.

Characteristics of an IOCDS

The definitions for channel paths, control units, I/O devices, and PCIe functions are processed by the IOCP and stored in an IOCDS. During initialization of the CPC, the definitions of a selected IOCDS are transferred to the hardware system area (HSA). The IOCDS is used to define the I/O configuration data required by the CPC to control I/O requests.

Channel paths in an IOCDS are assigned to one or more LPs. The characteristics of an IOCDS are:

- Using the IOCP RESOURCE statement, you define logical channel subsystems (CSSs) and the logical partitions that have access to the channel paths in a CSS.
- Using the IOCP RESOURCE statement, you can name logical partitions and assign MIF image ID numbers to them. MIF image ID numbers are necessary for FICON CTC definitions.
- Using the IOCP CHPID statement, you can assign a channel path as reconfigurable or shared.
- Using the IOCP CHPID statement, you can specify a Dynamic CHPID Management (DCM) channel path and the cluster to which the CHPID belongs. The CHPID is shareable among active LPs that have become members of the specified cluster.
- You can duplicate device numbers within a single IOCP input file, but the device numbers cannot be duplicated within an LP. See "IOCP coding specifications" on page 52.

Maximum number of logical partitions

The maximum number of LPs supported by the z15 for models T01 and LT1 is 85 and the maximum number for models T02 and LT2 is 40.

Determining the size of the I/O configuration

To determine the size of the current I/O configuration (number of control unit headers and devices), review the IOCDS Totals Report for the current IOCDS.

Maximum size of the I/O configuration

Limits within an I/O configuration exist for the following:

- Devices
- Control unit headers
- Physical control units

Z models

- The maximum number of control unit headers (CUHs) is 4096 per logical channel subsystem (CSS).
- The maximum number of physical control units is 8192.
- The maximum number of devices is 65280 per CSS for subchannel set 0.
- The maximum number of devices is 65535 per CSS for subchannel set 1.
- The maximum number of devices is 65535 per CSS for subchannel set 2.
- The maximum number of devices is 65535 per CSS for subchannel set 3, if the subchannel set is supported by the model.

Guidelines for setting up the I/O configuration

Follow these guidelines when setting up an I/O configuration.

1. Determine the number of LPs and in which logical channel subsystem (CSS) they exist.

2. For dynamic I/O configurations, include any logical partitions for which you do not yet have a meaningful name. These logical partitions are reserved until a subsequent dynamic I/O configuration change is made to assign them a name.

3. Determine if you want to move any channel paths among LPs. If you do, then these channel paths must be defined as reconfigurable in the IOCP CHPID statement. You cannot move a channel path from an LP in one CSS to an LP in another CSS.

4. Determine if you want to share any channel paths among LPs in the same CSS. If you do, then specify these channel paths as SHARED in the IOCP CHPID statement. This specification helps reduce the number of channel paths configured to a physical control unit and device. Make sure that the channel path type supports being shared.

5. Determine if you want to share any channel paths among LPs in different CSSs. If you do, then define these channel paths as spanned by specifying multiple CSS IDs in the PATH keyword of the IOCP CHPID statement. This specification further helps reduce the number of channel paths configured to a physical control unit and device. Make sure that the channel path type supports being spanned.

6. Within each LP, configure primary and backup paths from separate channel adapter cards.

7. Within each LP, configure primary and backup paths from separate self-timed interfaces (STIs).

Recovery considerations

When planning for recovery, consider the following I/O configuration guidelines.

- Assign channel paths to LPs as described in "Guidelines for setting up the I/O configuration" on page 37.
- Review the recoverability characteristics of the I/O configuration described in the section "Shared devices" on page 55.

Managing logical paths for FICON channels

This section describes logical paths, explains overall system considerations, and makes specific configuration recommendations.

Definition

A logical path is a logical connection between a control unit and a FICON channel (TYPE=FC). Logical paths are important because each sharing LP on a CPC requires that a logical path is established between a FICON channel and a control unit for I/O operations to occur.

Logical paths do **not** exist for coupling facility channel paths, CE LR, ICA SR coupling links, internal coupling channel paths (TYPE=ICP), Open Systems Adapter channel paths (TYPE=OSC, TYPE=OSD, TYPE=OSE, TYPE=OSM, TYPE=OSN, or TYPE=OSX), internal queued direct communication (HiperSockets) channel paths (TYPE=IQD), or fibre channel protocol channel paths (TYPE=FCP).

Control unit allocation of logical paths

Control units allocate logical paths to channels dynamically on a first-come-first-served basis. Control units do not manage the allocation of logical paths but instead allow channels to compete for logical paths until all the logical paths of the control unit are used.

Why manage logical paths?

The FICON environments (the use of FICON Express channels and FICON Directors) greatly enhances the connectivity potential for control units. In addition, you can define shared channels that can request additional logical paths. However, control units can only allocate a limited number of logical paths in relation to the number of logical paths that channels can request. In configurations where channels request more logical paths than a control unit can allocate, you must manage logical paths to help ensure that the I/O operations you intend take place.

The FICON Express offers increased connectivity in the same amount of physical space, and offer the possibility of increased performance. Up to 384 FICON Express8, Express16, Express16S, Express16S+, and Express16SA channels on a z15 model and can be employed to greatly expand connectivity and

throughput capability. The FICON connectivity solution is based on industry-standard Fibre Channel technology and uses our exclusive native FICON architecture. For detailed information, see *Input/Output Configuration Program User's Guide for ICP,* SB10-7172.

MIF example

Figure 4 on page 39 shows a FICON shared channel configuration on an MIF-capable CPC. In this example, all five LPs share each of four FICON channels attached to a DS8K. Each shared FICON channel represents five channel images corresponding to the five LPs. Each channel image requests a logical path to the DS8K. Again, you can avoid this situation by managing logical paths.
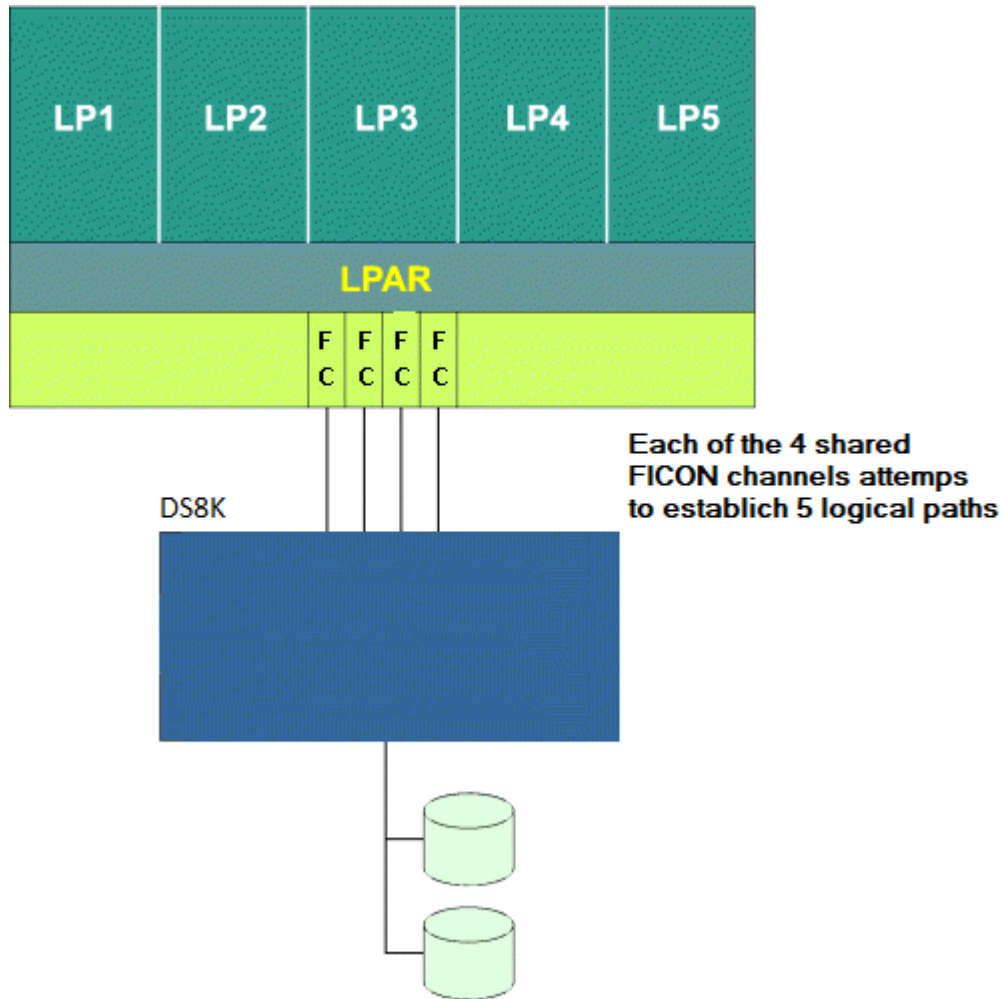

Figure 4. A shared FICON configuration that can benefit from better logical path management

Managing the establishment of logical paths

You can manage the establishment of logical paths between channels and control units. With proper planning, you can create I/O configuration definitions that allow control units in the configuration to allocate logical paths for every possible request made by channels in either of the following ways:

• Create a one-to-one correspondence between the logical path capacity of all control units in the physical configuration and the channels attempting to request them.

• Create I/O configurations that can exceed the logical path capacity of all or some of the control units in the physical configuration but, at the same time, provide the capability to selectively establish logical connectivity between control units and channels as needed.

This capability can be useful or even necessary in several configuration scenarios. See "Recommendations" on page 41.

Logical path considerations

You can better understand how to manage the establishment of logical paths by understanding the following:

- Control unit considerations
- Connectivity considerations
- Channel configuration considerations

Control unit considerations

Consider the following factors concerning the allocation of logical paths by control units:

- Control units allocate logical paths dynamically on a first-come-first-served basis.

  Control units do not manage the allocation of logical paths but instead allow channels to compete for logical paths until all the logical paths of the control unit are used.

- Control units vary in the number of logical paths they support.

Connectivity considerations

FICON system hardware, CPCs, and FICON Directors significantly affect the volume of logical path requests to a control unit as follows:

- Control units can attach to one or more ports on a Director or to additional ports on other Directors. Each Director port can dynamically connect to many other ports to which channels requesting logical paths are attached.
- For CPCs, each logical partition attaching to the same control unit compete for the logical paths of the control unit.
- In a configuration where control units are shared by different CPCs, I/O configuration definitions for individual control units are not coordinated automatically among the IOCDSs of the different CPCs. Each CPC competes for the logical paths of a control unit.
- Shared channels require the establishment of a logical path for each channel image corresponding to an active LP sharing the channel. This requirement can significantly increase the number of logical paths that a single channel requests.

Channel configuration considerations

The following configuration rules determine how logical paths are established for FICON channels.

- A channel initially attempts to establish logical paths:
  - If you perform POR, only those channels configured online to LPs that are activated at POR attempt to establish logical paths. Shared channels attempt to establish logical paths only for those activated LPs with the channel configured online.
  - When the LP is activated.
  - When configured online (if previously configured offline).
- A channel cannot establish a logical path or has its logical path removed when:
  - An LP is deactivated. A shared channel continues to operate for any other remaining activated LPs to which it is defined. Logical paths to those LPs remain established.
- A shared channel **cannot** establish a logical path to a control unit for an LP that **cannot** access any of the I/O devices on the control unit. In IOCP, the PARTITION or NOTPART keyword on the IODEVICE statement specifies which LPs can access a device.
- A channel that cannot initially establish a logical path can reattempt to establish a logical path if the channel detects or is notified of:
  - A change in the state of a control unit
  - A change in the state of a link or port
  - A dynamic I/O configuration change that frees previously allocated logical paths

- A channel cannot establish a logical path or has its logical path removed if:
  - The Director that connects the channel to the control unit blocks either the channel port or control unit port used in the path.
  - The Director that connects the channel to the control unit prohibits the dynamic connection or communication between the channel port and the control unit port used in the path.
  - A link involved in the path fails or is disconnected. When a shared channel is affected by a port being blocked, a dynamic connection or communication being prohibited, or a link failing or being disconnected, each LP sharing the channel is equally affected and all logical paths using the port or link (regardless of which LP they are associated) are removed.
  - The channel is configured offline. When a shared channel is configured offline for an LP, it continues to operate for any other LP that has the channel configured online. Logical paths to these other logical partitions remain established.
  - Power to the channel, control units, or Directors in the configuration is turned off.

Recommendations

Creating I/O configuration definitions where channels could request more logical paths to control units than the control units could support can be useful in the following scenarios:

- **Workload balancing**

  When a system image becomes overloaded, you might need to reassign a workload and the necessary logical paths (for example, its tape or DASD volumes, a set of display terminals, or a set of printers) to another system image that has available capacity.

- **Backup**

  When an outage occurs, you can move the critical application set (the program and associated data) and the necessary logical paths to a backup or standby CPC. This process is simple if the CPCs have identical I/O configurations.

In I/O configurations where channels can request more logical paths to control units than the control units can support, you can manage how logical paths are established by:

- Deactivating unneeded LPs.
- Configuring offline unneeded channels. For shared channels, configure offline unneeded channels on an LP basis.
- Limiting the number of LPs that can access the I/O devices attached to a control unit when the control unit attaches to shared channels. In IOCP, specify the PARTITION or NOTPART keyword on the IODEVICE statement for every I/O device attaching to a control unit so that 1 or more LPs *cannot* access any of the I/O devices.
- Using the Director to block ports or prohibit dynamic connections or communication between ports.
- Combinations of the options in this list.

To better understand how you can manage logical paths using these methods, consider the following examples.

Deactivating unneeded logical partitions

Deactivating unneeded LPs can prove useful for managing how logical paths are established on CPCs in some situations.

The system establishes logical paths only when an LP is activated. Deactivating an LP results in removal of those logical paths associated with the LP. This deactivation can greatly reduce the number of logical paths requested by the system at any given time.

In Figure 5 on page 42, if all five of the LPs each share all four of the FICON channels and all the LPs are activated, the DS8K would be requested to establish five logical paths for each of the four shared FICON channels (or a total of 20 logical paths). DS8K models support varying numbers of logical paths for FICON channel paths. One such model has a maximum of 1,280 logical paths for a Fiber Channel port in the

DS8K. It also has a maximum of 512 logical paths per control unit image of the DS8K. These maximums help to reduce the need for managing logical paths but the maximums can still be exceeded.

For example, if you want to reduce your logical paths by four and you used LP4 and LP5 as test LPs that did not need to be active concurrently, you could reduce the number of logical paths requested by four by not activating either LP4 or LP5. In this case, four LPs (LP1, LP2, LP3, and LP4 or LP5) configured to four shared FICON channels would request a total of 16 logical paths. Later, you could transfer logical paths between LP4 and LP5 by first deactivating one LP to remove its logical paths, then activating the other LP to use the freed logical paths.
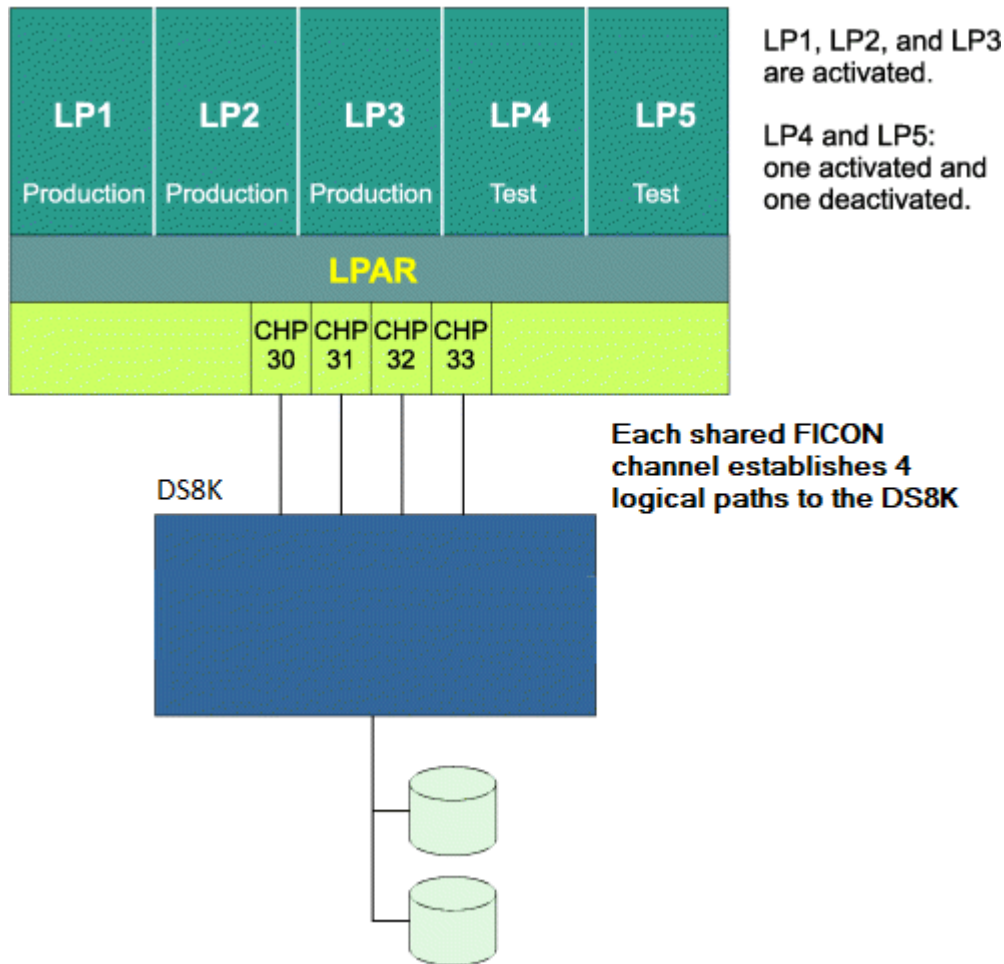

Figure 5. Deactivating unneeded logical partitions

Configuring offline unneeded channels or shared channels on an LP basis

You can configure offline unneeded channels or shared channels on an LP basis to manage how logical paths are established. In , all five LPs need to be active concurrently. If all five LPs had each of the four shared FICON channels configured online, 20 logical paths (four logical paths for each of the five LPs) would be requested.

However, if LP4 or LP5 (both test LPs) did not require four channel paths each to the DS8K, you could configure offline two of the four channel images used by LP4 and two of the four channel images used by LP5, reducing the total number of logical paths requested from 20 to 16.
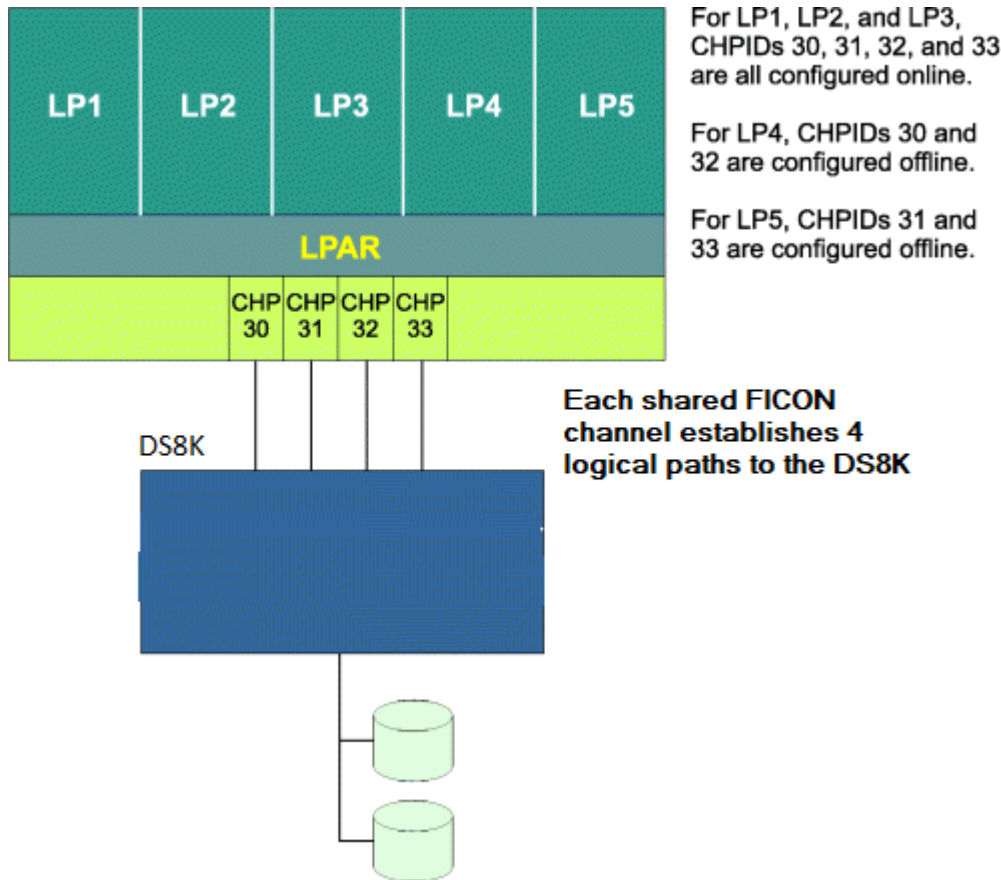
For LP1, LP2, and LP3, CHPIDs 30, 31, 32, and 33 are all configured online.

For LP4, CHPIDs 30 and 32 are configured offline.

For LP5, CHPIDs 31 and 33 are configured offline.

Each shared FICON channel establishes 4 logical paths to the DS8K

Figure 6. Configuring offline unneeded channels or shared channels on an LP basis

It is also possible to manage how logical paths are established by using IOCP or Hardware Configuration Definition (HCD) to create I/O configuration definitions that:

- Define a subset of LPs that have their corresponding channels configured online at power-on reset (POR) (CHPID access list)
- Allow LPs to configure online their channels at a later time (CHPID candidate list)

Use IOCP or HCD to define the access lists and candidate lists for channel paths to determine the configurability of a channel to an LP. This capability exists for both unshared and shared channels and can help automate and establish the configuration in . Additionally, HCD allows you to dynamically change the access list and candidate list for a channel path.

Defining devices to a subset of logical partitions

You can limit I/O device access from LPs to I/O devices assigned to shared channels by using IOCP or HCD to specify device candidate lists. By defining devices attached to a control unit to a subset of LPs, you can manage which LPs attempt to establish logical paths to the control unit through a shared channel.

If you define no devices to a control unit from a particular LP, the shared channel associated with the LP does not attempt to establish a logical path. However, if there is at least one device defined to the control unit for the shared channel associated with a particular LP, the shared channel for the LP attempts to establish a logical path to the control unit for the LP.

In , LP access to a series of control units is managed through use of the device candidate lists for the I/O devices attached to the control units. The shared FICON channel attempts to establish only one logical path to each of the control units even though 5 LPs share the channel.

In the example, the channel only attempts to establish a logical path for LP1 to the control unit defined as number 10 because only LP1 has a device defined to that control unit. Similarly, only LP2 can access

control unit 11, only LP3 can access control unit 12, only LP4 can access control unit 13, and only LP5 can access control unit 14.

Partial IOCP Deck for the Configuration

Following is a partial IOCP deck for the example in

```
CHPID PATH=30,SHARED

CNTLUNIT CUNUMBR=10,PATH=30
IODEVICE ADDRESS=VVVV,CUNUMBR=10,PART=LP1

CNTLUNIT CUNUMBR=11,PATH=30
IODEVICE ADDRESS=VVVV,CUNUMBR=11,PART=LP2

CNTLUNIT CUNUMBR=12,PATH=30
IODEVICE ADDRESS=VVVV,CUNUMBR=12,PART=LP3

CNTLUNIT CUNUMBR=13,PATH=30
IODEVICE ADDRESS=VVVV,CUNUMBR=13,PART=LP4

CNTLUNIT CUNUMBR=14,PATH=30
IODEVICE ADDRESS=VVVV,CUNUMBR=14,PART=LP5
```
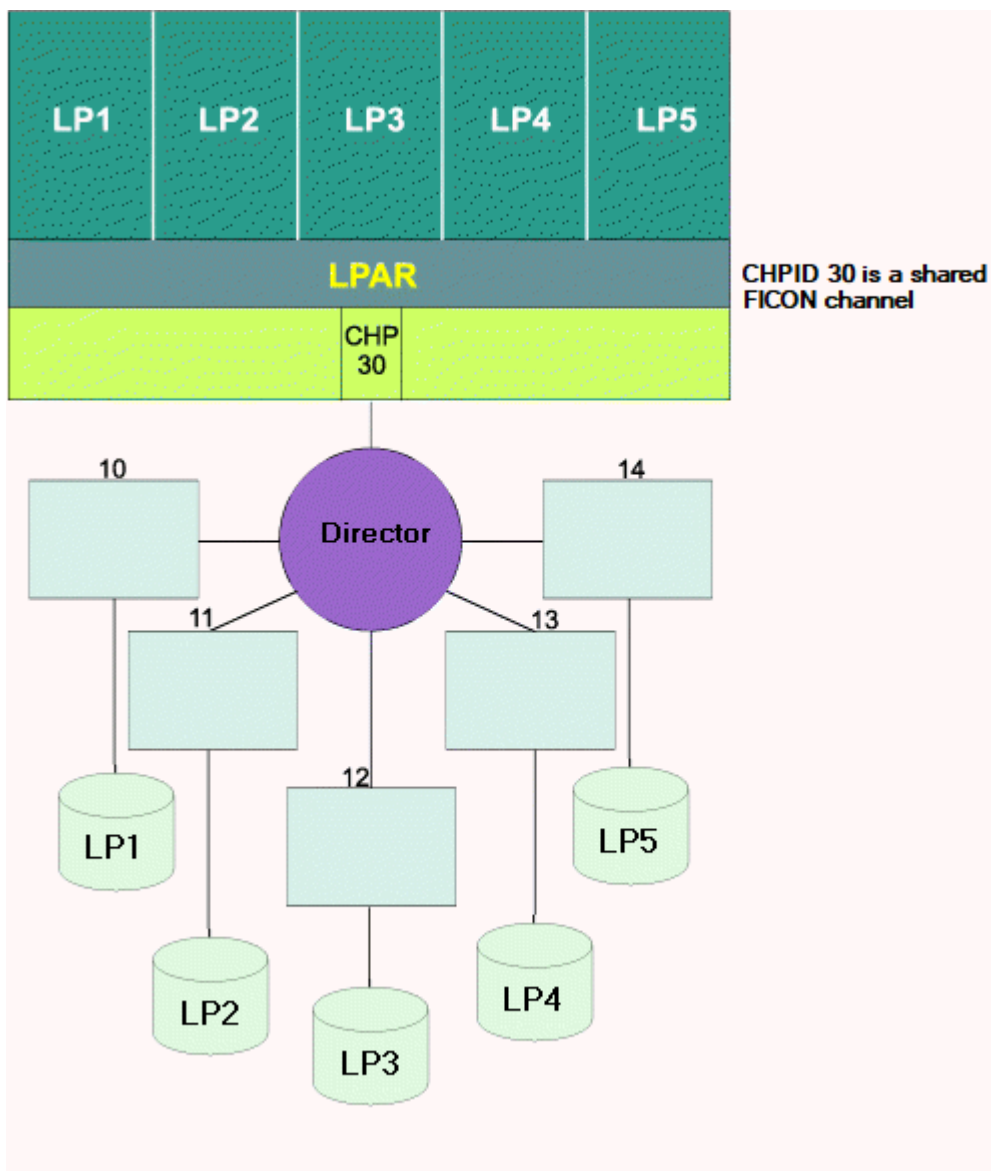

Figure 7. Defining devices to a subset of logical partitions

In Figure 8 on page 45, a DS8K is configured as 5 logical control units or control nit images and is defined as five control unit headers (CUHs). Each control unit image has a logical path limit which may need to be managed.
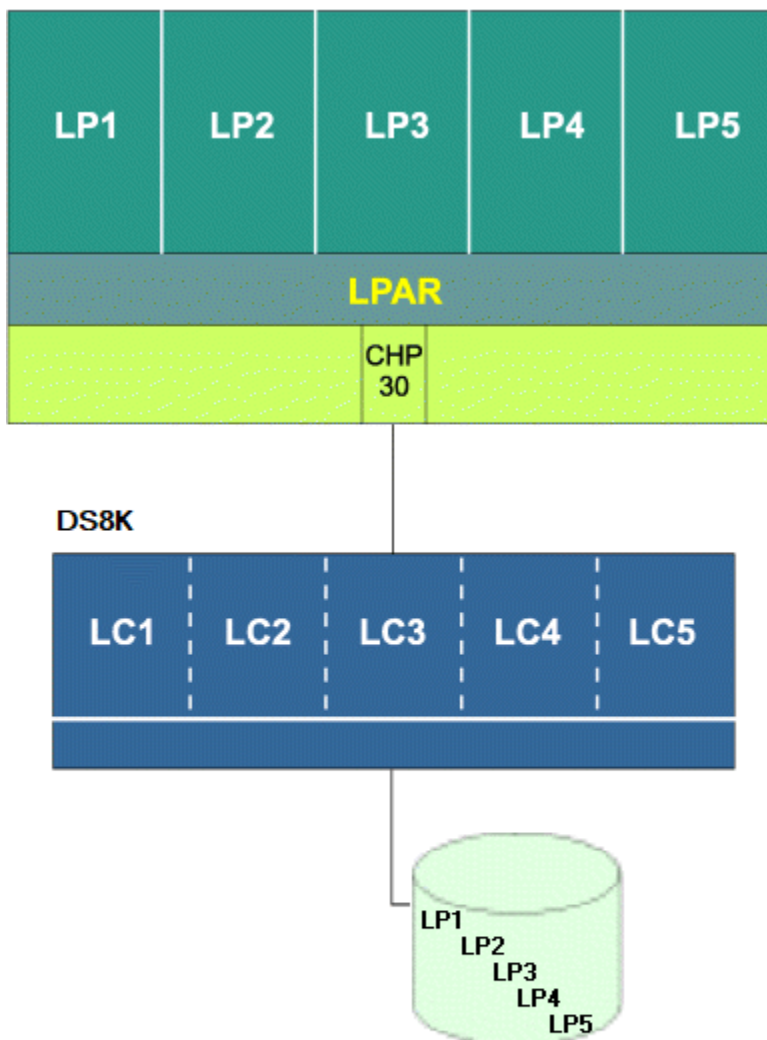

Figure 8. Defining devices to a subset of logical partitions

Even though 5 LPs share the FICON channel, the channel only attempts to establish a logical path for LP1 to CUH 0 because only LP1 has a device defined on that CUH. Similarly, only LP2 can access CUH 1, only LP3 can access CUH 2, only LP4 can access CUH 3, and only LP5 can access CUH 4.

Partial IOCP Deck for the Configuration

Following is a partial IOCP deck for the example in Figure 8 on page 45.

```
CHPID PATH=30,SHARED

CNTLUNIT CUNUMBR=10,PATH=30,CUADD=0
IODEVICE ADDRESS=VVVV,CUNUMBR=10,PART=LP1

CNTLUNIT CUNUMBR=11,PATH=30,CUADD=1
IODEVICE ADDRESS=WWWW,CUNUMBR=11,PART=LP2
```

```
CNTLUNIT CUNUMBR=12,PATH=30,CUADD=2
IODEVICE ADDRESS=XXXX,CUNUMBR=12,PART=LP3

CNTLUNIT CUNUMBR=13,PATH=30,CUADD=3
IODEVICE ADDRESS=YYYY,CUNUMBR=13,PART=LP4

CNTLUNIT CUNUMBR=14,PATH=30,CUADD=4
IODEVICE ADDRESS=ZZZZ,CUNUMBR=14,PART=LP5
```

Using a director to block ports or prohibit dynamic connections or communication

When FICON Directors are used in an I/O configuration, you can prevent channels from establishing logical paths or can remove established logical paths by either blocking a Director port or by prohibiting a dynamic connection or communication between two Director ports.

In terms of logical path removal, blocking a Director port connected to a channel produces a similar outcome to configuring offline a channel or all channel images of a shared channel. Blocking a Director port connected to a control unit prevents any logical path from being established to the attached control unit port.

You can more selectively prevent logical paths from being established by prohibiting a dynamic connection or communication between two FICON Director ports instead of blocking a Director port. By prohibiting a dynamic connection or communication between two Director ports, you can control which channels have connectivity to a control unit port rather than blocking all connectivity to the control unit port.

Prohibiting a dynamic connection or communication between two Director ports affects all channel images of a shared channel. The system does not establish any logical paths to the attached control unit port from any of the LPs that share the FICON channel.

You can prohibit dynamic connections or communication between Director ports by modifying the active configuration table. The active configuration table specifies the connectivity status of a port relative to the other ports on the Director. When a Director is first installed, it has a default configuration that allows any-to-any connectivity (every port can dynamically connect or communicate with every other port). If you require a different configuration, you can define and designate a different table to be the default configuration used at power-on of the Director. This table allows only those dynamic connections or communication necessary to establish the logical paths the configuration requires. Dynamic connections or communication necessary to establish other logical paths (for example, those necessary for backup configurations) would be prohibited by the default configuration of the Director.

shows an example of prohibiting dynamic connections. CPC1, CPC2, CPC3, and CPC4 are all production systems and CPC5 is a backup system to be used only if one of the other CPCs fail. If the default configuration used by the FICON Director prohibits all dynamic connections between CPC5 and the DS8K, the DS8K will only be requested to establish a total of 16 logical paths from the channels on CPC1, CPC2, CPC3, and CPC4. If one of four production CPCs fails, you could transfer the logical paths from the failing CPC to the backup CPC by prohibiting the dynamic connection to the failed CPC and allowing the dynamic connection to the backup CPC.

If a control unit is connected to more than one Director, it is necessary to coordinate allocation of the logical paths of the control unit across all the Directors. You can use the System Automation for z/OS (SA z/OS) to dynamically manage the Directors and logical paths by sending SA z/OS commands to reconfigure one or more Directors. SA z/OS then sends the appropriate operating system Vary Path requests. SA z/OS can also provide coordination between operating systems when logical paths are removed from one system and transferred to another system as a result of blocking Director ports or prohibiting Director dynamic connections or communication.
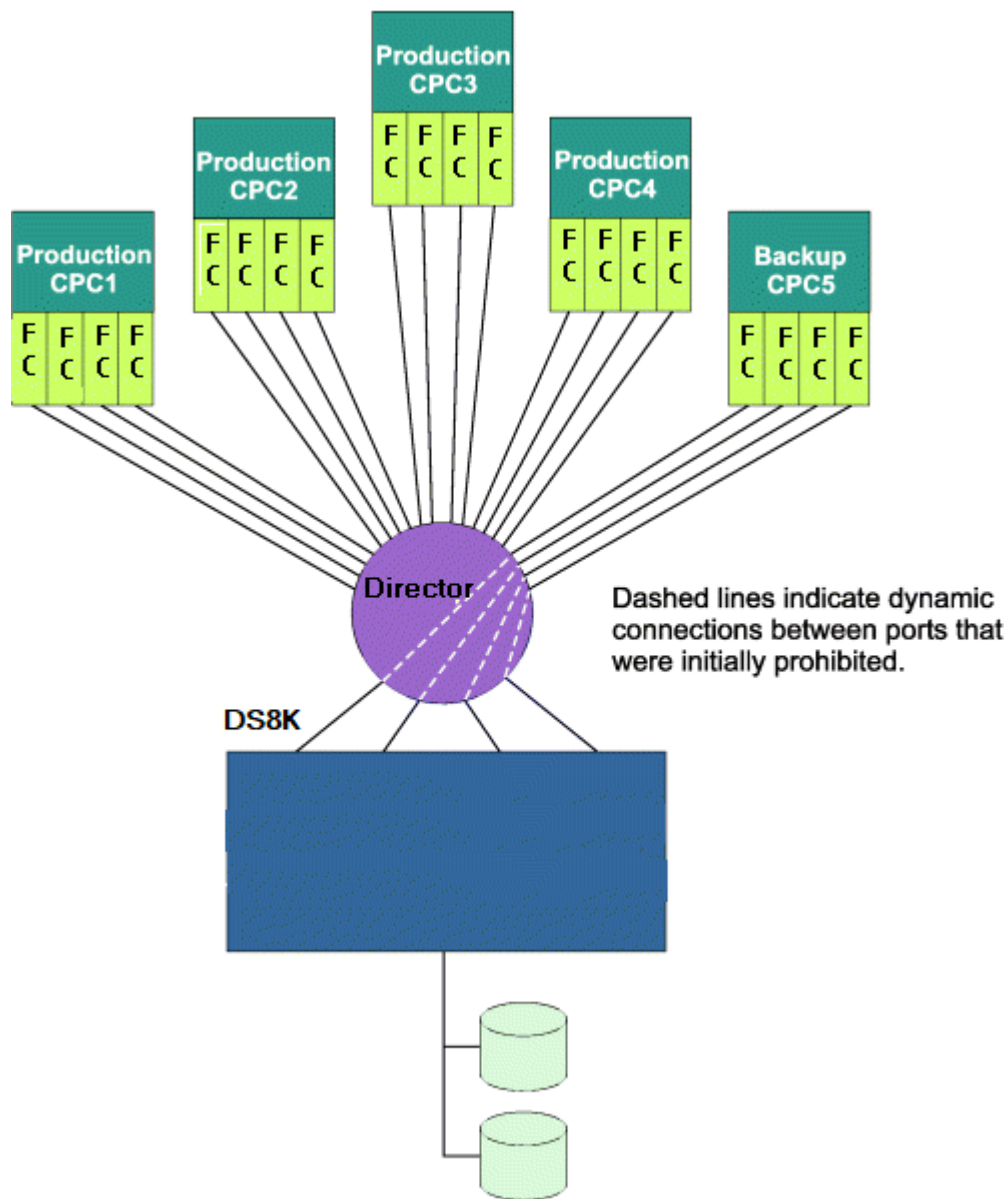
Figure 9. Using the FICON Director to manage logical paths by prohibiting dynamic connections

Shared channel overview

MIF allows channels to be shared among multiple LPs. Shared channels are configured to an LP giving the LP a channel image of the shared channel that it can use. Each channel image allows an LP to independently access and control the shared channel as if it were a physical channel assigned to the LP.

By providing the logical equivalent of multiple physical channels dedicated to multiple LPs, a shared channel can reduce hardware requirements without a corresponding reduction in I/O connectivity. This reduction in hardware requirements can apply to physical channels, Director ports, and control unit ports, depending on the configuration.

MIF performance planning considerations

Your installation can take advantage of MIF performance enhancements offered by:

- Understanding and using I/O-busy management enhancements
- Planning for concurrent data transfer
- Understanding examples of MIF consolidation

Planning for concurrent data transfer

Before you can consolidate channels, you must be aware of the channel requirements of the particular control units you are configuring. The number of channels needed is independent of the number of LPs on a system. The number of channels is based on the number of concurrent data transfers the control unit can handle. Although the recommended number of channels satisfies connectivity and performance requirements, additional channels can be added for availability.

Understanding examples of MIF consolidation

The following examples provide some general guidelines to show how MIF can help you consolidate and use hardware resources more efficiently:

FICON configurations

Figure 10 on page 48 shows how four shared FICON channels can replace 16 unshared (dedicated or reconfigurable) FICON channels and use 12 fewer control unit ports.
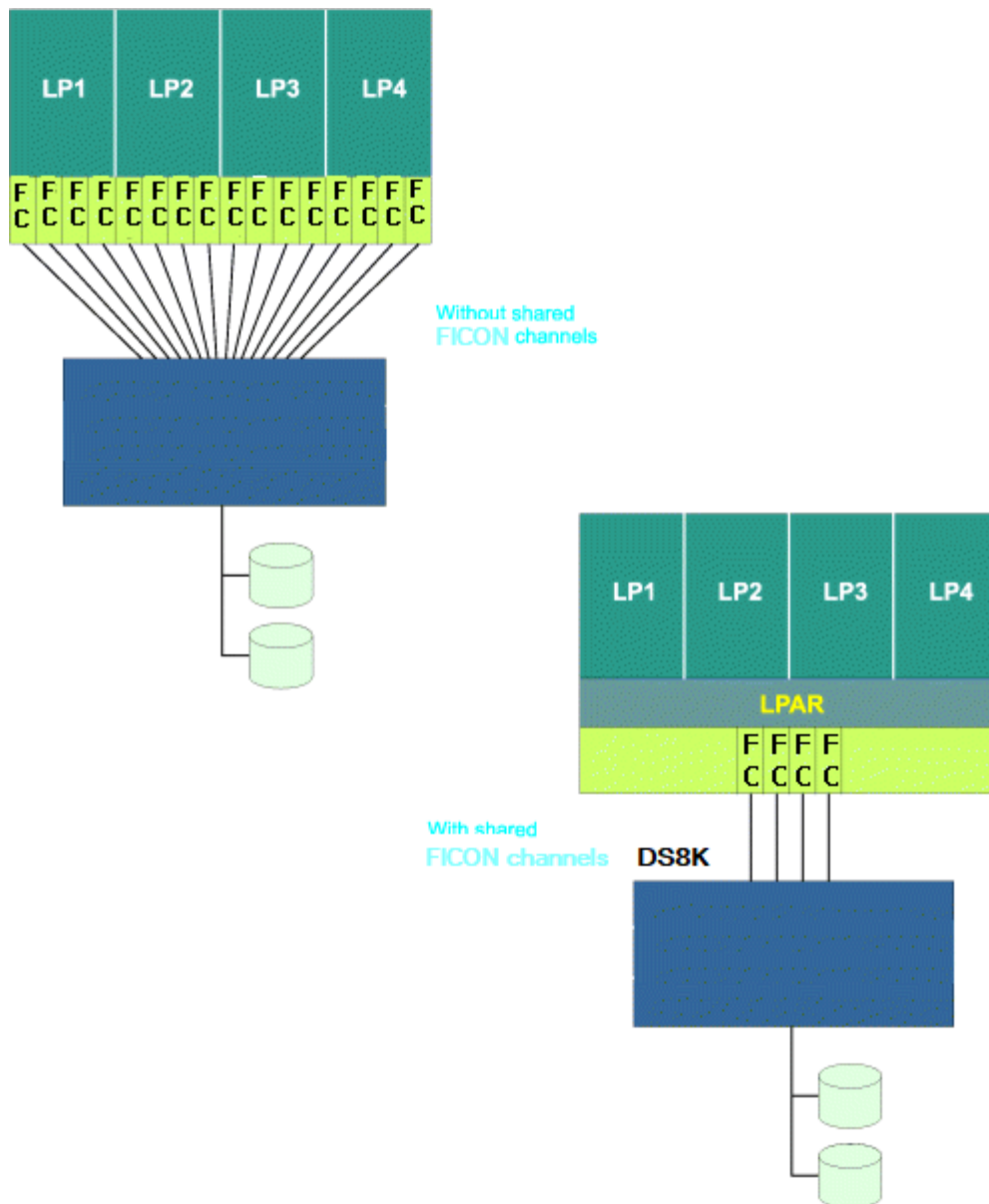
Figure 10. Consolidating FICON channels and FICON control unit ports

FICON Director configurations

shows how shared FICON channels can reduce FICON Director port requirements. In this example, two shared FICON channels replace 10 unshared (dedicated or reconfigurable) FICON channels and use eight fewer FICON Director ports without a reduction in I/O connectivity.
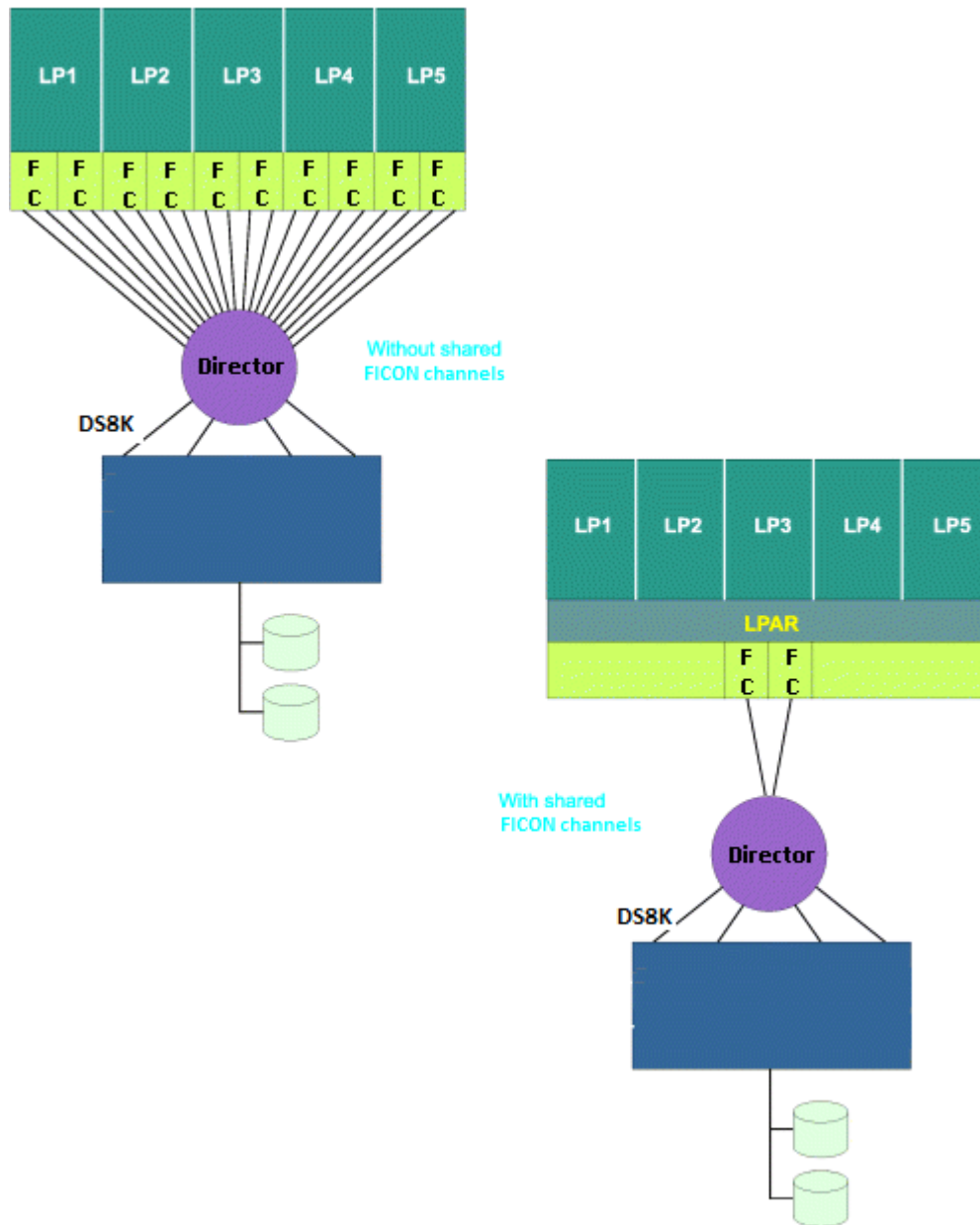

Figure 11. Consolidating FICON channels and FICON Director ports

FICON FC Configurations

shows how shared FICON channels can reduce the FICON channel requirements for FICON FC configurations. In this example, the CPC requires FC communications among all its LPs.
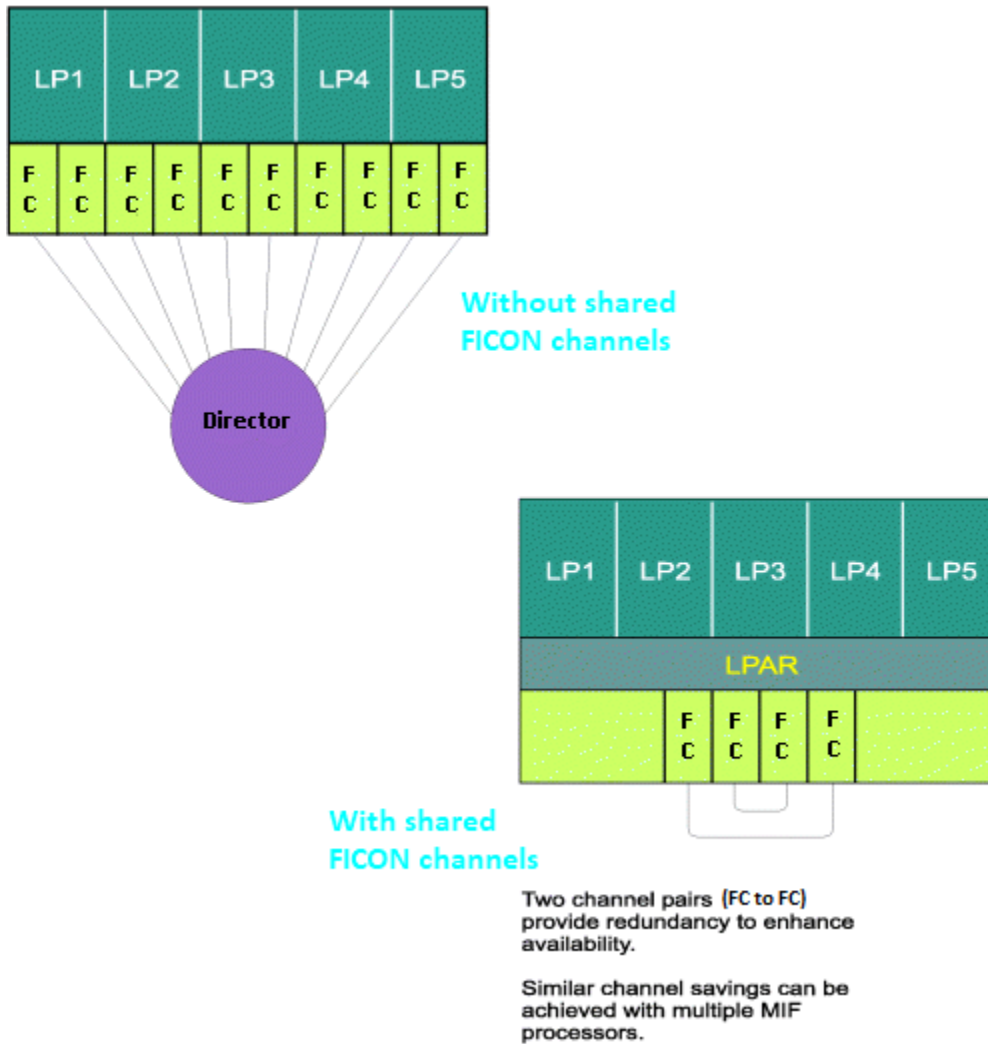
Figure 12. Consolidating FICON channels used for FICON FC communications

By using two shared FICON FC/FC pairs (4 shared FICON channels), you can:

- Replace five unshared FICON FC/FC pairs (10 unshared FICON channels) and the FICON Director used to connect them
- Provide full redundancy

I/O connectivity is maintained while hardware requirements (channels and a FICON Director) are reduced.

In situations where FICON FC communication is required among LPs that exist on two or more CPCs, shared channels can reduce even further channel and other hardware requirements and their associated cost.

FICON FC configurations are well-suited to take advantage of the consolidation benefits associated with shared channels. FC/FC pairs used for FICON FC communications have no limitation on the number of logical paths that can be established between them. The only limitations are the number of control units that can be defined for a FICON FC channel and the performance expectations you determine for your configuration.

Infrequently Used FICON Control Units

FICON control units not frequently used can use shared channels. You can attach such a control unit to a shared channel that is also attached to other, more frequently used control units without adding greatly to the channel utilization of the shared channel. A good example is the control unit within the Director.

Notes:

1. You cannot define a control unit (or multiple control units with common I/O devices) to a mixture of shared and unshared channel paths in the same IOCDS.

2. You cannot define more than one control unit with the same CUADD to the same link on a Director (or point-to-point) if the attaching CHPIDs are shared.

Understanding and using I/O-busy management enhancements

This section shows how the various FICON and MIF topologies offer improvements in managing I/O-busy conditions. Figure 13 on page 51 compares the point-to-point, switched point-to point, and MIF channel sharing topologies.
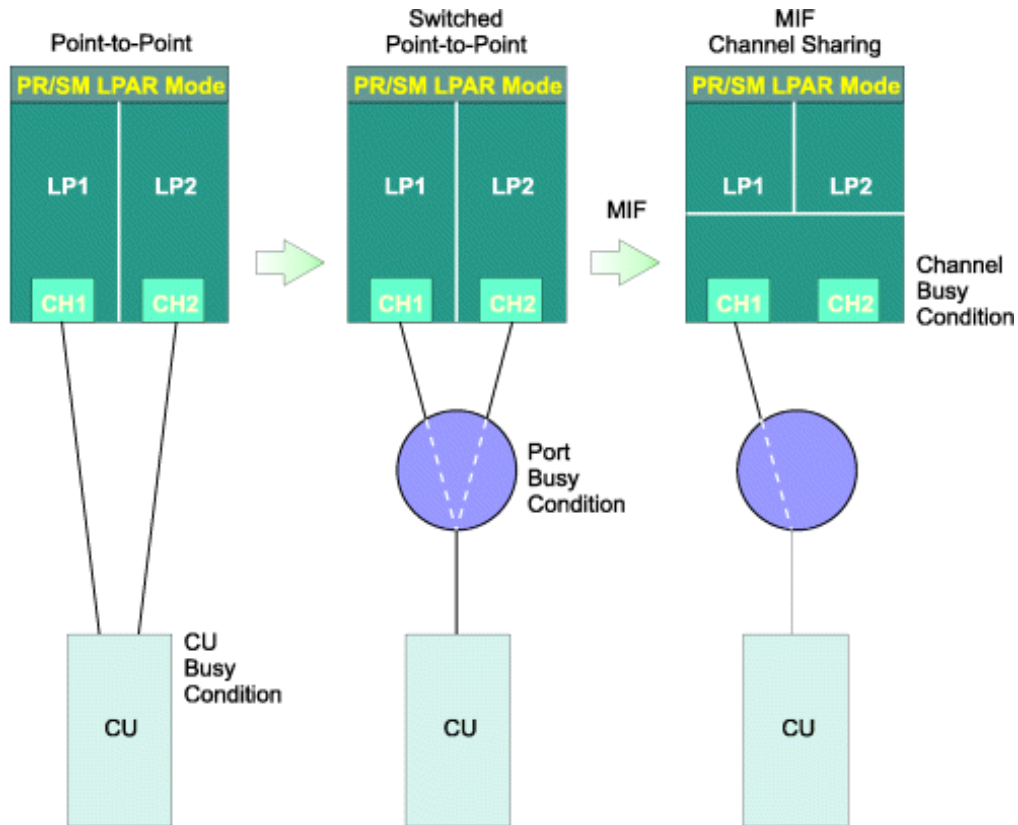
Figure 13. Progression of busy condition management improvements

Point-to-Point topologies

Concentrate I/O attempts at the control unit level and are distance-dependent. At the time of a control unit busy encounter, the control unit must present control unit busy status to the channel. Once the control unit is free, it presents a control unit no longer busy status to the channel. This process of presenting status to the channel requires control unit processing and many trips over the control unit to channel link.

Switched Point-to-Point topologies

Concentrate I/O attempts within the Director and therefore encounter switch port busies. The processing of switch port busies does not require any control unit involvement. Busies are handled by the FICON Director. Therefore, the control unit is effectively relieved of handling busy conditions and is able to handle more I/O requests. Because switch port busies require fewer trips over the FICON connection link, they are less sensitive to increased distances than control unit busy encounters.

MIF channel sharing

Moves busy management back into the channel subsystem, providing the most efficient management of busy conditions. Because multiple LPs access the same physical channel, I/O attempts are concentrated at the channel level. A CU busy or a switch port busy is handled as a channel busy.

Dynamically managed CHPIDs

A key aspect of the Intelligent Resource Director (IRD) provided by the WLM component of z/OS is Dynamic CHPID Management (DCM). DCM provides the ability to have the system automatically move the available channel bandwidth to where it is most needed. CHPIDs identified as managed in the IOCDS (by using CHPARM and IOCLUSTER keywords) are dynamically shared among z/OS images within an LPAR cluster.

Before DCM, available channels had to be manually balanced across I/O devices in an attempt to provide sufficient paths to handle the average load on every controller. Natural variability in demand means that some controllers at times have more I/O paths available than they need, while other controllers possibly have too few. DCM attempts to balance responsiveness of the available channels maximizing the utilization of installed hardware. Fewer overall channels are required because the DCM CHPIDs are more fully used. RMF provides a report showing the average aggregate utilization for all managed channels.

By using DCM, you now only have to define a minimum of one nonmanaged path and up to seven managed paths to each control unit (although a realistic minimum of two nonmanaged paths are recommended), with dynamic channel path management taking responsibility for adding additional paths as required. For more information about defining and using DCM, including detailed examples, see the *z/OS Intelligent Resource Director*, SG24-5952.

IOCP coding specifications

ICP IOCP can only generate an LPAR IOCDS. No IOCP invocation parameter is required to generate an LPAR IOCDS.

IOCP statements for ICP

The RESOURCE statement is used to specify all the logical partition names defined in a machine configuration. To plan for growth in the number of logical partitions in the configuration, one or more asterisks (*) might be used to specify that one or more logical partitions are to be reserved along with their associated CSS and MIF image IDs. A reserved LP can only be specified for a dynamic-capable IOCDS. A dynamic-capable IOCDS is built when using HCD on z/OS or z/VM or by specifying the IOCP CMS utility option DYN for z/VM. Space in the hardware system area (HSA) is allocated for reserved LPs but cannot be used until a dynamic I/O configuration change is made to assign a name to the LP. The following rules apply when specifying reserved LPs:

- A reserved LP must have a user-specified MIF image ID
- A reserved LP cannot have any channel paths assigned to it
- An IOCDS cannot contain only reserved LPs. At least one LP must be defined with a name.

Dynamic CHPID Management (DCM) channel paths defined for a given LPAR cluster are shareable among all active LPs that have joined that cluster. Other than DCM channel paths, you must assign each channel path to a logical partition in an LPAR IOCDS. For each DCM channel path, ICP requires the CHPARM keyword have a value of 01 and the IOCLUSTER keyword on a CHPID statement. All other channel paths require the PART|PARTITION, NOTPART, or SHARED keyword on all CHPID statements unless a channel path is defined as spanned by specifying multiple CSS IDs in the PATH keyword of the IOCP CHPID statement.

Use the CHPARM and IOCLUSTER keywords on the CHPID statement to specify channel paths reserved for the use of a particular LPAR cluster. A DCM channel path becomes available to a candidate logical partition when the LP is activated and joins the specified cluster.

Use the CHPID PART|PARTITION, NOTPART, and SHARED keywords to determine which:

- Channel paths are assigned to each LP

- Devices and control units are shared among LPs
- Channel paths are reconfigurable
- Channel paths are shared

Use the CHPID CPATH keyword to connect two internal coupling channels.

Use the CHPID PATH keyword to define a channel path as spanned to multiple CSSs. Spanned channel paths are also shared channel paths.

DCM channel paths are implicitly shared. Use of the IOCLUSTER keyword implies a null access list (no logical partition has the channel path brought online at activation) and a candidate list of all defined logical partitions. The IOCLUSTER keyword is mutually exclusive with the PART|PARTITION and NOTPART keywords.

All LP names that you specify in the CHPID statements must match the names specified in the RESOURCE statement. An IOCDS must have at least one LP name defined.

```
PARTITION={(CSS(cssid),{name|0}[,REC])|
          (CSS(cssid),access list)|
          (CSS(cssid),(access list)[,(candidate list)][,REC])|
        ((CSS(cssid),(access list)[,(candidate list)]),...)}

NOTPART={(CSS(cssid),access list)|
        ((CSS(cssid),(access list)[,(candidate list)]),...)}

IOCLUSTER=cluster_name

SHARED

CPATH=(CSS(cssid),chpid number)
```

Where:

name
 specifies the name of the LP that has authority to access the CHPID. The LP name is a 1 - 8 alphanumeric (0 - 9, A - Z) character name that must have an alphabetic first character. Special characters ($, #, @) are not allowed. A reserved LP cannot have any channel paths assigned to it.

 The following words are reserved and you cannot use them as LP names:

 PHYSICAL
 REC
 SYSTEM
 PRIM*nnnn* (where *nnnn* are digits)

 ICP IOCP supports a maximum of 85 LP names (models T01/LT1) and 40 LP names (models T02/LT2) for the CPC.

cluster_name
 specifies the name of an LPAR cluster that has authority to access the specified DCM CHPID. The name of the LPAR cluster is a one- to eight- alphanumeric character name (0-9, A-Z) that must have an alphabetic first character. Special characters ($, #, @) are not allowed.

REC
 specifies that the CHPID is reconfigurable. A reconfigurable CHPID must have an initial access list of one LP name. Its candidate list must consist of one or more LP names.

access list
 specifies the LPs that have initial access to the CHPID at the completion of the initial power-on reset. An LP name can only appear once in an access list.

 You can specify that no LPs access the channel path following LP activation for the initial POR of an LPAR IOCDS. Specifying 0 indicates a null access list.

candidate list
 specifies the LPs that have authority to access the CHPID. Any LP that is not in a CHPIDs candidate list cannot access the CHPID.

You can specify as many LP names as your CPC supports. However, the number of unique LP names specified in both the access list and candidate list can not exceed the number of LPs your CPC supports.

If you specify the candidate list, you do not need to specify again the LP names specified in the initial access list. The initial access list is always included in the candidate list.

An LP name can only appear once in a candidate list. If the candidate list is not specified, it defaults to all LPs in the configuration for reconfigurable and shared channels.

Note: It is highly recommended that a peer mode CHPID (CE LR, ICA SR coupling link, or ICP) have at most one coupling facility LP specified in its initial access list in order to avoid confusion on subsequent LP activations. A peer mode CHPID can be online to only one coupling facility LP at a time.

Using the SHARED keyword specifies that the channel paths on the CHPID statement are shared. More than one LP, at the same time, can access a shared CHPID. When CHPIDs are not shared, only one LP can access it. Although you can dynamically move a reconfigurable CHPID between LPs, it can only be accessed by 1 LP at any given time. CVC and CBY channel paths (TYPE keyword) cannot be shared. On CF only models and ICP channel paths cannot be shared.

The CPATH keyword is only valid for ICP, ICA SR coupling link, an Coupling Express LR channel paths (TYPE keyword) and required for all ICP definitions. CPATH specifies the connection between 2 ICPs at either end of a coupling link:

```
PATH=FE,TYPE=ICP,CPATH=FF,...
PATH=FF,TYPE=ICP,CPATH=FE,...
```

specifies that ICP channel path FF connects to ICP channel path FE. Every ICP channel path of a coupling facility must be connected to an ICP channel path of a z/OS LP. The connection needs to be specified for each channel path. ICP channel paths cannot connect to each other if they both have candidate lists with the same, single logical partition. This restriction prevents the definition of internal coupling channels in an PR/SM configuration with only one logical partition. Also, an ICP channel path cannot connect to itself.

The CPATH value for a ICA SR coupling link CHPID specifies the CSS and CHPID number this ICA SR coupling link CHPID connects with on the target system. For example:

```
PATH=C0,TYPE=ICA SR,CPATH=(CSS(1),D0),...
```

Defines an ICA SR coupling link CHPID, C0, on this system that connects with ICA SR coupling link CHPID D0 in CSS 1 on the remote system.

Shared devices using shared channels

MIF allows you to use shared channels when defining shared devices. Using shared channels reduces the number of channels required, allows for increased channel utilization, and reduces the complexity of your IOCP input.

Note: You cannot mix shared and unshared channel paths to the same control unit or device.
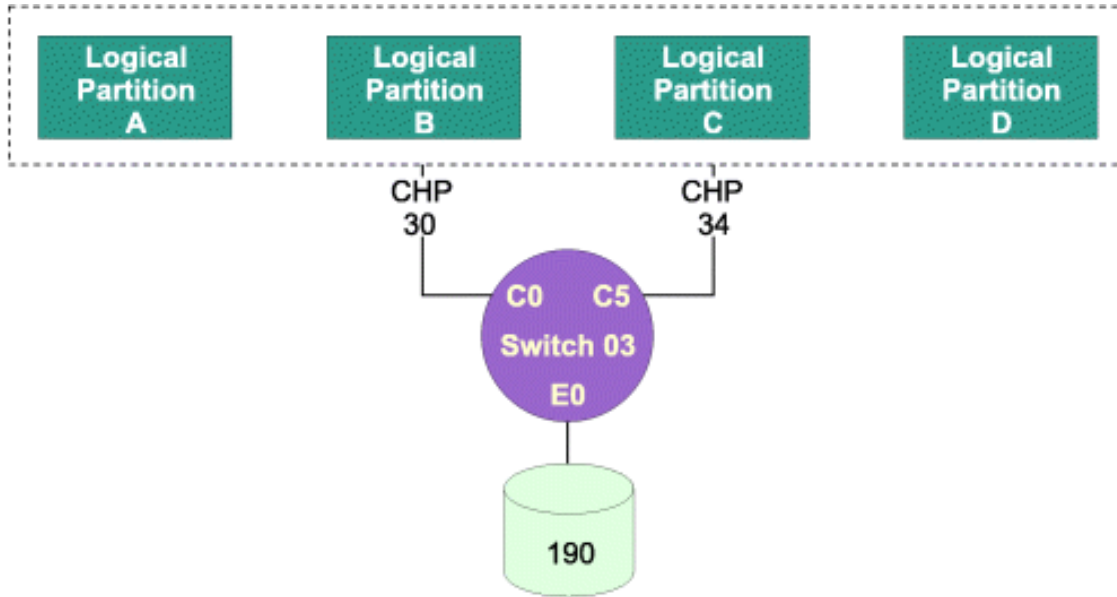
Following is an example of an IOCDS with a shared device.

Figure 14. Shared devices using shared FICON channels

Following is the IOCP coding for Figure 14 on page 55.

```
CHPID PATH=(30),TYPE=FC,SWITCH=03,SHARED . . .
CHPID PATH=(34),TYPE=FC,SWITCH=03,SHARED . . .
CNTLUNIT CUNUMBR=000,PATH=(30,34),UNITADD=((90)),LINK=(E0,E0),UNIT=xxx . . .
IODEVICE ADDRESS=(190),CUNUMBR=000,UNIT=xxx . . .
```

Shared devices using unshared channels

When coding an IOCP input file, the following specifications are allowed:

- Duplicate device numbers can be specified within a single IOCP input file, if device numbers are not duplicated within an LP.
- You can assign a maximum of eight channel paths from each LP to a device.

  Device sharing among LPs is accomplished by attaching multiple channel paths from each LP to a device.

The following section illustrates IOCP coding for IOCDSs when shared devices on unshared channels and duplicate device numbers are specified.

Shared devices

The following examples illustrate this concept by showing the physical connectivity of an I/O configuration for multiple LPs and the IOCP coding for the same configuration.

Using channels

Figure 15 on page 56 shows an example of an I/O configuration with a device shared by each of the four logical partitions. In this representation of a shared device, each logical partition views device 190 as part of its own I/O configuration. Notice the recoverability characteristics of this configuration: each logical partition has two channel paths to the shared device, each attached to a different storage director.
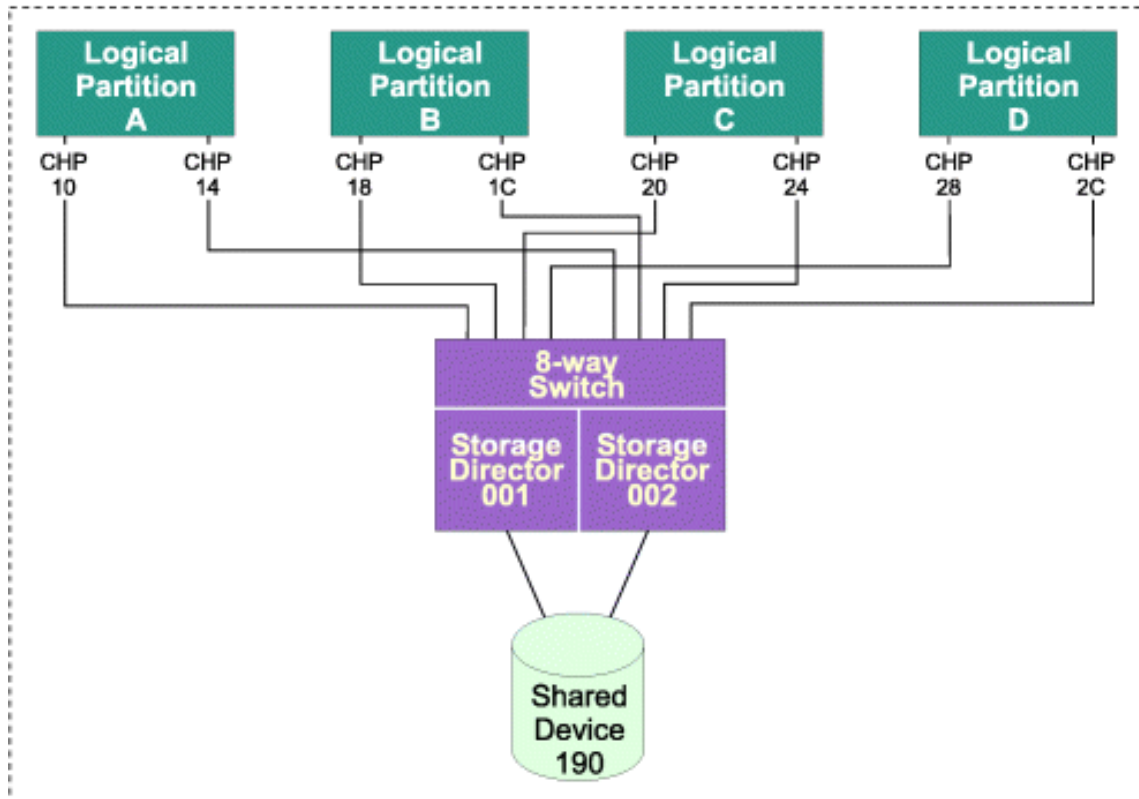
Figure 15. Physical connectivity of shared device 190

The following example shows the IOCP statement for .

```
CHPID PATH=(10),PART=(A,REC)
CHPID PATH=(14),PART=(A,REC)
CHPID PATH=(18),PART=(B,REC)
CHPID PATH=(1C),PART=(B,REC)
CHPID PATH=(20),PART=(C,REC)
CHPID PATH=(24),PART=(C,REC)
CHPID PATH=(28),PART=(D,REC)
CHPID PATH=(2C),PART=(D,REC)
CNTLUNIT CUNUMBR=0001,PATH=(10,18,20,28),UNITADD=((90)) . . .
CNTLUNIT CUNUMBR=0002,PATH=(14,1C,24,2C),UNITADD=((90)) . . .
IODEVICE ADDRESS=(190),CUNUMBR=(0001,0002) . . .
```

If 8 or less channels attach to the device, this method of defining the IOCP input provides greater flexibility because it allows you to move CHPIDs from one LP to another and eliminates possible conflicts (see).

shows an alternative method of defining the configuration. This method is required if there are greater than eight paths to the device. This logical representation has the same recoverability characteristics as the physical connectivity:

• Each LP has two channel paths to the shared device

• Each LP is attached to a different storage director

However, paths cannot be moved between the LPs.
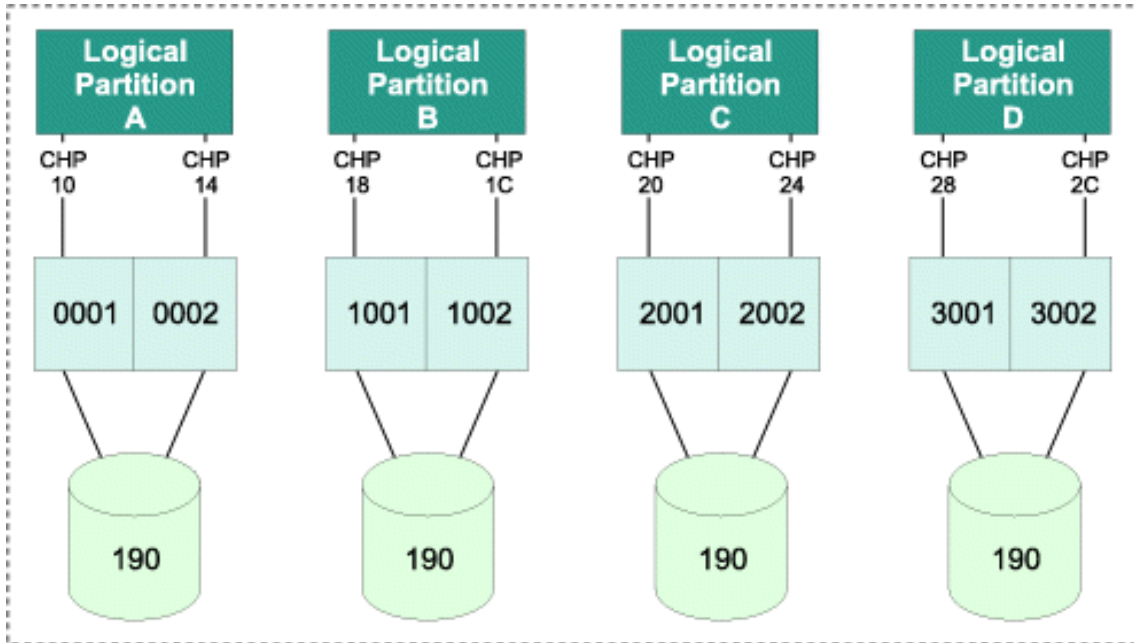
Figure 16. Logical view of shared device 190

The following example shows the IOCP statement for Figure 16 on page 57.

```
CHPID PATH=(10),PARTITION=(A), . . .
CHPID PATH=(14),PARTITION=(A), . . .
CNTLUNIT CUNUMBR=0001,PATH=(10),UNITADD=((90)) . . .
CNTLUNIT CUNUMBR=0002,PATH=(14),UNITADD=((90)) . . .
IODEVICE ADDRESS=(190),CUNUMBR=(0001,0002) . . .

CHPID PATH=(18),PARTITION=(B), . . .
CHPID PATH=(1C),PARTITION=(B), . . .
CNTLUNIT CUNUMBR=1001,PATH=(18),UNITADD=((90)) . . .
CNTLUNIT CUNUMBR=1002,PATH=(1C),UNITADD=((90)) . . .
IODEVICE ADDRESS=(190),CUNUMBR=(1001,1002) . . .

CHPID PATH=(20),PARTITION=(C) . . .
CHPID PATH=(24),PARTITION=(C) . . .
CNTLUNIT CUNUMBR=2001,PATH=(20),UNITADD=((90)) . . .
CNTLUNIT CUNUMBR=2002,PATH=(24),UNITADD=((90)) . . .
IODEVICE ADDRESS=(190),CUNUMBR=(2001,2002) . . .

CHPID PATH=(28),PARTITION=(D), . . .
CHPID PATH=(2C),PARTITION=(D), . . .
CNTLUNIT CUNUMBR=3001,PATH=(28),UNITADD=((90)) . . .
CNTLUNIT CUNUMBR=3002,PATH=(2C),UNITADD=((90)) . . .
IODEVICE ADDRESS=(190),CUNUMBR=(3001,3002) . . .
```

Duplicate device numbers for different physical devices

Figure 17 on page 58 illustrates a configuration where duplicate device numbers are used to represent a console (110) and a printer (00E) within each of four logical partitions.
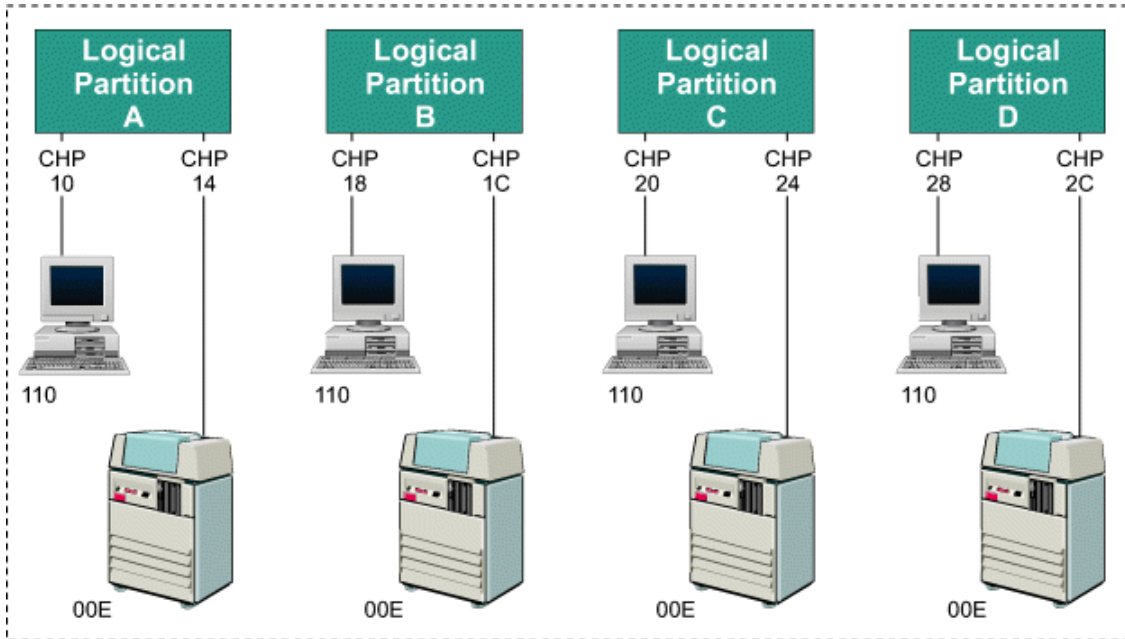
Figure 17. PR/SM configuration with duplicate device numbers

The following example shows the IOCP statement for . This IOCP coding example groups the input statements by logical partition. When coding IOCP, view the I/O devices from a logical partition perspective.

```
CHPID PATH=(10),PARTITION=(A), . . .
CHPID PATH=(14),PARTITION=(A), . . .
CNTLUNIT CUNUMBR=0011,PATH=(10),UNITADD=(10), . . .
CNTLUNIT CUNUMBR=0012,PATH=(14),UNITADD=(0E), . . .
IODEVICE ADDRESS=(110),CUNUMBR=(0011), . . .
IODEVICE ADDRESS=(00E),CUNUMBR=(0012), . . .

CHPID PATH=(18),PARTITION=(B), . . .
CHPID PATH=(1C),PARTITION=(B), . . .
CNTLUNIT CUNUMBR=0013,PATH=(18),UNITADD=(10), . . .
CNTLUNIT CUNUMBR=0014,PATH=(1C),UNITADD=(0E), . . .
IODEVICE ADDRESS=(110),CUNUMBR=(0013), . . .
IODEVICE ADDRESS=(00E),CUNUMBR=(0014), . . .

CHPID PATH=(20),PARTITION=(C), . . .
CHPID PATH=(24),PARTITION=(C), . . .
CNTLUNIT CUNUMBR=0015,PATH=(20),UNITADD=(10), . . .
CNTLUNIT CUNUMBR=0016,PATH=(24),UNITADD=(0E), . . .
IODEVICE ADDRESS=(110),CUNUMBR=(0015), . . .
IODEVICE ADDRESS=(00E),CUNUMBR=(0016), . . .

CHPID PATH=(28),PARTITION=(D), . . .
CHPID PATH=(2C),PARTITION=(D), . . .
CNTLUNIT CUNUMBR=0017,PATH=(28),UNITADD=(10), . . .
CNTLUNIT CUNUMBR=0018,PATH=(2C),UNITADD=(0E), . . .
IODEVICE ADDRESS=(110),CUNUMBR=(0017), . . .
IODEVICE ADDRESS=(00E),CUNUMBR=(0018), . . .
```

Eight IODEVICE statements are used, one for each console and one for each printer that has a duplicate device number. Device numbers 110 and 00E occur four times each; however, they are not duplicated within a logical partition. When coding an IOCP input file, remember that the unique device number rule applies for logical partitions in an IOCDS.

shows another example of a logical partition configuration in which the device number for a console (110) is duplicated for all four logical partitions.
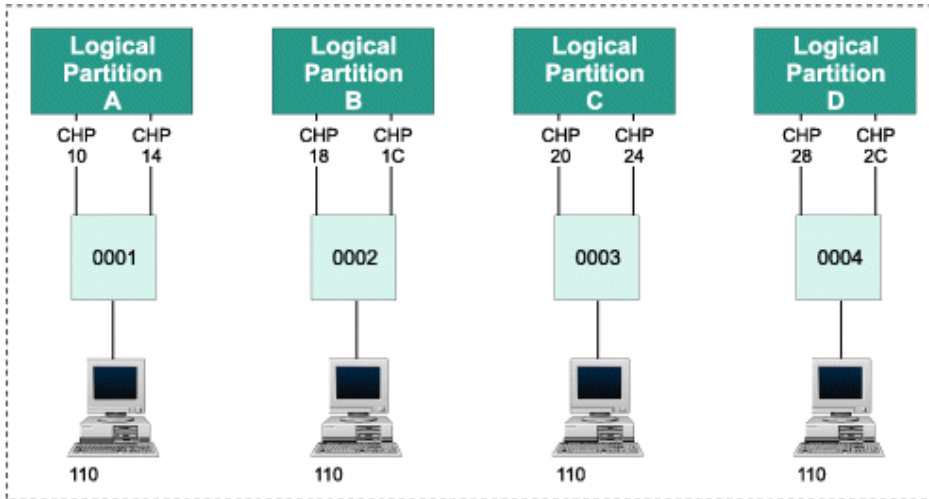
Figure 18. Duplicate device numbers for console

The following example shows the IOCP coding for the previous configuration. Four IODEVICE and four CNTLUNIT statements are used, one each for the console within each logical partition that has a duplicate device number.

```
CHPID PATH=(10),PARTITION=(A), . . .
CHPID PATH=(14),PARTITION=(A), . . .
CNTLUNIT CUNUMBR=0001,PATH=(10,14),UNITADD=((10)), . . .
IODEVICE ADDRESS=(110),CUNUMBR=(0001), . . .

CHPID PATH=(18),PARTITION=(B), . . .
CHPID PATH=(1C),PARTITION=(B), . . .
CNTLUNIT CUNUMBR=0002,PATH=(18,1C),UNITADD=((10)), . . .
IODEVICE ADDRESS=(110),CUNUMBR=(0002), . . .

CHPID PATH=(20),PARTITION=(C), . . .
CHPID PATH=(24),PARTITION=(C), . . .
CNTLUNIT CUNUMBR=0003,PATH=(20,24),UNITADD=((10)), . . .
IODEVICE ADDRESS=(110),CUNUMBR=(0003), . . .

CHPID PATH=(28),PARTITION=(D), . . .
CHPID PATH=(2C),PARTITION=(D), . . .
CNTLUNIT CUNUMBR=0004,PATH=(28,2C),UNITADD=((10)), . . .
IODEVICE ADDRESS=(110),CUNUMBR=(0004), . . .
```

Duplicate device number conflicts

IOCP allows duplicate device numbers in an IOCDS only if the duplicate device numbers do not occur in the same logical partition. Therefore, IOCP allows systems to use different logical partitions to integrate a processor complex without changing device numbers.

IOCP requires a unique device number for each device within a logical partition. When IOCP completes without error, the initial configuration contains no duplicate device number conflicts within a logical partition.

Conflicts can occur when the I/O configuration is modified. If a channel path is configured to a logical partition and devices attached to the channel path have device numbers that are already assigned in the receiving logical partition to other online channel paths, a conflict results.

When an I/O configuration is dynamically modified so the logical partition can gain access to a device not previously accessible, a device conflict can occur. The conflicts are detected when commands are processed that change the I/O configuration or when you attempt to activate the logical partition which has the device number conflict. A message displays identifying the error.

The identified device cannot be accessed while a conflict exists. Two types of conflict are possible:

• Conflicts between device numbers for the same device (a shared device)

• Conflicts between device numbers for different devices (unshared devices)

Activation fails if a duplicate device number conflict exists.

Examples of duplicate device number conflicts
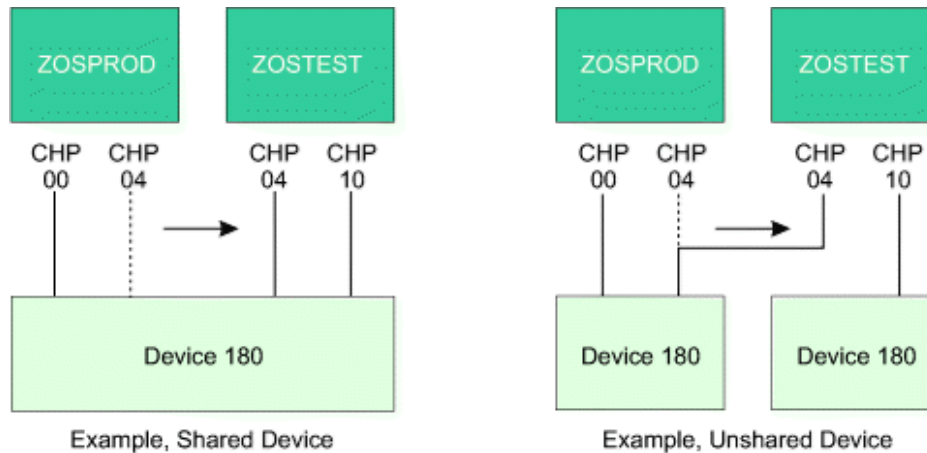
provides two examples of duplicate device number conflict.

Figure 19. Two examples of duplicate device number conflicts

The following example shows the IOCP statement for . Both examples use identical IOCP statements.

```
CHPID PATH=(00),PARTITION=(ZOSPROD,REC)
CHPID PATH=(04),PARTITION=(ZOSPROD,REC)
CNTLUNIT CUNUMBR=0001,PATH=(00,04),UNITADD=80
IODEVICE ADDRESS=180,CUNUMBR=0001

CHPID PATH=(10),PARTITION=(ZOSTEST)
CNTLUNIT CUNUMBR=0002,PATH=(10),UNITADD=80
IODEVICE ADDRESS=180,CUNUMBR=0002
```

Channel path 04 is reassigned from ZOSPROD to ZOSTEST in each example. This reassignment creates a duplicate device number conflict for device number 180 when the devices are connected to two different control units. This conflict occurs because a device numbered 180 exists on the original channel path 10. If such conflicts occur, the operator must know what configuration is wanted.

Shared device
> In the example on the left, the duplicate device numbers refer to the same device from different logical partitions (a new path to the same device has been moved to ZOSTEST). This situation might result in a performance problem because the control program in logical partition ZOSPROD cannot access the device from channel path 4.

Unshared Device
> In the example on the right, the duplicate device numbers refer to a different device from each logical partition (a new device has been moved to ZOSTEST). This situation might result in a data integrity problem because the control program in logical partition ZOSTEST cannot access the correct device from channel path 04.

Resolving duplicate device number conflicts

Consider options A, B, and C when planning the I/O configuration and the reconfigurability of channel paths. You can resolve duplicate device number conflicts by choosing one of the options:

A
> **Use the original channel path:**
>
> If the receiving logical partition does not need a new path to a shared device or does not need the new (unshared) device, take no action. The conflict is resolved by using only the original path (shared device) or the original device. (Access is still allowed to any non-conflicting devices on the newly configured channel path.)

In Figure 19 on page 60 , ZOSTEST can access device 180 only through channel path 10 if the operator takes no action in response to the conflict message.

B

Deconfigure the original channel path:

If the logical partition must have the reassigned channel path to a shared device or access to a new (unshared) device, the conflict is resolved by substituting the reassigned channel path for the original channel path. Do the following:

1. Configure offline the original channel path (CHP 10 in Figure 19 on page 60).
2. Configure offline and then online the reassigned channel path (CHP 04 in Figure 19 on page 60).
3. If necessary, configure online the original channel path (CHP 10 in Figure 19 on page 60). Another conflict message is issued because a new conflict has been created. The operator then ignores this conflict as described in option A. (Access is still allowed to any non-conflicting devices on the original channel path.)

In Figure 19 on page 60, ZOSTEST can access device 180 only through channel path 04 if the preceding steps are performed in response to the conflict message.

C

Change the I/O configuration:

Only option C provides a permanent resolution to a device number conflict.

If the logical partition must have access to all devices over the original channel path and the reassigned channel path (shared devices), or to a new device and the original device (unshared devices), do one of the following:

- Create a configuration with unique device numbers, if they are unshared devices.
- For shared devices, define a single device with access to all the channel paths attached to the physical control units.
- For a shared device assigned to unshared channel paths, change the channel paths to shared and consolidate the control units and device definitions to one each.
- If the device is assigned to shared channel paths, control access to the devices using their device candidate list.

The configuration can be activated by performing a POR or by performing a dynamic I/O configuration.

In Figure 19 on page 60 (shared device), ZOSTEST can access device 180 through CHP 04 and CHP 10 if CHP 04 is defined to ZOSTEST in the IOCDS.

In Figure 19 on page 60 (unshared device), ZOSTEST can access either device 180 (unshared device) if one or both of the devices are assigned a new device number in the IOCDS.

When a device number conflict exists, logical partitions fail to activate. This happens when one of the following conditions occurs:

- The receiving logical partition was deactivated when a channel path is reassigned
- The receiving logical partition is deactivated after a channel path is reassigned

Failure to activate can result if options A or B are used. If a logical partition fails to activate, use option B or C to resolve the conflict and to activate the logical partition.

In Figure 19 on page 60, if ZOSTEST is not active when CHP 04 is reassigned, or ZOSTEST is deactivated and then activated after CHP 04 is reassigned, ZOSTEST does not activate until the conflict over device 180 is resolved.

If you resolve the conflict by using option B do the following steps:

1. Establish the correct configuration by configuring offline one of the channel paths (CHP 04 or CHP 10)
2. Configure offline and then online the other channel path

If it is necessary to have access to other devices on the first channel path, the operator can configure online the first channel path while the LP is active. Ignore the messages issued at the hardware console.

The following IOCP statement example shows coding that removes duplicate device number conflicts for shared devices.

```
CHPID PATH=(00),PARTITION=(ZOSPROD,REC), . . .
CHPID PATH=(04),PARTITION=(ZOSPROD,REC), . . .
CHPID PATH=(10),PARTITION=(ZOSTEST), . . .
CNTLUNIT CUNUMBR=0001,PATH=(00,04),UNITADD=80
CNTLUNIT CUNUMBR=0002,PATH=(10),UNITADD=80
IODEVICE ADDRESS=180,CUNUMBR=(0001,0002)
```

RoCE planning considerations

There is added support for RDMA over Converged Ethernet (RoCE) for networking. There are 10GbE and 25GbE variants. These cards are normally installed in 2 (or more) increments per production partition, in different Resource Groups (see the eConfig report), so as to provide high Reliability and Serviceability. RoCE cards are specified in the IOCP input via FUNCTION statements, giving them a Function ID (FID) and an association with a single active partition at a time (they are serially reusable, but remain in one partition at a time). RoCE Express2 and later adapters require a FID to be defined on a port basis rather than an adapter PCHID basis. RoCE Express adapters have 2 ports, but FIDs can use either port. See the *Input/Output Configuration Program User's Guide for ICP*, SB10-7172.

The 25GbE RoCE Express2 and the new 25GbE RoCE Express2.1 with enhanced code can provide flexibility to consolidate adapters or increase overall speed continue to have high reliability and performance established on prior adapters. This offers one PCIe adapter, one PCHID, and two ports per feature. The benefit of the 25GbE RoCE (Remote Direct Memory Access (RDMA) over Converged Ethernet) Express2 (#430) and 25GbE RoCE Express2.1 (#450) are recognized when connected to another 25GbE RoCE Express2 (#430) or 25GbE RoCE Express2.1 (#450). You should plan for multiple features when upgrading from a 10GbE RoCE Express (#411). However, the 25GbE RoCE Express2 (#430) and 25GbE RoCE Express 2.1 (#450) can also be used as a traditional networking card for TCP/IP communication when connected to a 25GbE capable fabric.

There is added support for RDMA over Converged Ethernet (RoCE), 10GbPS links for networking. These cards are normally installed in 2 (or more) increments per production partition, in different Resource Groups (see the eConfig report), so as to provide high Reliability and Serviceability. RoCE cards are specified in the IOCP input via FUNCTION statements, giving them a Function ID (FID) and an association with a single active partition at a time (they are serially reusable, but remain one partition at a time. More details are in the *Input/Output Configuration Program User's Guide for ICP*, SB10-7172.

Example:

```
FUNCTION FID=10,PCHID=314,PART=((LP01),(LP02)),PNETID=NETWORK1
FUNCTION FID=11,PCHID=338,PART=((LP01),(LP02)),PNETID=NETWORK1
CHPID PCHID=380,PATH=(CSS(0,1,2,3),D0),TYPE=OSD,SHARED,PNETID=NETWORK1
CHPID PCHID=3B8,PATH=(CSS(0,1,2,3),D1),TYPE=OSD,SHARED,PNETID=NETWORK1
```

In this example, 2 10Gb RoCE cards are defined to partition LP01, and two OSA (OSD) cards are shared by the partitions (including LP01). z/OS CommManager setup requires they be paired, a conventional OSA card for setup, with a 10Gb RoCE. The pair of pairs is for RAS, as otherwise during certain system maintenance operations or card failures, there can be a loss of connectivity. The PNETID=NETWORK1 is a Physical Network ID, it is optional, but if specified, should be the same value. When used with a 10 Gb RoCE, the allowed OSA CHPID types are OSD, OSM, OSX. If there are more partitions that need to be RoCE connected, additional cards will need to be added, up to a maximum of 8 or 16, depending on the machine type and model.

Coupling facility planning considerations

The coupling facility provides shared storage and shared storage management functions for the sysplex (for example, high speed caching, list processing, and locking functions). Applications running on z/OS images in the sysplex define the shared structures used in the coupling facility.

The coupling facility allows applications, running on multiple z/OS images that are configured in a sysplex, to efficiently share data so that a transaction processing workload can be processed in parallel across the sysplex.

PR/SM allows you to define the coupling facility, which is a special logical partition (LP) that runs coupling facility control code. Coupling facility control code is Licensed Internal Code (LIC).

At LP activation, coupling facility control code automatically loads into the coupling facility LP from the Support Element hard disk. No initial program load (IPL) of an operating system is necessary or supported in the coupling facility LP.

Coupling facility control code runs in the coupling facility LP with minimal operator intervention. Operator activity is confined to the Operating System Messages task. PR/SM limits the hardware operator controls typically available for LPs to avoid unnecessary operator activity.

Coupling facility channel hardware provides the connectivity required for data sharing between the coupling facility and the CPCs directly attached to it. Coupling facility channels are point-to-point connections that require a unique channel definition at each end of the channel. See "Coupling facility channels" on page 77.

Dynamic I/O for Standalone Coupling Facility enables dynamic activation of a new or changed IODF on a standalone coupling facility CPC without requiring a re-IML or power-on reset (POR). For more information see "Dynamic activation of I/O configurations for stand-alone Coupling Facilities" on page 127.

Test or migration coupling configuration

You can run a test or migration coupling facility to test and develop data sharing applications. You can define a test or migration coupling facility LP on the same CPC where other LPs are:

- Running z/OS images connected to the coupling facility
- Running non-coupled production work

A single CPC configuration has the following consideration:

- Simultaneous loss of the coupling facility and any z/OS images coupled to it (a more likely possibility in a single CPC configuration) can potentially cause extended recovery times

You can define a test or migration coupling facility with or **without** coupling facility channel hardware. See "Defining internal coupling channels (TYPE=ICP)" on page 79 for information about how to define a test or migration facility **without** coupling facility channel hardware.

Production coupling facility configuration

It is recommended that you run your production applications on a sysplex that uses a production coupling facility configuration.

A properly configured production coupling facility configuration can reduce the potential for extended recovery times, achieve acceptable performance, and maximize connectivity to the coupling facility.

For production configurations, the use of one or more dedicated Coupling Facility engines is recommended; shared Coupling Facility engines are strongly discouraged. For more information, see "Coupling facility LPs using dedicated Central Processors (CPs) or dedicated Internal Coupling Facility (ICF) processors" on page 105.

Production coupling facility configuration for full data sharing

The preferred solution for a full data sharing (IMS, DB2®, VSAM/RLS) production parallel sysplex is a coupling facility configuration that consists of:

- One stand-alone coupling facility running as a single dedicated coupling facility LP to provide large capacity shared storage and maximum coupling facility channel connectivity (up to 64 coupling facility channels).

- A second stand-alone coupling facility, similarly configured, to reduce the possibility of a single point of failure. A second stand-alone coupling facility improves application subsystem availability by allowing fast recovery from one coupling facility to the other in the event of a coupling facility outage. Alternatively, an Internal Coupling Facility (ICF) feature can be used to provide the backup coupling facility. See "Internal Coupling Facility (ICF)" on page 64.

Notes:

1. The backup CF in the configuration must provide sufficient storage, processor, and connectivity resources to assume the workload of the other production CF in the event of its failure.

2. With the use of System-Managed CF Structure Duplexing for all relevant data sharing structures, it is possible to have a production data-sharing configuration that uses only 2 or more internal CFs, because duplexing avoids the "single point of failure" failure-isolation issue.

Production coupling facility configuration for resource sharing

A viable solution for a resource sharing (XCF Signaling, Logger Operlog, RACF®, BatchPipes®, Logger Logrec, Shared Tape, GRS, WLM Enclave Support, LPAR Clusters) production level parallel sysplex is a coupling facility configuration that consists of:

- One dedicated CF provides reduced cost of ownership without compromising sysplex availability or integrity.

- A second dedicated CF reduces the possibility of a single point of failure. A second CF improves application subsystem availability by allowing fast recovery from one coupling facility to the other in the event of a coupling facility outage.

These configurations offer the best performance, the best reliability, availability, and serviceability (RAS).

Note: The backup CF in the configuration must provide sufficient storage, processor, and connectivity resources to assume the workload of the other production CF in the event of its failure.

Internal Coupling Facility (ICF)

You can purchase and install one or more ICF features for use in coupling facility LPs. With this feature, the coupling facility runs on special ICF processors that no customer software can use. This feature allows the coupling facility function to be performed on the CPC without affecting the model group and thus without impacting software licensing costs for the CP resources used by the coupling facility. See "Considerations for coupling facilities using Internal Coupling Facility (ICF) processors" on page 105.

These features are ordered separately, and are distinguished at the hardware level from any general-purpose CPs, Integrated Features for Linux (IFLs), Integrated Information Processor (zIIPs). ICFs, IFLs, and zIIPs are perceived by the system as multiple resource pools.

With the CFCC Enhanced Patch Apply process, you can perform a disruptive install of new CFCC code on a CF image by deactivating and then reactivating the CF image, without the much greater disruption of a Power On Reset (POR) of the entire CPC that contains the CF image. Thus, availability is greatly improved.

Coupling facilities that reside on the same CPC as one or more z/OS parallel sysplex logical partitions are ideal for coupling resource sharing sysplexes (sysplexes that are not in production data sharing with IMS, DB2 or VSAM/RLS). You can simplify systems management by using XCF structures instead of FICON FC connections.

It is not recommended the use of coupling facilities that reside on the same CPC as one or more z/OS parallel sysplex logical partitions involved in data sharing that are in the same sysplex unless using System-Managed CF Structure Duplexing or Asynchronous Duplexing for Lock Structures for all relevant

data sharing structures. This is because of the possibility of double outages involving the simultaneous loss of a coupling facility image and one or more z/OS system images that are using the coupling facility for data sharing. Depending on the structure, a double outage can result in a significantly more involved recovery than a single outage of either a coupling facility or a z/OS image in isolation from one another. With System-Managed CF Structure Duplexing or Asynchronous Duplexing for Lock Structures for all relevant data sharing structures, it is possible to have a production data sharing configuration with a coupling facility image and one or more z/OS system images in the same sysplex on a single CPC. This is because duplexing provides the necessary failure isolation solution.

With the use of System-Managed CF Structure Duplexing for all relevant data sharing structures, it is possible to have a production data-sharing configuration that uses only 2 or more internal CFs, because duplexing avoids the "single point of failure" failure-isolation issue.

ICFs on stand-alone coupling facilities need configuration planning to account for storage and channels. The storage requirements for the CPC with an ICF installed likely increases, especially if software uses the coupling facility to provide additional function not available except when running a coupling facility in a parallel sysplex. For more information, see "Dynamic activation of I/O configurations for stand-alone Coupling Facilities" on page 127 .

Note: The number of ICFs on a z15 is limited only to the number of customer definable PUs for the model. There is a limit of 16 ICFs per LP.

Dynamic Coupling Facility Dispatching and Coupling Thin Interrupts

Coupling facility dispatching behavior on shared-engines is controlled via the dynamic coupling facility dispatching command (DYNDISP) for the coupling facility logical partition.

Available options for dynamic coupling facility dispatching are as follows:

- DYNDISP=OFF: The coupling facility does not do dynamic dispatching. It always retains control of the shared coupling facility processor from the time PR/SM dispatches it to the coupling facility image, until PR/SM undispatches it at end of timeslice. The coupling facility never voluntarily gives up control of the shared coupling facility processor. With this setting, a coupling facility using shared CPs would attempt to get all the CP resource it could even when there was no real work for it to do.

- DYNDISP=ON: The coupling facility uses a timer-based algorithm to determine when to voluntarily give up control of the shared coupling facility processor, and sets a timer interrupt to cause the image to be re-dispatched some number of milliseconds into the future; essentially, timer-based sharing of the processor. With DYNDISP=ON, the coupling facility monitors the request rate that is driving it and adjust its usage of CP resource accordingly. If the request rate becomes high enough, the coupling facility reverts to its original dispatching algorithm, constantly looking for new work. When the request rate lowers, the coupling facility again becomes more judicious in its use of CP resource.

- DYNDISP=THIN: The coupling facility voluntarily gives up control of the shared coupling facility processor whenever it runs out of work to do, relying on coupling thin interrupts to cause the image to get re-dispatched in a timely fashion when new work (or new signals) arrive at the coupling facility to be processed. This allows efficient sharing and timeslicing between the sharing coupling facility images and avoids many latencies inherent in polling-based techniques.

For more information on using DYNDISP options to share processors, see:

Coupling Thin Interrupts and Coupling Facility Performance in Shared Processor Environments

http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102400.

Note: DYNDISP=THIN is the Default setting for shared-engine coupling facilities.

With DYNDISP=THIN, thin interrupts are used to initiate dispatching the coupling facility. Thin interrupts are generated on the coupling facility when:

- A coupling facility command is received by a shared-engine coupling facility image
- A coupling facility signal is received by a shared-engine CF image (for example, arrival of a coupling facility-to-coupling facility duplexing signal)

- Completion of a coupling facility signal previously sent by the coupling facility (for example, completion of a coupling facility-to-coupling facility duplexing signal).

The interrupt causes the receiving partition to be dispatched by PR/SM, if it is not already dispatched. This allows the request, signal, or request completion to be recognized and processed in a more timely manner. Once the image is dispatched, existing poll for work logic in both CFCC and z/OS can be used largely as is to locate and process the work. The new interrupt simply expedites the re-dispatching of the partition. When using DYNDISP=THIN, the coupling facility will relinquish the processor as soon as all available pending work has been exhausted (or when PR/SM undispatches it off the shared processor, whichever comes first).

In back-up mode or in certain test configurations, the coupling facility has a very low request rate so it throttles back to very low CP usage. Using DYNDISP=THIN, the requests themselves will drive PR/SM to dispatch the coupling facility as requests arrive at the coupling facility with minimal delay that does not adversely affect the performance of the overall system. Since the coupling facility is not consuming more CP resource than it needs to, you can now set the processor weights for the coupling facility to a value high enough to handle the load if the coupling facility was to take over for a failing primary coupling facility. If the primary coupling facility does fail, the requests can be moved immediately to the back-up coupling facility which can then get the CP resource it needs automatically with properly defined LP weights.

Dynamic coupling facility dispatching is particularly useful in configurations where less than one CP of capacity is needed for use by a coupling facility. To enable dynamic coupling facility dispatching, use the DYNDISP coupling facility control code command. See "Coupling facility control code commands" on page 71. It is not recommended to have more than one coupling facility LP with DYNDISP=OFF sharing a given shared CF processor. When shared processors are used for CF images, it is recommended that each of the CF images be configured with DYNDISP=THIN. Alternatively, it is recommended that the sharing CF images that are expected to have low activity (for example, standby/back-up coupling facilities, or test coupling facilities) be configured with DYNDISP=THIN, and any single sharing CF image that is expected to have high activity be configured with DYNDISP=OFF. This will allow the low activity CF images to voluntarily give up the shared processor when they are not actively using it, which makes the shared processor more available for use by a single high activity CF image.

Coupling Facility Configuration	All Dedicated CPs or ICFs	Shared ICF	Shared CP	Shared ICF
CF Model	-	Z mainframe with only ICF processors	Z mainframe with 1 or more general-purpose CPs	Z mainframe with 1 or more general-purpose CPs
Dynamic Dispatching Default Value	Off	THIN	THIN	THIN

Asynchronous coupling facility duplexing for lock structures

The z15 supports the Asynchronous Coupling Facility (CF) duplexing to improve the performance of the duplexing of lock structures by:

- Reducing z/OS, CF, and link utilization overhead costs associated with synchronous duplexing of lock structures
- Improving performance with cross-site duplexing of lock structures at distance
- Maintaining robust failure recovery capability through the redundancy of duplexing.

Asynchronous CF duplexing for lock structures is designed to allow secondary structure updates to be performed asynchronously with respect to primary updates to:

- Drive out cross-site latencies that exist today when replicating CF data across distance
- Avoid the need for synchronous speed-of-light communication delays during the processing of every duplexed update operation.

Asynchronous CF duplexing for lock structures provides all of the same benefits that system-managed coupling facility structure duplexing provides (see "System-managed coupling facility structure duplexing" on page 67), but with better performance and lower overhead.

System-managed coupling facility structure duplexing

A set of parallel sysplex architectural extensions is provided for support of *system-managed* duplexing of coupling facility structures for high availability. All three structure types, cache, list, and locking, can be duplexed using this architecture.

Benefits of system-managed CF structure duplexing include:

- **Availability:** Faster recovery of structures by having the data already in the second CF.
- **Manageability and Usability:** A consistent procedure to set up and manage structure recovery across multiple exploiters
- **Cost Benefits:** Enables the use of non-stand-alone CFs (for example, ICFs) for all resource sharing and data sharing environments.

Preparations for CF duplexing includes the requirement to connect coupling facilities to one another with coupling links. The required CF-to-CF connectivity is bi-directional, so that signals can be exchanged between the CFs in both directions. A single peer-mode coupling link between each pair of CFs can provide the required CF-to-CF connectivity; however, for high availability at least two peer-mode links between each pair of CFs are recommended.

While peer-mode CHPIDs cannot be shared between multiple coupling facility images, they can be shared between a single coupling facility image and one or more z/OS images. At least two such links is recommended for high availability. In addition, ICR SR links can provide the ability to actually share the same physical link between multiple CF images. By defining multiple CHPIDs on the same physical ICR SR link, the individual CHPIDs can be defined for a single CF image while the physical link is being shared by multiple CF images.

Coupling Facility Virual Flash Memory Exploitation

Coupling facility Virtual Flash Memory provides a way to get high total storage capacity for a CF structure, without needing to define excessively large amounts of structure real memory. Coupling facility Virtual Flash Memory also provides resiliency and capacity in the event of such backups. Coupling facility Virtual Flash Memory provides a way to get high total storage capacity for a CF structure, without needing to define excessively large amounts of structure real memory.

Coupling facility Virtual flash Memory also provides resiliency and capacity in the event of such backups. Initial coupling facility Virtual Flash Memory exploitation is targeted for MQ shared queues application structures. It provides standby capacity to handle MQ shared queue buildups during abnormal situations, such as where *putters* are putting to the shared queue, but *getters* are transiently not getting from the shared queue. Virtual Flash Memory in the system is assigned to a CF partition via hardware definition panels, just like it is assigned to the z/OS partitions. The CFRM policy definition permits the desired maximum amount of Virtual Flash Memory to be used by a particular structure, on a structure-by-structure basis. Structure size requirements for real memory get somewhat larger at initial allocation time to accommodate additional control objects needed to make use of Virtual Flash Memory. The CFSIZER's structure recommendations will take these additional requirements into account, both for sizing the structure's Virtual Flash Memory usage itself, and for the related real memory considerations.

Single CPC software availability sysplex

For single CPC configurations, Z mainframe can use an ICF to form a single CPC sysplex, providing significant improvement in software continuous operations characteristics when running two z/OS LPs in data-sharing mode versus one large z/OS image. For these configurations, overall RAS is improved over that provided by a single z/OS image solution. Hardware failures can take down the entire single CPC sysplex, but those failures are far less frequent than conditions taking down a software image, and planned software outages are the predominant form of software image outages in any case. Forming a single CPC sysplex allows software updates to occur in a "rolling" IPL fashion, maintaining system availability throughout. An LPAR cluster is one example of a single CPC sysplex which has significantly

improved system availability over a single LP. For additional benefits provided by an LPAR cluster using IRD technology, see *z/OS Intelligent Resource Director*, SG24-5952.

Coupling facility nonvolatility

Continuous availability of the transaction processing workload in a coupling facility configuration requires continuous availability of the shared structures in the coupling facility. To help ensure this availability, you must provide an optional backup power supply to make coupling facility storage contents nonvolatile across utility power failures.

Nonvolatility choices

The following table indicates the optional non-volatility choices available and their capabilities:

Table 7. Nonvolatility choices for coupling facility LPs

Nonvolatility Choices	z15
Uninterruptible power supply (UPS) (See Notes [1])	Yes
Internal Battery Feature (IBF)	Yes
Local Uninterruptible Power Supply (LUPS) (See Notes [2])	Yes

Notes:

1. Optional uninterruptible power supply (UPS) provides a secondary power source for use during extended utility power outages allowing continuous coupling facility operation.

2. The optional Local Uninterruptible Power Supply supports 0 - 18 minutes of full power operation.

Setting the conditions for monitoring coupling facility nonvolatility status

In addition to installing an optional backup power supply to help ensure continuous availability, you **must** also set the conditions by which the coupling facility determines its volatility status. Software subsystems with structures defined in the coupling facility can monitor this status. Use the coupling facility control code MODE command as follows:

- MODE NONVOLATILE sets coupling facility volatility status to nonvolatile and should be used if a floor UPS is available to the CPC. Coupling facility control code does **not** monitor the installation or availability of UPS but maintains a nonvolatile status for the coupling facility.

- MODE VOLATILE sets coupling facility volatility status to volatile and should be used if no backup power supply is installed and available. Coupling facility control code maintains volatile status for the coupling facility even if a backup power supply is installed and available.

The coupling facility MODE setting is saved across power-on reset and activation of the coupling facility. You can use online help from the Operator Messages panel to get additional information about coupling facility control code commands.

Coupling facility mode setting

The following table summarizes the relationship between the coupling facility MODE setting, and the resulting conditions you can expect if utility power fails at your site.

Table 8. Coupling facility mode setting

CF MODE setting	Local UPS or IBF installed	Results on Utility Power Failure
VOLATILE	Yes	Ride out utility power failure on UPS/IBF. (Setting the mode to VOLATILE here would be a configuration error because there is local UPS or IBF to provide nonvolatility.)

CF MODE setting	Local UPS or IBF installed	Results on Utility Power Failure
VOLATILE	No	Machine down unless alternate floor level UPS/IBF provided.
NONVOLATILE	Yes	Ride out utility power failure on UPS/IBF.
NONVOLATILE	No	Machine down unless alternate floor level UPS/IBF provided. **Note:** This is the recommended setting when providing floor-wide UPS/IBF backup.

Table 8. Coupling facility mode setting (continued)

Note: Reflects the real-time status of the power (volatile or nonvolatile).

Coupling facility LP definition considerations

You can define coupling facility mode for an LP at the Hardware Management Console or Support Element console using the **Customize/Delete Activation Profiles** task.

You can define coupling facility LPs with shared or dedicated CPs on all Z mainframes. Coupling facility LPs must be defined with at least 1024 MB of central storage. See .

Coupling facility LPs do *not* support some LP definition controls typically available to other LPs. For coupling facility LPs, you *cannot* define:

- Reserved central storage (coupling facility LPs do *not* support dynamic storage reconfiguration)
- Cryptos
- Automatic load
- Automatic load address
- Automatic load parameters

Internal Coupling Facility (ICF)

You can install one or more internal coupling facility (ICF) features. See .

Coupling facility LP storage planning considerations

You must define at least 1024 MB of central storage for a coupling facility LP to activate.

This storage is reserved for coupling facility control code use and *cannot* be used for other purposes. Minimum storage size for coupling facilities is *primarily* a function of the coupling facility control code level. This implies that, over time, the minimum storage size required by a coupling facility on a particular machine can grow as new coupling facility control code updates are applied.

You must also define additional storage to accommodate the shared structures and dump space used by software subsystems using the coupling facility.

You can also define Virtual Flash Memory for use by the coupling facility logical partition.

Structures, dump space, and coupling facility LP storage

A coupling facility allocates storage for structures and for dump space based on values specified in the SIZE, INITSIZE, and DUMPSPACE parameters of the coupling facility resource management (CFRM) policy used for the coupling facility.

Structures consist of control objects and data elements. The control objects include entries and various other control structures used to manipulate the entries. The data elements store user data associated with structure entries.

Dump space is storage in the coupling facility set aside for use as a dump table when a structure dump is taken. Dump tables are used for application development and problem determination purposes.

Table 9. Maximum central storage for a z15 / LinuxONE III model

z15 Model	LinuxONE III Model	Maximum Central Storage (GB)
Model T02	Model LT2	8128
Model T01	Model LT1	8128

Estimating coupling facility structure sizes

Estimating coupling facility structure sizes is useful to system programmers to help ensure that there is enough coupling facility storage to meet application needs. The estimation of the minimum central storage requirements of your List or Cache structure has been superseded by the following two commands used by the CFSizer Utility to make space calculations in the coupling facility:

- **Compute list-structure parameters:** For computing the minimum central storage requirements for a List structure. (What z/OS calls *lock structures* are actually a special form of list structure.)
- **Compute cache-structure parameters:** For computing the minimum central storage requirements for a Cache structure.

Important: When implementing a new CFLEVEL in your configuration, redetermine the size of List and Cache structures using the CFSizer Utility and update the CFRM policy with the newly acquired values. For more information about the CFSizer Utility, see the following link: http://www.ibm.com/systems/support/z/cfsizer.

Dump space allocation in a coupling facility

Dump space is storage you define using the DUMPSPACE parameter in the coupling facility resource management (CFRM) policy. It is set aside for the creation of dump tables. Dump tables are portions or snapshots of a structure typically saved for application development or problem determination purposes. The coupling facility allocates dump space in multiples of the coupling facility storage increment.

Dump tables for several different structures can exist in dump space at the same time. The amount of storage in any one dump table depends on the following factors:

- Amount of information you want to save to a dump table

  The software subsystem can request the portions of a structure that are to be captured in the dump table. For example, lock tables, lists within a list structure, or directory entries belonging to particular storage classes or castout classes.

- Free dump space

  Structures share dump space. If a structure is using some of the dump space, other structures cannot use that portion of the dump space until it is released.

- Characteristics of the structure saved to a dump table

  When saving structure objects to a dump table, the amount of dump space used depends on the parameters specified for the structure. For example, list entry size, the number of list entries in a specified list, the number of directory entries in a castout class, or whether adjunct data is included.

The coupling facility can return the maximum requested dump space value. This value indicates the largest amount of dump space requested by the software subsystems using the coupling facility. This value allows you to adjust the amount of allocated dump space to better match actual usage.

Coupling facility LP activation considerations

At LP activation, coupling facility control code automatically loads into the coupling facility LP from Support Element hard disk. No initial program load (IPL) of an operating system is necessary or supported in the coupling facility LP.

All coupling facility channel path types targeted to be brought online are automatically configured online.

Note: All channel paths types that are targeted to be brought online will automatically be configured online if the coupling facility LP is redefined as a General mode LP.

Coupling facility shutdown considerations

Important: It is important to properly remove all structures from a coupling facility that will be permanently taken out of the sysplex before shutting down the coupling facility. Failure to remove all structures might result in a pending condition (or transitioning structure) when attempting to allocate structures in a new coupling facility. Reactivation of the old coupling facility might be required to resolve the condition.

A running coupling facility LP can contain structures and data important to a sysplex. Make sure that you take proper precautions to preserve these structures and data **before** a power off, power-on reset (POR), LP deactivation, or shutdown (using the coupling facility control code SHUTDOWN command) of the coupling facility LP. Coupling facility structures and data will **not** survive any of these actions. Loss of coupling facility structures and data occurs regardless of the volatility state of the coupling facility including situations where a battery backup is in place.

It is recommended to use the coupling facility SHUTDOWN command to shut down a CF image, because it performs a check to make sure that there are no allocated structure instances in the CF before proceeding to shut down the coupling facility.

For more information about removing, replacing, or shutting down a coupling facility, see *z/OS MVS Setting Up a Sysplex*, SA22-7625.

Coupling facility LP operation considerations

Coupling facility control code runs in the coupling facility LP with minimal operator intervention. Operator activity is confined to the **Operating Systems Messages** task. PR/SM limits the hardware operator controls typically available for LPs to avoid unnecessary operator activity.

Coupling facility LPs only support the following tasks typically available to the hardware console operator:

- Activate
- Deactivate
- Operating System Messages

Coupling facility control code commands

Coupling facility control code does provide a limited set of hardware operator controls unique to the coupling facility LP. These controls are available from the Operating System Messages window. From this window, you can enter the HELP command to display coupling facility control code command syntax.

Coupling facility control code provides the following commands for use in the coupling facility LP:

- CONFIGURE (configure coupling facility channel paths online or offline)
- CFDUMP (force a non-disruptive dump of the CF)
- CP (configure a central processor online or offline)
- DISPLAY (display coupling facility resource information)
- DYNDISP (turn dynamic coupling facility dispatching on or off or request the use of Thin Interrupts for a coupling facility LP). See <u>"Dynamic Coupling Facility Dispatching and Coupling Thin Interrupts" on page 65</u>
- HELP (display coupling facility control code command syntax)
- NDDUMP (sets or modifies nondisruptive dumping options)
- MODE (define coupling facility volatility mode)
- SHUTDOWN (shutdown coupling facility operation)
- TIMEZONE (sets timezone offset from Greenwich Mean Time for a coupling facility)

- TRACE (sets or modifies tracing options).

Notes:

1. Support for the CP and HELP coupling facility control code commands is available on all Z mainframes.

2. The settings established using DYNDISP, MODE and TIMEZONE commands are recorded in the policy file for the coupling facility. As a result, all values are persistent across resets, deactivations, and reactivations.

Coupling facility level (CFLEVEL) considerations

To support migration from one coupling facility level to the next, you can run different levels of the coupling facility concurrently as long as the coupling facility LPs are running on different CPCs. CF LPs running on the same CPC share the same coupling facility control code EC level.

When migrating CF levels, lock, list, and cache structure sizes might need to be increased to support new function. The amount of space needed for the current CFCC levels must be redetermined by visiting the CFSizer tool at http://www.ibm.com/systems/support/z/cfsizer.

CPC support for coupling facility code levels

The following table summarizes CPC support for the different coupling facility code levels.

CPC models	Coupling facility code level							
	Level 24	Level 23	Level 22	Level 21	Level 20	Level 19	Level 18	Level 17
8561 z15	EC P46603 MCL 001	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3907 z14	N/A	EC P41419 MCL 005	EC P42606 MCL 007	N/A	N/A	N/A	N/A	N/A
3906 z14		EC P41419 MCL 005	EC P42606 MCL 007	N/A	N/A	N/A	N/A	N/A
2965 z13s®	N/A	N/A	N/A	EC P08416 MCL 009	EC N98780 MCL 007	NA	NA	NA
2964 z13®	N/A	N/A	N/A	EC P08416 MCL 009	EC N98780 MCL 007	NA	NA	NA

Table 10. CPC support for coupling facility code levels

Notes:

1. Previous machine types that support the coupling facility are included for completeness and sysplex connectivity purposes.

2. The (Ver n.n.n) refers to the version of code installed in the Support Element. You can verify what version code you have by looking at the title bar on the Workplace window of your Support Element.

3. All MCLs cited in the table are current as of the publication date of this edition and can be superseded by later MCLs. The MCLs cited in the table are the latest service levels as of the current publication date of this edition. They do not necessarily reflect the minimum service level required for a particular CFLEVEL but rather the recommended service level. For the latest EC and MCL information, use the service support link to view the appropriate PSP bucket subset ID for hardware and software maintenance information.

Level 24 coupling facility

A level 24 coupling facility (CFLEVEL=24) provides the following enhancements:

- CFCC Fair Latch Manager

  – This is an enhancement to the internals of the Coupling Facility (CFCC) dispatcher to provide CF work management efficiency and processor scalability improvements, as well as improve the "fairness" of arbitration for internal CF resource latches across tasks.

- CFCC Message Path Resiliency Enhancement

  – CF Message Paths use a z/OS-provided system identifier (SYID) to uniquely identify which z/OS system image, and instance of that system image, is sending requests over a message path to the CF.

  – When a z/OS system IPLs, message paths are supposed to be deactivated via system reset, and their SYIDs are supposed to be cleared in the process; During IPL, z/OS will then re-activate the message paths with a new SYID that represents the new instance of z/OS that is currently using the paths

  – On rare occasions, a message path may not get deactivated during system reset / IPL processing, leaving the message path left active with the z/OS image's OLD, now-obsolete SYID. From the CF's perspective, the incorrect SYID persists, and prevents delivery of signals to the z/OS image currently using that message path.

  – This new resiliency mechanism will transparently recover for this "missing" message path deactivate (if and when that ever happens)

- CFCC Change Shared-Engine CF Default to DYNDISP=THIN

  – Make DYNDISP=THIN the default mode of operation for coupling facility images that use shared processors

Level 23 coupling facility

A level 23 coupling facility (CFLEVEL=23) provides the following enhancements:

- Asynchronous Cache Cross-Invalidation (XI)

  – Asynchronous Cache Cross-Invalidation (XI) is a sysplex capability for performance, scalability, and improved cross-site operation. This function allows the cache coherency messages that flow around the sysplex to maintain data integrity to be performed in an asynchronous fashion rather than synchronously. Exploitation must provide support to sync up with the asynchronous cross-invalidate messages at critical points in its processing, such as at transaction commit. The asynchronous protocol is expected to reduce CF cache structure service times and sysplex coupling overhead, particularly in sysplex environments that involve multiple sites with significant cross-site distances involved. The asynchronous protocol avoids some of the distance latencies associated with the communication of XI messages across inter-site distance.

Level 22 coupling facility

A level 22 coupling facility (CFLEVEL=22) provides the following enhancements:

- Notification Delay and Round Robin Support for List and Key-Range Monitoring
- CFCC Encryption Support
- CLTE Performance Enhancements
- Controller/Follower duplexing enhancements for Cache Structures
- CFCC dispatcher enhancements

Level 21 coupling facility

A level 21 coupling facility (CFLEVEL=21) provides the following enhancements:

- Asynchronous CF duplexing for lock structures when CFLEVEL 21 is at service level 02.16 or higher.
- A CF Dump Reason Code added to the dump header when a CF non-disruptive dump is taken. This allows for a quick evaluation of why the dump was taken.

- The coupling facility will provide identifying information to the service processor similar to what other operating systems running in other logical partitions currently provide.

CPC Support

See Table 10 on page 72 for a listing of the CPCs that support a level 21 coupling facility.

Software Corequisites

For a list of the software levels that use the function and levels that can coexist with CFLEVEL=21, see the "Summary of CFLEVEL Functions" section of the *z/OS MVS Setting Up a Sysplex* document.

Level 20 coupling facility

A level 20 coupling facility (CFLEVEL=20) provides the following enhancements:

- ICA SR coupling link support
- CFCC processing scalability support
- 256 coupling CHPIDs per CPC support
- Support for up to 141 ICF processors per z Systems server

  – The maximum number of logical processors in a Coupling Facility Partition remains at 16.

- Large Memory Support

  – Improve availability/scalability for larger CF cache structures and data sharing performance with larger DB2 Group Buffer Pools (GBP).

  – This support removes inhibitors to using large CF structures, enabling use of Large Memory to appropriately scale to larger DB2 Local Buffer Pools (LBP) and Group Buffer Pools (GBP) in data sharing environments.

  – CF structure size remains at a maximum of 1 TB

CPC Support

See Table 10 on page 72 for a listing of the CPCs that support a level 20 coupling facility.

Software Corequisites

For a list of the software levels that use the function and levels that can coexist with CFLEVEL=20, see the Summary of CFLEVEL Functions section of the *z/OS MVS Setting Up a Sysplex* document.

Level 19 coupling facility

A level 19 coupling facility (CFLEVEL=19) provides the following enhancements:

- Coupling Thin Interrupts

  – Improves the efficiency of environments where shared engines are used as Coupling Facilities. While dedicated engines continue to be recommended to obtain the best Coupling Facility performance, Coupling Thin Interrupts may help to facilitate the use of a shared pool of engines, helping to lower your hardware acquisition costs.

  – You may now experience Coupling Facility response time improvements or more consistent response times when using Coupling Facilities with shared engines. This may allow more environments with multiple Coupling Facility images to coexist in a server, and share Coupling Facility engines with reasonably good performance. The response time for asynchronous Coupling Facility requests may also be improved as a result of using Coupling Thin Interrupts on the z/OS host system, regardless of whether the Coupling Facility is using shared or dedicated engines.

- XI Detection

  – Cross-invalidiate and list notification error detection

    Cross-invalidate (XI) and list notification (LN) signals sent by a coupling facility will now receive meaningful replies from the target systems that provide a secondary message exception code and

additional diagnostics if the XI or LN experienced an error in its processing. The CF can take additional diagnostic steps like tracing relevant data and/or marking the structure damaged and taking a non-disruptive structure dump.

- Coupling facility flash exploitation

Coupling facility flash provides a way to get high total storage capacity for a CF structure without needing to define excessively large amounts of structure real memory. Coupling facility flash also provides resiliency and capacity in the event of such backups.

- Initial coupling facility flash exploitation is targeted for MQ shared queues application structures. It provides standby capacity to handle MQ shared queue buildups during abnormal situations, such as where *putters* are putting to the shared queue, but *getters* are transiently not getting from the shared queue. Flash memory in the CPC is assigned to a CF partition through hardware definition panels, just like it is assigned to the z/OS partitions. The CFRM policy definition permits the desired maximum amount of Flash memory to be used by a particular structure, on a structure-by-structure basis. Structure size requirements for real memory get somewhat larger at initial allocation time to accommodate additional control objects needed to make use of Flash memory. The CFSIZER's structure recommendations take these additional requirements into account, both for sizing the structure's Flash usage itself, and for the related real memory considerations.

Important: CF structures allocated in a CFLEVEL 19 coupling facility might need to be significantly larger than in previous CFLEVELs, in order to be allocated with a similar number of usable structure objects. It is highly recommended to use the CFSIZER tool: http://www.ibm.com/systems/support/z/cfsizer.

CPC Support

See for a listing of the CPCs that support a level 19 coupling facility.

Software Corequisites

For a list of the software levels that use the function and levels that can coexist with CFLEVEL=19, see the Summary of CFLEVEL Functions section of the *z/OS MVS Setting Up a Sysplex* document.

Level 18 coupling facility

A level 18 coupling facility (CFLEVEL=18) provides the following enhancements:

- CF cache write-around support for performance enhancement. DB2 can use a conditional write command during batch update/insert processing to decide which entries should be written to the GBP caches and which entries should be written around the cache to disk.
- CF cache registration attachment validation for error detection.
- CF large structure testing capability.
- CFCC non-disruptive coupling facility dumping support, for improved coupling facility serviceability
- CF structure expansion/contraction/reapportionment performance enhancement for list and cache structures
- Increase in the maximum number of CF structure instances (SIDs) per CF image from 1023 to 2047
- Support for greater than 32 connectors to a CF list/lock structure
- Increase in the number of coupling CHPIDs that can attach to a CF image from 64 to 128

Important: CF structures allocated in a CFLEVEL 18 coupling facility might need to be significantly larger than in previous CFLEVELs, in order to be allocated with a similar number of usable structure objects. It is highly recommended to use the CFSIZER tool: http://www.ibm.com/systems/support/z/cfsizer.

CPC Support

See for a listing of the CPCs that support a level 18 coupling facility.

Software Corequisites

For a list of the software levels that use the function and levels that can coexist with CFLEVEL=18, see the Summary of CFLEVEL Functions section of the *z/OS MVS Setting Up a Sysplex* document.

Level 17 coupling facility

A level 17 coupling facility (CFLEVEL=17) provides the following enhancements:

- CF cache write-around support for performance enhancement. DB2 can use a conditional write command during batch update/insert processing to decide which entries should be written to the GBP caches and which entries should be written around the cache to disk.

- CF cache registration attachment validation to protect against lost cross-invalidate signals.

- CFCC non-disruptive coupling facility dumping support, for improved coupling facility serviceability. Additional capabilities for the CFCC non-disruptive dumping support are provided to allow for capturing and collecting extended disagnostic structure data from CF structures that have encountered an error that are non-disruptive and performed in the background.

- CF structure expansion/contraction/reapportionment performance enhancement for list and cache structures

- Increase in the maximum number of CF structure instances (SIDs) per CF image from 1023 to 2047

- Support for greater than 32 connectors to a CF list/lock structure

- Increase in the number of coupling CHPIDs that can attach to a CF image from 64 to 128

Important: CF structures allocated in a CFLEVEL 17 coupling facility might need to be significantly larger than in previous CFLEVELs, in order to be allocated with a similar number of usable structure objects. It is highly recommended to use the CFSIZER tool: http://www.ibm.com/systems/support/z/cfsizer.

CPC Support

See for a listing of the CPCs that support a level 17 coupling facility.

Software Corequisites

For a list of the software levels that use the function and levels that can coexist with CFLEVEL=17, see the "Summary of CFLEVEL Functions" section of the *z/OS MVS Setting Up a Sysplex* document.

Coupling Facility Resource Management (CFRM) policy considerations

To define how to manage z/OS images and coupling facilities in the sysplex, you must specify hardware configuration information in the coupling facility resource management (CFRM) policy as follows:

- Coupling facility node descriptor information

  You must identify each coupling facility in the sysplex and the processor complex on which it is running. To do so, you must specify the following information in the CFRM policy:

  CFRM parameter
  Description

  PLANT
  Plant of manufacture

  SEQUENCE
  Machine sequence number

  SIDE
  Machine side

  TYPE
  Machine type

  MFG
  Manufacturer

  CPCID
  CPC identifier

  This information is available on the CPC Details panel. You can access the CPC Details panel by opening the CPC object that is running the coupling facility LP.

• LP information for the coupling facility

  For a coupling facility residing on a z15 model, the partition ID specified on the activation profile for the CF image on the Support Element or Hardware Management console must match the number specified in the PARTITION keyword of the CF statement in the policy information defined in the CFRM policy. It is recommended that the LP names for the CF LPs in IOCP input files match the names used in the NAME keyword in the CF statement in the CFRM policy.

  You can find the LP names in either the IOCP or HCD reports.

Coupling facility channels

Coupling facility channels are channels that use fiber optic cables (CE LR and ICA SR coupling links) or internal memory bus (ICP channel paths) to provide the connectivity for data sharing between a coupling facility and the central processor complexes (CPCs) or logical partitions (LPs) directly attached to it.

The class of CHPIDs, known as peer mode channels, provide both sender and receiver capability on the same link. Peer mode links come in these varieties: Internal Coupling channels, ICA (TYPE=CS5), and CE LR (TYPE=CL5). Each ICP, CS5, and CL5 channel can be configured as an unshared channel path to a single coupling facility or z/OS image, or as a shared channel path among several z/OS images and one coupling facility image.

Note: The following bulleted items only describe z/OS to coupling facility connections. However, they also apply to coupling facility duplexing connections (CF to CF).

Coupling facility channels:

• Require a point-to-point connection (direct channel attach between a CPC or LP and a coupling facility). Internal Coupling channels can only be used to connect a coupling facility and LPs on the same CPC.

• Can be used to connect a coupling facility to other LPs on the same CPC when a coupling facility is one of multiple LPs running on a single CPC. Internal coupling channels are recommended for these connections.

• Can be redundantly configured (two or more coupling facility channels from each CPC involved in coupling facility data sharing) to enhance availability and avoid extended recovery time. This does not apply to Internal Coupling channels.

• Require ICP, ICA SR, or CE LR coupling links channel path definition at the coupling facility end of a coupling facility channel connection.

• Require ICP, ICA SR, or CE LR coupling links channel path definition at the z/OS (and, for System-Managed CF Structure Duplexing, the coupling facility) end of a coupling facility channel connection.

• Require a ICP, ICA SR, or CE LR coupling link channel path be connected to a CE LR or ICA SR coupling link. You must define the ICP channel paths in pairs and you must connect each pair. You connect an ICP channel path to an ICP channel path by specifying the CPATH keyword on the CHPID statement for every ICP channel path.

Internal Coupling channel

The Internal Coupling channel emulates the coupling facility functions in LIC between images within a single system. Internal Coupling channel implementation is completely logical, requiring no channel or even cable hardware. However, a CHPID number must be defined in the IOCDS. Internal Coupling channels cannot be used for coupling connections to images in external systems.

Partitions with Internal Coupling channels can also have coupling facility channels which allow external system coupling. ICs, which use the system bus, are extremely fast (approximately 6 GB/second).

Internal Coupling channels have channel path type ICP (Internal Coupling Peer). Internal Coupling channels are identified by 2 CHPIDs representing the two ends of a coupling link. The rules that apply to the ICP CHPID type are the same as those which apply to external coupling link types, with the exception that the following functions are not supported:

• Service On/Off
• Reset I/O Interface

- Reset Error Thresholds
- Swap Channel Path
- CHPID Reassign
- Channel Diagnostic Monitor
- R/V
- Configuration Manager Vital Product Data (VPD)

Internal coupling channels have improved coupling performance over coupling facility channels.

Coupling Express LR (CE LR)

The z15 supports the Coupling Express LR (CE LR), a two-way ethernet-based, long-distance coupling card that utilizes a new coupling channel type: CL5. The CE LR is designed to drive distances up to 10 km unrepeated and up to 100 km with a qualified DWDM. The CE LR supports a link data rate of 10 Gbps and is designed to support 8 or 32 subchannels (devices) per CHPID and up to 4 CHPIDs per port. The maximum number of CE LR adapter features is 32 per z15. The CE LR resides in a PCIe I/O drawer card slot.

Note: The link data rates do not represent the performance of the links. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload.

The CE LR can only be used for coupling connectivity between Z mainframes, and the CE LR can only connect to another CE LR. It is recommended that you order CE LR on the z15 machines used in a Parallel Sysplex to help ensure long-distance coupling connectivity with future processor generations. The CE LR requires a 9u single-mode fiber cable. Refer to *Planning for Fiber Optic Links (FICON/FCP, Coupling Links,Open System Adapters, and zHyperLink Express)*, GA23-1408 and *Maintenance for Fiber Optic Links (FICON/FCP, Coupling Links, Open System Adapters, and zHyperLink Express)*, SY27-7696 which can be found in the Library section of Resource Link at: http://www.ibm.com/servers/resourcelink.

Integrated Coupling Adapter (ICA SR)

The z15 supports the Integrated Coupling Adapter (ICA SR), a two-way short distance coupling fanout that utilizes a new coupling channel type: CS5. The ICA SR utilizes PCIe Gen3 technology, with x16 lanes that are bifurcated into x8 lanes for coupling. The ICA SR is designed to drive distances up to 150m and support a link data rate of 8 GBps. It is also designed to support up to 4 CHPIDs per port and 8 subchannels (devices) per CHPID. The maximum number of ICA SR fanout features is limited to 40 per Z mainframes.

Note: The link data rates do not represent the performance of the links. The actual performance is dependent upon many factors including latency through the adapters, cable lengths, and the type of workload.

The ICA SR can only be used for coupling connectivity between Z mainframes, and the ICA SR can only connect to another ICA SR. It is recommended that you order ICA SR (#0172) on the z15 machines used in a Parallel Sysplex to help ensure short-distance coupling connectivity with future processor generations. The ICA SR fanout requires new cabling. For distances up to 100 m, clients can choose the OM3 fiber type. For distances up to 150m, clients must choose the OM4 fiber type. Refer to *Planning for Fiber Optic Links (FICON/FCP, Coupling Links, Open System Adapters, and zHyperLink Express)*, GA23-1408 and *Maintenance for Fiber Optic Links (FICON/FCP, Coupling Links, Open System Adapters, and zHyperlink Express)*, SY27-7696 which can be found in the Library section of Resource Link at: http://www.ibm.com/servers/resourcelink.

Coupling facility channels (TYPE=ICP, TYPE=CS5, or TYPE=CL5)

You can configure a (ICP, CS5, or CL5 channel path as:

- An unshared dedicated channel path to a single LP
- 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

- A shared channel path that can be shared between at most one coupling facility image and one or more z/OS images.

Shared coupling facility channel path recommendations

The following are recommended:

1. For shared coupling facility channel paths, make sure that only LPs that need to use the channel path have it configured online. Doing so eliminates unnecessary traffic on the channel path from those systems that have it online but do not have the attached coupling facility in the active CFRM policy.

2. These channel paths can result in 'Path Busy' conditions when another LP is using the path. This situation can result in delays in getting requests to the coupling facility on this path. The number of 'Path Busy' conditions can be found in the RMF CF Subchannel Activity report in the BUSY COUNTS column labeled PTH. As a guideline, if this count exceeds 10% of the total requests, you should consider *not* sharing the channel path or adding additional coupling facility channel paths.

Defining internal coupling channels (TYPE=ICP)

Internal coupling channels are virtual attachments and, as such, require no real hardware. However, they do require CHPID numbers and they do need to be defined in the IOCDS.

You must define an even number of ICP channel paths and you must connect them in pairs. A connected pair of ICP CHPIDs is called an *internal coupling link*. Both ends of an internal coupling link must specify the other ICP CHPID with which it is to communicate. Use the CPATH keyword in the CHPID statement to connect internal coupling CHPIDs (see "IOCP statements for ICP" on page 52).

It is suggested that you define a minimum of internal coupling channels. For most customers, it is suggested defining at least two internal coupling links for each coupling facility logical partition (LP) in your configuration. For instance, if your general-purpose configuration has several z/OS LPs and one CF LP, you would define two links (four ICP CHPIDs) shared by all the LPs in your configuration. If your configuration has several z/OS LPs and two CF LPs, you define four links (two links per CF LP).

Maximum recommended number of ICP CHPIDs

Real CPU resources are used to implement the link function of connected ICP CHPIDs. Production environments should limit the maximum number of internal coupling links that are defined for a CPC to optimize the internal coupling link function utilization of CPU resources. This maximum number of internal coupling links is based on the number of available physical cores on the CPC used by the z/OS and CF LPs. This maximum number of internal coupling links can be calculated by taking the number of CPs in the CPC that are used for general-purpose CPs and for ICF processors, and subtracting one from that total. For example: a CPC that consists of four general-purpose CPs and two ICF processors would have a maximum five (4 + 2 -1 = 5) internal coupling links recommended. This represents a maximum total of 10 ICP CHPIDs being defined.

Coupling channel path selection

Each coupling channel type is assigned to a performance selection tier. Those channels with native like performance characteristics occupy the same tier. There is no order within a tier. The tiers themselves are ordered with respect to performance. The assignment is made at the time channel initialization is completed. Channels in a mixed control unit will be assigned to different performance groups or tiers

Message path selection has an approach to use every available buffer of a channel in a higher performance tier before selecting any buffers from channels in a lower performance tier. When multiple channels within tiers are selectable, the firmware implements a round robin selection within those channels so that a single channel within a performance group is not over utilized and all paths within a tier see approximately equal usage.

The current performance selection tiers in order from highest to lowest:

1. Internal Coupling channel
2. Integrated Coupling Adapter (ICA SR)
3. Coupling Express LR (CE LR)

I/O configuration considerations

ICP IOCP supports coupling facility channel path definition on the z15.

With z/OS, HCD provides controls for defining coupling facility channels. HCD also automatically generates the control unit and device definitions associated with CE LR, ICA SR coupling links, or ICP channel paths.

Note: It is recommended that you use the Hardware Configuration Definition (HCD), when possible, to define the coupling facility channel configuration to the channel subsystem.

Linux operating system planning considerations

Linux is an open operating system with a wealth of applications which, in most cases, can run on a z15 with a simple recompile. The z15 includes features that provide an extremely cost-effective environment in which to run Linux.

Integrated Facility for Linux (IFL)

You can purchase and install one or more IFL features exclusively for Linux workloads (a single Linux image or z/VM Version 6.4 and later with only Linux guests) with no effect on the model designation. Consequently, no additional IBM operating system or middleware charges are incurred with the addition of this capacity unless that software is actually running in that additional capacity.

These features are ordered separately, and are distinguished at the hardware level from any general-purpose CPs, ICFs, or zIIPs, . CPs, ICFs, IFLs, and zIIPs, are perceived by the system as multiple resource pools.

With this feature, Linux, or z/VM Version 6.4 or later with only Linux guests, runs on IFLs. These IFLs cannot be used to run other IBM operating systems such as z/OS, z/VSE, or z/TPF. Only logical partitions specified as either Linux-Only Mode or z/VM Mode in their Activation profiles can be allocated IFLs. IFLs can be allocated as either dedicated or shared. z/VM 6.4 and later can run in a logical partition that includes IFLs and can dispatch Linux guest virtual IFLs on the IFL logical processors. z/VM 6.4 and later can also simulate IFLs for Linux guests, dispatching virtual IFLs on general-purpose logical processors (CPs).

z/VM Mode allows z/VM users to configure all CPU types on a logical partition. z/VM 6.4 and later versions support this mode, which provides increased flexibility and simplifies systems management, by allowing z/VM to manage guests to perform the following tasks all in the same z/VM LP:

- Operate Linux on Z on IFLs
- Operate z/VSE, z/TPF, and z/OS on CPs
- Offload z/OS system software overhead, such as DB2 or Java workloads, on zIIPs

For more information, see .

z/VM utilizing IFL features

z/VM utilizing IFL features provides an easy-to-use high-performance hypervisor that operates within a logical partition. It can create a significant number of Linux images. z/VM creates and manages Linux images quickly and easily, providing the ability to share resources, and supports an arbitrary number of internal networks that can be used for high-speed communication among Linux images.

IBM Secure Service Container planning considerations

The IBM Secure Service Container is a container technology through which you can quickly and securely deploy firmware and software appliances on Z mainframes. Unlike most other types of partitions, a Secure Service Container partition contains its own embedded operating system, security mechanisms,

and other features that are specifically designed for simplifying the installation of appliances, and for securely hosting them.

A Secure Service Container partition is a specialized container for installed and running specific firmware or software appliances. An appliance is an integration of operating system, middleware, and software components that work autonomously and provide core services and infrastructures that focus on consumability and security. Firmware appliances are delivered with the mainframe system; software appliances are delivers through software distribution channels.

The Z mainframes support several types of partitions. When system administrators define a partition, they specify characteristics that include processor resources, memory resources, and security controls. System administrators use the Hardware Management Console to define partition characteristics.

Secure Service Container supports the following firmware and software appliances:

- IBM Blockchain High Security Business Network. For more information, see the announcement at https://www.ibm.com/common/ssi/cgi-bin/ssialias?infotype=an&subtype=ca&appname=gpateam&supplier=897&letternum=ENUS216-491
- The IBM z Advanced Workload Analysis Reporter (IBM zAware) Software Appliance. For more information, go to the web page for IBM Operations Analytics Version 3.1 at http://www.ibm.com/software/products/en/ibm-operations-analytics-for-z-systems
- The IBM z Advanced Workload Analysis Reporter (IBM zAware) firmware appliance, which requires hardware feature code 0011.

When the Secure Service Container partition is activated, the sequence of events varies, depending which boot selection you specified on the **SSC** page of the image profile.

Secure Service Container appliance installer
Select this option if you want the partition start process to initialize the Secure Service Container Installer so you can install an appliance. This boot selection is the only option when you start a newly configured Secure Service Container partition for the first time. With this option, the Secure Service Container Installer is started automatically. When the start process competes, you can access the Secure Service Container Installer through your choice of browser.

Secure Service Container appliance
Select this option if you want the partition start process to effectively restart an installed appliance. If you previously used the Secure Service Container Installer to successfully install a firmware or software appliance, this boot selection becomes the default selection in the image profile for the Secure Service Container partition. In this case, the Secure Service Container Installer is rebooted, and the installed appliance is restarted in the Secure Service Container partition on this and all subsequent reboots, until you change the boot selection in the image profile.

For more information on IBM Secure Service Container, see the *Secure Service Container User's Guide*, SC28-6978.

Additional LP definition controls are available for use by Secure Service Container logical partitions. For details, refer to "Establishing Secure Service Container parameter descriptions" on page 145.

IBM System Recovery Boost planning considerations

System Recovery Boost can provide additional processing capacity during the limited-duration startup and shutdown "boost periods" in a variety of different ways, depending on the IBM z15 server configuration and on specific operating system exploitation capabilities. On subcapacity machine models (4xx, 5xx, 6xx, and Axx-Yxx), System Recovery Boost can temporarily boost the processing capacity of general purpose processors to run as if they were full-capacity processors, only in those system images that are actively experiencing a boost. This temporary increase in general purpose processing capacity for the boosting images is not visible for pricing purposes, so it does not lead to IBM software licensing cost increases. On servers with zIIP processors, System Recovery Boost is designed to help unlock the parallelism, capacity and acceleration potential of those zIIP processors by temporarily allowing general-purpose workload to run on them during the boost period, only in the system images that are actively experiencing a boost. This use of zIIP processor capacity applies only to operating systems that support

zIIPs (for example, z/OS), and to system images that actually have zIIP processing capacity defined to them.

System Recovery Boost can optionally, with a priced Boost temporary capacity record and "Turbo" offering requiring the use of feature codes #6802 and #9930, enable the temporary activation of additional physical zIIP processors on an IBM z15 server. Activation of these processors makes use of unused processing cores in the machine and provides more zIIP processing capacity that can then be used by boosting images to accelerate execution of their workload (both general-purpose workload and workload that is already zIIP eligible). When the Boost temporary capacity record is activated for use during, for example, a planned maintenance window or for a planned site-switch activity, up to 20 additional zIIP engines may become available for a period of up to 6hours for use on an IBM z15 server. This additional zIIP capacity is then shared across images in accordance with existing PR/SM management controls, making additional zIIP capacity available to individual system images. The system images that want to take advantage of this additional zIIP capacity should predefine reserved logical zIIP capacity in their PR/SM image profiles. The operating system can then bring those additional logical zIIP processors (with physical backing from the additional physical zIIPs that were activated) online for use during the boost period, thereby providing the image with increased zIIP capacity and parallelism to accelerate the workload.

System Recovery Boost requires operating system exploitation. z/OS will fully support the boosting of general purpose processors on subcapacity machine models to run as if they were full-capacity processors and also the use of zIIP processors to run general purpose work during boost periods. z/OS will support such boosts during a planned system shutdown and during system startup. The zIIP Boost can improve z/OS recovery time by making CP-only work eligible to run on zIIPs and CPs during the boost period. In order to exploit this capability, the logical partition must use shared processors, run with HiperDispatch=Yes, and have one or more logical zIIPs defined. Typically, you would define as many initial zIIPs as you need for normal operation. However, you may want to consider increasing the number of reserved logical zIIPs you have defined for the logical partition. These additional reserved logical zIIPs, up to what can be backed by physical zIIPs at the time, are automatically configured online by z/OS during the boost period for the logical partition. z/OS automatically configures these same logical zIIP processors offline at the end of the boost period. Stand-Alone Dump (SADMP) will support boosting of general purpose processors on subcapacity machine models only to provide additional processor capacity for use in capturing diagnostic information for system failures. z/VM and z/TPF will provide support for boosting of general purpose processors on subcapacity machine models only for both planned system shutdown and system startup. In the case of z/VM, the increased capacity of general purpose processors during a z/VM system startup or shutdown can be "inherited" by some second-level guests to provide additional processing capacity to accelerate the guest startup and/or shutdown processing. The z/VM running on IFL processors will not experience any capacity boost for IFLs.

IBM z Integrated Information Processor (zIIP)

The z Integrated Information Processor (zIIP), is the latest customer-inspired specialty engine for the Z mainframe. It provides a cost-effective workload reduction environment that is used by DB2 and other software products, some from ISVs. The zIIP is designed to help improve resource optimization and lower the cost of eligible workloads, enhancing the role of the mainframe as the data hub of the enterprise.

The execution environment of the zIIP accepts eligible work from z/OS 1.8 or higher, which manages and directs the work between the general-purpose processor and the zIIP. DB2 for z/OS V8 uses the zIIP capability for eligible workloads. The zIIPs are designed to free up general-purpose capacity which might be used by other workloads. Java workload can also now utilize zIIPs.

In addition to improving the utilization of existing resources, the zIIP might help you to use the z/OS and DB2 for z/OS qualities of service for data access and information management across your enterprise. It does this by making direct access to DB2 more cost effective and potentially reducing the need for many local copies of the data and the complexity that brings.

On the Z mainframe, you can purchase and install one or more zIIP features with no effect on the model designation. Consequently, no additional operating system or middleware charges are incurred with the addition of this capacity unless that software is actually running in that additional capacity. The zIIP

feature is ordered separately and is distinguished at the hardware level from any general-purpose Central Processors (CPs), Internal Coupling Facility (ICF) processors, or Integrated Features for Linux (IFLs).

z/VM 6.4 and later can run in a logical partition that includes zIIPs and can dispatch z/OS guest virtual zIIPs on the zIIP logical processors. z/VM 6.4 and later can also simulate zIIPs for z/OS guests, dispatching virtual zIIPs on general-purpose logical processors (CPs).

Concurrent patch

Concurrent patch is available on the Z mainframe. It is possible to apply BPC, UPC, Support Element (SE), Hardware Management Console, channel Licensed Internal Code (LIC), PR/SM, coupling facility control code, I390, and PU patches nondisruptively and concurrent with system operation. There can still be situations where a small percentage of patches is disruptive; however, all major LIC components now support concurrent patch.

Additionally, there is also support for multiple EC streams (one per major LIC component) further minimizing the number of disruptive patch sessions. On previous models, a single EC stream contained all major LIC components and provided a mandatory sequence for patch application. This could lead to situations where a disruptive patch belonging to one LIC component, for example, PU, could prevent you from applying all nondisruptive SE patches if one or more of the SE patches came after this channel patch in the patch application sequence.

Patches for each major LIC component have their own sequence for patch application. This means that disruptive patches belonging to one LIC component no longer stand as an obstacle to the concurrent application of nondisruptive patches belonging to another LIC component as long as the patches in question are not otherwise defined as corequisite for each other.

Dynamic capacity upgrade on demand

The z15 includes a function to dynamically increase the number of CPs, ICFs, IFLs, or zIIPs without an intervening IPL. A logical partition (LP) might be defined with both an initial and reserved amount of logical cores. This enables a definition where the number of logical cores for a logical partition is greater than the number of physical cores installed on the model. These reserved CPs are automatically in a deconfigured state at partition activation and can be brought online at any future time by the SCP operator command if the requested resource is available. To prepare for a nondisruptive upgrade, a Logical Partition simply needs to be defined and activated in advance with an activation profile indicating reserved CPs. This helps ensure that any planned logical partition can be as large as the possible physical machine configuration, nondisruptively.

With support available on the z15, the logical core definition for a logical partition can be dynamically changed without requiring a reactivation of the logical partition. This allows you to add to the definition of offline CPs (or any other supported processor types) dynamically should the need arise. If the system control program running in that logical partition supports this dynamic add capability, the additional offline CPs can be configured online in the same way as preplanned reserved CPs are brought online.

For more information about using dynamic CPU addition, see *z/OS MVS Planning: Operations* and *z/VSE Planning*.

The following example assumes a nondisruptive concurrent CP upgrade from an 8-Way to a 9-Way Server.
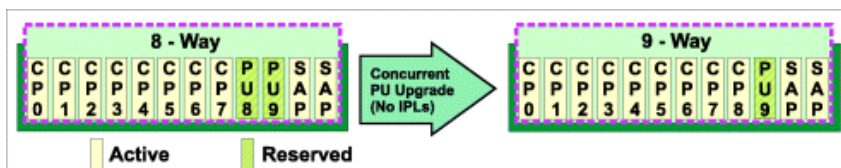
Figure 20. Nondisruptive concurrent CP upgrade

PR/SM shared partitions

PR/SM configurations supporting multiple partitions, all sharing CPs, support concurrent CP upgrades. PR/SM code, once signaled that one or more central processors have been made available to the configuration, will vary them online automatically into the *shared pool* of physical cores and begin full utilization of the added capacity. In the following example, three partitions sharing eight physical cores are able to share the increased capacity resulting from a nondisruptive upgrade to a 9- way server without any other configuration changes. In the second upgrade scenario, Partition 1 is additionally changed from a 4-way to a 5-way partition nondisruptively. The preparation for this is straightforward and easy. Simply define and activate logical Partition 1 with four initial and one reserved logical cores (see Figure 37 on page 137 for a similar example). At the time of the concurrent CP upgrade, the **z/VSE SCP** operator command or the **z/OS CONFIG CPU** command can be used in Partition 1 to configure a fifth CP online to the shared partition without interruption to any logical partition.
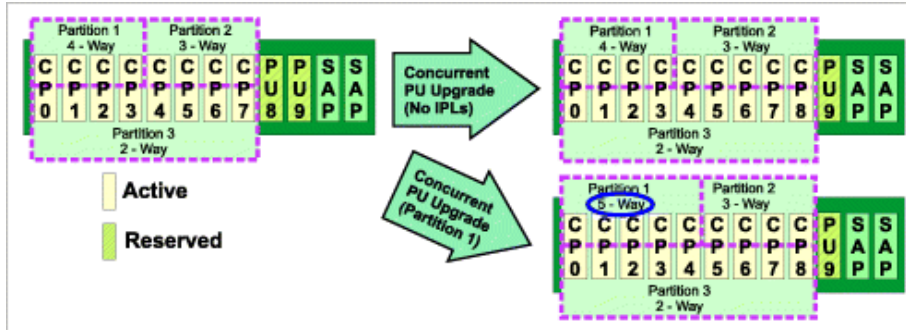
Figure 21. PR/SM shared partitions

Mixed shared and dedicated PR/SM partitions

As with configurations in which all PR/SM partitions share available CPs, those shared partitions within a mixed configuration also support concurrent CP upgrade. CPs are added, without disruption to any of the partitions, to the pool of physical cores shared among the shared CP LPs. In addition, partitions configured with dedicated CPs in a mixed environment can add new CP capacity while both the shared CP partitions and the dedicated CP partition run uninterrupted. To prepare for this ability, in the following example simply define Partition 3 as a three-way dedicated partition with two initial logical cores and one reserved logical core (see Figure 37 on page 137 for a similar example). The reserved logical core is offline automatically at the time of partition activation. At the time of the concurrent CP upgrade, the SCP operator command can be used in the dedicated partition to configure a third CP online to the dedicated partition without interruption to any logical partition.
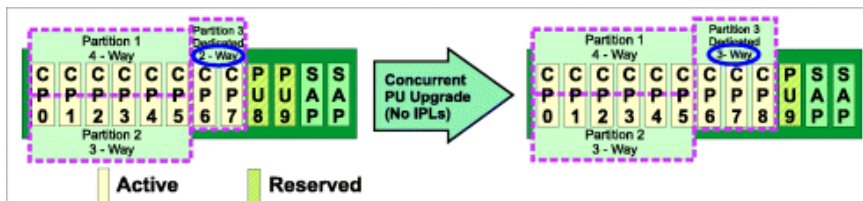
Figure 22. Mixed shared and dedicated PR/SM partitions

Multiple dedicated PR/SM partitions

Configurations in which all PR/SM partitions use dedicated CPs, where there is more than one dedicated partition, also support concurrent CP upgrade. CPs are added to the configuration, without disruption to any of the partitions, and can be brought online to a dedicated partition without an interruption. In the following example, all ICFs in the configuration are dedicated. The partition (ICF 1) is defined for the ability to be upgraded dynamically. To prepare for this ability, simply define ICF 1 as a two-way dedicated partition with one initial and one reserved logical core (see Figure 37 on page 137 for a similar example). At the time of the concurrent CP upgrade, the CF operator command can be used in the ICF 1 dedicated partition to configure a second ICF processors online to the dedicated partition without interruption to

any logical partition. Note the same type of procedure can be accomplished with logical partitions using dedicated general-purpose CPs.
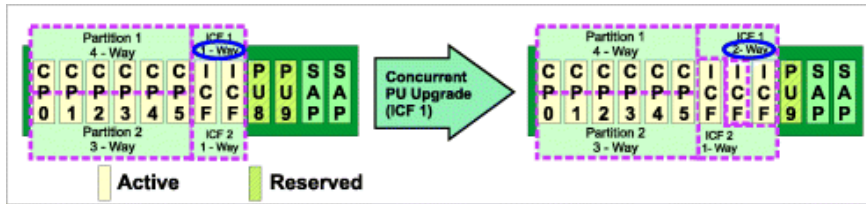
Figure 23. Multiple dedicated PR/SM partitions

Shared Internal Coupling Facility

Concurrent upgrade can be used to add a PU to a shared pool of PUs supporting existing Internal Coupling Facilities. In the following example, Partition 1 is defined for the ability to be upgraded dynamically. To prepare for this ability, simply define Partition 1 as a seven-way shared partition with six initial and one reserved logical cores (see Figure 37 on page 137 for a similar example). At the time of the concurrent CP upgrade, the CF operator command can be used in Partition one shared partition to configure a seventh ICF processors online to the shared partition without interruption to any logical partition. Partition 2 in this case could have also been defined as a seven-way shared ICF partition with one ICF configured offline. This would allow Partition 2 to grow concurrently without an outage as well.
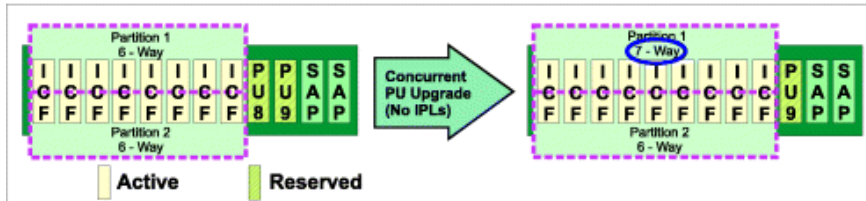
Figure 24. Shared internal coupling facility

Dynamic capacity upgrade on demand limitations

1. Inactive (spare) PUs can be added concurrently, dynamically providing nondisruptive upgrade of processing capability. They can be characterized, in any combination, as CPs, ICFs, IFLs, or zIIPs .

2. For shared CP PR/SM configurations, added CPs are brought into the pool of shared CPs, effectively increasing the number of physical cores to be shared among partitions. To increase the number of logical cores online to an LP, simply define the LP with both initial and reserved logical cores. Then all you need to do is configure on the extra logical core(s) after the concurrent upgrade.

3. All models can define logical partitions with as many reserved CPs as necessary. With thoughtful planning, there is never a need for a disruptive increase in the number of logical cores.

4. The maximum initially online logical core width that logical partitions can be defined and activated with at any point in time is as follows (the maximum defined logical core width is as great as the total number of CPs achievable with concurrent CPU upgrade):

   - Maximum initial logical cores defined for a dedicated partition **equals** physical cores online for the current model **minus** physical cores currently dedicated and online to other dedicated partitions **minus** the maximum online number of shared CPs among all the activated logical partitions that are using shared CPs.

   - Maximum initially online logical cores defined for a shared partition **equals** physical cores online for the current model **minus** physical cores currently dedicated and online to other partitions using dedicated CPs.

   - Maximum **total** logical cores (including initial and reserved) for any partition **equals** the number of physical cores achievable through concurrent CP upgrade.

- When a logical partition is defined to use ICFs, IFLs, or zIIPs these rules are applied against the installed processors of that type. The total number of all logical processors defined to the logical partition cannot exceed the maximum supported for a logical partition by the CPC, independent of processor type.

Concurrent Memory Upgrade

A Z mainframe includes a function to dynamically increase the amount of configured storage. Concurrent Memory Upgrade allows for a memory upgrade without changing hardware or experiencing an outage, provided there is enough spare memory existing on the memory cards. An IML is not required to use the previously unavailable storage. It is immediately available for allocation to logical partitions as central storage. The new storage can be allocated to either newly activated logical partitions or to already active logical partitions by using dynamic storage reconfiguration (see "Dynamic storage reconfiguration" on page 97). In planning for a concurrent memory upgrade, logical partition activation profiles should be defined with storage (central) specifications that include a reserved as well as an initial storage amount. Following the completion of the concurrent memory upgrade operation, issue the z/OS command CF STOR(E=1),ONLINE to bring the new memory online to already active LPs for which a reserved central storage amount was specified. For z/VM, following the completion of the concurrent memory upgrade operation, issue the z/VM SET STORAGE command, to bring the new memory online to already active LPs for which a reserved central storage amount was specified.

Capacity Backup Upgrade (CBU) capability

This orderable feature can be integrated with Geographically Dispersed Parallel Sysplex. This should reduce disaster recovery times by automating the Capacity Backup Upgrade process at a remote site.

Operationally, the planning considerations and limitations for Capacity Backup Upgrade, for mode purposes, are similar to those for Dynamic Capacity Upgrade on Demand. Planning is simple because reserved logical cores can be specified for logical partitions on the backup system. The logical cores can then be brought online, quickly and nondisruptively.

Concurrent Undo CBU is provided to dynamically remove from the configuration processors that had been added with Capacity Backup Upgrade. The server for disaster recovery with Capacity Backup Upgrade activated can now be restored to its base configuration without requiring a system outage.

When a disaster occurs, Capacity Backup Upgrade is intended to provide the extra capacity without disruption. When the disaster is over and normalcy is restored, Concurrent Undo CBU is intended to allow the system to be returned to its previous configuration without disruption.

Annual testing of a Capacity Backup Upgrade system is highly recommended. Invoking the Capacity Backup Upgrade configuration is nondisruptive and now returns to the original configuration, after the test, without disruption as well.

When a typical Capacity Backup Upgrade occurs, typically logical processors are configured online so that the system closely resembles the failing system. To prepare for Concurrent Undo CBU, the most expedient method is to simply configure offline all those logical processors that were configured online in support of the failing system.

In order for Concurrent Undo CBU to proceed, the restored original configuration must have a physical processor for each online dedicated logical processor. Theoretically, the only additional requirement is that at least one non-dedicated physical processor remains that matches each type (general purpose, ICF, IFL, or zIIP) of online shared logical processors. However, it is highly recommended that shared logical processors be configured offline so that the highest number of online shared logical cores for any active LP does not exceed the number of non-dedicated physical cores remaining. For further guidelines, see the rules governing the number of CPs that can be specified for an activating logical partition on page "Number of central processors" on page 104.

Enhanced Processor Drawer Availability

The Z mainframe is designed to allow a single processor drawer, in a multi-processor drawer server, to be concurrently removed from the server and reinstalled during an upgrade or repair action, while continuing to provide connectivity to the server I/O resources using a second path from a different processor drawer. To help minimize the impact on current workloads and applications, you should ensure that you have sufficient inactive physical resources on the remaining processor drawers to complete a processor drawer removal.

Enhanced processor drawer availability might also provide benefits should you choose not to configure for maximum availability. In these cases, you should have sufficient inactive resources on the remaining processor drawers to contain critical workloads while completing a processor drawer replacement. Contact your system support to help you determine the appropriate configuration. With proper planning, you might be able to avoid planned outages when using enhanced processor drawer availability.

Enhanced driver maintenance is another step in reducing the duration of a planned outage. One of the greatest contributors to downtime during planned outages is Licensed Internal Code (LIC) updates performed in support of new features and functions. When properly configured, the Z mainframe is designed to support activating select new LIC level concurrently. Concurrent activation of the select new LIC level is only supported at specific sync points (points in the maintenance process when LIC might be applied concurrently - MCL service level). Sync points might exist throughout the life of the current LIC level. Once a sync point has passed, you will be required to wait until the next sync point supporting concurrent activation of a new LIC level. Certain LIC updates are not supported by this function.

Preparing for Enhanced Processor Drawer Availability

This option determines the readiness of the system for the targeted processor drawer. The configured processors and the in-use memory are evaluated for evacuation from the targeted processor drawer to the unused resources available on the remaining processor drawers within the system configuration. In addition, the I/O connections associated with the targeted processor drawers are analyzed for any Single Path I/O connectivity.

There are three states which can result from the prepare option:

- The system is ready to perform the Enhanced Processor Drawer Availability for the targeted processor drawer with the original configuration.

- The system is not ready to perform the Enhanced Processor Drawer Availability due to conditions noted from the prepare step. See "Getting the system ready to perform Enhanced Processor Drawer Availability" on page 87 for more details.

- The system is ready to perform the Enhanced Processor Drawer Availability for the targeted processor drawer. However, processors were reassigned from the original configuration in order to continue. "Reassigning non-dedicated processors" on page 88 for details.

Getting the system ready to perform Enhanced Processor Drawer Availability

Review the conditions that are preventing the Enhanced Processor Drawer Availability option from being performed. There are tabs on the resulting panel for Processors, Memory, and for various Single Path I/O conditions. The tabs that have conditions preventing the perform option from being executed are displayed. Each tab indicates what the specific conditions are and possible options to correct the conditions.

The Processor tab displays, indicating the corrective actions suggested for the processor configuration. Following is a list of tabs that might appear for your particular configuration:

Processors
Use this tab to view the corrective actions required for the processor configuration conditions that are preventing the Perform Enhanced Processor Drawer Availability option from being performed for the targeted processor drawer. You might need to deactivate partitions or deconfigure processors to meet requirements as indicated by the window data.

Memory

Use this tab to view the corrective actions required for the memory configuration conditions that are preventing the Perform Enhanced Processor Drawer Availability option from being performed for the targeted processor drawer. You may need to deactivate partitions to meet requirements as indicated by the window data. The In-Use memory must be less than or equal to the available memory on the remaining processor drawers within the system.

Single Path I/O

Use this tab to view the corrective actions required for the single I/O configuration conditions that are preventing the Perform Enhanced Processor Drawer Availability option from being performed for the targeted processor drawer. You need to deconfigure all the PCHIDs that are indicated by the window data.

Single Domain

Use this tab to view the corrective actions required for the single I/O domain configuration conditions that are preventing the Perform Enhanced Processor Drawer Availability option from being performed for the targeted processor drawer. You need to change the alternate path to a different processor drawer or deconfigure the PCHID.

Single Alternate Path

Use this tab to view the corrective actions required for the single I/O configuration conditions that are preventing the Perform Enhanced Processor Drawer Availability option from being performed for the targeted processor drawer. You need to correct the alternate path error condition or deconfigure the PCHIDs.

Reassigning non-dedicated processors

Following is an example showing the Reassign Non-Dedicated Processors window. Use this window to change or accept the system processor assignments that are generated during the processing of the **Prepare for Enhanced Processor Drawer Availability** option. The processor values that are entered from this window will be the processor configuration used during the Perform Enhanced Processor Drawer Availability processing.

Attention: The values should never be altered without approval from the system programmer.

Reassign Non-Dedicated Processors - P00699CF

Accept or reassign non-dedicated processors without exceeding the LICCC count.

Warning: These values should only be reassigned under the direction of the System Programmer. Otherwise accept the provided default values.

Processor Type	Dedicated Count	Non-Dedicated Count	Processor Totals	LICCC Count
CPU	0	7	7	8
ICF	0	1	1	2
IFL	0	0	0	1
zIIP	0	2	2	2
SAP	5	0	5	6
Available to use		0	0	
Remaining Book Totals	5	12	17	

OK   Cancel   Help

Figure 25. Reassign non-dedicated processors window

Customer Initiated Upgrade (CIU)

This feature is designed to allow timely response to sudden increased capacity requirements by downloading and automatically applying a processor and/or memory upgrade using Resource Link and the support server.

Operationally, the planning considerations and limitations for CIU are like those for Dynamic Capacity Upgrade on Demand and Concurrent Memory Upgrade. Planning is simple because reserved logical cores can be specified for logical partition (LP)s in the original configuration. The logical cores can then be brought online, quickly and nondisruptively at the completion of the Concurrent Upgrade. Similarly, a reserved central storage amount can be specified for participating LPs and brought online non-disruptively following the completion of the concurrent memory upgrade.

Concurrent Processor Unit conversion

The Z mainframe supports concurrent conversion of different Processor Unit (PU) types. This capability is extended to Central Processor (CPs), Integrated Facility for Linux (IFLs), z Integrated Information Processor (zIIPs), and Internal Coupling Facility (ICFs) providing flexibility in configuring the system to meet the changing business environments.

Planning for nondisruptive install of crypto features

Crypto Express feature can be added to logical partitions non-disruptively using the **Change LPAR Cryptographic Controls** task. For more information, see "Changing LPAR cryptographic controls" on page 172. Logical partitions can either be configured in advance or dynamically with the appropriate domain indexes and Cryptographic numbers (see the Cryptographic Candidate List information under "Crypto parameter descriptions" on page 149).

If the customer plans to use ICSF or the optional cryptographic hardware, the CP Crypto Assist functions (CPACF DES/TDES) must be enabled. Many products take advantage of the cryptographic hardware using ICSF, so enabling CPACF is recommended. View the System Details panel to determine if the CPACF feature is installed. For more detailed information, see the *z/OS Cryptographic Services ICSF TKE Workstation User's Guide*.

It is important to remember that when non-disruptively installing crypto adapters, the default configuration of the card is coprocessor (CCA Coprocessor). To change the type configuration of a crypt adapter, use the Cryptographic Configuration window. Once the crypt adapter has been installed, and the previous preparations made, the crypt adapter is available to any partition that specifies their assigned Cryptographic numbers in the Candidate List of the Image Activation Profile. To bring the crypto adapter online, use the Config On /Off Window.

The Cryptographic Online List needs to be set up in the Image Activation Profile to reflect the crypto adapter that you want to bring online automatically during the partition activation. If the Cryptographic Candidate List does not reflect the crypto adapter that you plan to use, then these crypto adapters are not available to the logical partition after this partition is activated. If the Cryptographic Online List is set up properly, the crypto adapter is brought online when the partition is activated and available to the operating system. Otherwise, a Configure On is necessary to bring the crypto adapter online in the logical partition.

Chapter 3. The characteristics of logical partitions

This chapter provides a planning overview for defining logical partitions (LPs). Support for features, functions, and windows can differ depending on machine type, engineering change (EC) level, or machine change level (MCL). During IOCP execution, the names and numbers of the LPs are specified and channel paths are assigned to the LPs.

Performance considerations

The performance of an LP is controlled by:

- The number of logical cores online
- The CPs are dedicated to or are shared by the LP
- The processing weight assigned to each LP
- The processor running time interval

The total number of logical cores across all active LPs is one of the factors used to determine the dynamic dispatch interval. See "Processor running time" on page 119.

Use the RMF Partition Data Reports to determine the effective dispatch time for LPs. For more information about this RMF reporting enhancement see "RMF LPAR management time reporting" on page 179.

The greater the number of active logical cores relative to the number of physical cores configured, the smaller the dispatch interval.

Dedicated and shared central processors (CPs)

LPs can have CPs dedicated to them, or they can share CPs with other active LPs. Because the use of dedicated or shared CPs in an LP affects performance in several ways, the characteristics, limitations, and advantages of each should be carefully studied.

All processor types in a partition must be either shared or dedicated. You can not mix shared and dedicated processors in the same partition.

Dedicated and shared channel paths

A configuration defining shared channel paths offers additional capabilities over an equivalent configuration containing unshared channel paths, while maintaining comparable system performance.

ITR performance

The best ITR performance is achieved with dedicated LPs. To achieve optimal ITR performance in sharing LPs, keep the total number of logical cores online to a minimum. This reduces both software and hardware overhead.

Capped logical partitions

It is recommended that LPs be defined as capped LPs only when needed to support planned requirements. When a capped LP does not obtain needed CP resources, because it has reached its cap, activity for that LP is similar to a system running out of CP resources. Response time can be slower on systems which operate at their cap. For this reason, interactive response times can suffer when there is a mix of interactive and CP-intensive work in the same capped LP.

Recovery considerations

Resources should be defined to LPs so that any hardware failure has a minimal impact on the remaining active LPs.

For example, the failure of a physical core can cause the temporary loss of any logical core that was dispatched on the physical cores. In many instances, recovery of a logical core that was running on a failed physical core will take place automatically when an available spare physical core is dynamically brought into the configuration. Also, PR/SM is often able to transparently re-dispatch a shared logical core on a different physical core even when no spares are available. If a logical core is still lost, the LP owning the logical core can continue operating if it was running on an LP with at least two CPs dispatched on different physical cores, and if the control program that is active in the LP can recover from CP failures.

Determining the characteristics

The information in this section should help you determine the type and amount of CPC resources you need for each LP.

The total amount of resources that you can define for all LPs can exceed the configured resources.

Individual LP definitions are checked against the total resources installed. The actual allocation of these resources takes place only when the LP is activated. This design characteristic allows considerable flexibility when defining and activating LPs.

Control program support

Table 11 on page 92 summarizes the characteristics of the control programs that can be supported in an LP. See "Control program support in a logical partition" on page 6 for more information.

Some control programs require specific LP characteristics. For this reason consider all control programs before planning or defining LP characteristics.

Table 11. Control program support				
Control program	Control program operating mode	Maximum number CPs	Maximum central storage	Maximum number channels
z/OS 2.4	General	100	4 TB	256
z/OS 2.3	General	100	4 TB	256
z/OS 2.3	General	100	4 TB	256
z/OS 2.2	General	100	4 TB	256
z/OS 2.1	General	100	4 TB	256
z/VM 7.2	z/VM, LINUX-Only, or General	80^2	2 TB	256
z/VM 7.1	z/VM, LINUX-Only, or General	80^2	2 TB	256
z/VM 6.4	z/VM, LINUX-Only, or General	64^1	2 TB	256
z/VSE 6.2	General	10	32 GB	256
z/TPF 1.1	General	64	1 TB	256
Linux for z Systems	LINUX-Only or General	64	1 TB	256

Control program	Control program operating mode	Maximum number CPs	Maximum central storage	Maximum number channels
Table 11. Control program support (continued)				

Notes:

1. For z/VM 6.4 up to 64 CPs are supported without SMT and up to 32 are supported with SMT.

2. For z/VM 7.1 and 7.2 up to 80 CPs are supported without SMT and up to 40 are supported with SMT.

IOCDS requirements

You must use IOCP or BUILD IOCDS with HCD to create an LPAR IOCDS. You can specify the LP names and MIF image ID numbers in an LPAR IOCDS.

Logical partition identifier

The logical partition identifier (ID) is used as the third and fourth hexadecimal digits of the operand stored by the Store CPU ID instruction for each CP in the LP. Even though at most 85 (decimal) logical partitions can be defined, valid identifiers for LPs are X'00' through X'7F'. The LP identifier must be unique for each active LP.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to define the LP identifier for an LP. The Partition identifier field is located on the General page for the LP. (See).

Mode of operation

The mode of an LP depending on the model can be **General** , **LINUX-Only**, **z/VM**, **Coupling Facility**, or **SSC** .

The mode of an LP must support the mode of the control program loaded into it. **General** LPs support z/Architecture control programs. Coupling facility LPs support the coupling facility control code, **z/VM** LPs supports z/VM, and **LINUX-Only** LPs support Linux or z/VM. The Secure Service Container LPs supports the IBM zAware virtual appliance and supported software appliances.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to define the mode for an LP. The **Mode** field is located on the General page for the LP. (See).

Storage configurations

The definition of central storage to the LP depends on the size of the I/O configuration, the storage limits of the control program loaded into the LP and on the storage requirements of the applications running in the LP.

Storage resources

Use standard capacity-planning procedures to assess CPC storage requirements for LPs.

With dynamic storage reconfiguration (see), an appropriately defined LP can dynamically add storage to its configuration that is released when another active LP removes it from its configuration or is deactivated. Additionally, an LP can dynamically add storage to its configuration following an increase in configured storage via concurrent memory upgrade (see).

Single storage pool

With this function, all physical storage is dynamically designated by PR/SM as central storage as requirements of active logical partitions dictate.

As a result, the need to predesignate configured storage as central storage prior to IML on the **Storage** page of the Customize Activation Profile window is not necessary and is not provided. The system programmer now has greater flexibility when planning the division of storage in order to satisfy anticipated logical partition definitions. Single storage pool streamlines the planning effort because PR/SM automatically provides the correct storage designation for any configuration as the need arises. This feature is especially useful as it simplifies planning for migration to z/OS a 64-bit capable operation system.

A feature of this function is that the central storage addressability ranges are communicated on the logical partition storage allocation display of the **Storage Information** task.

Central storage

Central storage is defined to LPs before LP activation. When an LP is activated, storage resources are allocated in contiguous blocks. These allocations can be dynamically reconfigured. Sharing of allocated central storage among multiple LPs is not allowed.

Granularity of initial and reserved central storage amounts is dependent on the largest central storage amount (LCSA) for the LP as follows:

Table 12. Central storage granularity	
Largest Central Storage Amount	**Storage Granularity**
LCSA ≤ 512 GB	1 GB
512 GB < LCSA ≤ 1024 GB	2 GB
1024 GB < LCSA ≤ 2048 GB	4 GB
2048 GB < LCSA ≤ 4096 GB	8 GB
4096 GB < LCSA ≤ 8192 GB	16 GB
8192 GB <LCSA ≤ 16384 GB	32 GB

In support of 2 GB large pages, **all** logical partition origins and limits must be on a 2 GB boundary. In other words, the addressing range assigned to the LP will start and end on a 2 GB (2048MB) boundary.

The granularity from Table 12 on page 94 applies across the central input fields (Initial and Reserved fields). Use the larger of the initial and reserved central storage amounts to calculate storage granularity. For example, for an LP with an initial storage amount of 1280 GB and a reserved storage amount of 2848 GB, the central storage granularity of initial and reserved central storage fields is 8 GB, using the larger reserved storage amount to determine storage granularity. The granularity of the central storage origin is not determined from Table 12 on page 94; the initial storage origin input field granularity is fixed at 2 GB (or 2048 MB).

Note: The required granularity of initial and reserved central storage fields of an LP for which an *origin has been specified* is a minimum of 2 GB. In other words, the above table only applies when the LCSA is > 1024GB. For partitions smaller than that, the granularity is 2GB.

Additionally, when migrating to these models, you must convert the existing central storage amounts that are not in multiples of the storage granularity supported.

For example, an LP with an initial central storage size of 10368 MB on a previous model with 16 GB of storage installed would require conversion of the initial central storage size to a multiple of 1024 MB (64 MB granularity was valid on this prior model configuration).

Check your central storage definitions and consult your system programmer to ensure that they meet your processing needs.

Use the **Customize/Delete Activation Profiles** task available to open a reset or image profile to define the central storage for an LP. The Initial, Storage Origin, and Reserved fields are located on the Storage page for the LP. See Figure 42 on page 144 for an example.

Initial central storage

- The initial amount of central storage represents the amount of storage allocated to an LP when it is activated.

- You must specify a nonzero number for the initial amount of central storage to activate the LP.

- If no origin is specified for the initial value, the LP will be treated as an LP that owns its own **addressability range**. This means that no other LP can map over it.

Reserved central storage

- Only **General** , **z/VM**, and **Linux-Only** LPs can have nonzero amounts for this parameter. Coupling facility and Secure Service Container LPs cannot have reserved central storage amounts.

- The reserved amount of central storage defines the additional amount of central storage that can become available to an LP when no other activated LP has this reserved storage online. Reserved storage amounts are always offline after LP activation.

- Only LPs that specify a central storage origin can be allocated storage within the reserved central storage defined for another LP. Reserved central storage can only overlap LPs that specify an origin.

- For LPs that do not specify central storage origins, the LP's reserved storage is available to be brought online to the LP whenever there is any storage that is not being used by any other LPs.

- For LPs that do specify a central storage origin, if some of the reserved central storage is not available, the reserved central storage that is available can still be configured online starting at the reserved storage origin.

- Specifying a zero for the reserved size indicates the LP's central storage cannot get any larger than the initial size for the duration of the activation.

Central storage origin

- There must be enough contiguous central storage addressability to fit the reserved amounts, but the reserved amounts can be in use by other LPs (that specified origins for storage) at the time the LP is activated.

- The origin for central storage defines the starting gigabyte where the central storage *addressability range* begins for the LP. If enough storage is not available to satisfy the initial central storage request starting at this origin, LP activation will fail.

- For central storage, the specification of the origin parameter provides the only way to overlap storage definitions. For example, the reserved storage definition for one LP can overlap the storage definition of another LP when the origin parameter is specified for both LPs.

  The total amount of central storage *addressability* that can be used to map central storage for LPs is **at least** twice the amount of *customer* storage. You can review current LP storage allocations by using the **Storage Information** task available.

- Since the specification of a central storage origin is optional, there are two implementations of dynamic storage reconfiguration for reserved central storage:

  1. If you do not specify a central storage origin, the reserved central storage is available for the LP when there is sufficient physical storage available to meet the reserved central storage request.

     Storage can be reassigned between LPs that have noncontiguous address ranges. Reconfiguration is not limited to growing into an adjacent LP. LPs are eligible to use noncontiguous central storage if they have unique (not overlapping) storage range definitions. Specifying an origin is not required when defining a reserved storage amount for an LP, and by not specifying an origin, the LP is reserved an addressable storage range for its entire storage configuration (initial plus reserved amount). Not specifying an origin ensures no overlap. The reserved central storage is available for LPs that do not specify an origin when there is sufficient physical storage to meet the reserved central storage request.

  2. If you specify a central storage origin, the entire reserved central storage amount is only available for the LP when:

- – No other LP has this central storage address range, to which the reserved central storage is mapped, online.
- – There is sufficient physical storage available to meet the reserved central storage request.
- If no origin is specified for an LP, the system assigns the storage to the LP using a top-down first-fit algorithm. Storage that is part of reserved central storage for another activated LP is not available for any LPs that do not specify a central storage origin. The addressing for each logical partition must start and end on a 2 GB boundary. If the INITIAL or non-zero RESERVED amounts defined for a logical partition with no origin specified are not multiples of 2 GB, the **addressing** allocated for each of these specifications will be rounded up to 2 GB boundaries. This will be reflected to the partition in the amount of offline central storage for the partition.

For example, define a logical partition with 5 GB of INITIAL, 0 MB of RESERVED, and a system determined central storage origin. The 5GB requires a 1 GB granularity for allocation here but to meet the 2 GB addressing needs for the central storage, 6 GB of addressing is required for this logical partition. A single online central storage element is created with provisions for 6 GB of central storage but only the original 5 GB is actually allocated to the logical partition at activation time. The additional 1 GB could potentially be configured online for use at a later time via appropriate OS commands. The Figure 26 on page 96 displays what the z/OS D M=STOR command would display when this logical partition is first activated with this configuration.

```
D M=STOR
IEE174I 10.00.09 DISPLAY M 399
REAL STORAGE STATUS
ONLINE-NOT RECONFIGURABLE
          OM-5120M
ONLINE-RECONFIGURABLE
          NONE
PENDING OFFLINE
          NONE
      OM IN OFFLINE STORAGE ELEMENT(S)
      1024M UNASSIGNED STORAGE
STORAGE INCREMENT SIZE IS 1024M
```

Figure 26. Example of z/OS D M=STOR command output

IBM Virtual Flash Memory

The Virtual Flash Memory is designed to improve availability and handling of paging workload spikes when running z/OS V2.1, V2.2, or V2.3. The Virtual Flash Memory support with z/OS is designed to help improve system availability and responsiveness using Virtual Flash Memory across transitional workload events, such as market openings and diagnostic data collection.

The Virtual Flash Memory can also be used in coupling facility images to provide extended availability for workloads making use of WebSphere MQ Shared Queues structures. Virtual Flash Memory can help availability by reducing latency from paging delays that can occur at the start of the workday or during transitional periods. Virtual Flash Memory is also designed to help eliminate delays that can occur when collecting diagnostic data during failures.

The initial Virtual Flash Memory represents the amount of Virtual Flash Memory allocated to an activated logical partition. The maximum Virtual Flash Memory amount represents the maximum Virtual Flash Memory the logical partition is allowed. This means, if the initial and maximum amounts are specified, the maximum amount minus the initial amount is the Virtual Flash Memory amount that the logical partition's operating system can dynamically configure.

Notes:

1. The sum of the maximums for all activated logical partitions cannot exceed 128 TB.
2. When planning for Virtual Flash Memory, the maximum amount should take into account the Virtual Flash Memory growth required by the partition. If the maximum is reached and the partition wants more Virtual Flash Memory, the change is disruptive for the partition.

The Virtual Flash Memory can be configured in any of the following:

Initial and Maximum specified
> The initial amount allocated at activation and amount (initial and maximum) can be configured dynamically

Initial Specified
> The initial amount allocated at activation and the logical partition cannot dynamically configure Virtual Flash Memory.

Maximum specified
> No initial amount allocated during activation and maximum amount can be configured dynamically when required.

Initial and Maximum not specified
> No Virtual Flash Memory planned for the logical partition.

IBM Adapter for NVMe (LinuxONE only)

This feature provides support for the Non-Volatile Memory express (NVMe) communications protocol that was built specifically for solid-state drives (SSDs). This feature brings integrated storage to LinuxONE by allowing a procured SSD directly connected to the I/O subsystem through an IBM PCIe adapter card. This gives the ability to have embedded storage for various applications. The low latency and high I/O throughput of NVMe SSDs connected directly to the I/O backplane can help with memory-intensive workloads, real-time analytics, fast storage workloads (such as streaming, paging/sorting), and traditional application (such as relational databases).

Note: The use of non-tested SSDs within the LinuxONE system may produce unintended and/or unexpected results. Selection and purchase of the SSD for this NVMe application is important in order to provide cost-effective options. For a list of tested SSDs, see the IBM LinuxONE NVMe white page (73019373-USEN-00).

Dynamic storage reconfiguration

Dynamic storage reconfiguration allows central storage allocated to an LP to be changed while the LP is active. It is supported in Linux and **General** LPs running z/OS. Dynamic storage reconfiguration is not supported in coupling facility or Linux-Only LPs. z/VM 5.4 and later support dynamic storage reconfiguration (for central storage only) in any partition mode.

Dynamic storage reconfiguration provides the capability to reassign storage from one LP to another without the need to POR the CPC or IPL the recipient LP. Every LP has a storage range definition consisting of an origin, initial, and reserved amounts. The reserved value determines how much additional storage can be acquired using dynamic storage reconfiguration.

Storage is released when an LP is either deactivated or its reserved storage element is deconfigured. Additionally, you can release central storage in amounts smaller than the defined storage element size.

With dynamic storage reconfiguration, **General** LPs can have reserved amounts of central storage. This storage can become available to the LP if no other active LP has this reserved storage online. Reserved central storage can be made available to the LP by commands from the operating system console.

If the operating system running in the LP supports physical storage reconfiguration, use operating system commands to make the reserved storage available to the LP without disrupting operating system activities.

For z/OS, use the following command format to reconfigure central storage:

```
CF STOR(E=1),<OFFLINE/ONLINE>
```

Dynamic storage reconfiguration on the mainframes for central storage enables central storage to be reassigned between LPs that have noncontiguous address ranges. In this case, PR/SM can allocate a *hole* or some set of central storage addresses for which there is no backing physical storage assigned. Later, PR/SM is able to configure storage online to the LP by assigning some physical central storage to the hole.

For central storage, you can reconfigure central storage in amounts equal to the storage granularity supported by the CPC. For z/OS, use the following command format:

```
CONFIG STOR(nnM),<OFFLINE/ONLINE>
```

For more information on using the commands see *Device Drivers, Features, and Commands Reference*.

Central storage dynamic storage reconfiguration examples

Figure 27 on page 98 shows an example of central storage with dynamic storage reconfiguration capability. This figure shows LP-A, LP-B, and LP-C.
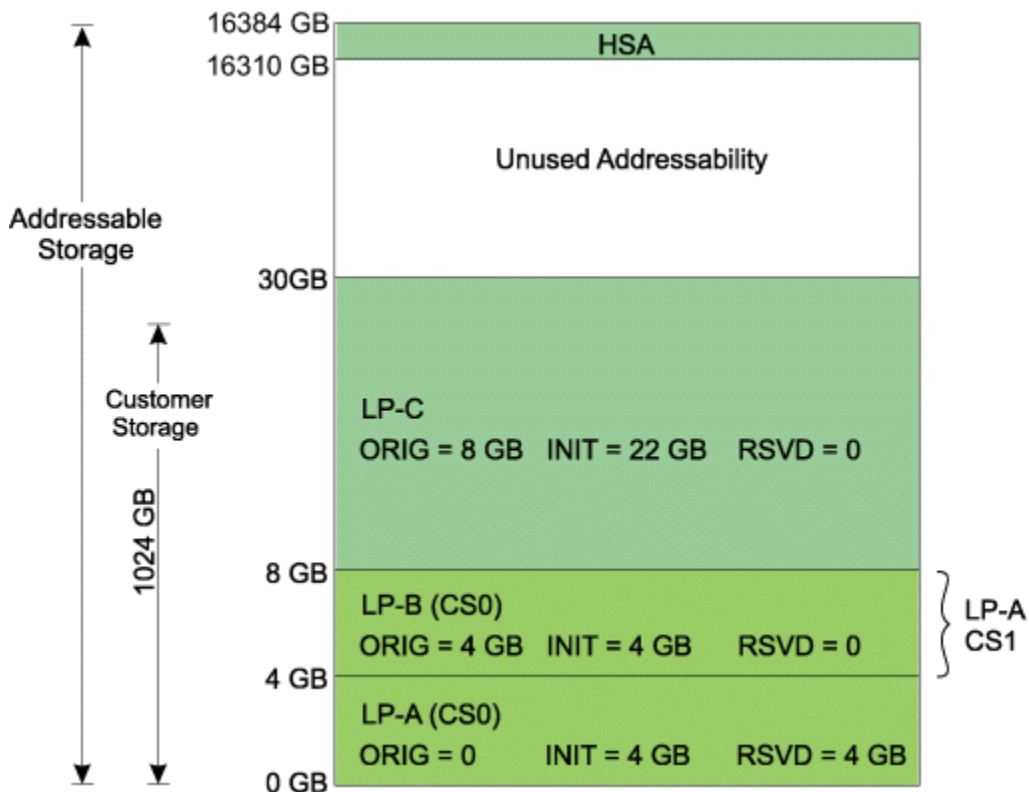
Figure 27. Central storage layout

To reconfigure central storage, deactivate LP-B to free up storage directly above LP-A. shows how you may reconfigure central storage in this case.
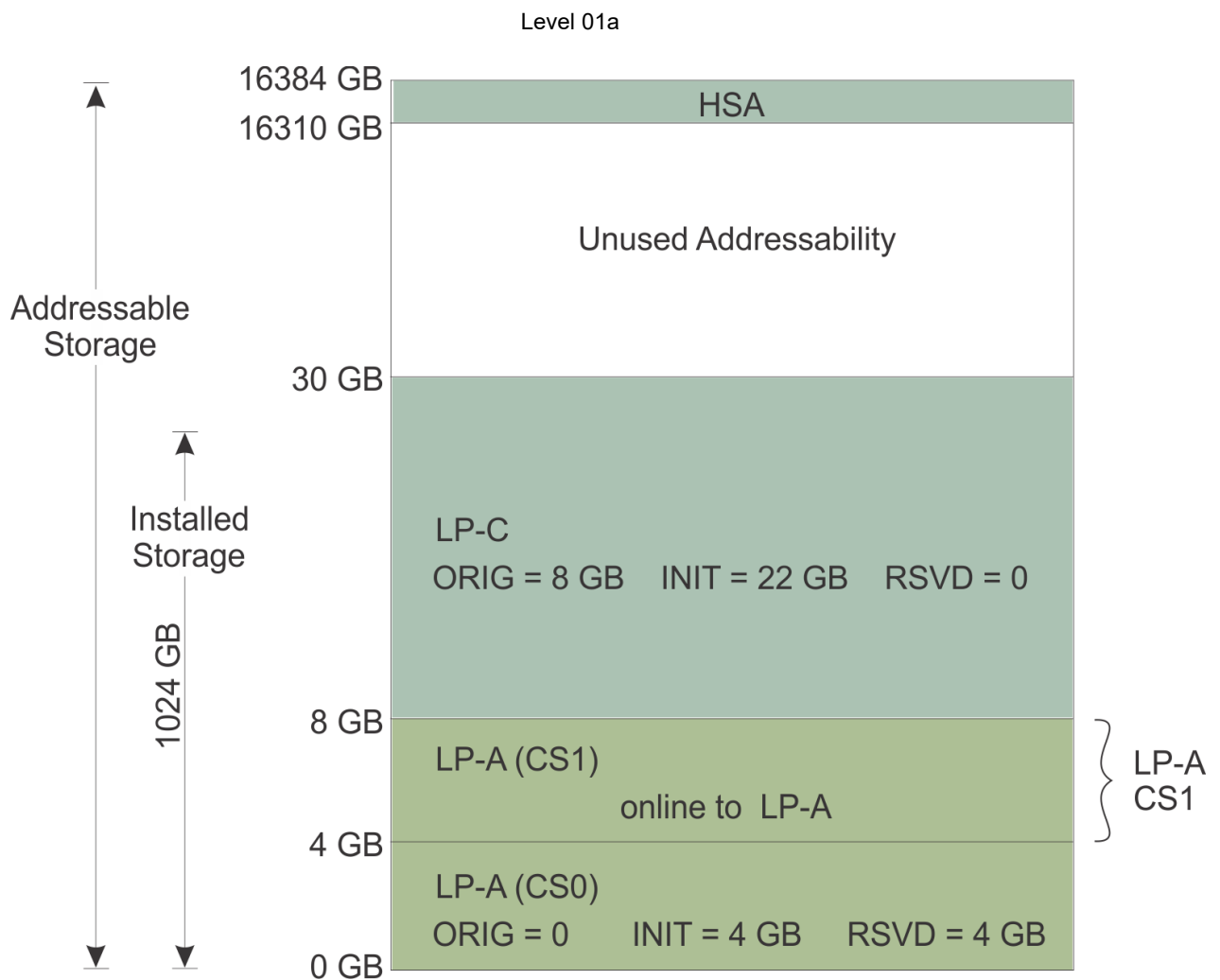
Figure 28. Reconfigured central storage layout

Figure 29 on page 100 is another example of dynamic reconfiguration of central storage. For this example, assume the amount of customer storage is 1 GB. The amount of *addressable* central storage used by hardware system area (HSA) is 74 GB in this example. The storage granularity is 1 GB. This leaves 16310 GB of central storage addressability to be allocated to LPs.

LP-A and LP-B are defined with an initial amount of 14 GB each of central storage, a reserved amount of 2 GB, and system determined central storage origins. LP-A and LP-B are activated and IPLed. At the completion of IPL, LP-A has its reserved central storage configured online by entering **CF STOR(E=1),ONLINE** from the z/OS software console for LP-A. Figure 29 on page 100 shows the resulting storage layout following these actions.
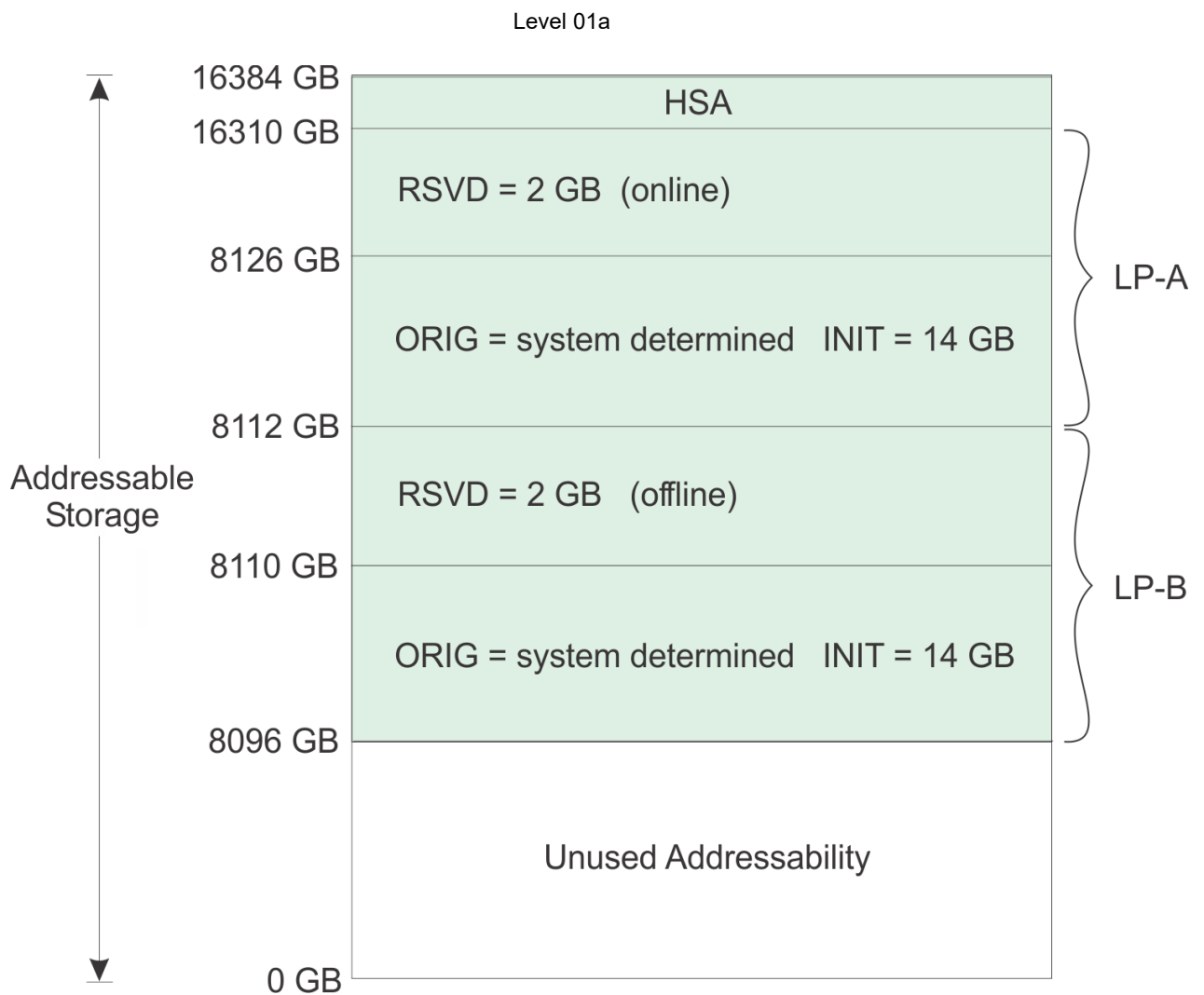
Figure 29. Initial central storage layout

At a later time, the reserved storage from LP-A can be reconfigured to LP-B.

Note: Both LPs should specify an RSU value of at least 2 (2048/1024) for reconfigurations of storage to work.

From the z/OS software console for LP-A, enter **CF STOR(E=1),OFFLINE**. Next, from the z/OS software console for LP-B, enter **CF STOR(E=1),ONLINE**. shows the resulting storage layout following these actions. The reserved storage is fully reconfigured without an outage to either LP. Also, the procedure can be reversed without an outage as well.

```
16384 GB ─┬─                    ┌──────────────────────────────┐
          │                     │            HSA               │
16310 GB ─┤                     ├──────────────────────────────┤  ⎫
          │                     │                              │  ⎪
          │                     │   RSVD = 2 GB   (offline)    │  ⎪
          │                     │                              │  ⎪
 8126 GB ─┤                     ├──────────────────────────────┤  ⎬ LP-A
          │                     │                              │  ⎪
          │                     │  ORIG = system determined    │  ⎪
          │                     │          INIT = 14 GB        │  ⎪
 8112 GB ─┤                     ├──────────────────────────────┤  ⎭
          │                     │                              │  ⎫
Addressable                     │   RSVD = 2 GB   (online)     │  ⎪
 Storage                        │                              │  ⎪
 8110 GB ─┤                     ├──────────────────────────────┤  ⎬ LP-B
          │                     │                              │  ⎪
          │                     │  ORIG = system determined    │  ⎪
          │                     │          INIT = 14 GB        │  ⎪
 8096 GB ─┤                     ├──────────────────────────────┤  ⎭
          │                     │                              │
          │                     │                              │
          │                     │      Unused Addressability   │
          │                     │                              │
    0 GB ─┴─                    └──────────────────────────────┘
```
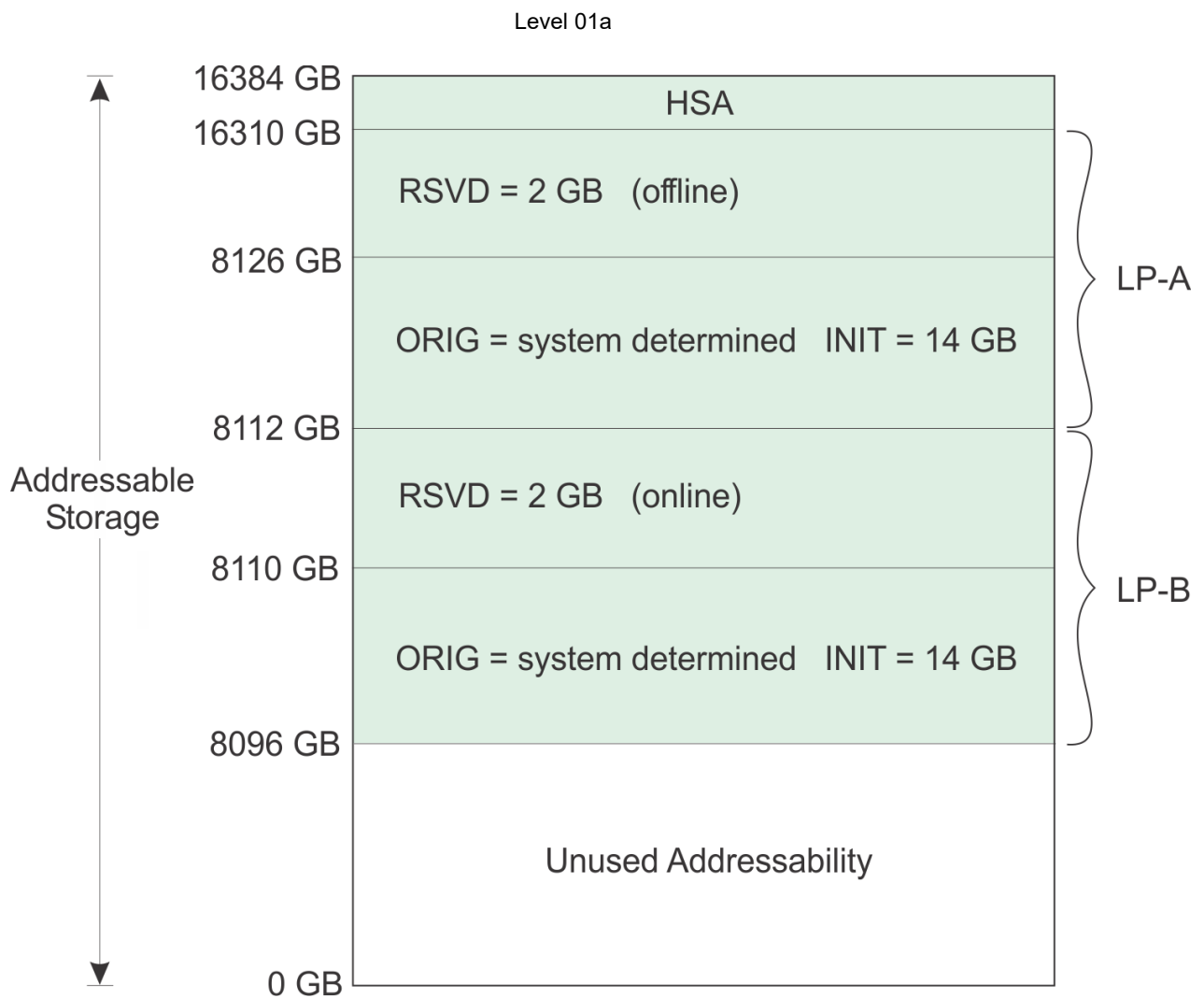
Figure 30. Central storage layout following reconfiguration

Recommendations for storage map planning

Planning storage maps as described below helps avoiding storage fragmentation and remove dependencies on the order of activation of LPs.

For more information and examples of storage map planning, see the publications listed for dynamic storage reconfiguration in "About this publication" on page xv.

Map all LPs that require an origin from the bottom up (that is, start with 0 MB and build upward).

If no origin is specified for an LP, the system assigns the storage to the LP using a top-down first-fit algorithm.

Operation considerations for dynamic storage reconfiguration

- Initial central storage for an LP is allocated on logical storage element 0 and any reserved central is allocated on logical storage element 1.
- Whenever a load clear or system reset clear is performed on an LP, it forces the reserved central storage element offline. This allows z/OS to flag the reserved storage element as reconfigurable to allow it to be later deconfigured from the LP. z/VM provides no facility to deconfigure storage while maintaining normal system operation.

  Note: When z/OS or z/VM is IPLed immediately after the LP is activated, the reserved central storage element is offline.

- Whenever z/OS is re-IPLed in an LP that has a reserved central storage element, a load clear or system reset clear followed by load normal should be performed to force the reserved central storage element offline.

- If additional storage has been made available to z/VM 6.4 or higher by the SET STORAGE operator command, then, when z/VM is restarted by a load normal operation or a software initiated re-IPL, that additional storage is used as is any other available online storage. For z/VM 6.4, any additional storage is offline.

- When a standalone dump is to be performed on an LP, perform a load normal (not a load clear) on that LP to keep the reserved storage element online and preserve the storage contents.

CPCs with the Sysplex Failure Manager (SFM)

The Sysplex Failure Manager (SFM) allows you to reset and reconfigure one or more logical LPs and their related storage. SFM allows workload redistribution from the failed primary system to the backup system without operator intervention. For a detailed description of how SFM works, see *z/OS MVS Setting Up a Sysplex*.

Note: These SFM reset/deactivate functions are not compatible with the z/OS AutoIPL function (which is supported on z/OS 1.10 and higher). For example, if AutoIPL is being used on a particular z/OS system (LP) to automatically re-IPL that system if it requests a disabled wait state to be loaded. It is not desirable to have that same LP be the target of one of these cross-partition Reset or Deactivation functions, because these actions prevent the system from re-IPLing itself successfully through AutoIPL.

To allow an LP to initiate these functions, use the **Customize/Delete Activation Profiles** task to open a reset profile to authorize an LP to issue instructions to other LPs. The **Cross partition authority** check box is located on the Security page for the LP.

The following functions exist for SFM:

- Cross Partition System Reset:

  This function causes a specified LP to be reset. The reset is accomplished via the RESETTIME(nnn) keyword in the SYSTEM statement of the z/OS SFM policy.

- Cross Partition Deactivation:

  This function causes a specified LP to be deactivated. The deactivation is accomplished via the DEACTTIME(nnn) keyword in the SYSTEM statement of the SFM policy, and also, the RECONFIG statement in the SFM policy with a specific TARGETSYS(sysname) specified.

- Cross Partition Nonspecific Deactivation:

  This function causes all logical partitions which are currently using any portion of the reconfigurable central storage of the issuing partition to be deactivated. The issuing partition is not deactivated. The nonspecific deactivation is accomplished via the RECONFIG statement in the SFM policy with a non-specific TARGETSYS(ALL) issued.

The Automatic Reconfiguration Facility (ARF) function is a hardware/LP function that is part of the cross-partition authority control setting. ARF functions are used by SFM policy functions within z/OS, when RESETTIME, DEACTTIME, or the RECONFIG statement is coded in the SFM policy.

TARGETSYS(ALL) examples

Specifying an Origin

Assume that the backup partition has specified an origin, minimal initial storage, and a large amount of reserved storage. Since the backup system does not own its complete addressable range, two other partitions, are defined in the reserved storage of the backup partition. See Figure 31 on page 103 for the storage layout before nonspecific deactivation.
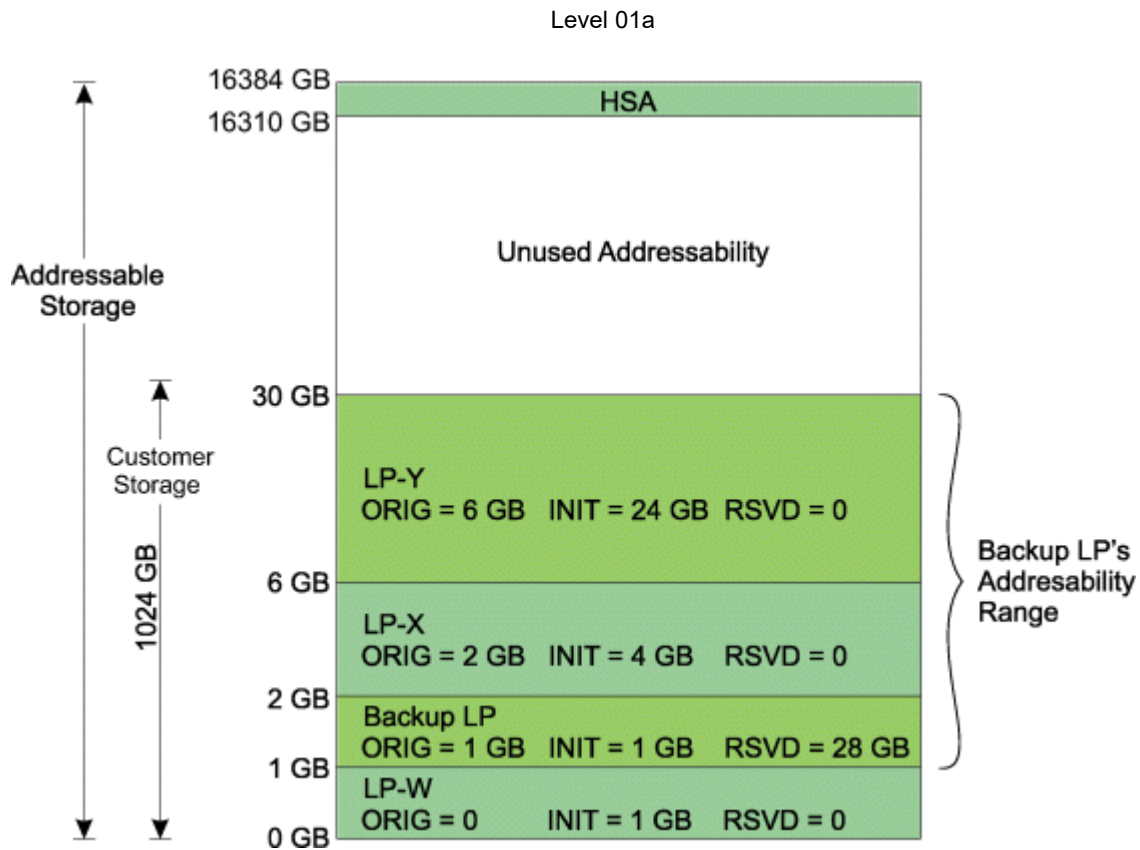
Figure 31. Backup partition layout before nonspecific deactivation

Assume that the backup LP has been given cross partition authority. See Figure 32 on page 103 for the storage layout at the completion of a takeover by the backup LP.
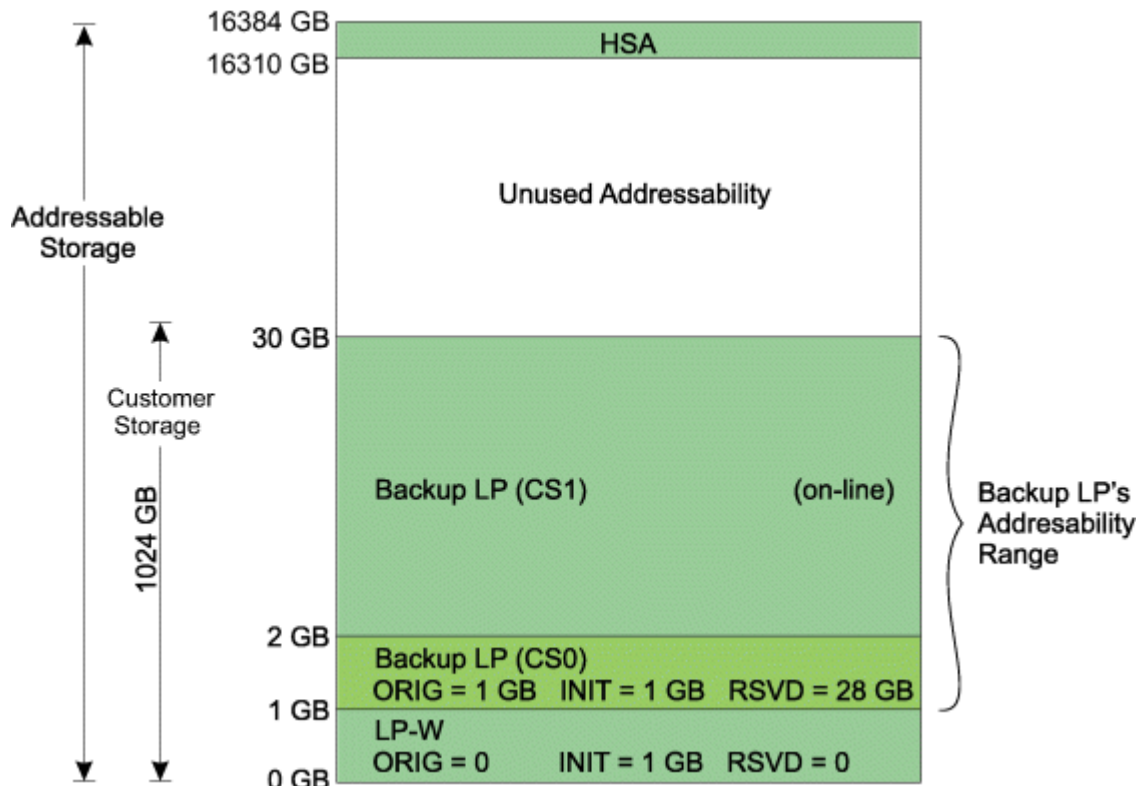
Figure 32. Backup partition layout after nonspecific deactivation

Chapter 3. The characteristics of logical partitions **103**

Number of central processors

The CPs defined to an LP are called logical cores. The total number of initial logical cores for each LP cannot be greater than the number of physical cores installed. CPs can be dedicated to LPs or shared by LPs (*sharing* LPs).

The number of CPs defined for an LP represents the number of logical cores on which the control program will dispatch work and is determined by several factors as described in "Maximum number of central processors" on page 104. If a logical partition enables SMT, each of these logical processor cores will have two CPUs defined. If the processor type supports SMT (for example zIIPs and IFLs), this creates two CPUs per logical core for the control program dispatch work.

On a z15, you can optionally install one or more Internal Coupling Facility (ICF) features for use by a coupling facility LP. See "Coupling facility LPs using dedicated Central Processors (CPs) or dedicated Internal Coupling Facility (ICF) processors" on page 105.

On a z15, you can optionally install one or more Integrated Facility for Linux (IFL) features for use by a Linux-Only LP. See "Processor considerations for Linux-only LPs" on page 105

You can also optionally install one or more zIIP feature for use by an LP. The total number of initial logical zIIPs for each LP cannot be greater than the number of physical zIIPs installed.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to define the number of CPs for an LP. The Number of processors field is located on the Processor page for the LP.

Maximum number of central processors

The maximum number of CPs that can be defined depends on:

- The number of CPs that are available.

  The maximum number of logical cores available for definition in a single LP is the total number of CPs achievable through concurrent CPU upgrade or 190, whichever is less.

  Notes:

  1. The maximum initial logical cores defined for a dedicated partition *equals* physical cores online for the current model *minus* physical cores currently dedicated and online to other dedicated partitions *minus* the maximum online number of shared CPs among all the activated logical partitions that are using shared CPs.

  2. The maximum initially online logical cores defined for a shared partition *equals* physical cores online for the current model *minus* physical cores currently dedicated and online to other partitions using dedicated CPs.

  3. The preceding rules for maximum CPs apply independently to each type (general purpose, ICF, IFL, or zIIP) of processor defined to the logical partition. For instance, when defining a logical partition to use a combination of general purpose CPs and zIIPs, the rules for the maximum initially online of each type of processor is calculated independently against what is currently installed and in use for that type of physical processor.

- The number of CPs that are supported by the required control program.

  Some control programs support as many as 100 CPs. The number of CPs defined for an LP should not exceed the number supported by the control program used for that LP. The stated maximum supported processors for a particular control program is applied against the sum of the counts of all processor types defined to the partition.

Workload requirements

The number of logical cores defined also depends on the workload requirements and the ability of the control program or application program to effectively use multiple logical cores.

- The number of CPs required to meet the peak demands of the LP

  When a sharing LP is activated, it should be assigned enough CPs to meet its peak demands and any immediate growth requirements.

Note: Too few CPs could limit the number of potential transactions, and too many active logical cores could affect performance. In addition to a number of initially online logical cores sufficient to meet the current peak demands of the LP, the definition should include the number of reserved logical cores required for possible growth requirements.

For an LP that uses dedicated CPs, the number of CPs required to meet peak demand should be determined.

- The number of CPs shared by an LP

The physical cores used by a sharing LP can be limited by the number of CPs defined to it. For example, on a six-way CPC, a two-way LP could never get more than a third of the CP resources.

On most machines, there are many possible configurations. For example, if a two-way CPC is to be used by three sharing LPs, configuration options include all two-way LPs, all one-way LPs, or a mix of one-way and two-way LPs. PR/SM manages logical cores according to the specified processor weights.

Three two-way LPs should only be considered if all LPs have peak demands exceeding the capacity of a one-way CPC. In this instance, the average requirements during peak periods should be examined carefully to ensure that the total does not exceed the capacity of the CPC.

Three 1-way LPs are optimal for ITR performance and should be used if no LP has peaks exceeding half of the capacity of a two-way CPC.

Processor considerations for Linux-only LPs

A Linux-Only mode LP can be allocated either general purpose CPs or IFLs. For optimum cost-effectiveness, IFLs should be used whenever possible. Choose IFL CPs on the Processor page of the **Customize/Delete Activation Profile** task. IFLs can be allocated to a Linux-Only mode LP as either dedicated or shared.

Utilization of IFLs will be included in the data reported for ICF processors in RMF Partition Data reports and other similar reports. There is no way to distinguish on such a report between data reported for ICF processors and IFLs.

A Linux-Only LP, whether allocated IFLs or general purpose CPs, will not support any of the traditional operating systems (such as z/OS, z/TPF, or z/VSE). Only Linux or z/VM with only Linux and OpenSolaris guests can run in a Linux-Only mode LP. Logical partitions defined as Linux-Only that attempt to load a traditional operating system will be system check-stopped.

Processor considerations for coupling facility LPs

You can define a coupling facility to use any one of the following CP definitions:

- One or more dedicated general purpose CPs
- One or more shared general purpose CPs
- One or more dedicated ICF processors
- One or more shared ICF processors

Any other combination of processor definitions is not supported.

Coupling facility LPs using dedicated Central Processors (CPs) or dedicated Internal Coupling Facility (ICF) processors

Important: It is strongly recommended using dedicated CPs or dedicated ICFs for production coupling facility LPs because coupling facility channel paths and requests have critical response time requirements. When the coupling facility is running on a dedicated CP or dedicated ICF processors, an *active wait* polling algorithm is used to look for coupling facility requests. This results in the fastest response time and throughput possible for coupling facility requests.

Considerations for coupling facilities using Internal Coupling Facility (ICF) processors

The following considerations apply to coupling facility LPs that use ICF processors.

- ICF processors are managed separately from general purpose CPs. Their presence and usage does not affect any of the definition or activation rules pertaining to LPs that use general purpose CPs.

- ICF processors appear in RMF Partition Data reports and other similar reports only when an LP is activated and assigned to use an ICF processors.

- Non-ICF work is never allowed to be run on an ICF processors. Only PR/SM (to manage the CP) or a logical partition defined to use ICFs is allowed on the ICF processors.

Coupling facility LPs using shared Central Processors (CPs) or shared Internal Coupling Facility (ICF) processors

Using shared CPs for coupling facility LPs has the following considerations. The same set of considerations for using shared CPs equally apply to using shared ICF processors.

- You should **not** cap the processing weights for a coupling facility LP. If you cap processing weights and you are running a coupling facility, PR/SM attempts to support the cap but may not be successful and the system can be less responsive.

- DYNDISP=THIN is the Default for shared processor coupling facilities, because this minimizes the "wasted" shared processor cycles consumed by the CF image to poll for work, compared to other DYNDISP options.

- A CF using DYNDISP=OFF with shared CPs continues to run in a polling loop so it always looks busy, even if there is little CF activity. If the CF is sharing a CP with a z/OS image that has low CPU demands, the CF could use the entire CP. If the CF is sharing CP(s) with another CF, both will look busy and the CPU resource will be split between them. For example, if a production CF and a test CF shared a CP with equal weights, each would only have use of one-half the CP. This would appear on the RMF CF Usage Summary Report as LOGICAL PROCESSORS: DEFINED 1 EFFECTIVE 0.5.

  Regardless of DYNDISP specifications, requests to a CF that is sharing a CP may have to wait if another image (z/OS or CF) is using the shared CP. In fact, the CF's logical cores may have to wait for several logical cores of other partitions to run before it runs again. This has the net effect of elongating response times. Requests to the CF that could typically take several microseconds in a dedicated environment can be elongated by many milliseconds, easily 12.5 to 25 milliseconds or more. If CF response time is not a consideration (for example, a test CF or a backup CF), the installation can further limit CF usage of CPU resource by enabling the Dynamic Dispatch option. With DYNDISP=ON, this will periodically pause the polling loop. The number and length of the pauses is determined by the activity in this CF. While this gives more CPU resource to the other partitions that are sharing the CP, it elongates the service time in this CF, in most cases, by a considerable amount. With DYNDISP =THININTERRUPT, the coupling facility voluntarily gives up control of the shared coupling facility processor whenever it runs out of work to do, relying on coupling thin interrupts to cause the image to get re-dispatched.

- Shared ICF and shared general purpose processors are each managed as separate "pools" of physical resources. As such, the processing weights assigned to logical partitions using shared ICF processors are totaled and managed separately from the total weights derived from all of the logical partitions using shared general purpose processors. Similarly, when the processor running time is dynamically determined by the system, the calculation of the dynamic running time is performed separately for each pool. If a user supplied run time is specified, this will apply to both processor pools.

- RMF identifies which logical and physical cores are of each type when both general purpose and ICF processors are present in the configuration on its partition data report.

- An uncapped coupling facility LP honors its processing weight up to a point. PR/SM attempts to help ensure that each logical core defined to a coupling facility LP gets at least 1 run time interval of service per every 100 milliseconds. For example, for a typical 12.5 millisecond run time interval, each logical core gets 1/8th of a physical cores. This can translate into a response time elongation that is several thousand times as long as a typical CF request using dedicated CPs.

  With dynamic coupling facility dispatch, the coupling facility will not necessarily consume entire run time intervals at low request rate times. In low request rate times, each coupling facility engine can consume far less than 1/8th of a physical cores. The CP resource consumption can be more in the 1-2% range. At higher request rates (for example, when the coupling facility is actually busy handling requests), the 1/8th minimum will again become effective.

Note: Anticipated processor usage by a coupling facility may spike much higher than what you would intuitively expect given the non-CF workload. For instance, system reset of a system or logical partition that was communicating with (connected to) a coupling facility can temporarily cause a considerable increase in the demands placed on the coupling facility.

- All requests to coupling facilities from z/OS LPs that share CPs with a coupling facility are treated internally to the machine as asynchronous requests. This is true even if the requests are to a coupling facility that the z/OS LP is not sharing CPs with. This conversion is transparent to z/OS but it can result in increased synchronous service times to the coupling facility as reported by RMF. As far as the operating system, RMF, and the exploiter code is concerned, the requests that are initiated synchronously by software are still being processed synchronously, and they show up as such on RMF reports.

- Choose a weight for the coupling facility LP based on the anticipated CP requirements of the coupling facility. When deciding how much CP resource should be given to each coupling facility logical core, consider the following:

  - When using dynamic CF dispatching including Thin Interrupts, (DYNDISP ON or THIN) the weight for the coupling facility can safely be set to a value that affords the proper CP resources to the coupling facility in times of the highest volume of requests to the coupling facility. In low request rate periods, the coupling facility will automatically throttle back on its CP usage making the CP resource available for redistribution to the other defined LPs in the configuration. Also, note that at low request rate times, RMF Coupling Facility Activity Reports will show some elongation in response times for requests to the coupling facility. With the low rate of these requests, overall system performance should not be noticeably impacted.

  - CP resource requirements vary depending on the coupling facility exploiter functions and your sysplex hardware and software configuration. As a general guideline, when the anticipated CP requirement of the coupling facility is less than one physical core, set the weight of the coupling facility LP so that the coupling facility logical core has 50% or more of a physical core resource. If you have less than a full physical core allocated to the coupling facility logical core, **this will result in elongation of response times**.

  Examine the RMF Coupling Facility Activity Report and tune the coupling facility LP weight until your performance requirements are met.

  Note: When examining the RMF Coupling Facility Activity Report, you may see elongated average response times. Usually, these are accompanied with high standard deviations as well. This indicates most requests are in the expected range with an occasional very elongated response that throws the average off.

  - Give more weight to the coupling facility LP for functions that have more stringent responsiveness requirements. For example, you can set the coupling facility weight higher for coupling facility exploiter functions such as IMS/IRLM. For IMS/IRLM, you may want to set the weight of the coupling facility LP so that each coupling facility logical core runs almost dedicated. For example, you may want to set a weight that will give each coupling facility logical core 95% or more of physical core resources. In another case, if the CF contains structures which are using System Managed Duplexing, you should set the weight of the coupling facility LP so that the coupling facility CP has at least 95% of a physical core. If the coupling facility has less than 95% of a physical core, there is a possibility that response from the partner of the duplexed structure will timeout and the duplexing of that structure will cease.

  - Less weight may be required for coupling facility exploiter functions, such as the JES2 Checkpoint data set. If your coupling facility is being used exclusively as a JES2 checkpoint, your coupling facility responsiveness requirements may be less stringent. If this is the case, try decreasing the weight of the coupling facility LP so that each coupling facility logical cores receives, for example, 40-50% of physical core resources.

  - As the total traffic (requests per second) to the coupling facility increases, there is a greater need for real coupling facility CP time. To a point, the increase in traffic may not require an increase in the coupling facility LP weight. This is because coupling facility active wait time turns into coupling facility busy time. You must monitor coupling facility utilization and adjust the LP weight to help ensure that your performance requirements are being met.

- Even in a test environment, the above guidelines should be followed. Optimally, a weight resulting in approximately 50% or more of a processor should be made for each coupling facility logical processor. *Failure to provide sufficient weight to a coupling facility may result in degraded performance, loss of connectivity, and possible loss of coupling links due to time-outs.* Dynamic CF dispatching must not be set to OFF. Additionally, processor resource capping of the coupling facility logical partition(s) logical cores must **not** be enabled.

Considerations for coupling facilities running on uniprocessor models

On a uniprocessor or smaller machine, It is strongly recommended that coupling facility LPs should not share general purpose CPs with non-CF workloads (for example, z/OS), even in a test environment. While it is possible to use general purpose CPs as CF processors and share the CPs with z/OS and other similarly defined CFs, this is not a recommended configuration. Using CPs as CF processors creates interaction between z/OS and the CF in terms of cache reuse and other factors that may impact performance. For production configurations, It is strongly recommended using one or more Internal Coupling Facility (ICFs) CPs for the coupling facility LPs.

For a test configuration, if this option is not available and coupling facility LPs must share a general purpose CP with non-CF workloads, adherence to the preceding recommendations for coupling facilities using shared CPs minimize undesirable consequences.

For greater detail on how to plan for and set up a parallel sysplex in a shared CP environment, see:

> **MVS/ ESA Parallel Sysplex Performance - LPAR Performance Considerations for Parallel Sysplex Environments** http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/FLASH10002).

For more information on using DYNDISP options to share processors, see:

> **Coupling Thin Interrupts and Coupling Facility Performance in Shared Processor Environments** http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102400

.

Processor considerations for z/VM mode LPs

You can define a logical partition to use one or more zIIPs, IFLs, and ICFs with z/VM 6.4 or higher with either of the following combinations:

- One or more dedicated general-purpose CPs and one or more dedicated zIIPs/IFLs/ICFs.
- One or more shared general-purpose CPs and one or more shared zIIPs/IFLs/ICFs.

The partition mode must be set to **z/VM** in order to allow zIIPs, IFLs, and ICFs to be included in the partition definition.

In a **z/VM** mode partition, z/VM will:

- Operate z/TPF, z/VSE, and z/OS guests on CPs.
- Operate Linux on Z machines guests on IFLs and optionally on CPs.
- Offload z/OS guest system software process requirements, such as DB2 workloads, on zIIPs, and optionally on CPs.
- Provide an economical Java execution environment for z/OS guests on zIIPs and optionally on CPs.
- Operate coupling facility virtual machines in support of a Parallel Sysplex test environment on ICFs and optionally on CPs.

For additional information about using these capabilities of z/VM, refer to *z/VM Running Guest Operating Systems*, SC24-6115.

Processor considerations for LPs with multiple CP types

You can define a logical partition to use one or more zIIPs with either of the following combinations:

- One or more dedicated general purpose CPs and one or more dedicated zIIPs
- One or more shared general purpose CPs and one or more shared zIIPs.

The mode specified for the logical partition must be set to **General** or **z/VM** in order to allow the definition of zIIPs.

Dedicated central processors

An LP can have CPs dedicated to it. When an LP that uses dedicated CPs is activated, a physical cores is assigned to each defined logical core. The LP then has exclusive use of its physical cores.

The physical cores that belong to an LP that uses dedicated CPs are always available for its use, but the capacity that is not used cannot be used by other LPs.

For coupling facility considerations, see "Processor considerations for coupling facility LPs" on page 105.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to define dedicated CPs for an LP. The Dedicated field is located in the Logical processor assignment group box on the Processor page for the LP.

Suitable workloads

Workloads best suited for logical partitions that use dedicated processors are those that maintain a fairly even external throughput rate (ETR) while using most of the capacity of the logical partition. Logical partitions with timing dependencies might require dedicated processors.

The installation goals for a logical partition that uses dedicated processors should be similar to the goals of the processor complex. For example, if the goal is an average processor utilization rate of 70%–85%, then the same goal should be used for a logical partition that uses dedicated processors.

Shared central processors

LPs can share CPs with other LPs. A sharing LP does not have exclusive use of the physical cores. There is no correspondence between the logical cores in a sharing LP and the physical cores on which the logical cores are dispatched (except on a one-way CPC). A logical core can be dispatched on any physical core and, on successive dispatches, the logical core can be dispatched on different physical cores.

The number of CPs available for use by sharing LPs is determined by adding the number of CPs already assigned to active, dedicated LPs and subtracting that sum from the total number of physical cores available.

The total of all logical cores for all sharing LPs can be larger than the number of physical cores serving the sharing LPs. For example, if four LPs are active on a six-way CPC and each LP is defined to have four CPs, the total number of online logical cores is 16.

For coupling facility considerations, see "Processor considerations for coupling facility LPs" on page 105.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to define shared CPs for an LP. The Dedicated processors check box is located in the **Logical processor assignment** group box on the Processor page for the LP.

Suitable workloads

Workloads best suited for sharing logical partitions are those that have a widely fluctuating ETR or would not fit well into the capacity of a dedicated logical partition. If a workload can use only a small portion of the capacity of a dedicated logical partition, redefine the logical partition to use shared processors to free the available capacity for use by other logical partitions.

A workload with a widely fluctuating ETR would experience peaks and valleys in its processor-utilization curve. Such fluctuations can occur over extremely short periods of time (minutes or seconds). This type of workload could take advantage of the time- and event-driven dispatching available. With event-driven dispatching, a sharing logical partition receives the resources required as needed and leaves the capacity free for other logical partitions when not needed.

When combining workloads on a processor complex by means of logical partitions, examine their average and peak requirements carefully. If the workloads fluctuate over very short intervals, the total capacity of the system must meet the sum of the average requirements for each workload. If processor utilization

fluctuates over longer periods, and the peak utilization periods for these workloads occurs simultaneously, then the total capacity of the logical partitions must meet the sum of the peak requirements for each workload.

Sharing logical partitions that use event-driven dispatching are better able to maintain high transaction rates with fluctuating demand while being responsive. However, the Internal Throughput Rate (ITR) for a sharing logical partition is lower than the ITR for a dedicated logical partition.

The capability to limit CPU usage for any or all logical partitions with shared processors is provided by the PR/SM capping function. The capping function enhances PR/SM workload balancing controls for environments with a requirement that the CPU resources for a logical partition be limited. Capped logical partitions are recommended for use when CPU resources must be limited for business reasons (in accordance with a contract), or when the impact that one logical partition can have on other logical partitions needs to be limited.

Processing weights

An LP with dedicated CPs is not affected by processing weights.

Processing weights are used to specify the portion of the shared CP resources allocated to an LP. Although PR/SM always manages sharing LPs according to the specified processing weights, there are times when an LP will receive either more or less than its processing share:

- An LP will receive more than its processing share when there is excess CP capacity, provided it has work to do and other LPs are not using their share.
- An LP will receive less than its processing share when its workload demand drops below the capacity specified by its weight.
- An LP will not receive more than its processing share when the CP resources for that LP are capped.

The recommended procedure is to specify processing weights to satisfy the peak requirements of the LPs.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to define the processing weight for a shared LP. The Processing weight field is located in the Logical processor assignment group box on the Processor page for the LP.

Use of processing weights

As an example, consider a system with 6 CPs and 3 LPs defined as follows:

```
        LP Name   Logical Cores  Weight
        -------   -------------  ------
        ZVSE           1          300
        ZOSTEST        6          100
        ZVM            2          900
```

Processing weights can range from 1 to 999 (weights of less than 2% difference are not considered significant) and are used as follows:

- The processing weights for all active, sharing LPs are added together. This total is considered to be 100% of the processing resource available to shared CPs.
- The share of processing resources for each LP is calculated by dividing the processing weight for each sharing LP by the total processing weight. For example, at peak CP utilization levels, the dispatcher allocates shared processing resources to each of the LPs as follows:

```
        ZVSE         300/1300 = 23.1%
        ZOSTEST      100/1300 =  7.7%
        ZVM          900/1300 = 69.2%
```

- The share of processing resource for each online logical core with HiperDispatch disabled in the logical partition is calculated by dividing the share for each LP by the number of online logical cores. The share for each logical core is as follows:

```
ZVSE        23.1/1 CP  = 23.1%
ZOSTEST      7.7/6 CPs =  1.3%
ZVM         69.2/2 CPs = 34.6%
```

These percentages are used to determine preemption priority for I/O interruptions. A lower priority logical core can be preempted when an I/O interruption is pending for a higher priority logical core when the following occurs:

– The higher priority logical core is further behind in its share, or

– The higher priority logical core is not as far ahead of its share as the lower priority logical core.

For example, the lower priority LP is receiving 15% more than its processing share, and the higher priority LP is receiving 10% more than its processing share.

As long as there is excess CP capacity, processing weights have no effect on the CP resources consumed. Weights affect processing when the number of logical cores that need processing time is greater than the number of physical cores available.

Processing weights and shared CP, ICF, IFL, and zIIP processors

Shared general purpose, ICF, IFL, and zIIP, processors are each managed as separate *pools* of physical resources. The processing weights assigned to logical partitions using shared ICF, IFL, zIIP, or general purpose processors are totaled and managed together only with the total weights from all of the logical partitions using the same processor type. The calculations shown in previous examples as well as the examples to follow are done independently for general purpose, ICF, IFL, or zIIP processors, on a machine that has them.

Processing weights for logical partitions with multiple shared CP types

When a logical partition is defined to use one or more shared general purpose CPs and one or more shared zIIPss, each of the types of logical processors is managed independently. The shared general purpose CPs compete with all other shared general purpose CPs defined in other logical partitions in the configuration. The zIIPs compete with all other shared zIIPs defined in other logical partitions. General purpose, zIIP processors each have a separate processing weight specified.

Note: If WLM weight management is being used for such a logical partition, only the weight of the shared general purpose processor portion of the logical partition will be altered by WLM. The specified weight for an LPs zIIP processors is unaltered by WLM.

Effects of processing weights

Several implications are derived from the rules described above. First, every time a sharing LP is activated or deactivated, the share of all other active LPs, using the same processor types, changes. This happens because the total of the processing weights has changed.

Because the processing share of LPs can vary, the actual utilization reported by monitors such as RMF can be different from the weights. In systems at less than 100% utilization, some LPs could receive greater than their share if other LPs are not using their share. The number of CPs defined also affects the maximum resource allocated to that LP. For example, an LP defined to have two CPs on a three-way CPC can never be allocated more than 67% of the CP resources no matter what its processing weight.

Capping processing weights

The PR/SM capping function provides the capability of limiting CPU resource usage for one or more processor types for one or more LP. The relative processing weight of a processor type for an LP is its capping value for that processor type.

A capped LP running at its cap for a capped processor type does not have access to the CP resources that are not utilized by other LPs. However, CP resources that are not used by a capped LP can be used by other LPs. Equitable distribution of CP resources is maintained.

Capping values can be dynamically adjusted. The capping function can be turned on and off independently for each defined processor type in an LP, and provides the capability of specifying capping for individual LPs without a re-IPL of the LP.

Use the **Change Logical Partition Controls** task to change the Capped setting for the specific logical partition. Selecting the **Initial Capping** check box turns the capping function on. **Initial Capping** is set independently for each processor type in the logical partition. Click **Save Running System** to have this option take effect immediately for an active partition. (**Save and Change** changes the running system and updates the partition's profile. **Save to Profiles** does not change the running system; it just saves the new definition to the partition's profile.) If you do not need to change a running system, the **Customize/Delete Activation Profiles** task can be used to open a reset or image profile to cap processing weight for an LP. This change would take effect when the partition is activated.

With HiperDispatch disabled, an LPs relative weight for a processor type is divided by the number of shared logical cores online of that type for the LP to give the share for each logical core. The goal of the PR/SM dispatcher is to give each logical core of that processor type its share of the total relative weight. Capping is done on a logical core basis.

An LPs share of CP resources for a processor type is determined by its weight for that processor type. The combined processing weights for all active LPs for a particular processor type are considered to be 100% of the available shared CP resources for that processor type. The activating and deactivating of LPs changes the amount of CP resources available to LPs, making the percentage of CP resources requested for each active LP a relative one, and not a fixed percentage of CP resources.

Note: If an extremely low processing weight is specified for a logical partition using capping, tasks such as **Reset Clear**, **Activate**, and **Deactivate** may fail due to a time-out. To prevent this problem, avoid use of capping in conjunction with low processing weights. A preferable solution is specification of processing weights that results in a 1/10 share (the suggested minimum) or greater of one physical core for a logical core. If the extremely low weight cannot be avoided, temporarily turn all capping off for the logical partition prior to activating, deactivating, resetting, or loading the logical partition. Restore the cap(s) following completion of these operations.

Enforcement of processing weights

Processing weight management

PR/SM enforces LP processing weights as follows:

- For LPs with processor resource capping, PR/SM enforces the processing weights to within 3.6% of the LP's physical cores share for logical cores entitled to 1/10 or more of one physical cores. Typically, PR/SM manages processing weights to within 1% of the LP's physical core share. See "Example 1. With processor resource capping" on page 113.

- For LPs *without* processor resource capping, PR/SM enforces the processing weights to within 3.6% of the LP's physical cores share for logical cores entitled to 1/2 or more of one physical cores. Typically, PR/SM manages the processing weights to within 1% of the LP's physical cores share. See "Example 2. Without processor resource capping" on page 113.

- If a logical cores falls outside the enforceable ranges for logical cores entitled to less than 1/10 of a physical cores using capping or less than 1/2 of a physical cores not using capping, PR/SM enforces the processing weights to within 3.6% of the total capacity of the shared physical cores resources. However, PR/SM should typically manage the processing weights to within 1% accuracy.

- Unused CP cycles to which a logical cores is entitled are made available to other logical cores in proportion to the weights set for the other logical cores.

- An uncapped coupling facility LP with shared CPs and coupling facility channels defined honors its processing weight up to a point. PR/SM attempts to help ensure that each logical cores defined to a coupling facility LP gets at least 1 run time interval of service per every 100 milliseconds. For example, for a typical 12.5 millisecond run time interval, each logical cores gets 1/8th of a physical cores.

Processing weight management examples

For the formulas used in the following two examples, see "Processing weight management formulas" on page 114.

Example 1. With processor resource capping

In the following example:

- Six physical cores are online
- All LPs are capped
- All LPs have sufficient workload demand to use their shares

Table 13. PR/SM processor weight management with processor resource capping and with HiperDispatch Disabled

LP	LCPs Online	Weight (1)	Weight per LCP	LCP % of PCP	3.6% Share	3.6% Total	Resulting Weight Range	Resulting Utilization Range
A	6	500	83.3	50	3.0	-	80.3 - 86.3	48.2% - 51.8%
B	3	480	160.0	96	5.8	-	154.2 - 165.3	46.3% - 49.7%
C(2) C(3)	3	20	6.7	4	0.2	- 12	6. - 6.9 0.0 -18.7	2.0% - 2.1% 0.0% - 5.6%
Total Capacity of the Shared PCP Resources	1000 / 6 PCPs = 166.7 = PCP capacity							

Legend:

LCP
    Logical cores
PCP
    Physical cores

Notes:

1. Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to set the processing weight for an LP. The Processing weight field is located on the Processor page for the LP.
2. The logical cores for LP C (2) fall outside the enforceable range because each logical cores share is only 4% of a physical core.
3. The LP's range is bounded by the line shown for LP C (3). Typically, though, even this configuration will see results comparable to the line shown for LP C (2).

Example 2. Without processor resource capping

In the following example:

- Six physical cores are online
- No LPs are capped
- All LPs have sufficient workload demand to use their shares

Table 14. PR/SM processor weight management without processor resource capping and with HiperDispatch Disabled

LP	LCPs Online	Weight (1)	Weight per LCP	LCP % of PCP	3.6% Share	3.6% Total	Resulting Weight Range	Resulting Utilization Range
A	6	500	83.3	50	3.0	-	80.3 - 86.3	48.2% - 51.8%
B	3	300	100.0	60	3.6	-	96.4 - 103.6	28.9% - 31.1%
C (2) C (3)	3	200	66.7	40	2.4	- 12	64.3 - 69.1 54.7 - 78.7	19.3% - 20.7% 16.4% - 23.6%
Total Capacity of the Shared PCP Resources	1000 / 6 PCPs = 166.7 = PCP capacity							

Legend

LCP
    Logical cores
PCP
    Physical cores

Notes:

1. Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to set the processing weight for an LP. The Processing weight field is located on the Processor page for the LP.

2. The logical cores for LP C (2) fall outside the enforceable range because each logical cores share is only 40% of a physical core.

3. The LP's range is bounded by the line shown for LP C (3). Typically, though, even this configuration will see results comparable to the line for LP C (2).

Processing weight management formulas

The following formulas were used to compute the values in the previous two examples:

$$Weight\ per\ LPC = \frac{LPCTL\ Weight}{Number\ of\ LCPs\ online}$$

$$LCP\ percent\ of\ PCP = \frac{LPCTL\ Weight}{Total\ of\ LPCTL\ Weights} \times \frac{Number\ of\ PCPs\ online}{Number\ of\ LCPs\ online}$$

$$3.6\ percent\ Share = Weight\ per\ LCP \times 3.6\ percent$$

$$3.6\ percent\ Total = \frac{Total\ Weight \times 3.6\ percent}{Number\ of\ LCPs\ online}$$

$$Resulting\ Weight\ Range = Weight\ per\ LCP \pm 3.6\ percent\ Share$$
$$or$$
$$Resulting\ Weight\ Range = Weight\ per\ LCP \pm 3.6\ percent\ Total$$

$$Resulting\ Utilization\ Range = Resulting\ Weight\ Range \times \frac{Number\ of\ LCPs\ online}{Total\ Weight}$$

Maintaining the same relative percentages of CPU resources

To maintain the same relative percentage of CP resources requested for a capped LP, processing weights should be readjusted immediately prior to, or immediately after, the activation or deactivation of an LP.

Processing weight values for use when specific LPs are activated or deactivated should be calculated in advance, and be readily available. It is recommended that a convenient method be developed for changing the processing weight values to readjust relative shares after an LP is activated or deactivated.

For example, if the sum of the weights of the active LPs totals 100, then the sum of the relative weights of the active LPs also totals 100. This provides an easy means for adjusting weights upon the activation or deactivation of LPs. Another good approach to maintaining the desired share for a capped LP is to also readjust the processing weights for LPs with the capping function turned off, as shown in Table 15 on page 115 for the LP ZOSTEST.

Table 15. Example of maintaining relative weight of a capped logical partition						
	Four LPs Active			**Three LPs Active**		
LP Name	**Status**	**Weight**	**Capped**	**Status**	**Weight**	**Capped**
ZVSE	A	30	No	D	-	-
ZOSPROD	A	40	No	A	64	No
ZOSTEST	A	20	Yes	A	20	Yes

When the sum of all the relative weights is maintained at 100, it is easy to recalculate weights when an LP is deactivated. After deactivating ZVSE, the weight for ZOSPROD can be changed to 64 to maintain the same relative weight of 20 for ZOSTEST, the capped LP.

Capping in a single logical partition

In order to use capping for an LP on a CPC where there is a need for only one active LP using shared CPs, you must define and activate a second *dummy* LP. The dummy LP must also be defined as using shared CPs. The weights of the two LPs can be adjusted to attain the desired cap for the one LP that will actually be used.

The *dummy* LP does not have to be capped. In most cases, the *dummy* LP does not need to have anything IPLed into it. If nothing is IPLed into the *dummy* LP, the *dummy* LP will not use any CP resources. In some situations where the single capped LP runs an extremely CP-intensive workload, it may be necessary to

run a program in the *dummy* LP to smooth distribution of CP resources to the LPs. The program can be a simple branch loop that spins to waste time. Without this program running in the *dummy* LP, the CP-intensive, capped LP can experience a lurching effect with its dispatched times. It will be capped to its weight properly, but it could get all of its allotted CP resource quickly and then wait for a period of time before it can run again.

If nothing is to be IPLed into the *dummy* LP, the *dummy* LP can be defined and activated with no channel paths or devices. A definition with no devices is desirable to prevent control unit logical paths from being established for the *dummy* LP. See "Managing logical paths for FICON channels" on page 38.

As alternative to this procedure for capping a single logical partition, consider using the absolute capping support described in "Absolute capping" on page 116 or capping a group of logical partitions described in "Absolute group capping" on page 116.

Absolute capping

Absolute capping provides an optional absolute capacity setting for logical partitions specified in absolute processor capacity (for example, 2.5 processors) in terms of cores. This setting is specified independently by processor type and provides an upper limit on the processor type in the partition as a whole at this capacity.

The shared partition's processing weight still dictates the logical partition priority compared to other shared logical partitions.

Absolute capping is most effective for absolute caps higher than what the partition's weight relative to other logical partitions capacity would deliver. In fact, absolute capping is not recommended to be set below what the logical partition's weight capacity would deliver.

Absolute capping is ideal for processor types and Operating Systems that WLM does not manage. It is not meant as a replacement for defined capacity or group capacity for z/OS.

Absolute group capping

Absolute group capping provides an optional absolute capacity setting for a group of logical partitions specified in absolute processor capacity (for example, 2.5 processors) in terms of cores. This setting is specified independently by processor type and provides an upper limit on the processor type in the group of partitions at this capacity.

The shared partitions processing weights still dictate the logical partition priorities compared to other shared logical partitions.

Absolute group capping is most effective for absolute caps higher than what the collective partition's weights relative to other logical partitions capacity would deliver.

Absolute group capping is ideal for processor types and operating systems that WLM does not manage. It is not meant as a replacement for group capacity for z/OS.

HiperDispatch and Shared Logical Partitions

The z15 provides a higher level of synergy between the PR/SM Hypervisor and z/OS software for managing logical core resource allocations, called HiperDispatch.

HiperDispatch is the true start of having z/OS software become aware of the topology of the machine, presented as a logical partition topology, to then provide dynamic affinity dispatching with regards to CP placement in the structure.

PR/SM has traditionally managed shared logical partitions as being *horizontally* polarized. That is, the processing weight for the logical partition is equally divided between all the online logical cores in the logical partition. In turn, the OS running in the logical partition is obligated to use all of those online logical cores equally in order to be assured of receiving its fair share of the physical core resources. HiperDispatch introduces a new, optional form of polarization for managing the shared logical cores of a logical partition called *vertical* polarization. z/OS running with HiperDispatch uses the logical partition topology information to decide how to group its logical cores to set up its work queues and to exploit the vertical configuration of logical cores to pack work into a smaller set of logical cores, optimizing processor cache usage.

With HiperDispatch there is a common understanding in the PR/SM Hypervisor and the z/OS software such that work can be concentrated on a smaller number of logical cores within the logical partition that reflects the actual assigned weight of the logical partition. With HiperDispatch enabled in a logical partition, z/OS will redispatch its tasks back to a smaller set of logical cores, and PR/SM in turn can dispatch those logical cores back to the same physical cores, optimizing use of the L1, L2, L3, and L4 caches. Work can still expand and flow into more logical processors dynamically should workload conditions warrant it.

Enabling HiperDispatch for z/OS Logical Partitions

HiperDipatch is enabled within a logical partition via software controls. To enable HiperDispatch under z/OS, specify a new keyword in the IEAOPTxx member of SYS1.PARMLIB. HiperDispatch=YES needs to be specified to turn on this mode of dispatching. The enablement is dynamic and can be changed without an IPL. Refer to the z/OS publications for further details. There are no new hardware controls or settings to enable use of HiperDispatch within a logical partition; however, WLM management of logical processors normally works best with **global performance data security** setting enabled. When global performance data security setting is disabled, one should usually be conservative in the number of logical processors defined in excess of the share of the type of processor (CP/zIIP) . The share of a type of processor in a partition is expressed as the number of physical processors that a partition is entitled to by the weight of the partition. For example a machine with 8 general purpose CPs and two partitions with weights of 600 and 400. The share for partition with weight 600 is 600 / (600 + 400) * 8 = 4.8 physical processors. If one defines 8 logical processors for this partition, 5 will be active at all times and up to 3 will be active only in periods that WLM determines 1 or more of the 3 has a good opportunity to increase the net capacity consumed for this partition. In the example, if the second partition is using its share of 3.2 physical processors, having one or more of the 3 excess logical processors in the first partition is not sensible in most cases since there is no excess *abandoned* share from the second partition.

For partitions with global performance data active, WLM is able to manage the excess logical processors that are active to intervals where other partitions have not completely used their share. When global performance data is disabled, WLM uses all logical processors at all times which may lead to attempts to consume processor that can not be achieved. The efficiency of the partition may be less with 8 logical processors active at all times in the example above. Defining 6 logical processors for the first partition is normally a better choice. If for some period of time the 8 logical processors are likely to use abandoned share from other partitions, 8 is a good number.

Allocating Processing Weights within a logical partition using HiperDipatch

Depending on the configuration of the logical partition running with HiperDipatch enabled, logical processors have high, medium or low vertical polarity. Polarity describes the amount of physical processor share vertical logical processors are entitled. The relative processing weight that is defined for the logical partition effectively defines the amount of physical processor cycles the logical partition is entitled.

Vertical polarity is measured by the ratio of a logical partition's current weight to the number of logical processors configured to the logical partition. High polarity processors have close to 100% CP share. Medium polarity processors have >0 to close to 100% shares and low polarity processors have 0% share (or very close to it). Medium polarity processors act as an *anchor* with sufficient weight reserved to allow the medium and low polarity processors to get dispatched in times of contention. The medium and low processors employ a new sharing algorithm to draw on this portion of the partition's processing weight for medium and low processors. As such, PR/SM reserves at least 1/2 a physical cores worth of weight in the medium polarity processor assignments, assuming the logical partition is entitled to at least that much service. High polarity logical cores will be assigned a physical processor to run on very similar to dedicated CPs but the shared high polarity CP can still give up the physical resource and allow other shared logical cores to use its excess cycles. The key here then becomes that the OS software sees the logical topology and tries to exploit the highly polarized logical cores for its work queues.

For example, consider a CPC that has 3 physical cores with 2 active logical partitions, each defined with 2 logical processors and each with a processing weight of 50. If the first logical partition enabled HiperDispatch, it would have 1 high polarity and 1 medium polarity logical cores.

```
50/100 x 3 = 1.5 physical cores
```

Effectively, one logical cores in the Hiperdispatch enabled logical partition is given an allocation of 33 1/3 for its portion of the partition's processing weight, this is the high polarity logical processor. This processor is also assigned a physical processor similar to a dedicated logical cores. The other logical cores, the medium polarity CP, is allocated 16 2/3 for its processing weight, effectively entitling it to 50% of one physical cores.

As a second example, suppose the same three-way processor now has 3 active logical partitions, each with 2 logical processors and each with a processing weight of 50. If the first logical partition enabled HiperDipatch, it would have 1 medium polarity and 1 low polarity logical cores. No high polarity logical cores are assigned because at least 1/2 a physical cores is kept in the medium/low pool.

```
50/150 x 3 = 1 physical core
```

In this case, one logical cores in the Hiperdispatch enabled logical partition is given the complete allocation of 50 for its portion of the partition's processing weight; this is the medium polarity logical processor. There are no high polarity processors in this example. The other logical cores, the low polarity CP, is allocated 0 for its processing weight. Note the allocations for the medium and lows are really for bookkeeping. The OS knows it has some capacity available for use by this set of logical cores but it should only be expanding into these beyond the count of medium CPs when there is excess capacity available in the system because some other logical partition is not demanding its allocation. When the mediums and lows demand CP resource, they will effectively share the processing weight that was allocated to the medium logical cores.

The logical partition's processing weight has a direct effect on the number of high polarity processors the logical partition will have when running with HiperDispatch. You should take care to set your processing weights to get your workload optimally allocated to the desired set of high polarity processors.

When a logical partition chooses to run with HiperDispatch enabled, the entire logical partition runs enabled. This includes all of its secondary processors such as zIIPs. It is the responsibility of the user to define processing weights to all of the processor types for these logical partitions that will achieve the desired level of vertical processor allocations for each type.

HiperDispatch was created primarily for logical partitions using shared logical cores, but HiperDispatch can be enabled in a logical partition using dedicated CPs. In this case, no change is seen in the way the PR/SM Hypervisor will treat the dedicated logical partition, but OS will have knowledge of the logical topology it has been assigned and it will localize the redispatch of its tasks to get optimal use of the processor caches.

z/VM 6.4, 7.1, and 7.2 HiperDispatch support

z/VM 6.4, 7.1, and 7.2 HiperDispatch improves CPU efficiency by causing the Control Program to run work in a manner that recognizes and exploits Z machine topology to increase the effectiveness of physical machine memory cache. This includes:

- Requesting the PR/SM Hypervisor to handle the partition's logical processors in a manner that exploits physical machine topology
- Dispatching virtual servers in a manner that tends to reduce their movement within the partition's topology
- Dispatching multiprocessor virtual servers in a manner that tends to keep the server's virtual CPUs logically close to one another within the partition's topology.

z/VM 7.2 HiperDispatch can also improve CPU efficiency by automatically tuning the LP's use of its logical CPUs to try to reduce multiprocessor effects. This includes:

- Sensing and forecasting key indicators of workload intensity and of elevated multiprocessor effect

- Autonomically tuning the z/VM system to reduce multiprocessor effects when it is determined that z/VM 7.2 HiperDispatch can help to improve CPU efficiency.

Processor running time

The processor running time is the length of continuous time (determined by the dispatch interval) allowed for the dispatch of a logical core.

When the processor running time is dynamically determined by the system, the calculation of the default running time is performed separately for general purpose and ICF processors. All logical processors that are shared, either general purpose or ICF, will be assigned a default running time but the values used for each type of processor may be different. If a user-defined run time is specified, the value applies to all shared general purpose and shared ICF processors. For shared logical partitions using HiperDispatch, any vertical high polarity logical processor's run time is not affected by this user-defined setting. Rather, vertical high polarity logical processors have a run time of 100 milliseconds.

The default value is dynamically calculated and changes when the number of active, scheduled logical core changes.

The default running time is determined using the formula:

$$\frac{25 \text{ milliseconds} \times (\text{number of physical shared CPs})}{(\text{total number of logical CPs not in stopped state for all sharing LPs})}$$

The default value is used whenever the processor running time is dynamically determined by the system. The run-time value can change whenever an LP is activated or deactivated and when a logical core stops or starts (for instance, when a logical cores is configured online or offline). The default processor running time is limited to the range of 12.5 to 25 milliseconds.

The logical cores might not use all of its run time because it goes into a wait state. With event-driven dispatching, when a logical cores goes into a wait state, the physical core is reassigned to another logical cores ready to do work. When a logical cores does not go into a wait state during its run time, it loses the physical cores when it reaches the end of its run time. Therefore, an LP with CP-bound work cannot permanently take control of the physical cores.

You can choose to set the runtime value yourself. However, when event-driven dispatching is enabled, it is generally recommended that the processor running time be dynamically determined. If event-driven dispatching is disabled, you should consider setting runtime values of 2 to 8 milliseconds. The recommended procedure is to start by using the default processor running time. That value should be acceptable when all sharing LPs have similar proportions of interactive work; for example, two LPs each running 40% - 70% of interactive work.

Adjustments to the runtime value might be necessary when one sharing LP contains a large proportion of CP-bound work and the other sharing LPs contain only short, interactive transactions. Degraded response time in one LP can indicate that the runtime value should be reduced to decrease the length of continuous time given to CP-bound work. The run-time value should be decreased by approximately 10% several times over several days while monitoring performance carefully. The processing weights should also be adjusted to favor the interactive LP. See "Processing weights" on page 110.

Use the **Customize/Delete Activation Profiles** task to open a reset profile to define processor running time. Select **Dynamically determined by the system** or **Determined by the user** on the Options page for the LP. See Figure 33 on page 132.

Workload manager LPAR CPU management of shared CPs

WLM's LPAR CPU Management component, together with the LPAR clustering technology of the z15, provides the ability to dynamically manage workloads within an LPAR cluster comprised of multiple logical z/OS images on a single z15. Each LP is assigned a *transaction* goal (desired response time) and an *importance* level. WLM monitors how well each LP is achieving its goals. A donor/receiver approach is utilized to reapportion CPU resources between LPs in the cluster. When WLM LPAR Weight CPU Management decides to change the weight of an LP, it adjusts the receiver LP and the donor LP by a percentage of the current weight of the receiver. WLM takes resources away from an LP that is over-achieving its target or has a workload that is less important (as defined by the installation). Any resource

given to a particular LP is taken away from another LP in the LPAR cluster. LPs whose workloads are of the same importance level should all have similar performance indexes (a measure of how closely the workload is meeting its defined goal).

One can think of the entire LPAR cluster as having a total processing weight. The total weight for an LPAR cluster is the sum of all the initial processing weights of all the LPs that have joined the cluster. As a new logical partition joins the cluster, its initial processing weight is added to that of the cluster. Though weights are adjusted from one LP to another, the total weight for the cluster is consistent. When an LP leaves the LPAR cluster, as when it is either system reset, deactivated, or re-IPLed, the initial processing weight, which it had been contributing to the LPAR cluster, is removed from the total weight available to the cluster. The weight removed from the cluster is not necessarily equal to the current weight for the exiting LP.

The optional minimum and maximum processing weights for an LP govern how much flexibility WLM has in adjusting weights from one LP in the cluster to another. The installation should assign a reasonably wide range of processing weights to each WLM managed LP. Assigning the same value for initial, minimum, and maximum weights effectively disables WLM LPAR CPU Management of processor weights.

Though logical cores of WLM managed LPs may need to be soft-capped (as for workload pricing, see "Workload charging by soft-capping to a defined capacity" on page 120), initial capping (traditional hardware capping) of these LPs is disallowed. Similarly, if an LP is WLM managed, its logical cores must be shared, not dedicated. For more information regarding internals of WLM CPU Management, see the IBM Redbook *z/OS Intelligent Resource Director*

Workload charging by soft-capping to a defined capacity

Workload charging introduces the capability to pay software license fees based on the size of the LP the product is running in, rather than on the total capacity of the CPC. The capability is enabled by the LPAR clustering technology of the z15 together with the License Manager component of z/OS. Each LP is assigned a ***defined capacity*** by the installation in terms of Millions of Service Units (MSUs).

WLM helps ensure that the rolling 4-hour average CPU utilization for the LP does not exceed this amount by tracking the CPU utilization for the logical partition. If the 4-hour average CPU consumption of the LP exceeds the defined capacity of the LP, WLM dynamically activates LP capping (soft-capping). When the rolling 4-hour average dips below the defined capacity, the soft-cap is removed.

WLM *will not* dynamically adjust the defined capacity for an LP. This is the responsibility of the installation. If an LP consistently exceeds its defined capacity, the license certificates and the defined capacity of the LP should be adjusted to reduce the amount of time the LP is soft-capped. If you have a configuration where the LP weights move significantly from one LP to another according to shift, then you must license the products in each LP at the highest capacity that will be used by that LP.

Defined capacity and the use of soft-capping by software applies to general purpose processors only. Initial Capping for general purpose processors is not supported for an LP that uses defined capacity and soft-capping. Initial capping for all other processor types is supported in such an LP.

Workload charging by soft-capping to a group capacity

Workload charging introduces the capability to pay software license fees based on the size of a group of the LPs the product is running in, rather than on the total capacity of the CPC. The capability is enabled by the LPAR clustering technology of the Z mainframe together with the License Manager component of z/OS. Each LP is assigned to a group with a **group capacity** by the installation in terms of Millions of Service Units (MSUs).

WLM helps ensure that the rolling 4-hour average CPU utilization for the group does not exceed this amount by tracking the CPU utilization for the group of logical partitions. If the 4-hour average CPU consumption of the group exceeds the group capacity of the group, WLM dynamically activates LP capping (soft-capping) in one or more of the members of the group. When the rolling 4-hour average dips below the group capacity, the soft-cap(s) is removed.

WLM *will not* dynamically adjust the group capacity for a group. This is the responsibility of the installation. If an LP or set of LPs in the group consistently exceeds its group capacity, the license certificates and the defined group capacity should be adjusted to reduce the amount of time the LPs are

soft-capped. If you have a configuration where the LP weights move significantly from one LP to another according to shift, then you must license the products in each LP at the highest capacity that will be used by that LP or set of LPs in the group.

Group capacity and the use of soft-capping by software applies to general purpose processors only. Initial capping for general purpose processors is not supported for an LP that uses group capacity and soft-capping. Initial capping for all other processor types is supported in such an LP.

Defined capacity and group capacity can be used together for a logical partition. WLM manages the logical partition accordingly, taking into account both definitions when present.

Recommendations on setting up an LPAR cluster

- An LPAR cluster is a collection of two or more logical partitions, on a particular CPC, that are part of the same parallel sysplex. LPAR clusters do not span CPCs as do parallel sysplexes. Though the member LPs of an LPAR cluster will all be in the same parallel sysplex, all members of a parallel sysplex might not be members of the same LPAR cluster. A given parallel sysplex can have member LPs that belong to multiple LPAR clusters, each on a different CPC.

- Identify logical partitions on the CPC that will be in the cluster (members of the same parallel sysplex). A single CPC can have several LPAR clusters just as a single CPC can have many LPs, each having membership in a different parallel sysplex.

- It is recommend to allocate shared CPs and enablement of WLM management for cluster members (see Note 1). The number of initially online CPs should be maximized to provide optimum flexibility to WLM. The number of reserved CPs defined should be the maximum allowed for an LP in your configuration minus the number of initially online CPs. See "Number of central processors" on page 104 for additional information on central processors.

- Establish an initial weight for each LP in the cluster. This will be the weight for the LP immediately after it is activated (see Note 2). Triple digit values should be used, wherever possible, for initial weights because WLM reapportions weights on a percentage basis. The total weight of the cluster will equal the sum of all the initial weights of its member LPs. Leave the minimum and maximum weights blank or make the range as wide as possible (optimally 1 to 999) to provide WLM maximum flexibility as it distributes CPU resource among the cluster members.

- Enable each LP in the cluster for WLM management.

- To enable DCM of managed channel paths for a logical partition, the name specified on the IOCLUSTER keyword for managed channel paths in the IOCDS must match the sysplex name of the software running in the logical partition. See "Dynamically managed CHPIDs" on page 52 for more information on the IOCLUSTER keyword.

- Calculation to estimate the number of cache structures that can be supported:

  The number of cache buffer data items that can be maintained locally in a logical partition is directly proportional to the number of online central storage pages in the LP. Each cache buffer or data item needs a local cache vector space bit. A heuristic value of 4 bits per online central storage 4K page is assigned by the system to each exploiting logical partition.

  The number of cache buffers supportable is easily calculated. Multiply the number of online central storage pages, in the z/OS logical partition exploiting the cache vector space, by two to get the number of cache vector bytes provided.

  For instance, if an exploiting logical partition has 32 GB of central storage online the amount of cache vector space provided would be 32 * 1024 (MB per GB) * 256 (pages per MB) * 2 (nybbles per byte) = number of bytes provided. For an LP with 32 GB 16777216 bytes, or 16 MB is provided. The size of the cache vector for an LP whose central storage definition includes the capability of using Dynamic Storage Reconfiguration will grow when the reserved storage amount is configured online.

Notes:

1. Logical partitions in a sysplex that have dedicated CPs can join a cluster but will not be enabled for WLM LPAR Weight and Vary CPU Management. They can derive I/O management benefits, however, from WLM's Dynamic CHPID Management capability.

2. Though the current weight of a logical partition can be changed (increased or decreased) by WLM once it joins an LPAR cluster, the initial weight is restored when (on IPL) it leaves the cluster. Software can then rejoin the same or a different cluster, again donating its initial weight to the sum available for the entire cluster.

Enabling management of Linux shared logical processors by WLM's LPAR CPU management component

1. For the target Linux LP, use the Options page (see Figure 41 on page 143) on the Customize Image Profiles window to specify the CP management cluster name of the intended LPAR cluster. Also select **Enable Workload Manager** on the Processor page (see Figure 37 on page 137) on the Customize Image Profiles window.

2. In the WLM policy statement, define a new service class giving it a velocity goal and an importance. Also, in the SYSH subsystem (create SYSH subsystem if one does not already exist), define a classification rule with the attribute **SY** (system name). Associate the system name of the target Linux logical partition and the service class you just defined.

3. IPL the Linux 2.4 kernel and provide a system name of the Linux logical partition by executing **insmod hwc_cpi.o system_name=xxxxxxx**, where **xxxxxxx** matches the system name communicated to WLM in Step 2.

Note: WLM is only able to manage Linux shared logical processors running on general purpose CPs. Management of Linux shared logical processors running on Integrated Facility for Linux (IFL) processors is not supported. The **Linux** mode partitions participating in a CP management cluster should be system reset through the Support Element and Hardware Management Console following a Linux shutdown command to help ensure accurate cleanup by WLM's LPAR CPU Management component.

Defining shared channel paths

Before defining shared channel paths, consider the following:

- Only CTC, CNC, ICP, OSC, OSD, OSE, OSM, OSN, OSX, FC, FCP, IQD, CS5, and CL5 channel paths can be shared. CVC and CBY channel paths *cannot* be shared. On CF only models ICP and channel paths *cannot* be shared.

- A failure in a shared channel path or I/O device can affect more than one LP; therefore, critical I/O devices (for example, DASD containing vital data sets) still need multiple channel paths. You can provide multiple shared channel paths (up to 8) to critical I/O devices.

- Using shared channel paths does not reduce the number of logical paths needed at a control unit. A control unit requires a logical path for each active LP that can access I/O devices through a shared channel path.

There are three possible channel path modes:

shared
A channel path that can be configured online to one or more LPs at the same time. One or more LPs can access I/O devices at the same time using this channel path. Spanned channel paths are shared by LPs in multiple logical channel subsystems (CSSs). Unspanned channel paths can only be shared by LPs in the same CSS.

reconfigurable
An unshared channel path you can reconfigure offline from one LP, then online to another, within the same CSS. Only one LP can access I/O devices on this channel path at a time.

dedicated
An unshared and non-reconfigurable channel path. Only one LP can access I/O devices on this channel path.

You cannot mix unshared and shared CHPIDs to a device.

Channel path access and candidate lists

If a channel path is either shared or reconfigurable, you can specify which LPs have access to that channel path. Use the channel path access list with or without the channel path candidate list.

Channel Path Access List

An LP has ***initial*** access to a channel path if the LP is on that channel path's access list.

For the first power-on reset with an LPAR IOCDS, the access list defines which LPs will initially have the channel path configured online. Reconfigurable and shared CHPIDs may later be accessed by LPs not in the access list. Subsequent power-on resets with the same LPAR IOCDS will have any reconfigurations applied and the LPs may have access to channel paths other than those that were specified in the initial access lists for the channel paths. See "Channel path reconfiguration and logical partition activation" on page 128 for more information on assigning channel paths to an LP.

Channel Path Candidate List

An LP can ***gain*** access to a channel path if the LP is on that channel path's candidate list. An LP is allowed to configure a channel path online if the LP is in that channel path's candidate list.

I/O device candidate list

The I/O device candidate list specifies the LPs which can access the device. You can use the I/O device candidate list to restrict LP access to I/O devices on shared channel paths. If you do not specify an I/O device candidate list, all LPs that share the channel paths, to which the device is attached, can have access to the device. For coupling facility devices, the device candidate list is not supported.

Procedure for defining shared channel paths

To share channel paths, use the following general procedure:

1. Select which I/O devices to access through shared channel paths

2. Select the LPs that will share channel paths

   - Specify the desired channel paths as shared

   - Use the access list and candidate list of the desired shared channel path to select which LPs can access that shared channel path.

3. For each I/O device that will be on a shared channel path, you can use the I/O device's candidate list to restrict which LPs can access that I/O device.

4. Make physical connections to the shared channel paths and the I/O devices.

5. Update the software I/O configuration for the control programs running in LPs that can access devices through shared channels.

6. Use IOCP or HCD to create an IOCDS that defines the I/O configuration, including shared channel paths, to the CPC channel subsystem. The channel subsystem includes all channel paths, control units, and I/O devices accessible by all LPs.

Communicating by means of FICON CTC

You can connect shared or unshared CNC channel paths to shared or unshared CTC channel paths for the purpose of CTC communication. The connected CNC and CTC channel paths can be on the same CPC (to communicate between the LPs on that CPC) or on different CPCs.

You can connect shared or unshared FC channel paths to shared or unshared FC channel paths for the purpose of CTC communication. The connected FC channel paths can be on the same CPC (to communicate between the LPs on that CPC) or on different CPCs.

CTC communications involving a shared channel path require the specification of control unit logical addresses. The control unit logical address identifies the MIF image ID number for the LP to which the channel path is to communicate. If the remote channel path is not shared, the logical address must be zero or not specified. If the remote channel path is shared, the logical address must equal the desired MIF image ID number for the LP that the shared channel path can access. If the remote channel path is a shared FC channel path and the target LP has a CSS ID other than zero, then the logical address must equal the combination of the desired CSS ID and the MIF image ID for the LP with which you want to communicate. For example, if the remote channel path is shared within CSS 1 and you want to communicate with the LP that has MIF image ID 5, specify logical address 15. You must define a control unit and a control unit logical address for every LP that you want to communicate with.

Dynamic CHPID management (DCM) considerations

DCM CHPIDs used by Workload Manager (WLM) to optimize I/O throughput across an LPAR cluster are identified in the IOCDS by specifying CHPARM equals 01 and IOCLUSTER equals name (where name is an 8-byte EBCDIC cluster identifier).

All DCM CHPIDs are inherently shareable by all LPs but reserved for use by the WLM enabled members of the specified cluster. At completion of an LP activation, its DCM CHPIDs will be in a deconfigured state. When an LP joins a cluster, the DCM CHPIDs in that cluster become available for use by WLM but are not brought online until the need for greater throughput arises. System reset of an LP that was a member of a cluster causes each of the DCM CHPIDs that were online to the LP to be deconfigured. For information on how to define DCM CHPIDs, see "IOCP statements for ICP" on page 52. For allocation rationale of DCM CHPIDs, see redbook *z/OS Intelligent Resource Director*, SG24-5952.

I/O priority recommendations

Channel subsystem I/O priority queuing is used by z/OS WLM component to dynamically manage the priority of I/O operations for given workloads based on the performance goals for these workloads as specified in the WLM policy.

Channel subsystem I/O priority queuing is used by z/VM to manage the priority of I/O operations performed by guests. The VM Resource Manager adjusts guest I/O priorities based on the performance goals of the associated workloads. It is recommended that you establish a range of I/O priorities for z/VM logical partitions that is sufficient to enable effective discrimination among guest I/O requests of different importance.

In order to provide WLM the greatest flexibility in managing I/O requests across members of an LPAR cluster, it is highly recommended that you establish the same range of priorities for each member of an LPAR cluster. A range of eight priorities (from minimum to maximum) is optimum. If a range greater than eight is specified, only the top eight will be utilized by WLM.

Non-WLM managed LPs should be assigned I/O priorities according to their importance relative to the LPs that are members of LPAR clusters. Unless it is running z/VM (in which case the recommendations above should be followed), a non-WLM managed LP should be assigned equal values for minimum and maximum I/O priorities.

Security-related controls

You can define security-related controls for a logical partition.

Global performance data control authority

This control limits the ability of a LP to view CP activity data for other LPs. Logical partitions with control authority for global performance data can view CP utilization data and Input/Output Processor (IOP) busy data for all of the LPs in the configuration. Additionally, gathering of channel measurements requires selection of this parameter.

Note: see "Enabling HiperDispatch for z/OS Logical Partitions" on page 117 for considerations for Hiperdispatch and global performance data.

With the exception of an LP that is a member of a WLM Cluster, an LP without control authority for the performance data can view only the CP utilization data for that LP.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to enable global performance data control for an LP. The Global performance data control selection is located on the Security page for the LP.

I/O configuration control authority

This control can limit the ability of the LP to read or write any IOCDS in the configuration locally or remotely. LPs with control authority for the I/O configuration data can read and write any IOCDS in the configuration, and can change the I/O configuration dynamically.

Additionally, this control allows the OSA Support Facility to control OSA configuration for other LPs and allows access to certain STP data.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to enable I/O configuration control for an LP. The Input/output (I/O) configuration control selection is located on the Security page for the LP.

Cross-partition authority

This control can limit the capability of the LP to issue certain control program instructions that affect other LPs. LPs with cross-partition authority can issue instructions to perform a system reset of another LP, deactivate any other LP, and provide support for the automatic reconfiguration facility.

The automatic reconfiguration facility permits a backup LP to deactivate a primary LP if a problem is detected in the primary LP. The backup LP can then configure online, storage resources that become available when the primary LP is deactivated. See "CPCs with the Sysplex Failure Manager (SFM)" on page 102.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to enable cross-partition authority for an LP. The **Cross partition authority** selection is located on the Security page for the LP.

Logical partition isolation

This control reserves reconfigurable unshared channel paths for the exclusive use of an LP. Channel paths assigned to an isolated LP are not available to other LPs and remain reserved for that LP when they are configured offline.

Use the **Release** task to release an unshared channel path from an isolated LP.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to enable isolation for an LP. The Logical partition isolation selection is located on the Security page for the LP.

Using IOCP, you can control access to channel paths using the channel path candidate list. Access to I/O devices on shared channel path can be further controlled through the I/O device candidate list.

Enable the partition to receive commands from other partitions

This control enables the selected active logical partition to receive BCPii commands from other active logical partitions. You can select either:

- **All partitions** if you want the selected logical partitions to receive BCPii commands from all active logical partitions.
- **Selected partitions** if you want to remove or add selected logical partitions to receive BCPii commands from the logical partition.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to enable the logical partition to receive commands. The **Enable the partition to receive commands from other partitions** selection is located on the Security page for the LP.

Enable the partition to send commands

This control enables the selected active logical partition to send BCPii commands to other active logical partitions.

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile to enable the logical partition to send commands. The **Enable the partition to send commands** selection is located on the Security page for the LP.

Basic counter set

The basic counter set includes counts of central processing unit cycles, instructions executed, and directory-write and penalty cycles for level-1 instruction and data caches. For more information about the basic counter set, see *The Set-Program-Parameter and CPU-Measurement Facilities*, SA23-2260.

Use the **Customize/Delete Activation Profiles** task to authorize the use of the basic counter set for an LP. The **Basic counter set authorization control** selection is located on the Security page for the LP.

Problem state counter set

The problem state counter set includes counts of central processing unit cycles, instructions executed, and directory-write and penalty cycles for level-1 instruction and data caches only when the processor is in problem state. For more information about the problem state counter set, see *The Set-Program-Parameter and CPU-Measurement Facilities*, SA23-2260.

Use the **Customize/Delete Activation Profiles** task to authorize the use of the problem state counter set for an LP. The **Problem state counter set authorization control** selection is located on the Security page for the LP.

Crypto activity counter set

The crypto activity counter set can be used to identify the crypto activities contributed by the logical CPU. It includes counters related to PRNG, SHA, DEA, AES, and ECC functions. For more information about the crypto activity counter set, see *The Set-Program-Parameter and CPU-Measurement Facilities*, SA23-2260.

Use the **Customize/Delete Activation Profiles** task to authorize the use of the crypto activity counter set for an LP. The **Crypto activity counter set authorization control** selection is located on the Security page for the LP.

Extended counter set

The extended counters provide information about hardware facilities and structures that are specific to a machine family. The extended counters are designed to expand upon information provided by the basic counter set. For more information about the extended counter set, see *The Set-Program-Parameter and CPU-Measurement Facilities*, SA23-2260.

Use the **Customize/Delete Activation Profiles** task to authorize the use of the extended counter set for an LP. The **Extended counter set authorization control** selection is located on the Security page for the LP.

Basic sampling

With basic sampling, samples are taken and stored at the end of each sampling interval.

Use the **Customize/Delete Activation Profiles** task to authorize the use of basic sampling for an LP. The **Basic sampling authorization control** selection is located on the Security page for the LP.

Diagnostic sampling

With diagnostic sampling, samples are taken and stored at the end of each sampling interval.

Use the **Customize/Delete Activation Profiles** task to authorize the use of diagnostic sampling for an LP. The **Diagnostic sampling authorization control** selection is located on the Security page for the LP.

CPACF key management operations

This control enables the CP Assist Cryptographic Functions (CPACF) key management operations to permit AES, DEA, and ECC key import functions.

Use the **Customize/Delete Activation Profiles** task to change the key import functions setting for CPACF when the logical partition is activated. The **CPACF key management operations control** selection is located on the Security page for the LP.

Dynamic I/O configuration

Dynamic I/O configuration, available with z/OS and z/VM, provides the capability of changing the currently active I/O configuration. Using dynamic I/O configuration logical partitions, channel paths, control units, devices, and PCIe functions of the currently active I/O configuration can be added, deleted, or modified without requiring a power-on reset and an IPL for the change to take effect. Changes made to the

currently active I/O configuration can be saved and the IOCDS that reflects these changes can be written and made the active IOCDS.

Dynamic I/O configuration does *not* support the following:

- Changing MIF image ID numbers

Use the **Customize/Delete Activation Profiles** task to open a reset profile to enable dynamic I/O configuration for the CPC.

- On the General page, select a dynamic IOCDS from the IOCDS list
- Then select the Dynamic page for the IOCDS and select the **Allow dynamic changes to the channel subsystem input/output (I/O) definition** check box.

Managing dynamic I/O configuration

For detailed information about changing the I/O configuration dynamically, refer to *z/OS Hardware Configuration Definition Planning*, GA22-7525, *z/VM CP Planning and Administration*, SC24-6271, and *z/VM I/O Configuration*, SC24-6291.

Planning and operation considerations

Guidelines for planning and operating in a dynamic I/O environment are detailed in the appropriate z/OS, and z/VM publications. The planning and operation considerations described in this section are additional guidelines that apply to a dynamic I/O environment.

You should plan carefully to avoid confusion when moving CHPIDs and eliminate duplicate device situations should a backout be required when a power-on reset is performed.

- Prepare for a backout situation.

  Before changing the currently active I/O configuration, prepare for a possible backout situation. Record the current CHPID assignments and the planned I/O changes.

  It is important to prepare for a backout situation. A backout situation occurs when changes are made to the I/O configuration but the changes are not saved prior to performing a power-on reset. If the changes are not saved, the CHPID assignments prior to the first dynamic I/O change takes effect.

  In addition CHPID assignments after a backout will also reflect any changes made by hardware operator tasks or control program commands.

- Avoid CHPID reconfigurations concurrent with dynamic I/O changes

  Do not perform hardware operator tasks and system control program commands to cause a CHPID to be moved from one LP to another or to give or take access to a shared channel path while dynamic I/O changes are being made. Use these commands only after dynamic I/O changes are saved.

Dynamic activation of I/O configurations for stand-alone Coupling Facilities

Coupling Facilities (CFs) provide locking, caching, and list services between coupling-capable z/OS® processors, and are a significant component of highly available Parallel Sysplex configurations. Dynamic I/O for Standalone Coupling Facility enables dynamic activation of a new or changed IODF on a standalone coupling facility CPC without requiring a re-IML or power-on reset (POR). Stand-alone CF servers can seamlessly make hardware-only dynamic I/O configuration changes on behalf of the CF partitions that reside there without requiring a disruptive reset. This capability both improves client workload availability and minimizes the risks associated with relocation of CF structures.

This capability requires Version 2.14.1 or later firmware support on the coupling facility CPC as well as the CPC where the HCD system is running. If you are planning to use the Dynamic I/O for Standalone Coupling Facility capability on a CPC, you **must** use HCD to configure your IODF/IOCDS appropriately for that CPC. Then, IML the Coupling Facility CPC with that IOCDS in order to use the Dynamic I/O Standalone Coupling Facility capability for future dynamic IO operations. No IODF/IOCD updates are required on the CPC where the HCD is running. For more information, see *z/OS HCD User's Guide*, SC34-2669.

Dynamic activation of I/O configurations requires an activation service on the stand-alone coupling facility server. This activation service is FW running in firmware partitions. This firmware partition must be

defined as one of the reserved partitions, partition B in the highest supported logical channel subsystem. The name of the partition must be MCS_1. With z15, this partition is automatically defined. The partition does not support any attached I/O and is automatically managed by the FW.

Assigning channel paths to a logical partition

Channel paths are defined in the IOCDS. Channel paths that are specified as reconfigurable can be moved among LPs.

Channel paths assigned in the IOCDS to **General** LPs can be shared, reconfigurable, or dedicated.

Channel paths that are specified as shared can be accessed by one or more LPs at the same time. Unshared channel paths that are defined to a LP as not reconfigurable are dedicated to that LP.

Coupling facility logical partitions

Channel paths (CE LR, ICA SR coupling links, and ICP) assigned in the IOCDS to coupling facility LPs can be online to only one coupling facility at a time. A coupling facility can be assigned CE LR, ICA SR coupling links, or ICP channel paths.

Note: The CE LR, ICA SR coupling links, and ICP channel paths that are online to a single coupling facility are shareable by multiple non-coupling facility logical partitions.

Channel path reconfiguration and logical partition activation

When the **Configure On/Off** task or an equivalent system control program command is run successfully for a channel path in an *active* LP, the channel path is configured to the LP at that time.

When a successful **Configure On/Off** task completes for a channel path in an LP that is *not active*, the channel path is not actually configured online to the LP at that time. Rather, the channel path is *targeted* to be configured online to that LP when the LP is activated.

When an LP is deactivated, all shared channel paths configured to it at the time of deactivation are targeted to be configured online to it when it is subsequently activated. Unshared channels that were last configured to this LP and were not yet reconfigured to another LP are also targeted to be configured online to this LP at its next activation. However, the targeting of a channel path may change prior to the next activation due to channel reconfiguration commands, dynamic I/O configuration changes, or POR.

Channel paths can also be targeted to be configured online to an LP by using dynamic I/O configuration. See "Dynamic I/O configuration effects on channel path reconfiguration" on page 128 for more details.

PR/SM manages lists of targeted channel paths for each LP on an IOCDS basis so that all channel paths that are targeted to be configured online to an LP will be automatically configured online when that LP is activated. Exceptions to this rule are:

- Targeted CE LR, ICA SR coupling links, and ICP channel paths online to another coupling facility are not automatically configured online when that LP is activated.

  Note: The CE LR, ICA SR coupling links, and ICP channel paths can only be online to one active coupling facility LP at a time. If such a channel path is targeted to a coupling facility LP but is already online to another coupling facility LP, then it will be removed (deconfigured) from the activating LP.

- The targeted channel path is in single channel service mode or otherwise broken.

Dynamic I/O configuration effects on channel path reconfiguration

If a channel path is dynamically added to the configuration using dynamic I/O configuration, all the LPs in the channel path access list (for a shared channel path) or the one LP in the channel path access list (for an unshared channel path) are targeted to have this new channel path configured online. The dynamic I/O configuration change does not bring the channel path online.

The channel path is configured online to the targeted LP when one of the following occurs:

- The system control program running in the targeted LP issues the appropriate reconfiguration command to configure the channel path online. For z/OS, this would be:

CF CHP(*nn***),ONLINE**

For z/VM, this would be:

VARY ON CHPID nn

For Linux on Z Systems, this would be:

echo on > /sys/devices/css0/chp0.nn/status

- The **Configure On/Off** task is used to configure the channel path online to the targeted LP while the LP is active.
- A power-on-reset is done with the IOCDS that has the new dynamic changes defined in it without the changes being made active using HCD z/OS, or z/VM commands. Following a power-on-reset, activation of the targeted LP will configure the channel path online.

Automatic channel path reconfiguration

PR/SM records shared channel path configurations and unshared channel path reconfigurations and uses the information to modify the initial targeting of channel paths that are defined in the IOCDS. This information is maintained on an IOCDS basis.

When a particular IOCDS is used in a POR for the first time after it has been written, the definitions in that IOCDS are used to determine the assignment of channel paths to LPs according to the channel path access lists that are defined. All previous information about channel configuration associated with this IOCDS is discarded. The exception to this rule is when a newly written IOCDS is first used as part of a dynamic I/O change to the system. (For example, the new IOCDS is used as a result of a "Switch IOCDS for Next POR" action by HCD, or the new IOCDS is the target of the ACTIOCDS= parameter of the z/OS ACTIVATE command.) When a new IOCDS is used in this manner, the current state of channel configurations is preserved and immediately associated with the newly written IOCDS.

Over time, the list of channel paths targeted to be configured online to a given LP can be changed by system control program configuration commands, configure on tasks, or dynamic I/O configuration commands issued through HCD or z/VM. Similarly, reconfigurable unshared channel paths can be moved from one LP to another using the same commands; changing the owner of the unshared channel path. For activated coupling facility LPs, you can change the channel paths targeted to be configured online using coupling facility control code commands. Automatic channel path reconfiguration restores all of the latest changes for each POR with the IOCDS automatically.

Automatic channel path reconfiguration does ***not*** preserve the online/offline state of unshared channel paths (reconfigurable or dedicated). Rather, at POR time, all unshared channel paths are targeted to come online to the LP that last owned it. For dedicated channel paths, this owner never changes but for reconfigurable channel paths the owner can change and is remembered.

Following a POR, a channel path that is targeted to come online to a LP will be physically online to that LP and usable at the completion of activation for the LP.

Automatic load for a logical partition

Select the **Load during activation** check box if you want the control program to be loaded automatically each time the LP is activated. Use the **Customize/Delete Activation Profiles** task to open a image profile. The **Load during activation** selection is located on the Load page for the LP. See .

Coupling Facility Logical Partitions
Do not support automatic load because the coupling facility control code is automatically loaded and made operational at LP activation. No IPL of an operating system is necessary.

z/OS LPs
Specify parameter (PARM) information if desired. Byte 2 and byte 3 of the PARM field are used to select an IOCONFIG member for this IPL if you do not want the default (00).

Defining logical partitions

Before using this section you should first read "Determining the characteristics" on page 92. This section describes the windows, parameters, and tasks, you can use to define LP definitions.

Sample tasks and windows explained in this section reference tasks and windows available from the Support Element console. The Hardware Management Console (HMC) and Support Element (SE) information can be found on the console help system.

LP definitions are saved across a power-on reset and are used during each power-on reset. You can use reset, image, and load profiles to modify LP definitions. Use image and load profiles to modify LP definitions after the CPC has been activated. Use the **Customize/Delete Activation Profiles** task to open a reset, image, and load profiles.

You can change reset, image, and load profiles at any time. However, some of these definition parameters cannot affect the running system if the affected LP is currently activated. See "Changing logical partition definitions" on page 159.

LP definition parameters fall into the following categories:

Global reset profile definitions

- Enable global input/output I/O priority queuing
- Automatic I/O interface reset
- Processor running time
- LP automatic activation order

General

- Logical partition identifier
- LP mode of operation
- Enable logical partition time offset

Processor characteristics

- Dedicated and Not dedicated CPs, ICFs, zIIPs, and IFLs
- Initial, minimum, and maximum weight
- Absolute capping
- Number of initial and reserved processors

Security characteristics

- Global performance data control
- Input/output configuration control
- Cross partition authority
- Logical partition isolation
- BCPii permissions
- Counter facility security options
- Sampling security options
- CPACF key management operations

Storage

- Central storage
- Virtual flash memory

Secure Service Container

- Boot Selection

- Master user ID
- Master password
- Network adapters
- DNS servers

Time Offset

- Time offset

Load information

- Load during activation
- Load type
- Load address
- Use of dynamically changed address
- Use of dynamically changed parameter
- Load parameter
- Time-out value

Cryptographic characteristics

- Assigned domains
- Assigned cryptos

Note: Coupling Facility partitions cannot take on these characteristics.

Options

- Minimum and maximum input/output (I/O) priority queuing values
- Defined capacity

Parameter descriptions

Enable global input/output (I/O) priority queuing
Select this option to enable I/O priority queuing. Selecting this option causes I/O requests to be prioritized according to the values specified on the Options page of the Image profile.

Automatic input/output (I/O) interface reset
Enables automatic system reset of a logical partition on behalf of a software initiated recovery action.

Dynamically determined by the system
Select this option to allow the system to dynamically calculate the length of CP time dispatched. The default is selected for this check box.

Processor running time

Determined by the user
Select this option to manually set the length of CP time dispatched.

Running time
Selecting the **Determined by the user option** requires that you enter a running time amount from 1 - 100 milliseconds.

Do not end the timeslice if a partition enters a wait state
When the processor running time is determined by the user through this profile, check this box to indicate whether logical processors lose their share of running time when their logical partition enters a wait state. This check box only gets enabled (not grayed out) if the user first sets his own (**Determined by the user**) processor running time. We are NOT removing processor running time, nor the selections, only this selection when **Determined by the user** is selected.

Set power saving
Select the energy management power saving option for the CPC upon performing the power-on reset:

Custom energy management
Select this option to use the current power saving settings.
Emergency high performance
Select this option to use the high performance setting with no power saving.

Global reset profile definitions

Use the **Customize/Delete Activation Profiles** task to open a reset profile.

Options page definitions

Open the Options page to define the following LP characteristics:

• Enable global input/output (I/O) priority queuing

• Automatic I/O interface reset

• Processor running time
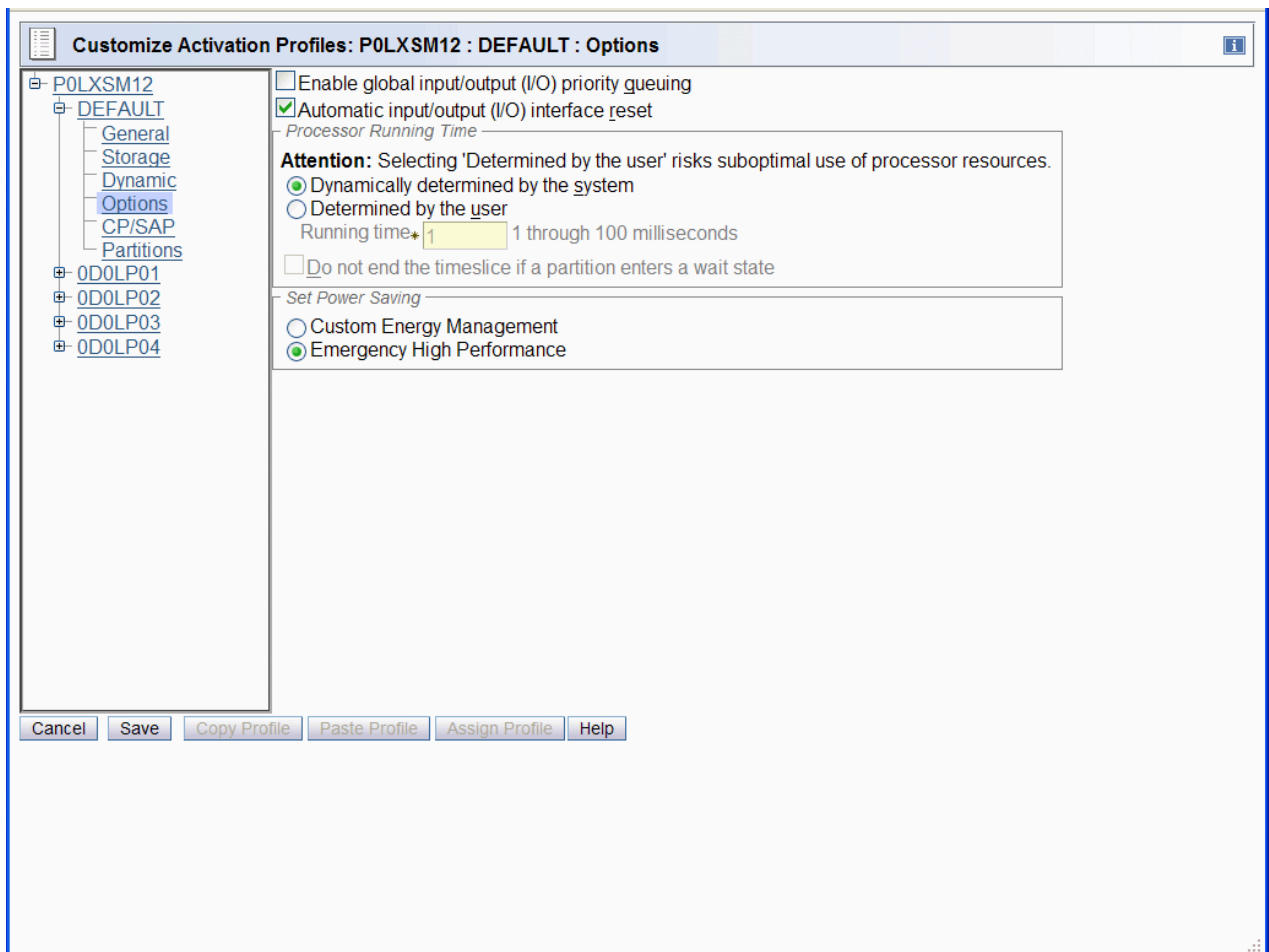
Figure 33. Options page, reset profile

Partitions page definitions

Use the **Customize/Delete Activation Profiles** task to open a reset profile. Open the Partitions page to define the following LP characteristics:

• LP automatic activation order
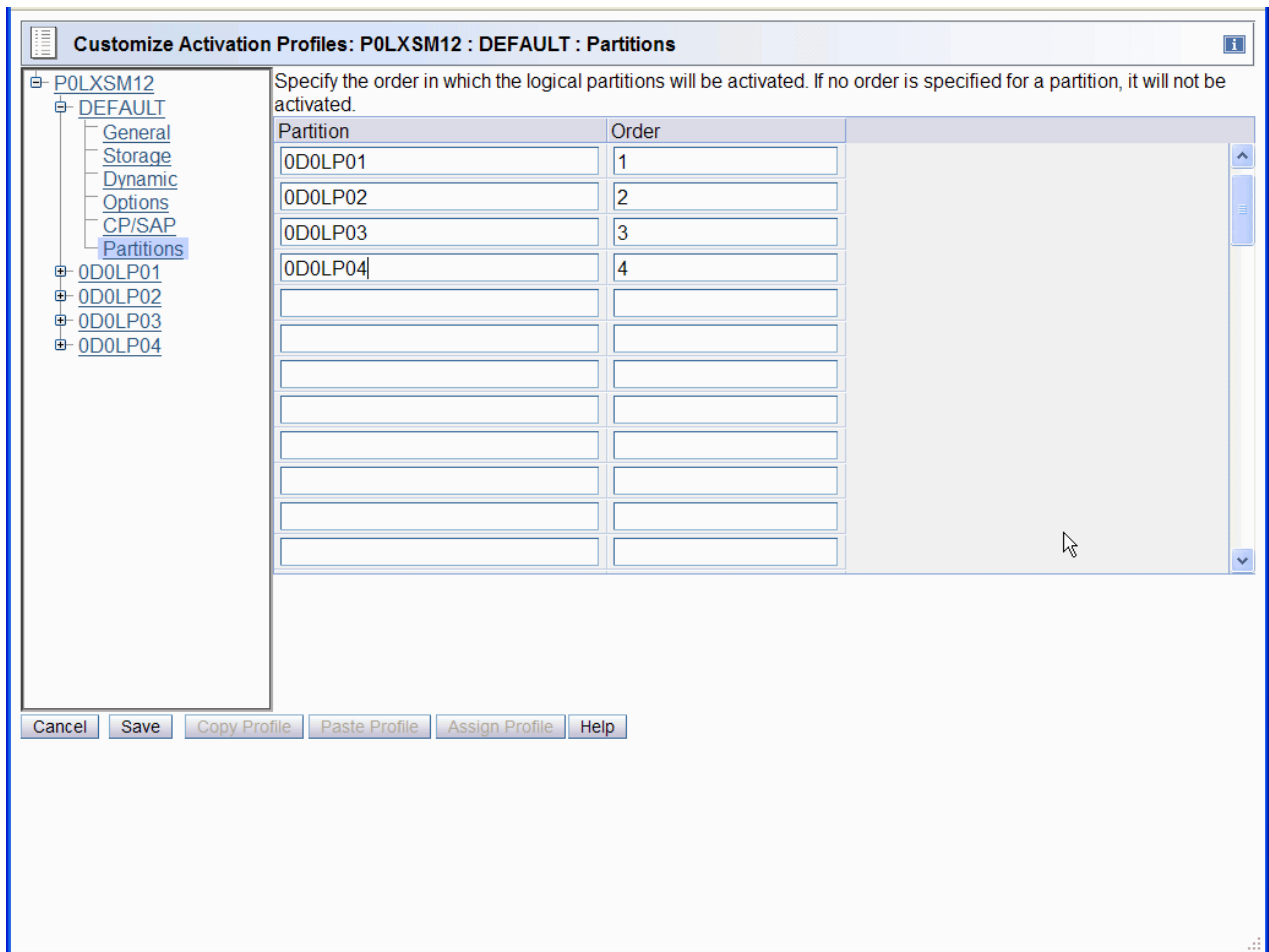
Figure 34. Partitions page, reset profile

Parameter descriptions

Partition
 LP name.

Order
 Enter a number indicating when the LP will be activated in the automatic activation order.

General

Use the **Customize/Delete Activation Profiles** task to open an image profile for an LP. Open the General page to define the following LP characteristics:

- Logical partition identifier
- LP mode of operation
- Clock type assignment

Figure 35. General page, image profile with **SSC** *mode selected*

Parameter descriptions

Partition identifier

Enter a hex value (X'00' through X'7F') for the LP. This parameter identifies the LP and is used as the third and fourth hexadecimal digit of the operand stored by the Store CPU ID Instruction for each logical core in the LP. The partition identifier must be unique for each active LP.

Mode

Select an LP operating mode from this scrollable list - **General**, **Linux-Only**, **SSC**, **z/VM**, or **Coupling facility** mode.

Clock type assignment

Select a time source for setting the logical partition's time-of-day (TOD) clock:

Standard time of day

Select this option to set the logical partition's clock to the same time set for the CPC's time source (either the CPC TOD clock or an external time reference, such as the STP).

Logical partition time offset

Select this option to set the logical partition's clock using an offset from the time of day supplied by its time source. Then use the Time Offset window to set the offset.

Ensure that the image profile data conforms to the current maximum LICCC configuration

Select this option to ensure that the image profile data conforms to the current maximum Licensed Internal Code Configuration Control (LICCC) configuration.
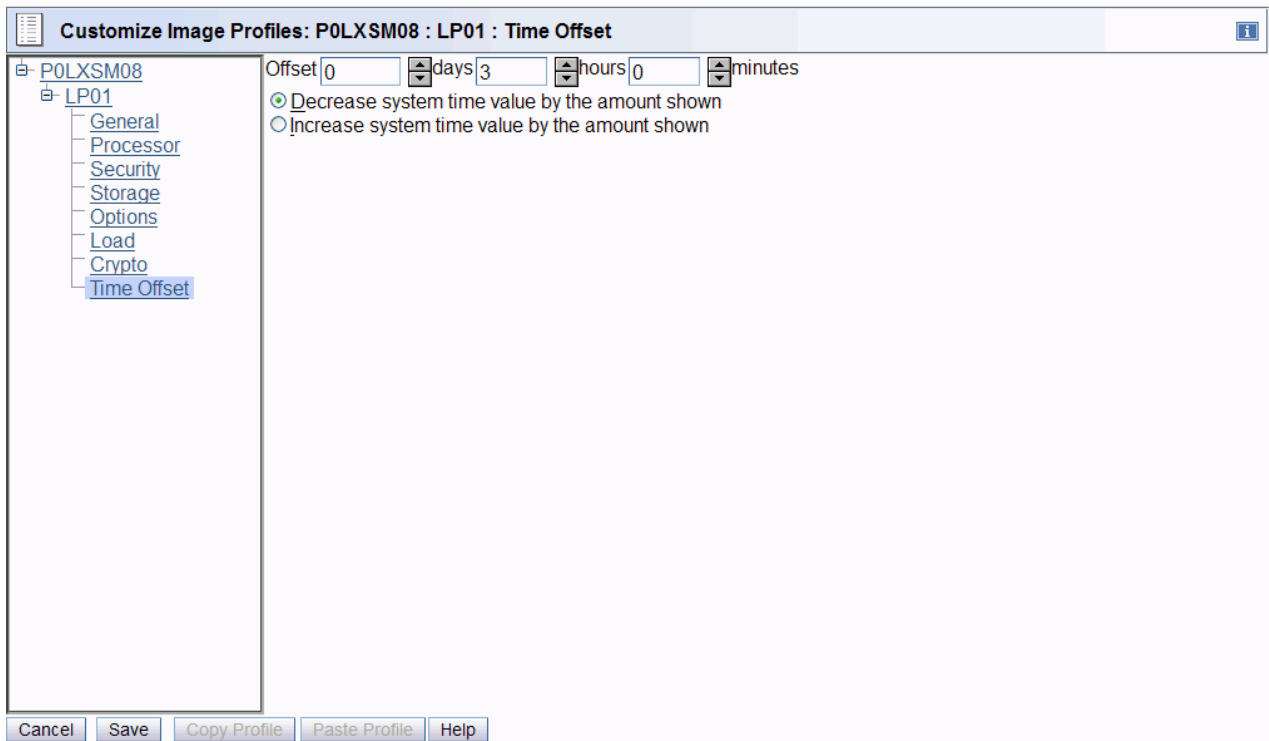
Figure 36. Time offset, image profile

Parameter descriptions

Offset

Type or spin to the number of days, hours, and minutes you want to set as the offset from the time of day supplied by its time source. You can set an offset within the following range:

- 0 to 999 days
- 0 to 23 hours
- 0, 15, 30, or 45 minutes

Decrease system time value by the amount shown

Select this choice to set the logical partition's clock back from the time of day supplied by its time source by the number of days, hours, and minutes in the offset. Use this setting to provide a local time zone WEST of UTC.

Increase system time value by the amount shown

Select this choice to set the logical partition's clock ahead of the time of day supplied by its time source by the number of days, hours, and minutes in the offset. Use this setting to provide a local time zone EAST of UTC or a date and time in the future.

Processor Characteristics

Table 16 on page 136 shows all logical partitions, required characterized PUs, and operating systems, and which PU characterizations can be configured to a logical partition image. The available combinations of dedicated (DED) and shared (SHR) processors are also shown. For all combinations, a logical partition can also have reserved processors defined, allowing nondisruptive logical partition upgrades.

Table 16. LP mode and PU usage			
Logical partition mode	**PU type**	**Operating systems**	**PUs usage**
General	CPs	z/Architecture operating systems	CPs DED *or* CPs SHR
	CPs or zIIPs	z/OS	CPs DED *and/or* zIIPs DED *or* CPs SHR *and/or* zIIPs SHR
General	CPs	z/TPF	CPs DED *or* CPs SHR
	CPs	z/VSE	CPs DED *or* CPs SHR
Coupling facility	ICFs *or* CPs	CFCC	ICFs DED *or* ICFs SHR, *or* CPs DED *or* CPs SHR
Linux only	IFLs or CPs	Linux z/VM	IFLs DED *or* IFLs SHR, *or* CPs DED *or* CPs SHR
z/VM	CPs, IFLs, zIIPs ICFs	z/VM	All PUs must be either SHR *or* DED
SSC	CPs, IFLs	SSC	IFLs DED or IFLs SHR, or CPs DED or CPs SHR

Processor page definitions

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile for an LP. Open the Processor page to define the following LP characteristics:

- Dedicated or shared general purpose CPs
- Dedicated or shared internal coupling facility (ICF) CPs
- Number of initial and reserved processors (number of processors is the number of cores when a logical partition uses SMT)
- Initial processing weight
- Initial weight capping
- Workload manager enablement
- Minimum processing weight
- Maximum processing weight
- Absolute capping

Figure 37. General mode logical partition with shared CPs and zIIPs

Figure 38. Customization for a Linux-only mode logical partition with shared Integrated Facilities for Linux (IFLs). There can be both an initial and reserved specification for the IFLs.

Figure 39. Customization for a coupling facility mode logical partition with shared central processors. There can be both an initial and reserved specification for the Central Processors.

Parameter descriptions

Note: Depending on the processor page, (see Figure 37 on page 137, Figure 38 on page 137 and Figure 39 on page 138) some of the following parameters may not be present.

Group Name
If you choose to assign the logical partition (or image) to a group, select a defined group from the list.

Dedicated processors
Select this option if you want to select all processors to be dedicated when the LP is activated. You can then specify the number of initial and reserved processors for each.

Not dedicated processor details
Select the processor type to display details such as Initial processing weight, Initial capping, and Enable workload manager.

Dedicated central processors
Select this option if you want the general purpose CPs that are allocated for the LP to be dedicated when the LP is activated.

Not dedicated central processors
Select this option if you want the general purpose CPs that are allocated for the LP to be shared when the LP is activated.

Dedicated integrated facility for Linux
If Integrated Facility for Linux (IFL) is supported and installed in the Central Processor Complex (CPC), select **Dedicated integrated facility for Linux** if you want an IFL processor dedicated to each logical processor.

Not dedicated integrated facility for Linux
If you want the logical processors to share not dedicated integrated facility for Linux (Integrated Facility for Linux (IFL) processors that are not already dedicated to other activated logical partitions when this logical partition is activated), select **Not dedicated integrated facility for Linux**.

Dedicated internal coupling facility processors
If internal coupling facility processors are supported by and installed in the Central Processor Complex (CPC), select **Dedicated internal coupling facility processors** if you want one dedicated to each logical processor.

Not dedicated internal coupling facility processors

Note: All processors assigned to a coupling facility partition should be dedicated to that logical partition if it is used for primary production workload.

If you want the logical processors to share not dedicated internal coupling facility processors (internal coupling facility processors that are not already dedicated to other activated logical partitions when this logical partition is activated), select **Not dedicated internal coupling facility processors**.

Initial processing weight
Enter a value between 1 - 999 to set the processing weight for the type of processor for an LP. The default value is 10.

Initial capping
Select this option to cap the CP resources for the type of processor for an LP. Capping has no effect on LPs with dedicated CPs.

Enable Workload Manager
Select this option so that CPU and I/O resources can be managed by WLM using IRD clustering technology.

Minimum processing weight
Select this option to establish a minimum processing weight that WLM will allocate to the LP. Do not specify a value here unless you determine a true need for it in your configuration. Specifying a value here can needlessly constrain what WLM can do to optimize the management of your workload.

Maximum processing weight
Select this option to establish a maximum processing weight that WLM will allocate to this LP. Do not specify a value here unless you determine a true need for it in your configuration. Specifying a value here can needlessly constrain what WLM can do to optimize the management of your workload.

Absolute capping
Select this optional fixed cap on the partition's shared logical processors of this processor type. This is specified in processor units (cores).

Security characteristics

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile for an LP. Open the Security page to define the following LP characteristics:

• Partition Security Options

• BCPii Permissions

• Counter Facility Security Options

• Sampling Facility Security Options

• CPACF Key Management Operations

Figure 40. Security page, image profile

Security parameter descriptions

The following logical partition reset or image profile security options can be defined.

Partition security options:

Global performance data control
Select this option to allow the LP to view the CPU utilization data and the Input/Output Processor (IOP) data for all LPs in the configuration. Not selecting this option only allows the LP to view its own CPU utilization data. Additionally, gathering of FICON channel measurements requires selection of this parameter. The default is selected.

Note: An LP running a level of RMF that supports FICON requires control authority even if no FICON is installed.

Input/output (I/O) configuration control
Select this option to allow the LP to read or write any IOCDS in the configuration and to make dynamic I/O changes. Additionally, this parameter allows the OSA Support Facility for z/OS, z/VM, and z/VSE to control OSA configuration for other LPs. Access to certain STP data is also managed by this option. The default is selected. If a z/VM guest image is managed as a virtual server in an ensemble, you must enable the **Input/Output (I/O) configuration control** option.

Cross partition authority
Select this option to allow the LP to issue control program commands that affect other LPs; for example, perform a system reset of another LP, deactivate an LP, or provide support for the automatic reconfiguration facility. The default is not selected.

Logical partition isolation
Select this option to reserve unshared reconfigurable channel paths for the exclusive use of the LP. The default is not selected.

BCPii Permissions:

Enable the partition to send commands
Select this option enable the selected partition to send BCPii commands. When selected, the active logical partition can send BCPii commands to other active logical partitions.

Enable the partition to receive commands from other partitions
Select this option to enable the selected partition to receive BCPii commands from other partitions. When selected, the active logical partition can receive BCPii commands from other active logical partitions.

> **All partitions**
> Select this option if you want the selected logical partition to receive BCPii commands from all the active logical partitions.

> **Selected partitions**
> Select this option if you want to remove or add selected logical partitions to receive BCPii commands from the logical partition.

>> **Add**
>> To add a system and logical partition to receive BCPii commands from the logical partition, click **Add**.

>> **Remove**
>> To remove a selected logical partition to receive BCPii commands from the logical partition, click **Remove**.

Counter facility security options:

Basic counter set authorization control
Select this option to authorize the use of the basic counter set. This set includes counts of central processing unit cycles, instructions executed, and directory-write and penalty cycles for level-1 instruction and data caches.

Problem state counter set authorization control
Select this option to authorize the use of the problem state counter set. This set includes counts of central processing unit cycles, instructions executed, and directory-write and penalty cycles for level-1 instruction and data caches only when the processor is in problem state.

Crypto activity counter set authorization control
Select this option to authorize the use of the crypto activity counter set. This set includes counters for a central processing unit related to PRNG, SHA, DEA, AES, and ECC function counts.

Extended counter set authorization control
Select this option to authorize the use of the extended counter set. The extended counters provide information about hardware facilities and structures that are specific to a machine family. The extended counters are designed to expand upon information provided by the Basic Counter Set.

Sampling facility security options:

Basic sampling authorization control
Select this option to authorize the use of the basic sampling function. Samples are taken and stored at the end of each sampling interval.

Diagnostic sampling authorization control
Select this option to authorize the use of the diagnostic sampling function. Samples are taken and stored at the end of each diagnostic interval.

CPACF key management operations:

Permit AES key functions
If the CPACF feature is installed, this option displays. Select this option to allow an Advanced Encryption Standard (AES) key to be wrapped using the CPACF TDES wrapping key.

Permit DEA key functions
If the CPACF feature is installed, this option displays. Select this option to allow a Data Encryption Algorithm (DEA) key to be wrapped using the CPACF TDES wrapping key.

Permit ECC key functions
If the CPACF feature is installed, this option displays. Select this option to allow a Elliptical Curve Cryptography (ECC) key to be wrapped using the CPACF TDES wrapping key.

Establishing optional characteristics

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile for an LP. Open the Options page to define the following LP characteristics:

- Minimum input/output (I/O) priority
- Maximum input/output (I/O) priority
- Defined capacity
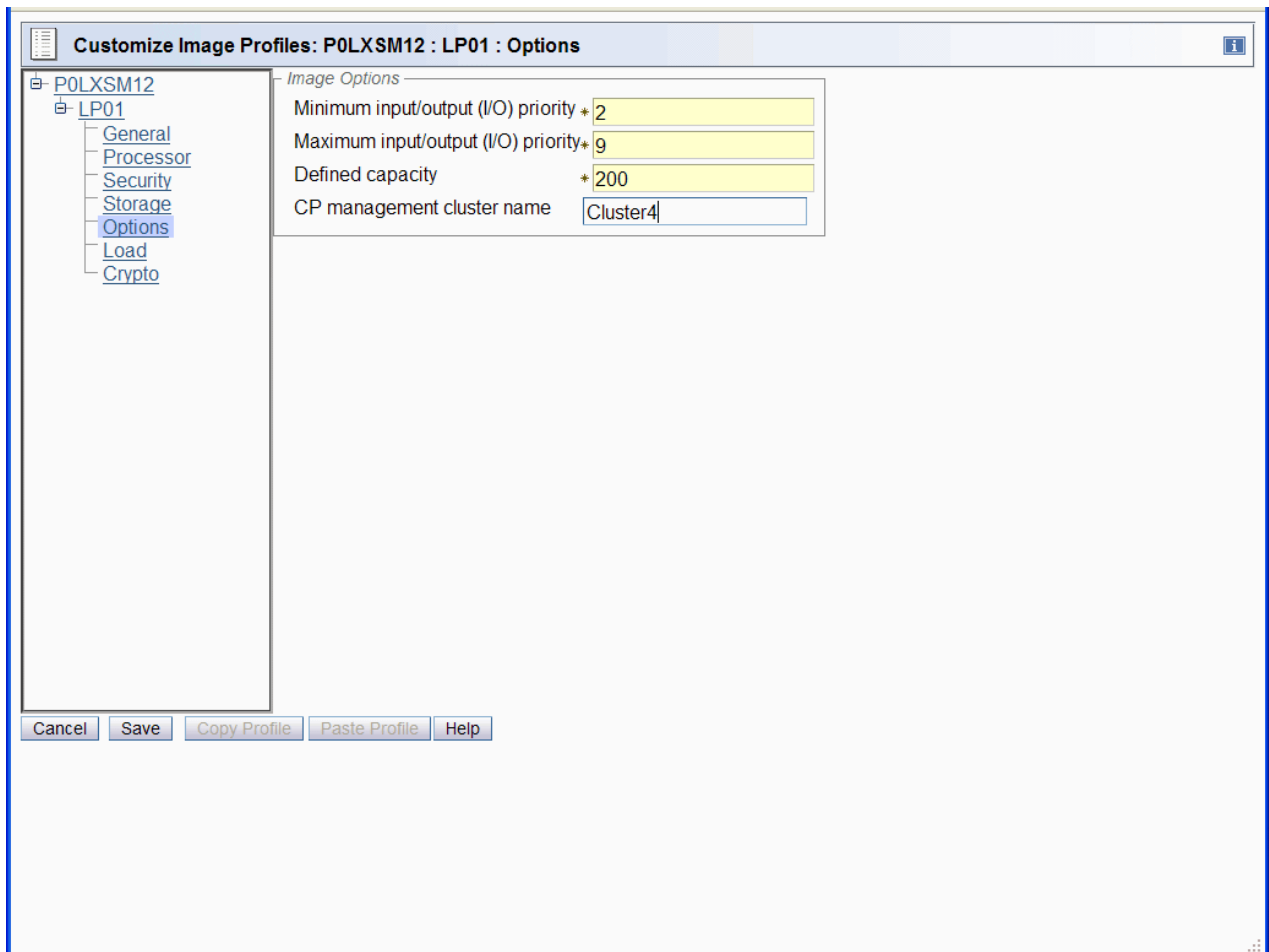
Figure 41. Options page, image profile

Image options parameter descriptions:

Minimum input/output (I/O) priority
Enter the minimum priority to be assigned to I/O requests from this logical partition.

Maximum input/output (I/O) priority
Enter the maximum priority to be assigned to I/O requests from this logical partition.

Defined capacity
Enter the upper bound in terms of millions of service units (MSUs) beyond which the rolling 4-hour average CPU utilization cannot proceed.

CP management cluster name
Enter the name of the Sysplex Cluster of which this logical partition is made a member. z/OS will not IPL if the name defined in the Image Profile does not match the sysplex name with which the IPLing system is associated.

Storage characteristics

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile for an LP. Open the Storage page to define the following LP characteristics:

- Central storage
- Virtual Flash Memory

Figure 42. Storage page, image profile

Central storage parameter descriptions:

See "Central storage" on page 94 for a discussion of the appropriate entries for these fields.

Initial
Enter, from the selection, the initial amount of central storage to be allocated to the LP at activation.

Reserved
Enter, in MB, GB, or TB, the amount of additional central storage requested for the LP. The reserved storage space is storage that can be dynamically brought online to the LP at some point after LP activation. Entering 0 limits central storage to the initial amount for the duration of the LP activation. Enter a value that is compatible with the storage granularity supported by your CPC.

Storage origin
If **Determined by the user** is selected, enter, in MB, GB, or TB, the central storage origin for the LP. When the LP is activated, it is allocated at the origin you specify here. Enter a value that is compatible with the storage granularity supported by your CPC.

Determined by the system
Select this option if you want the system to allocate where the LP storage resides.

Determined by the user
Select this option if you want to allocate where the LP storage resides.

Virtual flash memory parameter descriptions:

Initial
Enter the initial amount of Virtual Flash Memory to be allocated to the LP at activation.

Maximum
Enter the maximum amount of Virtual Flash Memory for the LP.

Establishing Secure Service Container parameter descriptions

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile for a Secure Service Container. Select the partition that is to function as the IBM zAware partition. Define the selected partition characteristics:

- Boot selection
- Host name
- Master user ID
- Master password
- Confirm master password
- Network Adapters
- IPv4 gateway
- IPv6 gateway
- DNS Servers

Figure 43. Secure Service Container page

Secure Service Container parameter descriptions

The profile tree view contains the **SSC** page. When the Secure Service Container partition is activated, the sequence of events varies, depending which boot selection you specified on the **SSC** page of the image profile.

Boot selection

Secure Service Container installer
Select this option if you want the partition start process to initialize the Secure Service Container Installer so you can install an appliance. This boot selection is the only option when you start a newly configured Secure Service Container partition for the first time. With this option, the Secure

Service Container Installer is started automatically. When the start process competes, you can access the Secure Service Container Installer through your choice of browser.

Secure Service Container
Select this option if you want the partition start process to effectively restart an installed appliance. If you previously used the Secure Service Container Installer to successfully install a firmware or software appliance, this boot selection becomes the default selection in the image profile for the Secure Service Container partition. In this case, the Secure Service Container Installer is rebooted, and the installed appliance is restarted in the Secure Service Container partition on this and all subsequent reboots, until you change the boot selection in the image profile.

Host name
Enter the host name for the partition.

Master user ID
Enter the user ID to be used as the default master user ID.

Master password
Enter the password for the master user ID. A master password can have a minimum of 8 characters and a maximum of 256 characters.

Confirm master password
Re-enter the password exactly as you typed it for the Master password field.

IPv4 gateway
Use the network adapter table to view and change an IPv4 address and detail settings for the selected network adapters.

IPv6 gateway
Use the network adapter table to view and change an IPv6 address and detail settings for the selected network adapters.

DNS Servers
Use the DNS servers table to add, edit, or remove the IP address for the IPv4 or IPv6 address.

Load information

Use the **Customize/Delete Activation Profile**s task to open an image or load profile for an LP. Open the Load page to define the following LP characteristics:

- Load during activation
- Load address
- Load parameter
- Load timeout value

Figure 44. Load page, image profile

Load parameter descriptions

Load during activation
Selecting this option allows initial program load of the operating system to occur automatically at LP activation. The default is not selected.

Load type
You can select the following:

Standard load
To perform the load on the logical partition

SCSI load
To IPL from a device that requires a SCSI load

SCSI dump
To IPL a standalone dump program from a device that requires a SCSI load.

NVMe load
To IPL from a device that requires a NVMe load

NVMe dump
To IPL a standalone dump program from a device that requires a NVMe load.

Enable Secure Boot for Linux
To verify the signature of the load program and distributor's signature match , select **Enable Secure Boot for Linux**.

Clear the main memory on this partition before loading it
Select this option on the Load Profile page to clear main memory storage before a load.

Note: Available when **Standard load**, **SCSI load**, or **NVMe load** are selected. Clearing partitions with larger amounts of main memory storage may take longer.

Load address
Enter the hex address of the I/O device containing the operating system to be loaded automatically at LP activation.

Use dynamically changed address
Select this option if you want to use the load address from a dynamically changed I/O configuration. This option and the **Load address** option are mutually exclusive.

Load parameter
Enter a 1 to 8 character, optional IPL load parameter for loading an operating system on each activation of the LP. This is useful for loading z/OS, or z/VSE. Valid characters for a load parameter are:

- At (@)
- Pound (#)
- Dollar ($)
- Blank character
- Period (.)
- Decimal digits 0 through 9
- Capital letters A through Z .

Use dynamically changed parameter
Select this option if you want to use the load parameter from a dynamically changed I/O configuration. This option and the **Load address** option are mutually exclusive.

Time-out value
Enter a time-out value in seconds to limit the amount of time for successful completion of the operating system load.

Worldwide port name
Specify the Worldwide Port Number identifying the Fibre Channel port of the SCSI target device (according to the FCP/SCSI-3 specifications). This is a 64-bit binary number designating the port name, represented by 16 hexadecimal digits. This is required for SCSI IPL or SCSI Dump.

If the selected load type is **Standard load**, **NVMe load**, or **NVMe dump** this field is unavailable.

Logical unit number
Specify the number of the logical unit as defined by FCP (according to the FCP/SCSI-3 specifications). This is the 64-bit binary number designating the unit number of the FCP I/O device, represented by 16 hexadecimal digits. This field is required for SCSI IPL or SCSI Dump.

If the selected load type is **Standard load**, **NVMe load**, or **NVMe dump** this field is unavailable.

Boot program selector
Specify the DASD partition number in decimal notation.

If the selected load type is **Standard load**, this field is unavailable.

Boot record logical block address
Specify the load block address. This is a 64-bit binary number, represented by 16 hexadecimal characters, designating the logical block address of a boot record on the FCP-load device. If no block address is specified, the logical-block address of the boot record is assumed to be zero. This feature could be used to IPL using a second or backup boot record, in case the original one is corrupted or overwritten by accident.

If the selected load type is **Standard load**, this field is unavailable.

Operating system specific load parameters
Specify a variable number of characters to be used by the program that is loaded during SCSI IPL or SCSI Dump. This information will be given to the IPLed operating system and will be ignored by the machine loader. The IPLed operating system (or standalone dump program) has to support this feature. Any line breaks you enter are transformed into spaces before being saved.

If the selected load type is **Standard load**, this field is unavailable.

Cryptographic characteristics

Use the **Customize/Delete Activation Profiles** task to open a reset or image profile for an LP.

Note: To verify the active settings for the cryptographic characteristics use the **View LPAR Cryptographic Controls** task (For information regarding the View LPAR Cryptographic Controls page, see "Reviewing and changing current logical partition cryptographic controls" on page 171.)
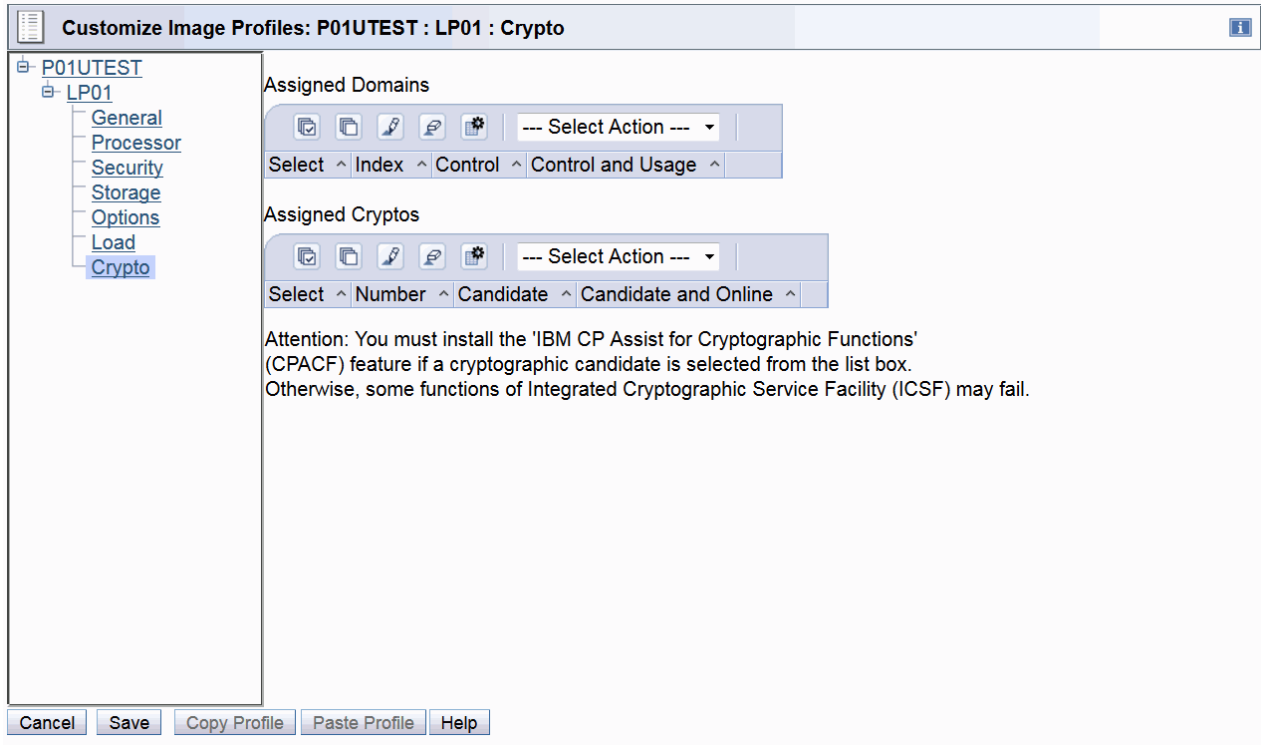
Figure 45. Crypto page, image profile
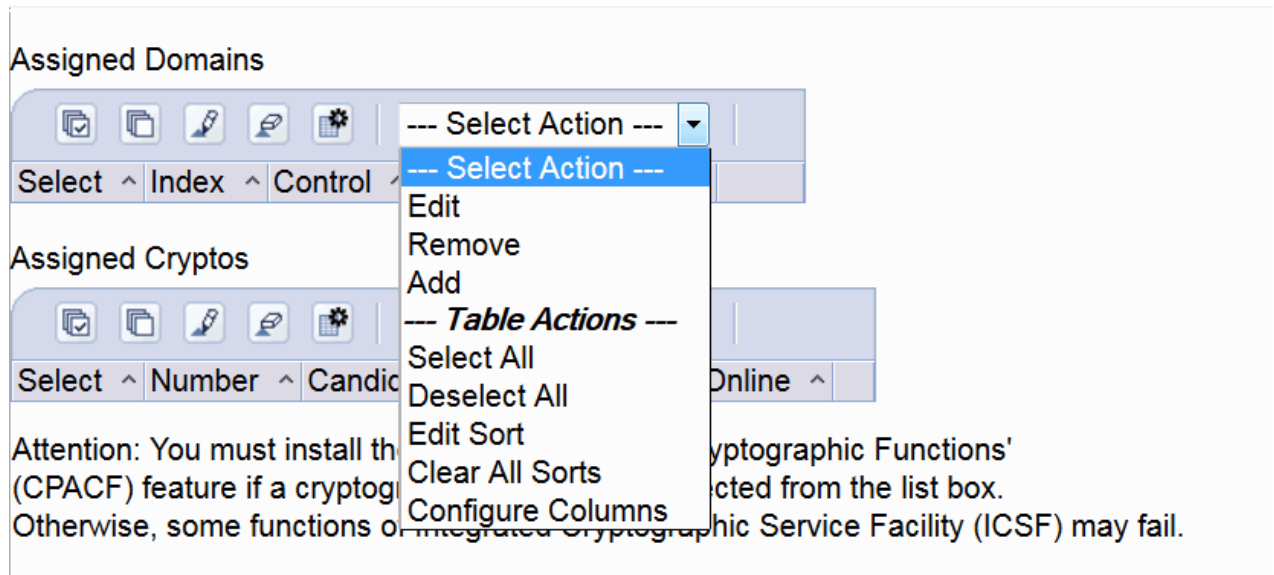
Figure 46. Add, Remove Cryptos

Crypto parameter descriptions

The following crypto options in the image profile can be specified.

Assigned domains:

The assigned domains table displays the control domains and control and usage domain indexes which can be modified in the logical partition.

Control domain

The logical partition's control domains are those cryptographic domains which remote secure administration functions can be established and administered from the logical partition when set up as the TCP/IP host for the TKE Workstation.

If you are setting up the host TCP/IP in this logical partition for communicating with the TKE Workstation, the partition will be used as a path to this and the other domains' keys. Indicate all the domains you want to access, including this partition's own domain, from this partition as control domains.

Note: You can manage both master keys and operational keys from a TKE Workstation.

For more TKE Workstation information, refer to the *z/OS Cryptographic Services ICSF TKE Workstation User's Guide.*

Control and Usage Domain

The logical partition's control and usage domains are domains in the cryptos that can be used for cryptographic functions. The usage domains cannot be removed if the crypto is online. A logical partition's control domains can include the usage domains of other logical partitions. Assigning multiple logical partitions usage domains as control domains of a single logical partition allows using it to control their software setup.

If running z/OS, one of the usage domain index(es) selected must match the domain number entered in the Options dataset when starting this partition's instance of ICSF. As of z/OS 1.2 the usage domain specification in the Options dataset is only required if multiple usage domain index(es) are selected.

If running z/VM in a logical partition with guests, such as Linux or z/OS, a range of usage domain indices should be selected when assigning access to the cryptographic accelerator or coprocessor. A range will allow more than one guest to have dedicated or shared access to the cryptographic queues. For further information, see the *z/VM CP Planning and Administration* and *z/VM Running Guest Operating Systems* documents.

The Usage Domain assignment, in combination with the Cryptographic Number must be unique across all partitions defined to the CPC. If you assign Usage Domain 1 on Crypto Adapter 1 to LP 1 and LP 11, then the first of those LPs to be activated is given access to that Usage Domain and the second LP cannot be activated because that Usage Domain is already assigned and no longer available.

The maximum number of LPs that can have a cryptographic adapter assigned depends on how many cryptographic adapters are available. Each cryptographic adapter can support up to 16 Usage Domains, so the maximum number of LPs that can be assigned cryptographic hardware is 16, the number of adapters. In this case, each LP would be assigned one and only one adapter. There would be no additional crypto capacity for these LPs nor would there be any redundancy in case of a failure on one of the crypto devices. The Crypto Express feature has a single adapter available.

LPs that are predominately for development or testing may only need a single adapter assigned to provide functionality. Production LPs will likely need at least two adapters assigned for redundancy and may need multiple adapters assigned to provide acceptable performance and throughput of the crypto workload.

The type configuration of the adapter (coprocessor, accelerator or EP11 mode) means that the adapter can only be used in that mode. If you only have a small EP11 mode workload on a single LP, configuring an adapter in EP11 mode means that the adapter will only be used for EP11 mode for that particular LP. If you need two EP11 adapters (for redundancy) those adapters will be unavailable for any other work.

Consider an environment where you have multiple LPs supporting various types of workload (coprocessor, accelerator and EP11 mode) and the assignment of adapters and Usage Domains across those LPs.

You can have Crypto Express hardware installed. The following example includes each type of device. Your configuration will likely be much simpler, probably with only one type of adapter, or at most two types.

Table 17. Example Selection of Usage Domain Assignment

Feature	Adapter	Crypto Configuration	Crypto Number	Crypto Label
Crypto Express	1	X - Coprocessor	06	X06
Crypto Express	1	A- Accelerator	07	X07
Crypto Express	1	P-EP11	08	P08
Crypto Express	1	P-EP11	09	P09

Note: For availability reasons, it is recommended that at least two cryptographic adapters with the same capability be assigned to each partition that executes cryptographic operations. Because accelerators do not contain any internal security data (cryptographic keys), all accelerators are equivalent. Coprocessors and EP11 Coprocessors will contain cryptographic keys and it is recommended that at least two coprocessors with the appropriate domains and cryptographic keys be assigned to a logical partition (LP) that requires secure key operations.

There are multiple LPs which require access to crypto hardware. These include LPs that will only perform SSL workload and require access only to accelerators. Some LPs only need access to secure key coprocessors. These LPs may perform SSL workload, but volume is sufficiently low that work can be performed on the coprocessor and an accelerator is not required. Other LPs perform both secure key work and sufficient SSL workload that also assigning an accelerator makes sense. There is also a need for a VM environment with multiple guests sharing access to cryptographic adapters. There are multiple LPs that will be performing EP11 workload.

Table 18. Example Selection of Usage Domain Assignment

LP and Crypto Use	Usage Domain Assignment	Adapter Assignment	Second Assigned Adapter (for capacity and/or redundancy)
PRODSSL0 Prod SSL only	UD=0	A00	A04
PRODSSL0 Prod SSL only	UD=1	A00	A04
PRODCOM2 Prod SSL & secure key	UD=2	A00 & X02	A04 & A05
PRODSSL3 Prod SSL only	UID=0	A01	A07
PRODSSL4 Prod SSL only	UD=3	A00	A01
PRODSSLF Prod SSL only	UD=3	A04	A07
PRODCOM3 Prod SSL & secure key3	UD=4	A00 & X05	A04 & X06
TESTSEC1 secure key only	UD=5	X02	X05
PRODSEC2 secure key only	UD=6	X03	X06
TESTSSL9 Test SSL only	UD=6	A07	A04
TESTSEC1 secure key only	UD=5,12	X03	X06
DEVVM	UD=7,8,9,10	X05	

Table 18. Example Selection of Usage Domain Assignment (continued)			
LP and Crypto Use	**Usage Domain Assignment**	**Adapter Assignment**	**Second Assigned Adapter (for capacity and/or redundancy)**
TESTPKCS	UD=0	P08	P09

For example:

- The LP PRODSSL0 only performs SSL work and therefore only needs an accelerator. It is assigned A00 (A for Accelerator and Crypto Number 00). And for redundancy, it is also assigned A04 (A for Accelerator and Crypto Number 04) to provide additional capacity and/or redundancy in case the first card fails.

  This LP is also assigned Usage Domain 0. Even though an accelerator does not have master keys loaded, it is still assigned a Usage Domain. This means that Usage Domain 0 on cryptographic adapters 00 and 04 are no longer available to be used by any other LPs.

- The PRODSSL1 LP also only performs SSL workload so it can be assigned the same two adapters (A00 and A04), however this LP will be assigned Usage Domain 1.

- The PRODSSL3 is a third LP that only performs SSL workloads and therefore only needs to have an accelerator assigned. We assign A07 (on a Crypto Express) to this LP. This LP is assigned Usage Domain 0.

  Note: This is not in conflict with the PRODSSL0 LP, which also uses Usage Domain 0, but on different crypto adapters.

- The PRODCOM2 has a different workload requirement. It performs SSL workload and will benefit from having an accelerator assigned, but it also performs secure key operations and must have a coprocessor assigned as well. It is also assigned to use A00 and A04. Iit is sharing the same accelerator as PRODSSL0 and PRODSSL1, but with a different Usage Domain.

- The PRODSSLF is another LP with only SSL workload. It has been assigned A07 (Crypto Express) and uses Usage Domain 3 on those two cards. Even though PRODSSL4 and PRODSSLF are both using Usage Domain 3, they use it on two different adapters, so there is no conflict.

- The PRODCOM3 LP is similar to PRODCOM2 in that it has combined SSL workload and secure key work, so it is assigned X06 (Crypto Express) for coprocessors. It is assigned Usage Domain 4 on these cards.

- The PRODSEC1 is a secure key only partition. If it does any SSL work, it's a trivial amount that can be handled by the secure key cards without impacting the other secure key work going on in the LP.

- The PRODSEC2 is a secure key only partition, possibly running in a Sysplex with PRODSEC1. It is assigned Usage Domain 6 on X06 (Crypto Express). If these two LPs are using the same ICSF repositories, the same master keys will be loaded into both Usage Domains.

- The TESTSSL9 is an LP for testing new applications that only require System SSL. It is assigned A07 (Crypto Express). It is assigned two accelerators not for throughput, but for redundancy. Testing can continue even if one of the two accelerators should have a problem. It is assigned Usage Domain 6 on these cards.

- The TESTSEC1 LP is primarily intended for testing applications that require secure key technology. It has two Usage Domains assigned, but only one can be used at a time. It is assigned two coprocessors, and X06 (Crypto Express), and it will normally use Usage Domain 5 on those cards. However, in an emergency, this LP can also be IPL'd as a production LP to provide additional capacity for the Sysplex that includes PRODSEC1 & PRODSEC2. That is, if the workload on the Sysplex exceeds the capacity of those two LPs, then this test LP could be shut down and another copy of the production LP IPL'd here. In this configuration, it would still use X03 and X06, however ICSF would point to Usage Domain 12 and the same key repositories as PRODSEC1 and PRODSEC2. This Usage Domain would contain the same master key as used by PRODSEC1 and PRODSEC2, so it can access the key material in the shared ICSF repositories.

- The DEVVM LP is a development LP that provides multiple guest operating systems, which require access to a secure key device. The LP is also assigned 4 Usage Domains (7 through 10). Presumably there will be four guests running in this LP, each assigned there own unique Usage Domain via the VM User Directory. Since this environment is only for development, there is no backup crypto adapter assigned. If X05 is unavailable, secure key work will stop on these guests.

- The TESTPKCS LP is for testing new PKCS #11 applications running on the EP11 coprocessor. It is assigned to use P08 (Crypto Express) and P09 (Crypto Express) and Usage Domain 0 on these coprocessors. Once these PKCS #11 applications have been sufficiently tested and are ready for production, there would be another LP defined, PRODPKCS, which would also be assigned P08 and P09, but would have a different Usage Domain assigned.

Assigned cryptos

The assigned cryptos table displays the cryptographic candidate list and cryptographic candidate and online list which can be modified in the logical partition.

Candidate

The Candidate identifies the cryptographic numbers that are eligible to be accessed by this logical partition. Select from the list the number(s), from 0 to 15, that identify the coprocessor or accelerator to be accessed by this partition.

When the partition is activated, an error condition is not reported if a cryptographic number selected in the Assigned Cryptos table is not installed in the system. Selecting a cryptographic number that is not installed prepares the settings in the active partition in the event that you wish to nondisruptively install the crypto in the future.

A Crypto Express contains single adapter which can be configured as a CCA coprocessor, EP11 coprocessor, or an accelerator. The default configuration is CCA coprocessor. A crypto adapter can be shared across all partitions utilizing usage domains.

It is possible to select all 16 candidate numbers (0-15) even before a crypto feature is installed. When a new crypto feature is installed and its cryptographic number(s) have been previously selected in the Candidate list of an active partition, it can be configured on to the partition from the Support Element using the **Configure On/Off** task.

Selecting all the values will not cause a problem if you have 16 or fewer partitions in your configurations that will be using the Crypto Express feature. If you have more than 16 partitions that require access to cryptographic coprocessors or accelerators, carefully assign the cryptographic numbers across the partitions in conjunction with unique Usage Domain selections.

Table 19. Example Selection of Crypto Numbers

Feature	Adapter	Crypto Configuration	Type	Crypto Number
Crypto Express7S/6S/5S 1	1	Accelerator	A	00
	1	Accelerator	A	01
Crypto Express7S/6S/5S 2	1	Coprocessor	X	02
	1	Coprocessor	X	03
Crypto Express7S/6S/5S 3	1	Accelerator	A	04
	1	Coprocessor	X	05
Crypto Express7S/6S/5S 4	1	Coprocessor	X	06
	1	Accelerator	A	07

It is recommended that at least two cryptographic adapters of the same type and capability be assigned to each partition that executes cryptographic operations. Because accelerators do not

contain any internal security data (cryptographic keys), all accelerators are equivalent. coprocessors, on the other hand, will contain cryptographic keys and it is recommended that at least two coprocessors with the appropriate domains and cryptographic keys be assigned to a logical partition (LP) that requires secure key operations.

Table 20. LP & crypto assignments

LP & Crypto Use	Usage Domain Assignment	Logical Partition Assignment	Backup Required? Specify 2nd Logical Partition
ACME0 Prod SSL only[1,2]	UD=0	A00	A04
ACME1 Prod SSL only[1]	UD=1	A00	A04
ACME2 Prod SSL & secure	UD=2	A00 & X02	A04 & X05
ACME3 Prod SSL only[2]	UD=0	A01	A07
.... SSL only	UD=3...10	A00	A01
ACMEF Prod SSL only	UD=0	A04	A07
ACM17 Prod SSL & secure[3]	UD=4	A00 & X05	A01 & X06
ACM18 Test SSL & secure[3]	UD=5, 2[4]	A00 & X02	A04 & X05
ACM19 Test SSL only	UD=6	A07	A04
ACM5VM Prod VM	UD=7, 8, 9, 10	A07 & X05	

Notes:

1. LPs ACME0 and ACME1 both use Accelerator cards A00 and A04, however, they use two different Usage Domains on these cards.

2. LPs ACME0 and ACME3 both use Usage Domain 0, but they use them on different accelerator cards, A00/A04 and A01/A07.

3. LPs ACM17 and ACM18 both use Crypto Coprocessor X05, but they use different Usage Domains on those cards, so there is no conflict.

4. ACM18 has two Usage Domains assigned, but only one can be used at a time. Normally, this TEST LP will provide SSL and Secure support for the Test environment using Usage Domain 5 on crypto accelerator cards A00 and A01, and crypto coprocessor card X02. By defining this LP with access to Usage Domain 2 it can be a backup LP for ACME2. If and when there is a problem with LP ACME2, that operating system can be IPL'd in this LP, with the ICSF started task pointing to UD=2, and it will be able to access the cryptographic keys for ACME2, which are stored in Usage Domain 2 on X05.

Note: It is important to make the correct crypto number assignments from the Assigned Cryptos table for each of these logical partitions to avoid assignment conflicts.

If the customer plans to use ICSF or the optional cryptographic hardware, the CP Crypto Assist functions (CPACF DES/TDES) must be enabled. Many IBM products will take advantage of the cryptographic hardware using ICSF, so enabling CPACF is recommended. See the *z/OS ICSF Administrator's Guide* and the *z/OS ICSF System Programmer's Guide* for complete information.

Candidate and Online

The Candidate and Online identify the cryptographic numbers that are automatically brought online during logical partition activation. The cryptographic numbers selected in the assigned table must also be selected in the Candidate.

When the logical partition activation is complete, installed Cryptographic features that are in the Candidate column but not in the Candidate and Online Column are in a *configured off* state (Standby). They can be later configured *on* to the partition using the **Configure On/Off** task.

When the partition is activated, an error condition is not reported if the cryptographic number selected from the assigned table is not installed in the system. The cryptographic number is ignored and the activation process continues.

If a cryptographic number selected from the assigned cryptos table has been configured off to the partition, it is automatically configured back on during the next partition activation.

Creating a logical partition group profile

Creating a group, or grouping logical partitions, is a way to assign more than one activation profile to an object, rather than changing the object's assigned activation profile every time you want to activate it differently. Grouping creates copies of objects on the Support Element workspace. The objects can be the CPC or its images. Different groups can contain the same object, such as the CPC, but the object's settings in one group can be customized independently of its settings in other groups. One such setting is the activation profile assigned to the object.
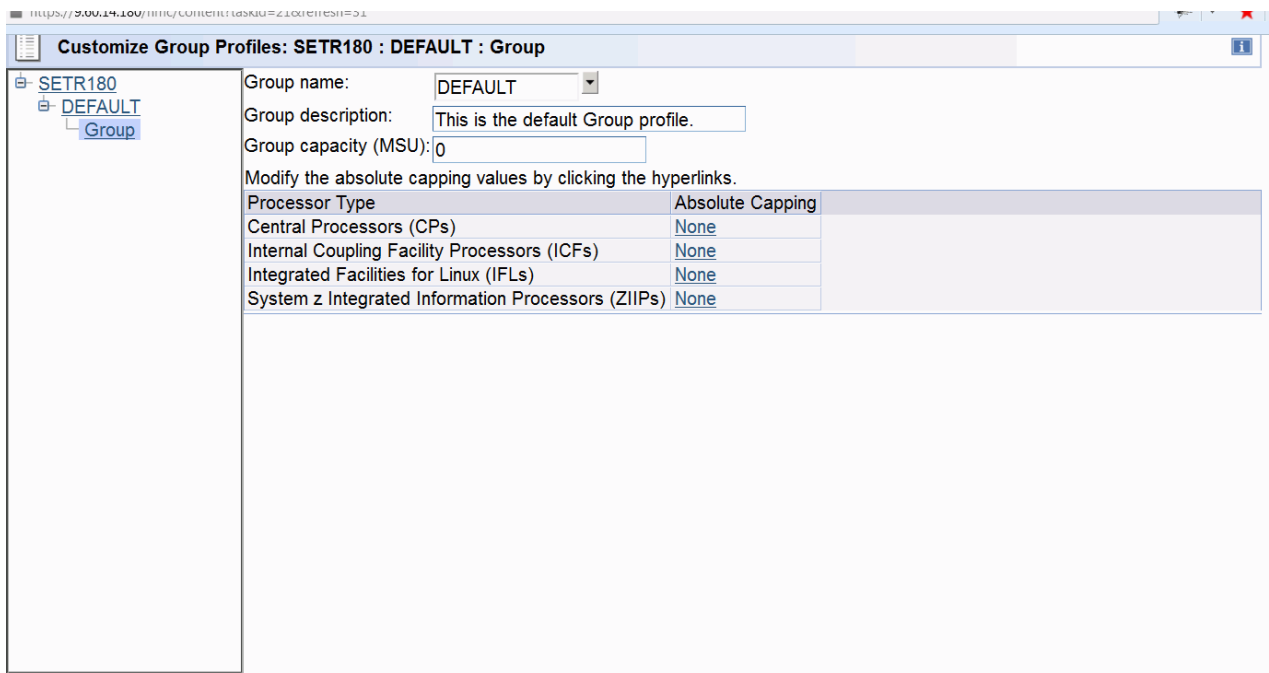
Figure 47. Customize Group Profiles window
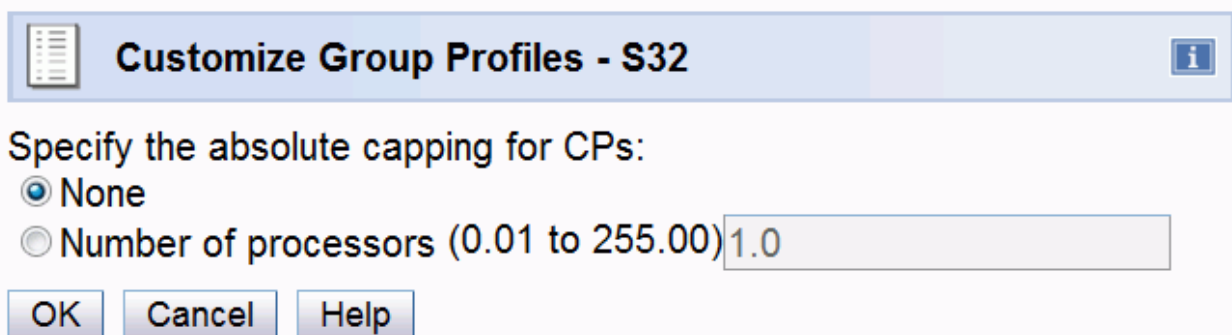
Figure 48. Edit absolute capping

On the Customize Group Profiles window, select the **Group name** list to select a group to use as a template for the new group, or use DEFAULT if no other groups exist. Then enter a unique name for the logical partition in the Group name field. Enter a description of the new group name in the **Group description** field. Click **Save** to save the new group profile.

For the **Group Capacity**, enter the upper bound in terms of millions of service units (MSUs). A value of 0 indicates the setting is unused. Refer to *z/OS MVS Planning: Workload Management* and *z/OS Planning for Subcapacity Pricing* for help in choosing an appropriate value for this field, based on your workload needs.

For the **Absolute Capping**, use this field to change the absolute capping of logical partitions in a group that share processors. The absolute capping can be None or a number of processors value from 0.01 to 255.0. To change an absolute capping for a processor type for a group, select the current absolute capping setting in its field and click the hyperlink to display the next Edit Absolute Capping window. Specify the absolute capping for the selected processor type to indicate the new setting.

Enabling Input/Output priority queuing

Use the **Enable I/O Priority Queuing** task to either enable or disable I/O priority queuing for the entire CPC.
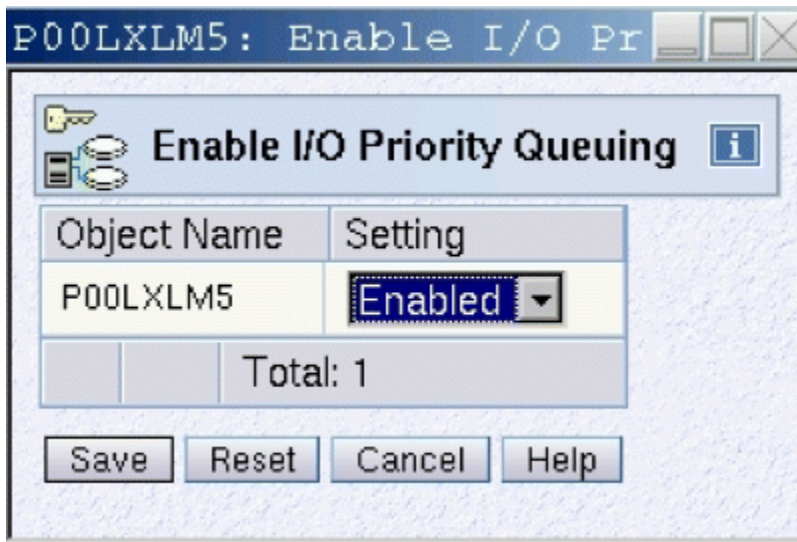
Figure 49. Enabling I/O priority queuing

Changing logical partition Input/Output priority queuing values

Use the **Change Logical Partition I/O Priority Queuing** task to set the minimum and maximum I/O priority queuing values for logical partitions.

Figure 50. Change Logical Partition I/O priority queuing

Note: Minimum I/O Priority and **Maximum I/O Priority** should be specified as a range of values that give software some ability to make choices. All logical partitions in a given LPAR cluster should be given the same range of values so that Workload Manager can optimize I/O throughput across the LPAR cluster.

If the software in the logical partition does not have an understanding of I/O Priority Queuing, the system programmer should set the Minimum and Maximum I/O priorities to the same value. The value chosen is assigned to that logical partition as a constant priority value relative to all other logical partitions. This way even logical partitions that do not employ IRD technologies can benefit from this support.

Parameter descriptions

Minimum input/output (I/O) priority
    Enter the minimum priority to be assigned to I/O requests from this logical partition.

Maximum input/output (I/O) priority
    Enter the maximum priority to be assigned to I/O requests from this logical partition.

Moving unshared channel paths

You can move reconfigurable channel paths owned by one LP to another LP.

Moving unshared channel paths from a z/OS system

1. Select the LP that owns the channel path to display channel path information for the LP.
2. Move the channel path from the z/OS console.

   a. Enter **CF CHP(***nn***),OFFLINE**, where *nn* is the number of the desired channel path, from the z/OS operator console that has the reconfigurable channel path online.

   b. Enter **CF CHP(***nn***),ONLINE**, where *nn* is the number of the desired channel path, from the z/OS operator console that is the target LP.

   Notes:

   1) If a channel path is configured offline while the LP is isolated and remains offline when you change the LP's isolation status from enabled to disabled, the channel path must be configured offline again. See "Releasing reconfigurable channel paths" on page 158.

   2) If the channel path you want to move is currently assigned to a deactivated LP, you must configure it offline from the hardware console.

   3) If the LP is not running a system control program that supports channel path configuration commands, you can move the channel path from the hardware console.

Moving a channel path from the hardware console

1. Select the LP that owns the channel path to display channel path information for the LP.
2. Select a reconfigurable channel path. Open the **Reassign Channel Path** task.
3. The Reassign a Channel Path window displays the targeted channel path, the current owning LP, and a list of target LPs you can reassign the channel path. Select the LP that you want to reassign the channel path.
4. Click **Reassign** and confirm the action to release the channel path.
5. When the *Requested operation is complete* message displays, click **OK**.

Releasing reconfigurable channel paths

Use this procedure when the owning LP has LP isolation enabled.

1. Select the LP that owns the channel path to display channel path information for the LP.
2. Select a reconfigurable channel path. Open the **Release** task.

   A *Confirm the Action* message displays warning you that the channel paths will be released and made available for reassignment to other LPs. Confirming the action releases the channel path and a *Requested operation is complete* message displays. Click **OK** to complete the task.

Configuring shared channel paths

Verify the status of the channel path for each LP to which you plan to configure the channel path by opening each LP's CHPIDs Work Area.

Enter **CF CHP(***nn***),ONLINE** (where *nn* is the number of the desired CHPID) from each z/OS operator console to which the CHPID is to be brought online.

If the operating system running in the LP is not z/OS, use the **Configure On/Off** task to configure the CHPID online. The shared channel path will be physically configured when the first LP configures the channel path online.

Note: Dynamically managed channel paths can be configured *Off* but **cannot** be configured *On* from the CHPID Operations task list.

Deconfiguring shared channel paths

Verify the status of the channel path for each LP to which you plan to configure the channel path.

Enter **CF CHP(***nn***),OFFLINE** (where *nn* is the number of the desired CHPID), from each z/OS operator console from which the CHPID is to be taken offline.

If the operating system running in the LP is not z/OS, use the **Configure On/Off** task to configure the CHPID offline. The shared channel path will be physically deconfigured when the last LP that had the channel path online configures it offline.

Removing shared channel paths for service

1. Enter (from the z/OS console that has the shared channel path online) **CF CHP(***nn***) OFFLINE**, where *nn* is the number of the desired CHPID.
2. Use the Toggle all off option in the **Configure On/Off** task to remove the CHPID.

   The Toggle all off option detaches the CHPID from all LPs that it is currently attached to regardless of the target LP. A CHPID that is shared by multiple LPs is detached from all LPs without forcing you to detach it individually from each LP.

   Toggle all off is also valid for unshared CHPIDs and it is mutually exclusive of the **Release** task.
3. Use the **Service On/Off** task to remove the CHPID.

Changing logical partition definitions

You can make changes to LP definitions that are available dynamically to a running LP or that are available at the next LP activation.

Changes available dynamically to a running LP

The following changes are available dynamically to a running LP:

- Using the **Change Logical Partition Controls** task, you can change the following LP definitions:
  - Defined capacity
  - Workload Manager enablement
  - Initial processing weight
  - Minimum processing weight
  - Maximum processing weight
  - Current capping
  - Initial capping
  - Absolute capping
  - Number of Dedicated Processors
  - Number of Not dedicated Processors
  - Processor running time (globally applicable to logical cores of all shared LPs)
  - Global enablement of event-driven dispatching
- Using the **Change Logical Partition Input/Output (I/O) Priority Queuing Controls** task, you can change the following LP definitions:
  - Minimum input/output (I/O) priority
  - Maximum input/output (I/O) priority

- Using the **Change Logical Partition Security** task, you can change the following LP definitions:
  - Performance data control
  - Input/Output configuration control
  - Cross partition authority
  - BCPii permissions
  - Partition isolation
  - Basic counter
  - Problem state counter set authorization control
  - Extended counter set authorization control
  - Crypto activity counter set authorization control
  - Extended counter set authorization control
  - Basic sampling authorization control
  - Diagnostic sampling authorization control
  - CPACF key management operations
- Using the **Logical Processor Add** task, you can do the following:
  - Increase the number of reserved CPs for a processor type
  - Add a new processor type not yet in use for that partition
  - Increase the number of initial CPs for a processor type
  - Change the running system or save the changes to profiles.
- Using the **Change LPAR Cryptographic Controls** task, you can:
  - Add unassigned crypto(s) and domain(s) to a logical partition for the first time. (Update the Candidate list and Online List)
  - Edit assigned crypto(s) and domain(s) types to a logical partition already using cryptos and domains (Update the Candidate list and Online List)
  - Remove crypto(s) and domain(s) from a logical partition.
  - Zeroize or clear the cryptographic secure keys for a given usage domain in a partition.

Changes available at the next LP activation

The following changes are available at the next LP activation:

- Use the IOCP RESOURCE statement to specify the MIF image ID numbers assigned to the LPs. The only way to change the specified MIF image ID numbers is by creating a new IOCDS.
- Use the Partitions page in the reset profile to change the:
  - LP automatic activation order
- Use the **General** page in the image profile to change the:
  - Logical partition identifier
  - Mode of the LP
  - Enablement of the Logical partition time offset
- Use the **Processor** page in the image profile to change the:
  - Number of CPs, ICFs, IFLs, or zIIPs
  - Whether or not CPs are dedicated to an LP
  - Weight capping
  - Workload Manager enablement
  - Initial processing weight

- Initial capping enablement
- Minimum processing weight
- Maximum processing weight
- Absolute capping

• Use the **Options** page in the image profile to change the:

- Minimum input/output (I/O) priority
- Maximum input/output (I/O) priority
- Defined capacity
- CP Management Cluster

• Use the **Security** page in the image profile to change the:

- Global performance data control
- Input/output configuration control
- Cross partition authority
- Logical partition isolation
- Enable the partition to send commands
- Enable the partition to receive commands from other partitions
- Basic counter set authorization control
- Problem state counter set authorization control
- Crypto activity counter set authorization control
- Extended counter set authorization control
- Basic sampling authorization control
- Diagnostic sampling authorization control
- Permit AES key functions
- Permit DEA key functions
- Permit ECC key functions

• Use the **Storage** page in the image profile to change the:

- Central storage definitions
- Virtual flash memory definitions

• Use the **SSC** page in the image profile to change the:

- Boot selection
- Host name
- Master user ID
- Master password
- Network adapters
- DNS servers

• Use the **Time Offset** page in the image profile to change the:

- Logical partition time offset

• Use the **Load** page in the image or load profile to change the:

- Automatic load data

• Use the **Crypto** page in the image profile to change the:

- Assigned domains
- Assigned cryptos

Changes available at the next Power-On Reset (POR)

The following changes are available at the next power-on reset.

- Use the Options page in the reset profile to change the:
  - Global enablement of input/output (I/O) priority queuing
  - Processor running time interval
  - Enablement of event-driven dispatching
  - Automatic input/output (I/O) interface reset

Chapter 4. Monitoring the activities of logical partitions

This chapter describes the tasks and windows that can be used to monitor LP activity. It also provides LP performance information and provides guidelines and suggestions for planning a recovery strategy for operation.

Reviewing current storage information

Use the **Storage Information** task to open the **Storage Information** task to display LP storage information for LPs currently activated on the CPC.

For this example, assume the amount of customer storage is 5184 GB (5308416 MB). The amount of addressable central storage used by hardware system area (HSA) is 256 GB (262144 MB), leaving 4928 GB (5046272 MB) of addressable storage for allocation to LPs.

Figure 51. Storage information task

Reviewing partition resource assignments

Use the **View Partition Resource Assignment** task to open the **View Partition Resource Assignment** task to view the mapping of active logical partitions and associated processor information. The active logical partitions are identified at the top of the table and the Node and Chip numbers associated with each active logical partition are identified on the left. The Chip displays the processor Chip number associated with the Node and lists the processor types associated with each active logical partitions. The physical processor types for the active logical partitions may have conditions identified as Dedicated or Shared.
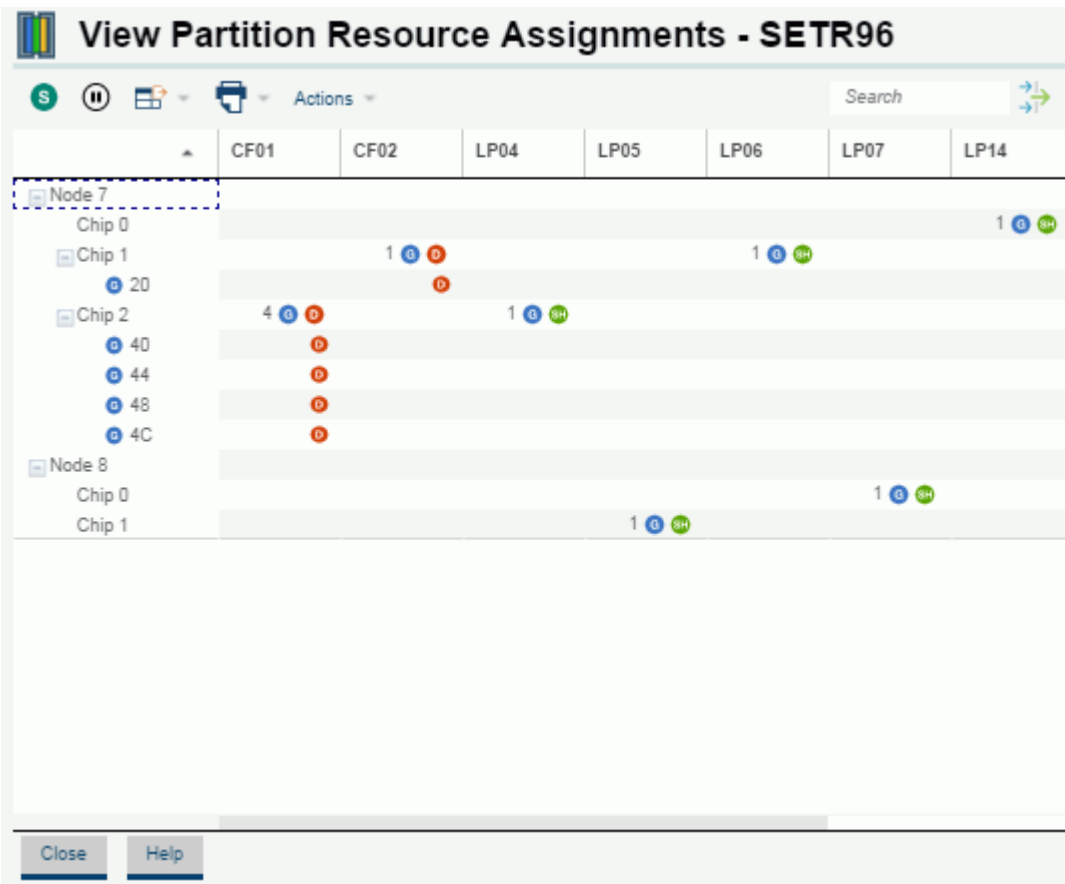
Figure 52. View Partition Resource Assignments

Reviewing and changing current logical partition controls

Use the **Change Logical Partition Controls** task to display and modify logical partition controls for a logical partition.

The Current Weight for an active shared logical partition will always show nonzero. For non-WLM managed logical partitions, the current weight will always be equal to the initial processing weight. The current weight for Workload Manager (WLM) managed logical partitions is the value WLM has currently assigned to the logical partition. Current Capping will always be indicated for an active logical partition that has Initial Capping. Current capping will also be indicated for a logical partition which has been dynamically soft-capped by the WLM.

The Absolute Capping allows you to change the absolute capping of logical partitions in a group that share processors. The absolute capping can be None or a number of processors value from 0.01 to 255.0. To change an absolute capping for a processor type for a group, select the current absolute capping setting in its field and click the hyperlink to display the next Edit Absolute Capping window. Specify the absolute capping for the selected processor type to indicate the new setting.

Figure 53. Change Logical Partition Controls task

Figure 54. Edit absolute capping

Reviewing status of Secure Execution

Secure Execution protects data confidentiality and integrity when running multiple Linux virtual machines in a logical partition. A system is enabled for Secure Execution when the feature code is installed on the Support Element. An LPAR opts-in to use the Secure Execution support. The **Image Details** task displays the status of Secure Execution for a specific logical partition.

Figure 55. Determining Secure Execution status

LP01 Details - LP01
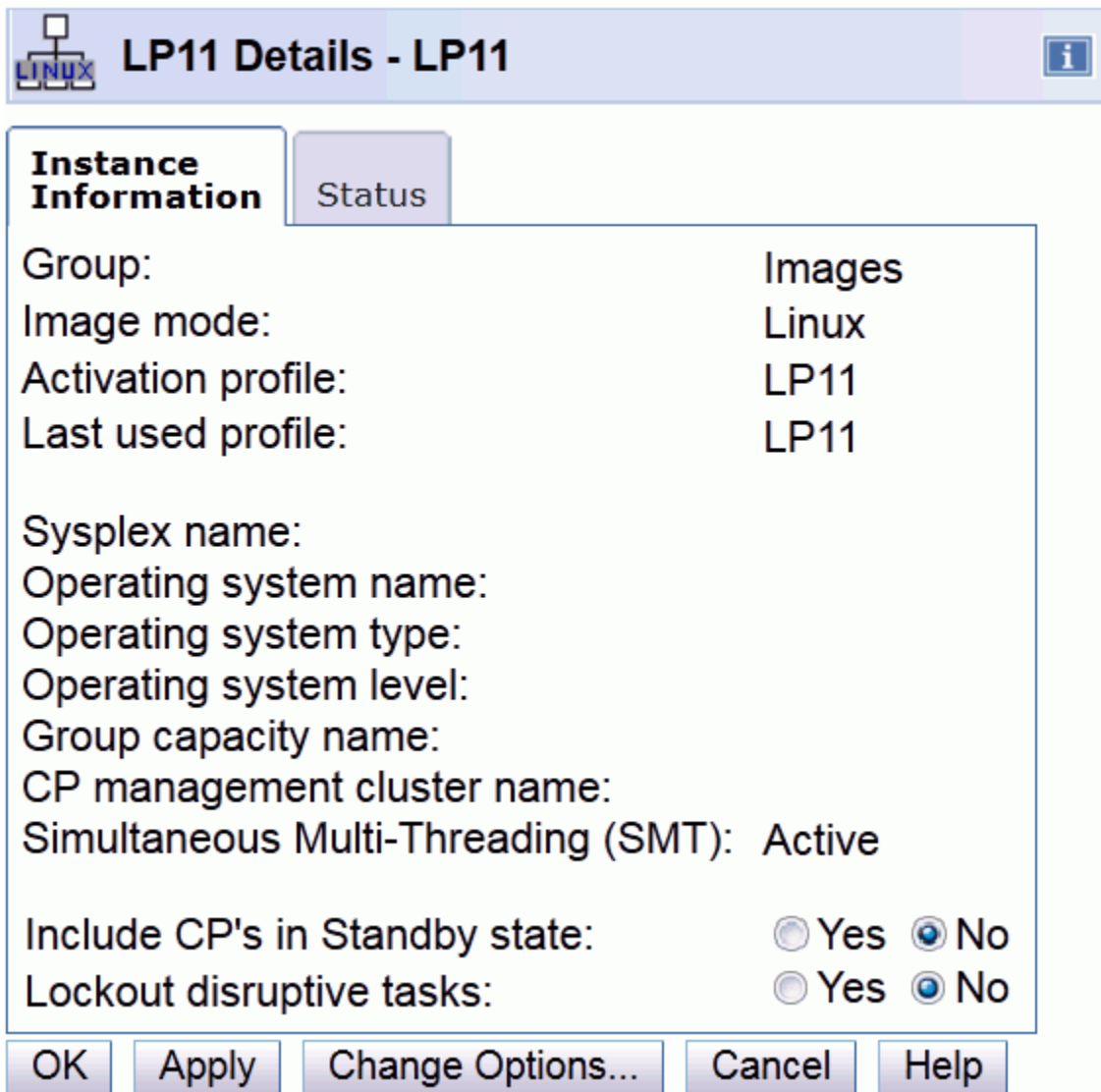
Instance Information | Status

Group:	Images
Image mode:	Not set
Activation profile:	LP01
Last used profile:	LP01
Sysplex name:	
Secure Execution for Linux:	Off
Operating system name:	
Operating system type:	
Operating system level:	
Group capacity name:	
CP management cluster name:	
Simultaneous Multi-Threading (SMT):	Inactive
Include CP's in Standby state:	○ Yes ● No
Lockout disruptive tasks:	○ Yes ● No

OK | Apply | Change Options... | Cancel | Help

Reviewing status of Simultaneous Multi-Threading (SMT)

Simultaneous Multi-Threading (SMT) is for workloads of the Integrated Facility for Linux (IFL) and the IBM z Integrated Information Processors (zIIP). Some logical partitions may be running with SMT enabled and other may not. Use the **Image Details** task to determine if SMT is enabled on your system.
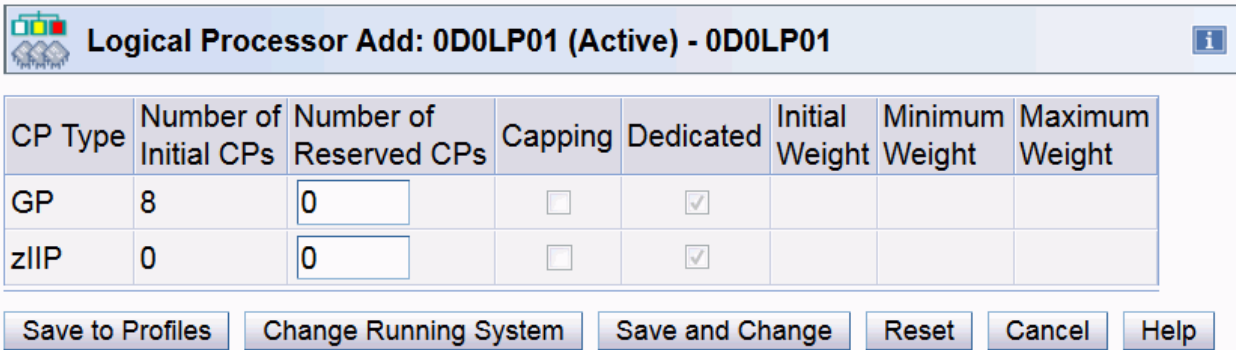
Figure 56. Determining Simultaneous Multi-Threading

Reviewing and adding logical processors

Use the **Logical Processor Add** task to select logical processor definitions to be changed dynamically on the system, in the image profile, or both. Dynamic changes will take effect without performing a reactivation of the logical partition. This tasks allows you to:

- Increase the initial and/or reserved values for installed logical processor type(s)

- Add a reserved value and set weight and capping indicators for logical processor type(s) that have not yet been installed and have no reserved CPs defined

- Increase the reserved value for logical processor type(s) that have not been installed and already have reserved CP(s) defined

The partition status (active and inactive) is indicated in the window title, along with the logical partition name. If the logical partition is active, the current settings are displayed. If the logical partition is inactive, the settings contained in the image profile displays.
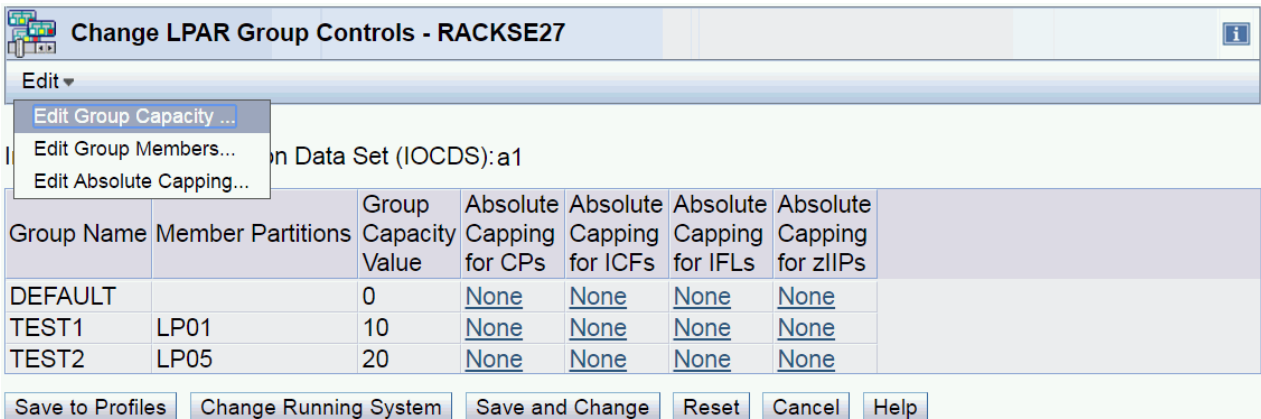
Figure 57. Logical Processor Add task

Reviewing and changing current logical partition group controls

Use the **Change LPAR Group Controls** task to define LP group capacity limits or edit the absolute capping of logical partitions. The group capacity limit allows you to specify one or more groups of LPs on a server, each with its own capacity limit. This is designed to allow z/OS to manage the groups that the sum of the LPs' CPU utilization within a group will not exceed the group's defined capacity. Each Logical partition in a group can still optionally continue to define an individual logical partition capacity limit.

The LP group capacity limits may help provision a portion of a Z mainframe to a group of logical partitions allowing the CPU resources to float more readily between those logical partitions, resulting in productive use of *white space* and higher server utilization.

The Absolute Capping allows you to change the absolute capping of logical partitions in a group that share processors. The absolute capping can be **None** or a number of processors value from 0.01 to 255.0. To change an absolute capping for a processor type for a group, select the current absolute capping setting in its field and click the hyperlink to display the next Edit Absolute Capping window. Specify the absolute capping for the selected processor type to indicate the new setting. Absolute capping is managed by PR/SM, independent of the operating system running in the capped logical partitions. This can be useful for capping a collection of partitions where z/OS managed group capacity limits cannot be used.
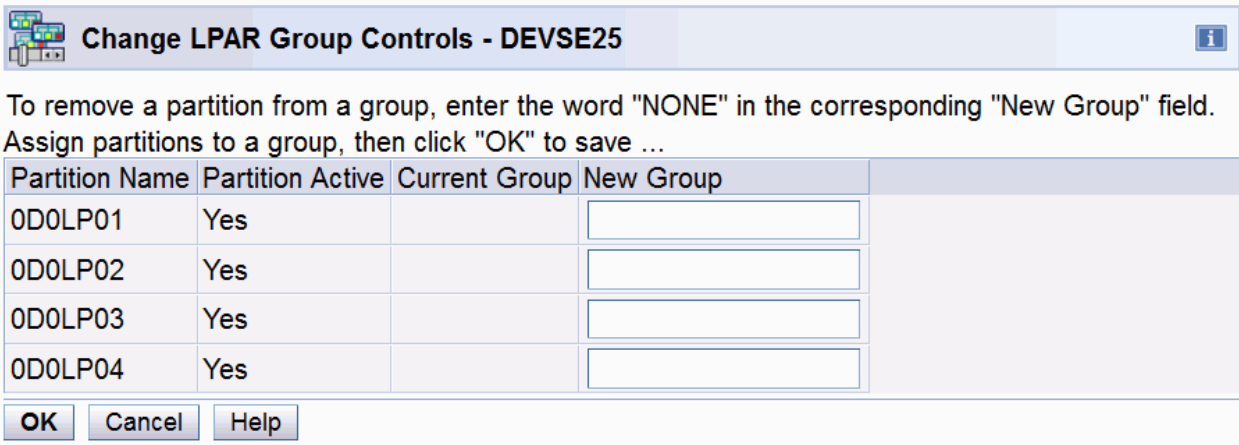
Figure 58. Change LPAR Group Controls task
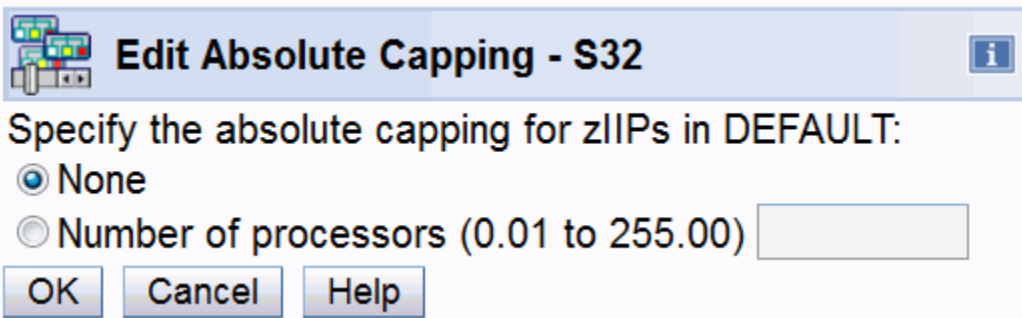
Figure 59. Edit group members

Figure 60. Edit absolute capping

For information about creating a group, see "Creating a logical partition group profile" on page 155.

For information about how workload management and workload license charges relate to the Group Capacity setting, see *z/OS MVS Planning: Workload Management*, SA22-7602.

Reviewing and changing current logical partition security

Use the **Change Logical Partition Security** task to display and modify security controls for an LP.

Figure 61. Change logical partition security task

Figure 62. Configure logical partition BCPii permissions

Figure 63. Add partition to receive BCPii commands from the active logical partition

Reviewing and changing current logical partition cryptographic controls

View LPAR cryptographic controls

Use the **View LPAR Cryptographic Controls** task to display crypto characteristics for an LP. The **Summary** tab displays the current crypto configuration settings for all active and inactive partitions in the CPC. (The inactive partition information displayed is a result of the settings selected in the Image Activation profile.) The tab with the name of the partition displays the current crypto configuration for that active partition.
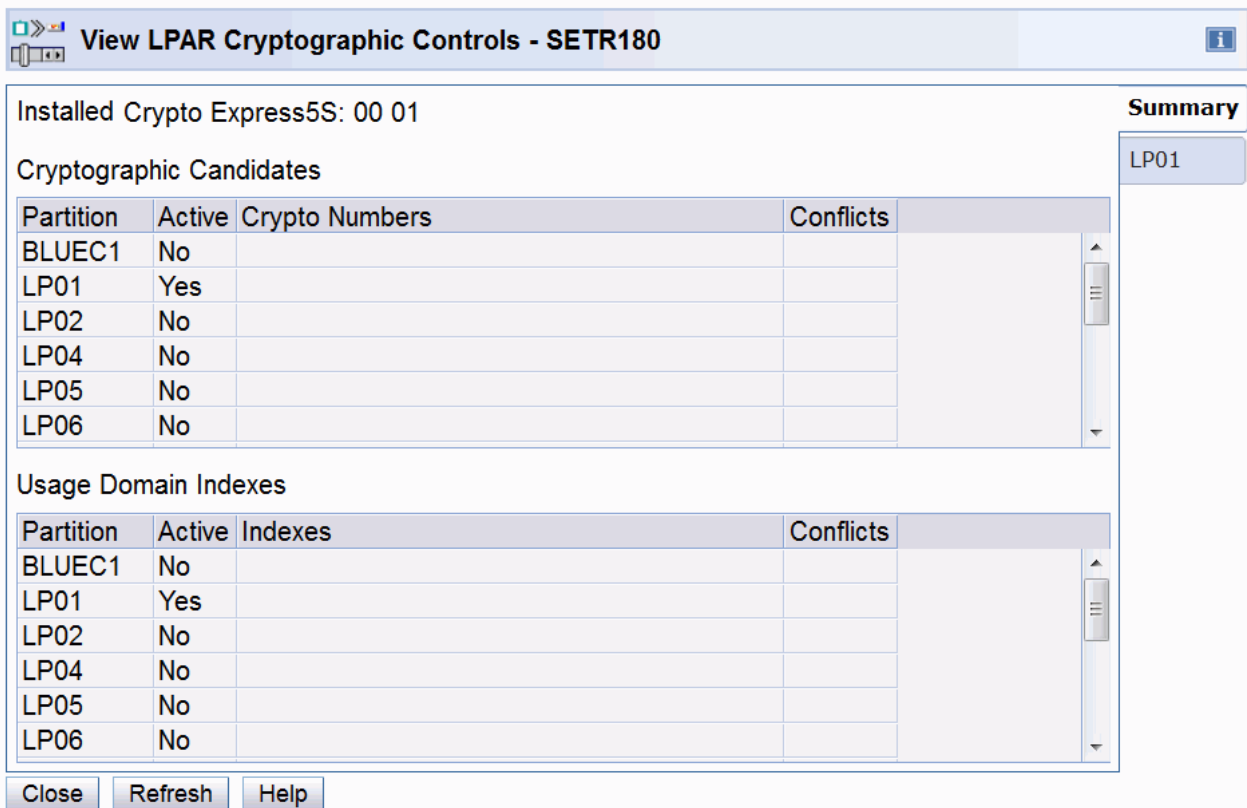
Figure 64. View LPAR cryptographic controls window (summary tab)
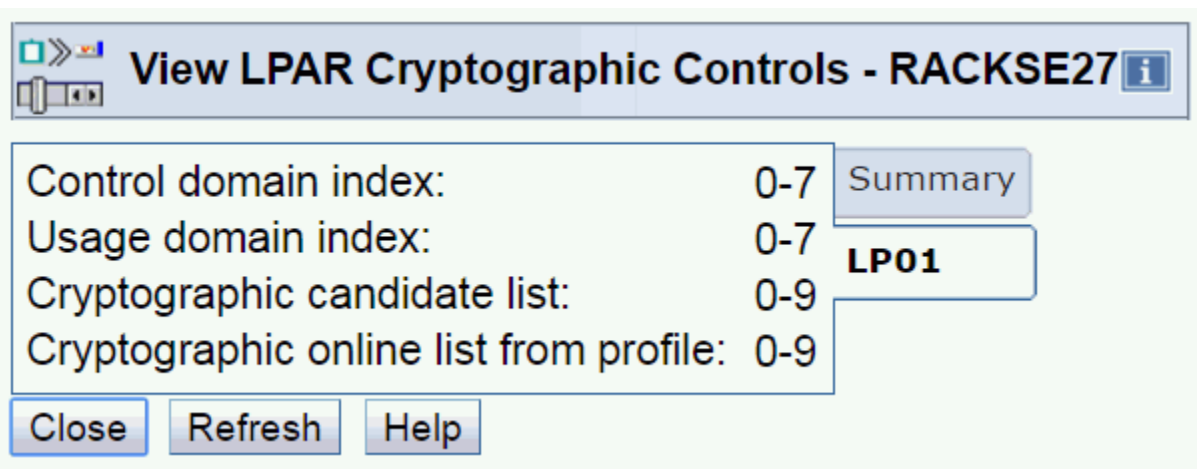
Figure 65. View LPAR cryptographic controls (showing tab containing crypto configuration information for an active partition)

Changing LPAR cryptographic controls

Use the **Change LPAR Cryptographic Controls** task to make changes to the crypto configuration of an active partition without affecting the operating status of the partition. This allows you to update your crypto configuration without reactivating the logical partition. You can add cryptos to a partition, delete cryptos from a partition, and/or move a crypto from one partition to another using the following task:

Figure 66. Change LPAR Cryptographic Controls task

There are requirements for adding, removing and moving a crypto:

- The crypto that is assigned to this partition must be configured offline before the removal of the crypto from the partition can be done. If the crypto was operating as a coprocessor, and a removal of a crypto is being done, the user is given the opportunity to remove the cryptographic keys from the partition associated with the selected usage domain, using the usage domain zeroize function.

- If the crypto is not online to the partition, but the associated PCHID is online and operating, the usage domain zeroize action can immediately be done. If the crypto assigned to the partition and the PCHID are both offline, the usage domain zeroize action will be pending until the next time this crypto is brought online.

- To move a crypto from one partition to another requires you to perform two steps.

  - Remove the crypto from the first partition.
  - Then, add it to the second partition.

- After a crypto is added to a partition, the crypto needs to be brought online using the **Configure On/Off** task.

Note: Changes made using the **Change LPAR Cryptographic Controls** task can be made to both active and inactive partitions. When performed on an inactive partition, the changes are made to the image activation profile, since the partition is not active.

The cryptographic assigned domains table displays the current settings of the usage domain indexes and control domain indexes which can be modified in the logical partition and the image profile.

Control Domain
    A logical partition's control domains are those cryptographic domains for which remote secure administration functions can be established and administered from this logical partition. This logical

partition's control domains must include its usage domains. For each index selected in the usage domain index list, you must select the same index in the control domain index list.

Control and Usage Domain

A logical partition's usage domains are domains in the cryptos that can be used for cryptographic functions. The usage domains cannot be removed if they are in use by the partition. The usage domains you select for this logical partition must also be selected in the control domain index.

The assigned crypto table displays the current settings of the cryptographic candidate list and cryptographic online list settings which can be modified in the logical partition and/or the image profile:

Candidate

The candidate list identifies which cryptos are eligible to be assigned to the active logical partition. If a card is not installed in the slot, it will not be available to the logical partition. However, if a card is installed in a slot specified in the candidate list, it can immediately be made available to the logical partition.

Candidate and Online

The online list identifies which cryptos will be brought online at the next activation. Changes to the online list do not affect the running system.

To commit your changes, use one of the following:

Save to Profiles

Select this if you want new settings to take effect whenever the logical partition is activated with the modified profile. This changes the cryptographic settings in the logical partition's image profile. The settings take effect whenever the logical partition is activated with its image profile.

Change Running System

Select this if you want the new settings to take effect in the active logical partition immediately. This changes the cryptographic settings in the logical partition without reactivating the partition. The new settings remain in effect for the logical partition until you either dynamically change the settings again or reactivate the partition.

Note: This button can be used for an active logical partition only. For an inactive partition, this button is disabled.

Save and Change

Select this if you want the new settings to take effect immediately and whenever the logical partition is activated with the modified profile. **Save and Change**:

- Saves a logical partition's cryptographic control settings in its image profile. The settings take effect whenever the logical partition is activated with its image profile.

- Changes the cryptographic settings in the logical partition without reactivating the partition. The new settings remain in effect for the logical partition until you either dynamically change the settings again or reactivate the partition.

Note: This button can be used for an active logical partition only. For an inactive partition, this button is disabled.

Reset

Select this to revert the settings back to their original values.

When a crypto with it's associated usage domains are removed from a partition, this partition no longer has access to the cryptographic keys. But if this crypto is then assigned to a different partition utilizing the same usage domains as before, then this new partition will have access, possibly unintentional access, to the cryptographic keys. Therefore, when a crypto is removed from an active partition, the Usage Domain Zeroize window is displayed, providing the opportunity to clear the cryptographic keys associated with the given usage domain(s).
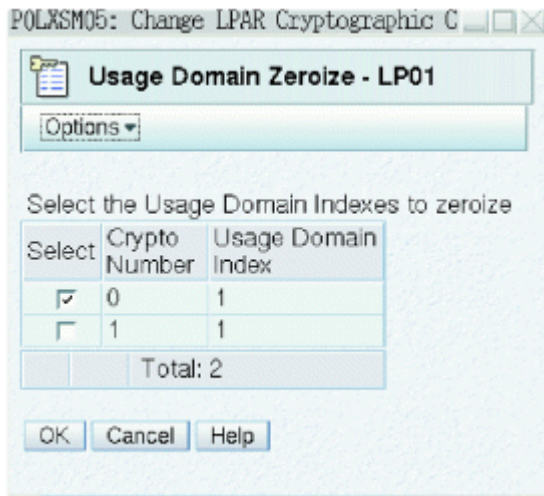
Figure 67. Usage domain zeroize
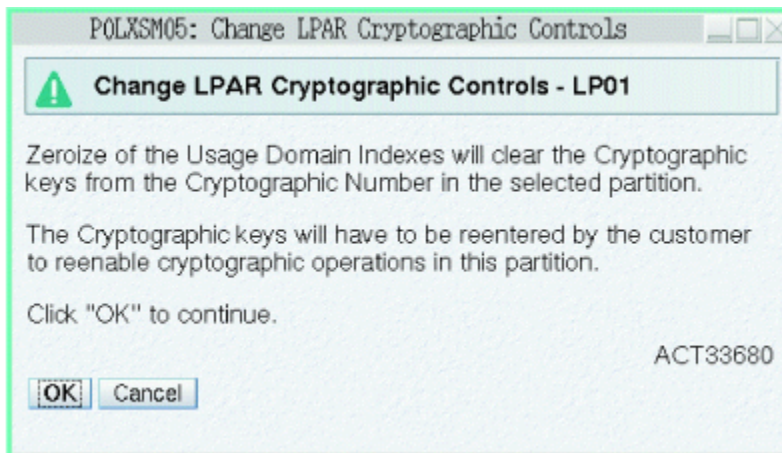
Figure 68. Message received from change LPAR cryptographic controls

Note: If the crypto device you remove is the last remaining one, a caution displays that all cryptographic candidates have been removed from the partition, which removes the partition's access to all cryptos.

Cryptographic configuration

The opportunity to clear or zeroize the cryptographic keys associated with a usage domain is available when removing a crypto using the **Change LPAR Cryptographic Controls** task. The zeroize of the cryptographic keys can also be performed using the **Cryptographic Configuration** task.
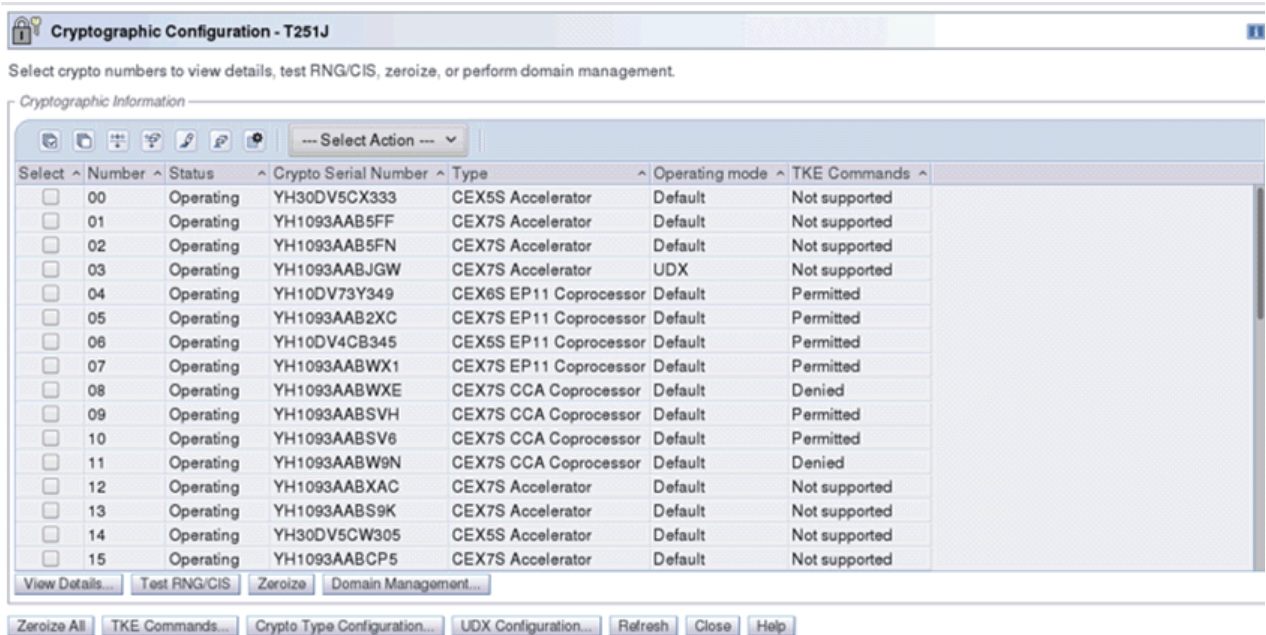
Figure 69. Cryptographic configuration window

The **Cryptographic Configuration** task can be used to configure and monitor the cryptos that are installed on your system. Use this window to perform various tasks, including:

- Checking the status and details of the Crypto Express feature
- Testing the Random Number (RN) generator of the CCA Coprocessor
- Manually zeroizing cryptos and erasing configuration data from the Support Element
- If you select one or more cryptographic devices and click **Domain Management**, a window displays to select the specific usage domains to clear the master keys. If you click **Zeroize**, the master keys will be cleared for all domains on the selected cards. If you click **Zeroize All** the master keys are cleared for all devices that are configured or online for the CPC.

   Note: The **Zeroize** and **Zeroize All** reset the TKE commands flag to Not Supported. If you want to continue to manage a crypto adapter using TKE, you will need to click **TKE Commands**.

- Run Customer Initiated Selftest of the EP11 Coprocessor.
- Manually clear the cryptographic keys within the given usage domain(s).
- Import and activate a UDX file configuration.
- Indicate whether to permit TKE commands for processing on the selected CCA Coprocessor.
- Indicate the crypto type configuration for the Crypto Express feature.

Note: With Crypto Express, Segment 3 of the Common Cryptographic Architecture (CCA) supports Concurrent Driver Upgrade (CDU). In other words, Segment 3 of the CCA can be updated dynamically, and therefore does not require a configure on/off to activate these changes. (However, there may be cases where an update of Segment 3 CCA may be disruptive, these will be identified in the documentation for the MCL). If Segment 1 or 2 is updated, the crypto must be restarted (for example, Configure Off/On) in order to utilize the firmware updates.

Usage domain zeroize

You can clear or zeroize the cryptographic keys associated with an available usage domain when removing a crypto using the **Change LPAR Cryptographic Controls** task. Perform zeroize of the cryptographic keys using the **Cryptographic Configuration** task.
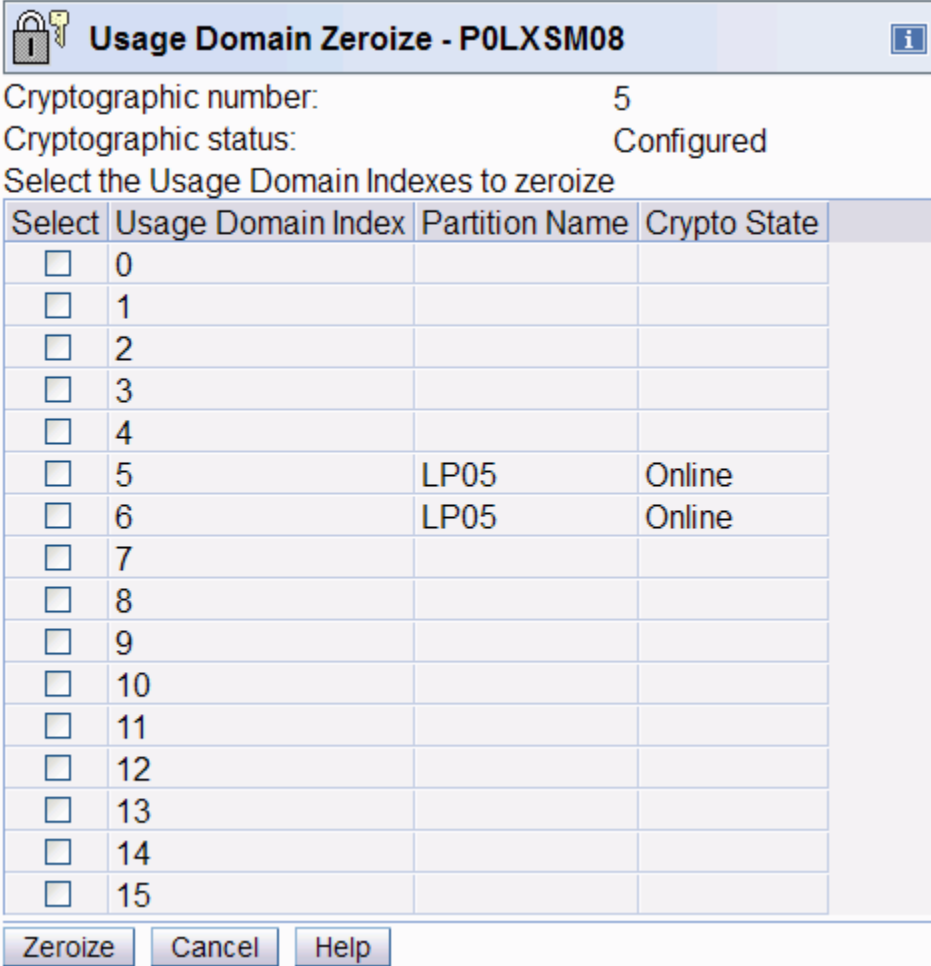
The removal of a crypto from a logical partition could make those cryptographic keys available to another LP, if the crypto and usage domains are then reassigned to a new partition. This can occur when:

- Removing the crypto from the candidate list in the Activation profile, then reassigning the crypto to another active partition.
- Moving a crypto using the **Change LPAR Cryptographic Controls** task.

The Usage Domain Zeroize in the **Cryptographic Configuration** task offers the opportunity to clear the cryptographic keys when desired, not just when the cryptographic settings are modified using the **Change LPAR Cryptographic Controls** task.

It is recommended that the usage domain zeroize be performed with the crypto offline, but it is not required. When performing the usage domain zeroize with the crypto offline, the zeroize of the usage domain index(es) selected is deferred until the selected crypto is configured online, using the **Configure On/Off** task.

On the Cryptographic Configuration window, select the crypto and click **Domain Management**. The Usage Domain Zeroize window displays:

Figure 70. Usage domain zeroize window

On the Usage Domain Zeroize window, select the usage domain index(es) that have the cryptographic keys you wish to zeroize, then click **Zeroize**. The zeroize of the Usage Domain Index(es) is deferred until the selected Cryptographic number is configured online (Configure On). When the selected cryptographic number is configured online, the selected Usage Domain indexes is cleared of cryptographic keys.

Crypto type configuration

This window displays what configuration type for the selected Crypto Express feature is currently configured on your system. The Crypto Express feature must be deconfigured prior to changing the crypto configuration type and it must be deactivated from any LP using it before it can be deconfigured. Specify the crypto configuration type for the Crypto Express feature installed in your system. If changing from a

CCA Coprocessor to an accelerator, you can zeroize the cryptographic keys in the CCA Coprocessor when the crypto is configured online.

On the Cryptographic Configuration window, select the crypto and click **Crypto Type Configuration**. The Crypto Type Configuration window displays:

For a Crypto Express feature select:

• CCA Coprocessor

• EP11 Coprocessor

• Accelerator
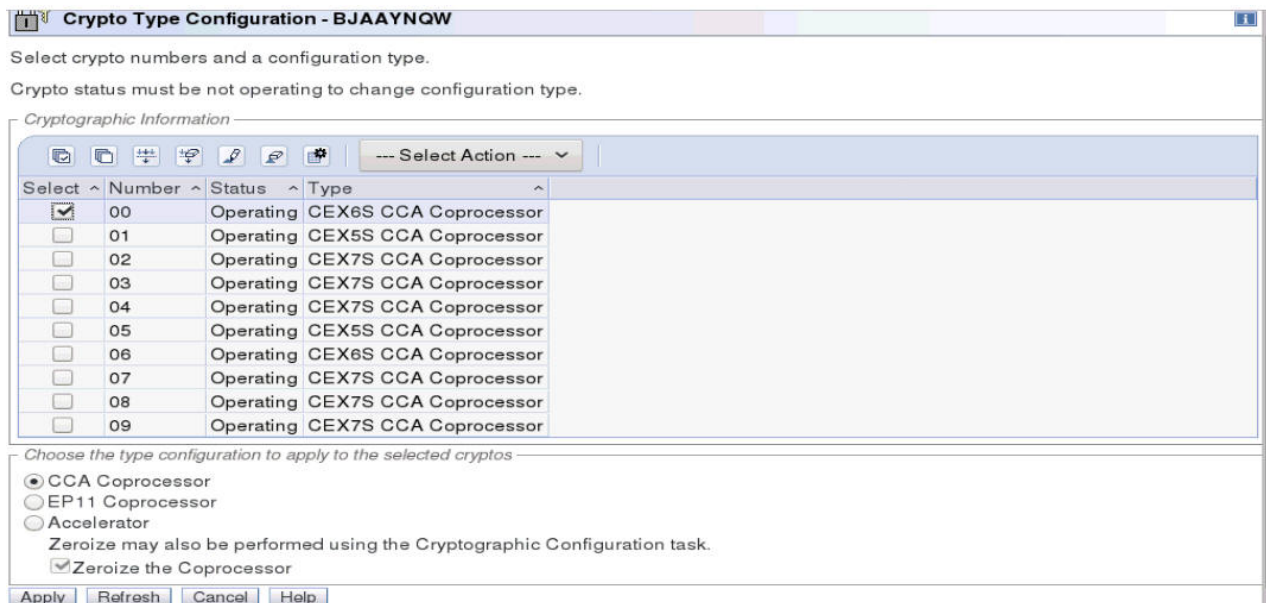
Figure 71. Crypto type configuration window

Note: The TKE Workstation is required for key management of the EP11 Coprocessor.

Reviewing and changing logical partition I/O priority values

Use the **Change Logical Partition I/O Priority Queuing** task available to display and modify I/O priority queuing values for one or more LPs.

Figure 72. Change Logical Partition I/O priority queuing window

Logical partition performance

Contact your support system for assistance in planning configuration capacity requirements.

The performance of LPs depends on:

- The operating systems and their workloads
- The configuration of the LPs
- The configuration of the CPC
- The performance tuning parameters

You can contact an service representative who has access to a proprietary performance planning tool (LPARCE from CPSTOOLS at WSCVM) to assist you in this task.

RMF LPAR management time reporting

RMF processor utilization reporting includes LPAR Management Time on the Partition Data Report. RMF provides information about all shared logical cores that remain active for the duration of the reporting interval. RMF can also be run on a single LP if additional information is required.

There are two types of LPAR management time reporting: time spent managing the LP, and time spent managing the physical configuration. With LPAR Management Time reporting, the time used to manage an LP can be separated from the time used by the workload in an LP.
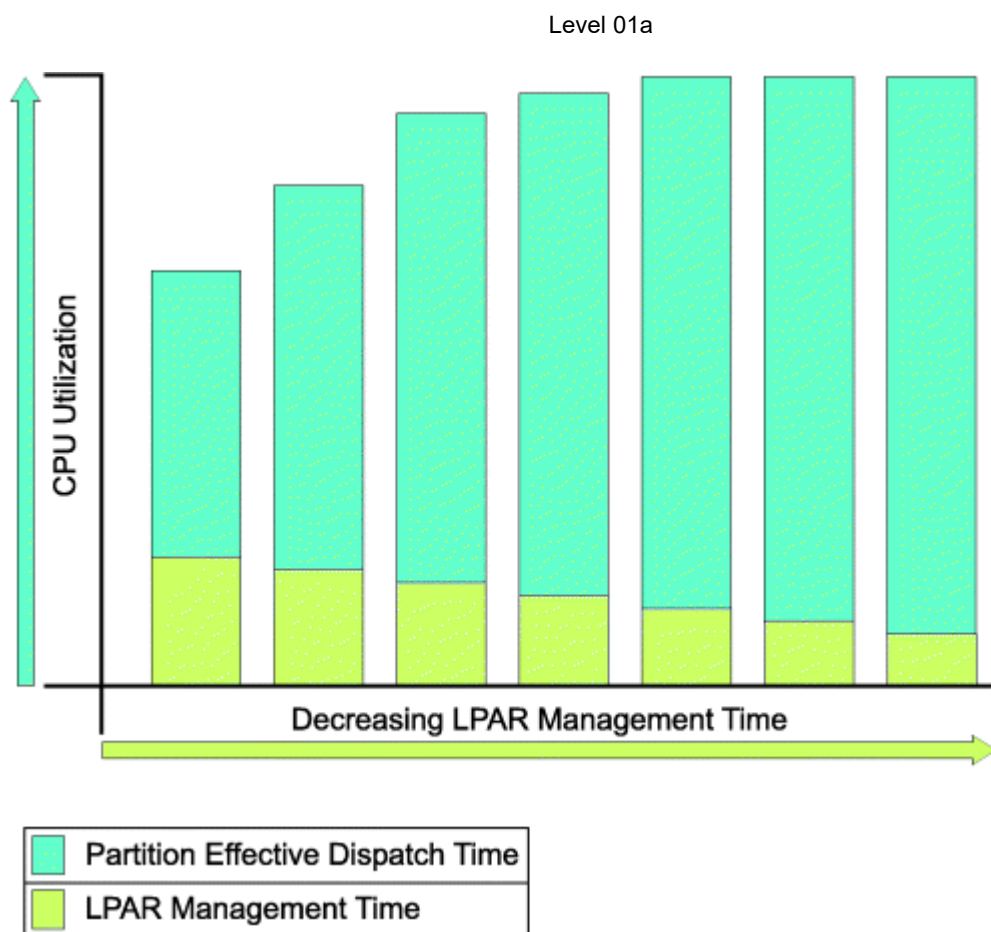
Figure 73. ETR increasing with CPU utilization

is an example of how information from the enhanced RMF LPAR Management Time reports can be used. This example shows the LPAR Management Time and the partition effective dispatch time for each LP. As CPU utilization increases, LPAR management time decreases, and the external throughput rate (ETR) increases.

Using the RMF report, CPU-utilization calculations can be based on an LP's effective dispatch time. The effective dispatch time can be determined by excluding LPAR Management Time from the time that a physical core is assigned to a logical core.

Note: For more information about RMF support, the RMF Partition Data Report, and LPAR Management Time see the *Resource Measurement Facility User's Guide*.

Dedicated and shared central processors

Generally, LPs with dedicated CPs require fewer processing resources. The internal throughput rate (ITR) of LPs with dedicated CPs is higher than that of identically defined LPs with shared CPs. It is recommended that dedicated processors (ICFs or CPs) be used in a coupling facility partition that is used in a production configuration.

Generally, the capability to recognize CP resources that are not used and to balance workloads can result in an improved ETR for LPs with shared CPs. In particular, for workloads that exhibit fluctuations in processing demands, the consolidation of several systems as LPs with shared CPs on a CPC can improve responsiveness and increase ETR.

CPENABLE

The z/OS CPENABLE parameter can be used to provide selective enablement for I/O interrupts. On the Z mainframe, the best ITR is achieved when the fewest CPs are enabled for I/O interrupts. Selective enablement for I/O and the CPENABLE parameter are described in the *z/OS MVS Initialization and Tuning Reference*, SA22-7592.

For logical partitions using dedicated processors, it is recommended that a CPENABLE=(10,30) setting to minimize the number of CPs handling I/O interrupts in an LP.

A CPENABLE=(0,0) setting, enabling all CPs to handle I/O interrupts, can cause performance degradation for logical partitions with dedicated CPs.

For logical partitions sharing CP resources, it is recommended that a CPENABLE=(10,30) setting. There is a negligible ITR impact between a CPENABLE=(0,0) setting versus a CPENABLE=(10,30) setting for logical partitions sharing CP resources.

Start Interpretive Execution (SIE) performance

PR/SM implementation uses CPC hardware mechanisms that are also used to implement the SIE facility when operating in **General** mode. Therefore, if SIE is executed in an LP, SIE performance is reduced relative to operation in **General** mode. The performance reduction depends on the frequency of SIE invocation and should be evaluated for each application.

Recovery strategy

Recovery planning requires that the appropriate planners and technical support personnel understand the recovery strategy.

In planning for recovery, consider the following guidelines:

- Recovery is considered successful if an LP is able to perform useful work and critical application programs remain operational (even if one or more LPs are disabled) after a failure occurs. Recoverability depends on the extent and location of a failure and the ability of the operating system to handle the failure.

- Develop a recovery strategy for the specific installation that addresses the specific hardware and applications of that installation. For example, if a failure occurs that might disrupt multiple LPs or a critical LP, the operator should be able to determine what must remain active and what can be deactivated, and to perform the appropriate recovery procedures.

- The operator should follow established local procedures for reporting problems and for performing recovery procedures. It is recommended that recovery procedures be ranked from least to most disruptive. The operator should know what procedures to follow if any or all LPs do not respond to the recovery actions directed to it.

- Assign channel paths to LPs as described in the guidelines under "Guidelines for setting up the I/O configuration" on page 37.

- Define resources to LPs so that any hardware failure has a minimal impact on any LP that must remain active.

  For example, the failure of a physical core causes the temporary loss of the logical core that was dispatched on the physical core. The LP owning that logical core may continue running if it was running on an LP with at least two CPs and if the operating system can recover from a CP failure.

Operation considerations

If an individual LP is affected, the following recovery actions (ranked from least to most disruptive) should be considered when planning recovery procedures.

- If an affected LP continues to operate with a problem, allow it to do so.

- If the operating system in an LP remains active but processing is interrupted, consider the applicability of a restart or IPL.

- Perform a stand-alone dump and IPL the operating system.

- Perform an orderly shutdown of the operating system in an affected LP. If the LP is not critical, allow the other LPs to continue.

If all LPs are affected, or if a critical LP is affected and did not respond to the recovery actions directed only to it, the following recovery actions (also ranked from least to most disruptive) should be considered when planning recovery procedures.

- Perform an orderly shutdown of all the operating systems. Activate the LPs and IPL the operating systems.
- Perform an orderly shutdown of all the operating systems. Perform the most appropriate recovery action (for example, in response to a hardware failure). Perform a power-on reset.
- If a power-on reset fails to initialize the LPs, perform a power-on reset to attempt a recovery. IPL the most critical operating system.

Application preservation

The application preservation facility enhances system availability and provides additional protection from CP failures. This support is available on models with two or more central processors (CPs). Using application preservation, the system moves an application in process from a failing CP to another operating CP. Both the CP failure and the move are transparent to the application.

There are no software corequisites when running with shared CPs. For LPs using shared CPs, even 1-way LPs can survive a CP failure without experiencing a failure to the application or the LP, providing the 1-way LP is being run on a model with more than one physical core.

Transparent sparing

Transparent sparing takes LP recovery one step further by combining the benefits of application preservation and concurrent CP sparing to allow for the transparent recovery of an LP and its applications (including CF LPs). Transparent sparing uses hardware to handle the recovery, requires no software support or operator intervention, and is effective for both shared and dedicated LP environments.

Transparent sparing configures a spare PU (processor unit) to replace a failed CP, ICF, IFL, SAP, or zIIP. Each of the z15 models has 2 spare PUs.

Appendix A. Coupling facility control code support

Coupling facility control code is Licensed Internal Code (LIC) that supports the following coupling facility limits:

Table 21. Coupling facility limits at different coupling facility code levels

Coupling Facility Limit	Coupling Facility Code Level							
	Level 24	Level 23	Level 22	Level 21	Level 20	Level 19	Level 18	Level 17
Maximum number of CPs	16	16	16	16	16	16	16	16
Storage increment	1 MB	1 MB	1 MB	1 MB	1 MB	1 MB	1 MB	1 MB
Structure ID limit	2047	2047	2047	2047	2047	2047	2047	2047
Retry buffer limit	1799	1799	1799	1799	1799	1799	1799	1799
Facility information	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes	64 bytes
Maximum list element characteristic	4	4	4	4	4	4	4	4
Maximum lock table entry characteristic	5	5	5	5	5	5	5	5
User identifier limit	255	255	255	255	255	255	255	255
Maximum data area element characteristic	4	4	4	4	4	4	4	4
Local cache identifier limit	255	255	255	255	255	255	255	255
Storage class limit	63	63	63	63	63	63	63	63

Coupling Facility Limit	Coupling Facility Code Level							
	Level 24	Level 23	Level 22	Level 21	Level 20	Level 19	Level 18	Level 17
Castout class limit	1024	1024	1024	1024	1024	1024	1024	1024
Notification-Delay Limit (NDL)	X'FFFFFFFF'	X'FFFFFFFF'	X'FFFFFFFF'	X'FFFFFFFF'	X'FFFFFFFF'	X'FFFFFFFF'`	X'FFFFFFFF'	X'FFFFFFFF'

Table 21. Coupling facility limits at different coupling facility code levels (continued)

Legend

Maximum number of CPs
Indicates the maximum number of CPs that can be used by a coupling facility logical partition.

Storage increment
Indicates the granularity with which storage allocation requests are rounded, to the amount shown in the table for a particular CFCC level.

Structure ID limit
Cache and list structure ID. Effectively, this limit defines the maximum number of coupling facility structure instances that a coupling facility at this level may contain.

Retry buffer limit
Retry buffer range upper limit.

Facility information
This area contains coupling facility control code release and service level information

Maximum list element characteristic
The size of a list structure list element in bytes equals 256 * (2 ** list element characteristic), for example, 256 * (2**4) = 4K.

Maximum lock table entry characteristic
The size of a lock table entry in bytes equals 2 ** lock table entry characteristic.

User identifier limit
The maximum number of users and list notification vectors that can be attached to a list structure.

Maximum data area element characteristic
The size of a data area element in bytes equals 256 * (2 ** data area element characteristic, for example, 256 * (2**4) = 4K).

Local cache identifier limit
The maximum number of local caches that can be attached to a cache structure.

Storage class limit
Storage classes are in the range 1 to the value shown in the table.

Castout class limit
Castout classes are in the range 1 to the value shown in the table.

Notification-Delay Limit (NDL)
The maximum time delay value that can be specified for a list structure's notification delay. The notification delay is the time delay between the initial notification performed to a single selected user in response to a monitored object (sublist, list, or key-range) becoming non-empty, and the subsequent notification(s) that are performed to the remaining monitoring users.

Appendix B. Developing, building, and delivering a certified system

This appendix is intended to provide guidance in setting up, operating, and managing a secure environment using Z mainframe PR/SM when the CPC is operating in logically partitioned (LPAR) mode. It is primarily for the security administrator, but can also be useful to other involved operations technical support personnel.

Creating Common Criteria-Based evaluations

In October 1998, after two years of intense negotiations, government organizations from the United States, Canada, France, Germany, and the United Kingdom signed a historic mutual recognition arrangement for Common Criteria-based evaluations. This arrangement, officially known as the Arrangement of the Recognition of Common Criteria Certificates in the field of IT Security, was a significant step forward for government and industry in the area of IT product and protection profile security evaluations. The partners in the arrangement share the following objectives in the area of Common Criteria-based evaluation of IT products and protection profiles:

- To help ensure that evaluations of IT products and protection profiles are performed to high and consistent standards and are seen to contribute significantly to confidence in the security of those products and profiles
- To increase the availability of evaluated, security-enhanced IT products and protection profiles for national use
- To eliminate duplicate evaluations of IT products and protection profiles, and
- To continuously improve the efficiency and cost-effectiveness of security evaluations and the certification/validation process for IT products and protection profiles.

The purpose of this arrangement is to advance those objectives by bringing about a situation in which IT products and protection profiles which earn a Common Criteria certificate can be procured or used without the need for them to be evaluated and certified/validated again. It seeks to provide grounds for confidence in the reliability of the judgment on which the original certificate was based by declaring that the Certification/Validation Body associated with a Participant to the Arrangement shall meet high and consistent standards. The Arrangement specifies the conditions by which each Participant will accept or recognize the results of IT security evaluations and the associated certifications/validations conducted by other Participants and to provide for other related cooperative activities.

The PR/SM functionality and assurances have been evaluated and certified at an EAL5 level of assurance. This assurance enables PR/SM to meet stringent requirements for confidentiality of processed information including requirements mandated by the federal government and the banking industry.

The Certification/Validation Body which performs the evaluations of PR/SM is Bundesamt Fuer Sicherheit Informationstechnik (BSI). The BSI issued certificate IDs for the most recent PR/SM evaluations are: BSI-DSZ-CC-1048, BSI-DSZ-CC-1101, BSI-DSZ-CC-1109, and BSI-DSZ-CC-1133. Additional information, including the Security Target forming the base document for the evaluation is available at the BSI website: http://www.bsi.bund.de/EN/TheBSI/thebsi_node.html.

This appendix must be used in conjunction with other pertinent manuals supplied with the IBM Z mainframe to give a security administrator all the required information to configure and operate a logically partitioned (LPAR) mode system in a secure manner. This appendix provides instruction on the correct use of the system so that a secure environment is created and maintained. It defines and explains the parameters, settings, and commands recommended, including references to those sections in the manuals being discussed in "Trusted facility library" on page 195.

Functional characteristics

PR/SM is a cornerstone of IBM's server security. PR/SM's logical partitioning facility enables the resources of a single physical Z mainframe to be divided and shared by distinct logical machines, each capable of running its own operating system.

The security administrator can configure one or more distinct logical machines to ensure complete isolation from one another; one logical machine cannot gain knowledge about any other logical machine's available I/O resources or performed operations. Logical Partitions configured in this manner will be referred to as *Isolated Logical Partitions* throughout the remainder of this appendix.

A single physical Z mainframe allows any combination of Isolated and non-Isolated logical partitions to be configured. The non-Isolated logical partitions can be configured in any manner supported by Z mainframes. Any level of sharing or cooperation among the non-Isolated logical partitions (for example, Parallel Sysplex) is permitted and will not have any impact on the Isolated logical partitions.

Logical partitions are defined, and the I/O resources of the overall physical computing system are pre-allocated by the security administrator. I/O allocation is an integral part of the process of defining a total system configuration, and must be completely performed before that system configuration can be initialized. This preallocation is done by executing the Input/Output Configuration Program (IOCP) or Hardware Configuration Definition (HCD) to create a hardware-specific data set, called an Input/Output Configuration Data Set (IOCDS), of the I/O resources and their allocation to specific logical partitions. PR/SM allocates an entire resource, such as an I/O channel path or a contiguous region of storage. At no time is any real resource allocated to more than one Isolated logical partition. Each complete I/O resource allocation is called a *configuration*. During the period between processor initialization, several IOCDS configurations can be stored, but only one is in effect at any time. The configuration becomes effective as part of the power-on reset sequence. In order to change the active configuration it is necessary to perform an activation of the hardware.

The preceding paragraph deliberately omits any discussion of Dynamic I/O Configuration, Dynamic CHPID management, Reconfigurable channel paths (CHPIDs), I/O resource sharing using Multiple Image Facility (MIF) or Intelligent Resource Director (IRD), because each of them has characteristics that, if inappropriately used, can compromise the secure capability of PR/SM. Cautions and requirements relating to their use are included throughout this appendix.

The remainder of the logical partition's resources are defined by the security administrator prior to the activation of the logical partition. These resources include storage size, number of logical processors, scheduling parameters, and security controls, which can be specified by the security administrator using the appropriate interfaces on the Hardware Management Console and Support Element. Many of the control and security parameters can be changed at any time and takes effect dynamically with few exceptions (for example, specifying dedicated processors for a partition will only take effect if the partition is not yet activated.) Logical partition definitions take effect at logical partition activation, and generally are static while the partition they pertain to is active.

When a resource is allocated to a logical partition, it is set to its architecturally-defined reset state. Channel paths are reset, main storage is zeroed.

Trusted configuration

This section describes the actions the Security Administrator must take to help ensure that the computer system is configured for a secure mode of operation. The contents of this section specify the configuration of the evaluated product. Any deviation from the specified configuration will not be consistent with that of the evaluated product and may result in partitions that do not provide strict separation.

Subsequent sections in this document detail the security related characteristics of the evaluated product as well as security configurations that were not included in the evaluation. These details are provided to explain and highlight the differences between the various security settings. Nevertheless, to insure strict separation of Isolated logical partitions, only the configuration specified in this section should be used.

The Licensed Internal Code level of the evaluated configuration is specified in a Common Criteria related document called the Security Target. The installed LIC level for a HMC/CPC can be determined via the **View Console Information/System Information** task available in the HMC Management/Change Management Task List of the HMC/SE. A User ID with its authority based on the default SERVICE User ID must be used to display the complete configuration information.

Note: All configuration requirements listed in subsequent sections are mandatory regardless of whether the term *must* or *should* is used.

- The hardware and any networks used to connect the hardware must be physically secure. Access to I/O devices must be restricted to authorized personnel. The Hardware Management Console must be physically protected from access other than by authorized system administrators.

- The network used for HMC/SE communications should be physically separate from the logical partition data networks.

- Any FTP Servers utilized for Firmware Management must be physically secure and restricted to authorized personnel.

- Devices must be configured so that no device is accessible by any partition other than the partition to be isolated (although they may be accessible by more than one channel path).

- Each I/O (physical) control unit must be allocated to only one Isolated logical partition in the current configuration.

- The Security Administrator must not reconfigure a channel path owned by an Isolated partition unless all attached devices and control units are attached to that path only.

- The Security Administrator should ensure that all devices and control units on a reconfigurable path owned by an Isolated partition are reset before the path is allocated to another partition.

  Note: This reallocation is NOT permitted for any devices/control units which retain customer data after being reset, unless the reallocation is permitted by the customer's security practices.

- No channel paths may be shared between an Isolated partition and any other partition(s).

- Logically partitioned (LPAR) mode must be selected as the mode of operation for the CPC.

- Dynamic I/O Configuration changes must be disabled.

- For Isolated partitions, Workload Manager must be disabled so that CPU and I/O resources are not managed across partitions.

- An Isolated partition must not be configured to enable hipersockets (Internal Queued Direct I/O).

- An Isolated partition must not be configured to enable SMC-L virtual network connections (internal). No Function Identifier (FID) of Type ISM should be specified in the IOCDS.

- Partitions must be prevented from receiving performance data from resources that are not allocated to them (Global Performance Data Control Authority must be disabled).

- At most one partition can have I/O Configuration Control Authority (for example, no more than one partition must be able to update any IOCDS) and this partition must be administered by a trustworthy administrator (i.e. the administrator of this partition is considered to be the Security Administrator). I/O Configuration Control should be enabled for a single, specific logical partition only during the short period of time when it is permitted to write a new IOCDS.

- The Security Administrator must ensure that write access is disabled for each IOCDS, unless that IOCDS is to be updated (the current IOCDS must not be updated).

- The Security Administrator must verify any changed IOCDS after a power-on reset with that IOCDS, before any partitions have been activated (the Security Administrator may determine whether the IOCDS has been changed by inspecting the date of the IOCDS).

- No partition should have Cross-partition Control Authority (for example; No partition should be able to reset or deactivate another partition).

- No Isolated partitions may have coupling facility channels which would allow communication to a Coupling Facility partition.

- Isolated partition must not have network connections which would allow communication to IBM zAware running in a Secure Service Container mode partition.

- The 'Use dynamically changed address' and 'Use dynamically changed parameter' checkboxes must not be selected in the Image or Load profile.

- The Hardware Management Console's Customizable Data Replication service should be disabled.

- Product Engineering (PE) access to the HMC/SE should normally be disabled but can be permitted for brief periods to allow PE diagnostic work.

- No Enterprise Directory Server (LDAP) Definitions should be created on the Hardware Management Console or the Support Element.

- The Hardware Management Console and the Support Element API setting for the Simple Network Management Protocol (SNMP) API should be disabled.

- The Hardware Management Console service for the Web Services Application Programming Interface should be disabled.

- The System BCPii Permissions must be disabled to prevent BCPii commands from being received.

- All partitions must have their BCPii Permissions disabled to prevent BCPii commands from being sent and received.

PR/SM characteristics

- There is a Hardware Management Console (HMC) and Support Element (SE) from which the system can be operated. Therefore the system administrators of the system must be cleared for the highest security classification of work being performed on the system.

- Hardware-related operations for each logical partition will be conducted from the HMC or SE. Operations are invoked by selecting the desired CPC image (representing a logical partition) and invoking the desired task.

- For enhanced integrity of execution, locking of partitions is recommended. The partition must then be unlocked before other disruptive operations can be performed on that partition. Lock a logical partition by selecting the CPC image representing the logical partition and invoking the **Image Details** task. Select **Lockout disruptive tasks** to **Yes** and click **Apply** . You can use this same procedure to unlock a logical partition by setting the Lockout disruptive tasks radio button to No and saving the setting. Locking a logical partition can prevent accidentally performing disruptive tasks on it.

- When entering values on an Hardware Management Console or Support Element window, values are not recognized by the system until you save the data and confirm the changes appearing on the screen.

- The Security Log records system operator actions and responses for operations that are security relevant. The entries are in chronological order and provide an audit log. Entries also include a user (system administrator) identifier when appropriate.

- The Security Log, when full, will be pruned to 67% of its capacity. The oldest entries are deleted. Care should be taken to periodically off load the security log to insure that no records are lost.

- When the security log is successfully off loaded to removable media, the active log is pruned so that it does not exceed 20% of its capacity. If the active security log is below 20%, then no entries are removed. If it is above 20%, then enough active security log entries are removed (from oldest to newest) to reduce the size of the active security log to 20%. The oldest entries are still in the offloaded log.

- The security log on both the Hardware Management and Support Element is 30 megabytes. Entries range from 40 bytes to 400 bytes.

- Open the **View Security Logs** tasks from the Hardware Management Console and Support Element to view the security log.

Central storage

Throughout this document there are statements that state "Sharing of allocated central storage among multiple logical partitions is not allowed", and " ... it becomes available to the logical partition if no other logical partition is using the same storage. This is because PR/SM has a mechanism that detects

conditions where sharing was defined (where address ranges overlap), and rejects such requests. PR/SM licensed internal code (LIC) and hardware rigidly enforce the *no-sharing* rule at logical partition definition, during logical partition activation, during logical partition reconfiguration, and during logical partition execution. PR/SM monitors each instruction's storage accesses for validity; accesses outside the logical partition's defined storage are not permitted to proceed.

Only storage increments within the logical partition's storage allocation as defined in the activation profile can be placed offline. For z/OS System Control Program (SCP) partitions, storage is varied off and on line by using the z/OS CONFIG (CF) operator command. See *z/OS MVS System Commands* for further detail. While processing this command, MVS™ must interact with PR/SM, through a service call instruction, to request that the storage be varied. Because storage cannot be varied without PR/SM involvement, no way exists to circumvent the validity checking PR/SM does to confine a partition occupant within the storage limits defined for the logical partition.

I/O security considerations

IOCDS considerations

Chapter 2, "Planning considerations," on page 35 contains a very thorough discussion of I/O configuration-related topics. It should be read in its entirety before reading the following security considerations.

When the IOCDS does not specify any sharing, I/O devices are owned solely by the logical partitions that own the channel paths that are attached to them. Even if a channel path has been designated as reconfigurable, that channel path cannot be removed from a logical partition unless the channel path has first been taken offline from within that logical partition. For z/OS System Control Program (SCP) partitions, this is done with the SCP operator command CONFIG (CF). For partitions containing other SCPs, the Channel Operations task list must be used. Use the **Configure Channel Path On/Off** task to configure channel paths that are online. Use the **Release I/O Path** task to release the channel paths that are assigned to logical partitions that have the Logical Partition Isolation security control enabled or, use the **Reassign I/O Path** task to reconfigure a CHPID in one step.

I/O sharing should never be allowed for Isolated logical partitions. If the IOCDS were to specify I/O sharing, it would be indicated in the Input/Output Configuration Program's Configuration Reports (see the *Input/Output Configuration Program User's Guide for ICP*).

Isolated logical partitions must never define channel paths as shared in the IOCDS. Specification of a shared channel path can compromise the security of the Isolated logical partitions in the installation. A shared channel path is defined by specifying one of the following on the CHPID statement:

- SHARED keyword
- NOTPART keyword
- PARTITION keyword with more than one logical partition in the access list
- IOCLUSTER keyword
- PATH keyword with more than one CSS ID (for example, a spanned channel path)

Use of a shared channel path allows the possibility of two partitions having access to the same I/O control units and devices. This is in contradiction to the policy of strict separation. Additionally, the use of shared channels may facilitate some form of covert signaling. However, if covert signaling is not perceived to be a significant threat, it is highly recommended that each use of a shared channel be carefully analyzed for its possible effect on the installations security policy. Although a shared channel path is defined to be shared, **none** of the devices that are connected to it need to be shared among logical partitions. When devices are assigned to a single logical partition, they cannot be accessed by any other logical partition.

Low-speed devices (such as SCP Operator's Consoles) are especially inviting targets for sharing a single channel path using multiple image facility (MIF).

If you choose to share channel paths between Isolated logical partitions, and their access to specific devices attached to that channel path must be restricted, I/O Device Candidate Lists are the means for restricting access to devices. The default, if no I/O Device Candidate List is specified, is that all partitions sharing the MIF channel path, also share access to all attached devices. Such free access is incompatible with the concept of a secure platform that provides disjoint, non-communicating logical partitions, and is therefore not recommended. We recommend that when sharing is specified for a CHPID, all the associated, attached I/O devices (IODEVICE statement) must have a candidate list specified. Following a rule of always specifying a device's partition explicitly prevents unexpected results from defaults being applied. For further details on I/O device candidate list, refer to the discussion of the IODEVICE statement's PARTITION parameter in the *Input/Output Configuration Program User's Guide for ICP*.

Sharing of channel paths is controlled by the **SHARED** parameter, and the partition names specified in the **PARTITION** and **NOTPART** parameters for each channel path definition (CHPID statement) in the IOCDS. If the PARTITION parameter specifies multiple partition names, it specifies that this particular CHPID is shared among the named partitions. If a NOTPART parameter is used, it implies the sharing characteristic. However, if a NOTPART parameter includes all partition names but one, in both access and candidate lists, no sharing is permitted. Devices attached to a shared CHPID are restricted to the partitions included in the device candidate list (specified in the IODEVICE PARTITION parameter). If the IOCDS does not specify sharing, then no sharing of CHPIDs will take place.

Operational considerations

Global, system-wide control of Dynamic I/O Configuration is provided by the **I/O Configuration Control Authority**. Use of this facility does not result in strict separation of partitions and was not included in the evaluated product. At most, only use the **Customize/Delete Activation Profiles** task to open a reset or image profile to enable I/O Configuration control for a logical partition. The I/O configuration control selection is located on the **Security** page for the logical partition. See Input/output (I/O) configuration control for more information.

Logical partitions may also be defined with their Logical Partition Isolation security control enabled. For such logical partitions, an offline, reconfigurable CHPIDs cannot be assigned to another logical partition unless the **Release I/O Path** task is invoked (or the **Reassign I/O Path** task) by the System Administrator from the SE or HMC. These tasks are available from the Channel Operations task list. The CHPID statement's candidate list can be used to limit the "mobility" of a reconfigurable channel. The system will only accept configure on commands for CHPIDS in partitions specified in the candidate list of the target channel path.

All channel path reconfiguration procedures should be specifically described in the secure installation's procedures. Any not described, must not be permitted. While developing these procedures, consideration must be given to the security implications of defining a channel path (and its attached devices) to be reconfigurable. Specifically, which from-to pairs of logical partitions are valid? (When this has been established, candidate lists are the means for implementing this aspect of the installation's security policy). In the process of reconfiguration, could data be passed from one logical partition to another via one of the attached devices? What other procedural controls must be followed in order that your organization's security policy is maintained? What operator actions are required to reconfigure this specific channel path in a secure manner? Lastly, careful attention to the IOCDS language rules relating to the CHPID REC parameter is necessary to achieve the result desired.

Channel path reassignments which result from executing configure CHPID actions, are remembered by the system by recording these changes on the SE hard drive and associating them with the IOCDS (the IOCDS itself is not changed). These changes to channel path assignments (I/O Configuration) take effect whenever the logical partitions are again activated. If the IOCDS is rewritten (by invoking HCD or IOCP), the channel path reassignments are erased (at the first Activation using that newly rewritten IOCDS).

When a channel path is deconfigured from a logical partition, each subchannel (an internal structure that provides the logical appearance of an I/O device, and is uniquely associated with one I/O device) for which this channel path is the only (remaining) online path, is removed from the logical partition. Before the subchannels are removed, they are drained and disabled. Subsequently the channel path is reset. If the channel path being deconfigured is the last channel path to a device, that device is also reset. Actions directed to a removed subchannel result in a condition code=3 (not operational).

At that very first use of a newly created IOCDS, activation configures all channel paths to the logical partitions as defined by the IOCDS. The subsequent movements of reconfigurable channel paths, from one logical partition to another, is remembered by the system. During subsequent activations, as each logical partition is activated, if a channel path was (previously) moved out of a logical partition, the channel path is taken offline to that logical partition; if a channel path was moved into a logical partition, the channel path is brought on line to that logical partition. These logical configuration changes can be viewed by performing the following steps:

- Go to the Configuration task list
- Select **Input/Output (I/O) Configuration** task
- Select the IOCDS which is marked **Active**
- Select the **View** pulldown
- In the **View** pulldown, select **Channel Path Configuration**
- Select a **PCHID**
- Select the **View** pulldown
- In the **View** pulldown, select **CHPID information**

The Security Administrator can repeat the final three steps shown above to see all defined CHPIDS and determine which partition(s) each is assigned to, whether they are Dedicated, Shared or Reconfigurable, and the type of each CHPID. In a secure installation, CHPIDs must not be shared among Isolated logical partitions.

Careful review of installation security guidelines must precede using the **Swap Channel Path** procedure. All devices attached to the channel path being switched in, may be accessible to the logical partition. This caution does not apply to a truly "spare" channel path, one with no devices currently defined or attached.

Input/Output Configuration Data Set (IOCDS)

An IOCDS defines the logical partitions by name, allocates I/O resources to each of them, and specifies the security characteristics of those I/O resources. The following list describes the security-relevant parameters of each type of IOCDS source statement.

Statement Type
   Discussion

ID
   No security-relevant parameters.

RESOURCE
   Assign logical partition names and MIF image IDs so that explicit control is asserted, and maximum checking of following IOCDS source statements is enabled.

CHPID

- Use **PARTITION** parameter to specify which logical partition each channel path is allocated to.
- Don't use the **SHARED** parameter for Isolated logical partitions without study of the security implications.
- Don't use the **REC** parameter without study of the security implications.
- Specify whether the channel path is **REConfigurable**, and specify which logical partitions are to have access (using logical partition names in the candidate list).
- Don't use the **IOCLUSTER** keyword for any Isolated logical partitions.

CNTLUNIT
   Specification of the **PATH** parameter must be accorded care so that a secure configuration results.

IODEVICE
   Specification of the **CUNUMBR** parameter must be accorded care so that a secure configuration results.

LPAR Input/Output configurations

- In general, I/O devices must not be shared by isolated logical partitions, since they can be used to pass information from one partition to another. There may be special cases, such as an output-only device which an installation may consider sharable after careful review of any related security risks, and defining related security procedures and processes.

- The PCHID Summary Report, Channel Path Identifier (CHPID) Summary Report and I/O Device Report produced by the Input/Output Configuration Program must be thoroughly examined by the Security Administrator for indications of unwanted sharing or reconfigurability of channels and devices.

- A thorough review of the actual physical connections/links of the I/O configuration must be performed to establish that the physical configuration is identical to that specified in the IOCDS source file. Specific attention should be given to devices with multiple device path capability, to help ensure that one device (or control unit) does not (accidentally) connect to more than one isolated logical partition's channel paths.

- All IOCDSs should be write-protected except for the few minutes during which they are actually updated.

- The time stamps of the production-level IOCDSs should be recorded. By selecting the CPC and invoking the **Input/Output (I/O) Configuration** task, a display of the available IOCDSs will be generated. Periodic audits should be made to assure that the IOCDSs have remained unchanged.

Activation

A reset profile includes information for activating a CPC and its images (logical partitions).

- In the reset profile, after selecting an LPAR IOCDS (A0-A3) deemed valid for secure operation using the **Input/Output (I/O) Configuration** task, the operating mode selected must be logically partitioned (LPAR).

- Dynamic I/O changes can be disabled on the Dynamic Page of the **Power-on Reset** task displayed during Power on Reset of the CPC. Ensuring the **Allow dynamic changes to the channel subsystem input/output (I/O) definition)** is not selected, disables dynamic I/O for the CPC. Globally disabling dynamic I/O configuration narrows the control of the I/O configuration control parameter to only controlling a logical partition's reading and writing of IOCDS.

- Workload Manager (found on the Processor page of the Image profile) should not be enabled for Isolated logical partitions.

Enabling Workload Manager (WLM) enables Dynamic CHPID Management (DCM) to optimize I/O throughput across an LPAR cluster by sharing CHPIDs among partitions who have joined the cluster. Any logical partition that is WLM enabled may join the cluster and therefore share CHPIDs reserved for use by members of the specified cluster. Furthermore, partitions within a cluster may issue a special DIAGNOSE instruction to obtain information about other partitions within the same cluster even when Global Performance Data Authority is not enabled. See for more information.

Security controls

- A logical partition's initial security settings are set in the image profile used to activate it. Afterward, the **Change Logical Partition Security** task can be used to view or change the settings. Changes must be saved in the profile in order to have them available for subsequent use. Security settings are saved by the system across activations for the current configuration. Therefore, if the same configuration is used, Security settings need not be reentered (but should be checked).

- The following Logical Partition Security Controls settings are required for a secure mode of operation:

  - **ISOLATION should be enabled**. This option binds the partition's allocated I/O configuration to it, even when a Channel Path (CHPID) is in an offline state. An overt, auditable operator action is required to unbind an item of the I/O configuration and move it to another partition.

  - **I/O CONFIGURATION CONTROL should be disabled for every partition**. By negating this option, the partitions are prevented from accessing (read or write) the existing IOCDS data sets, or

dynamically altering the current I/O configuration. IOCDSs can be a means to surreptitiously pass data between partitions. In addition, dynamic alteration of the current I/O configuration can result in a partition having access to data that it is not authorized to access. Dynamic I/O Configuration is supported by the Hardware Configuration Definition (HCD) product for the z/OS or z/VM operating system.

Note: I/O Configuration control should be enabled for a single, specific logical partition only during the short period of time when it is permitted to write a new IOCDS. Only the IOCDS to be written should have its write-protection temporarily reset. All other IOCDSs should remain write-protected during an IOCDS update operation. The Security Administrator should remain logged onto the console until the IOCDS update is complete, and the IOCDS update authority is disabled.

Note: Neither Isolation nor I/O Configuration Control option has any effect on the sharing of CHPIDS or I/O Devices. Sharing is enabled by parameters of the CHPID statement used in the definition of the IOCDS.

– **GLOBAL PERFORMANCE DATA AUTHORITY should be disabled for every partition**. This recommendation is based on a desire to block any possibility of a partition extracting meaning from another partition's performance data.

– **CROSS-PARTITION CONTROL should be disabled for every partition**. Enabling cross-partition control permits one partition to disrupt processing in other partitions, resulting in the threat of denial of service to those partition's users. When cross-partition control is disabled, Automatic Reconfiguration Facility (ARF) is disabled. ARF uses the cross-partition control capability of PR/SM. ARF is not generally appropriate in a tightly managed, secure system.

Reconfiguring the system

Deconfiguration

The recommended way to deconfigure objects owned by a logical partition is to first deconfigure the object from the operating system's point of view, and when necessary (z/OS interacts with PR/SM to complete the reconfiguration process, other operating systems may not), use the Hardware Management Console's tasks to request PR/SM to deconfigure the identical object. The z/OS operating system expects operations personnel to use the HMC/SE based configuration tasks to request deconfiguration of a logical partition object.

Reconfiguration

If the objects are not presently part of the logical partition's configuration, they must be made available to the partition with the use of facilities at the Hardware Management Console. Channel Paths (CHPIDs) may be made available to the target logical partition using the CHPID Operations tasks; reserved storage may be available, but if it isn't, it can be made available by a Security Administrator action by updating the Image profile's Storage page of the **Customize/Delete Activation Profiles** task. There are many operational considerations relating to reconfiguration that are covered in greater detail in the *z/OS MVS Planning: Operations* document and the *z/OS MVS Programming: Resource Recovery*.

The following elements can be reconfigured from the z/OS Operator's console using a CONFIG command. Such a reconfiguration is limited to the objects owned by the logical partition:

• Logical Processors

• Central Storage

• Channel Paths

See *z/OS MVS System Commands* for further detail on the CONFIG command.

z/OS is aware of the logical partition objects it owns, and interacts with PR/SM to reconfigure them using the service call instruction. This Execution of this instruction results in a mandatory interception which causes every use thereof to be mediated by PR/SM. PR/SM mediates the instruction to limit the scope of such requests to the objects that the security administrator defined for the specific logical partition.

Audit trail

All security-relevant events initiated from the HMC/SE by the System Administrator will be written to the security log. When these logs become full, they are *pruned*. This means that the oldest are deleted and the log is reduced to 67% of its capacity. The log has the capability to store many weeks worth of security relevant events under normal system operation.

To insure the no security relevant information is lost, the security log should be offloaded periodically to removable media provided with the processor. When the security log is successfully off loaded to removable media, the active log is pruned so that it does not exceed 20% of its capacity. If the active security log is below 20%, then no entries are removed. If it is above 20%, then enough active security log entries are removed (from oldest to newest) to reduce the size of the active security log to 20%. The oldest entries are still in the offloaded log.

Recovery planning

You should read "Recovery strategy" on page 181, and then adapt it to your configuration's requirements for security and processing priorities. Installation-specific recovery procedures must be developed and documented in advance, always giving consideration to where the sensitive data will be after each recovery scenario has completed.

Service and maintenance

Many secure accounts are hesitant about enabling remote service. Consideration should be given to enabling outbound RSF calls that contain the data necessary to automatically dispatch a service representative. Since there is considerable customizing capability provided, RSF can probably be tailored to match your installation's security policy and practices.

This product has support for the concurrent service and maintenance of hardware. The following can be serviced concurrently while normal customer operations continue:

- Power supplies
- Channel cards
- Licensed Internal Code (LIC)
- Processor Drawer

When service is performed on the above-listed elements of the processor, the physical elements are logically and electrically isolated from the remaining portions of the system still in use. This is begun by first logging on the HMC with a SERVICE ID and then performing the desired maintenance or service task. Refer to the *Service Guide* for information on how to perform the desired task.

Note: Before placing a reconfigurable or shared channel path into a service configuration, record the logical partition name(s) that it's currently assigned to. This will assure that after service is complete, the channel path will be returned to the logical partition(s) to which it was previously allocated, even if different operations personnel are now in charge.

When a partial or complete configuration is surrendered for service or maintenance the following recommendations should be followed:

- The IOCDSs should remain write-protected.
- All installation configuration data should be, or has been previously, saved. The installation configuration data should be restored, and the initial activation must be fully manual. When activation completes, use the following procedure to check the active I/O configuration:
  - go to the **Configuration** task list
  - select **Input/Output (I/O) Configuration** task
  - select the IOCDS which is marked **Active**
  - select the **View** pulldown
  - in the **View** pulldown, select **Channel Path Configuration**

- Prior to giving the system to the service representative for disruptive maintenance, it is advisable to idle the partitions (perform an orderly shutdown of the applications and control programs occupying the partitions, followed by stopping each partition) rather than deactivating them. Doing this allows the system to perform automatic (re)activation on the subsequent activation. Automatic activation offers fewer opportunities for human error to affect the controlling parameters of the system, and hence is more secure.

After completion of a disruptive service operation, the CPC should be selected and the I/O Configuration task invoked to display a list of the available IOCDSs. Use this list to check the IOCDS time stamps against the values recorded the last time the IOCDSs were updated. This is to help ensure that the IOCDSs remain unchanged.

The Hipersocket Sniffer debugging tool should not be used in normal operations, as it could cause security problems. For more information, see the IBM HiperSockets Network Traffic Analyzer (HS NTA) Frequently Asked Questions document, found at http://www.ibm.com/systems/z/hardware/networking/products.html#hipersockets.

Logical processors

A logical core may be taken offline as the result of an z/OS operator entering an z/OS CONFIG command to take one (or more) CPs offline. When this is done, z/OS performs the work necessary to no longer dispatch work on the CP(s), and then executes a service call instruction to request that PR/SM take the logical cores(s) offline. See *z/OS MVS System Commands* for further detail on the CONFIG command. Lastly, a logical core may be taken off line at the next activation of the partition by reducing the number of CPs defined for a logical partition in the image profile for the logical partition.

The maximum number of logical processors for each logical partition is defined at logical partition activation, and remains fixed for the duration of the activation. Each of these logical cores is represented by a data structure that is associated only with its specific logical partition. There are no circumstances where a logical core can be "transferred" to another logical partition, nor is there a capability within the system to accomplish this.

When a logical core is taken offline, the data structure that represents it is marked as "offline", and continues to be maintained in PR/SM-accessible storage, remaining absolutely bound to its logical partition for the duration of that partition's activation. An offline logical core presents a checkstopped status when interrogated by the other logical cores in the partition. An offline logical core can be restored to the online status by issuing an z/OS CONFIG command. z/OS uses the service call instruction to request PR/SM bring an offline logical core back on line. If successful, z/OS prepares its control structures to add the CP to its pool of available resources.

Initial Program Load

An Initial Program Load (IPL) resets a logical partition to prepare it for loading an operating system, and then loads the operating system using the specified IPL address and IPL parameter. The IPL address and IPL parameter are normally entered manually in the image activation profile or the load profile. However, a partition with I/O Configuration authority has the capability of dynamically changing an IPL address and IPL parameter. This could potentially cause an unintended operating system to be loaded into the partition. To prevent this, the 'Use dynamically changed address' and 'Use dynamically changed parameter' check boxes must not be selected in the Image or Load profile of an Isolated logical partition.

Trusted facility library

Use the manuals listed in this section as needed for background or supplemental information.

Check the edition notices in the beginning of each manual for correspondence to the appropriate driver level. For manuals that do not specify driver level, the highest (most recent) suffix (last two digits of the form number) is required.

Table 22. Trusted facility library for PR/SM	
Title	**Order Number**
Hardware Management Console (HMC) and Support Element (SE) information can be found on the console help system. **Note:** When operating a certified system, the console help should be the primary reference source for any information pertaining to security related tasks.	
Secure Service Container's User's Guide	SC28-7005
Service Guide for Hardware Management Consoles and Support Elements	GC28-6983
Service Guide for Trusted Key Entry Workstations	GC28-6980
8562 Installation Manual	GC28-7009
8562 Installation Manual for Physical Planning (IMPP)	GC28-7011
Stand-Alone IOCP User's Guide	SB10-7173
Input/Output Configuration Program User's Guide for ICP IOCP	SB10-7172
8562 Safety Inspection	GC28-7007
SNMP Application Programming Interfaces	SB10-7171
Security Architecture: Securing the Open Client/Server Distributed Enterprise	SC28-8135
MVS Planning: Security	GC28-1439
Introducing Enterprise Systems Connection	GA23-0383
z/OS Hardware Configuration Definition: User's Guide	SC33-7988
z/OS MVS System Commands	SA22-7627
z/OS MVS Planning: Operations	SA22-7601
z/OS MVS Recovery and Reconfiguration Guide	SA22-7623
z/OS Cryptographic Services ICSF TKE Workstation User's Guide	SA23-2211
zVM: Secure Configuration Guide	SC24-6230
Security for Linux on System z	SG24-7728

Appendix C. Notices

This information was developed for products and services offered in the US.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

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

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

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprise is entirely coincidental.

Trademarks

IBM, the IBM logo, and ibm.com® are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other

companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at http://www.ibm.com/legal/copytrade.shtml.

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

Java is a trademark or registered trademarks of Oracle and/or its affiliates.

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

Class A Notices

The following Class A statements apply to this IBM product. The statement for other IBM products intended for use with this product will appear in their accompanying manuals.

Federal Communications Commission (FCC) Statement

Note: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference, in which case the user will be required to correct the interference at his own expense.

Properly shielded and grounded cables and connectors must be used in order to meet FCC emission limits. IBM is not responsible for any radio or television interference caused by using other than recommended cables and connectors or by unauthorized changes or modifications to this equipment. Unauthorized changes or modifications could void the user's authority to operate the equipment.

This device complies with Part 15 of the FCC rules. Operation is subject to the following two conditions: (1) this device may not cause harmful interference, and (2) this device must accept any interference received, including interference that may cause undesired operation.

Industry Canada Compliance Statement

This Class A digital apparatus complies with Canadian ICES-003.

Avis de conformité à la réglementation d'Industrie Canada

Cet appareil numérique de la classe A est conforme à la norme NMB-003 du Canada.

European Community Compliance Statement

This product is in conformity with the protection requirements of EU Council Directive 2014/30/EU on the approximation of the laws of the Member States relating to electromagnetic compatibility. IBM cannot accept responsibility for any failure to satisfy the protection requirements resulting from a non-recommended modification of the product, including the fitting of non-IBM option cards.

This product has been tested and found to comply with the limits for Class A Information Technology Equipment according to European Standard EN 55032. The limits for Class A equipment were derived for commercial and industrial environments to provide reasonable protection against interference with licensed communication equipment.

European Community contact:
IBM Deutschland GmbH
Technical Regulations, Department M372
IBM-Allee 1, 71139 Ehningen, Germany
Tele: +49 (0) 800 225 5423 or +49 (0) 180 331 3233
email: halloibm@de.ibm.com

Warning: This is a Class A product. In a domestic environment, this product may cause radio interference, in which case the user may be required to take adequate measures.

VCCI Statement - Japan

この装置は、クラス A 情報技術装置です。この装置を家庭環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。　　　　　　　VCCI-A

The following is a summary of the Japanese VCCI statement above:

This is a Class A product based on the standard of the VCCI Council. If this equipment is used in a domestic environment, radio interference may occur, in which case the user may be required to take corrective actions.

Japan JIS C 61000-3-2 Compliance

（一社）電子情報技術産業協会 高調波電流抑制対策実施
要領に基づく定格入力電力値： Knowledge Centerの各製品の
仕様ページ参照

For products less than or equal to 20 A per phase, the following statement applies:

高調波電流規格　JIS C 61000-3-2 適合品

For products greater than 20 A, single-phase, the following statements apply:

高調波電流規格　JIS C 61000-3-2 準用品

本装置は、「高圧又は特別高圧で受電する需要家の高調波抑制対策ガイドライン」対象機器（高調波発生機器）です。
回路分類：6（単相、ＰＦＣ回路付）
換算係数：0

For products greater than 20 A per phase, three-phase, the following statements apply:

高調波電流規格　JIS C 61000-3-2 準用品

本装置は、「高圧又は特別高圧で受電する需要家の高調波抑制対策ガイドライン」対象機器（高調波発生機器）です。
回路分類　：5（3相、ＰＦＣ回路付）
換算係数　：0

Electromagnetic Interference (EMI) Statement - People's Republic of China

声　明

此为 A 级产品，在生活环境中，
该产品可能会造成无线电干扰，
在这种情况下，可能需要用户对其
干扰采取切实可行的措施。

Declaration: This is a Class A product. In a domestic environment, this product may cause radio interference, in which case the user may need to perform practical action.

Electromagnetic Interference (EMI) Statement - Taiwan

警告使用者：
這是甲類的資訊產品，在
居住的環境中使用時，可
能會造成射頻干擾，在這
種情況下，使用者會被要
求採取某些適當的對策。

The following is a summary of the Taiwan EMI statement above:

Warning: This is a Class A product. In a domestic environment, this product may cause radio interference, in which case the user will be required to take adequate measures.

IBM Taiwan Contact Information:

台灣IBM 產品服務聯絡方式：
台灣國際商業機器股份有限公司
台北市松仁路7號3樓
電話：0800-016-888

Electromagnetic Interference (EMI) Statement - Korea

이 기기는 업무용(A급)으로 전자파적합등록을 한 기기이오니
판매자 또는 사용자는 이 점을 주의하시기 바라며, 가정외의
지역에서 사용하는 것을 목적으로 합니다.

Germany Compliance Statement

Deutschsprachiger EU Hinweis: Hinweis für Geräte der Klasse A EU-Richtlinie zur Elektromagnetischen Verträglichkeit

Dieses Produkt entspricht den Schutzanforderungen der EU-Richtlinie 2014/30/EU zur Angleichung der Rechtsvorschriften über die elektromagnetische Verträglichkeit in den EU-Mitgliedsstaaten und hält die Grenzwerte der EN 55032 Klasse A ein.

Um dieses sicherzustellen, sind die Geräte wie in den Handbüchern beschrieben zu installieren und zu betreiben. Des Weiteren dürfen auch nur von der IBM empfohlene Kabel angeschlossen werden. IBM übernimmt keine Verantwortung für die Einhaltung der Schutzanforderungen, wenn das Produkt ohne Zustimmung von IBM verändert bzw. wenn Erweiterungskomponenten von Fremdherstellern ohne Empfehlung von IBM gesteckt/eingebaut werden.

EN 55032 Klasse A Geräte müssen mit folgendem Warnhinweis versehen werden:

"Warnung: Dieses ist eine Einrichtung der Klasse A. Diese Einrichtung kann im Wohnbereich Funk-Störungen verursachen; in diesem Fall kann vom Betreiber verlangt werden, angemessene Maßnahmen zu ergreifen und dafür aufzukommen."

Deutschland: Einhaltung des Gesetzes über die elektromagnetische Verträglichkeit von Geräten

Dieses Produkt entspricht dem "Gesetz über die elektromagnetische Verträglichkeit von Geräten (EMVG)". Dies ist die Umsetzung der EU-Richtlinie 2014/30/EU in der Bundesrepublik Deutschland.

Zulassungsbescheinigung laut dem Deutschen Gesetz über die elektromagnetische Verträglichkeit von Geräten (EMVG) (bzw. der EMC EG Richtlinie 2014/30/EU) für Geräte der Klasse A

Dieses Gerät ist berechtigt, in Übereinstimmung mit dem Deutschen EMVG das EG-Konformitätszeichen - CE - zu führen.

Verantwortlich für die Einhaltung der EMV Vorschriften ist der Hersteller:
International Business Machines Corp.
New Orchard Road
Armonk, New York 10504
Tel: 914-499-1900

Der verantwortliche Ansprechpartner des Herstellers in der EU ist:
IBM Deutschland GmbH
Technical Regulations, Abteilung M372
IBM-Allee 1, 71139 Ehningen, Germany
Tel: +49 (0) 800 225 5423 or +49 (0) 180 331 3233
email: halloibm@de.ibm.com

Generelle Informationen:

Das Gerät erfüllt die Schutzanforderungen nach EN 55024 und EN 55032 Klasse A.

Electromagnetic Interference (EMI) Statement - Russia

ВНИМАНИЕ! Настоящее изделие относится к классу А.
В жилых помещениях оно может создавать радиопомехи, для
снижения которых необходимы дополнительные меры

Index

Sysplex Failure Manager (SFM) *(continued)*
     description 102
system-managed coupling facility structure duplexing 67

T

TARGETSYS(ALL) examples 102
time offset, image profile 135
TOD clock 32
trademarks 198
transparent sparing 182
trusted configuration 186

U

uniprocessor models, processor considerations 108
unshared channel paths, moving 158
upgrade, capacity on demand 83
Usage Domain Index 150
usage domain zeroize 174

V

View LPAR Cryptographic Controls 171
virtual flash memory
     definition 1, 96

W

weights, processing
     enforcement 112
workload
     balancing 41
     requirements, CP 104
workload charging by soft-capping to a defined capacity 120
Workload Manager 122
workload manager LPAR CPU management of shared CPs
119

X

XRF 29

Z

z Integrated Information Processor (zIIP) 82
z/VM guest coupling simulation 5
z/VM mode LPs, processor considerations 108
zeroizing a domain 174
zIIP 82

IBM®