zEnterprise 196
System Overview

SA22-1086-03

Level 03c

zEnterprise 196

System Overview

SA22-1086-03

Note:
Before using this information and the product it supports, read the information in "Safety" on page xi, Appendix D, "Notices," on page 157, and *IBM Systems Environmental Notices and User Guide*, Z125-5823.

This edition, SA22-1086-03, applies to the IBM zEnterprise 196 (z196) and replaces SA22-1086-02.

These might be a newer version of this document in a PDF file available on **Resource Link**. Go to http://www.ibm.com/servers/resourcelink and click **Library** on the navigation bar. A newer version is indicated by a lowercase, alphabetic letter following the form number suffix (for example: 00a, 00b, 01a, 01b).

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 IBM® 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 System z® models can use I/O cards such as, ESCON®, FICON®, Open Systems Adapter (OSA), InterSystem Channel-3 (ISC-3), 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.

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 publication describes the design, components, functions, features, and capabilities of the IBM zEnterprise® 196 models. It is intended for executives, data processing managers, data processing technical staff, consultants, and vendors who wish to exploit z196 advantages.

You should be familiar with the various publications listed in "Prerequisite publications" and "Related publications." A glossary and an index are provided at the back of this publication.

What is included in this publication

This publication contains the following chapters and appendices:

Revisions

A technical change to the text is indicated by a vertical line to the left of the change.

Prerequisite publications

Before reading this publication you should be familiar with IBM z/Architecture, IBM S/390, and IBM Enterprise Systems Architecture/390 (ESA/390) as described in the following publications:

- *z/Architecture Principles of Operation*, SA22-7832
- *Enterprise System Architecture/390 Principles of Operation*, SA22-7201

Related publications

Important:
Please ensure that you are using the most recent version of all related documentation.

Other IBM publications that you will find helpful and that you should use along with this publication are in the following list. You can access these books from *Resource Link* under the **Library** section.

- *System z Application Programming Interfaces*, SB10-7030
- *System z Application Programming Interfaces for Java*, API-JAVA
- *System z Common Information Model (CIM) Management Interface*, SB10-7154
- *System z Hardware Management Console Web Services API* , SC27-2616-00
- *zEnterprise System Capacity on Demand User's Guide*, SC28-2605
- *System z CHPID Mapping Tool User's Guide*, GC28-6900
- *System z ESCON and FICON Channel-to-Channel Reference*, SB10-7034

- *System z Hardware Management Console Operations Guide*, SC28-6905
- *System z Stand-Alone Input/Ouput Configuration Program (IOCP) User's Guide*, SB10-7152
- *System z Input/Output Configuration Program User's Guide for ICP IOCP*, SB10-7037
- *zEnterprise 196 Installation Manual*, GC28-6890
- *zEnterprise 196 Installation Manual for Physical Planning*, GC28-6897
- *zEnterprise System Processor Resource/Systems Manager Planning Guide*, SB10-7155
- *zEnterprise System Support Element Operations Guide*, SC28-6906
- *System z Small Computer Systems (SCSI) IPL - Machine Loader Messages*, SC28-6839
- *System z Planning for Fiber Optic Links (ESCON, FICON, Coupling Links, and Open System Adapters)*, GA23-0367
- *zEnterprise System Service Guide for Trusted Key Entry Workstations*, GC28-6901
- *zEnterprise 196 Service Guide*, GC28-6892
- *System z Coupling Links I/O Interface Physical Layer*, SA23-0395
- *System z Service Guide for Hardware Management Consoles and Support Elements*, GC28-6861
- *System z Maintenance Information for Fiber Optics (ESCON, FICON, Coupling Links, and Open System Adapters)*, SY27-2597
- *System z Fibre Channel Connection (FICON) I/O Interface Physical Layer*, SA24-7172
- *Fiber Optic Cleaning Procedures*, SY27-2604
- *Set-Program-Parameter and the CPU-Measurement Facilities*, SA23-2260
- *CPU-Measurement Facility Extended Counters Definition for z10 and z196*, SA23-2261

Ensemble publications

The following publications provide overview, planning, performance, and Hardware Management Console (HMC) and Support Element task information about creating and managing an ensemble.

- *zEnterprise System Introduction to Ensembles*, GC27-2609
- *zEnterprise System Ensemble Planning and Configuring Guide*, GC27-2608
- *zEnterprise System Ensemble Performance Management Guide*, GC27-2607
- *zEnterprise System Hardware Management Console Operations Guide for Ensembles*, GC27-2615
- *zEnterprise System Support Element Operations Guide*, SC28-6906
- *zEnterprise BladeCenter Extension Installation Manual Model 002*, GC27-2610
- *zEnterprise BladeCenter Extension Installation Manual Model 002 for Physical Planning*, GC27-2611
- *z/VM Systems Management Application Programming*, SC24-6234
- *z/VM Connectivity*, SC24-6174
- *z/VM CP Planning and Administration*, SC24-6178

Parallel sysplex publications

A **Parallel Sysplex** system consists of two or more z/OS images coupled by coupling links to a common Coupling Facility and synchronized by a common time source, such as Server Time Protocol (STP) or a Sysplex Timer. A Parallel Sysplex can be used to present a single image to the user. A Parallel Sysplex can use the coupling facility to provide data sharing among the systems participating in the Parallel Sysplex.

The following publications provide additional information to help you understand and prepare for a Parallel Sysplex that uses coupling facility for data sharing purposes.

- *z/OS Parallel Sysplex Application Migration*, SA22-7662
- *z/OS Parallel Sysplex Overview: Introducing Data Sharing and Parallelism in a Sysplex*, SA22-7661
- *z/OS MVS Setting Up a Sysplex*, SA22-7625

OSA publications

The following publications provide additional information for planning and using the OSA-Express features:

- *zEnterprise, System z10, System z9 and zSeries Open Systems Adapter-Express Customer's Guide and Reference*, SA22-7935
- *System z10 Open Systems Adapter-Express3 Integrated Console Controller Dual-Port User's Guide*, SC23-2266

Cryptographic publications

The following publications provide additional information about the cryptographic function:
- *z/OS Integrated Cryptographic Service Facility Trusted Key Entry PCIX Workstation User's Guide*, SA23-2211
- *z/OS Integrated Cryptographic Service Facility Administrator's Guide*, SA22-7521
- *z/OS Integrated Cryptographic Service Facility Application Programmer's Guide*, SA22-7522
- *z/OS Integrated Cryptographic Service Facility Messages*, SA22-7523
- *z/OS Integrated Cryptographic Service Facility Overview*, SA22-7519
- *z/OS Integrated Cryptographic Service Facility System Programmer's Guide*, SA22-7520

IBM Smart Analytics Optimizer for DB2 for z/OS, V1.1 publications

The following publications provide additional information about the IBM Smart Analytics Optimizer:
- *IBM Smart Analytics Optimizer for DB2 for z/OS Quick Start Guide*, GH12-6915
- *IBM Smart Analytics Optimizer for DB2 for z/OS Installation Guide*, SH12-6916
- *IBM Smart Analytics Optimizer for DB2 for z/OS Stored Procedures and Messages Reference*, SH12-6917
- *IBM Smart Analytics Optimizer Studio User's Guide*,
- SH12-6919
- *IBM Smart Analytics Optimizer for DB2 for z/OS Getting Started*, GH12-6953

IBM DB2 Analytics Accelerator for z/OS V2.1 publications

The following publications provide additional information about the IBM DB2 Analytics Accelerator for z/OS V2.1:
- *IBM DB2 Analytics Accelerator for z/OS, V2.1, Quick Start Guide*, GH12-6957
- *IBM DB2 Analytics Accelerator for z/OS, V2.1, Installation Guide*, SH12-6958
- *IBM DB2 Analytics Accelerator for z/OS, V2.1, Stored Procedures Reference*, SH12-6959
- *IBM DB2 Analytics Accelerator Studio, V2.1, User's Guide*, SH12-6960
- *IBM DB2 Analytics Accelerator for z/OS, V2.1, Getting Started*, GH12-6961

Miscellaneous publications

The following publications provide additional miscellaneous information:
- *IBM Enterprise Storage Server Host Systems Attachment Guide*, SC26-7446
- *IBM Enterprise Storage Server Introduction and Planning Guide, 2105 and Models E10 and E20*, GC26-7444
- *Server Time Protocol Planning Guide*, SG24-7280
- *Server Time Protocol Implementation Guide*, SG24-7281
- *Getting Started with InfiniBand on System z10 and System z9*, SG24-7539

Related web sites

The following web sites provide additional z196 information:

Resource Link
http://www.ibm.com/servers/resourcelink

Resource Link is a key element in supporting the z196 product life cycle. Some of the main areas include:
- *Education*
- *Planning*
- *Library*
- *CHPID Mapping Tool*
- *Customer Initiated Upgrade (CIU)*

Supported operating systems information
http://www.ibm.com/systems/z/os/

Parallel Sysplex and coupling facility information
http://www.ibm.com/systems/z/pso/

FICON information
http://www.ibm.com/systems/z/hardware/connectivity

Open Systems Adapter information
http://www.ibm.com/systems/z/hardware/networking/index.html

Linux on System z information
- *http://www.ibm.com/systems/z/os/linux*
- *http://www.ibm.com/developerworks/linux/linux390/*

Note: When searching, specify "Linux" instead of "All of dW."

IBM WebSphere DataPower Integration Appliance XI50 information
http://www.ibm.com/software/integration/datapower/xi50

Additional online information

Online information about defining tasks and completing tasks associated with z196 is available on the Hardware Management Console and the Support Element. This information is available under the Library category on the Hardware Management Console Welcome screen or the Support Element Welcome page:

- Coupling Facility Control Code (CFCC) commands
- Coupling Facility Control Code (CFCC) messages
- Hardware Management Console Operations Guide
- Support Element Operations Guide
- Hardware Management Console Operations Guide for Ensembles.

Help is available for panels, panel options, and fields on panels.

Engineering change (EC) level considerations

Future enhancements available for z196 models may be dependent on the EC level of the Central Processor Complex (CPC) and/or Hardware Management Console. Additionally, some enhancements may further be dependent on the Microcode Load (MCL) level of the EC on the CPC and/or Hardware Management Console. The required MCL level will be available to the IBM field representative.

EC levels can be tracked by accessing Resource Link, *http://www.ibm.com/servers/resourcelink*. Go to **Tools –> Machine Information**.

Accessibility

This publication is in Adobe Portable Document Format (PDF) and should be compliant with accessibility standards. If you experience difficulties using this PDF file you can request a web-based format of this publication. 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 request, 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.

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).

Summary of changes

Summary of changes for the *zEnterprise 196 System Overview*, SA22-1086.

Table 1. Summary of changes

Release Level	Date	Changes in Level
03c	3/2014	This revision contains the following technical changes: • Common Cryptographic Architecture (CCA) enhancements • IBM WebSphere DataPower Integration Appliance XI50 for zEnterprise (DataPower XI50z) firmware V5.0 support for zBX Model 002.
03b	3/2012	This revision contains editorial changes and the following technical changes: • IBM Smart Analytics Optimizer withdrawal from marketing for new build and MES zBX • Increase in the number of select IBM System x blades you can have in a zBX.
03a	12/2011	This revision contains editorial changes and the following technical changes: • Server/Application State Protocol (SASP) support for load balancing • Integration of HiperSockets network with the intraensemble data network (IEDN) • Support for HiperSockets completion queue function.
03	10/2011	This revision contains editorial changes and the following technical changes: • Microsoft Windows support on selected IBM System x blades • IBM DB2 Analytics Accelerator for z/OS V2.1 support • Hardware Management Console Web Services (Web Services) APIs support • GDPS updates • Ability to use APIs to access Unified Resource Manager function • EAL5+ certification • Ability to monitor and display intraensemble data network (IEDN) network resources.
02a	07/2011	This revision contains corrections to the PSP bucket subset IDs listed in Appendix A, “zEnterprise 196 Version 2.11.1 purpose and description,” on page 147.
02	07/2011	This revision contains editorial changes and the following technical changes: • PCIe I/O drawer support • New HCA3-O, HCA3-O LR, and PCIe fanout cards for InfiniBand coupling • FICON Express8S feature • OSA-Express4S feature • IBM zEnterprise BladeCenter Extension (zBX) support for IBM System x blades • TKE 7.1 feature • Crytographic enhancements • HMC enhancements.
01	02/2011	This revision contains editorial changes and the following technical changes: • IBM WebSphere DataPower Integration Appliance XI50 for zEnterprise (DataPower XI50z) support

Chapter 1. Introduction

The zEnterprise System is designed and built on the precepts of Smarter Computing – an era in which systems are optimized for clients' business needs. Enterprises are evolving. They are no longer deploying systems as fast as possible; they are instead matching the systems architecture to specific workload requirements. zEnterprise embodies this concept and embraces multiple technology platforms – mainframe, POWER® Systems, and System x – to maximize efficiency and transform IT economics. It offers clients a strong architectural foundation, optimized total cost of ownership, and performance levels to meet or exceed business needs.

IBM zEnterprise 196 (z196) represents the latest state of the art scalable server designed and optimized for growth and large-scale consolidation as well as the premier platform to host traditional mission critical workloads along side new application workloads. z196 provides an advanced combination of scalability, availability, reliability, security, and virtualization.

z196 is a total systems approach to delivering advanced heterogeneous systems management functions. Under the management of one system – z196 delivers higher total system capacity and memory; z196 allows you to integrate system and workload management across a multisystem, multitier, and multiarchitectural environment; and z196 enables you to efficiently monitor and management your energy consumption.

Figure 1. z196

With a range of five models, z196 delivers enriched functions over the previous System z systems:
- Increased processing power
- Improved availability with fault tolerant RAIM memory

- Increased I/O capacity
- Increased granularity for I/O infrastructure by allowing I/O drawers, PCIe I/O drawers, and I/O cages
- Increased options for specialty engines
- Enhanced security
- Enhanced network and On Demand offerings
- Enhanced system management
- Enhanced virtualization
- Enhanced energy monitoring and management.

z196 allows virtualization of resources such as:

- Sharing without boundaries.
- Empowerment of business by applying intelligence to adapt and optimize to changing requirements.
- Smart and secure management of global transactions.
- Positioning the mainframe at the center of a heterogeneous on-demand infrastructure.
- Creation and deletion of virtual servers from a central point of control.

To address the growing complexity of fiber optic connectivity in the Information Technology (IT) infrastructure, IBM Site and Facilities Services offers scalable fiber optic cabling services to satisfy Smarter Computing infrastructure requirements at both the product-level and the enterprise-level. See Chapter 8, "Cabling," on page 111 for more information. You can also access Resource Link at *http://www.ibm.com/servers/resourcelink* and click **Services** on the navigation bar for the network cabling services.

zEnterprise 196 highlights

z196 provides:

- **Up to 80 customer processor units (PUs)**
- **Up to 15 PUs available for subcapacity use**
- **IBM System z Integrated Information Processor (zIIP)**
  zIIP is a specialty engine designed to help free-up general computing capacity and lower overall total cost of computing for select data and transaction processing workloads. Using a zIIP can help free capacity on the general-purpose processor.
- **IBM System z Application Assist Processor (zAAP)**
  zAAP is a specialized processor unit that provides a strategic Java™ execution environment, which enables clients to integrate and run new Java-based web applications alongside core z/OS® business applications and backend database systems.
- **Integrated Facility for Linux (IFL)**
  An IFL is a specialty engine that provides additional processing capacity exclusively for Linux on System z workloads.
- **Internal Coupling Facility (ICF)**
  An ICF is a specialty engine that provides additional processing capability exclusively for the execution of the Coupling Facility Control Code (CFCC) in a coupling facility partition.
- **Up to 3 TB of Redundant array of independent memory (RAIM)**
  RAIM technology provides protection at the dynamic random access memory (DRAM), dual inline memory module (DIMM), and memory channel level.
- **Maximum available real memory:**
  – Up to 3.0 TB of maximum available memory available on z196 (depending on the model)
  – Up to 1.0 TB of maximum available memory supported on z196 in any LPAR.
- **Up to 768 GB of memory per book**
- **IBM zEnterprise BladeCenter® Extension (zBX) support**
  zBX is a separate machine, machine type 2458 Model 002, attached to a z114 or z196 and can be viewed

as a logical extension of z114 or z196, respectively. With this support, heterogeneous applications distributed across multiple environments can be configured and processed in a single zEnterprise environment.

With zBX, you can:

– Process CPU intensive DB2® queries, found in Business Intelligence and Data Warehousing applications, using the IBM Smart Analytics Optimizer for DB2 for z/OS, V1.1 (IBM Smart Analytics Optimizer).

  Note: On November 1, 2011, IBM Smart Analytics Optimizer was withdrawn from marketing for new build and MES zBX. It has been replaced with IBM DB2 Analytics Accelerator for z/OS, a workload optimized, appliance add-on that attaches to the zEnterprise System.

– Enable application integration with System z transaction processing, messaging, and data serving capabilities using select IBM POWER7® blades (supporting AIX®) and/or select IBM System x® blades (supporting Linux and Microsoft Windows).

– Secure your Service Oriented Architecture (SOA) and Web environments, simplify your connectivity infrastructure, provide multiple levels of XML optimization, and govern your evolving IT architecture using IBM WebSphere® DataPower® Integration Appliance XI50 for zEnterprise (DataPower XI50x).

- **IBM zEnterprise Unified Resource Manager (Unified Resource Manager)**
  Unified Resource Manager is part of the Hardware Management Console that manages an ensemble. This includes providing energy monitoring and management, goal-oriented policy management, increased security, virtual networking, virtual server lifecycle management, and data management for the physical and logical resources of an ensemble.

  An ensemble is a collection of one to eight zEnterprise CPCs, including any optionally attached zBX, that are managed as a single logical virtualized system by the Unified Resource Manager.

  These management capabilities are available using the Hardware Management Console (HMC) user interface and application programming interfaces (APIs).

- **IBM DB2 Analytics Accelerator for z/OS V2.1**
  IBM DB2 Analytics Accelerator for z/OS V2.1 is a workload-optimized, LAN-attached applicance based on Netezza® technology. It is a blending of System z and Netezza technologies that delivers unparalleled mixed workload performance for addressing complex analytic business needs.

  IBM DB2 Analytics Accelerator for z/OS V2.1 is not integrated into a zBX and is not managed by Unified Resource Manager. It does not require or exploit zEnterprise ensemble capabilities.

- **Reduction in the preplanning requirements by:**
  – Providing a fixed HSA (16 GB)
  – Reducing the number of Power-on-Resets
  – Allowing dynamic add/remove of a new LPAR to a new or existing LCSS.

- **Up to a maximum of 128 coupling CHPIDs and 96 physical extended distance coupling links (1x IFB or ISC-3)**

- **Coupling using InfiniBand®**
  A zEnterprise to zEnterprise connection, using HCA3-O fanout cards, is provided by:
  – A 12x Infiniband fiber optic link using the 12x IFB3 protocol if four or fewer CHPIDs are defined per HCA3-O port. If more than four CHPIDs are defined per HCA3-O port, the 12x IFB protocol is used. This 12x Infiniband fiber optic link has a link data rate of 6 GBps (Gigabytes) and a maximum link distance of 150 meters (492 feet). The 12x IFB3 protocol improves service times.

  A zEnterprise to zEnterprise, zEnterprise to System z10®, or System z10 to System z10 connection is provided by:
  – A 12x InfiniBand fiber optic link using the 12x IFB protocol with a link data rate of 6 GBps (Gigabytes) and a maximum link distance of 150 meters (492 feet)
  – A 1x InfiniBand fiber optic link using the 1x IFB protocol with a link data rate of 2.5 Gbps (Gigabits) and a maximum unrepeated distance of 10 kilometers (6.2 miles) or a maximum repeated distance of 100 kilometers (62 miles)

  - A 1x InfiniBand fiber optic link using the 1x IFB protocol with a link data rate of 5 Gbps (Gigabits) maximum unrepeated distance of 10 kilometers (6.2 miles) or a maximum repeated distance of 100 kilometers (62 miles).

  A zEnterprise to System z9® or System z10 to System z9 connection is provided by:
  - A 12x Infiniband fiber optic link using the 12x IFB protocol with a link data rate of 3 GBps (Gigabytes) and a maximum link distance of 150 meters (492 feet).

- **InterSystem Channel-3 (ISC-3)**
  The ISC-3 link is a member of the family of coupling link options. ISC-3 has a link data rate of 2 Gbps. It is used by coupled servers to pass information back and forth over high speed links in a Parallel Sysplex® environment for unrepeated distances up to 10 km (6 miles) and repeated distances up to 100 km (62 miles).
- **Optical fanout cards (HCA2-O, HCA2-O LR, HCA3-O, HCA3-O LR)**
  HCA2-O, HCA2-O LR, HCA3-O, and HCA3-O LR cards are used for coupling using InfiniBand on a zEnterprise.

  HCA2-O supports 12x InfiniBand at 3 GBps or 6 GBps. HCA2-O LR supports 1x InfiniBand at 2.5 Gbps or 5 Gbps. HCA3-O supports 12x InfiniBand at 6 GBps. HCA3-O LR supports 1x InfiniBand at 2.5 or 5 Gbps.
- **Copper fanout cards (HCA2-C and PCIe)**
  HCA2-C fanout card supports two ports. Each port uses a 12x InfiniBand copper cable (6 GBps in each direction) providing a connection to the I/O drawer or I/O cage in a z196.

  PCIe fanout cards supports two ports. Each port uses an x16 PCIe Gen2 copper cable (8 GBps in each direction) providing a connection to the PCIe I/O drawer.
- **Server Time Protocol function** provides:
  - The only time synchronization for zEnterprise using the FSP/STP card.
  - Going away signal sent when entering a failed (check stopped) state
  - Multisite sysplex distance to 100 km.
  - Coexistence of non-zEnterprise servers and coupling facilities (CFs) synchronized in an ETR network with servers and CFs that are synchronized with Coordinated Server Time (CST)
  - Concurrent migration from an ETR network
  - Messages over ISC-3 links and InfiniBand (IFB) links
  - NTP client support. The NTP client attaches to an NTP server that provides time accuracy across heterogeneous platforms in an enterprise
  - Enhanced accuracy to an external time source using pulse per second (PPS) output from NTP server
  - Use of the Hardware Management Console as an NTP server configured for use as the external time source
  - Continuous availability of NTP servers used as an external time source
  - Enhanced STP recover when the Internal Battery Feature is in use
  - Ability to save the STP configuration and time information across Power® on Resets (POR) or power outages for a single or dual server STP-only CTN
  - Automation of STP CTN reconfiguration using the System z application programming interface (API)
  - Ability to notify z/OS when events related to accessing an external time source occur.
- **Geographically Dispersed Parallel Sysplex™ (GDPS®)** is an integrated, automated application and data availability solution designed to provide the capability to manage the remote copy configuration and storage subsystem(s), automate Parallel Sysplex operational tasks, and perform failure recovery from a single point of control, thereby helping to improve application availability.
- **Internet Protocol Version 6 (IPv6) support**
  IPv6 is available for the Hardware Management Console and Support Element customer network, the TKE network connection to operating system images, OSA-Express4S, OSA-Express3, OSA-Express2, and HiperSockets™. IPv6 is the protocol designed by the Internet Engineering Task Force (IETF) to replace Internet Protocol Version 4. IPv6 expands the IP address space from 32 bits to 128 bits enabling a far greater number of unique IP addresses.

- **Server/Application State Protocol (SASP) support for load balancing**
  Using the performance data about the virtual servers that it manages, the Unified Resource Manager can provide load balancing recommendations to the configured external routers (also called load balancers). These recommendations enable the external routers to distribute incoming work more effectively among virtual servers that are defined to a load balancing group. To receive these recommendations from the Unified Resource Manager, the external routers must support the Server/Application State Protocol (SASP) communication mechanism.

  The Load Balancing Report lists the load balancing groups and group members receiving work requests and lists the recommended weights for each load balancing group member.

  SASP support for load balancing is available for these types of virtual servers:
  - Virtual servers running AIX on a POWER blade
  - Virtual servers running Windows or Linux on a System x blade
  - z/VM guests running Linux.
- **Power estimating and monitoring functions:**
  - Power Estimator tool on Resource Link.
  - Monitoring of power consumption and thermal loading using the **Activity** task and the **Monitors Dashboard** task on the Hardware Management Console.

    The **Monitors Dashboard** task also allows you to export this data to a read-only spreadsheet format and to create histograms showing processor usage, channel usage, power consumption, or input air temperature data over a specified time interval. You can also set thresholds for processor usage, channel usage, power consumption, and input air temperature. This provides an indication when the values are outside your specified threshold.
  - Support for IBM Systems Director Active Energy Manager™ for x86, IBM Systems Director Active Energy Manager for POWER, and IBM Systems Director Active Energy Manager for Linux on System z, which can monitor power and thermal data for zEnterprise, as well as other systems.
- **Historical view of power, temperature, and utilization data of your system**
  Using the **Environment Efficiency Statistics** task on the Hardware Management Console and Support Element, you can display a historical view of system power consumption, system temperature, blade CPU utilization, and CP utilization data. This data will assist you in monitoring the performance of your system.
- **Energy consumption reduction**
  You can reduce the energy consumption of your system by enabling power saving mode or setting a peak power consumption limit. To enable power saving mode, use the **Set Power Saving**, **Customize Schedule Operations**, or **Customize/Delete Activation Profiles** Hardware Management Console and Support Element tasks, the Active Energy Manager (AEM), or the Simple Network Management Protocol (SNMP), Common Information Model (CIM), or Hardware Management Console Web Services (Web Services) APIs. To limit the peak power consumption of zBX blades, use the **Set Power Cap** task.
- **Capacity on Demand functions, which include:**
  - Ability to perform a permanent LICCC upgrade while temporary resources are active
  - Ability to install and activate multiple temporary records at any given time
  - Ability to activate partial resources on a single temporary record
  - Disaster recovery solutions:
    Capacity for Planned Events (CPE) - Short range - 3 days
    Capacity Backup (CBU) - Long range - 90 days
  - Capacity provisioning, which provides a means of managing your processing capacity based on business needs
  - Ability to prepay for On/Off CoD upgrades
  - Ability to set spending limits when ordering an On/Off record
  - Ability to order permanent unassigned engines
  - Ability to order an administrative On/Off test record, which allows you to order, download, activate, and deactivate On/Off upgrades without actually setting real capacity or incurring costs
  - Automatic renewal of On/Off CoD records
  - Automatic installation of up to four CPE and CBU records on an initial z196 order

  - 45 available subcapacity settings.
- **HiperSockets** provides high-speed communications between partitions on a zEnterprise, System z10, and System z9 running several different operating systems, including z/OS, z/VM®, z/VSE®, and Linux on System z. HiperSockets requires no physical cabling. A single logical partition can connect up to 32 HiperSockets. HiperSockets can also be used with the intraensemble data network (IEDN) for data communications.
- **HiperDispatch** helps provide increased scalability and performance of higher *n*-way and multibook z196 systems by improving the way workload is dispatched across the server. HiperDispatch accomplishes this improvement by recognizing the physical processor where the work was started and then dispatching subsequent work to the same physical processor. This intelligent dispatching helps reduce the movement of cache and data, and improves CPU time and performance.

  Support to dynamically optimize the CPU-to-book allocation of physical processor (PUs).

  HiperDispatch is available with zEnterprise PR/SM™, System z10 PR/SM, and z/OS functions.
- **Large page support (1 MB pages)** provides performance improvement for a select set of applications, primarily long running memory access intensive applications.
- **Reduced impact of planned and unplanned server outages** through:
  - Enhanced book availability
  - Redundant I/O interconnect
  - Enhanced driver maintenance
  - Dynamic oscillator card switchover
  - Dynamic FSP/STP card switchover
  - Program directed re-IPL
  - System-initiated CHPID reconfiguration
  - Concurrent HCA fanout card hot-plug and rebalance.
- **Enhanced driver maintenance** allows Licensed Internal Code (LIC) updates to be performed in support of new features and functions. When properly configured, z196 is designed to support activating a selected new LIC level concurrently. Certain LIC updates are not supported by this function.
- **Enhanced book availability** allows a single book, in a multibook server, to be concurrently removed from the server and reinstalled during an upgrade or repair action. Enhanced book availability is an extension of the support for Concurrent Book Add (CBA).
- **Flexible memory** provides the additional resources to maintain a constant level of memory.
- **Redundant I/O interconnect** helps maintain critical connections to devices. z196 allows a single book, in a multibook server, to be concurrently removed and reinstalled during an upgrade or repair, continuing to provide connectivity to the server I/O resources using a second path from a different book.
- **Up to 60 logical partitions (LPARs)**
- **Server consolidation**

  The expanded capacity and enhancements to the I/O infrastructure facilitates the consolidation of multiple servers into one z196 with increased memory and LPARs, which might allow you to reduce the number of servers while hosting additional applications.

  z196 provides the ability to define up to four logical channel subsystems (LCSS). Each LCSS can support up to 256 CHPID definitions and 15 LPARs (up to a maximum of 60 LPARs per system).
- **Water-cooling option**
  A water-cooling option is available to cool the server using customer supplied chilled water and a special water circulation unit within the frame.
- **Top exit cabling**
  z196 provides the ability to route I/O cables through the top of the frame.
- **Frame bolt-down kits**
  Two optional bolt-down kits (one for refrigerated/air-cooled models and one for water-cooled models)

are available to help secure the frames and its contents from damage when exposed to vibrations and shocks. The kits supply parts to cover raised floor heights from 9-13 inches, 12-22 inches, and 12-36 inches.

- **High Voltage DC universal input option**
  Ability to operate z196 using high voltage DC power (380-570 volts) in addition to AC power. The direct high voltage DC design improves data center energy efficiency by removing the need for any conversion.
- **PCIe I/O drawer**
  The PCIe I/O drawer is a PCIe based infrastructure. The PCIe I/O drawer provides increased port granularity and improved power efficiency and bandwidth over the I/O drawers.
- **ESCON (16 ports) supporting 240 channels**
- **FICON Express8S, FICON Express8 and FICON Express4**

  Note: FICON Express4 features can only be carried forward. FICON Express8 features can be carried forward or ordered on MES using an RPQ for certain machine configurations.

  FICON Express8S features:
  – FICON Express8S 10KM LX (2 channels per feature)
  – FICON Express8S SX (2 channels per feature)

  FICON Express8 features:
  – FICON Express8 10KM LX (4 channels per feature)
  – FICON Express8 SX (4 channels per feature)

  FICON Express4 features:
  – FICON Express4 10KM LX (4 channels per feature)
  – FICON Express4 4KM LX (4 channels per feature)
  – FICON Express4 SX (4 channels per feature)

  Enhancements:
  – T10-DIF support for FCP channels for enhanced reliability
  – High Performance FICON for System z (zHPF) for FICON Express8S, FICON Express8, and FICON Express4 features (CHPID type FC)
  – Extension to zHPF multitrack operations removing the 64 kB data transfer limit
  – Assigning World Wide Port Names (WWPNs) to physical and logical Fibre Channel Protocol (FCP) ports using the WWPN tool
- **OSA-Express4S, OSA-Express3, and OSA-Express2**

  Note: OSA-Express2 features can only be carried forward. OSA-Express3 10 Gigabit Ethernet and OSA-Express3 Gigabit Ethernet features can be carried forward or ordered on MES using an RPQ for certain machine configurations.

  OSA-Express4S features:
  – OSA-Express4S GbE LX (2 ports per feature)
  – OSA-Express4S GbE SX (2 ports per feature)
  – OSA-Express4S 10 GbE LR (1 port per feature)
  – OSA-Express4S 10 GbE SR (1 port per feature)

  OSA-Express3 features:
  – OSA-Express3 GbE LX (4 ports per feature)
  – OSA-Express3 GbE SX (4 ports per feature)
  – OSA-Express3 1000BASE-T Ethernet (4 ports per feature)
  – OSA-Express3 10 GbE LR (2 ports per feature)
  – OSA-Express3 10 GbE SR (2 ports per feature)

  OSA-Express2 features:
  – OSA-Express2 GbE LX (2 ports per feature)
  – OSA-Express2 GbE SX (2 ports per feature)
  – OSA-Express2 1000BASE-T Ethernet (2 ports per feature)

  Enhancements:

  - OSA-Express3 1000BASE-T Ethernet (CHPID type OSM) provides connectivity to the intranode management network (INMN)
  - OSA-Express4S and OSA-Express3 10 GbE (CHPID type OSX) provide connectivity and access control to the intraensemble data network (IEDN)
  - Inbound workload queuing (IWQ) and IWQ for Enterprise Extender (EE) for OSA-Express4S and OSA-Express3.
- **Cryptographic options:**
  - Configurable Crypto Express3 feature.
  - CP Assist for Cryptographic Function (CPACF), which delivers cryptographic support on every PU with data encryption/decryption. CPACF also provides a high performance secure key function that ensures the privacy of key material used for encryption operations.

    CPACF support includes AES for 128-, 192-, and- 256-bit keys; SHA-1, SHA-224, SHA-256, SHA-384, and SHA-512 for message digest; PRNG, DES, and TDES

    CPACF supports the following Message-Security Assist 4 instructions: Cipher Message with CFB (KMF), Cipher Message with Counter (KMCTR), Cipher Message with OFB (KMO), and Compute Intermediate Message Digest (KIMD)

    Using the Support Element, you can enable or disable the encrypt DEA key and encrypt AES key functions of the CPACF.
  - Elliptic Curve Cryptography (ECC) and RSA public-key cryptography support
  - User Defined Extension (UDX) support.
  - Remote loading of ATMs and POS keys.
  - Dynamically add, move, or delete a Crypto Express3 feature to or from an LPAR.
  - Cryptographic migration wizard on TKE for migrating configuration data from one Cryptographic coprocessor to another Cryptographic coprocessor.
  - The tamper-resistant hardware security module, which is contained within the Crypto Express3 is designed to meet the FIPS 140-2 Level 4 security requirements for hardware security modules.
- **Fiber Quick Connect (FQC)**, an optional feature, is a fiber harness integrated in the z196 frame for a "quick" connect to ESCON and FICON LX channels.
- **Simple Network Management Protocol (SNMP) Client Libraries 3.0 support**
- **Common Information Model (CIM) API support**
- **Hardware Management Console Web Services (Web Services) API support**
- **CFCC Level 17 support**
- **TKE 7.1 Licensed Internal Code (LIC) support**
- **z/VM-mode partition (LPAR) support** to contain processor types (CPs, IFLs, zIIPs, zAAPs, and ICFs)
- **Plan ahead memory**, an optional feature, allows you to preplan to future memory upgrades. The memory upgrades can be made nondisruptively and also concurrently.
- **Worldwide Port Name (WWPN) tool**

  The WWPN tool assists you in preplanning and setting up your Storage Area Networks (SANs) environment. The WWPN tool can calculate WWPNs for virtual and physical ports before the installation of z196. Therefore, you can be up and running much faster after the server is installed. This tool applies to all FICON channels defined as CHPID type FCP (for communication with SCSI devices). The WWPN tool is located on Resource Link.
- **Support to control user access to the Hardware Management Console** using a pattern name that defines:
  - Search criteria used to identify specific user IDs
  - LDAP server used for authentication
  - Hardware Management Console user ID template used to identify logon permissions for the user IDs using this template
  - List of Hardware Management Consoles that can be accessed.
- **EAL5+ certification**

  The z196 has achieved a Common Criteria certification at an EAL5+ level for the security of its LPARs that run under the control of the Processor Resource/Systems Manager™ (PR/SM).

- **Enhanced security using digital signatures**

  Digitally Signed Firmware (Licensed Internal Code) support provided by the Hardware Management Console and the Support Element. This support provides the following benefits:

  – Ensures that no malware can be installed on System z products during firmware updates (such as, transmission of MCL files, delivery of code loads, and restoration of critical data)

  – Designed to comply to FIPS (Federal Information Processing Standard) 140-2 Level 1 for Cryptographic LIC (Licensed Internal Code) changes.

- **Auditability function**Hardware Management Console/Support Element tasks are available to generate, view, save, and offload audit reports (**Audit & Log Management** task), to set up a schedule for generating, saving, and offloading audit information (**Customize Scheduled Operations** task), to receive email notifications for select security log events (**Monitor** task), and to remove the predefined password rules to prevent them from being mistakenly used (**Password Profiles** task).

  You can also manually offload or set up a schedule to automatically offload Hardware Management Console and Support Element log files, which can help you satisfy audit requirements.

z196 models

z196 (machine type 2817) is offered in five models. The model naming is representative of the maximum number of customer configurable processor units (PUs) in the system. PUs are delivered in single engine increments orderable by feature code. The model number also reflects the number of books installed. See Table 2. It lists the z196 models, maximum number of books allowed, and the corresponding feature codes.

Table 2. z196 model feature codes

Model number	Maximum number of books allowed	Refrigerated/Air-cooled feature code	Water-cooled feature code
M15	One book	FC 1125	FC 1130
M32	Two books	FC 1126	FC 1131
M49	Three books	FC 1127	FC 1132
M66	Four books	FC 1128	FC 1133
M80	Four books	FC 1129	FC 1134

The following table lists the five z196 models and some of their characteristics, such as range of PUs allowed, the memory range of each model, and the number of I/O cages, I/O drawers, and PCIe I/O drawers that can be installed. These values are the same for both the refrigeration/air-cooled and the water-cooled models. The table lists the maximum values. These values are affected by the number of fanout cards ordered and available.

Table 3. Characteristics of the z196 models

Models	Processor Units (PUs)	Memory[1]	I/O cages [2]	I/O drawers PCIe I/O drawers
M15	1 to 15	32 GB to 704 GB	0 to 2	0 to 6 0 to 5
M32	1 to 32	32 GB to 1520 GB	0 to 2	0 to 6 0 to 5
M49	1 to 49	32 GB to 2288 GB	0 to 2	0 to 6 0 to 5
M66	1 to 66	32 GB to 3056 GB	0 to 2	0 to 6 0 to 5
M80	1 to 80	32 GB to 3056 GB	0 to 2	0 to 6 0 to 5

Table 3. Characteristics of the z196 models (continued)

Models	Processor Units (PUs)	Memory[1]	I/O cages [2]	I/O drawers PCIe I/O drawers
Note:				
1. A portion of this memory (16 GB) is delivered and reserved for HSA.				
2. With RPQ 8P2506, up to three I/O cages can be ordered.				

The system model number does not reflect the number of processors that are enabled for use. It reflects only the maximum number of customer-used processors that can be enabled when purchased.

The CP features offered have varying levels of capacity. The capacity setting is based on the quantity and type of CP feature. It is identified by a ***model capacity indicator***. The model capacity indicator identifies the number of active CPs rather than the total physical PUs purchased and identifies the type of capacity. The model capacity indicators are identified as 7*xx*, 6*xx*, 5*xx*, and 4*xx*, where *xx* is the number of active CP features. 7*xx* is a full capacity identifier. 6*xx*, 5*xx*, and 4*xx* are subcapacity identifiers. Subcapacity CP features provide reduced capacity relative to the full capacity CP feature. While you can only have up to 15 subcapacity CPs, you are not limited to one book.

For example, model capacity indicator “700” indicates no active CPs at full capacity. Model capacity indicator “510” indicates 10 active CP5 type PUs.

Table 4. Model capacity indicators

	PU Type (Capacity Identifier)	Allowable Quantity (*xx* =) [1]
Full capacity	CP7 (7*xx*)	00 - 15 00 - 32 00 - 49 00 - 66 00 - 80
Subcapacity	CP6 (6*xx*)	1 - 15 [2]
Subcapacity	CP5 (5*xx*)	1 - 15 [2]
Subcapacity	CP4 (4*xx*)	1 - 15 [2]
Note:		
1. Depends on the 2817 model.		
2. For all 2817 models.		

Performance

With the expanded capacity of z196 and enhancements to the I/O infrastructure, IBM continues to facilitate the consolidation of multiple servers into one z196 with a substantial increase in:

- Available memory
- Advanced virtualization technologies
- LPARs
- Speed using InfiniBand
- Available processors in a single footprint
- 5.2 GHz high frequency z196 quad core processor chip.

IBM's Large Systems Performance Reference (LSPR) method provides comprehensive z/Architecture® processor capacity data for different configurations of Central Processor Units across a wide variety of system control program and workload environments. The LSPR ratios defining the capacity for z196 are associated with the 701 - 780 model names as discussed in the previous section.

For more information about LSPR, see *http://www.ibm.com/servers/resourcelink/lib03060.nsf/pages/lsprindex?OpenDocument*.

Resource Link

Resource Link is a key component in getting your z196 server up and running and maintained. Resource Link provides: customized planning aids, a CHPID Mapping Tool, Customer Initiated Upgrades (CIU), power estimation tool, and education courses. Refer to Appendix B, "Resource Link," on page 149 for detailed information about Resource Link and all the functions that it can assist you with your z196.

Fiber optic cabling

To serve the cabling needs of System z customers, IBM Site and Facilities Services has fiber optic cabling services available whether the requirements are product-level or enterprise-level. These services consider the requirements for the protocols and media types supported on z196 (for example, ESCON, FICON, OSA-Express) whether the focus is the data center, the Storage Area network (SAN), the Local Area Network (LAN), or the end-to-end enterprise.

The IBM Site and Facilities Services is designed to deliver convenient, packaged services to help reduce the complexity of planning, ordering, and installing fiber optic cables. The appropriate fiber cabling is selected based upon the product requirements and the installed fiber plant.

See to Chapter 8, "Cabling," on page 111 for additional information.

z/Architecture

The z196, like its predecessors, support 24, 31 and 64-bit addressing, as well as multiple arithmetic formats. High-performance logical partitioning via Processor Resource/System Manager (PR/SM) is achieved by industry-leading virtualization support provided by z/VM. The z/Architecture also provides key technology features such as HiperSockets and the Intelligent Resource Director, which result in a high speed internal network and an intelligent management with dynamic workload prioritization and physical resource balancing.

IBM's z/Architecture or a characteristic of a particular implementation includes:
- New high-frequency z196 processor chip (5.2 Ghz operation in system)
- Out-of-order execution of instructions
- Hardware accelerators on the chip for data compression, cryptographic functions and decimal floating point
- Integrated SMP communications
- Instructions added to z196 chip to improve compiled code efficiency
- Enablement for software/hardware cache optimization
- Support for 1MB segment frame
- Full hardware support for Hardware Decimal Floating-point Unit (HDFU)
- 64-bit general registers
- 64-bit integer instructions. Most ESA/390 architecture instructions with 32-bit operands have new 64-bit and 32- to 64-bit analogs
- 64-bit addressing is supported for both operands and instructions for both real addressing and virtual addressing
- 64-bit address generation. z/Architecture provides 64-bit virtual addressing in an address space, and 64-bit real addressing.
- 64-bit control registers. z/Architecture control registers can specify regions and segments, or can force virtual addresses to be treated as real addresses
- The prefix area is expanded from 4K to 8K bytes
- Quad-word storage consistency
- The 64-bit I/O architecture allows CCW indirect data addressing to designate data addresses above 2GB for both format-0 and format-1 CCWs

- The 64-bit SIE architecture allows a z/Architecture server to support both ESA/390 (31-bit) and z/Architecture (64-bit) guests and Zone Relocation is expanded to 64-bit for LPAR and z/VM
- 64-bit operands and general registers are used for all cryptographic instructions
- The implementation of 64-bit z/Architecture can help reduce problems associated with lack of addressable memory by making the addressing capability virtually unlimited (16 exabytes).

For more detailed information about z/Architecture and a list of the supported instructions and facilities, see the *z/Architecture Principles of Operation*. To determine what facilities are present in your configuration, you can use the STORE FACILITY LIST EXTENDED instruction. Information about how to use this instruction is described in the *z/Architecture Principles of Operation*.

Upgrade progression

You can upgrade within the z196 family. However, an upgrade to model M80 or an upgrade from refrigeration/air-cooled model to a water-cooled model cannot be performed concurrently.

You can also upgrade from a System z9 EC model or System z10 EC model to an IBM zEnterprise 196 model.

An upgrade includes all frames, cages, drawers, support cards, and new I/O features.

Unsupported features/functions

This section lists the features/functions that are **not** supported on z196 and a recommended alternative, if applicable.

- FICON Express® and FICON Express2
- ETR feature
- Crypto Express2
- ICB-4 links
- ICB-3 links
- ICB-2 links
- ISC-3 links in Compatibility Mode
- OSA-Express
- OSA-Express2 10 GbE LR
- PCIXCC and PCIC

Chapter 2. Hardware characteristics

This chapter describes the hardware features and functions for the five z196 (machine type 2817) models: M15, M32, M49, M66, and M80

Note: You can also see *zEnterprise 196 Installation Manual for Physical Planning*, available on Resource Link at *http://www.ibm.com/servers/resourcelink*, for initial system physical planning requirements.

System frame configuration

The z196 frames (the A frame and Z frame) are enclosures built to Electronic Industry Association (EIA) standards. The A frame and Z frame (also called the zCPC) make up the z196 configuration shown in the following figure:

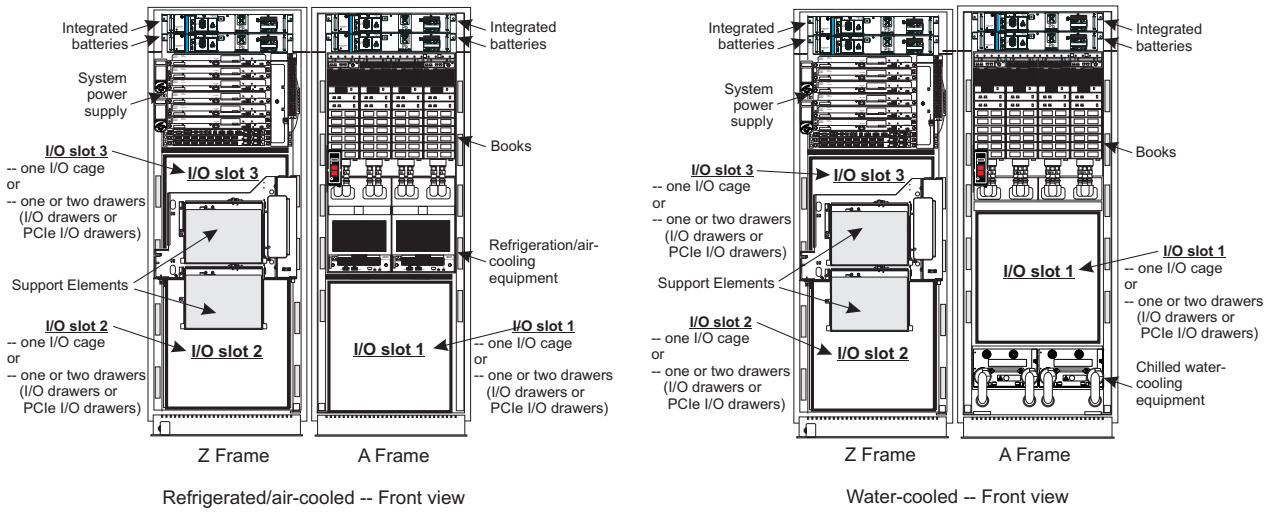


Figure 2. z196 sample frame configuration - front view

The basic A frame consists of:
- Two oscillator (OSC) cards
- Two Flexible Service Processor/Server Time Protocol (FSP/STP) cards
- One to four books
- Input/Output (I/O) cage, I/O drawers, or PCIe I/O drawers for channel attachment capability
- Cooling equipment.

The basic Z frame consists of:
- System power supply
- Two internal Support Elements (Support Elements)
- Optional I/O cages, I/O drawers, or PCIe I/O drawers.

Both the A frame and the Z frame can include the optional Internal Battery Features (IBFs). They are installed in pairs in the top of the A frame and Z frame. There can be up to six IBFs installed. See "Internal Battery Feature" on page 25 for more information.

Note: If all book positions are not used, the following are required:
- Air block book for book positions not containing a physical book

- FSP/STP cards air flow book for unpopulated positions.

Oscillator (OSC) card

Two quarter high oscillator (OSC) cards (FC FPH604) are required on z196. These oscillator cards serve as a primary card and a backup card. If the primary oscillator card fails, the backup card detects the failure and continue to provide the clock signal preventing an outage due to an oscillator failure.

FSP/STP cards

Two FSP/STP cards (FC FPH603) are required on z196. If the active FSP/STP card fails, there is an automatic switchover to the other FSP/STP card. Each card contains one Pulse per second (PPS) port. A cable connection from the PPS port on the FSP/STP card to the PPS output of the Network Time Protocol (NTP) server is required when the z196 is using STP and configured in an STP-only CTN using NTP with pulse per second as the external time source.

FSP/STP cables are customer supplied.

Books

A z196 system consists of one to four books. The books are referred to as the first book, second book, third book, and fourth book. Each book consists of:

- One multichip module (MCM)
- 30 memory dual inline memory module (DIMM) slots
- Up to eight HCA fanout cards (depending on the number of books installed)
- Two FSP cards
- Three distributed converter assembly (DCA) cards. The DCA cards are mounted directly to the rear of the book.

The books are connected using a point-to-point SMP network. Point-to-point connectivity allows the books to be plugged from left to right, but due to airflow and heating restrictions, a plugging sequence will be specified. This design provides growth paths up to a 80 engine system where each of the 80 PUs have full access to all system resources, specifically memory and I/O. The plugging sequence is — the first book plugs into slot 6, the second book plugs into slot 15, the third book plugs into slot 10, and the fourth book plugs into slot 1. Figure 3 on page 15 displays the book positions in the z196.

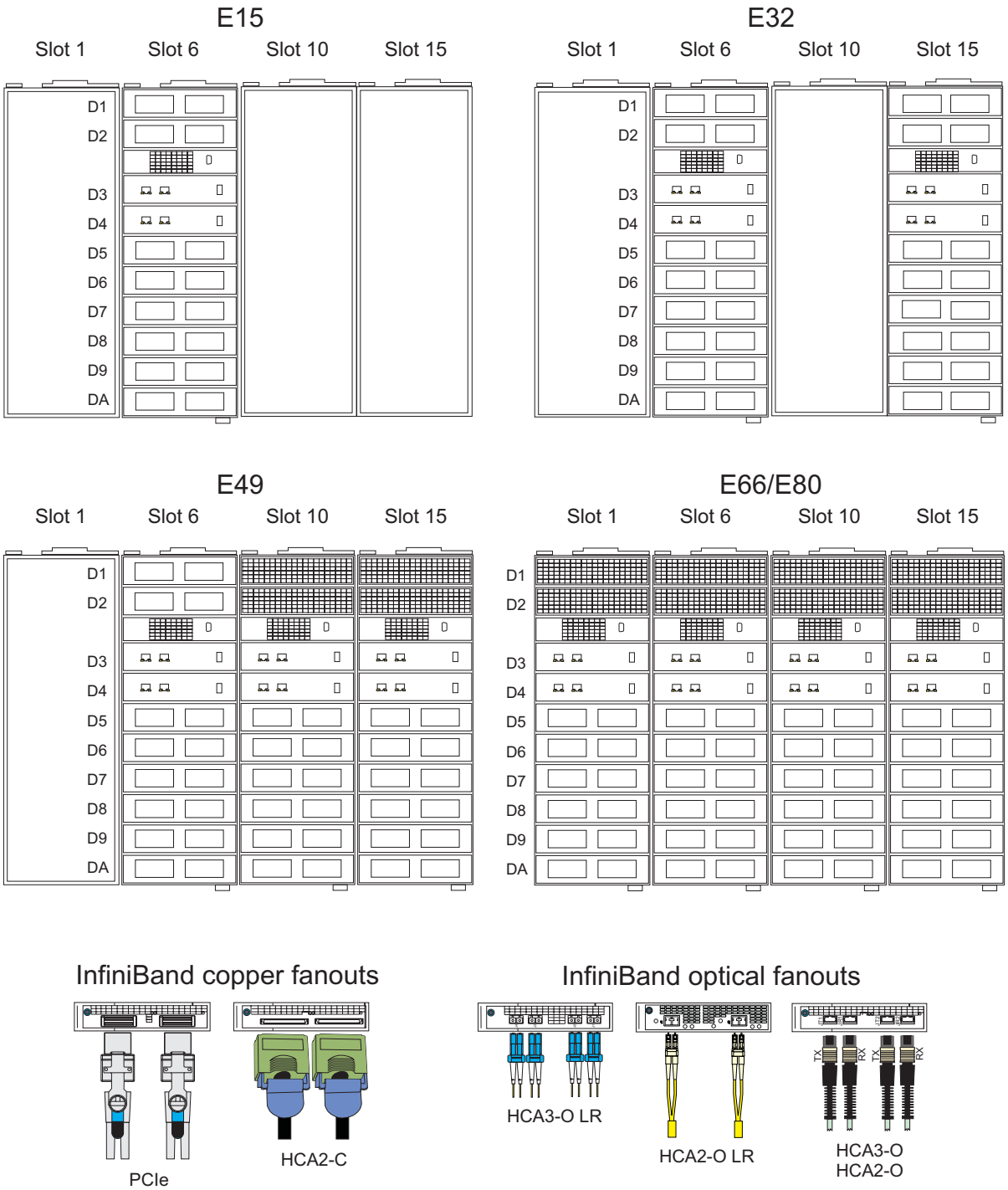


Figure 3. z196 book positions

The number of processor books in your system determines which z196 model you have. Figure 4 on page 16 identifies these models. Figure 4 on page 16 also identifies available PUs, the number of SAPs and spares provided to you as standard PUs, and the MCM size for each book on each z196 model.

	First Book				Second Book				Third Book				Fourth Book			
Model	Avail CPs	SAPs	Spare	MCM Size	Avail CPs	SAPs	Spare	MCM Size	Avail CPs	SAPs	Spare	MCM Size	Avail CPs	SAPs	Spare	MCM Size
M15	15	3	2	20												
M32	16	3	1	20	16	3	1	20								
M49	16	3	1	20	16	3	1	20	17	3	0	20				
M66	16	3	1	20	16	3	1	20	17	3	0	20	17	3	0	20
M80	20	3	1	24	20	3	1	24	20	4	0	24	20	4	0	24

Figure 4. Book characteristics

Multichip module (MCM)

The z196 models M15, M32, M49, and M66 utilize a 20 PU MCMs. The Model M80 utilizes up to four 24 PU MCMs.

A processor unit (PU) is the generic term for the z/Architecture processor on the Multichip Module (MCM) that can be characterized as a:

- Central Processor (CP) to be used by the operating system
- Internal Coupling Facility (ICF) to be used by the coupling facility control code (CFCC)
- Integrated Facility for Linux (IFL)
- Additional System Assist Processors (SAPs) to be used by the CSS
- IBM System z Integrated Information Processor (zIIP)
- IBM System z Application Assist Processor (zAAP).

Table 5 lists the number of physical processor units (PUs) per z196 model. For each z196 model, you must select at least one CP, IFL, or ICF as shown in Table 5. Any remaining PUs can be assigned as additional SAPs or additional spares, or can be assigned to optional functions such as ICFs, IFLs, zIIPs, zAAPs, CPs, CIU, On/Off CoD, CBU, or CPE engines, as indicated in Table 5.

Table 5. PUs per z196 model

Model	Books	PUs per MCM	Active PUs			zAAPs [2]	zIIP [2]	SAPs Std	SAPs Opt	Spare PUs	CBU, CIU, On/Off CoD, CPE PUs Available
			CPs [1]	ICF [1]	IFL / uIFLS [1]						
M15	1	20	0 - 15	0 - 15	0 - 15	0 - 7	0 - 7	3	0 - 4	2	0 - 15
M32	2	20	0 - 32	0 - 16	0 - 32	0 - 16	0 - 16	6	0 - 10	2	0 - 32
M49	3	20	0 - 49	0 - 16	0 - 49	0 - 24	0 - 24	9	0 - 15	2	0 - 49
M66	4	20	0 - 66	0 - 16	0 - 66	0 - 33	0 - 33	12	0 - 20	2	0 - 66
M80	4	24	0 - 80	0 - 16	0 - 80	0 - 40	0 - 40	14	0 - 18	2	0 - 80

Notes:

1. Only one CP, ICF, or IFL is required for any model. The total number of PUs purchased cannot exceed the total number available for that model.
2. One CP must be installed with or before any zIIPs or zAAPs that are installed. The total number of zIIPs or zAAPs purchased must be less than or equal to the number of active and unassigned CPs purchased on a single machine.
3. PU selection is completed by identifying the number of features when ordering.
4. Minimum storage is 16 GB for all models.

Central Processor (CP): A Central Processor (CP) is a PU that has the z/Architecture and ESA/390 instruction sets. It can run z/VM, z/OS, z/VSE, z/TPF, and Linux on System z operating systems and the Coupling Facility Control Code (CFCC). z196 processors operate only in LPAR mode; consequently all CPs are dedicated to a partition **or** shared between partitions. Reserved CPs can also be defined to a logical partition, to allow for nondisruptive image upgrades.

All CPs within a configuration are grouped into a CP pool. Any z/VM, z/OS, z/VSE, z/TPF, and Linux on System z operating systems can run on CPs that were assigned from the CP pool. Within the capacity of the book, CPs can be concurrently added to an existing configuration permanently by using CIU or CUod, or temporarily by using On/Off CoD, CBU. and CPE.

Internal Coupling Facility (ICF): An ICF provides additional processing capability exclusively for the execution of the Coupling Facility Control Code (CFCC) in a coupling facility LPAR. Depending on the model, optional ICF may be ordered. ICFs can only be used in coupling facility logical partitions. However, it can be shared **or** dedicated, because only CFCC runs on these PUs. The use of dedicated processors is strongly recommended for production coupling facility use. Software Licensing charges are not affected by the addition of ICFs. For more information, refer to “Coupling facility” on page 91.

Integrated Facility for Linux (IFL): An IFL feature provides additional processing capacity exclusively for Linux on System z workloads with no effect on the z196 model designation. An IFL can only be used in Linux on System z or z/VM LPARs. However, it can be shared or dedicated because only Linux on System z software runs on these CPs.

IFL is an optional feature for z196. Up to 80 IFL features may be ordered for z196 models, depending upon the server model and its number of maximum unused PUs.

Software licensing charges are not affected by the addition of IFLs. For more information on software licensing, contact your IBM representative.

The IFL enables you to

- Add processing capacity dedicated to running Linux on System z on a z196 server.
- Run multiple Linux on System z images independently of the traditional z/Architecture, with associated savings of IBM z/Architecture.
- Define many virtual Linux on System z images on fewer real z196 resources.

As with any change in the LPAR configuration of a processor, the introduction of additional resources to manage may have an impact on the capacity of the existing LPARs and workloads running on the server. The size of the impact is dependent on the quantity of added resources and the type of applications being introduced. Also, one should carefully evaluate the value of sharing resources (like CHPIDs and devices) across LPARs to assure the desired balance of performance, security, and isolation has been achieved.

System z Applications Assist Processor (zAAP): The System z Application Assist Processor is a specialized processor unit that provides a Java execution environment for a z/OS environment. This enables clients to integrate and run new Java-based web application alongside core z/OS business applications and backend database systems, and can contribute to lowering the overall cost of computing for running Java technology-based workloads on the platform.

zAAPs are designed to operate asynchronously with the CPs to execute Java programming under control of the IBM Java Virtual Machine (JVM). This can help reduce the demands and capacity requirements on CPs.

The IBM JVM processing cycles can be executed on the configured zAAPs with no anticipated modifications to the Java application. Execution of the JVM processing cycles on a zAAP is a function of the Software Developer's Kit (SDK) 1.4.1 for zEnterprise, System z10, System z9, zSeries, z/OS, and Processor Resource/Systems Manager (PR/SM).

Note: The zAAP is a specific example of an assist processor that is known generically as an Integrated Facility for Applications (IFA). The generic term IFA often appears in panels, messages, and other online information relating to the zAAP.

z/VM 5.4 or later supports zAAPs for guest exploitation.

System z Integrated Information Processor (zIIP): The IBM System z Integrated Information Processor (zIIP) is a specialty engine designed to help improve resource optimization, enhancing the role of the server as the data hub of the enterprise. The z/OS operating system, on its own initiate or acting on the direction of the program running in SRB mode, controls the distribution of work between the general purpose processor (CP) and the zIIP. Using a zIIP can help free capacity on the general purpose processor.

z/VM 5.4 or later supports zIIPs for guest exploitation.

System Assist Processor (SAP): A SAP is a PU that runs the channel subsystem Licensed Internal Code (LIC) to control I/O operations. One of the SAPs in a configuration is assigned as a Master SAP, and is used for communication between the book and the Support Element. All SAPs perform I/O operations for all logical partitions.

A standard SAP configuration provides a very well balanced system for most environments. However, there are application environments with very high I/O rates (typically some z/TPF environments), and in this case additional SAPs can increase the capability of the channel subsystem to perform I/O operations. Additional SAPs can be added to a configuration by either ordering optional SAPs or assigning some PUs as SAPs. Orderable SAPs may be preferred since they do not incur software charges, as might happen if PUs are assigned as SAPs.

z/VM-mode LPARs: z196 allows you to define a z/VM-mode LPAR containing a mix of processor types including CPs and specialty processors (IFLs, zIIPs, zAAPs, and ICFs). This support increases flexibility and simplifies systems management by allowing z/VM 5.4 or later to manage guests to operate Linux on System z on IFLs, operate z/VSE and z/OS on CPs, offload z/OS system software overhead, such as DB2 workloads, on zIIPs, and provide an economical Java execution environment under z/OS on zAAPs, all in the same VM LPAR.

Memory

Each z196 has its own processor memory resources. This memory can consist of both **central** and **expanded** storage.

Central Storage (CS): Central storage consists of main storage, addressable by programs, and storage not directly addressable by programs. Nonaddressable storage includes the Hardware System Area (HSA). Central storage provides:

- Data storage and retrieval for the Processor Units (PUs) and I/O
- Communication with PUs and I/O
- Communication with and control of optional expanded storage
- Error checking and correction.

Part of central storage is allocated as a fixed-sized Hardware System Area (HSA), which is not addressable by application programs. See “Hardware System Area (HSA)” on page 19 for further information.

In z/Architecture, storage addressing is 64 bits, allowing for an addressing range up to 16 exabytes. Consequently, all central storage in a z196 (up to 3056 GB) can be used for central storage.

Key-controlled storage protection provides both store and fetch protection. It prevents the unauthorized reading or changing of information in central storage.

Each 4 KB block of storage is protected by a 7-bit storage key. For processor-initiated store operations, access key bits 0-3 from the active program status word (PSW) are compared with bits 0-3 from the storage key associated with the pertinent 4 KB of storage to be accessed. If the keys do not match, the central processor is notified of a protection violation, the data is not stored, and a program interruption occurs. PSW key 0 matches any storage key. The same protection is active for fetch operations if bit 4 of the storage key (the fetch protection bit) is on. See *zEnterprise System Processor Resource/Systems Manager Planning Guide* for more information about central storage.

Expanded Storage (ES): Expanded storage can optionally be defined on z196. It is controlled by the control program, which can transfer 4 KB pages between expanded storage and central storage. The control program can use expanded storage to reduce the paging and swapping load to channel-attached paging devices in a storage-constrained environment and a heavy-paging environment.

The z196 models offer a flexible storage configuration which streamlines the planning effort by providing a single storage pool layout at IML time. The storage is placed into a single pool which can be dynamically converted to ES and back to CS as needed. LPARs are still specified to have CS and optional ES as before. Activation of LPARs as well as dynamic storage reconfigurations causes the LPAR to convert the storage to the type needed.

The control program initiates the movement of data between main storage (the addressable part of central storage) and expanded storage. No data can be transferred to expanded storage without passing through main storage. With z196 models, a **dedicated move page engine** assists in efficiently transferring data between main and expanded storage. See *zEnterprise System Processor Resource/Systems Manager Planning Guide* for more information about expanded storage.

Memory cards: Up to 30 memory cards reside within a book. The physical card capacity of each card is 4 GB, 16 GB, and 32 GB. The sum of enabled memory on each card is the amount available for use in the system.

The following list contains some general rules for memory.

- Each book can contain different physical capacities from other books.
- The physical memory card capacity is not required to be identical on each book installed in the system.
- Larger capacity cards can be used for repair actions and manufacturing substitution. LICCC dials down to the ordered size.
- Memory downgrades are not supported.
- Minimum memory orderable on any model is 32 GB.
- Memory can only be upgraded in 16 GB increments between the defined minimum and maximum.
- Changing an installed physical memory card is not a disruptive action using Enhanced Book Availability.
- LICCC dialing is used to offer concurrent memory upgrades within the physical memory card installed.

Hardware System Area (HSA): The HSA contains the Licensed Internal Code (LIC) and configuration-dependent control blocks. HSA is not available for program use. The HSA has a fixed size of 16 GB. Customer storage is not reduced due to HSA size increase on a GA upgrade within the same machine family because an additional 16 GB is always delivered and reserved for HSA.

Error Checking and Correction (ECC): Data paths between central storage and expanded storage (if configured), and between central storage and the central processors and channels are checked using either parity or Error Checking and Correction (ECC). Parity bits are included in each command or data word. ECC bits are stored with data in central storage. ECC codes apply to data stored in and fetched from central storage. Memory ECC detects and corrects single bit errors. Also, because of the memory structure design, errors due to a single memory chip failure are corrected. Unrecoverable errors are flagged for follow-up action. ECC on z196 models is performed on the memory data bus as well as memory cards.

Fanout cards

z196 can have one to four books. Each book has eight (8) fanout slots and each fanout slot can accommodate one fanout card. There are six fanout cards that will plug into the z196 – HCA2-O fanout card, HCA2-O LR fanout card, HCA2-C fanout card, HCA3-O fanout card, HCA3-O LR fanout card, PCIe fanout card.

The HCA2-O, HCA2-O LR, HCA3-O, HCA3-O LR fanout cards are used for coupling using fiber optic cabling. The maximum number of HCA2-O, HCA2-O LR, HCA3-O, HCA3-O LR features is 16.

The HCA2-O fanout supports a two-port 12x IFB coupling link with a link data rate of 3 GBps (if attached to a z9) and 6 GBps (if attached to a z10 or zEnterprise), and a maximum distance of 150 meters (492 feet). The HCA2-O LR fanout supports a two-port 1x IFB coupling link with a link data rate of either 5.0 Gbps or 2.5 Gbps and a maximum unrepeated distance of 10 kilometers (6.2 miles) and a maximum repeated distance of 100 kilometers (62 miles).

The HCA2-C fanout card supports two ports. Each port uses a 12x InfiniBand copper cable (6 GBps in each direction) providing a connection to an I/O cage or I/O drawer.

The HCA3-O fanout supports a two-port 12x IFB coupling link with a link data rate of 6 GBps and a maximum distance of 150 meters (492 feet). The HCA3-O fanout also supports the 12x IFB3 protocol if four or less CHPIDs are defined per port. The 12x IFB3 protocol provides improved service times. An HCA3-O fanout can communicate with an HCA2-O fanout on z196, z114, or z10™.

The HCA3-O LR fanout is designed to support four-port 1x IFB coupling link with a link data rate of 5.0 Gbps and a maximum unrepeated distance of 10 kilometers (6.2 miles) or a maximum repeated distance of 100 kilometers (62 miles). With DWDM, the HCA3-O LR fanout supports a four-port 1x IFB coupling link with a link data rate of either 2.5 or 5 Gbps. An HCA3-O LR fanout can communicate with a HCA2-O LR fanout on z196, z114, or z10.

The PCIe fanout card provides PCIe interface and is used to connect to the PCI-IN I/O interconnect cards in the PCIe I/O drawer.

The following is a list of InfiniBand connections from z196 to a z196, z114, z10 EC, z10 BC, z9 EC, or z9 BC:

- HCA3-O fanout card on a z196 can connect to a:
  - HCA3-O fanout card on a z196 or z114
  - HCA2-O fanout card on a z196, z114, z10 EC, or z10 BC
- HCA2-O fanout card on a z196 can connect to a:
  - HCA3-O fanout card on a z196 or z114
  - HCA2-O fanout card on a z196, z114, z10 EC, or z10 BC
  - HCA1-O fanout card on a z9 EC or z9 BC
- HCA3-O LR fanout card on a z196 can connect to a:
  - HCA3-O LR fanout card on a z196 or z114
  - HCA2-O LR fanout card on a z196, z114, z10 EC, or z10 BC

- HCA2-O LR fanout card on a z196 can connect to a:
  - HCA3-O LR fanout card on a z196 or z114
  - HCA2-O LR fanout card on a z196, z114, z10 EC, or z10 BC

The number of fanout cards allowed in each book depends on the number of books installed. See Table 6 for the maximum number of fanout cards allowed.

Table 6. Maximum number of fanout cards per book per structure

Book Structure	Maximum # of fanout cards allowed in the first book	Maximum # of fanout cards allowed in the second book	Maximum # of fanout cards allowed in the third book	Maximum # of fanout cards allowed in the fourth book	Maximum # of fanout cards in the structure
One-book structure	8	-	-	-	8
Two-book structure	8	8	-	-	16
Three-book structure	8	6	6	-	20
Four-book structure	6	6	6	6	24

The fanout cards are inserted in a specific sequence from bottom-to-top in the book – first are the HCA2-C fanout cards and PCIe fanout cards used for I/O, then the HCA3-O LR fanout cards and the HCA2-O LR fanout cards, and last the HCA3-O fanout cards and HCA2-O fanout cards.

When configuring the machine to use the best RAS characteristics, we want to use IFBs equally from each installed fanout card in each book and plug each I/O type equally across those fanout cards in those books.

Figure 5 on page 22 is a sample configuration showing connections from the fanout cards in the books on the z196 to another book or processor drawer, an I/O cage, an I/O drawer, or a PCIe I/O drawer.

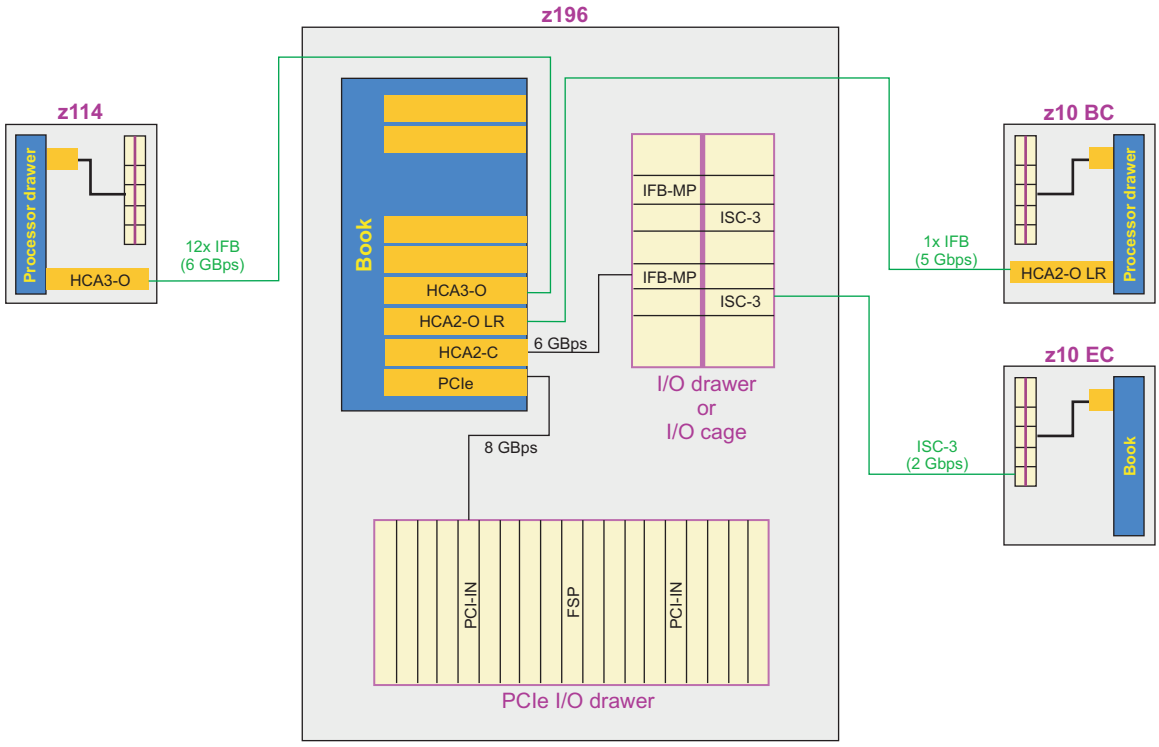


Figure 5. HCA fanout connections

Distributed Converter Assembly (DCA) cards

The Distributed Converter Assembly (DCA) cards are DC-to-DC converter cards in the books that convert -350 volts DC to logic voltages. There are a three DCA cards per book.

I/O cages, I/O drawers, and PCIe I/O drawers

The z196 supports up to two I/O cages (RPQ 8P2506) is required to order the third I/O cage) and six I/O drawers, five PCIe I/O drawers, or a combination of six I/O drawers and PCIe I/O drawers. The number of I/O cages, I/O drawers, or PCIe I/O drawers can be driven by the total number of cards in a configuration or by the cage or drawer power budget.

Table 7 identifies the plugging sequence (using the EIA locations on the frame) for the I/O cages, I/O drawers, and PCIe I/O drawers.

Table 7. I/O cage, I/O drawer, and PCIe I/O drawer plugging sequence

	EIA locations	
	Refrigeration/air cooled	**Water cooled**
I/O cage	A01 -> Z01 -> Z15	A08 -> Z01 -> Z15
I/O drawer	Z22 -> Z15-> Z08 -> Z01 -> A08 -> A01	Z22 -> Z15 -> Z08 -> Z01 -> A15 -> A08
PCIe I/O drawer	A01 -> A08 -> Z01 -> Z08 -> Z15 -> Z22	A08 -> A15 -> Z01 -> Z08 -> Z15 -> Z22

The I/O slots support ESCON, FICON Express4, FICON Express8, FICON Express8S, ISC-3, OSA-Express2, OSA-Express3, OSA-Express4S, and Crypto Express3 features. An I/O cage, I/O drawer, and PCIe I/O drawer allows you to add channels up to the amount supported by a particular I/O cage, I/O drawer, and PCIe I/O drawer and books. There is a system maximum of 384 channels.

Note: An addition or removal of an I/O cage is disruptive; however, an addition or removal of an I/O drawer or PCIe I/O drawer can be performed concurrently.

See Figure 11 on page 48 for the layout of an I/O cage. See Figure 10 on page 47 for the layout of an I/O drawer. See Figure 12 on page 49 for the layout of an PCIe I/O drawer.

I/O features

The I/O cards that are supported in z196 are shown in Table 8. There are a total of 28 I/O slots per I/O cage, 8 I/O slots per I/O drawer, and 32 I/O slots per PCIe I/O drawer.

See Chapter 5, "I/O connectivity," on page 59 for more detailed information about the I/O channels and adapters.

Note:

1. The Crypto Express3 feature uses I/O slots. Each Crypto Express3 feature has two PCIe adapters. The Crypto Express3 features do not have ports and do not use fiber optic cables. They are not defined in the IOCDS and, therefore, do not receive a CHPID number. However, they are assigned a PCHID.
2. The HCA2-C fanout card and PCIe fanout card are used for internal connection from a book to the I/O cage, I/O drawer, or PCIe I/O drawer. The HCA2-O, HCA2-O LR, HCA3-O, and HCA3-LR fanout cards are not I/O features. They are fanout cards plugged into the books.

Table 8. Channels, links, ports, and adapters summary per system

Feature	Maximum features	Maximum connections	Channels/ Links/ Adapters per feature	Purchase increment
16-port ESCON (FC 2323)[1]	16	240 channels [8]	16 channels[2]	4 channels
FICON Express8S 10KM LX (FC 0409) [1] FICON Express8S SX (FC 0410) [1]	160	320 channels	2 channels	2 channels
FICON Express8 10KM LX (FC 3325)[1] FICON Express8 SX (FC 3326)[1]	72 [8]	288 channels [8]	4 channels	4 channels
FICON Express4 10KM LX (FC 3321)[1, 7] FICON Express4 4KM LX (FC 3324)[1, 7] FICON Express4 SX (FC 3322)[1, 7]	72 [8]	288 channels [8]	4 channels	4 channels
OSA-Express4S GbE LX (FC 0404)[10] OSA-Express4S GbE SX (FC 0405)[10]	48 [9]	96 ports	2 ports	1 feature
OSA-Express4S 10 GbE LR (FC 0406)[10] OSA-Express4S 10 GbE SR (FC 0407)[10]	48 [9]	48 ports	1 port	1 feature
OSA-Express3 GbE LX (FC 3362)[10] OSA-Express3 GbE SX (FC 3363)[10]	24	96 ports	4 ports	4 ports
OSA-Express3 1000BASE-T Ethernet (FC 3367)[10]	24	96 ports	4 ports	4 ports
OSA-Express3 10 GbE LR (FC 3370)[10] OSA-Express3 10 GbE SR (FC 3371)[10]	24	48 ports	2 ports	2 ports
OSA Express2 GbE LX (FC 3364)[7, 10] OSA Express2 GbE SX (FC 3365)[7, 10] OSA Express2 1000BASE-T Ethernet (FC 3366)[7, 10]	24	48 ports	2 ports	2 ports
Crypto Express3 (FC 0864)[6]	8	16 PCIe adaptrs	2 PCIe adaptrs	4 PCIe adaptrs[6] 2 PCIe adaptrs[6]
ISC-3 [1]	12	48 links	4 links	1 link
12x IFB (HCA2-O) [1, 5]	8 [3] 16 [4]	16 links [3] 32 links [4]	2 links	2 links

Table 8. Channels, links, ports, and adapters summary per system (continued)

Feature	Maximum features	Maximum connections	Channels/ Links/ Adapters per feature	Purchase increment
12x IFB (HCA3-O) [1, 5]	8 [3] 16 [4]	16 links [3] 32 links [4]	2 links	2 links
1x IFB (HCA2-O LR) [1, 5]	8 [3] 16 [4]	16 links [3] 32 links [4]	2 links	2 links
1x IFB (HCA3-O LR) [1, 5]	8 [3] 12 [4]	32 links [3] 48 links [4]	4 links	4 links

Notes:

1. A minimum of one I/O feature (ESCON or FICON) or one coupling link feature (12x InfiniBand, 1x InfiniBand, or ISC-3) is required.
2. Each ESCON feature has 16 channels, of which a maximum of 15 can be activated. One is reserved as a spare.
3. Applies to Model M15.
4. Applies to Models M32, M49, M66, and M80.
5. IFBs are not included in the maximum feature count for I/O slots, but they are included in the CHPID count.
6. The initial order for Crypto Express3 is two features (four PCIe adapters). After the initial order, the minimum order is one feature (2 PCIe adapters).
7. This feature can only be carried forward.
8. RPQs may be available to increase these numbers.
9. For every OSA-Express3 feature in the configuration, the OSA-Express4S maximum number of features is reduced by two.
10. Maximum number of PCHIDs for combined OSA-Express4S, OSA-Express3, and OSA-Express2 features is 48.

IFB-MP and PCI-IN cards

The IFB-MP card can only be used in the I/O cage or I/O drawer. The IFB-MP cards provide the intraconnection from the I/O cage or I/O drawer to the HCA2-C fanout card in the books.

The PCI-IN card can only be used in the PCIe I/O drawer. The PCI-IN cards provide the intraconnection from the PCIe I/O drawer to the PCIe fanout card in the books.

Distributed Converter Assembly (DCA) cards

The Distributed Converter Assembly (DCA) cards are DC-to-DC converter cards in the I/O cages that convert -350 volts DC to logic voltages. There are two DCA cards in each I/O cage and I/O drawer.

PSC24V card

The PSC24V card is a Power Sequence Control (PSC) card used to turn on/off specific control units from the z196. The PSC24V card in the I/O cage or I/O drawer provides the physical interface between the cage controller and the PSC boxes, located outside the I/O cage or I/O drawer in the system frame. Only one PSC24V card is required on z196. Each card has two jacks that are used to connect to the PSC boxes.

The PSC feature is not supported on the PCIe I/O drawer, so the PSC24V card is not available with the PCIe I/O drawer.

Note: The PSC24V card is **not** hot pluggable.

For information about PSC, see “Power sequence controller” on page 30.

Support Element

The z196 is supplied with two integrated laptop computers that function as a primary and alternate Support Elements. Mounted inside the "Z" frame, the Support Elements communicate with the CPC and each other through the service network. The Support Element sends hardware operations and management controls to the Hardware Management Console for the CPC and allows for independent and parallel operational control of a CPC from the Hardware Management Console. The second, or alternate, Support Element is designed to function as a backup and to preload Support Element Licensed Internal Code. See Figure 2 on page 13 to view the location of the Support Elements on a z196.

The Support Element contains the following:

- Licensed Internal Code for the CPC.
- Hardware system definitions for the CPC (contained in the reset, image, and load profiles for the CPC and IOCDs).
- Battery-powered clock used to set the CPC time-of-day (TOD) clock at power-on reset. In STP timing mode, the CPC TOD clock is initialized to Coordinated Server Time (CST).
- Two 1 GB SMC Ethernet hubs (FC 0070) to manage the Ethernet connection between the Support Elements and the Hardware Management Console.

  The SMC hubs are offered as part of the initial order or as a Manufacturing Engineering Specification (MES). They are shipped automatically on every order unless you deselected FC 0070.
- An ethernet LAN adapter or LAN on board to connect the Support Element to the CPC through the power service network.

For more detailed information on the Support Element, refer to the Chapter 9, "Hardware Management Console and Support Element," on page 115 or to the *zEnterprise System Support Element Operations Guide*.

System power supply

The system power supply provides the control structure to support the z196 power requirements for the books, three I/O cages, and six I/O drawers.

The z196 power subsystem basic components include:

- Bulk Power Assembly (BPA) - provides the prime power conversion and high voltage DC distribution.
- Bulk Power Controller (BPC) - is the main power controller and cage controller for the BPA.

  The BPC is the principal control node for the z196 diagnostic/service and power/cooling system. It is the cage controller for the BPA cage and connects to both ethernet service networks.
- Bulk Power Distribution (BPD) - distributes -350 VDC and RS422 communications to logic cage power Field Replaceable Units (FRUs)
- Bulk Power Fan (BPF) - is a cooling device
- Bulk Power Regulator (BPR) - is the main front end power supply that converts line voltage (DC and AC) to regulated -350 VDC
- Bulk Power Enclosure (BPE) - is the metal enclosure that contains the back plane
- Bulk Power Hub (BPH) - is the Ethernet hub for system control and monitoring. BPH contains 32 ports (8 1-Gigabit Ethernet ports and 24 10/100 Ethernet ports).
- Internal Battery Feature (IBF) - provides battery power to preserve processor data if there is a power loss.
- Distributed Converter Assemblies (DCAs).

Internal Battery Feature

The optional Internal Battery Feature (FC 3212) provides the function of a local uninterruptible power source. It has continuous self-testing capability for battery backup which has been fully integrated into the diagnostics, including Remote Service Facility (RSF) support.

The IBF provides battery power to preserve processor data if there is a power loss on both of the AC or DC supplies.

In the event of input power interruption to the system, the IBF provides sustained system operation for the times listed in the following table.

Table 9. System IBF hold up times

	Number of I/O units[1]						
Model	**0**	**1**	**2**	**3**	**4**	**5**	**6**
M15 - 1 book	7 min[2]	5.3 min[2]	4.3 min[2]	9 min[2]	13 min[2]	10.2 min[2]	10.1 min[2]
M32 - 2 books	9.3 min[2]	14 min[2]	13 min[2]	10 min[2]	8 min[2]	7.5 min[3]	6.7 min[3]
M49 - 3 books	10.1 min[2]	9.5 min[2]	8.8 min[3]	7 min[3]	6 min[3]	5.4 min[3]	5 min[3]
M66 - 4 books	6.9 min[3]	6.4 min[3]	6 min[3]	5 min[3]	4.3 min[3]	4 min[3]	3.5 min[3]
M80 - 4 books	6.9 min[3]	6.4 min[3]	6 min[3]	52 min[3]	4.3 min[3]	4 min[3]	3.5 min[3]

Note:

1. An I/O unit is defined as follows:
   - I/O cage = 2 I/O units
   - I/O drawer or PCIe I/O drawer = 1 I/O unit

   For example, 5 I/O units can be composed of 5 (combination of I/O drawers and PCIe I/O drawers), 1 I/O cage + 3 (combination of I/O drawers and PCIe I/O drawers), or 2 I/O cages + 1 (I/O drawer or PCIe I/O drawer).
2. Single-line power cord pair
3. Two line power cord pair
4. The hold up times are valid for batteries three years old or less that have seen normal service life (2 or less complete discharges per year) with the system input power at N+1 mode of operation.

If the IBF is ordered, they must be installed in pairs. Two to six battery units are used depending on the number of Bulk Power Regulators (BPRs) required.

The IBF is fully integrated into the server power control/diagnostic system that provides full battery charge, and test and repair diagnostics. For more information about the IBF, see *zEnterprise 196 Installation Manual for Physical Planning*.

Cooling equipment

z196 provides two options for cooling the MCM – refrigeration/air-cooled or water-cooled. z196 also supports an integrated rear frame heat exchanger with the water-cooled option.

Modular Refrigeration Unit (MRU)

The Modular Refrigeration Units (MRUs) are the main components of the z196 refrigeration/air-cooled subsystem. These units are located in the A frame. Each MRU cools the MCM in one or two books. The first MRU controls the first book and the third book. The second MRU controls the second book and the fourth book. Therefore, if you install a second book, you need a second MRU.

Modular Water Unit (MWU)

The Modular Water Units (MWUs) are the main components of the z196 water-cooled subsystem. These units are located in the A frame. The water-cooled system requires a rear frame heat exchanger to support the addition of water delivery and distribution. The rear frame heat exchanger removes heat from non-MCM components in the system.

The MWUs are fully redundant. Each MWU cools the MCMs within all books that are installed. The integrated rear frame heat exchanger provides backup for the water-cooled option.

Internet Protocol Version 6

IPv6 is the protocol designed by the Internet Engineering Task Force (IETF) to replace Internet Protocol Version 4 (IPv4) to satisfy the demand for additional IP addresses. IPv6 expands the IP address space from 32-bits to 128-bits enabling a far greater number of unique IP addresses.

IPv6 is available for the Hardware Management Console and Support Element customer network, the Trusted Key Entry (TKE) workstation network connection to operating system images, OSA-Express4S, OSA-Express3, OSA-Express2, and HiperSockets.

The HMC and Support Elements are designed to support customer internal and open networks that are configured to use only IPv6 addresses, only IPv4 addresses, or a combination of the two.

IPL from an alternate subchannel set

z196 allows you to IPL from subchannel set 1 or subchannel set 2, in addition to subchannel set 0. Devices used early during IPL processing can now be accessed using subchannel set 1 or subchannel set 2. This is intended to allow the users of Metro Mirror (PPRC) secondary devices defined using the same device number and a new device type in an alternate subchannel set to be used for IPL, IODF, and stand-alone dump volumes when needed.

Multiple subchannel sets (MSS)

The multiple subchannel sets structure allows increased device connectivity for Parallel Access Volumes (PAVs). Three subchannel sets per Logical Channel Subsystem (LCSS) are designed to enable a total of 65,280 subchannels in set-0 and the addition of 64K - 1 subchannels in set-1 and set-2. Multiple subchannel sets is supported by z/OS V1.12 and Linux on System z. This applies to the ESCON, FICON, and zHPF protocols.

LPAR mode

LPAR mode is the mode of operation for the z196. It allows you to:
- Define ESA/390, ESA/390 TPF, coupling facility, z/VM-mode, and Linux-only LPARs
- Define and use up to 1024 GB in a single LPAR.
- Dynamically reconfigure storage between LPARs.

You can define and activate up to 60 LPARs for each CPC.

After you define and activate an ESA/390 or ESA/390 TPF LPAR, you can load a supporting operating system into that LPAR.

Processor Resource/Systems Manager (PR/SM) enables logical partitioning of the CPC.

Resources for each LPAR include:
- Processor Units (CPs, ICFs, IFLs, zIIPs, or zAAPs)
- Storage (central storage and expanded storage)
- Channels.

Processor units

On z196, PUs can be used within an LPAR as Central Processors (CPs), Internal Coupling Facilities (ICFs), Integrated Facilities for Linux (IFLs), System z Integrated Information Processor (zIIPs), or System z Application Assist Processors (zAAPs). The initial allocation of CPs, ICFs, IFLs, zIIPs, and zAAPs to an LPAR is made when the LPAR is activated.

Within an LPAR on z196 models, they can be used as follows:

- CPs can be dedicated to a single LPAR or shared among multiple LPARs. The use of CP resources shared between LPARs can be limited and modified by operator commands while the LPARs are active. CPs that are dedicated to an LPAR are available only to that LPAR.
- ICFs, IFLs, and zAAPs are available as orderable features on z196 models for use in an LPAR. ICFs are available as a feature for use in a coupling facility partition (see "Internal Coupling Facility (ICF)" on page 17 for additional information). IFLs are available as a feature for running Linux on System z (see "Integrated Facility for Linux (IFL)" on page 17 for additional information). zAAPs are available as a feature for providing special purpose assists that execute Java programming under control of the IBM Java Virtual Machine (JVM) (see "System z Applications Assist Processor (zAAP)" on page 17 for additional information).

Storage

Before you can activate LPAR, you must define central storage and optional expanded storage to the LPARs. See "Central Storage (CS)" on page 18 and "Expanded Storage (ES)" on page 19 for more information.

All installed storage is initially configured as central storage. Each individual LPAR is limited to a maximum of 1 TB of central storage. When an LPAR is activated, the storage resources are allocated in contiguous blocks.

For z196 models, LPAR central storage granularity is a minimum of 256 MB and increases as the amount of storage defined for the LPAR increases. You can dynamically reallocate storage resources for z/Architecture and ESA/390 architecture LPARs using **Dynamic Storage Reconfiguration**. Dynamic storage reconfiguration allows both central and expanded storage allocated to an LPAR to be changed while the LPAR is active. It provides the capability to reassign storage from one LPAR to another without the need to POR the CPC or IPL the recipient LPAR. For more information, see *zEnterprise System Processor Resource/Systems Manager Planning Guide*.

Note: You cannot share allocated central storage or expanded storage among multiple LPARs.

Expanded storage granularity for LPARs is fixed at 256 MB.

Channels

You can allocate channels to LPARs as follows:

- **Dedicated channels**

  Dedicated channels are unshared channels and can only be used by one LPAR. All channel types supported by the model can be allocated as dedicated channels.
- **Reconfigurable channels**

  Reconfigurable channels are unshared channels that can be moved among LPARs within an LCSS but can only belong to one LPAR at a given time. All channel types supported by the model can be allocated as reconfigurable channels.
- **Shared channels**

  The Multiple Image Facility (MIF) allows channels to be shared among multiple LPARs in a Logical Channel Subsystem (LCSS). Shared channels are configured to an LPAR giving the LPAR a channel image of the shared channel that it can use. Each channel image allows an LPAR to independently access and control the shared channel as if it were a physical channel assigned to the LPAR. For more information, see "Multiple Image Facility (MIF)" on page 55.

  You can define the channels, shown in Table 12 on page 44, as shared among multiple LPARs within an LCSS so that the shared channels can be accessed by more than one LPAR in an LCSS at the same time.

  On z196 models with coupling facility LPARs, CFP, CIB, and ICP channels can be shared by many ESA LPARs and one coupling facility partition.

- **Spanned channels**

  Spanned channels are channels that are configured to multiple Logical Channel Subsystems (LCSSs) and are transparently shared by any or all of the configured LPARs without regard to the LCSS to which the LPAR is configured.
- **Device Sharing**

  You can share a device among LPARs by:
  – Using a separate channel for each LPAR
  – Using a shared channel
  – Using a spanned channel.

LPAR time offset support

LPAR Time Offset support provides for the optional specification of a fixed time offset (specified in days, hours, and quarter hours) for each LPAR activation profile. The offset, if specified, is applied to the time that an LPAR receives from the Current Time Server (CTS) in a Coordinated Timing Network (CTN).

This support can be used to address the customer environment that includes multiple local time zones with a Current Time Server (CTS) in a CTN.

It is sometimes necessary to run multiple Parallel Sysplexes with different local times and run with the time set to GMT=LOCAL. This causes the results returned in the store clock (STCK) instruction to reflect local time. With LPAR time offset support, LPARs on each z196 in a Parallel Sysplex that need to do this can specify an identical time offset that will shift time in the LPAR sysplex members to the desired local time. Remaining LPARs on the z196 can continue to participate in current date production Parallel Sysplexes using the same CTS with the time provided by the CTS.

This function is supported by all in service releases of z/OS.

For more information about LPARs, see *zEnterprise System Processor Resource/Systems Manager Planning Guide* and *System z Input/Output Configuration Program User's Guide for ICP IOCP*.

Server Time Protocol (STP)

Server Time Protocol (STP) (FC 1021) provides the means for multiple zEnterprise, System z10, and System z9 servers to maintain time synchronization with each other. STP is designed to synchronize servers configured in a Parallel Sysplex or a basic sysplex (without a coupling facility), as well as servers that are not in a sysplex.

STP uses a message-based protocol to transmit timekeeping information over externally defined coupling links between servers. STP distributes time messages in layers (called stratums). The timekeeping information is needed to determine the Coordinated Server Time (CST) at each server. The coupling links used to transport STP messages include ISC-3 links configured in peer mode and IFB links. These links can be the same links already being used in a Parallel Sysplex for coupling facility communications.

For more details about Server Time Protocol, refer to "Server Time Protocol (STP)" on page 95.

For hardware and software requirements, see the STP website located at *http://www.ibm.com/systems/z/advantages/pso/stp.html*.

Hardware Management Console (HMC)

The Hardware Management Console (HMC) is a desktop PC. The HMC performs system management tasks or performs both system management tasks and ensemble management tasks. The HMC provides a single point of control and single system image for those CPCs (nodes) defined to it. (A single CPC, including any optionally attached zBX, is called a node.)

When managing an ensemble, a pair of HMCs are required – the primary HMC and the alternate HMC. The HMC managing the nodes in an ensemble is referred to as the primary HMC. The primary HMC can also manage CPCs that are not member of an ensemble. The alternate HMC is used as backup. If the primary HMC fails, the alternate HMC will inherit the role of the primary HMC.

A HMC, other than the primary HMC or the alternate HMC, can manage CPCs that are in an ensemble. However, it cannot perform any ensemble management tasks.

The HMC can manage up to 100 CPCs. However, only eight of these CPCs can be a member of an ensemble managed by that HMC. The other CPCs can be members of an ensemble managed by other HMCs. A CPC, that is not a member of an ensemble, can be managed by up to 32 HMCs. A single node can be a member of only one ensemble.

The HMCs utilize VLAN and an included PCI Express Ethernet adapter for handling both single and dual Ethernet configuration. The HMC is supplied with two Ethernet ports.

The physical location of the Hardware Management Console hardware features (standard and/or optional) are dictated by the specific PC. Some features can be mutually exclusive with other features depending on the PC model. Each CPC must be connected to at least one Hardware Management Console on the same network as the Support Elements of the CPC.

For more detailed information on the Hardware Management Console, refer to Chapter 9, "Hardware Management Console and Support Element," on page 115 or to the *System z Hardware Management Console Operations Guide*.

Top exit cabling

For z196, you can optionally route all I/O cables; ESCON, FICON, OSA-Express, 12x InfiniBand, 1x InfiniBand, and ISC-3 cables; and 1000BaseT Ethernet cables from I/O cages, I/O drawers, or PCIe I/O drawers through the top of the frame. This option (FC 7942) improves the airflow, therefore improving efficiency. The power cables can only exit through the bottom of the frame to the floor.

Extensions are added to each corner of the frame with this option.

Bolt-down kit

Two optional bolt-down kits (one for refrigerated/air-cooled models and one for water-cooled models) are available to help secure the frames and its contents from damage when exposed to vibrations and shocks. The kits supply parts to cover raised floor heights from 9-13 inches, 12-22 inches, and 12-36 inches.

Power sequence controller

The power sequence controller (PSC) is available on the z196 models. The PSC feature consists of two PSC boxes, a PSC24V card and card holder, and PSC cables. Each PSC box provides power control interfaces to external devices.

A single PSC pair supports one I/O cage or four I/O drawers. No more than two PSC features can be installed in any configuration. For the Z frame containing any I/O cages, the number of PSC features is restricted to one.

For water cooled models, the PSC boxes cannot be placed in the A frame.

IBM zEnterprise BladeCenter Extension (zBX)

zBX, machine type 2458 (model number 002), is a hardware infrastructure that consists of a BladeCenter chassis attached to a z114 or a z196. zBX can contain optimizers (IBM Smart Analytics Optimizer for DB2 for z/OS, V1.1 (IBM Smart Analytics Optimizer) and IBM WebSphere DataPower Integration Appliance XI50 for zEnterprise (DataPower XI50z)) and IBM blades (select IBM POWER7 blades and select IBM System x blades (supporting Linux and Microsoft Windows).

The IBM Smart Analytics Optimizer processes and analyzes CPU intensive DB2 queries. It provides a single enterprise data source for queries providing fast and predictable query response time. For a quick step-through of prerequisite, installation, and configuration information about the IBM Smart Analytics Optimizer, see the *IBM Smart Analytics Optimizer for DB2 for z/OS Getting Started*.

DataPower XI50z is used to help provide multiple levels of XML optimization, streamline and secure valuable service-oriented architecture (SOA) applications, and provide drop-in integration for heterogeneous environments by enabling core Enterprise Service Bus (ESB) functionality, including routing, bridging, transformation, and event handling.

DataPower XI50z firmware V5.0 for the zBX Model 002 adds support for the following items:

- WebSphere Service Registry and Repository (WSRR) subscription can help to distinguish similar Saved Search Queries and support automatic synchronization and enforcement between WSRR and DataPower. This can provide more consumable and centralized service level agreement (SLA) management.
- DataPower XI50z V5.0 appliances support the IETF Open Authorization (OAuth) 2.0 protocol. Using the OAuth protocol decreases the need to share your credentials with third parties. IT provides an authorization service separate and apart from the resource owner. OAuth is focused on the emerging Web 2.0 infrastructure and the popularity of APIs that exist to provide customizable access to an organization's applications. For example, eBay™ provides an API to provide enhanced shopping experiences by integrating with third-party applications. Twitter™ and Facebook™ provide APIs that extend their applications by providing content sharing capabilities. Each of these integrations requires focused attention on all aspects of security and the need to consider all access to be untrusted until proven otherwise.
- Authentication, authorization, and auditing (AAA) is a framework within the WebSphere DataPower firmware. DataPower takes advantage of AAA extensively to support the OAuth 2.0 protocol. AAA is used to authenticate both the resource owner's and OAuth client's identities. It is also used for authorizing a request. In release 3.8.1, DataPower introduced form-based authentication, which is tied closely with web application firewall. As of the V5.0 firmware release, the support is expanded to other service objects. DataPower XI50z V5.0 can act to protect access to resources when defined as a Policy Enforcement Point (PEP) for a resource server that is receiving and authorizing OAuth 2.0 requests.
- DataPower XI50z V5.0 firmware improves processing power with extended memory support by removing some of the limitations for large files.

Note: For additional details on DataPower XI50z firmware V5.0, see the WebSphere DataPower Integration Appliance Information Center at:
http://pic.dhe.ibm.com/infocenter/wsdatap/v5r0m0/index.jsp?topic=%2Fcom.ibm.dp.xi.doc%2Fwelcome.htm

The IBM POWER7 blades and IBM System x blades enable application integration with System z transaction processing, messaging, and data serving capabilities.

The IBM POWER7 blades, the IBM Smart Analytics Optimizer, the DataPower XI50z, and the IBM System x blades, along with the z196 central processors, can be managed as a single logical virtualized system by the Unified Resource Manager.

zBX configuration

A zBX Model 002 configuration can consist of one to four zBX racks (Rack B, Rack C, Rack D, and Rack E) depending on the number of zBX blades. Each IBM POWER7 blade, IBM Smart Analytics Optimizer, and IBM System x blade require one blade slot. Each DataPower XI50z requires two adjacent blade slots. See Figure 6.

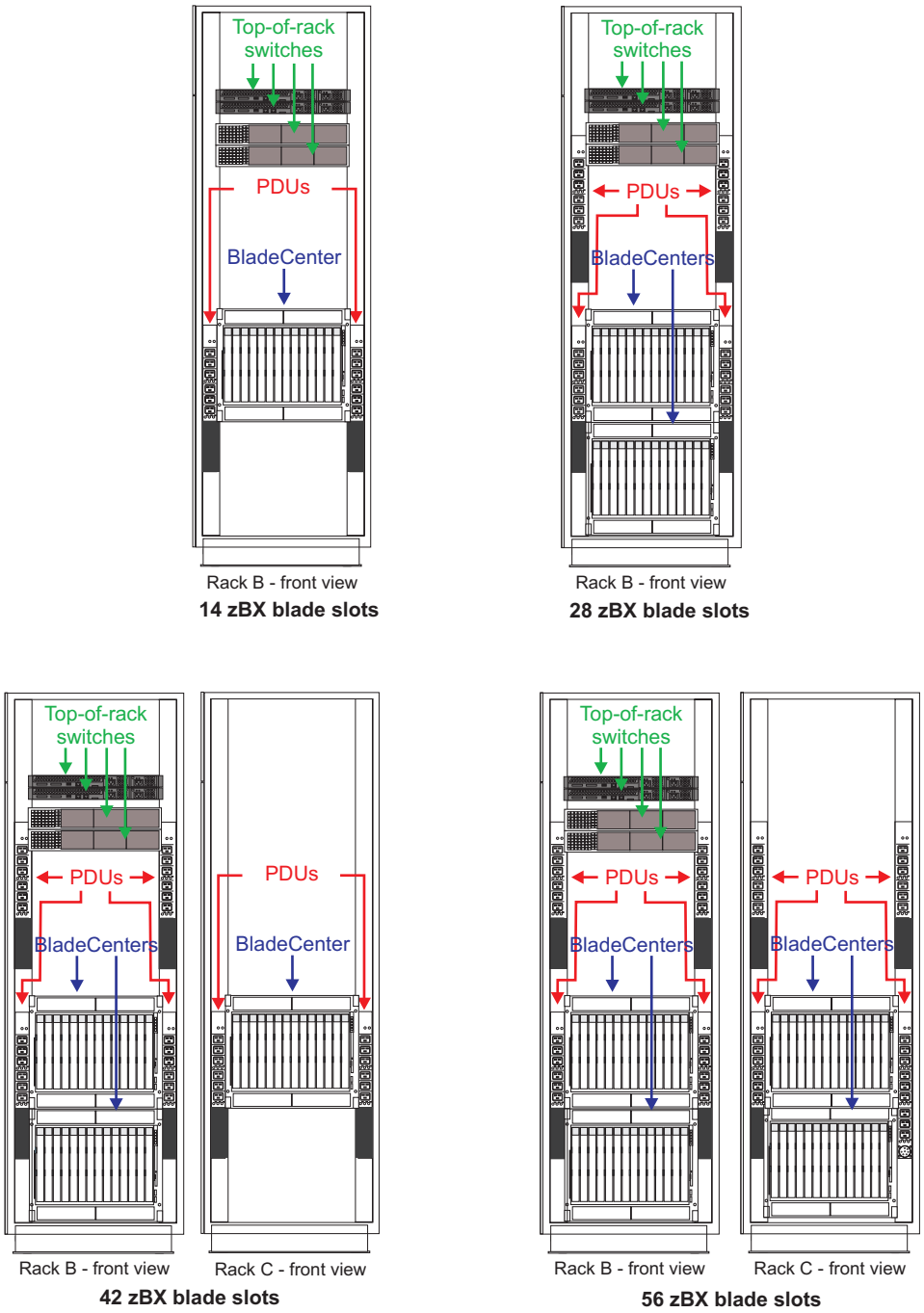


Figure 6. 14, 28, 42, and 56 zBX blade slots (Part 1 of 4)

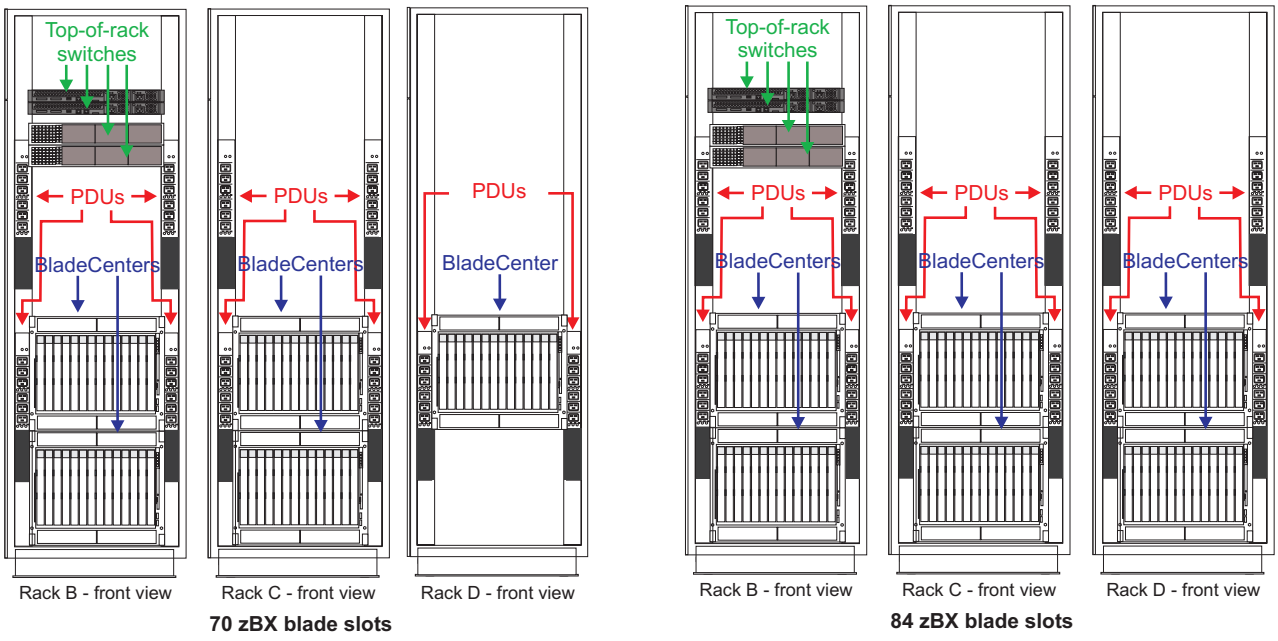


Figure 7. 70 and 84 zBX blade slots (Part 2 of 4)

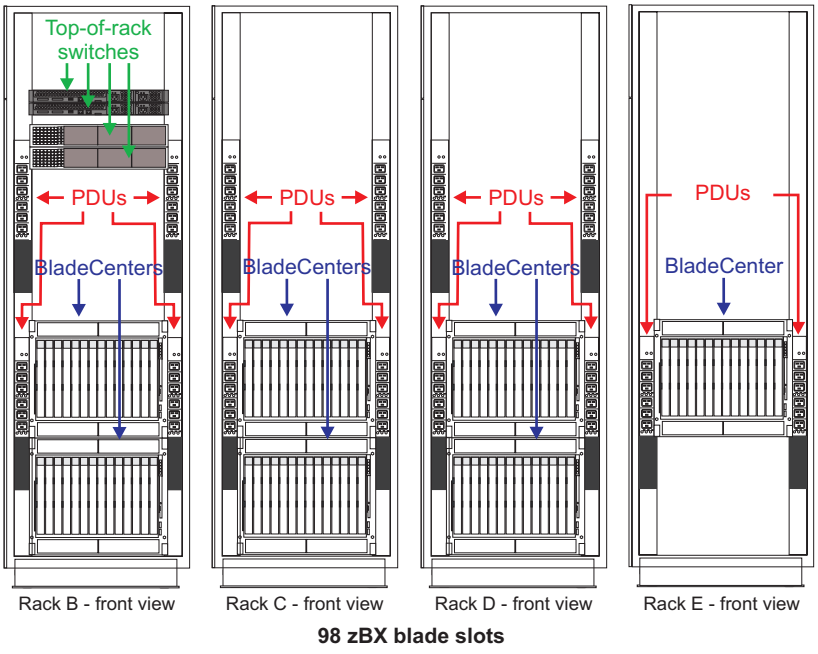


Figure 8. 98 zBX blade slots (Part 3 of 4)

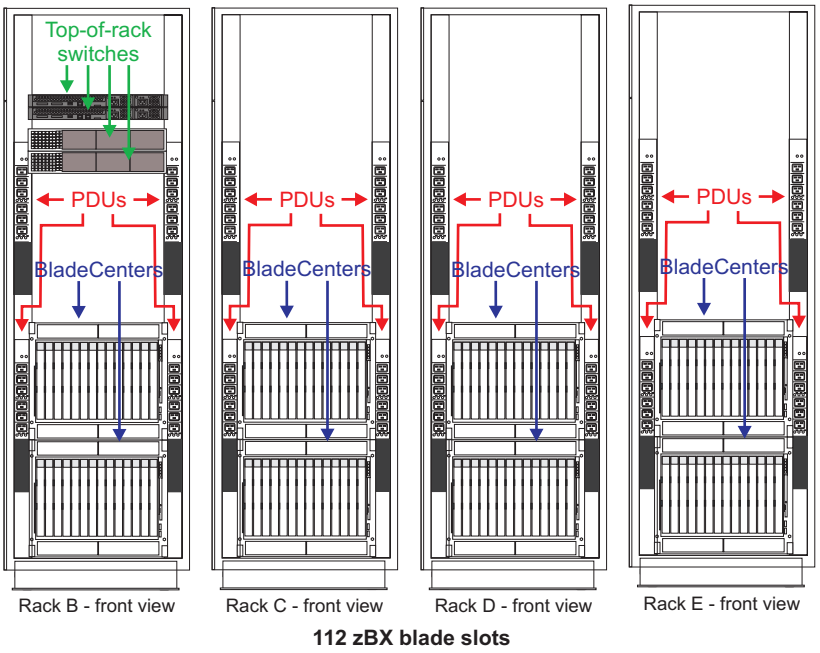


Figure 9. 112 zBX blade slots (Part 4 of 4)

Each zBX rack consists of:

- Top-of-rack (TOR) switches – management TOR switches and intraensemble data network TOR switches (only located in the first rack, Rack B)
- Power distribution units (PDUs) (2 per BladeCenter)
- One or two BladeCenters per rack
- Optional rear door heat exchanger
- Optional acoustic door.

Top-of-rack (TOR) switches

There are two management TOR switches and two intraensemble data network (IEDN) TOR switches in the first rack. If more than one rack is needed, the additional racks do not require management or intraensemble data network TOR switches. These switches are located near the top of the rack and are mounted from the rear of the rack.

The management TOR switches provide a 1000BASE-T Ethernet connection to z196 operating at 1 Gbps. One switch connects the zBX to the bulk power hub (BPH) port on the z196. For redundancy, a switch on the other set of management TOR switches connects to the BPH 1 Gbps port on the other side of the z196.

The management TOR switches also provide connectivity to the management modules and the switch modules located on the BladeCenter. The management modules monitor the BladeCenter. The information gathered is reported to the Support Element on z196 using the connectivity set up over the intranode management network (INMN). These connections are also configured for redundancy.

The intraensemble data network TOR switches provide connectivity for application data communications within an ensemble. Data communications for workloads can flow over the IEDN within and between nodes of an ensemble. This is provided by redundant connections between the OSA-Express3 10 GbE SR, OSA-Express3 10 GbE LR, OSA-Express4S 10 GbE SR, or OSA-Express4S 10 GbE LR cards in the I/O drawer, I/O cage, or PCIe I/O drawer in the z196 and the intraensemble data network TOR switches in the zBX.

The intraensemble data network TOR switches also provide zBX to zBX IEDN communications and external customer network communications.

Power Distribution Unit (PDU)

The power distribution units (PDUs) provide the connection to the main power source, the power connection to the intranode management network and intraensemble data network top-of-rack switches, and the power connection to the BladeCenter. The number of power connections needed is based on the zBX configuration. A rack contains two PDUs if only one BladeCenter is installed. A rack contains four PDUs if two BladeCenters are installed.

BladeCenter

The BladeCenter is a type H chassis. It is configured for redundancy to provide the capability to concurrently repair its components.

An IBM BladeCenter consists of:

Blade slot

The BladeCenter can contain IBM POWER7 blades, IBM System x blades, the IBM Smart Analytics Optimizer, and the DataPower XI50z. Within a BladeCenter, the IBM Smart Analytics Optimizer cannot be mixed with any other type of zBX blade. However, the IBM POWER7 blades, the DataPower XI50z, and the IBM System x blades can reside in the same BladeCenter.

Each IBM POWER7 blade, IBM Smart Analytics Optimizer, and IBM System x blade require one blade slot. Each DataPower XI50z requires two adjacent blade slots.

Power module and fan pack

The BladeCenter contains up to four hot-swap and redundant power supply modules with load-balancing and failover capabilities. The power supply also contains fan packs.

Switch modules

The switch modules are the interface to the BladeCenter. The BladeCenter provides up to two high-speed switch module bays and four traditional switch module bays.

- 10 GbE switch modules (FC 0605)

  (Bays 7 & 9) These switches are part of the intraensemble data network, which is used for application data communication to the node. The switches are connected to the intraensemble data network top-of-rack switches, which are connected to either two OSA-Express3 10 GbE ports or two OSA-Express4S 10 GbE ports on the z196.

  Up to a combination of eight z114s and z196s can connect to a zBX through an intraensemble data network.
- 1000BASE-T Ethernet switch modules operating at 1 Gbps

  (Bays 1 and 2) These switches are part of the intranode management network. They assist in providing a path for the Support Element to load code on a zBX blade.
- 8 GbE Fibre Channel switch modules (FC 0606)

  (Bays 3 & 4) These switches provide each zBX blade with the ability to connect to Fibre Channel (FC) Disk Storage.

Management modules

The management modules monitor the BladeCenter. The information gathered by the management modules is reported to the Support Element using the connectivity set up over the intranode management network.

Blowers

Two hot-swap and redundant blowers standard. There are additional fan packs on power supplies.

Rear door heat exchanger

The heat exchanger rear door (FC 0540) is an optional feature on the zBX. The heat exchanger is intended to reduce the heat load emitted from the zBX. The rear door is an air to water heat exchanger.

Acoustic door

The acoustic door (FC 0543) is an optional feature on the zBX. The acoustic door is intended to reduce the noise emitted from the zBX.

Storage

Storage is provided by the customer outside of the zBX racks via the Fibre Channel (FC) Disk Storage. The required number of SFPs per switch module depends on the number of BladeCenters. There is a connection from the FC switch modules (Bays 3 and 4) in the BladeCenter to the ports in the FC Disk Storage.

Display networking resources associated with the IEDN

You can use the **Network Monitors Dashboard** task to monitor network metrics and to display statistics for the networking resources associated with the IEDN. You can also view performance of the IEDN resources to validate the flow of traffic.

Time coordination for zBX components

z114 and z196 provide the capability for the components in zBX to maintain an approximate time accuracy of 100 milliseconds to an NTP server if they synchronize to the Support Element's NTP server at least once an hour.

Entitlement and management of zBX racks, BladeCenters, and zBX blades

For z196, your IBM representative must identify the number of select IBM POWER7 blades and select IBM System x blades you might use and the number of zBX blades you might need for the IBM Smart Analytics Optimizer and DataPower XI50z.

The maximum number of select IBM POWER7 blades you can have is 112. The maximum number of select IBM System x blades you can have is 56. The maximum number of DataPower XI50z blades you can have is 28. The maximum number of IBM Smart Analytics Optimizers you can have is 56, and they must be ordered in the following increments: 7, 14, 28, 42, 56. The total zBX capacity cannot exceed 112 zBX blades.

Management of the zBX blades is provided by the HMC and Support Element. You can use the Support Element to add and remove a zBX blade. You can also add entitlement to a zBX blade, remove entitlement from a zBX blade, or transfer entitlement from one zBX blade to another zBX blade.

Note: On November 1, 2011, IBM Smart Analytics Optimizer was withdrawn from marketing for new build and MES zBX. It has been replaced with IBM DB2 Analytics Accelerator for z/OS, a workload optimized, appliance add-on that attaches to the zEnterprise System.

Ensemble

With zEnterprise, you can create an ensemble. An ensemble is a collection of one to eight zEnterprise nodes, and each node is a single z114 or z196 with or without an attached zBX. An ensemble delivers a logically integrated and managed view of the zEnterprise infrastructure resources. A zEnterprise node can be a member of only one ensemble.

The ensemble is managed by the Unified Resource Manager, which is Licensed Internal Code (LIC) that is part of the HMC. The Unified Resource Manager performs tasks that provide a single, cohesive

management context applied across all managed objects of the ensemble. See "Unified Resource Manager" on page 127 for more information on the Unified Resource Manager.

For an ensemble, you must have two HMCs:
- A primary HMC managing the resources of one ensemble (and managing CPCs that are not part of an ensemble)
- An alternate HMC, which will become the primary HMC if the HMC currently managing the ensemble fails.

For more information about ensembles, see the *zEnterprise System Introduction to Ensembles*.

IBM DB2 Analytics Accelerator for z/OS V2.1

IBM DB2 Analytics Accelerator for z/OS V2.1 is a workload optimized, appliance add-on that logically plugs into DB2 for z/OS on a z196 or z114, and uses Netezza technology to perform high speed, complex DB2 queries. This disk-based accelerator speeds the response time for a wide variety of complex queries that scan large tables. Efficient data filtering by early SQL projections and restrictions is performed using a Field Programmable Gate Array (FPGA).

Similar to IBM Smart Analytics Optimizer for DB2 for z/OS V1.1, IBM DB2 Analytics Accelerator for z/OS V2.1:
- Provides access to data in terms of authorization and privileges (security aspects) is controlled by DB2 and z/OS (Security Server)
- Uses DB2 for z/OS for the crucial data management items, such as logging, backup/recover, enforcing security policies, and system of record
- Provides no external communication to the IBM Smart Analytics Optimizer for DB2 for z/OS V1.1 beyond DB2 for z/OS
- Is transparent to applications.

Enhancements provided with IBM DB2 Analytics Accelerator for z/OS V2.1 include:
- Extended acceleration to significantly larger number of queries
- Expansion of the size of the data to be accelerated
- Improved concurrent query execution
- Incremental update by partition
- DB2 for z/OS V9 and DB2 for z/OS V10 support.

Communication to a z196 or z114 is provided through an OSA-Express3 10 GbE SR, OSA-Express3 10 GbE LR, OSA-Express4S 10 GbE SR, or OSA-Express4S 10 GbE LR connection.

IBM DB2 Analytics Accelerator for z/OS V2.1 is not integrated into a zBX and is not managed by Unified Resource Manager. It does not require or exploit zEnterprise ensemble capabilities.

Additional features/functions supported

In addition to the standard and optional features previously listed, the design of the z196 also provides the following functions:

Monitoring and estimating power consumption and temperature

You can use the Hardware Management Console (HMC) or the Active Energy Manager to monitor the power consumption and the internal temperature of a CPC.

Using the HMC

You can use the **Activity** task and the **Monitors Dashboard** task on the HMC to monitor the following:

- Power consumption of a zCPC, CPC, BladeCenters, and blades
- Internal temperature of a zCPC CPC, and BladeCenters
- Processor usage of a zCPC, a CPC, BladeCenters, blades, CPs, ICFs, IFLs, zIIPs, zAAPs, SAPs, virtual servers, and LPARs
- Memory usage of virtual servers and blades
- Shared and non-shared channel usage
- Relative humidity of the air entering the system
- Dewpoint – the air temperature at which water vapor will condense into water
- Amount of heat removed from the system by either chilled water or forced air.

Note: Some of this data is only available under certain conditions.

The **Activity** task displays the information in a line-oriented format. The **Monitors Dashboard** task displays the information in a dashboard format that uses tables and graphs.

Using the **Monitors Dashboard** task, you can export the data displayed in the window to a read-only spreadsheet format (.csv file). For a selected CPC, you can also create histograms that display processor usage, channel usage, power consumption, and input air temperature data over a specified time interval.

Using the Active Energy Manager

In addition to providing the power consumption and temperature of a specific CPC, Active Energy Manager also provides the aggregated temperature and power for a group of systems or a complete data center. Active Energy Manager allows you to display this data in a format that shows trends over a specified time interval.

Before using Active Energy Manager, you must enable the SNMP or Web Services APIs, and, if using SNMP, you must define a community name for Active Energy Manager. This action is specified on the **Customize API Settings** task on the HMC. Once you have configured the SNMP or Web Services support on the HMC, you must set up Active Energy Manager so it can communicate to the HMC. You can perform this setup, within Active Energy Manager, by defining it as an SNMP or Web Services device. Once the setup is complete, the Active Energy Manager can communicate to the HMC.

Active Energy Manager is a plug-in to IBM Director.

For more information, see the IBM Systems Software Information Center website (*http://publib.boulder.ibm.com/infocenter/eserver/v1r2/index.jsp*). Expand **IBM Systems Software Information Center** located in the navigation pane on the left, select **Product listing**, then select **IBM Director extension: Active Energy Manager** from the product listing.

Power estimation tool

You can estimate the power consumption of a specific z196 model and its associated configuration using the Power Estimation tool. The exact power consumption for your machine will vary. The purpose of the tool is to produce an estimation of the power requirements to aid you in planning for your machine installation. This tool is available on Resource Link.

Reducing power consumption

zEnterprise provides the capability for you to reduce the energy consumption of a system component (zBX blade, zBX BladeCenter, zCPC, CPC) or group of components. You can reduce the energy consumption by enabling power saving mode or limiting the peak power consumption.

To enable power saving mode, use any of the following methods:

- The **Set Power Saving** task to manually enable the power saving mode
- The **Customize Scheduled Operations** task to set up a schedule defining when you want to turn on power saving mode
- SNMP, CIM, or Web Services APIs
- Active Energy Manager (AEM)
- The **Customize/Delete Activation Profiles** task to enable power saving mode at activation time.

Note: A zCPC can be placed in power saving mode only once per day. If you try to set a zCPC in power saving mode more than once in any given day, the **Power Saving** field for the zCPC will contain "Not Supported" and will be disabled.

To limit the peak power consumption of zBX blades, use the **Set Power Cap** task to enter the power cap value in watts (W).

Displaying historical power, temperature, and utilization data

You can use the **Environmental Efficiency Statistics** task to display a historical view of power, temperature, and utilization data of your z196. Reviewing this data over a period of time can help you track the performance of your system and make appropriate changes to improve your system's performance. The data displays in both table and graph format.

When using the **Environmental Efficiency Statistics** task, you identify:

- A start date
- The number of days (from one to seven) of information you want to display. This includes the start date.
- The type of data you want displayed:
  - System power consumption (in kW and Btu/hour)
  - System temperature (in Celsius and Fahrenheit)
  - Average utilization of all central processors
  - Average CPU utilization of all blades.

You can also export this data to a read-only spreadsheet format (.csv file).

This data is not saved on your system forever. Therefore, if you want to monitor this data for a period of time, you can use the export function to save the data to a .cvs file.

Preplanning and setting up the Storage Area Network (SAN) environment

The WWPN tool assists you in preplanning and setting up your Storage Area Networks (SANs) environment before the installation of your z196. Therefore, you can be up and running much faster after the server is installed.

The WWPN tool assigns WWPNs to virtual and physical FCP ports, which is required to set up your SAN, and creates a binary configuration that can be imported by your system.

This tool applies to all FICON channels defined as CHPID type FCP (for communication with SCSI devices). The WWPN tool can be downloaded from Resource Link under the **Tools** section.

Chapter 3. Software support

This chapter describes the software support for the z196. This information applies to z196 systems running in LPAR mode. The following table displays a summary of the minimum supported operating systems levels for the z196 models.

Table 10. Supported operating systems for z196

Operating System	ESA/390 (31-bit)	z/Architecture (64-bit)
z/OS Version 1 Release 11 [5, 8], 12 [5, 8], 13 [5, 8]	No	Yes
z/OS Version 1 Release 8 [1], 9 [2], 10 [3] with IBM Lifecycle Extension for z/OS V1.8, V1.9, and V1.10	No	Yes
Linux on System z [4, 7]: Red Hat RHEL 5, 6 and Novell SUSE SLES 10, 11	No	Yes
z/VM Version 5 Release 4 [4, 5, 9] z/VM Version 6 Release 1 [4, 5, 9] z/VM Version 6 Release 2 [4, 5, 9]	No [6]	Yes
z/VSE Version 4 Release 2 and later [5, 10]	No	Yes
z/TPF Version 1 Release 1	No	Yes

Note:

1. z/OS V1.8 support was withdrawn September 30, 2009. However, with the z/OS Lifecycle Extension for z/OS V1.8 (5638-A01), z/OS V1.8 supports z196. Talk to your IBM representative for details. No exploitation of z196 functions is available with z/OS V1.8. Certain functions and features of the z196 require later releases of z/OS. For a complete list of software support, see the 2817DEVICE Preventive Planning (PSP) bucket. For more information on the IBM Lifecycle Extension for z/OS V1.8, see Software Announcement 209-180 (RFA 53080) dated June 9, 2009.
2. z/OS V1.9 support was withdrawn September 30, 2010. After that date, the z/OS Lifecycle Extension for z/OS V1.9 (5646-A01) is required for z196. Talk to your IBM representative for details. No exploitation of z196 functions is available with z/OS V1.9. Certain functions and features of the z196 require later releases of z/OS. For a complete list of software support, see the 2817DEVICE Preventive Planning (PSP) bucket. For more information on the IBM Lifecycle Extension for z/OS V1.9, see Software Announcement 210-027 dated May 11, 2010.
3. z/OS V1.10 supports z196; however, z/OS V1.10 support will be withdrawn September 30, 2011. After that date, the z/OS Lifecycle Extension for z/OS V1.10 (5656-A01) is required for z196. Talk to your IBM representative for details. Certain functions and features of the z196 require later releases of z/OS. For a complete list of software support, see the 2817DEVICE Preventive Planning (PSP) bucket. For more information on the IBM Lifecycle Extension for z/OS V1.10, see Software Announcement 211-002, dated February 15, 2011.
4. Compatibility support for listed releases. Compatibility support allows OS to IPL and operate on z196.
5. With PTFs.
6. z/VM supports 31-bit and 64-bit guests.
7. RHEL is an abbreviation for Red Hat Enterprise Linux. SLES is an abbreviation for SUSE Linux Enterprise Server.
8. Refer to the z/OS subset of the 2817DEVICE Preventive Service Planning (PSP) bucket prior to installing zEnterprise.
9. Refer to the z/VM subset of the 2817DEVICE Preventive Service Planning (PSP) bucket prior to installing zEnterprise or IPLing a z/VM image.
10. Refer to the z/VSE subset of the 2817DEVICE Preventive Service Planning (PSP) bucket prior to installing zEnterprise.

Any program written for z/Architecture or ESA/390 architecture mode can operate on CPCs operating in the architecture mode for which the program was written, provided that the program:

- Is not time-dependent.

- Does not depend on the presence of system facilities (such as storage capacity, I/O equipment, or optional features) when the facilities are not included in the configuration.
- Does not depend on the absence of system facilities when the facilities are included in the configuration.
- Does not depend on results or functions that are defined as unpredictable or model dependent in the *z/Architecture Principles of Operation* or in the *Enterprise System Architecture/390 Principles of Operation*.
- Does not depend on results or functions that are defined in this publication (or, for logically partitioned operation, in the *zEnterprise System Processor Resource/Systems Manager Planning Guide*) as being differences or deviations from the appropriate *Principles of Operation* publication.
- Does not depend on the contents of instruction parameter fields B and C on interception of the SIE instruction.

Any problem-state program written for ESA/390 architecture mode can operate in z/Architecture mode provided that the program complies with the limitations for operating in ESA/390 mode and is not dependent on privileged facilities which are unavailable on the system.

Chapter 4. Channel subsystem structure

A channel subsystem (CSS) structure for z196 is designed for 256 channels. With the scalability benefits provided by z196, it is essential that the channel subsystem (CSS) structure is also scalable and permits "horizontal" growth. This is facilitated by allowing more than one logical channel subsystem (LCSS) on a single z196.

Table 11. Channel, port, adapter maximums

Type	z196 Maximum
ESCON	16 cards / 240 channels
FICON Express8S	160 features / 320 channels
FICON Express8	72 features / 288 channels
FICON Express4	72 features / 288 channels
OSA-Express4S GbE [1]	48 features / 96 ports [7]
OSA-Express4S 10 GbE [1]	48 features / 48 ports [7]
OSA-Express3 GbE [1]	24 features / 96 ports
OSA-Express3 10 GbE [1]	24 features / 48 ports
OSA-Express3 1000BASE-T [1]	24 features / 96 ports
OSA-Express2 GbE [1]	24 features / 48 ports
OSA-Express2 1000BASE-T [1]	24 features / 96 ports
IC link	32 links
ISC-3 [2]	12 mother cards / 48 links
12x IFB (HCA2-O) [2]	8 features / 16 links [3] 16 features / 32 links [4]
12x IFB (HCA3-O) [2]	8 features / 16 links [3] 16 features / 32 links [4]
1x IFB (HCA2-O LR) [2]	8 features / 16 links [3] 16 features / 32 links [4]
1x IFB (HCA3-O LR) [2]	8 features / 32 links [3] 12 features / 48 links [4]
Crypto Express3 [5, 6]	8 cards / 16 PCIe adapters

Note:

1. Maximum number of PCHIDs for combined OSA-Express4S, OSA-Express3, and OSA-Express2 features is 48.
2. Maximum number of coupling CHPIDs (ISC-3 and IFB) is 64. Each coupling feature cannot exceed its individual maximum limit (shown in the table).
3. Applies to Model M15.
4. Applies to Models M32, M49, M66, and M80.
5. The maximum number of combined Crypto Express3 features is eight.
6. The initial order for Crypto Express3 is two features.
7. For every OSA-Express3 feature in the configuration, the OSA-Express4S maximum number of features is reduced by two.

The CSS structure offers the following:

- Four logical channel subsystems (LCSSs)
  - Each LCSS can have up to 256 channels defined

  - Each LCSS can be configured with one to 15 logical partitions (cannot exceed 60 LPARs per system).
- Spanned channels are shared among logical partitions across LCSSs. For more information on spanned channels, refer to Table 12 and to "Spanned channels" on page 56.
- Channel paths, control units, and devices that can be dynamically added, changed, and deleted in multiple LCSSs.

Note: One operating system image supports up to a maximum of 256 Channel Path Identifiers (CHPIDs).

The I/O Subsystem (IOSS) continues to be viewed as a single Input/Output Configuration Data Set (IOCDS) across the entire system with up to four LCSSs. Only one Hardware System Area (HSA) is used for the multiple LCSSs.

A CHPID is a two-digit hexadecimal number that identifies a channel path in the CPC. A Physical Channel Identifier (PCHID) is a three-digit number that identifies the physical location (cage, slot, card port) for a channel path in the CPC. An adapter ID (AID) is a two-digit hexadecimal number that identifies HCA3-O, HCA3-O LR, HCA2-O or HCA2-O LR fanout cards. CHPIDs are associated with ports on an adapter and the AID is used in that definition.

The CHPID Mapping Tool can help you map your PCHIDs to the CHPID definitions in your IOCP source statements. The tool will provide you with a new report with your CHPID assignment in addition to the PCHID values. The CHPID Mapping Tool is available from Resource Link, *http://www.ibm.com/servers/resourcelink*, as a standalone PC-based program. For more information on the CHPID Mapping Tool, CHPIDs, PCHIDs or AIDs, refer to *System z CHPID Mapping Tool User's Guide*.

IOCP channel, link, and adapter definitions

The following table lists the channel and link types as defined in an IOCDS that are used with z196 systems.

Table 12. Channels, links, and adapters with CHPID type

Channels/Links/Adapters	CHPID type	May be defined as Shared	May be defined as Spanned
ESCON channels:			
Connection Channel (ESCON architecture) Channel-to-Channel (connects to CNC)	CNC CTC	yes yes	no no
ESCON channels connected to converter:			
Conversion Channel (ESCON to Parallel Block Multiplexer (BL)) Conversion Channel (ESCON to Parallel Byte Multiplexer (BY))	CVC CBY	no no	no no
FICON channels — native FICON, zHPF, or CTC for attachment to FICON channels on System z servers, directors, control units, and printers	FC	yes	yes
Fibre Channel Protocol (FCP) for communicating with SCSI devices	FCP	yes	yes
ISC-3 peer mode links (connects to another ISC-3)	CFP	yes	yes
IC peer links (connects to another IC)	ICP	yes	yes
IFB peer links (connects to another IFB)	CIB	yes	yes
HiperSockets	IQD	yes	yes
OSA adapters using QDIO architecture: TCP/IP traffic when Layer 3, Protocol-independent when Layer 2	OSD	yes	yes
OSA adapters using non-QDIO architecture for TCP/IP and/or SNA/APPN/HPR traffic	OSE	yes	yes

Table 12. Channels, links, and adapters with CHPID type (continued)

Channels/Links/Adapters	CHPID type	May be defined as Shared	May be defined as Spanned
OSA-ICC: OSA 1000BASE-T Ethernet adapters for TN3270E, non-SNA DFT, IPL CPCs, and LPARs, OS system console operations	OSC	yes	yes
OSA-Express for NCP: NCPs running under IBM Communication Controller for Linux (CDLC)	OSN	yes	yes
OSA-Express3 10 GbE LR , OSA-Express3 10 GbE SR, OSA-Express4S 10 GbE LR, OSA-Express4S 10 GbE SR adapters for intraensemble data network (IEDN)	OSX	yes	yes
OSA-Express3 1000BASE-T Ethernet adapters for intranode management network (INMN)	OSM	yes	yes

Each of these channel types requires that a CHPID be defined, even if it is an internal channel and no physical hardware (channel card) exists. Each channel, whether a "real" channel or a virtual (such as HiperSockets) must be assigned a unique CHPID within the LCSS. You can arbitrarily assign a number within the X'00' to X'FF' range. Real channels require a PCHID value to be defined. Most of these channel types can be shared and used concurrently among multiple LPARs within the same LCSS. Refer to "Multiple Image Facility (MIF)" on page 55 for more information on shared channels.

AIDs are used for InfiniBand connections.

Coupling link peer channels

You may define an ISC-3 feature as CFP and an IFB link as CIB. Any available/unused CHPID may be defined as ICP.

You can configure a CFP, ICP, or CIB channel path as:

- An unshared dedicated channel path to a single logical partition.
- An unshared reconfigurable channel path that can be configured to only one logical partition at a time but which can be dynamically moved to another logical partition by channel path reconfiguration commands. Reconfigurable support for CFP, CIB, and ICP is limited to two coupling facility logical partitions total. One coupling facility logical partition in the initial access list and one other coupling facility partition in the candidate list.
- A shared or spanned channel path that can be concurrently used by the logical partitions to which it is configured. A peer channel cannot be configured to more than one coupling facility logical partition at a time, although it can be configured to multiple z/Architecture or ESA/390 logical partitions in addition to the single coupling facility logical partition.
- Timing-only links. These are coupling links that allow two servers to be synchronized using Server Time Protocol (STP) messages when a coupling facility does not exist at either end of the coupling link.

  Note: CHPID type ICP is not supported for a timing connection.

Each ICP channel path must specify which ICP channel path it is logically connected to.

The z196 models support dynamic I/O configuration for all peer channel path types.

Subchannel connectivity

With four Logical Channel Subsystems comes more subchannels. There is a maximum of 65280 subchannels per LCSS for subchannel set 0, 65535 subchannels per LCSS for subchannel set 1, and 65535 subchannels per LCSS for subchannel set 2.

Subchannel set 0 allows definitions of any type of device (bases, aliases, secondaries, etc.). Subchannel set 1 and subchannel set 2 are designated for disk alias devices (of both primary and secondary devices) and metro mirror secondary devices.

z196 allows you to IPL from subchannel set 1 and subchannel set 2 in supported operating systems such as z/OS.

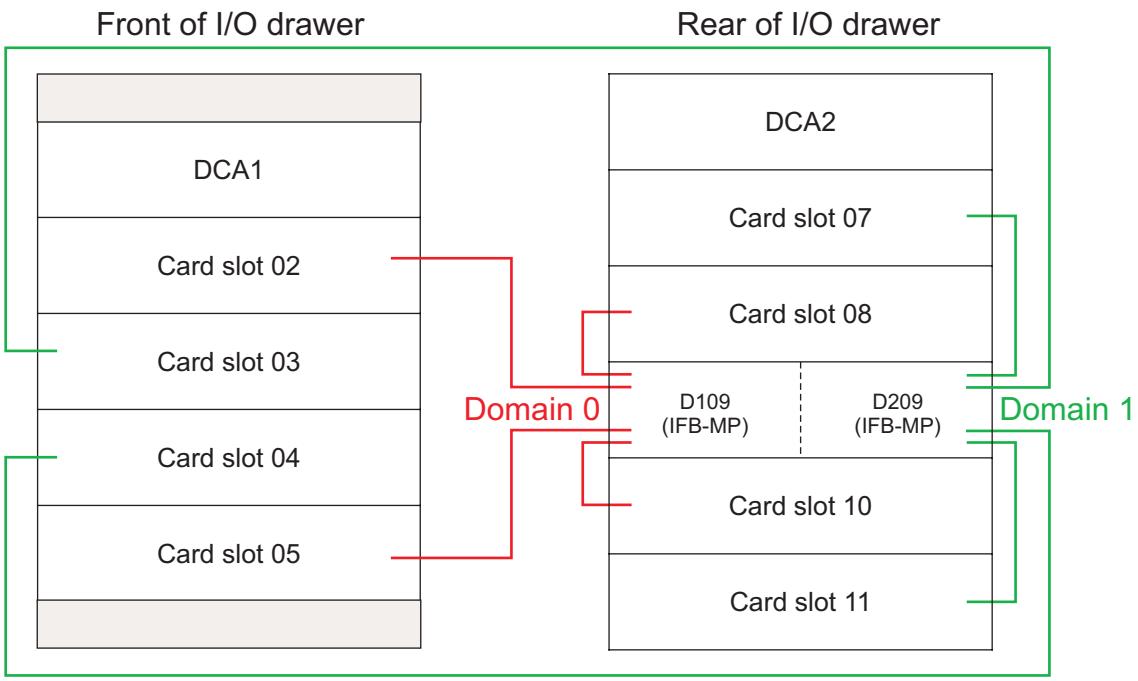
With four Logical Channel Subsystems you can have:
- Up to a maximum of 65280 devices/subchannels per LCSS for subchannel set 0
- Up to a maximum of 65535 devices/subchannels per LCSS for subchannel set 1
- Up to a maximum of 65535 devices/subchannels per LCSS for subchannel set 2
- Up to a maximum of 785400 devices for four LCSSs (four LCSSs times 65280 + 65535 + 65535 subchannels for each LCSS.)

Each LPAR can access all the devices in its assigned LCSS.

This capability relieves the I/O device configuration constraints experienced by large system configurations.

Guidelines for maximum availability

When configuring devices with multiple paths to the same CPC, select any of the channel paths from any I/O card shown in Figure 10 on page 47, Figure 11 on page 48, and Figure 12 on page 49 that:
- Are available on the CPC you are defining
- Are the correct type (FICON, ESCON, etc.) to meet the control unit, coupling facility, or network attachment requirements
- Satisfy the rules regarding the mixing of channel types to a control unit.

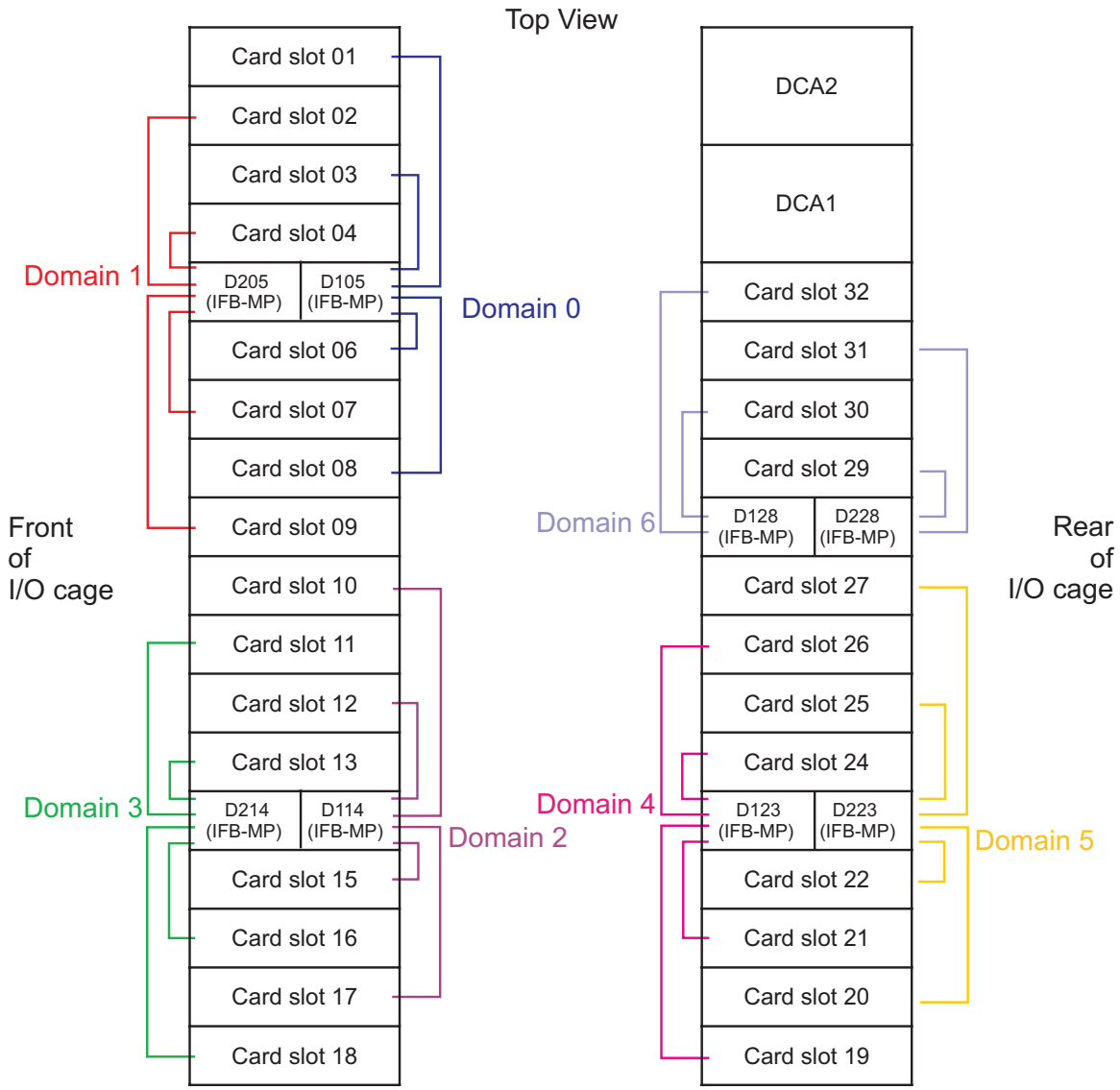


Notes:

D109 IFB-MP card location controls Domain 0 (card slots 02, 05, 08, and 10).

D209 IFB-MP card location controls Domain 1 (card slots 03, 04, 07, and 11).

Figure 10. I/O drawer layout



Notes:

D105 IFB-MP card location controls Domain 0. Domain 0 consists of card slots 01, 03, 06, and 08.

D205 IFB-MP card location controls Domain 1. Domain 1 consists of card slots 02, 04, 07, and 09.

D114 IFB-MP card location controls Domain 2. Domain 2 consists of card slots 10, 12, 15, and 17.

D214 IFB-MP card location controls Domain 3. Domain 3 consists of card slots 11, 13, 16, and 18.

D123 IFB-MP card location controls Domain 4. Domain 4 consists of card slots 19, 21, 24, and 26.

D223 IFB-MP card location controls Domain 5. Domain 5 consists of card slots 20, 22, 25, and 27.

D128 & D228 IFB-MP card locations control Domain 6. Domain 6 consists of card slots 29, 30, 31, and 32.

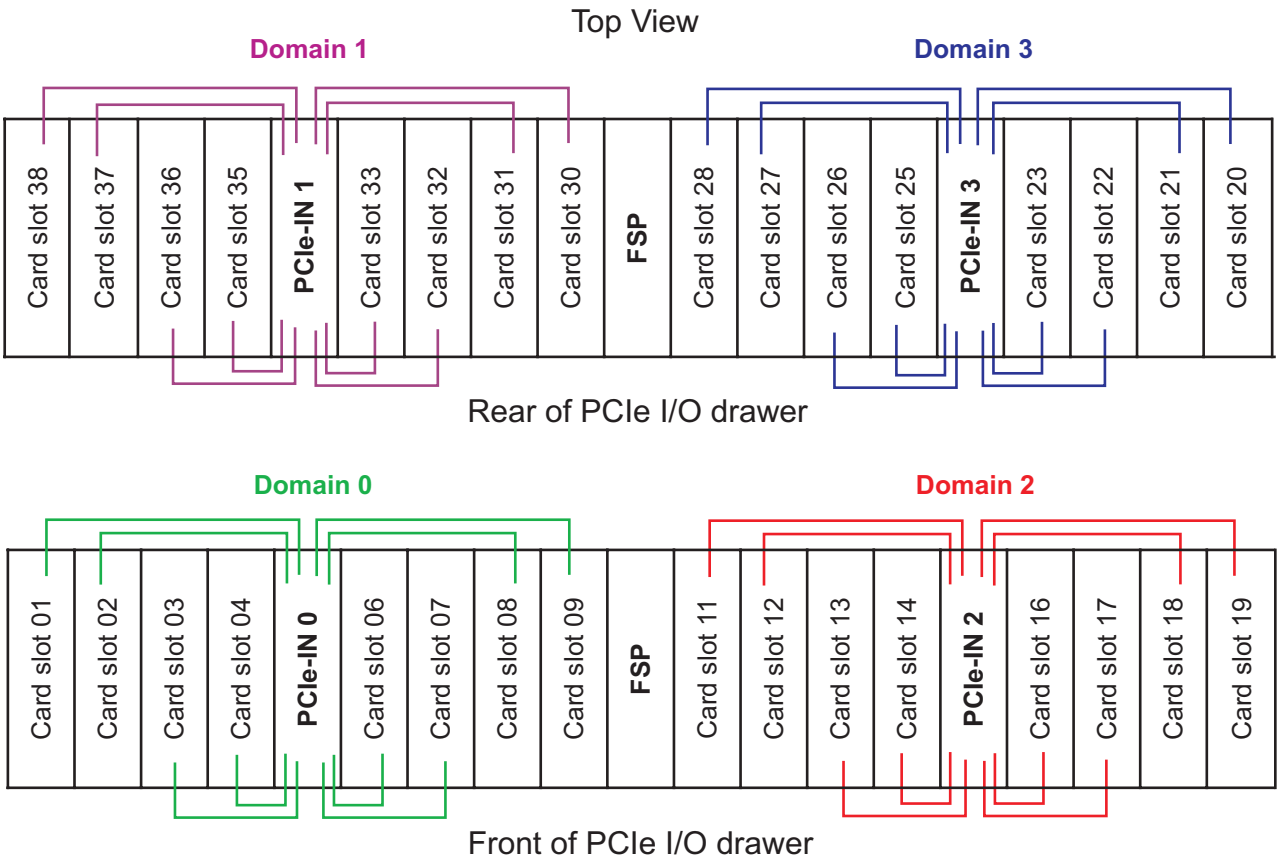
Figure 11. I/O cage layout

Legend:

IFB-MP
InfiniBand Multiplexer

D1 In the I/O drawer, D1 represents the half-high daughter card located in the left side of the slot. In the I/O cage, D1 represents the half-high daughter card located in the top half of the slot.

D2 In the I/O drawer, D2 represents the half-high daughter card located in the right side of the slot. In the I/O cage, D2 represents the half-high daughter card located in the bottom half of the slot.



Notes:

PCIe-IN card 0 controls Domain 0 (card slots 01, 02, 03, 04, 06, 07, 08, and 09).

PCIe-IN card 1 controls Domain 1 (card slots 30, 31, 32, 33, 35, 36, 37, and 38).

PCIe-IN card 2 controls Domain 2 (card slots 11, 12, 13, 14, 16, 17, 18, and 19).

PCIe-IN card 3 controls Domain 3 (card slots 20, 21, 22, 23, 25, 26, 27, and 28).

Figure 12. PCIe I/O drawer layout

Legend:

PCI-IN
PCIe Interconnect

For maximum availability of the device, OSA network, or coupling facility on z196, you should consider the following guidelines:

- Choose channels plugged in different I/O domains.

  With an I/O drawer or an I/O cage, an I/O domain contains four channel cards controlled by a single IFB-MP card. The IFB-MP card provides connection to the CPC. For example, for an I/O drawer, the domain for the IFB-MP card in D109 controls slots 02, 05, 08, and 10. (See Figure 10 on page 47.) For an I/O cage, the domain for the IFB-MP card in D105 controls slots 01, 03, 06, and 08. (See Figure 11 on page 48.) With a PCIe I/O drawer, an I/O domain contains eight channel cards controlled by a single PCI-IN card. For example, the domain for PCI-IN card 0 controls slots 01, 02, 03, 04, 06, 07, 08, and 09. (See Figure 12 on page 49.)

  Note: This is also recommended for optimum performance of your most heavily-used I/O devices.

  When choosing the I/O domains to use, whether from different cages or drawers or the same cage or drawer, consider using a combination of I/O domains from different books. When you must use IFB links from the same book, try to use IFB links from different HCA fanout cards on that book. Refer to your PCHID report to determine which IFB links belong to which HCA fanout cards on a book. If you have multiple paths to the device and multiple domains available that have the correct channel type, spreading the paths across as many books and HCAs as possible is also advisable.

  Redundant I/O interconnect is a function that allows one IFB-MP to back up another IFB-MP (for an I/O drawer or I/O cage) or one PCI-IN to back up another PCI-IN (for a PCIe I/O drawer) in case of a failure or repair. For example, in the I/O cage, the IFB-MP cards in slot 05 (same is true for those in slots 14, 23, and 28) back up each other. In the I/O drawer, the IFB-MP cards in slot 09 back up each other. In the PCIe I/O drawer, the PCI-IN card 0 and PCI-IN card 1 back up each other. Therefore, in the event an InfiniBand cable, fanout card, or book fails, the remaining IFB-MP card (or PCI-IN card) will control both domains. There are failures (for example, the IFB-MP card or PCI-IN card) that may prevent the redundant takeover, which is why it is advisable to spread your paths over multiple domains.

  When configuring Coupling using InfiniBand (CIB) links for the same target CPC or coupling facility, use InfiniBand links that originate from different books and different HCA fanout cards on those books. This eliminates the HCA fanout card and the book as a single point of failure where all connectivity would be lost.
- If you define multiple paths from the same IFB link, distribute paths across different channel cards. Also, if you define multiple coupling facility channels to the same coupling facility or to the same ESA image, distribute paths across different coupling facility channel adapter cards or different coupling facility daughter cards.

With z196, each SAP handles FICON work on an on-demand basis. That is, as FICON work for any channel arrives, the next available SAP will handle that request. It does not matter if it is an outbound request or an inbound interrupt, the next available SAP will handle the FICON work.

For the other channel types, the z196 automatically balances installed channel cards across all available SAPs. The processor attempts to assign an equal number of each channel card type to each available SAP. While all channels on a given I/O card are always in the same SAP, it is not predictable which I/O cards will be assigned to which SAPs. However, there are two exceptions. First, HCAs used for coupling are always given affinity to a SAP on the local drawer. Second, if an OSA channel is defined as OSD, OSM, or OSX or a FICON channel is defined as FCP, these channels use QDIO architecture and, therefore, do not actually use any SAP resource during normal operations.

For all channel types, simply follow the preceding recommendations for configuring for RAS, and the SAPs will handle the workload appropriately.

Planning for channel subsystem

This section contains information to aid in the planning for maximum channel subsystem availability on z196. It addresses ESCON, FICON, and OSA channels; ISC-3, IFB, and IC links; and HiperSockets. The information shows the major components of the channel subsystem and suggests ways to configure the CSS for maximum availability.

The overall process of assigning CHPIDs, PCHIDs, and AIDs begins when you order z196 or an MES to an existing machine. After placing the order, the configurator prepares a report (PCHID report) detailing the physical location for each channel in the machine. This report shows PCHID and AID assignments.

PCHID assignments

There are no default CHPIDs assigned. You are responsible for assigning the logical CHPID number to a physical location, identified by a PCHID number. You can complete this task in using either IOCP or HCD. The CHPID Mapping Tool may be used to help with these assignments. (Refer to "CHPID Mapping Tool" on page 54 for more information.)

You will use the data in the CFReport, which you can either obtain from your representative or retrieve from Resource Link, and the IOCP input for assigning PCHID values using the CHPID Mapping Tool.

Table 13 lists the PCHID assignments for slots in the I/O cages. Table 14 on page 52 lists the PCHID assignments for slots in the I/O drawers. Table 15 on page 52 lists the PCHID assignments for slots in the PCIe I/O drawers. Only the active ports on an installed card are actually assigned a PCHID. The remainder are unused.

Except for ESCON sparing, the cards in the I/O cage, I/O drawer, and PCIe I/O drawer are assigned a PCHID starting with the first value in the range for the slot and drawer where the card is located.

For ISC-3 cards, the first daughter is assigned the first two PCHID values of the slot. The second daughter is assigned the slot value plus 8 for the first port and plus 9 for the second port.

OSA-Express4S GbE LX and OSA-Express4S GbE SX cards have two ports, but only one PCHID is assigned.

Crypto cards are assigned one PCHID even though they have no ports.

Table 13. PCHID assignments for I/O cage

	PCHID Range		
Cargo Slot	**Cage 1/Bottom A**	**Cage 2/Bottom Z**	**Cage 3/Top Z**
1	100 - 10F	300 - 30F	500 - 50F
2	110 - 11F	310 - 31F	510 - 51F
3	120 -12F	320 - 32F	520 - 52F
4	130 - 13F	330 - 33F	530 - 53F
6	140 - 14F	340 - 34F	540 - 54F
7	150 - 15F	350 - 35F	550 - 55F
8	160 - 16F	360 - 36F	560 - 56F
9	170 - 17F	370 - 37F	570 - 57F
10	180 - 18F	380 - 38F	580 - 58F
11	190 - 19F	390 - 39F	590 - 59F
12	1A0 - 1AF	3A0 - 3AF	5A0 - 5AF
13	1B0 - 1BF	3B0 - 3BF	5B0 - 5BF

Table 13. PCHID assignments for I/O cage (continued)

Cargo Slot	PCHID Range		
	Cage 1/Bottom A	**Cage 2/Bottom Z**	**Cage 3/Top Z**
15	1C0 - 1CF	3C0 - 3CF	5C0 - 5CF
16	1D0 - 1DF	3D0 - 3DF	5D0 - 5DF
17	1E0 - 1EF	3E0 - 3EF	5E0 - 5EF
18	1F0 - 1FF	3F0 - 3FF	5F0 - 5FF
19	200 - 20F	400 - 40F	600 - 60F
20	210 - 21F	410 - 41F	610 - 61F
21	220 - 22F	420 - 42F	620 - 62F
22	230 - 23F	430 - 43F	630 - 63F
24	240 - 24F	440 - 44F	640 - 64F
25	250 - 25F	450 - 45F	650 - 65F
26	260 - 26F	460 - 46F	660 - 66F
27	270 - 27F	470 - 47F	670 - 67F
29	280 - 28F	480 - 48F	680 - 68F
30	290 - 29F	490 - 49F	690 - 69F
31	2A0 - 2AF	4A0 - 4AF	6A0 - 6AF
32	2B0 - 2BF	4B0 - 4BF	6B0 - 6BF

Table 14. PCHID assignments for I/O drawers

I/O drawer slot	PCHID Range					
	Drawer 1 Z22B	**Drawer 2 Z15B**	**Drawer 3 Z08B**	**Drawer 4 Z01B**	**Drawer 5 A08B (MRU) A15B (MWU)**	**Drawer 6 A01B (MRU) A08B (MWU)**
2	580 - 58F	500 - 50F	380 - 38F	300 - 30F	180 - 18F	100 - 30F
3	590 - 59F	510 - 51F	390 - 39F	310 - 31F	190 - 19F	110 - 31F
4	5A0 - 5AF	520 - 52F	3A0 - 3AF	320 - 32F	1A0 - 1AF	120 - 32F
5	5B0 - 5BF	530 - 53F	3B0 - 3BF	330 - 33F	1B0 - 1BF	130 - 33F
7	5C0 - 5CF	540 - 54F	3C0 - 3CF	340 - 34F	1C0 - 1CF	140 - 34F
8	5D0 - 5DF	550 - 55F	3D0 - 3DF	350 - 35F	1D0 - 1DF	150 - 35F
10	5E0 - 5EF	560 - 56F	3E0 - 3EF	360 - 36F	1E0 - 1EF	160 - 36F
11	5F0 - 5FF	570 - 507F	3F0 - 3FF	370 - 37F	1F0 - 1FF	170 - 37F

Table 15. PCHID assignments for PCIe I/O drawers

PCIe I/O drawer slot	PCHID Range					
	Drawer 1 Z22B	**Drawer 2 Z15B**	**Drawer 3 Z08B**	**Drawer 4 Z01B**	**Drawer 5 A08B (MRU) A15B (MWU)**	**Drawer 6 A01B (MRU) A08B (MWU)**
1	580 - 583	500 - 503	380 - 383	300 - 303	180 - 183	100 - 103
2	584 - 587	504 - 507	384 - 387	304 - 307	184 - 187	104 - 107
3	588 - 58B	508 - 50B	388 - 38B	308 - 30B	188 - 18B	108 - 10B
4	58C - 58F	50C - 50F	38C - 38F	30C - 30F	18C - 18F	10C - 10F

Table 15. PCHID assignments for PCIe I/O drawers (continued)

PCIe I/O drawer slot	PCHID Range					
	Drawer 1 Z22B	Drawer 2 Z15B	Drawer 3 Z08B	Drawer 4 Z01B	Drawer 5 A08B (MRU) A15B (MWU)	Drawer 6 A01B (MRU) A08B (MWU)
6	590 - 593	510 - 513	390 - 393	310 - 313	190 - 193	110 - 113
7	594 - 597	514 - 517	394 - 397	314 - 317	194 - 197	114 - 117
8	598 - 59B	518 - 51B	398 - 39B	318 - 31B	198 - 19B	118 - 11B
9	59C - 59F	51C - 51F	39C - 39F	31C - 31F	19C - 19F	11C - 11F
11	5A0 - 5A3	520 - 523	3A0 - 3A3	320 - 323	1A0 - 1A3	120 - 123
12	5A4 - 5A7	524 - 527	3A4 - 3A7	324 - 327	1A4 - 1A7	124 - 127
13	5A8 - 5AB	528 - 52B	3A8 - 3AB	328 - 32B	1A8 - 1AB	128 - 12B
14	5AC - 5AF	52C - 52F	3AC - 3AF	32C - 32F	1AC - 1AF	12C - 12F
16	5B0 - 5B3	530 - 533	3B0 - 3B3	330 - 333	1B0 - 1B3	130 - 133
17	5B4 - 5B7	534 - 537	3B4 - 3B7	334 - 337	1B4 - 1B7	134 - 137
18	5B8 - 5BB	538 - 53B	3B8 - 3BB	338 - 33B	1B8 - 1BB	138 - 13B
19	5BC - 5BF	53C - 53F	3BC - 3BF	33C - 33F	1BC - 1BF	13C - 13F
20	5C0 - 5C3	540 - 543	3C0 - 3C3	340 - 343	1C0 - 1C3	140 - 143
21	5C4 - 5C7	544 - 547	3C4 - 3C7	344 - 347	1C4 - 1C7	144 - 147
22	5C8 - 5CB	548 - 54B	3C8 - 3CB	348 - 34B	1C8 - 1CB	148 - 14B
23	5CC - 5CF	54C - 54F	3CC - 3CF	34C - 34F	1CC - 1CF	14C - 134F
25	5D0 - 5D3	550 - 553	3D0 - 3D3	350 - 353	1D0 - 1D3	150 - 153
26	5D4 - 5D7	554 - 557	3D4 - 3D7	354 - 357	1D4 - 1D7	154 - 157
27	5D8 - 5DB	558 - 55B	3D8 - 3DB	358 - 35B	1D8 - 1DB	158 - 15B
28	5DC - 5DF	55C - 55F	3DC - 3DF	35C - 35F	1DC - 1DF	15C - 15F
30	5E0 - 5E3	560 - 563	3E0 - 3E3	360 - 363	1E0 - 1E3	160 - 163
31	5E4 - 5E7	564 - 567	3E4 - 3E7	364 - 367	1E4 - 1E7	164 - 1367
32	5E8 - 5EB	568 - 56B	3E8 - 3EB	368 - 36B	1E8 - 1EB	168 - 16B
33	5EC - 5EF	56C - 56F	3EC - 3EF	36C - 36F	1EC - 1EF	16C - 16F
35	5F0 - 5F3	570 - 573	3F0 - 3F3	370 - 373	1F0 - 1F3	170 - 173
36	5F4 - 5F7	574 - 577	3F4 - 3F7	374 - 377	1F4 - 1F7	174 - 177
37	5F8 - 5FB	578 - 57B	3F8 - 3FB	378 - 37B	1F8 - 1FB	178 - 17B
38	5FC - 5FF	57C - 57F	3FC - 3FF	37C - 37F	1FC - 1FF	17C - 17F

AID assignments

HCA2-O, HCA2-O LR, HCA3-O, and HCA3-O LR fanout cards used for coupling are identified by adapter IDs (AIDs) rather than PCHIDs.

CHPID numbers need to be associated with ports on an adapter, and the AID is used for this purpose. It is your responsibility to assign CHPIDs using either IOCP or HCD. The CHPID assignment is done by associating the CHPID number to an AID and port. You cannot use the CHPID Mapping Tool to assign CHPID values to AIDs.

You cannot change an AID assignment. After an AID is assigned, if an optical fanout card is moved on a z196, the AID value moves with it.

Table 16 shows the initial AID assignments for the ports on the HCA fanout cards plugged into the books.

Table 16. AID assignments for HCA fanout cards

Book	Frame slot number	Fanout card slots	AIDs
First book	06	D1-DA	08-0F
Second book	15	D1-DA	18-1F
Third book	10	D1-DA	10-17
Fourth book	01	D1-DA	00-07
Note: The terms "first book," "second book," "third book," and "fourth book" refers to the plugging sequence of the book. (ie. the first book plugged in, the second book plugged in, etc.)			

Each fanout slot is allocated one AID number. (Remember that slots D3 and D4 do not have fanout cards plugged into them; therefore, they are not assigned AIDs.) For example, for the first book, the allocation is:

Fanout slot	AID
D1	08
D2	09
D5	0A
D6	0B
D7	0C
D8	0D
D9	0E
DA	0F

Note: These AID assignments can only be predicted for a new build machine. For an MES to an existing z196, the VPD contains the AID assigned to each installed HCA2-O, HCA2-O LR, HCA3-O, or HCA3-O LR. The VPD also contains the AID that is assigned to all other slots in existing drawers. If a new HCA fanout is added to the drawer, the AID from the VPD should be used.

PCHID report

The PCHID report from the configurator provides details on the placement of all the I/O features in your order. Your representative will provide you with this report. Using this report and the guidelines listed in "Guidelines for maximum availability" on page 46, you can plan the configuration of your I/O.

Note: If you use the CHPID Mapping Tool to aid you in assigning PCHIDs to CHPIDs, the tool will provide you with a new report with your CHPID assignment in addition to the PCHID values.

Other resources available are the *System z Input/Output Configuration Program User's Guide for ICP IOCP* and the CHPID Mapping Tool. These resources are available on Resource Link.

CHPID Mapping Tool

The CHPID Mapping Tool is a Java-based standalone application available from IBM Resourse Link, and it must be downloaded to your personal computer for use. Once downloaded, you can make CHPID assignments without further internet connections. As part of the CHPID Mapping Tool process, you will need a CFReport (which you can download from Resource Link or obtain from your representative) and an IOCP file.

Note: The CHPID Mapping Tool does not assign AID values.

The intent of the CHPID Mapping Tool is to ease installation of new z196 processors and for making changes to an already installed z196 processor either to make slight changes to the mapping or as part of an MES action to add or remove channel features on the processor.

z196 **does not** have default CHPIDs assigned to ports as part of the initial configuration process. It is your responsibility to perform these assignments by using the HCD/IOCP definitions and optionally the CHPID Mapping Tool. The result of using the tool is an IOCP deck that will map the defined CHPIDs to the corresponding PCHIDs for your processor. However, there is no requirement to use the CHPID Mapping Tool. You can assign PCHIDs to CHPIDs directly in IOCP decks or through HCD, but it is much easier to use the tool to do the channel mapping and the tool can help make PCHID to CHPID assignments for availability.

For more information on the CHPID Mapping tool refer to any of the following:
- *System z CHPID Mapping Tool User's Guide*
- *CHPID Mapping Tool* on Resource Link.

Multiple Image Facility (MIF)

The Multiple Image Facility (MIF) allows channel sharing among multiple LPARs and optionally shares any associated I/O devices configured to these shared channels. MIF also provides a way to limit the logical partitions that can access a reconfigurable channel, spanned channel, or a shared channel to enhance security.

With multiple LCSSs, the CSS provides an independent set of I/O controls for each logical channel subsystem called a CSS image. Each logical partition is configured to a separate CSS image in order to allow the I/O activity associated with each logical partition to be processed independently as if each logical partition had a separate CSS. For example, each CSS image provides a separate channel image and associated channel path controls for each shared channel and separate subchannel images for each shared device that is configured to a shared channel.

With MIF, you can configure channels as follows:
- **ESCON (TYPE=CNC, TYPE=CTC), FICON (TYPE=FC or TYPE=FCP), ISC-3 peer (TYPE=CFP), IC peer (TYPE=ICP), IFB peer (TYPE=CIB), HiperSockets (TYPE=IQD), and OSA (TYPE=OSC, TYPE=OSD, TYPE=OSE, TYPE=OSN, TYPE=OSX, or TYPE=OSM)**.

  You can configure a channel path as:
  - An unshared dedicated channel path to a single LPAR.
  - An unshared reconfigurable channel path that can be configured to only one logical partition at a time it can be moved to another logical partition within the same LCSS.
  - A shared channel path that can be concurrently used by the ESA/390 images or CF logical partitions within the same LCSS to which it is configured.

With MIF and multiple channel subsystems, shared and spanned channel paths can provide extensive control unit and I/O device sharing. MIF allows all, some, or none of the control units attached to channels to be shared by multiple logical partitions and multiple CSSs. Sharing can be limited by the access and candidate list controls at the CHPID level and then can be further limited by controls at the I/O device level.

For example, if a control unit allows attachment to multiple channels (as is possible with a 3990 control unit), then it can be shared by multiple logical partitions using one or more common shared channels or unique unshared channel paths.

Spanned channels

With multiple LCSSs, transparent sharing of internal (ICs and HiperSockets) and external (FICON, ISC-3, IFB, OSA) channels across LCSSs is introduced, extending Multiple Image Facility (MIF). MIF allows sharing of channel resources across LPARs. ICs, HiperSockets, FICON, ISC-3s, IFBs, and OSA features can all be configured as MIF spanning channels.

Spanning channels is the ability for the channels to be configured to multiple LCSSs, and be transparently shared by any or all of the configured LPARs without regard to the Logical Channel Subsystem to which the partition is configured. For information on the channel CHPID types and spanning capabilities, refer to Table 12 on page 44.

You can configure the following as a spanned channel:

- **FICON (TYPE=FC or TYPE=FCP), ISC-3 peer (TYPE=CFP), IC peer (TYPE=ICP), IFB peer (TYPE=CIB), HiperSockets (TYPE=IQD), and OSA (TYPE=OSC, TYPE=OSD, TYPE=OSE, TYPE=OSN, TYPE=OSX, or TYPE=OSM)**

  They can be shared by LPARs in different logical channel subsystems.

Internal coupling and HiperSockets channels

Internal coupling (IC) channels and HiperSockets 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. The CHPID type for IC channels is ICP; the CHPID type for HiperSockets is IQD.

- It is suggested that you define a minimum number of ICP CHPIDs for Internal Coupling. For most customers, IBM suggests defining just one ICP for each coupling facility (CF) LPAR in your configuration. For instance, if your z196 configuration has several ESA LPARs and one CF LP, you would define one pair of connected ICP CHPIDs shared by all the LPARs in your configuration. If your configuration has several ESA LPARs and two CF logical partitions, you still would define one connected pair of ICP CHPIDs, but one ICP should be defined as shared by the ESA images and one of the CF LPARs, while the other ICP is defined as shared by the ESA LPARs and the other CF LPAR. Both of these examples best exploit the peer capabilities of these coupling channels by using the "sending" and "receiving" buffers of both channels. If your ESA images and CF images are in different CSSs and you want to exploit the optimal use of ICP then your ICP CHPIDs must be defined as spanned.
- Each IQD CHPID represents one internal LAN. If you have no requirement to separate LAN traffic between your applications, only one IQD CHPID needs to be defined in the configuration. If the partitions sharing the LAN are in different LCSSs your IQD CHPID must be defined as spanned.

IOCP considerations

ICP IOCP supports z196 and multiple LCSSs. Refer to *System z Input/Output Configuration Program User's Guide for ICP IOCP* for more information.

IOCP allows you to define controls for multiple channel subsystems. This includes changes to the way you define LPARs, channel paths, and I/O devices.

LPAR definition

Use the RESOURCE statement to define LCSSs and the logical partitions in each LCSS. You can also assign a MIF image ID to each LPAR. If you do not specify a MIF image ID using the RESOURCE statement, ICP IOCP assigns them. Any LPARs not defined will be reserved and available to be configured later using dynamic I/O.

Channel path definition

You can define shared channel paths in addition to dedicated and reconfigurable channel paths. The CHPID statement has an additional SHARED keyword to accomplish this. You can also define spanned channel paths using the PATH keyword. You can define:

- All channel paths as dedicated or reconfigurable.
- Only CNC, CTC, FC, FCP, CFP, ICP, IQD, CIB, OSC, OSD, OSE, OSN, OSX, and OSM channel paths as shared.
- Only FC, FCP, CFP, ICP, IQD, CIB, OSC, OSD, OSE, OSN, OSX, and OSM channel paths as spanned.

ICP IOCP provides access controls for spanned, shared or reconfigurable channel paths. Parameters on the PART | PARTITION or NOTPART keyword on the CHPID statement allow you to specify an access list and a candidate list for spanned, shared and reconfigurable channel paths.

The access list parameter specifies the logical partition or logical partitions that will have the channel path configured online at logical partition activation following the initial power-on reset of an LPAR IOCDS. For exceptions, refer to *zEnterprise System Processor Resource/Systems Manager Planning Guide*.

The candidate list parameter specifies the LPARs that can configure the channel path online. It also provides security control by limiting the logical partitions that can access shared or reconfigurable channel paths.

Note: PR/SM LPAR manages the channel path configuration across POR. Refer to *zEnterprise System Processor Resource/Systems Manager Planning Guide*.

I/O device definition

You can specify either the optional PART | PARTITION keyword or the optional NOTPART keyword on the IODEVICE statement to limit device access by logical partitions for devices assigned to shared ESCON, FICON, or OSA channels, or HiperSockets. (The IODEVICE candidate list is not supported for shared CFP, CIB, or ICP CHPIDs.)

By limiting access to a subset of logical partitions, you can:

- Provide partitioning at the device level.
- Provide security at the device level.
- Better manage the establishment of logical paths.

Hardware Configuration Definition (HCD) considerations

HCD provides the capability to make both dynamic hardware and software I/O configuration changes. It also provides:

- An online, interactive way to more usably manage the I/O configuration than IOCP.
- The capability to define the I/O configuration for dynamic or nondynamic I/O configuration purposes.

HCD allows you to define LPAR controls for defining LPARs, channel paths, and I/O devices. The following HCD panels (or corresponding HCM dialogs) support these controls.

Add Partition
Allows explicit definition of LPARs and associated LPAR numbers.

Define Access List
Allows definition of initial access list for channel path access control of shared and reconfigurable channel paths.

Define Candidate List (for channel paths)
Allows definition of candidate list for channel path access control of shared and reconfigurable channel paths.

Define Candidate List (for devices)
Allows definition of candidate list for device access control for devices assigned to shared channels.

Add Processor
Allows you to determine the capabilities of a CPC.

Add Channel Path
Operation mode field allows definition of a channel path as dedicated, reconfigurable, or shared.

Define Device / Processor
Additional field to specify candidate list.

Chapter 5. I/O connectivity

This chapter discusses the channels associated with the z196 I/O connectivity. You can also refer to Table 8 on page 23 for a summary of the I/O channel characteristics.

FICON and FCP channels

The FICON Express channel uses the industry standard Fibre Channel Standard as a base. It is an upper layer protocol that maps the channel architecture on the general transport vehicle used throughout the industry for such other upper layer protocols as SCSI, IPI, and IP, among others. This transport vehicle includes the physical definition, the transmission protocol, and signalling protocol that is the same for all of the other upper layer protocols.

The FICON Express8S, FICON Express8, and FICON Express4 features conform to the Fibre Connection (FICON) architecture, the High Performance FICON on System z (zHPF) architecture, and the Fibre Channel Protocol (FCP) architecture, providing connectivity between any combination of servers, directors, switches, and devices (control units, disks, tapes, printers) in a Storage Area Network (SAN).

There are two CHPID types that can be specified using IOCP or HCD. Each channel has its own unique CHPID type:

- CHPID type FC — native FICON, High Performance FICON for System z (zHPF), and channel-to-channel (CTC)
- CHPID type FCP — Fibre Channel Protocol (FCP) for communication with SCSI devices

FICON builds upon the strengths of ESCON. The FICON implementation enables full duplex data transfer. So data travels both directions simultaneously, rather than the ESCON half duplex data transfer. Furthermore, concurrent I/Os can occur on a single FICON channel, a fundamental difference between FICON and ESCON. The data rate drop is minimal with FICON even at distances up to 10 km.

Native FICON supports up to 64 concurrent I/O operations. ESCON supports one I/O operation at a time.

In conjunction with the Fibre Channel protocol (FCP), N_Port ID Virtualization (NPIV) is supported, which allows the sharing of a single physical FCP channel among operating system images.

FICON Express8S features

There are two FICON Express8S features for z196 — FICON Express8S 10KM LX and FICON Express8S SX. These features can only be used in a PCIe I/O drawer.

Each FICON Express8S feature has two channels per feature. Each of the two independent channels supports a link data rate of 2 Gbps (gigabits per second), 4 Gbps, or 8 Gbps, depending upon the capability of the attached switch or device. The link speed is autonegotiated point-to-point. A link data rate of 1 Gbps is not supported. Each channel utilizes small form factor pluggable optics (SFPs) with LC duplex connectors. The optics allow each channel to be individually repaired without affecting the other channels.

Each FICON Express8S feature supports cascading (the connection of two FICON Directors in succession) to minimize the number of cross-site connections and help reduce implementation costs for disaster recovery applications, GDPS, and remote copy.

The FICON Express8S features:

- Provide increased bandwidth and granularity for the SAN
- Support 8 GBps PCIe interface to the PCIe I/O drawer
- Provide increased performance with zHPF and FCP protocols
- Provide increased port granularity with two channels/ports per feature
- Include a hardware data router for increased performance for zHPF.

The FICON Express8S features for z196 include:

- **FICON Express8S 10KM LX (FC 0409)**

  FICON Express8 10KM LX utilizes a long wavelength (LX) laser as the optical transceiver and supports use of a 9/125 micrometer single mode fiber optic cable terminated with an LC duplex connector.

  FICON Express8 10KM LX supports distances up to 10 km (kilometers) (6.2 miles).

  FICON Express8 10KM LX (CHPID type FC or FCP) can be defined as a spanned channel and can be shared among LPARs within and across LCSS.

  The sending and receiving transceiver must be the same type, LX.
- **FICON Express8S SX (FC 0410)**

  FICON Express8S SX utilizes a short wavelength (SX) laser as the optical transceiver and supports use of a 50/125 micrometer multimode fiber optic cable or a 62.5/125 micrometer multimode fiber optic cable terminated with an LC duplex connector.

  Note: You cannot mix 50 and 62.5 micron multimode fiber optic cabling in the same link.

  For details about the unrepeated distances for FICON Express8S SX, refer to *System z Planning for Fiber Optic Links (ESCON, FICON, Coupling Links, and Open System Adapters)*.

  FICON Express8S SX (CHPID type FC or FCP) can be defined as a spanned channel and can be shared among LPARs within and across LCSS.

  The sending and receiving transceiver must be the same type, SX.

FICON Express8 features

There are two FICON Express8 features for z196 — FICON Express8 10KM LX and FICON Express8 SX. These features can only be used in an I/O drawer or I/O cage.

FICON Express 8 features can be carried forward or ordered on MES using RPQ 8P2534. During an MES, if slots are available in an I/O cage or I/O drawer and no slots are available in a PCIe I/O drawer, RPQ 8P2534 is used to order these features.

Each FICON Express8 feature has four channels per feature. Each of the four independent channels supports a link data rate of 2 gigabits (Gbps), 4 Gbps, or 8 Gbps per second, depending upon the capability of the attached switch or device, with autonegotiation to 2, 4, or 8 Gbps depending upon the capability of the attached device. A link data rate of 1 Gbps is not supported. Each channel utilizes small form factor pluggable optics (SFPs) with LC duplex connectors. The optics allow each channel to be individually repaired without affecting the other channels.

Each FICON Express8 feature supports cascading (the connection of two FICON Directors in succession) to minimize the number of cross-site connections and help reduce implementation costs for disaster recovery applications, GDPS, and remote copy.

The FICON Express8 features for z196 include:

- **FICON Express8 10KM LX (FC 3325)**

  All the channels on a single FICON Express8 10KM LX feature are the same type, 10KM LX. FICON Express8 10KM LX utilizes a long wavelength (LX) laser as the optical transceiver and supports use of a 9/125 micrometer single mode fiber optic cable terminated with an LC duplex connector.

  FICON Express8 10KM LX supports distances up to 10 km (6.2 miles).

FICON Express8 10KM LX (CHPID type FC or FCP) can be defined as a spanned channel and can be shared among LPARs within and across LCSS.

- **FICON Express8 SX (FC 3326)**

  All the channels on a single FICON Express8 SX feature are the same type, SX. FICON Express8 SX utilizes a short wavelength (SX) laser as the optical transceiver and supports use of a 50/125 micrometer multimode fiber optic cable or a 62.5/125 micrometer multimode fiber optic cable terminated with an LC duplex connector.

  Note: You cannot mix 50 and 62.5 micron multimode fiber optic cabling in the same link.

  For details about the unrepeated distances for FICON Express8 SX, refer to *System z Planning for Fiber Optic Links (ESCON, FICON, Coupling Links, and Open System Adapters)*.

FICON Express4 features

FICON Express4 features can only be carried forward.

FICON Express4 features can only be used in an I/O drawer or I/O cage.

The FICON Express4 features for the z196 include:

- **FICON Express4 10KM LX (FC 3321)**

  FICON Express4 10KM LX has four channels per feature. It is designed to support unrepeated distances up to 10 km (6.2 miles) over single mode fiber optic cabling.
- **FICON Express4 4KM LX (FC 3324)**

  FICON Express4 4KM LX has four channels per feature. It is designed to support unrepeated distances up to 4 km (2.5 miles) over single mode fiber optic cabling.
- **FICON Express4 SX (FC 3322)**

  FICON Express4 SX has four channels per feature. It is designed to carry traffic over multimode fiber optic cabling.

All channels on a single FICON Express4 feature are of the same type: 4KM LX, 10KM LX, or SX.

FICON Express4 supports a 4 Gbps link data rate with auto-negotiation to 1, 2, or 4 Gbps for synergy with existing switches, directors, and storage devices.

Note: You need to ensure that the tactical as well as the strategic requirements for your data center, Storage Area Network (SAN), and Network Attached Storage (NAS) infrastructures are taken into consideration as you employ 2 Gbps and beyond link data rates.

Mode Conditioning Patch (MCP) cables are only supported at the 1 Gbps link data rate.

Channel consolidation using FICON Express8

You can consolidate your FICON Express4 and FICON Express8 channels onto fewer FICON Express8S channels while maintaining and enhancing performance. You can also migrate ESCON channels to FICON Express8 channels. Contact your IBM representative for assistance.

Name server registration

Registration information is provided on the name server for both FICON and FCP, which enhances problem determination, analysis, and manageability of the storage area network (SAN).

High Performance FICON for System z (zHPF)

High Performance FICON for System z (zHPF) is an extension to the FICON architecture and is designed to improve the performance of small block and large block data transfers. zHPF supports multitrack operations and the transfer of greater than 64 kB of data in a single operation, resulting in higher throughputs with lower response times.

zHPF applies to all FICON Express8S, FICON Express8, and FICON Express4 features (CHPID type FC) on z196.

Discover and automatically configure devices attached to FICON channels

z196 provides a function, z/OS discovery and autoconfiguration (zDAC), that discovers and automatically configures control units and devices that are accessible to z196, but not currently configured.

This function performs a number of I/O configuration definition tasks for new and changed control units and devices attached to FICON channels. The proposed configuration incorporates the current contents of the I/O definition file (IODF) with additions for newly installed and changed control units and devices.

This function is designed to help simplify I/O configuration on z196 running z/OS and reduce complexity and setup time.

This function is supported by z/OS V1.12 with PTFs and applies to all FICON features (CHPID type FC) on z196.

The MIDAW facility

The Modified Indirect Data Address Word (MIDAW) facility is designed to improve FICON performance. The MIDAW facility:

- Improves FICON performance for extended format data sets. Non-extended data sets can also benefit from MIDAW.
- Reduces FICON channel and control unit overhead.

Multipath Initial Program Load (IPL)

If I/O errors occur during the IPL, z/OS on z196 allows the system to attempt an IPL on alternate paths, if the paths are available. The system will attempt the IPL on an alternate path until all paths have been attempted or until the IPL is successful.

This function is applicable for all FICON features with CHPID type FC.

Fibre channel analysis

You can use the **Fibre Channel Analyzer** task on the HMC to identify fiber optic cabling issues in your Storage Area Network (SAN) fabric without contacting IBM service personnel. All FICON channel error information is forwarded to the HMC where it is analyzed to help detect and report the trends and thresholds for all FICON channels on z196. This report shows an aggregate view of the data and can span multiple systems.

This applies to FICON channels exclusively (CHPID type FC).

Fibre Channel Protocol (FCP) for SCSI devices

Fibre Channel (FC) is a computer communications protocol that attempts to combine the benefits of both channel and network technologies. Fibre Channel made the biggest impact in the storage arena, specifically using Small Computer System Interface (SCSI) as an upper layer protocol.

Fibre Channel is broken down into five layers: FC-0, FC-1, FC-2, FC-3, and FC-4. The layers define the following functions:
- **FC-0** defines the physical characteristics
- **FC-1** defines the character encoding and link maintenance
- **FC-2** defines the frame format, flow control, classes of service
- **FC-3** defines the common services

FICON and FCP implement those layers, unchanged.
- **FC-4** defines the upper layer protocol mapping which includes SCSI as well as Fibre Channel - Single Byte-2 (FC-SB-2), which is FICON.

The Fibre Channel Protocol (FCP) capability, supporting attachment to Small Computer Systems Interface (SCSI) is based on the Fibre Channel (FC) standards defined by the INCITS, and published as ANSI standards. SCSI devices in Linux on System z environments, as well as SCSI devices defined to z/VM and z/VSE, are based on the Fibre Channel standards. FC is an upper layer fibre channel mapping of SCSI on a common stack of Fibre Channel physical and logical communication layers. HIPPI, IPI, IP, and FICON (FC-SB-2) are other examples of upper layer protocols.

SCSI is an industry-standard protocol that is supported by a wide range of controllers and devices that complement the System z9, System z10, and zEnterprise storage attachment capability through FICON and ESCON channels. FCP channels on System z9, System z10, and zEnterprise are provided to enable operating systems on System z9, System z10, and zEnterprise to access industry-standard SCSI storage controllers and devices.

FCP is the base for open industry-standard Fibre Channel networks or Storage Area Networks (SANs).

Fibre Channel networks consist of servers and storage controllers and devices as end nodes, interconnected by Fibre Channel switches, directors, and hubs. While switches and directors are used to build Fibre Channel networks or fabrics, Fibre Channel loops can be constructed using Fibre Channel hubs. In addition, different types of bridges and routers may be used to connect devices with different interfaces (like parallel SCSI). All of these interconnects may be combined in the same network.

For information about the configurations supported by the FCP channel, refer to “Configurations” on page 64.

An FCP channel is defined in the IOCP as channel type FCP and is available on FICON features.

FCP channels support full-fabric support. The FCP full-fabric support means that multiple numbers of directors/switches can be placed between the server and FCP/SCSI device, thereby allowing many hops through a storage network for I/O connectivity.

In addition, for FCP channels, a high integrity fabric solution is not required but is recommended. If an FCP Interswitch Link (ISL) is moved, data could potentially be sent to the wrong destination without notification.

The FICON Express4, FICON Express8, and FICON Express8S features, when defined as CHPID type FCP in the IOCP, support storage controllers and devices with an FCP interface in z/VM, z/VSE, and Linux on System z environments.

Each port on a single FICON card can be configured individually and can be a different CHPID type.

FCP channels support T10-DIF

System z FCP has implemented support of the American National Standards Institute's (ANSI) T10 Data Integrity Field (T10-DIF) standard. With this support, data integrity protection fields are generated by the operating system and propagated through the storage area network (SAN). System z helps to provide added end-to-end data protection between the operating system and the storage device.

An extension to the standard, Data Integrity Extensions (DIX), provides checksum protection from the application later through the host bus adapter (HBA), where cyclical redundancy checking (CRC) protection is implemented.

T10-DIF support by the FICON Express8S and FICON Express8 features, when defined as CHPID type FCP, is exclusive to z114 and z196. Exploitation of the T10-DIF standard is required by the control unit.

Configurations

Storage controllers and devices with an FCP interface can be directly attached to zEnterprise (point-to-point connection), or by using Fibre Channel switches or directors. A storage controller or device with an appropriate FCP interface may be attached to each port of a FICON feature, or of a Fibre Channel switch or director.

In addition, the following devices and controllers can be attached to each port on a Fibre Channel switch or director:

- **FC-AL controllers or devices, and FC-AL hubs**

  If the switch or director supports the Fibre Channel Arbitrated Loop (FC-AL) protocol, devices implementing this protocol may be attached to that port and accessed from System z9, System z10, or zEnterprise. Devices typically implementing the FC-AL protocol are tape units and libraries, and low-end disk controllers.

  If the switch or director does not support the FC-AL protocol, you can also install a FC-AL bridge between the switch or director and the FC-AL controller or device.

  If more than one FC-AL controller or device should be attached to a FC-AL switch port, it is convenient to use a Fibre Channel hub, where multiple devices with a FC-AL interface can be directly attached to that hub.
- **Fibre-Channel-to-SCSI bridges**

  Fibre-Channel-to-SCSI bridges can be used to attach storage controllers and devices implementing the electrical, parallel SCSI interface. Different types of Fibre-Channel-to-SCSI bridges may support different variants of the parallel SCSI interface, such as Low Voltage Differential (LVD), High Voltage Differential (HVD), Single Ended, wide (16-bit) versus narrow (8-bit) interfaces, and different link speeds.

Each FCP channel (CHPID) can support up to 480 subchannels, where each subchannel represents a communication path between software and the FCP channel. Refer to “Channel and device sharing” on page 65 for more information.

Host operating systems sharing access to an FCP channel can establish in total up to 2048 concurrent connections to up to 510 different remote Fibre Channel ports associated with Fibre Channel controllers.

The total number of concurrent connections to end devices, identified by logical unit numbers (LUNs), must not exceed 4096.

I/O devices

The FCP channel implements the FCP standard as defined by the INCITS Fibre Channel Protocol for SCSI (FCP), and Fibre Channel Protocol for SCSI, Second Version (FCP-2), as well as the relevant protocols for the SCSI-2 and SCSI-3 protocol suites. Theoretically, each device conforming to these interfaces should work when attached to an FCP channel as previously defined. However, experience tells us that there are small deviations in the implementations of these protocols. Therefore, it is advisable to do appropriate conformance and interoperability testing to verify that a particular storage controller or device can be attached to an FCP channel in a particular configuration (i.e. attached via a particular type of Fibre Channel switch, director, hub, or Fibre-Channel-to-SCSI bridge).

Also, for certain types of FCP and SCSI controllers and devices, specific drivers in the operating system may be required in order to exploit all the capabilities of the controller or device, or to cope with unique characteristics or deficiencies of the device.

Information about switches and directors qualified for IBM System z FICON and FCP channels is located on Resource Link (*http://www.ibm.com/servers/resourcelink*) on the **Library** page under "Hardware products for servers."

Addressing

FCP channels use the Queued Direct Input/Output (QDIO) architecture for communication with the operating system. IOCP is only used to define the QDIO data devices. The QDIO architecture for FCP channels, derived from the QDIO architecture that had been defined for communications via an OSA card, defines data devices that represent QDIO queue pairs, consisting of a request queue and a response queue. Each queue pair represents a communication path between an operating system and the FCP channel. It allows an operating system to send FCP requests to the FCP channel via the request queue. The FCP channel uses the response queue to pass completion indications and unsolicited status indications to the operating system.

IOCP is not used to define the actual Fibre Channel storage controllers and devices, nor the Fibre Channel interconnect units such as switches, directors, or bridges. IOCP is only used to define the QDIO data devices. The Fibre Channel devices (end nodes) in a Fibre Channel network are addressed using World Wide Names (WWNs), Fibre Channel Identifiers (IDs), and Logical Unit Numbers (LUNs). These addresses are configured on an operating system level, and passed to the FCP channel together with the corresponding Fibre Channel I/O or service request via a logical QDIO device (queue).

Channel and device sharing

An FCP channel can be shared between multiple operating systems, running in a logical partition or as a guest operating system under z/VM. Under z/VM, multiple z/VM, CMS, Linux on System z, and z/VSE guests are able to share SCSI channels and devices using z/VM Fixed Block Architecture (FBA) emulation. To access the FCP channel, each operating system needs one FCP device on an FCP channel.

Each FCP channel can support up to 480 QDIO queue pairs. This allows each FCP channel to be shared among 480 operating system instances.

Channel and device sharing using NPIV: N_Port ID Virtualization (NPIV) allows the sharing of a single physical FCP channel and attached devices, logical units, among operating system images, whether in logical partitions or as z/VM guests in virtual machines. This is achieved by assigning a unique WWPN to each subchannel that is defined for an FCP Channel using IOCP.

Each operating system instance using such a subchannel and its associated QDIO queues therefore also uses its own WWPN. When the operating system image starts using its subchannel, the FCP channel performs a login to the Fibre Channel fabric and acquires a unique Fibre Channel ID, also called N_Port ID. This ID is used in all further Fibre Channel communication that is done on behalf of this operating system image.

Access controls based on the assigned WWPN can be applied in the SAN environment, using standard mechanisms such as zoning in FC switches and Logical Unit Number (LUN) masking in the storage controllers. You can configure the SAN prior to the installation of a new machine using the WWPN tool available on Resource Link.

NPIV exploitation requires a Fibre Channel director or switch that supports the NPIV standard. If such a director or switch is installed, NPIV mode can be enabled for the FCP channel that attaches to this Fibre Channel switch or director through the Support Element. This enablement can be done on logical partition base, i.e., per FCP channel image.

NPIV is not supported in a point-to-point topology.

Channel and device sharing without NPIV: Without NPIV support, multiple operating system images can still concurrently access the same remote Fibre Channel port through a single FCP channel. However,

Fibre Channel devices or logical units, identified by their LUNs, cannot be shared among multiple operating system images through the same FCP channel.

Positioning

FCP and SCSI are industry-standard protocols, which have been implemented by many vendors in a large number of different types of storage controllers and devices. These controllers and devices have been widely accepted in the market place and proven to be adequate to meet the requirements regarding reliability, availability, and serviceability (RAS) in many environments.

However, it must be noted that there are some advanced, unique RAS characteristics of zEnterprise storage attachments based on ESCON and FICON attachments, using channel programs (and the Extended Count Key Data (ECKD™) protocol in the case of disk control units), that may not be readily available in such an FCP or SCSI based world. Therefore, whenever there are very stringent requirements regarding isolation, reliability, availability, and serviceability, a conscious decision must be made whether FCP attached storage controllers and devices or FICON or ESCON attached control units should be used. Customers requiring the more robust RAS characteristics should choose FICON or ESCON channels.

SCSI Initial Program Load (IPL)

This function allows you to IPL an operating system from an FCP-attached disk, to execute either in a logical partition or as a guest operating system under z/VM. In particular, SCSI IPL can directly IPL a z196 operating system that has previously been installed on a SCSI disk. Thus, there is no need for a classical channel (ESCON or FICON) attached device, such as an ECKD disk control unit, in order to install and IPL a z196 operating system. The IPL device is identified by its Storage Area Network (SAN) address, consisting of the WWPN of the disk controller and the Logical Unit Number (LUN) of the IPL device.

You can also IPL a standalone-dump program from an FCP channel attached SCSI disk. The standalone-dump program can also store the generated dump data on such a disk.

SCSI IPL in z/VM allows Linux on System z, z/VSE, and other guest operating systems that support SCSI IPL to be IPLed from FCP-attached SCSI disk, when z/VM is running on a z196. Therefore, z/VM, z/VSE, and Linux on System z guests may be started and run completely from FCP channel attached disk in your hardware configuration.

z/VM provides the capability to install z/VM from a DVD to an Enterprise Storage Server® (ESS) SCSI disk emulated as a Fixed Block Architecture (FBA) disk as well as an Enterprise Storage Server from a DVD to a 3390 disk. Thus, z/VM and its Linux on System z guests may be started and run completely from FCP disks on your hardware configuration. Refer to z/VM subset of the 2817DEVICE Preventive Service Planning (PSP) bucket for any service required for z/VM support for SCSI IPL.

z/VM supports SCSI-attached disks to be used for installation, IPL, and operations such as storing dumps, and other functions, while continuing to provide support for ESCON-attached of FICON-attached disk or tape.

z/VM SCSI support allows a Linux on System z server farm and z/VSE to be deployed on z/VM in a configuration that includes only SCSI disks.

z/VM provides the capability to dump Linux on System z guests to FCP-attached SCSI disks. Benefits include:

- More guest virtual memory can be dumped because SCSI disks can be larger than ECKD disks
- Avoids the need to convert a VMDUMP into a Linux tool format
- Allows the same SCSI dump mechanisms to be used when running Linux for System z in an LPAR and in a z/VM virtual machine.

For Linux on System z support for SCSI IPL, refer to this website: *http://www.ibm.com/developerworks/linux/linux390/*.

z/VSE supports FCP-attached SCSI disks for installation and IPL. For z/VSE SCSI support, refer to the appropriate z/VSE publications (for example, *z/VSE Administration*).

For additional information on:

- How to use SCSI IPL for a logical partition, refer to the *zEnterprise System Support Element Operations Guide* or to the *System z Hardware Management Console Operations Guide*
- Messages that can show up on the operating systems console on the SE or Hardware Management Console, refer to *System z Small Computer Systems (SCSI) IPL - Machine Loader Messages*
- How to use SCSI IPL for a z/VM guest, refer to *http://www.vm.ibm.com/pubs* for appropriate z/VM publications
- How to prepare a Linux on System z IPL disk or a Linux on System z dump disk, refer to *http://www.ibm.com/developerworks/linux/linux390/* for appropriate Linux on System z publications.

ESCON channels

The ESCON channel provides a 17 MBps link data rate between host and control units for I/O devices. ESCON supports half-duplex data transfers over 62.5 multimode fiber optic cabling.

ESCON can only be used in an I/O drawer or I/O cage.

The ESCON channel provides a light-emitting diode (LED) light source for fiber optic cables. It can extend up to 3 kilometers (1.86 US miles), a range that can be further extended to 6 or 9 kilometers (km) by retransmission through one or two ESCON directors.

With the availability of two LCSSs, you can define a maximum of 240 ESCON channels on your z196 up to a maximum of 16 features per system (360 channels up to a maximum of 24 features with RPQ 8P2507). The maximum number of configurable channels is 256 per LCSS and per operating system image. The high density ESCON feature has 16 ports, 15 of which can be activated for your use. One port is always reserved as a spare, in the event of a failure of one of the other ports. When four ports are ordered, two 16-port ESCON features are installed and two ports are activated on each feature. After the first pair, ESCON features are installed in increments of one. ESCON channels continue to be ordered in increments of four channels.

ESCON supports these operating system environments: z/OS, z/VM, z/VSE, z/TPF, and Linux on System z.

ESCON channels affect the performance of the channel subsystem. Maximizing channel subsystem performance is an important consideration in configuring I/O devices to a z196 general purpose model CPC. Channel subsystem performance depends on the factors described in this chapter.

For an explanation of basic ESCON channel concepts, refer to *Introducing Enterprise Systems Connection*. For detailed information about synchronous and nonsynchronous I/O operation, refer to *Storage Subsystem Library Introduction to Nonsynchronous Direct Access Storage Subsystems*.

Multiple Initial Program Load (IPL)

If I/O errors occur during the IPL, z/OS on z196 allows the system to attempt an IPL on alternate paths, if the paths are available. The system will attempt the IPL on an alternate path until all paths have been attempted or until the IPL is successful.

This function is applicable for all FICON features with CHPID type FC.

ESCON converter operation

You can configure ESCON converter channels (attached to a parallel converter - for example, the IBM 9034 or the Optica 34600 FXBT) for block and byte multiplexer mode of operation. The data mode of operation is determined by the multiplexer mode (byte or block). This is selected for specific channels when either the CPC or an LPAR is initialized.

As many as eight channel paths are available to attach to any I/O device. During any I/O operation, one of the available channel paths is selected. Channel path selection is a hardware function rather than a function of the system control program.

At the start of an I/O operation, a central processor signals the channel subsystem that an I/O operation is needed. An I/O request is posted to a queue; meanwhile, instruction execution in the central processor continues.

Channel multiplexing modes

The data mode of operation is determined by the multiplexer mode (block or byte). This is selected for specific channels when either the CPC or a logical partition is initialized.

Block Multiplexer Mode of Operation: In block multiplexer mode of operation, a device stays connected to a channel continuously during the transfer of a full block of data.

Block multiplexer mode of operation allows a control unit to present “channel end” and to disconnect from a channel at the completion of a specified operation. “Device End” is presented at a later point. During the interval between “channel end” and “device end” another device attached to the same channel can be started or can complete an operation that is ready. However, if the second device does connect to the same channel during this interval, the first device may find the channel busy when it tries to reconnect, and then the first device must wait for service.

ESCON can be configured for block multiplexer mode of operation. In block multiplexer mode of operation, ESCON channels configured as CVC channel paths can operate in either interlock (high-speed transfer) mode or in data-streaming mode. They can also be attached to control units that operate in high-speed transfer or in data-streaming mode. Data rates can be as high 4.5 MBps for ESCON CVC channel paths.

Byte multiplexer mode of operation: Byte interleave mode of operation allows the execution of multiple I/O operations concurrently. Byte multiplexer mode permits several relatively slow-speed I/O devices to operate at the same time. Each addressed device requesting service is selected for transfer of a byte or a group of bytes to or from main storage. Bytes from multiple devices are interleaved on the channel and routed to or from the desired locations in main storage.

The load that a byte multiplexer channel can sustain is variable. It is governed by I/O device performance factors such as the data transfer rate, device buffers, number of bytes per data burst on the channel, channel program requirements, synchronized mechanical motion, and priority sequence position on the I/O interface.

ESCON converter channels (defined as CBY) can be configured for byte multiplexer mode of operation. In byte multiplexer mode of operation, ESCON channels configured as CBY channel paths can operate in either byte multiplexer mode or in burst mode. CBY channels require a 9034 ESCON converter. Byte multiplexer mode permits several relatively slow-speed I/O devices to operate at the same time.

Refer to the 2817IO subset id in the 2817DEVICE upgrade ID of the preventive service planning (PSP) bucket for prerequisite 9034 EC level information.

Byte multiplexer mode and burst mode: A byte multiplexer channel can be monopolized by one I/O device (burst mode) or shared by many I/O devices (byte multiplexer mode). The number of bytes transferred at a time in byte multiplexer mode can be one (single byte transfers) or more than one

(multibyte transfers). Most control units that operate in byte multiplexer mode can also operate in burst mode. A manually set switch at the control unit determines whether the control unit operates in burst mode or byte multiplexer mode.

Some devices offer a choice of how many bytes are transferred during a single data transfer sequence in byte multiplexer mode.

Because most of the time spent in a data-transfer control sequence is for control, increasing the burst size (the number of bytes transferred per sequence) results in a relatively small increase in the total channel busy time for the sequence. Also, increasing the burst size reduces the number of data transfer sequences required. The net effect is a significant improvement in channel efficiency and a higher allowable data rate.

Burst mode, although most effective in the use of channel resources, can cause another device on the byte multiplexer channel to exceed its critical time. From the perspective of the control unit, burst mode occurs when the time contributed by the control unit in a transfer sequence is more than 32 microseconds. (Refer to the *Enterprise System Architecture/390 System 360 and System 370 I/O Interface Channel to Control Unit OEMI*.)

If the device configuration guidelines are followed for byte multiplexer channels on a general purpose model CPC, deferred accesses are minimized and data transfer sequences exceeding 32 microseconds are acceptable when large burst sizes are specified.

Most class-2 and class-3 devices that can operate in burst mode should be attached to block multiplexer channels for better performance.

I/O operations control

ESA/390 and z/Architecture I/O operations are performed by executing a channel program that consists of one or more chained Channel Command Words (CCWs). Each CCW contains a command and other information that is used by both the channel and control unit in executing the I/O operation.

Channel commands are segmented into six basic categories with many variations based on control unit type. A channel program is initiated and controlled by executing one or more of the ESA/390 and z/Architecture I/O instructions described below. I/O interruptions may result during the execution of a channel program to notify the CP of progress or completion.

Channel commands

The six basic channel commands are:

Write Initiates the transfer of data from main storage to an I/O device.

Read Initiates the transfer of data from an I/O device to main storage.

Read Backward
Initiates the transfer of data from an I/O device to main storage, storing data bytes in reverse order.

Control
Specifies operations such as set tape density, rewind tape, advance paper in a printer, or sound an audible alarm.

Sense Requests information from a control unit. The information contains unusual conditions detected during the last I/O operation and detailed device status.

Transfer in Channel (TIC)
Specifies the location in main storage where the next CCW in the channel program is to be fetched. The TIC command provides branching between CCWs in noncontiguous storage areas. A TIC command cannot specify a CCW containing another TIC command.

ESA/390 and z/Architecture mode I/O instructions

In ESA/390 mode or z/Architecture mode, any CP can initiate I/O operations with any I/O device and can handle I/O interruptions from any I/O device. Each I/O device is assigned a unique device number, and is associated with one subchannel.

The CPs communicate with devices by specifying the appropriate subchannel. The subchannel uses the assigned device address to communicate with the device over one or more channel paths. The device number provides a path-independent means to refer to a device for use in operator messages or at the time of IPL.

For descriptions of these instructions, refer to the *Enterprise System Architecture/390 Principles of Operation* or *z/Architecture Principles of Operation* manual.

The I/O instructions for operation in ESA/390 mode or z/Architecture mode are:

- Start Subchannel (SSCH)
- Test Subchannel (TSCH)
- Clear Subchannel (CSCH)
- Halt Subchannel (HSCH)
- Resume Subchannel (RSCH)
- Store Subchannel (STSCH)
- Modify Subchannel (MSCH)
- Test Pending Interruption (TPI)
- Reset Channel Path (RCHP)
- Set Channel Monitor (SCHM)
- Store Channel Report Word (STCRW)
- Cancel Subchannel (XSCH)
- Set Address Limit (SAL)
- Store Channel Path Status (STCPS).

The SSCH instruction specifies an operation request block, which designates the channel program.

Chaining operations

Following the transfer of information over a channel designated by a Channel Command Word (CCW), an operation initiated by the Start Subchannel (SSCH) instruction can be continued by fetching a new CCW. Fetching a new CCW immediately following the completion of the previous CCW is called **chaining**. Chaining is described in more detail in the *Enterprise System Architecture/390 Principles of Operation* or *z/Architecture Principles of Operation*.

CCWs located in contiguous areas of central storage (successive doubleword locations) can be chained. Chains of CCWs located in noncontiguous storage areas can be coupled for chaining purposes by using a Transfer in Channel command. All CCWs in a chain refer to the I/O device specified in the original instruction.

The type of chaining (data or command) is specified by chain-data and chain-command flag bits in the CCW.

Data chaining
When the data transfer specified by the current CCW is finished, data chaining causes the operation to continue by fetching a new CCW and using the storage area defined by the new CCW. Execution of the operation at the I/O device is not affected.

Command chaining
Each time a new CCW is fetched during command chaining, a new I/O operation is specified. The new operation is initiated when the device end signal for the current operation is received, unless suspension is specified in the new CCW. When command chaining takes place, the completion of the current operation does not cause an I/O interruption.

I/O interruptions

I/O interruptions report the completion of I/O operations to the CPs, error and time-out conditions, and progress.

Ending status information about the operation is available to the control program at the end of the I/O operation. When an I/O operation is completed, an I/O interruption request is sent to a central processor. When the request is honored, an I/O interruption occurs and places the central processor under control of the I/O new program status word (PSW). Until an I/O interruption condition is honored, it is called a pending I/O interruption.

Errors detected by the channel subsystem are reported to the CPs as I/O interruptions or machine-check interruptions. I/O interruptions report the following hardware-related conditions:

- Interface Control Check (IFCC) - For example, interface tag errors and time-outs.
- Channel Control Check (CCC) - For example, parity, decode, or control errors.
- Channel Data Check (CDC) - For example, a parity error detected in central storage.

Machine-check interruptions include the following:

- Unrecoverable errors (retry is unsuccessful).
- Persistent errors (retry can be attempted, but the error threshold is exceeded).
- Serious channel element errors that require immediate reporting or cannot be reported as an IFCC or CCC with an I/O interruption.

Resets

An I/O system reset is issued to all channels, and the channels signal a system reset to all attached I/O devices.

An I/O system reset:

- Stops all subchannel operations.
- Resets interruptions and status in all subchannels.

An I/O system reset occurs as part of:

- Channel subsystem power-on reset.
- Initial program load.
- System reset.

A channel issues a selective reset to a specific I/O device in response to an IFCC, CCC, or as part of execution of the clear subchannel instruction. The status of the specific device is reset.

I/O interface protocol

The I/O interface protocol is determined by the interface sequencing operations selected for specific control units and their associated devices that are attached to the channel.

Channel-to-Channel connection

The Channel-to-Channel (CTC) function simulates an I/O device that can be used by one system control program to communicate with another system control program. It provides the data path and synchronization for data transfer between two channels. When the CTC option is used to connect two channels that are associated with different system, a loosely coupled multiprocessing system is established. The CTC connection, as viewed by either of the channels it connects, has the appearance of an unshared I/O device.

The CTC is selected and responds in the same manner as any I/O device. It differs from other I/O devices in that it uses commands to open a path between the two channels it connects, and then synchronizes the operations performed between the two channels.

ESCON CTC support: The parallel I/O CTC architecture defines two operating modes for CTC communication: basic mode and extended mode. ESCON CTC support for both of these modes is available.

ESCON channels (using link-level and device-level protocols): You can achieve ESCON channel-to-channel connections between CPCs with ESCON or FICON Express channels if one of the ESCON channels is defined to operate in channel-to-channel (CTC) mode.

ESCON channels that operate in CTC mode (extended mode or basic mode) can be defined as shared ESCON channels. For more information, refer to "Multiple Image Facility (MIF)" on page 55.

For detailed information about the ESCON channel-to-channel adapter, refer to *Enterprise Systems Architecture/390 ESCON Channel-to-Channel Adapter*.

Channel time-out functions

The optional time-out function described here applies only to ESCON channels that attach to a 9034 ESCON converter channel.

Each channel path has I/O interface time-out functions that time the control unit delays in completing the following I/O interface sequences:

- A 6-second time-out for all selection and status presentation sequences. A time-out occurs if the sequence is not complete within 6 seconds.
- A 30-second time-out for data transfer. A time-out occurs if a byte of data is not transferred within 30 seconds.

If a time-out occurs, the channel terminates the I/O request to the control unit and generates an IFCC interruption.

The time-out function detects malfunctions in control units and I/O devices that can cause the channel path to be unusable to other control units and I/O devices. The time-out function is specified as active or inactive for a device by IOCP when the IOCDS is created.

Control unit (CU) priority on an I/O interface

CU priority on an I/O interface applies only to ESCON channels attached to a 9034 ES connection converter channel.

CU priority on the I/O interface of a channel depends on the order in which they were attached. If the CUs are connected to the "select out" line, the first CU has the highest priority. If the CUs are attached to the "select in" line, the priority sequence is reversed. CUs attached to the "select out" line have priority over CUs attached to the "select in" line.

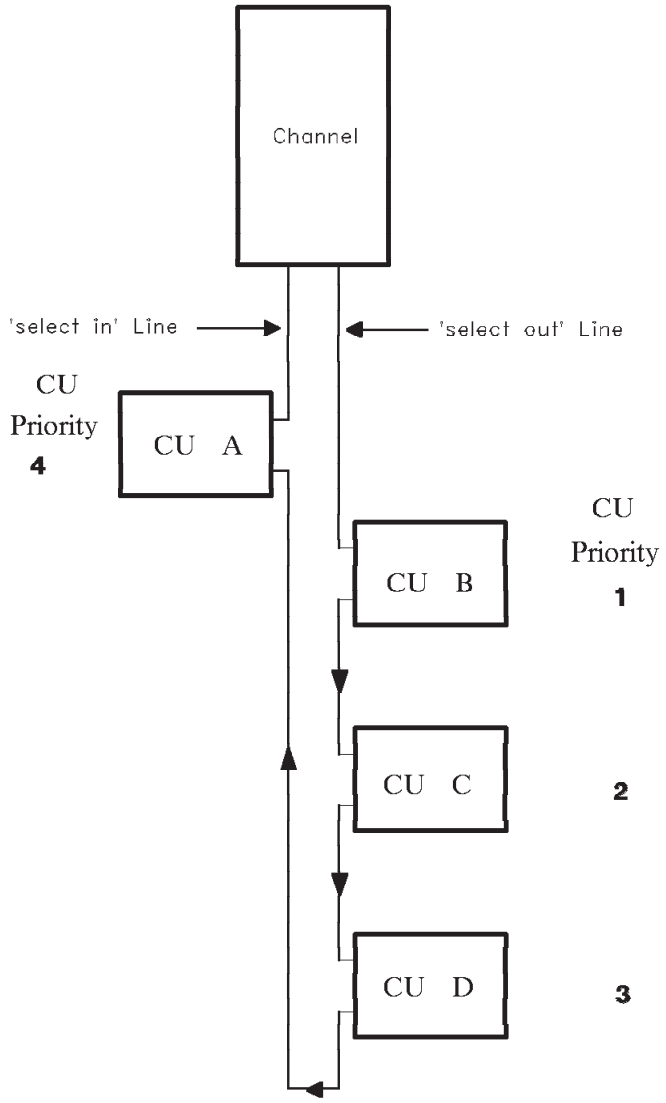


Figure 13. Control Unit (CU) priority on ESCON channels attached to a 9034 ES connection converter

Dynamic reconnection

The channel subsystem permits dynamic reconnection of I/O devices that have the dynamic-reconnection feature installed and that are set up to operate in a multipath mode, such as the IBM 3390 Direct Access Storage Model A14 or A22. Dynamic reconnection allows the device to reconnect and continue a chain of I/O operations using the first available channel path (one of as many as eight possible channel paths defined in an IOCP parameter). The selected path is not necessarily the one used initially in the I/O operation.

ESCON channel performance

Channel subsystem performance can be examined by observing two measurements:

- Response time (the amount of time taken to complete an I/O operation).
- Throughput (the number of I/O operations an I/O subsystem can complete in a given amount of time).

Channel subsystem response time and throughput can be divided into four major components:

- **Queuing and setup time**
  - The time taken for a channel path, control unit, and device to become available.
  - The time taken for a channel to send the I/O operation commands to the control unit.
- **Control unit and device time**

  The time required by the control unit and device to prepare for the transfer of data for the I/O operation. For example, a non-cached DASD control unit may have to wait for the DASD's seek and latency times before being ready to accept or send data.
- **Data transfer time**

  The time it takes to transfer the data for the I/O operation.
- **Completion time**

  The time it takes for the channel and control unit to post the status of and end the I/O operation.

Factors affecting ESCON channel performance

Factors that affect the various components of performance include:
- Synchronous or nonsynchronous type of operation
- Data transfer rate
- Attached device characteristics
- Channel subsystem workload characteristics.

Synchronous and nonsynchronous I/O operation: For detailed information about concepts described in this section, refer to *Storage Subsystem Library Introduction to Nonsynchronous Direct Access Storage Subsystems*.

Synchronous operation

Most DASD devices in a parallel environment transfer data synchronously. Synchronous operation requires that the channel, control unit, and device be active at the same time.

All work involved in ending an operation and advancing to the next operation must be completed before the DASD head reaches the next record (commonly referred to as the inter-record gap). If this does not occur, a rotational positional sensing/sensor (RPS) miss or an overrun is generated and the operation must wait for one DASD revolution before continuing.

Nonsynchronous operation

Nonsynchronous operation removes the requirements of synchronous operation. During nonsynchronous operation, the channel, control unit, and device do not have to be active at the same time to perform an I/O operation; thereby:
- Increasing DASD storage potential (by reducing inter-record gap).
- Allowing the channel and control units to be separated by longer distances.
- Eliminating command overruns.
- Reducing response time (by reducing RPS misses).
- Permitting the channel to perform other operations during the time it would normally wait for the device (this increases the throughput of the system).

Extended count key data (ECKD) channel programs are required to gain the benefits of nonsynchronous I/O operations. Count key data (CKD) channel programs are supported, but without the benefit of nonsynchronous operation. CKD channel-program performance could be degraded relative to ECKD channel programs in a nonsynchronous environment.

Data transfer rate: One of the factors that affects channel performance is the data transfer rate. The I/O subsystem data rate is the data transfer rate between processor storage and the device during an I/O operation.

The I/O subsystem data rate is made up of three components:
- **Channel data rate**

The rate that the channel transfers data between the transmission link and processor storage during an I/O operation. For ESCON channels, the link speed is 20 MBps and the channel data rate is 17 MBps at 0 distance. The data rate increases with distance.

- **Control unit data rate**

  The rate that the control unit transfers data between the control unit and the transmission link during an I/O operation.

- **Device data rate**

  The rate of data transfer between the control unit and the device. This rate depends on the control unit and device you use.

The I/O subsystem data rate is the lowest of the channel data rate, the control unit data rate, and the device data rate. In cases where the data comes from the control unit or is stored on the control unit and not directly to the device (for example, a cache read), the I/O subsystem data rate is the lower of the two: channel data rate or the control unit data rate.

The I/O subsystem data rate affects only the data transfer portion of the response time for an I/O operation. Response time and throughput both improve (response time decreases and throughput increases).

I/O device characteristics: The characteristics of devices attached to a channel subsystem can have a substantial effect on performance. Device characteristics such as caches, buffers, and data transfer rates all affect response time and throughput.

Channel subsystem workload characteristics: The performance of a specific I/O configuration varies based on the workload characteristics of that configuration. Two significant factors that determine workload characteristics and affect response time and throughput are channel program characteristics and cache-hit rates.

Channel program characteristics

Channel program characteristics affect channel subsystem performance. ESCON channel subsystems using link-level and device-level protocols perform nonsynchronous data transfers, and should use extended count key data (ECKD) channel programs.

Count key data (CKD) channel programs run in an ESCON environment, but may increase response times and reduce throughput due to lost DASD rotations.

Channel programs that contain indirect data address words (IDAWs), Transfer in Channel commands (TICs), and chained data commands, or that have poorly-aligned data boundaries, cause longer storage-response and increase channel subsystem response times.

Chained data commands increase response time due to an additional interlocked exchange between the channel and control unit. Refer to “ESCON performance characteristics” on page 76 for more information.

The amount of data to be transferred per I/O operation affects throughput. As the amount of data transferred per I/O operation increases (the ratio of data transferred to overhead improves), throughput improves.

Cache-hit rates

For control units which implement caches, cache-hit rates affect the channel subsystem performance. As the cache-hit rate increases, response time and throughput improve. The cache-hit rate is the percentage of times when data needed for a read operation is in the control unit's cache. For example, a cache-hit rate of 70% means that the required data is in the cache for 7 out of 10 read operations.

The cache-hit rate is significant because data is transferred out of the cache at the control unit's maximum data transfer rate, while data from the device is transferred at lower device speeds. This means that the higher the cache-hit rate, the better the response time and the better the throughput.

ESCON performance characteristics

With ESCON channels you need to consider the distance between the channel and the control unit since this affects the setup and completion times of an I/O operation. As the distance between the channel and the control unit increases, the response time increases and the throughput decreases. Channel and control unit utilization also increases as distance between the channel and control unit increases.

The speed of data transfer through fiber optic cable is subject to the Propagation delay time is determined by two factors: the speed of light through the optical fiber (which is fixed), and the length of the fiber optic link. Propagation delay time increases as the distance between elements in a fiber optic environment increase.

Interlocked exchange affects response time. Interlocked exchange requires that the channel (or control unit) wait for a response from the control unit (or channel) before proceeding with the next step of an I/O operation. As distance increases, the interlocked-exchange response time increases because of longer propagation delay times.

The throughput and response time for a shared ESCON channel are comparable to that of an unshared ESCON channel with comparable workload.

OSA channels

OSA channels include all OSA-Express2, OSA-Express3, and OSA-Express4S features.

z196 supports a maximum number of 48 features and 96 ports for the combined OSA-Express4S, OSA-Express3, and OSA-Express2 features.

Note: Unless noted differently, throughout this section, the term “OSA features” refers to all the OSA-Express4S, OSA-Express3 and OSA-Express2 features.

Supported CHPID types

OSA channels support the following modes of operation:

- CHPID type OSD
  - OSA-Express4S, OSA-Express3, or OSA-Express2 feature is running in QDIO mode.

    QDIO mode is the preferred architecture on z196 for high-speed communication, helping to reduce host interruptions and improve response time.
  - TCP/IP traffic when Layer 3
  - Protocol-independent when Layer 2
- CHPID type OSE
  - OSA-Express3 or OSA-Express2 feature is running in non-QDIO mode
  - SNA/APPN/HPF and/or TCP/IP passthru (LCS)
  - OSA-Express2 1000BASE-T Ethernet or OSA-Express3 1000BASE-T Ethernet is required.
- CHPID type OSC
  - OSA-Integrated Console Controller (OSA-ICC)
  - TN3270E, non-SNA DFT to IPL CPCs and LPARs
  - Operating system console operations
  - OSA-Express2 1000BASE-T Ethernet or OSA-Express3 1000BASE-T Ethernet is required.
- CHPID type OSN
  - OSA-Express for Network Control Program (NCP)

  - Supports channel data link control (CDLC) protocol. This provides connectivity between System z operating systems and IBM Communication Controller for Linux (CCL).

    CCL allows you to keep data and applications on the mainframe operating systems while moving NCP function to Linux on System z. CCL on System z helps to improve network availability by replacing token-ring networks and ESCON channels with an Ethernet network and integrated LAN adapters on zEnterprise, OSA-Express3 GbE or 1000BASE-T Ethernet features, or OSA-Express2 GbE or 1000BASE-T Ethernet features.
  - Requires the configuring to be done on a port-by-port basis
  - Used exclusively for internal communication, LPAR-to-LPAR
  - CHPID type OSN is not supported on the OSA-Express4S GbE features.
- CHPID type OSX
  - Provides connectivity and access control to the intraensemble data network (IEDN) from zEnterprise to zBX
  - Supported for OSA-Express3 10 GbE SR, OSA-Express3 10 GbE LR, OSA-Express4S 10 GbE SR, and OSA-Express4S 10 GbE LR only.
- CHPID type OSM
  - Provides connectivity to the intranode management network (INMN) from zEnterprise to Unified Resource Manager functions
  - Supported for OSA-Express3 1000BASE-T Ethernet only. Each z196 in an ensemble must have a pair of OSA-Express3 1000BASE-T Ethernet connections to the Bulk Power Hub (BPH) operating at 1 Gbps.

For more detailed information on these CHPID types and operating modes, refer to *zEnterprise, System z10, System z9 and zSeries Open Systems Adapter-Express Customer's Guide and Reference*.

OSA/SF

The Open Systems Adapter Support Facility (OSA/SF) is a host-based tool to support and manage the OSA features operating in QDIO (CHPID type OSD), non-QDIO mode (CHPID type OSE), or for OSA-Express for NCP (CHPID type OSN). The OSA/SF is used primarily to manage all OSA ports, configure all OSA non-QDIO ports, and configure local MACs.

One OSA/SF application can communicate with all OSA features in a hardware complex. OSA/SF communicates with an OSA feature through a device predefined on the OSA feature. The device type is OSAD.

OSA/SF is not required to set up the OSA features in QDIO mode (CHPID type OSD). However, it can be used to set up MAC addresses and set adapter speed. For channels (CHPID type OSN), OSA/SF does not provide any configuration management capabilities but provides capabilities only for operations management.

OSA/SF includes a Java-based Graphical User Interface (GUI) in support of the client application. The Java GUI is independent of any operating system/server (transparent to operating system), and is expected to operate wherever the Java 1.4 runtimes are available.

Interoperability testing has been performed for Windows 2000, Windows XP, and Linux on System z.

Use of the GUI is optional; a REXX command interface is also included with OSA/SF. OSA/SF has been, and continues to be, integrated in z/OS, z/VM, and z/VSE and runs as a host application. For OSA/SF, Java GUI communication is supported via TCP/IP only.

The Layer 3 OSA Address Table (OAT) displays all IP addresses registered to an OSA port.

OSA/SF has the capability of supporting virtual Medium Access Control (MAC) and Virtual Local Area Network (VLAN) identifications (IDs) associated with OSA-Express2, OSA-Express3, and OSA-Express4S features configured as a Layer 2 interface.

These OSA/SF enhancements are applicable to CHPID type OSD, OSE, and OSN.

For more detailed information on OSA/SF, refer to *zEnterprise, System z10, System z9 and zSeries Open Systems Adapter-Express Customer's Guide and Reference*.

OSA-Express4S features

The OSA-Express4S features are PCIe based cards used only in the PCIe I/O drawers.

Similar to OSA-Express3 features, OSA-Express4S features are designed for use in high-speed enterprise backbones, for local area network connectivity between campuses, to connect server farms to z196, and to consolidate files servers onto z196. The workload can be Internet protocol (IP) based or non-IP based. All OSA-Express4S features are hot-pluggable. Each port can be defined as a spanned channel and can be shared among LPARs within and across LCSS.

OSA-Express4S provides the following enhancements compared to OSA-Express3:

- Port granularity for increased flexibility allowing you to purchase the right number of ports to help satisfy your application requirements and to better optimize for redundancy.
- 8 GBps PCIe interface to the PCIe I/O drawer
- Reduction in CPU utilization by moving the checksum function for LPAR-to-LPAR traffic from the PCIe adapter to the OSA-Express4S hardware.

The OSA-Express4S features includes:

- **OSA-Express4S Gigabit Ethernet (GbE) LX (FC 0404)**

  OSA-Express4S GbE LX has one CHPID per feature and two ports associated with a CHPID. Supports CHPID type: OSD

  OSA-Express4S GbE LX uses a 9 micron single mode fiber optic cable with an LC duplex connector and a link data rate of 1 Gbps. It is designed to support unrepeated distances of up to 5 km (3.1 miles).

  The sending and receiving transceiver must be the same type, LX.
- **OSA-Express4S Gigabit Ethernet (GbE) SX (FC 0405)**

  OSA-Express4S GbE SX has one CHPID per feature and two ports associated with a CHPID. Supports CHPID type: OSD

  OSA-Express4S GbE SX uses a 50 or 62.5 micron multimode fiber optic cable with an LC duplex connector and a link data rate of 1 Gbps. The supported unrepeated distances vary:
  - With 50 micron fiber at 500 MHz-km: 550 meters (1804 feet)
  - With 62.5 micron fiber at 200 MHz-km: 273 meters (902 feet)
  - With 62.5 micron fiber at 160 MHz-km: 220 meters (772 feet)

  The sending and receiving transceiver must be the same type, SX.
- **OSA-Express4S 10 Gigabit Ethernet (GbE) Long Reach LR (FC 0406)**

  OSA-Express4S 10 GbE LR has one CHPID per feature and one port associated with a CHPID. Supports CHPID types: OSD and OSX

  OSA-Express4S 10 GbE LR uses a 9 micron single mode fiber optic cable with an LC duplex connector and a link data rate of 10 Gbps. It is designed to support unrepeated distances of up to 10 km (6.2 miles).

  The sending and receiving transceiver must be the same type, LR.
- **OSA-Express4S 10 Gigabit Ethernet (GbE) Short Reach SR (FC 0407)**

  OSA-Express4S 10 GbE SR has one CHPID per feature and one port associated with a CHPID. Supports CHPID types: OSD and OSX

OSA-Express4S 10 GbE SR uses a 50 or 62.5 micron multimode fiber optic cable with an LC duplex connector and a link data rate of 10 Gbps. The supported unrepeated distances vary:
- With 50 micron fiber at 2000 MHz-km: 300 meters (984 feet)
- With 50 micron fiber at 500 MHz-km: 82 meters (269 feet)
- With 62.5 micron fiber at 200 MHz-km: 33 meters (108 feet)

The sending and receiving transceiver must be the same type, SR.

OSA-Express3 features

All OSA-Express3 features are hot-pluggable.

OSA-Express3 10 Gigabit Ethernet and OSA-Express3 Gigabit Ethernet features can be carried forward or ordered on MES using RPQ 8P2534. During an MES, if slots are available in an I/O cage or I/O drawer and no slots are available in a PCIe I/O drawer, RPQ 8P2534 is used to order these features.

OSA-Express3 features can only be used in an I/O drawer or I/O cage.

The OSA-Express3 features includes:

- **OSA-Express3 Gigabit Ethernet (GbE) LX (FC 3362)**

  OSA-Express3 GbE LX has two CHPIDs per feature and two ports associated with a CHPID. Supports CHPID types: OSD and OSN

  The OSA-Express3 GbE LX uses a 9 micron single mode fiber optic cable with an LC duplex connector and a link data rate of 1000 Mbps (1 Gbps). However, OSA-Express3 GbE LX also accommodates the reuse of existing multimode fiber (50 or 62.5 micron) when used with a pair of mode conditioning patch (MCP) cables. It is designed to support unrepeated distances of up to 5 km (3.1 miles). If using MCP cables, the supported unrepeated distance is 550 meters (1804 feet).
- **OSA-Express3 Gigabit Ethernet (GbE) SX (FC 3363)**

  OSA-Express3 GbE SX has two CHPIDs per feature and two ports associated with a CHPID. Supports CHPID types: OSD and OSN

  The OSA-Express3 GbE SX uses a 50 or 62.5 micron multimode fiber optic cable with an LC duplex connector and a link data rate of 1000 Mbps (1 Gbps). The supported unrepeated distances vary:
  - With 50 micron fiber at 500 MHz-km: 550 meters (1804 feet)
  - With 62.5 micron fiber at 200 MHz-km: 273 meters (902 feet)
  - With 62.5 micron fiber at 160 MHz-km: 220 meters (772 feet)
- **OSA-Express3 1000BASE-T Ethernet (FC 3367)**

  OSA-Express3 1000BASE-T Ethernet has two CHPIDs per feature and two ports associated with a CHPID. Supports CHPID types: OSD, OSE, OSC, OSN, and OSM

  The OSA-Express3 1000BASE-T Ethernet uses a EIA/TIA Category 5 or Category 6 Unshielded Twisted Pair (UTP) cable with an RJ-45 connector and a maximum length of 100 meters (328 feet). It supports a link data rate of 10, 100, or 1000 Mbps; half duplex and full duplex operation modes; and autonegotiations to other speeds.
- **OSA-Express3 10 Gigabit Ethernet (GbE) Long Reach (LR) (FC 3370)**

  OSA-Express3 10 GbE LR has two CHPIDs per feature and one port associated with a CHPID. Supports CHPID types: OSD and OSX

  OSA-Express3 10 GbE LR uses a 9 micron single mode fiber optic cable with an LC duplex connector and a link data rate of 10 Gbps. It is designed to support unrepeated distances of up to 10 km (6.2 miles).

  OSA-Express3 10 GbE LR does not support autonegotiation to any other speed. It supports 64B/66B coding.
- **OSA-Express3 10 Gigabit Ethernet (GbE) Short Reach (SR) (FC 3371)**

  OSA-Express3 10 GbE SR has two CHPIDs per feature and one port associated with a CHPID. Supports CHPID types: OSD and OSX

OSA-Express3 10 Gigabit Ethernet SR uses a 50 or 62.5 micron multimode fiber optic cable with an LC duplex connector and a link data rate of 10 Gbps. The supported unrepeated distances vary:
– With 50 micron fiber at 2000 MHz-km: 300 meters (984 feet)
– With 50 micron fiber at 500 MHz-km: 82 meters (269 feet)
– With 62.5 micron fiber at 200 MHz-km: 33 meters (108 feet)

All OSA-Express3 features support full duplex operation and standard frames (1492 bytes) and jumbo frames (8992 bytes).

OSA-Express2 features

All OSA-Express2 features are hot-pluggable.

OSA-Express2 features can only be carried forward to z196.

OSA-Express2 features can only be used in an I/O drawer or I/O cage.

OSA-Express2 features include:

- **OSA-Express2 Gigabit Ethernet (GbE) LX (FC 3364)**

  OSA-Express2 GbE LX has two CHPIDs per feature and one port associated with a CHPID. Supports CHPID types: OSD and OSN

  The OSA-Express2 GbE LX uses a 9 micron single mode fiber optic cable with an LC duplex connector and a link data rate of 1000 Mbps (1 Gbps). However, OSA-Express2 GbE LX also accommodates the reuse of existing multimode fiber (50 or 62.5 micron) when used with a pair of mode conditioning patch (MCP) cables. It is designed to support unrepeated distances of up to 5 km (3.1 miles). If using MCP cables, the supported unrepeated distance is 550 meters (1804 feet).
- **OSA-Express2 Gigabit Ethernet (GbE) SX (FC 3365)**

  OSA-Express2 GbE SX has two CHPIDs per feature and one port associated with a CHPID. Supports CHPID types: OSD and OSN

  The OSA-Express2 GbE SX uses a 50 or 62.5 micron multimode fiber optic cable with an LC duplex connector and a link data rate of 1000 Mbps (1 Gbps). The supported unrepeated distances vary:
  – With 50 micron fiber at 500 MHz-km: 550 meters (1804 feet)
  – With 62.5 micron fiber at 200 MHz-km: 273 meters (902 feet)
  – With 62.5 micron fiber at 160 MHz-km: 220 meters (772 feet)
- **OSA-Express2 1000BASE-T Ethernet (FC 3366)**

  OSA-Express2 1000BASE-T Ethernet has two CHPIDs per feature and one port associated with a CHPID. Supports CHPID types: OSD, OSE, OSC, and OSN

  The OSA-Express2 1000BASE-T Ethernet uses a EIA/TIA Category 5 Unshielded Twisted Pair (UTP) cable with an RJ-45 connector and a maximum length of 100 meters (328 feet). It supports a link data rate of 10, 100, or 1000 Mbps over a copper infrastructure; half duplex and full duplex operation modes; and autonegotiations to other speeds.

All OSA-Express2 features support full duplex operation and standard frames (1492 bytes) and jumbo frames (8992 bytes).

OSA-Express4S, OSA-Express3 and OSA-Express2 supported functions

Note: Throughout this section, the term “OSA” refers OSA-Express4S, OSA-Express3, and OSA-Express2.

Query and display your OSA-Express4S and OSA-Express3 configuration

OSA-Express4S and OSA-Express3 provides the capability for z/OS to directly query and display your current OSA-Express4S and OSA-Express3 configuration information using the TCP/IP command, **Display OSAINFO**. This command allows the operator to monitor and verify your current

OSA-Express4S and OSA-Express3 configuration, which helps to improve the overall management, serviceability, and usability of OSA-Express4S and OSA-Express3.

This function is supported by z/OS and applies to OSA-Express4S (CHPID types OSD and OSX) and OSA-Express3 (CHPID types OSD, OSX, and OSM).

Optimized latency mode

Optimized latency mode helps to improve the performance of z/OS workloads by minimizing response times for inbound and outbound data when servicing remote clients.

Optimized latency mode applies to OSA-Express4S and OSA-Express3 (CHPID type OSD (QDIO) and CHPID type OSX).

Inbound workload queuing

To improve performance for business critical interactive workloads and reduce contention for resource created by diverse workloads, OSA-Express4S and OSA-Express3 provides an inbound workload queuing function.

The inbound workload queuing (IWQ) function creates multiple input queues and allows OSA-Express4S and OSA-Express3 to differentiate workloads "off the wire" and assign work to a specific input queue (per device) to z/OS. With each input queue representing a unique type of workload and each workload having unique service and processing requirements, the inbound workload queuing function allows z/OS to preassign the appropriate processing resources for each input queue. As a result, multiple concurrent z/OS processing threads can process each unique input queue (workload) avoiding traditional resource contention. In a heavily mixed workload environment, this function reduces the conventional z/OS processing required to identify and separate unique workloads, which results in improved overall system performance and scalability.

The types of z/OS workloads that are identified and assigned to unique input queues are:

- z/OS sysplex distributor traffic – network traffic, which is associated with a distributed virtual internet protocol address (VIPA), is assigned a unique input queue. This allows the sysplex distributor traffic to be immediately distributed to the target host.
- z/OS bulk data traffic – network traffic, which is dynamically associated with a streaming (bulk data) TCP connection, is assigned to a unique input queue. This allows the bulk data processing to be assigned the appropriate resources and isolated from critical interactive workloads.
- z/OS Enterprise Extender traffic – network traffic, which is associated with SNA high performance routing, is assigned a unique input queue. This improves the device and stack processing and avoids injecting latency in the SNA workloads.

The z/OS sysplex distributor traffic and z/OS bulk data traffic workloads are supported by z/OS V1.12 and z/VM 5.4 or later for guest exploitation. The z/OS Enterprise Extender traffic workload is supported by z/OS V1.13 and z/VM 5.4 or later for guest exploitation.

Inbound workload queuing applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX).

Dynamic LAN idle

The OSA LAN idle timer value defines how long OSA will hold packets before presenting the packets to the host. The LAN idle function now allows the host OS to dynamically update the existing LAN idle timer values (defined within the QIB) while the specific QDIO data device is in the QDIO active state.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID type OSD).

Dynamic LAN idle timer function is exploited by z/OS V1.8 with PTFs or later and z/VM 5.4 or later for guest exploitation.

OSA-Express Network Traffic Analyzer

The OSA-Express Network Traffic Analyzer is a diagnostic tool used to copy frames as they enter or leave an OSA adapter for an attached host. This facility is controlled and formatted by the z/OS Communications Server, but the data is collected in the OSA at the network port. Because the data is collected at the Ethernet frame level, you can trace the MAC headers for packets. You can also trace ARP packets, SNA packets, and packets being sent to and from other users sharing the OSA.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID type OSD).

To enable the OSA-Express Network Traffic Analyzer, you must be running with a minimum of z/OS V1.8 with PTFs or later.

Queued Direct I/O Diagnostic Synchronization (QDIOSYNC)

Queued Direct I/O Diagnostic Synchronization provides the ability to coordinate and simultaneously capture software (z/OS) and hardware (OSA) traces. This function allows the host operating system to signal the OSA feature to stop traces and allows the operator to capture both the hardware and software traces at the same time. You can specify an optional filter that alters what type of diagnostic data is collected by the OSA adapter. This filtering reduces the overall amount of diagnostic data collected and therefore decreases the likelihood that pertinent data is lost.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID type OSD).

To use the Queued Direct I/O Diagnostic Synchronization facility, you must be running with a minimum of z/OS V1.8 with PTFs or later.

Dynamic link aggregation for the z/VM environment

This function dedicates an OSA port to the z/VM 5.4 or later operating system for link aggregation under z/VM Virtual Switch-controlled link aggregation. Link aggregation (trunking) is designed to allow you to combine multiple physical OSA ports of the same type into a single logical link. You can have up to eight OSA ports in one virtual switch. This increases bandwidth and permits nondisruptive failover in the event that a port becomes unavailable. This function also supports dynamic add/remove of OSA ports and full-duplex mode (send and receive).

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) in Layer 2 mode in QDIO mode and OSA-Express2 (CHPID type OSD) in Layer 2 mode in QDIO mode.

Multiple Image Facility (MIF) and spanned channels

OSA features support the Multiple Image Facilitly (MIF) for sharing channels across LPARs. Then can be defined as a spanned channel to be shared among LPARs within and across LCSS.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (all CHPID types).

QDIO data connection isolation

QDIO data connection isolation provides protection for workloads (servers and clients) hosted in a virtual environment from intrusion or exposure of data and processes from other workloads.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID type OSD).

Layer 2 (Link Layer) support

OSA features can support two transport modes when using CHPID type OSD (QDIO): Layer 2 (Link Layer) and Layer 3 (Network or IP Layer). Layer 2 support can help facilitate server consolidation and will allow applications that do not use IP protocols to run on z196.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID type OSD).

640 TCP/IP stacks

Increasing the TCP/IP stacks allows you to host more Linux on System z images. OSA supports 640 TCP/IP stacks or connections per dedicated CHPID, or 640 total stacks across multiple LPARs using a shared or spanned CHPID when priority specification is disabled.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID type OSD).

Large send

Large send improves performance by offloading TCP packet processing from the host to the TCP/IP stack. Offloading allows the host to send IP datagrams up to 60K in size. The IP datagram is controlled by the host TCP/IP stack. Sending larger data blocks reduces host processor utilization while increasing network efficiencies.

Large send function of IPv4 packets is available for all in-service releases of z/OS, Linux on System z, and z/VM for guest exploitation. This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID type OSD).

Large send function for IPv6 packets is supported on z/OS. It is not supported for LPAR-to-LPAR packets. Large send support for IPv6 packets applies to the OSA-Express4S features (CHPID types OSD and OSX).

Concurrent LIC update

Allows you to apply LIC updates without requiring a configuration off/on, thereby minimizing the disruption of network traffic during the update.

This function applies to OSA-Express4S (CHPID types OSD and OSX), OSA-Express3 (CHPID types OSD, OSX, OSM, and OSN), and OSA-Express2 (CHPID types OSD and OSN).

Layer 3 virtual MAC

The z/OS Layer 3 Virtual MAC (VMAC) function simplifies the network infrastructure and facilitates IP load balancing when multiple TCP/IP instances are sharing the same OSA port or Media Access Control (MAC) address. With Layer 3 VMAC support, each TCP/IP instance has its own unique "virtual" MAC address instead of sharing the same universal or "burned in" OSA MAC address. Defining a Layer 3 VMAC provides a way for the device to determine which stack, if any, should receive a packet, including those received for IP addresses that are not registered by any TCP/IP stack. With Layer 3 VMAC in a routed network, OSA appears as a dedicated device to the particular TCP/IP stack, which helps solve many port-sharing issues.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID type OSD).

Layer 3 Virtual MAC function is supported by z/OS V1.8 with PTFs or later and z/VM V5.4 or later for guest exploitation.

Jumbo frames

When operating at 1 Gbps (fiber or copper) and 10 Gbps (fiber), use of jumbo frames (8992 bytes) are supported.

This function applies to OSA-Express4S and OSA-Express3 (CHPID types OSD and OSX) and OSA-Express2 (CHPID types OSD).

HiperSockets

HiperSockets "network within the box" functionality allows high speed any-to-any connectivity among OS images within the z196 server without requiring any physical cabling. This "network within the box" concept minimizes network latency and maximizes bandwidth capabilities between z/VM, Linux on System z, z/VSE, and z/OS images (or combinations of these) to enable optimized business and ERP solutions within a single server. These images can be first level (i.e. directly under LPAR), or second level images (i.e. under z/VM). Up to 32 separate internal LANs can be configured within a server thereby allowing OS images to be grouped according to the function they provide. These groupings are independent of sysplex affiliation.

Separate HiperSockets LANs are mainly required if some logical partitions need to be isolated from other logical partitions. Each LAN is configured as an CHPID type IQD.

In addition the number of communication queues is 4096 and each queue can have three subchannels. If you want the internal LANs shared between partitions in different LCSSs then the channel must be spanned. For more information on spanned channels, refer to "Spanned channels" on page 56.

Asynchronous delivery of data

The HiperSockets completion queue function allows both synchronous and asynchronous transfer of data between logical partitions. With the asynchronous support, during high-volume situations, data can be temporarily held until the receiver has buffers available in its inbound queue. This provides end-to-end performance improvement for LPAR to LPAR communications.

The HiperSockets completion queue function is available for HiperSockets on Linux on System z. Refer to *http://www.ibm.com/developerworks/linux/linux390/* for more information on Linux on System z support.

HiperSockets network integration with IEDN

z196 supports the integration of HiperSockets network with the existing intraensemble data network (IEDN). This extends the reach of the HiperSockets network outside the CPC to the entire ensemble, appearing as a single, Layer 2. Because HiperSockets and IEDN are both "internal" System z networks, the combination allows System z virtual servers to use the optimal path for communications.

CHPID type for HiperSockets is IQD. However, IEDN IQD CHPID (and IQDX) is used to refer to the IQD CHPID with the functional support for the IEDN.

Each CPC can have only one IQD CHPID (IEDN IQD CHPID) defined to enable HiperSockets communication to the virtual server.

This support is available for HiperSockets on z/OS.

Broadcast support

Internet Protocol Version 6 (IPv6) broadcast packets are supported over HiperSockets internal LANs. TCP/IP applications that support IPv6 broadcast, such as OMPROUTE when running Routing Information Protocol Version 1 (RIPv1), can send and receive broadcast packets over HiperSockets interfaces.

IPv4 and IPv6 broadcast support is available for HiperSockets on z/OS, z/VM V5.4 or later, and Linux on System z. Refer to *http://www.ibm.com/developerworks/linux/linux390/* for more information on Linux on System z support.

IPv6 support

HiperSockets supports Internet Protocol Version 6 (IPv6). IPv6 expands the IP address space from 32 bits to 128 bits to enable a greater number of unique IP addresses in support of the proliferation of devices, such as cell phones and PDAs, now connecting to the Internet.

IPv4 and IPv6 support is available for HiperSockets on z/OS, z/VM, z/VSE and Linux on System z. IPv6 on z/VSE requires z/VSE 4.2 with PTFs.

VLAN support

Virtual Local Area Networks (VLANs), IEEE standard 802.1q, is supported in HiperSockets in a Linux on System z environment. VLANs increase bandwidth and reduce overhead by allowing networks to be organized for more optimum traffic flow. The network is organized by traffic patterns rather than physical location. This allows traffic to flow on a VLAN connection over HiperSockets and between HiperSockets and OSA.

HiperSockets Network Concentrator

HiperSockets Network Concentrator simplifies network addressing between HiperSockets and OSA allowing seamless integration of HiperSockets-connected operating systems into external networks, without requiring intervening network routing overhead, thus helping to increase performance and simplify configuration.

HiperSockets Network Concentrator is implemented between HiperSockets, OSA, and Linux on System z. The Network Concentrator provides support for unicast, broadcast, and multicast. For more information, refer to *http://www.ibm.com/developerworks/linux/linux390/*.

HiperSockets Network Traffic Analyzer

The HiperSockets Network Traffic Analyzer Trace facility is used to diagnose problems in a HiperSockets network. As data flows over an IQD channel, the HiperSockets Network Traffic Analyzer captures and analyzes each packet. The captured data can be displayed immediately or written to a file.

The captured data includes packets being sent to and from other users sharing the Hipersockets channel, such as logical partitions with z/OS, Linux on System z, z/VSE, or z/VM and z/VM guests.

To use this function, the level of authorization for the HiperSockets network traffic analyzer must be selected. This authorization determines the scope of the tracing. Then a HiperSockets tracing device must be activated on your system. This is performed by the operating system of the owning partition.

Setting the authorization level is performed on the Support Element using the **Network Traffic Analyzer Authorization** task. The levels of authorization are as follows:

- No traffic on any IQD channel for the selected server can be traced
- No traffic on the selected IQD channel can be traced
- All traffic on the selected IQD channel can be traced. (This traces all traffic flowing between all the logical partitions using this IQD CHPID.)
- Customized traffic flow between selected logical partitions can be traced.

  From the Customize a HiperSockets NTA Logical Partition Authorization List window, select the logical partition that will be authorized to set up, trace, and capture the HiperSockets network traffic. Then select all eligible partitions to be traced. Only the traffic flowing between the selected eligible partition or partitions will be traced.

The Support Element issues security logs to create an audit trail of the HiperSockets network traffic analyzer tracing activity.

Layer 2 (Link Layer) support

HiperSockets supports two transport modes on the z196: Layer 2 (Link Layer) and Layer 3 (Network and IP Layer). HiperSockets in Layer 2 mode can be used by Internet Protocol (IP) Version 4 or Version 6 and non-IP protocols (such as AppleTalk, DECnet, IPCX, NetBIOS, or SNA).

Each HiperSockets device has its own Layer 2 MAC address and allows the use of applications that depend on a Layer 2 address such as DHCP servers and firewalls. LAN administrators can configure and maintain the mainframe environment in the same fashion as they do in other environments. This eases server consolidation and simplifies network configuration.

The HiperSockets device performs automatic MAC address generation to create uniqueness within and across logical partitions and servers. MAC addresses can be locally administered, and the use of Group MAC addresses for multicast and broadcasts to all other Layer 2 devices on the same HiperSockets network is supported. Datagrams are only delivered between HiperSockets devices using the same transport mode (Layer 2 with Layer 2 and Layer 3 with Layer 3).

A HiperSockets Layer 2 device may filter inbound datagrams by VLAN identification, the Ethernet destination MAC address, or both. This reduces the amount of inbound traffic, leading to lower CPU utilization by the operating system.

As with Layer 3 functions, HiperSockets Layer 2 devices can be configured as primary or secondary connectors or multicast routers enabling high performance and highly available Link Layer switches between the HiperSockets network and an external Ethernet.

HiperSockets Layer 2 is supported by Linux on System z and by z/VM guest exploitation.

For hardware and sofware requirements, refer to the z/OS, z/VM, z/VSE subsets of the 2817DEVICE Preventive Service Planning (PSP) bucket prior to installing z196.

Multiple Write facility

HiperSockets allows the streaming of bulk data over a HiperSockets link between LPARs. The receiving LPAR can process a much larger amount of data per I/O interrupt. This function is transparent to the operating system in the receiving LPAR. HiperSockets Multiple Write facility, with fewer I/O interrupts, is designed to reduce CPU utilization of the sending and receiving LPAR.

HiperSockets Multiple Write facility is supported in the z/OS environment.

Chapter 6. Sysplex functions

This chapter describes the following z196 sysplex functions:

Parallel Sysplex

IBM Parallel Sysplex makes use of a broad range of hardware and software products to process, in parallel, a transaction processing workload across multiple z/OS images with direct read/write access to sharing data.

The Parallel Sysplex allows you to manage a transaction processing workload, balanced across multiple z/OS images running on multiple Central Processor Complexes (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 Parallel Sysplex using a coupling facility by making it possible to dynamically reapportion the workload across the remaining active Parallel Sysplex participants. Additionally, you can dynamically add processing capacity (CPCs or LPs) during peak processing without disrupting ongoing workload processing.

z196 CPC support for the Parallel Sysplex consists of having the capability to do any or all of the following:
- Configure IC links and define them as CHPID type ICP (peer link - connects to another IC)
- Install ISC-3 links and define them as CHPID type CFP (peer link - connects to another ISC-3)
- Install 12x IFB links (connects zEnterprise to zEnterprise, System z10 or System z9 and connects System z10 to System z10 or System z9) and define them as CHPID type CIB
- Install 1x IFB links (connects zEnterprise to zEnterprise or System z10 and connects System z10 to System z10) and define them as CHPID type CIB
- Define, as an LPAR, a portion or all of the CPC hardware resources (CPs, ICFs, storage, and coupling connections) for use as a coupling facility that connects to z/OS or another CF
- Connect to a coupling facility for data sharing or resource sharing
- Define an Internal Coupling Facility (ICF).

z196 supports a maximum of 128 coupling CHPIDs for all link types (IFBs, ICs, and active ISC-3s per server).

The z196 models provide the following support for the Parallel Sysplex:
- The z196's Parallel Sysplex support consists of supporting coupling facilities on z196, supporting attachment to remote coupling facilities via various type of coupling links, supporting Server Time Protocol (STP) for purposes of time synchronization, and supporting various ancillary CPC functions used by Parallel Sysplex support.
- Internal coupling links can be used to connect either z/OS images to coupling facilities (CFs) or CF images to other CF images within a z196. IC links have the advantage of providing CF communication at memory speed and do not require physical links.

These various interconnect formats provide the connectivity for data sharing between a coupling facility and the CPCs or logical partitions directly attached to it.

Parallel Sysplex coupling link connectivity

z196 supports IC, ISC-3 and IFB for passing information back and forth over high speed links in a Parallel Sysplex environment. These technologies are all members of the family of coupling connectivity options available on z196. With Server Time Protocol (STP), coupling links can also be used to exchange timing information. Refer to "Server Time Protocol (STP)" on page 95 for more information about Server Time Protocol. Refer to Table 17 for a summary of the coupling link options.

Table 17. Coupling link options

Link type	Maximum links	
	M15[1]	M32[2], M49[3], M66[3], M88[3]
1x IFB (HCA3-O LR)	32*	48
12x IFB and 12x IFB3 (HCA3-O)	16*	32
1x IFB (HCA2-O LR) [4]	16*	32
12x IFB (HCA2-O)	16*	32
ISC-3	48	
IC	32	

Note:

1. z196 M15 supports a maximum of 72 extended distance links (24 1x IFB and 48 ISC-3) with no 12x IFB links*.
2. z196 M32 supports a maximum of 96 extended distance links (48 1x IFB and 48 ISC-3) plus four 12x IFB links*.
3. z196 M49, M55, M80 supports a maximum of 104 extended distance links (48 1x IFB and 48 ISC-3) plus 8 12x IFB links*.
4. Carried forward only.

* Uses all available fanout slots. Allows no other I/O or coupling.

Notes:

1. ISC-3 and IFB links require a point-to-point connection (direct channel attach between a CPC and a coupling facility).
2. ISC-3 and IFB links can be redundantly configured (two or more ISC-3 or IFB links from each CPC to enhance availability and avoid extended recovery time.
3. z196 is designed to coexist in the same Parallel Sysplex environment with (n-2) server families. This allows a z196 to coexist with the z10 and z9 servers. Connectivity to z990 or z890 is not supported.

Refer to Figure 14 on page 89 for an illustration of these coupling links.

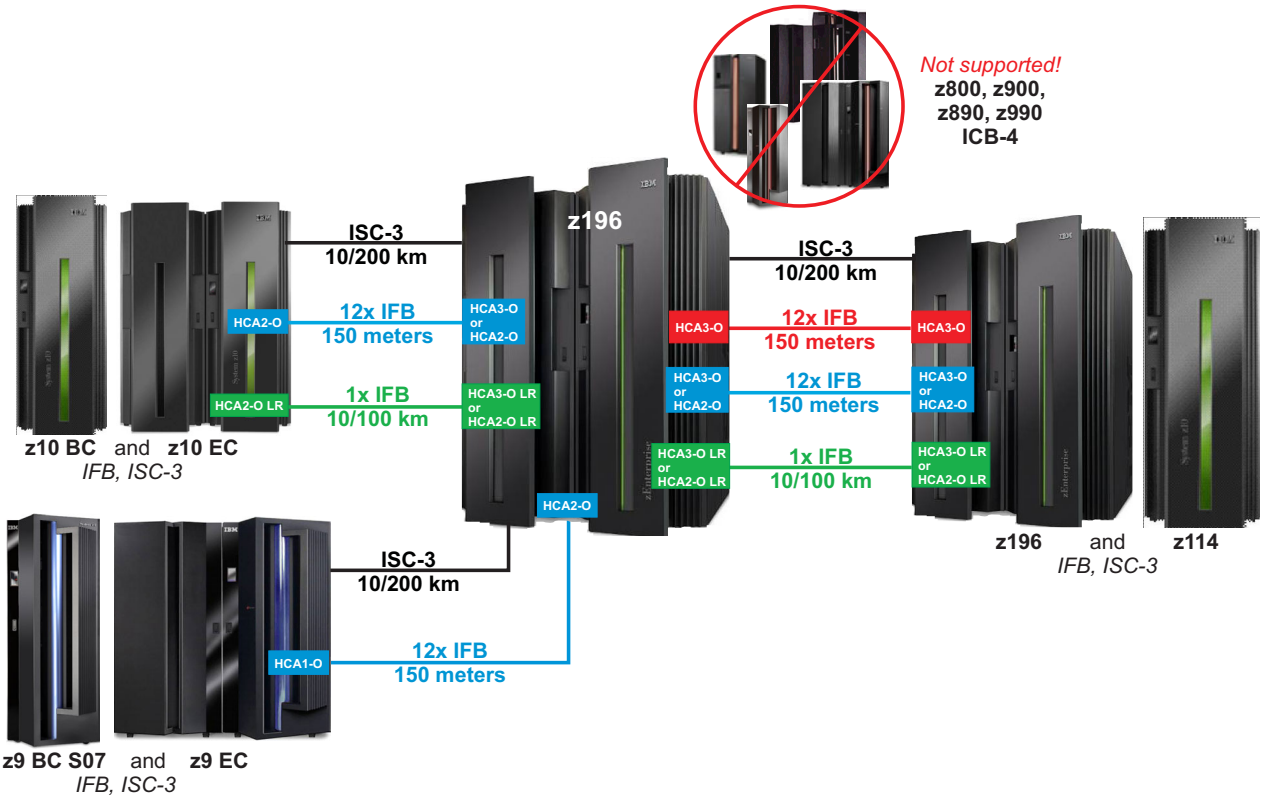


Figure 14. Coupling link connectivity

When coupling within a z196 server, the IC channel can be shared among several LPARs and one coupling facility partition.

ISC-3 links

The ISC-3 feature, with a link data rate of 2 Gbps, is a member of the family of coupling link options available on z196. The ISC-3 feature is used by coupled systems to pass information back and forth over high speed links in a Parallel Sysplex environment. When STP is enabled, ISC-3 links can be used to transmit STP timekeeping information to other z196s as well as z114, z10 EC, z10 BC, z9 EC, and z9 BC servers. ISC-3 links can also be defined as Timing-only links.

ISC-3 links support a maximum unrepeated fiber distance of 10 kilometers (6.2 miles) and a maximum repeated distance of 100 kilometers (62 miles) when attached to a qualified Dense Wavelength Division Multiplexer (DWDM). The list of qualified DWDM vendors is available on Resource Link, (*http://www.ibm.com/servers/resourcelink*), located under the "Hardware products for server" on the **Library** page.) RPQ 8P2197 is required for a maximum unrepeated fiber distance of 20 km. RPQ 8P2340 is required for repeated fiber distances in excess of 100 kilometers.

The z196 ISC-3 feature is compatible with ISC-3 features on z114, z10 EC, z10 BC, z9 EC, and z9 BC servers. ISC-3 (CHPID type CFP) can be defined as a spanned channel and can be shared among LPARS within and across LCSSs. z196 supports 48 ISC-3 links in peer mode – 12 features (four links per feature).

The ISC-3 feature is composed of:
- One Mother card (ISC-M), FC 0217
- Two Daughter cards (ISC-D), FC 0218.

Each daughter card has two ports or links, for a total of four links per feature. Each link is activated by using the Licensed Internal Code Configuration Control (LICCC) and can only be ordered in increments

of one. The ISC-D is not orderable. Extra ISC-M cards can be ordered in increments of one, up to a maximum of 12 or the number of ISC-D cards, which ever is less. When the quantity of ISC links (FC 0219) is selected, the appropriate number of ISC-M and ISC-D cards is selected by the configuration tool. Each port operates at 2 Gbps.

Each port utilizes a Long Wavelength (LX) laser as the optical transceiver, and supports use of a 9/125-micrometer single mode fiber optic cable terminated with an industry standard small form factor LC duplex connector. The ISC-3 feature accommodates reuse (at reduced distances) of 50/125-micrometer multimode fiber optic cables when the link data rate does not exceed 1 Gbps. A pair of mode conditioning patch cables are required, one for each end of the link.

InfiniBand (IFB) coupling links

There are two types of InfiniBand coupling links supported by z196, each supporting a point-to-point topology:

- 12x InfiniBand coupling links
- 1x InfiniBand coupling links

The 12x IFB coupling links are used to connect a zEnterprise to either a zEnterprise or a System z10 with a link data rate of 6 GBps. The 12x IFB coupling links are also used to connect a zEnterprise to a System z9 or to connect a System z10 to either a System z10 or a System z9 with a link data rate of 3 GBps. The 12x IFB coupling links support a maximum link distance of 150 meters (492 feet) – three meters are reserved for intraserver connection.

The 12x IFB coupling links initialize at 3 GBps and auto-negotiate to a higher speed (6 GBps) if both ends of the link support the higher speed. For example, when a zEnterprise is connected to a System z9, the link auto-negotiates to the highest common data rate – 3 GBps. When a zEnterprise is connected to a zEnterprise, the link auto-negotiates to the highest common data rate – 6 GBps.

The 12x IFB coupling links host channel adapter (HCA) fanout cards are as follows:

- HCA3-O fanout card on the zEnterprise
- HCA2-O fanout card on the zEnterprise or System z10
- HCA1-O fanout card on the System z9

12x IFB coupling links support use of a 50 micron OM3 multimode fiber optic cable with MPO connectors. The HCA3-O, HCA2-O, and HCA1-O fanout cards contain two ports. Each port has an optical transmitter and receiver module.

A 12x IFB coupling link using the 12x IFB3 protocol is used to connect a zEnterprise to a zEnterprise when using HCA3-O fanout cards and if four or fewer CHPIDs are defined per HCA3-O port. If more than four CHPIDSs are defined per HCA3-O port, the 12x IFB protocol is used. The 12x IFB3 protocol improves service times.

The 1x IFB coupling links are used to connect a zEnterprise to either a zEnterprise or a System z10 or to connect a System z10 to a System z10 with a link data rate of 5 Gbps. (When attached to a qualified Dense Wavelength Division Multiplexer (DWDM), the link data rate is 2.5 or 5 Gbps). The list of qualified DWDM vendors is available on Resource Link, (*http://www.ibm.com/servers/resourcelink*), located under "Hardware products for server" on the **Library** page.) The 1x IFB coupling links support a maximum unrepeated distance of 10 kilometers (6.2 miles) and the maximum repeated distance is 100 kilometers (62 miles) when attached to a qualified DWDM. RPQ 8P2340 is required for unrepeated fiber distance of excess of 10 kilometers or repeated fiber distances in excess of 100 kilometers.

The 1x IFB coupling links host channel adapter (HCA) fanout cards are as follows:

- HCA3-O LR fanout card on the zEnterprise
- HCA2-O LR fanout card on the zEnterprise or System z10

1x IFB coupling links support use of 9 micron single mode fiber optic cables with LC duplex connectors. The HCA3-O LR fanout card supports four ports, and the HCA2-O LR fanout card supports two ports.

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

When STP is enabled, IFB links can be used to transmit STP timekeeping information to other z196 systems, as well as z114, z10 EC, and z10 BC servers. IFB links can also be defined as Timing-only links.

The CHPID type assigned to InfiniBand is CIB. Up to 16 CHPID type CIB can be defined to an HCA3-O, HCA3-O LR, or HCA2-O fanout card distributed across two ports as needed. Up to 16 CHPID type CIB can be defined to an HCA3-O LR fanout card distributed across four ports as needed. The ability to define up to 16 CHPIDs allows physical coupling links to be shared by multiple sysplexes. For example, one CHPID can be directed to one Coupling Facility, and another CHPID directed to another Coupling Facility on the same target server, using the same port. Note that if more than four CHPIDs are defined per HCA3-O port, the 12x IFB3 protocol will not be used and service times will be reduced.

1x IFB links (both HCA2-O LR and HCA3-O LR fanout cards) support up to 32 subchannels per CHPID. This provides improved link utilization and coupling throughput at increased distances between the coupling facility (CF) and the operating system or between CFs without having to increased the number of CHPIDs per link for 1x IFB or adding ISC-3 links.

IC links

Internal coupling (IC) links are used for internal communication between coupling facilities defined in LPARs and z/OS images on the same server. IC link implementation is totally logical requiring no link hardware. However, a pair of CHPID numbers must be defined in the IOCDS for each IC connection. IC channels cannot be used for coupling connections to images in external systems.

IC links will have CHPID type of ICP (Internal Coupling Peer). The rules that apply to the CHPID type ICP are the same as those that apply to CHPID type CFP (ISC-3 peer links), with the exception that the following functions are not supported:
- Service On/Off
- Reset I/O Interface
- Reset Error Thresholds
- Swap Channel Path
- Channel Diagnostic Monitor
- Repair/Verify (R/V)
- Configuration Manager Vital Product Data (VPD).

IC channels have improved coupling performance over ISC-3 and IFB links. IC links also improve the reliability while reducing coupling cost. Up to 32 IC links can be defined on z196. However, the recommendation is to connect any two images using two IC links, each of which as two CHPIDs type ICP (one at each end), for a total of four CHPIDs type ICP.

Refer to “Internal coupling and HiperSockets channels” on page 56 for recommendations on CHPID usage.

Coupling facility

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

PR/SM LPAR allows you to define the coupling facility, which is a special logical partition that runs Coupling Facility Control Code (CFCC). Coupling Facility Control Code is Licensed Internal Control Code (LICC). It is not an operating system. z196 supports a 64-bit CFCC.

When the CFCC is loaded by using the LPAR coupling facility logical partition activation, the z/Architecture CFCC is always loaded. However, when CFCC is loaded into a coupling facility guest of z/VM, the ESA architecture or z/Architecture CFCC version is loaded based on how that guest is running.

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

CFCC runs in the coupling facility logical partition with minimal operator intervention. Operator activity is confined to the **Operating System Messages** task. PR/SM LPAR limits the hardware operator controls usually available for LPARs to avoid unnecessary operator activity. For more information, refer to *zEnterprise System Processor Resource/Systems Manager Planning Guide*.

A coupling facility link provides the connectivity required for data sharing between the coupling facility and the CPCs directly attached to it. Coupling facility links are point-to-point connections that require a unique link definition at each end of the link.

CFCC considerations

CFCC code can be delivered as a new **release level** or as a **service level** upgrade within a particular release level. Typically, a new release level is delivered as part of an overall system level driver upgrade and requires a reactivate of the CFCC partition in order to utilize the new code. Service level upgrades are delivered as LIC and are generally concurrent to apply.

Note: On rare occasions, we may be required to deliver a disruptive service level upgrade.

To support migration from one CFCC level to the next, you can run different levels of the coupling facility code concurrently in different coupling facility LPARs on the same CPC or on different CPCs. Refer to “CFCC LIC considerations” for a description of how a CFCC release or a service level can be applied.

When migrating CF levels, the lock, list, and cache structure sizes may increase to support new functions. This adjustment can have an impact when the system allocates structures or copies structures from one coupling facility to another at different CFCC levels.

For any CFCC level upgrade, you should always run the CFSIZER tool which takes into account the amount of space needed for the current CFCC levels. The CFSIZER tool is available at *http://www.ibm.com/systems/support/z/cfsizer/*.

CFCC LIC considerations

CFCC LIC can be marked as Concurrent or Disruptive to activate.

CFCC Concurrent LIC maintenance and upgrades can be performed concurrently while the z/OS images connected to it continue to process work and without requiring a POR or a deactivate of the LPAR image of the server on which the coupling facility is located. When applying concurrent CFCC LIC, the code is immediately activated on all of the coupling facility images that are defined on the CPC.

CFCC Disruptive LIC maintenance and new release level upgrades must be applied disruptively. Once the code is installed, the LPAR images on which the coupling facility resides must be deactivated/reactivated requiring z/OS images that are connected to this coupling facility to deallocate CF structures.

The alternative to deallocating CF structures in the CF image being patched would be to move the structures on the coupling facility to a backup coupling facility in the Parallel Sysplex, recycle the coupling facility LPAR image, and move the structures back again once the new code has been activated. This process significantly enhances the overall sysplex availability characteristics of disruptive CFCC LIC.

To support migration of new release or service levels that are marked as disruptive, you have the option to selectively activate the new LIC to one or more coupling facility images running on z196, while still running with the previous level active on other coupling facility images. For example, if you have a coupling facility image that supports a test Parallel Sysplex and a different coupling facility image that supports a production Parallel Sysplex on the same z196, you can install the new LIC to the z196, but may only choose to deactivate/activate the test coupling facility image to utilize and test the new CFCC code. Once you are confident with the new code, you can then selectively deactivate/activate all of the other coupling facility images on the same CPC.

CFCC Level 17

CFCC Level 17 provides the following:

- Increases the number of structures that can be allocated in a CF image from 1023 to 2047. This function permits more discrete data sharing groups to operate concurrently and satisfies the need for environments that require a large number of structures to be defined.
- Supports the ability to capture nondisruptive CFCC diagnostic dumps.
- Supports more connectors to CF list and lock structures.
- Supports CF cache write-around to improve performance. 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.
- Supports CF cache register attachment validation for error detection.
- Supports CF large structure testing

CFCC Level 17 includes the support introduced in previous CFCC levels.

CFCC Level 16

CFCC Level 16 provides the following enhancements:

- Coupling Facility duplexing protocol enhancements provide faster service time when running System-managed CF structure duplexing by allowing one of the duplexing protocol exchanges to complete asynchronously. More benefits are seen as the distance between the CFs becomes larger, such as in a multisite Parallel Sysplex.
- CF subsidiary list notification enhancements provided to avoid false scheduling overhead for Shared Message Queue CF exploiters.

CFCC Level 16 includes the support introduced in previous supported CFCC levels.

CFCC Level 15

CFCC Level 15 provides the following:

- Increase in the allowable tasks in the coupling facility from 48 to 112.
- RMF™ measurement improvements.

CFCC Level 15 includes the support introduced in previous CFCC levels.

CFCC Level 14

CFCC Level 14 provides dispatcher and internal serialization mechanisms enhancements to improve the management of coupled workloads from all environments under certain circumstances.

CFCC Level 14 includes the support introduced in previous CFCC levels.

CFCC Level 13

CFCC level 13 provides Parallel Sysplex availability and performance enhancements. It provides changes that affect different software environments that run within a Parallel Sysplex. For example, DB2 data sharing is expected to see a performance improvement, especially for cast-out processing against very large DB2 group buffer pool structures.

CFCC Level 13 includes the support introduced in previous CFCC levels.

CFCC Level 12

CFCC level 12 provides support for the following functions:

- **64-bit addressing**

  The 64-bit addressing supports larger structure sizes and eliminates the 2 GB "control store" line in the coupling facility. With this support, the distinction between 'control store' and 'non-control store' (data storage) in the coupling facility is eliminated, and large central storage can be used for all coupling facility control and data objects.
- **48 internal tasks**

  Up to 48 internal tasks for improved multiprocessing of coupling facility requests.
- **System-managed CF Structured duplexing (CF duplexing)**

  CF duplexing is designed to provide a hardware assisted, easy-to-exploit mechanism for duplexing CF structure data. This provides a robust recovery mechanism for failures such as loss of single structure or CF, or loss of connectivity to a single CF, through rapid failover to the other structure instance of the duplex pair. Refer to "System-managed CF structure duplexing" on page 97 for more information.

CFCC Level 12 includes the support introduced in previous CFCC levels.

CFCC Level 11

CFCC Level 11 provides support for the following function:

- **System-managed CF structured duplexing (CF duplexing)**

  CF duplexing is designed to provide an S/390® G5/G6 model, hardware assisted, easy-to-exploit mechanism for duplexing CF structure data. This provides a robust recovery mechanism for failures such as loss of single structure or CF, or loss of connectivity to a single CF, through rapid failover to the other structure instance of the duplex pair. Refer to "System-managed CF structure duplexing" on page 97 for more information.

CFCC Level 11 includes the support introduced in previous CFCC levels.

Coupling connection considerations

There are several limits regarding coupling connections to be aware of when ordering and configuring these resources. Refer to Table 17 on page 88 for information on these link limits.

If individual link limits are exceeded, IOCP issues caution messages and HCD issues errors. Refer to *System z Stand-Alone Input/Ouput Configuration Program (IOCP) User's Guide* for details.

I/O configuration considerations

For z196, HCD provides the following function to support coupling facility definition:

- Controls for defining coupling facility channels. HCD also automatically generates the control unit and device definitions associated with CFP, CIB, and ICP channel paths, when they are connected to their respective peer or receiver channel paths. All IC channels paths must be connected.
- Controls for defining a logical partition as either a coupling facility or an operating system logical partition. HCD also allows the definition of the logical partition as both so its usage does not have to be prespecified. This allows the flexibility of usage to be determined at logical partition activation. This

way, if a partition is used one day as a coupling facility and the next day as a z/OS image logical partition, the I/O definitions do not need to change. Additionally, you must use these controls when defining a new logical partition in HCD.

IBM recommends that if you know a logical partition is used exclusively for ESA or exclusively for a coupling facility that you define it that way. This supplies the best HCD rule checking. They also recommend that you use the HCD when possible, to define the coupling facility channel configuration to the channel subsystem.

Server Time Protocol (STP)

Server Time Protocol (STP) (orderable feature code 1021) provides the means for multiple zEnterprise, System z10, and System z9 servers to maintain time synchronization with each other. STP is designed to synchronize servers configured in a Parallel Sysplex or a basic sysplex (without a coupling facility), as well as servers that are not in a sysplex.

Server Time Protocol is a server-wide facility that is implemented in the Licensed Internal Code (LIC) of zEnterprise, z10, and z9 servers and CFs and presents a single view of time to Processor Resource/Systems Manager (PR/SM). STP uses a message-based protocol to transmit timekeeping information over externally defined coupling links between servers. The coupling links used to transport STP messages include ISC-3 links configured in peer mode and IFB links. These links can be the same links already being used in a Parallel Sysplex for coupling facility communications.

By using the same links to exchange timekeeping information and coupling facility messages in a Parallel Sysplex, STP can scale with distance. Servers exchanging messages over short distance links, such as 12x IFB links, are designed to meet more stringent synchronization requirements than servers exchanging messages over long distance links, such as ISC-3 and 1x IFB (distances up to 100 km), where the synchronization requirements are less stringent. If your requirements are to extend the distance to greater than 100 km, submit RPQ 8P2340.

The STP design introduces a concept called Coordinated Timing Network (CTN). A Coordinated Timing Network (CTN) is a collection of servers and coupling facilities that are time synchronized to a time value called Coordinated Server Time. The concept of a Coordinated Timing Network fulfills two key goals:

- Concurrent migration from an existing External Time Reference (ETR) network to a time network using STP. The z196 does not support attachment to a Sysplex Timer, however, it can participate in a Mixed CTN that has either a z10 or z9 synchronized to the Sysplex Timer. This maintains the capability for enterprises to concurrently migrate from an existing ETR network to a Mixed CTN and from a Mixed CTN to an STP-only CTN.
- Capability of servers that cannot support STP to be synchronized in the same network as servers that support STP (z114, z196, z10 EC, z10 BC, z9 EC, and z9 BC).

A CTN can be configured in two ways:

- **Mixed CTN** - Allows the coexistence of servers that are synchronized to a Sysplex Timer in an External Time Reference (ETR) timing network with servers synchronized using the STP message based protocol. In a Mixed CTN, the Sysplex Timer provides the timekeeping for the network. Each server must be configured with the same CTN ID. The Sysplex Timer console is used for time related functions, and the HMC is used to initialize or modify the CTN ID and monitor the CTN status.
- **STP-only CTN** - In an STP-only CTN, the Sysplex Timer does not provide time synchronization for any of the servers in the timing network. Each server must be configured with same CTN ID. The HMC provides the user interface for all time related functions, such as time initialization, time adjustment, and offset adjustment. The HMC or Support Element must also be used to initialize or modify the CTN ID and network configuration.

z196 is designed to coexist in the same CTN with (n-2) server families. This allows a z196 to participate in the same CTN with z10 and z9 servers, but not with z990 or z890 servers.

In an STP-only CTN, you can:

- Initialize the time manually or use an external time source to keep the Coordinated Server Time (CST) synchronized to the time source provided by the external time source (ETS).
- Configure access to an ETS so that the CST can be steered to an external time source. The ETS options are:
  - Dial-out time service (provides accuracy of 100 milliseconds to ETS)
  - NTP server (provides accuracy of 100 milliseconds to ETS)
  - NTP server with pulse per second (PPS) (provides accuracy of 10 microseconds to PPS)
- Initialize the time zone offset, daylight saving time offset, and leap second offset.
- Schedule changes to offsets listed above. STP can automatically schedule daylight saving time based on the selected time zone.
- Adjust time by up to +/- 60 seconds.

As previously stated, STP can be used to provide time synchronization for servers that are not in a sysplex. For a server that is not part of a Parallel Sysplex, but required to be in the same Coordinated Timing Network (CTN), additional coupling links must be configured in order for the server to be configured in the CTN. These coupling links, called Timing-only links, are coupling links that allow two servers to be synchronized using STP messages when a coupling facility does not exist at either end of the coupling link. Use HCD to define Timing-only links and generate an STP control unit.

The benefits of STP include:

- Allowing clock synchronization without requiring the Sysplex Timer and dedicated timer links. This reduces costs by eliminating Sysplex Timer maintenance costs, power costs, space requirements, and fiber optic infrastructure requirements.
- Supporting a multisite timing network of up to 200 km over fiber optic cabling, thus allowing a sysplex to span these distances. This overcomes the limitation of timer-to-timer links being supported only up to 40 km.
- Potentially reducing the cross-site connectivity required for a multisite Parallel Sysplex. Dedicated links are no longer required to transport timing information because STP and coupling facility messages may be transmitted over the same links.

STP enhancements

z196 and z114 have introduced an STP recovery enhancement to improve the availability of the STP-only CTN. The new generation of host channel adapters (HCA3-O or HCA3-O LR), introduced for coupling, have been designed to send a reliable unambiguous “going away signal” to indicate that the server on which the HCA3-O or HCA3-O LR is running is about to enter a failed (check stopped) state. When the “going away signal” sent by the Current Time Server (CTS) in an STP-only Coordinated Timing Network (CTN) is received by the Backup Time Server (BTS), the BTS can safely take over as the CTS without relying on the previous recovery methods of Offline Signal (OLS) in a two-server CTN or the Arbiter in a CTN with three or more servers.

This enhancement is exclusive to z114 and z196 and is available only if you have an HCA3-O or HCA3-O LR on the Current Time Server (CTS) communicating with an HCA3-O or HCA3-O LR on the Backup Time Server (BTS). The STP recovery design that has been available is still available for the cases when a “going away signal” is not received or for other failures besides a server failure.

Some of the other notable STP enhancements released since STP became general availability are:

- Improved availability when an Internal Battery Feature (IBF) is installed. If an Internal Battery Feature (IBF) is installed on your z196, STP can receive notification that power has failed and that the IBF is

engaged. When STP receives this notification from a server that has the role of PTS/CTS, STP can automatically reassign the role of the Current Time Server (CTS) to the Backup Time Server (BTS), thus automating the recovery action and improving availability.

- Save the STP configuration and time information across Power on Resets (POR) or power outages for a single or dual server STP-only CTN. This means you do not need to reinitialize the time or reassign the PTS/CTS role for a single server STP-only CTN or the Preferred Time Server (PTS), Backup Time Server (BTS), or Current Time Server (CTS) roles for a dual server STP-only CTN across Power on Resets (POR) or power outage events.
- Supporting the configuration of different NTP servers for the Preferred Time Server (PTS) and the Backup Time Server (BTS), which improves the availability of NTP servers used as an external time source.
- An Application Programming Interface (API) on the HMC to automate the assignment of the Preferred Time Server (PTS), Backup Time Server (BTS), and Arbiter.

System-managed CF structure duplexing

A set of architectural extensions to the Parallel Sysplex is provided for the support of system-managed coupling facility structure duplexing (CF duplexing) of coupling facility structures for high availability. All three structure types (cache structures, list structures, and locking structures) can be duplexed using this architecture.

Support for these extensions on z196 is concurrent with the entire System z family of servers. It also requires the appropriate level for the exploiter support of CF duplexing. CF duplexing is designed to:

- Provide the necessary base for highly available coupling facility structure data through the redundancy of duplexing
- Enhance Parallel Sysplex ease of use by reducing the complexity of CF structure recovery
- Enable some installations to eliminate the requirement for standalone CFs in their Parallel Sysplex configuration.

For those CF structures that support use of CF duplexing, customers have the ability to dynamically enable (selectively by structure) or disable the use of CF duplexing.

The most visible change for CF duplexing is the requirement to connect coupling facilities to each other with coupling links. The required connectivity is bidirectional with a peer channel attached to each coupling facility for each remote CF connection. A single peer channel provides both the sender and receiver capabilities; therefore, only one physical link is required between each pair of coupling facilities. If redundancy is included for availability, then two peer mode links are required. However, this connectivity requirement does not necessarily imply any requirement for additional physical links. Peer mode channels can be shared between ICF partitions and local z/OS partitions, so existing links between servers can provide the connectivity between both:

- z/OS partitions and coupling facility images
- Coupling facility images.

One of the benefits of CF duplexing is to hide coupling facility failures and structure failures and make total loss of coupling facility connectivity incidents transparent to the exploiters of the coupling facility. This is handled by:

- Shielding the active connectors to the structure from the observed failure condition so that they do not perform unnecessary recovery actions.
- Switching over to the structure instance that did not experience the failure.
- Reestablishing a new duplex copy of the structure at a specified time. This could be as quickly as when the coupling facility becomes available again, on a third coupling facility in the Parallel Sysplex, or when it is convenient for the customer.

System messages are generated as the structure falls back to simplex mode for monitoring and automation purposes. Until a new duplexed structure is established, the structure will operate in a simplex mode and may be recovered through whatever mechanism provided for structure recovery prior to the advent of CF duplexing.

As the two instances of a system-managed duplex structure get update requests, they must coordinate execution of the two commands to ensure that the updates are made consistently to both structures. Most read operations do not need to be duplexed.

z/OS operator commands display the status of the links for problem determination. In addition, the Resource Management Facility (RMF) provides the performance management aspects about the CF-CF connectivity and the duplexed structures. Together, these enable the installation to manage and monitor the coupling facility configuration and new structure instances resulting from CF duplexing.

For more information on CF duplexing, you can refer to the technical white paper, *System-Managed CF Structure Duplexing* at the Parallel Sysplex web site, *http://www.ibm.com/systems/z/pso/*.

GDPS

In business, two important objectives for survival are systems that are designed to provide continuous availability and near transparent disaster recovery (DR). Systems that are designed to deliver continuous availability combine the characteristics of high availability and near continuous operations to deliver high levels of service – targeted at 24 x 7.

To attain high levels of continuous availability (CA) and near transparent disaster recovery (DR), the solution should be based on geographical clusters and data mirroring. These technologies are the backbone of the GDPS solution. GDPS offers the following solutions based on the underlying mirroring technology:

- GDPS/PPRC – based on IBM System Storage® Metro Mirror (formally called Peer-to-Peer Remote Copy, or PPRC) synchronous disk mirroring technology.
- GDPS/PPRC Hyperswap Manager – based on the same disk mirroring technology as GDPS/PPRC.
- GDPS/XRC – based on IBM System Storage z/OS Global Mirror (formally called Extended Remote Copy, or XRC) asynchronous disk mirroring technology.
- GDPS/Global Mirror – based on IBM System Storage Global Mirror technology, which is a disk subsystems based asynchronous form of remote copy.
- GDPS Metro/Global Mirror – a three-site solution that provides continuous availability (CA) across two sites within metropolitan distances and disaster recovery (DR) to a third site at virtually unlimited distances. It is based on a cascading mirroring technology that combines Metro Mirror and Global Mirror.
- GDPS Metro/z/OS Global Mirror – a three-site solution that provides continuous availability (CA) across two sites within metropolitan distances and disaster recovery (DR) to a third site at virtually unlimited distances. It is based on a multitarget mirroring technology that combines Metro Mirror and z/OS Global Mirror.
- GDPS/Active-Active – based on software-based asynchronous mirroring between two active production sysplexes running the same applications with the ability to process workloads in either site.

  Note: The initial offering of GDPS/Active-Active only supports software replication for DB2 and IMS workloads.

GDPS is an integrated, automated application and data availability solution designed to provide the capability to manage the remote copy configuration and storage subsystem(s), automate Parallel Sysplex operational tasks, and perform failure recovery from a single point of control, thereby helping to improve

application availability. GDPS is independent of the transaction manager (e.g., CICS® TS, IMS™, WebSphere) or database manager (e.g., DB2, IMS, and VSAM) being used, and is enabled by means of key IBM technologies and architectures.

For more details on GDPS, refer to the GDPS website located at *http://www.ibm.com/systems/z/advantages/gdps/index.html*.

GDPS/PPRC

GDPS/PPRC is a near CA or DR solution across two sites separated by metropolitan distances. The solution is based on the Metro Mirror (also known as PPRC) synchronous disk mirroring technology. It is designed to manage and protect IT services by handling planned and unplanned exception conditions, and maintain data integrity across multiple volumes and storage subsystems. By managing both planned and unplanned exception conditions, GDPS/PPRC can help to maximize application availability and provide business continuity.

GDPS/PPRC includes automation to manage remote copy pairs, automation to invoke CBU, and automation to restart applications on the recovery site.

GDPS/PPRC can deliver the following capabilities:

- Near continuous availability or disaster recovery solution across two sites separated by metropolitan distances (distance between sites limited to 200 fiber km). Optionally, applications managed end-to-end is provided by the Distributed Cluster Management (DCM) and Tivoli® System Automation Application Manager or the Distributed Cluster Management (DCM) and Veritas Cluster Server (VCS).
- Recovery Time Objective (RTO) less than one hour
- Recovery Point Objective (RPO) of zero

GDPS/PPRC HyperSwap Manager

GDPS/PPRC HyperSwap® Manager is a near CA solution for a single site or entry level DR solution across two sites separated by metropolitan distances, and is based on the same mirroring technology as GDPS/PPRC. It is designed to extend Parallel Sysplex availability to disk subsystems by delivering the HyperSwap capability to mask disk outages caused by planned disk maintenance or unplanned disk failures. It also provides monitoring and management of the data replication environment.

In the multisite environment, GDPS/PPRC HyperSwap Manager provides an entry level disaster recovery offering. Because GDPS/PPRC HyperSwap Manager does not include the systems management and automation capabilities of GDPS/PPRC, the short RTO with GDPS/PPRC is not achievable.

GDPS/PPRC HyperSwap Manager can deliver the following capabilities:

- Near continuous availability solution within a single site
- Disaster recovery solution across two sites seperated by metropolitan distances (distance between sites limited to 200 fiber km)
- RTO of zero (within a single site)
- RPO of zero (within a single site)

GDPS/XRC

GDPS/XRC is a DR solution across two sites separated by virtually unlimited distance between sites. The solution is based on the z/OS Global Mirror (also known as XRC) asynchronous disk mirroring technology. It involves a System Data Mover (SDM) that is found only in z/OS, with supporting code in the primary disk subsystems.

GDPS/XRC includes automation to manage remote copy pairs and automates the process of recovering the production environment with limited manual intervention, including invocation of CBU. This provides significant value in reducing the duration of the recovery window and requiring less operator interaction.

GDPS/XRC can deliver the following capabilities:

- Disaster recovery solution across two sites at virtually unlimited distances
- RTO less than one hour
- RPO less than one minute

GDPS/Global Mirror

GDPS/Global Mirror is a DR solution across two sites separated by virtually unlimited distance between sites. The solution is based on the Global Mirror technology, which is a disk subsystems based asynchronous form of remote copy. Global Mirror enables a two-site disaster recovery solution for z/OS and open systems environments and is designed to maintain a consistent and restartable copy of data at a remote site.

GDPS/Global Mirror can deliver the following capabilities:

- Disaster recovery solution across two sites at virtually unlimited distances
- Continuous availability or disaster recovery solution at unlimited distance using GDPS/Global Mirror Distributed Cluster Management (DCM) and Veritas Cluster Server (VCS)
- RTO less than one hour
- RPO less than one minute

GDPS/Active-Active

GDPS/Active-Active is a solution for an environment consisting of two site or more sites, separated by unlimited distances, running the same applications and having the same data with cross-site workload monitoring and balancing.

The GDPS/Active-Active solution is intended to have multiple configurations. However, at this time, GDPS/Active-Active consists of the Active/Standby configuration. With the Active/Standby configuration, the workload managed by GDPS/Active-Active will be active in one site, receiving transactions routed to it by the workload distribution mechanism. When database updates are made, those changes are asynchronously transmitted from the active instance of the workload to the standby instance of the workload. At the standby site, the standby instance of the workload is active and ready to receive work. The updated data from the active site is then applied to the database subsystem running in the standby site in near real time.

Note: The initial offering of GDPS/Active-Active only supports software replication for DB2 and IMS workloads.

GDPS/Active Standby configuration of the GDPS/Active-Active solution can deliver the following capabilities:

- Near continuous availability or disaster recovery solution across two sites separated by virtually unlimited distance.
- RTO less than one minute
- RPO less than one minute.

For more details on GDPS, refer to the GDPS website located at *http://www.ibm.com/systems/z/advantages/gdps/index.html*.

Intelligent Resource Director (IRD)

Intelligent Resource Director (IRD) is a function that optimizes your workload's resource utilization of the z196 across multiple logical partitions.

It strengthens key z196 and z/Architecture platform technologies, including z/OS Workload Manager, Processor Resource/Systems Manager (PR/SM) (logical partitioning hardware technology) and Parallel Sysplex Clustering technology. This powerful combination provides the ability to dynamically manage workloads within multiple logical operating system images executing on a single z196, as a single large-scale computer resource, with dynamic workload management and physical resource balancing built into the native operating system and underlying hardware.

With IRD, z/OS WLM will exploit Parallel Sysplex technologies to monitor performance of workloads on multiple images against those workload goals. z/OS WLM will then interact with the PR/SM hipervisor, directing PR/SM to dynamically adjust the physical CPU and I/O resource allocation of the hardware across the multiple operating system instances, without requiring Parallel Sysplex data-sharing to achieve these benefits, and totally transparent to customer workload applications.

IRD not only combines PR/SM, z/OS WLM, and Parallel Sysplex for LPAR CPU management, but it also includes Dynamic Channel Path Management (DCM) and I/O (Channel) Subsystem Priority to increase business productivity.

Through IRD technology extensions, the Parallel Sysplex will be able to dynamically change system image weights, reconfigure channels on the fly, and vary logical processors on and offline dynamically to maximize overall throughput across all of the system images to enable the most critical business application of highest priority to get the resources (CPU and I/O) it needs.

LPAR CPU management (clustering)

An LPAR cluster is the subset of the systems in a Parallel Sysplex that are running as logical partitions on the same server.

LPAR CPU management allows dynamic adjustment of processor resources across partitions in the same LPAR cluster. Through the z/OS WLM policy, installations specify the business importance and goals for their workloads. WLM will then manage these sets of logical partitions to provide the processor resources needed for the work to meet its goals based on business importance.

LPAR CPU management requires z/OS WLM goal mode and a coupling facility structure which contains critical status information enabling cross-partition management of CP and I/O resources.

LPAR CPU management can manage Linux on System z on an LPAR running on regular CPs, but not on IFLs.

I/O priority queuing (IOPQ)

I/O subsystem priority queuing extends the classic strengths of I/O priority queuing by addressing other challenges that are not currently handled by existing I/O priority schemes.

For example, prior to I/O subsystem priority queuing, discretionary work in one partition could dominate channels shared with business critical work in another partition. With this new function, z/OS WLM and the Hardware Management Console set priorities that will be used to give the business-critical work higher priority access to the channels. This in turn may allow customers that do not exploit MIF, in order to prevent such problems, to be able to do so now and may lead to reduced overall channel requirements. These new capabilities will help provide optimal workload management.

The range of I/O weights for each logical partition is set within the Hardware Management Console. WLM adjusts the I/O weights within this range. It can be a fixed range, in which WLM would play no part.

Dynamic channel path management (DCM)

This portion of IRD is a combination of hardware strengths and software flexibility. Paths can now be managed between the processor and the control units in the system. Dynamic Channel Path Management (DCM) enables the system to respond to ever changing channel requirements by moving channels from lesser used control units to more heavily used control units as needed. DCM can manage control units connected to ESCON channels.

When used with z/OS Workload Manager (z/OS WLM) in goal mode, z/OS WLM is able to direct Dynamic Channel Path Management to move channels to help business critical work achieve its goals. This also helps reduce the requirement for greater than 256 channels.

I/O priority queuing and Dynamic Channel Path Management (DCM) benefit the Parallel Sysplex environment, with increased benefit in a multiimage environment (Parallel Sysplex). Although Parallel Sysplex data sharing is not required for IRD, the benefits of combining the two are unsurpassed.

I/O priority queueing has no value in a single-system environment.

Table 18. IOPQ in a single-system environment

IRD function	Require CF?	Require goal mode?	Value in single-system cluster
LPAR CPU Mgmt	Yes	Yes	Little (Vary Logical CP)
DCM	Yes	No	Yes
IOPQ	No	No	No
Note: Both DCM and IOPQ do have more value with goal mode.			

Workload manager (WLM)

With the z196, workload manager (WLM) provides industry leading partitioning and workload management. Maximum utilization of all system resources is enabled through dynamic, automatic allocation of processor, memory, and I/O resources across partitions based on real time workload demand and customer policy.

Workload manager on the z196 provides end-to-end management of transactions, from the web-browser to data storage then back to the web-browser. Workload manager can exploit Cisco routers and facilitate dynamic and automatic self-management of data based on business priorities.

Using IBM's discrete server technology with the z196 and z/OS, installations may take advantage of workload based pricing to further reduce the cost of computing as applications continue to grow by using:

- Software pricing based on what you define, not what capacity has been installed.
- Common pricing for many cross-platform products.
- License manager, which simplifies and centralizes via a standard licensing certificate to control software usage billing.

Workload based pricing is adopted by many tools vendors, and provides for 'rigid' management within a flexible system.

EAL5 certification

z196 has achieved a Common Criteria certification at an EAL5+ level for the security of its LPARs that run under the control of the Processor Resource/Systems Manager (PR/SM).

Chapter 7. Cryptography

z196 offers a number of standard and optional hardware-based encryption features. These features include:

- CP Assist for Cryptographic Function (CPACF)
- Configurable Crypto Express3.

CPACF delivers cryptographic support for Data Encryption Standard (DES), Triple Data Encryption Standard (TDES), Advanced Encryption Standard (AES), Secure Hash Algorithm (SHA), and Pseudo Random Number Generation (PRNG).

The Crypto Express3 feature (FC 0864) combines the functions of a coprocessor (for secure key encrypted transactions) and accelerator (for Secure Sockets Layer (SSL) modes into a single feature.

Support for CPACF is also available through the Integrated Cryptographic Service Facility (ICSF). ICSF is a component of z/OS that is designed to transparently use the CPACF and Crypto Express3 functions to balance the workload and satisfy the bandwidth requirements of the applications.

Products that include any of the cryptographic feature codes contain cryptographic functions that are subject to special export licensing requirements by the US Department of Commerce. It is the your responsibility to understand and adhere to these regulations whenever moving, selling, or transferring these products.

The cryptographic features are eligible for export under License Exception ENC as retail items to all end users in all countries except the embargoed, subject to the usual customer screening. The dormant cards themselves, without the enabling software, are also eligible for export an NLR (No License Required) to all customers in all countries except the embargoed, subject to the usual screening.

CP Assist for Cryptographic Function (CPACF)

CPACF is available on z196. The CPACF provides a set of symmetric cryptographic functions that focus on the encryption/decryption function of clear key operations for SSL, Virtual Private Network (VPN), and data storing applications not requiring FIPS 140-2 level 4 security. Each CPACF is shared between two processor units (PUs), which can be designated as various specialty engine types (CPs, IFLs, zIIPs, zAAPs).

The CPACF function is activated using a no-charge enablement feature (FC 3863) and offers the following support on every CFACF:

- Data Encryption Standard (DES)
- Triple data Encryption Standard (TDES)
- Advanced Encryption Standard (AES) for 128-bit, 192-bit, and 256-bit keys
- Secure Hash Algorithms: SHA-1, SHA-224, SHA-256, SHA-384, and SHA-512
- Pseudo Random Number Generation (PRNG).

The DES, TDES, and AES functions use clear key values.

The DES, TDES, AES, and PRNG functions require enablement of the CPACF function (no charge FC 3863) for export control. The CPACF for SHA-1, SHA-224, SHA-256, SHA-384, and SHA-512 are shipped enabled.

Message Security Assist (MSA) instructions for invoking CPACF function for DES, TDES, AES, PRNG, SHA-1, SHA-256, and SHA-512 are found in the *z/Architecture Principles of Operation*.

z196 also supports the following Message Security Assist 4 instructions:

- Cipher Message with CFB (KMF)
- Cipher Message with Counter (KMCTR)
- Cipher Message with OFB (KMO)
- Extension for GHASH function on Compute Intermediate Message Digest (KIMD).

Details on these MSA instruction are also found in the *z/Architecture Principles of Operation*.

Protected key CPACF

When using CPACF for high performance data encryption, CPACF also helps to ensure that key material is not visible to applications or the operating systems. The keys are stored in HSA. This support requires Crypto Express3.

Enablement and disablement of DEA key and AES key functions

Using the **Customize Activation Profile** task on the Support Element, you can enable the encrypt DEA key and encrypt AES key functions of the CPACF to import a clear key, then disable the encrypt DEA key and encrypt AES key functions to protect the CPACF from further imports. The CPACF feature must be installed to use the DEA key and AES key functions on the Support Element.

Crypto Express3

The Crypto Express3 feature (FC 0864) is a tamper-sensing, tamper-responding, programmable cryptographic card. The Crypto Express3 feature is designed to satisfy high-end server security requirements. Each Crypto Express3 feature contains two PCIe adapters. Each PCIe adapter can be configured as a coprocessor or an accelerator.

Crypto Express3 provides a PCI Express (PCIe) interface to the host. Dual processors operate in parallel to support the Common Cryptographic Architecture (CCA) with high reliability.

Although each Crypto Express3 feature occupies an I/O slot and each feature is assigned PCHID values (two PCHIDs for Crypto Express3), they do not use Channel Path Identifiers (CHPIDs). They use cryptographic numbers.

All LPARs can have access to the Crypto Express3 features, if the image activation profile configures the Crypto to the LPAR. Cryptos can be dynamically added, moved, or deleted to or from LPARs without affecting the operating state of the LPAR.

The Crypto Express3 functions include:

- Support for concurrent internal code changes on segment 3 to add/update a CCA application
- Dynamic power management to maximize RSA performance while keeping within temperature limits of the tamper-responding package
- Lock step checking of dual CPUs for enhanced error detection and fault isolation of cryptographic operations
- Updated cryptographic algorithms used in firmware loading with the TKE
- Cryptographic key exchanges between IBM CCA and non-CCA servers
- Secure remote key loading of encryption keys to ATMs, point of sale terminals (POS), and PIN entry devices
- PIN generation, verification, and translation functions
- Elliptic Curve Cryptography (ECC) key generation and key management and digital signature generation and verification
- Keyed-Hash Message Authentication Code (HMAC) for message authentication using secure hash functions with either secure keys or clear text keys

- CCA key token wrapping method to support key bundling requirements for Triple-DES keys while minimizing application program changes
- Secure cryptographic key generation, installation, and distribution using both public and secret key cryptographic methods.
- Consolidation and simplification using a single cryptographic feature
- Public key cryptographic functions
- Hardware acceleration for Secure Sockets Layer (SSL) and Transport Layer Security (TLS) protocols
- User Defined Extension (UDX)
- 13-19 Personal Account Numbers (PANs)
- Secure (encrypted) keys for AES-128, AES-192, and AES-256
- 4096-bit key RSA management capability and clear key RSA acceleration
- For Crypto Express3:
  Maximum number of features per server: 8
  Number of cryptos per feature: 2
  Maximum number of cryptos per server: 16
  Number of domains per crypto: 16
  Number of active LPARs per server: 60

Crypto Express3 can be configured as an accelerator or a CCA coprocessor. CCA is the default configuration.

Accelerator

The **Crypto Express3** accelerator is used for SSL acceleration.

The Crypto Express3 accelerator supports:
- Clear key RSA acceleration.
- Offloading compute-intensive RSA public-key and private-key cryptographic operations employed in the SSL protocol.

CCA coprocessor

The **Crypto Express3** CCA coprocessor is used for secure key encrypted transactions. This is the default configuration.

The Crypto Express3 coprocessor supports:
- AES PIN support for the German banking industry organization, DK
- New Message Authentication Code (MAC) support using the AES algorithm. The Cipher-based MAC (CMAC) is supported.
- User Defined Extension (UDX) simplification for PKA Key Translate
- Export Triple Data Encryption Standard (TDES) key under Advanced Encryption Standard (AES) transport key
- Diversified Key Generation Cipher Block Chaining (CBC) support
- Initial PIN Encrypting KEY (IPEK) support
- Remote Key Export (RKX) key wrapping method support
- Integration of User Defined Extensions (UDX) into CCA
- Highly secure cryptographic functions, use of secure encrypted key values, and user defined extensions (UDX)
- Secure and clear-key RSA operations
- Elliptic Curve Cryptography (ECC) function for clear keys, internal EC keys, and AES Key Encrypting Keys (KEK)
- AES Key Encrypting Keys, supporting exporter, importer, and cipher key types
- Use of decimalization tables in computing PINs
- RSA - Optimal Asymmetric Encryption Padding (OAEP) method with SHA-256

- TR-31 wrapping method for secure key exchange. This method encrypts key material and authenticates the key and attributes
- Derived Unique Key Per Transaction (DUKPT) for Message Authentication Code (MAC) and encryption keys. This crypto function is supported by z/OS and z/VM.
- Secure Cipher Text Translate2 (CTT2) to securely change the encryption key of ciphertext from one key to another key. The decryption of data and reencryption of data happens entirely inside the secure module on the Crypto Express4S feature. This crypto function is supported by z/OS and z/VM.
- Random Number Generation (RNG) in the coprocessor conforms to the Deterministic Random Bit Generator (DRBG) requirements using the SHA-256 based DBRB mechanism. This crypto function is supported by z/OS and z/VM.
- APIs to improve support of EMV (Europay, MasterCard and VISA) card applications that support American Express cards. This crypto function is supported by z/OS and z/VM.
- In order to comply with cryptographic standards, including ANSI X9.24 Part 1 and PCI-HSM, a key must not be wrapped with a key weaker than itself. CCA provides methods for wrapping all keys with sufficient strength. You can configure the coprocessor to ensure your system meets the key wrapping requirements. It can be configured to respond in one of three ways when a key is wrapped with a weaker key: ignore weak wrapping (the default), complete the requested operation but return a warning message, or prohibit weak wrapping altogether. This crypto function is supported by z/OS and z/VM.

User-defined extensions

User-Defined Extensions to the Common Cryptographic Architecture (CCA) are supported on Crypto Express3. For unique customer applications, Crypto Express3 will support the loading of customized cryptographic functions on z196. Support is available through ICSF and the Cryptographic Support for z/OS web deliverable. Under a special contract with IBM, as a Crypto Express3 customer, you will gain the flexibility to define and load custom cryptographic functions yourself. This service offering can be requested by referring to the IBM "Cryptocards" website, then selecting the **Custom Programming** option.

The following is required for UDX support:

- One or more Crypto Express3 features
- A Hardware Management Console
- A TKE workstation, if the UDX requires access control point
- z/VM 5.4 or later for guest exploitation
- z/OS V1R10 or later
- Cryptographic support for z/OS V1.10 or later web deliverable, and PTFs.

If you use a User Defined Extension (UDX) of the Common Cryptographic Architecture (CAA), you should contact your local UDX provider for an application update before ordering a new z196 or before migrating or activating your UDX application. Your UDX application must be migrated to CCA level 4.2.0 or higher before activating it on z196 using Crypto Express3.

See *http://www.ibm.com/security/cryptocards/* for more details.

Trusted Key Entry (TKE)

The Trusted Key Entry (TKE) workstation (FC 0841) and the TKE 7.1 Licensed Internal Code (FC 0867) are optional features that provide a basic key management system for ICSF. TKE includes one Cryptographic coprocessor, which can be logged on with a passphrase or a logon key pair, to be used by security administrators for key identification, exchange, separation, update, backup, and management. Additionally, optional Smart Card Readers can be attached as a secure way to store keys. Up to 10 TKE workstations can be ordered.

TKE 7.1 functions include:

- Elliptic Curve Cryptography (ECC) master key support
- Grouping of domains across one or more host cryptographic coprocessors. This allows you to run domain-scoped commands on every domain in the group using one command or to run host cryptographic adapter scoped commands on every adapter in the group using one command.
- Stronger cryptography encryption for TKE inbound/outbound authentication. This includes:
  - Ability to issue certificates with 2048-bit key strength
  - Encryption of sensitive data sent between the TKE and Crypto Express3 host cryptographic coprocessors using a 256-bit AES key
  - Signing of transmission requests with a 2048-bit signature key, if the host coprocessor is a Crypto Express3 coprocessor
  - Signing of replies sent by a Crypto Express3 coprocessor on the host with a 4096-bit key
- Support for TKE workstation audit records to be sent to a System z host and saved on the host as z/OS System Management Facilities (SMF) records. The TKE workstation audit records are sent to the same TKE host transaction program that is used for TKE operations.
- Support for decimalization tables for each domain on a host cryptographic adapter, used in computing PINs
- Support for AES importer, AES exporter KEK, and cipher operational keys
- Ability for TKE smart card on a TKE workstation with 7.1 code to hold up to 50 key parts
- Support to display the privileged access mode ID and the TKE local cryptographic adapter ID on the TKE console
- Requirement for the TKE local cryptographic adapter profile to have access to each TKE application
- Ability to generate multiple key parts of the same type at one time
- Availability of a master key and operational key loading procedure.

The Trusted Key Entry (TKE) workstation supports four users:

- Auto-logged user, which provides tasks to perform basic operations
- Admin user, which provides setup and configuration tasks
- Auditor user, which provides tasks related to configuring and viewing the audited security events
- Service user, which provides tasks for servicing the TKE workstation.

The orderable TKE features are:

- TKE 7.1 code (FC 0867) **and** TKE Workstation (FC 0841)
- TKE Smart Card Readers (FC 0885)
- TKE Additional Smart Cards (FC 0884)

The TKE workstations require the TKE 7.1 code **and** the TKE unit that contains Ethernet capability and PCIe card. The TKE workstation supports an USB flash memory drive as a removable media device.

Trusted Key Entry (TKE) with Smart Card Readers

Support for an optional Smart Card Readers attached to the TKE 7.1 workstation allows the use of smart cards that contain an embedded microprocessor and associated memory for key storage. Access to and use of confidential data on the smart card is protected by a user-defined Personal Identification Number (PIN).

Wizard for migrating cryptographic configuration data

A wizard on TKE is available to help you migrate Cryptographic configuration data from one Cryptographic coprocessor to a different Cryptographic coprocessor. Using the migration wizard will reduce the number of steps it takes to migrate data, therefore minimizing user errors and decreasing the duration of the migration.

The target Cryptographic coprocessor must have the same or greater capabilities as the Cryptographic coprocessor from which the data is migrating.

RMF monitoring

The Cryptographic Hardware Activity report provides information about the activities in Crypto Express3 features. The request rate (number of requests per second) is reported per adapter. In addition, the utilization (how much of the interval the feature is busy) and the average execution time of all operations is reported.

FIPS certification

The tamper-resistant hardware security module, which is contained within the Crypto Express3 is designed to meet the FIPS 140-2 Level 4 security requirements for hardware security modules.

Remote loading of ATM and POS keys

Remote key loading refers to the process of loading Data Encryption Standard (DES) keys to Automated Teller Machines (ATMs) and Point of Sale (POS) devices from a central administrative site. These enhancements provide two important new features:

- Ability to load initial keys to an ATM or a POS device from a remote location
- Enhanced capabilities for exchanging keys with non-CCA cryptographic systems.

Chapter 8. Cabling

z196 utilizes Small Form Factor (SFF) connectors for ESCON (MT-RJ), FICON, Gigabit Ethernet, 10 Gigabit Ethernet, ISC-3, and 1x InfiniBand. All support LC Duplex connectors. The 12x InfiniBand fanout supports an MPO connector. The speed of the link is determined by the architecture and ranges from 17 MBps (ESCON); 10, 100, or 1000 Mbps (1000BASE-T Ethernet); 1 Gbps (Gigabit Ethernet); 1, 2, or 4 Gbps, as well as 2, 4, or 8 Gbps (FICON); 10 Gbps (10 Gigabit Ethernet); 2.5 or 5 Gbps (1x InfiniBand); to 3 or 6 GBps (12x InfiniBand). Each feature has its own unique requirements, unrepeated distance, and link loss budget.

Fiber optic cables for z196 are available via IBM Site and Facilities Services.

IBM Site and Facilities Services has a comprehensive set of scalable solutions to address IBM cabling requirements, from product-level to enterprise-level for small, medium, and large enterprises. IBM Site and Facilities Services is designed to deliver convenient, packaged services to help reduce the complexity of planning, ordering, and installing fiber optic cables. The appropriate fiber cabling is selected based upon the product requirements and the installed fiber plant.

The services include:
- **IBM Facilities Cabling Services** — fiber transport system
- **IBM IT Facilities Assessment, Design, and Construction Services** — optimized airflow assessment for cabling

These services take into consideration the requirements for all of the protocols and media types supported on the zEnterprise, System z10, and System z9 (for example, ESCON, FICON, coupling links, OSA-Express) whether the focus is the data center, the Storage Area Network (SAN), the Local Area Network (LAN), or the end-to-end enterprise.

Under **IBM Facilities Cabling Services**, there is the option to provide IBM Fiber Transport System (FTS) trunking commodities (fiber optic trunk cables, fiber harnesses, panel-mount boxes) for connecting to the z196, z114, z10 EC, z10 BC, z9 EC, and z9 BC. IBM can reduce the cable clutter and cable bulk under the floor. An analysis of the channel configuration and any existing fiber optic cabling is performed to determine the required FTS trunking commodities. IBM can also help organize the entire enterprise. This option includes enterprise planning, new cables, fiber optic trunking commodities, installation, and documentation.

Under **IBM IT Facilities Assessment, Design, and Construction Services**, the Optimized Airflow Assessment for Cabling option provides you with a comprehensive review of your existing data center cabling infrastructure. This service provides an expert analysis of the overall cabling design required to help improve data center airflow for optimized cooling, and to facilitate operational efficiency through simplified change management.

For additional information on cabling, you can refer to any of the following:
- *zEnterprise 196 Installation Manual for Physical Planning*
- Resource Link (*http://www.ibm.com/servers/resourcelink*), under **Services** from the navigation bar.

Fiber Quick Connect (FQC) for ESCON and FICON LX cabling

Fiber Quick Connect (FQC), an optional feature on z196, is available for all ESCON (62.5 micron multimode fiber) and FICON LX (single-mode fiber) channels. FQC is designed to significantly reduce the amount of time required for on-site installation and setup of fiber optic cabling. FQC eases the addition of, moving of, and changes to ESCON and FICON LX fiber optic cables in the data center, and FQC may reduce fiber connection time by up to 80%.

FQC is for factory installation of IBM Facilities Cabling Services - Fiber Transport System (FTS) fiber harnesses for connection to channels in the I/O drawer. FTS fiber harnesses enable connection to FTS direct-attach fiber trunk cables from IBM Global Technology Services.

FQC, coupled with FTS, is a solution designed to help minimize disruptions and to isolate fiber cabling activities away from the active system as much as possible.

IBM provides the direct-attach trunk cables, patch panels, and Central Patching Location (CPL) hardware, as well as the planning and installation required to complete the total structured connectivity solution. For example, for ESCON, four trunks (each with 72 fiber pairs) can displace up to 240 fiber optic jumper cables, which is the maximum quantity of ESCON channels in one I/O drawer. This significantly reduces fiber optic jumper cable bulk.

On the CPL panels, you can select the connector to best meet your data center requirements. Small form factor connectors are available to help reduce the floor space required for patch panels.

Prior to the server arriving on-site, CPL planning and layout is done using the default CHannel Path IDdentifier (CHPID) report and the documentation showing the CHPID layout and how the direct-attach harnesses are plugged.

Note: FQC supports all of the ESCON channels and all of the FICON LX channels in all of the I/O drawers of the server.

Cabling responsibilities

Fiber optic cables ordering, cable planning, labeling, and placement are the customer responsibilities for new installations and upgrades. Fiber optic conversion kits and Mode Conditioning Patch (MCP) cables are not orderable as features on z196. Representatives will not perform the fiber optic cabling tasks without a service contract.

The following tasks are required to be performed by the customer prior to machine installation:

- All fiber optic cable planning.
- All purchasing of correct, qualified, fiber cables.
- All installation of any required Mode Conditioning Patch (MCP) cables.
- All installation of any required Conversion Kits.
- All routing of fiber optic cables to correct floor cutouts for proper installation to server.
  - Use the Physical Channel Identifier (PCHID) report or the report from the Channel Path Identifier (CHPID) Mapping Tool to accurately route all cables.
- All labeling of fiber optic cables with PCHID numbers for proper installation to server.
  - Use the PCHID report or the report from the CHPID Mapping Tool to accurately label all cables.

Additional service charges may be incurred during the machine installation if the preceding cabling tasks are not accomplished as required.

Cable ordering

Fiber optic cables for the z196 are available from IBM Site and Facilities Services.

The following table lists the channel card feature codes and associated cabling information available on z196.

Table 19. Channel card feature codes and associated connector types and cable types

Feature Code	Feature Name	Connector Type	Cable Type
2324[1]	ESCON channel	MT-RJ	62.5 micron MM
3321 [1]	FICON Express4 10KM LX	LC duplex	9 micron SM
3322	FICON Express4 SX	LC duplex	50, 62.5 micron MM
3324 [1]	FICON Express4 4KM LX	LC duplex	9 micron SM
3325[1]	FICON Express8 10KM LX	LC duplex	9 micron SM
3326	FICON Express8 SX	LC duplex	50, 62.5 micron MM
0409	FICON Express8S 10KM LX	LC duplex	9 micron SM
0410	FICON Express8S SX	LC duplex	50, 62.5 micron MM
0163	HCA2-O	12x MPO	50 micron, OM3 12x IB-DDR
0168	HCA2-O LR	LC duplex	9 micron SM
0170	HCA3-O LR	LC duplex	9 micron SM
0171	HCA3-O	12x MPO	50 micron, OM3 12x IB-DDR
0219	ISC-3	LC duplex	9 micron SM
3364	OSA-Express2 GbE LX	LC duplex	9 micron SM [2]
3365	OSA-Express2 GbE SX	LC duplex	50, 62.5 micron MM
3366	OSA-Express2 1000BASE-T Ethernet	RJ-45	EIA/TIA Category 5 Unshielded Twisted Pair (UTP)
3362	OSA-Express3 GbE LX	LC duplex	9 micron SM [2]
3363	OSA-Express3 GbE SX	LC duplex	50, 62.5 micron MM
3367	OSA-Express3 1000BASE-T Ethernet	RJ-45	EIA/TIA Category 5 Unshielded Twisted Pair (UTP)
3370	OSA-Express3 10 GbE LR	LC duplex	9 micron SM
3371	OSA-Express3 10 GbE SR	LC duplex	50, 62.5 micron MM
0404	OSA-Express4S GbE LX	LC duplex	9 micron SM [2]
0405	OSA-Express4S GbE SX	LC duplex	50, 62.5 micron MM
0406	OSA-Express4S 10 GbE LR	LC duplex	9 micron SM
0407	OSA-Express4S 10 GbE SR	LC duplex	50, 62.5 micron MM

Notes:
1. If this is an initial order and FQC is selected, the ESCON (FC 2324) and FICON (FC 0404, 3321, 3323, 3324, 3325) counts do not apply and are zeroed out.
2. Accommodates reuse of existing multimode fiber (50 or 62.5 micron) when used with a pair of mode conditioning patch (MCP) cables.

Refer to the **Services** section of Resource Link for additional information.

Cabling report

When the Fiber Quick Connect feature is ordered, a second part of the PCHID report is provided to document the connections between the ESCON and FICON LX channels and the MTP couplers. Figure 15 on page 114 shows an example of the cabling portion of the report.

```
---------------------- Fiber Trunking Section ----------------------

Cage   Slot   F/C    Brkt       Type PCHID/Harn.Leg
A01B   3      3325   1F(1.1)R   C    120/1-1 121/1-2 122/1-3 123/1-4
A01B   4      3325   1F(1.2)R   C    130/1-5 131/1-6 132/2-1 133/2-2
.
.
.
A01B   32     2323   1R(10.12)L C    2B0/10-1 2B1/10-2 2B2/10-3
                                      2B3/10-4 2B4/10-5 2B5/10-6
                                      2B6/11-1 2B7/11-2 2B8/11-3
                                      2B9/11-4 2BA/11-5 2BB/11-6
                                      2BC/12-1 2BD/12-2
.
.
.
Z01B   3      3325   1F(3.3)L   C    320/3-1 321/3-2 322/3-3 323/3-4
Z01B   4      3325   1F(3.4)L   C    330/3-5 331/3-6 332/4-1 333/4-2
.
.
.
Z22B   4      3325   1F(1.1)L   D    5A0/1-1 5A1/1-2 5A2/1-3 5A3/1-4
Z22B   5      3325   1F(1.2)L   D    5B0/1-5 5B1/1-6 5B2/2-1 5B3/2-2
Z22B   7      3325   1R(1.1)R   D    5C0/1-1 5C1/1-2 5C2/1-3 5C3/1-4
Z22B   8      3325   1R(1.2)R   D    5D0/1-5 5D1/1-6 5D2/2-1 5D3/2-2
Z22B   10     2323   1R(3.5)R   D    5E0/3-1 5E1/3-2 5E2/3-3 5E3/3-4
                                      5E4/3-5 5E5/3-6 5E6/4-1 5E7/4-2
                                      5E8/4-3 5E9/4-4 5EA/4-5 5EB/4-6
                                      5EC/5-1 5ED/5-2

Legend:
A25B Top of A frame
Z22B Top of Z frame
A01B Bottom of A frame
Z01B Bottom of Z frame
3325 FICON Express8 10KM LX
2323 ESCON Channel 16 Ports
D I/O Drawer
C I/O Cage
```

Figure 15. Cabling section of the PCHID report sample

The columns in this part of the report represent the following data:

Cage Displays the location of the I/O cages, I/O drawers, and PCIe I/O drawers.

Slot Displays the I/O slot where the harness is plugged.

F/C Displays the feature code of the channel card where the harness is plugged.

Brkt Displays the MTP bracket that the harness plugs into (an **F** indicates the front of the frame, an **R** before the value in parenthesis indicates the rear of the frame, an **R** after the value in parenthesis indicates the right of the frame, an **L** indicates the left of the frame).

Type Identifies whether an I/O cage, I/O drawer, or PCIe I/O drawer is installed. (**C** represents an I/O cage. **D** represents an I/O drawer. **S** represents a PCIe I/O drawer.)

PCHID/Harn-Leg
Displays the PCHID number port harness is plugged into, the harness number based on the MTP coupler the harness is plugged to, the harness leg that is plugged into the port.

Chapter 9. Hardware Management Console and Support Element

The z196 includes a Hardware Management Console and two internal Support Elements located on the "Z" frame. The second Support Element, the alternate Support Element, is standard on z196 and is configured the same as, and serves as an alternate to, the primary Support Element.

The Hardware Management Console is configured with a firewall to limit network access in and out. By default, no external connections are allowed through the firewall. As objects are defined to the Hardware Management Console application, the necessary firewall rules are added to allow for communications to and from these objects. Firewall rules are also added to allow external user connections, access by Product Engineering, and the customization of network settings to allow specific applications.

The Hardware Management Console communicates with each CPC through the CPC's Support Element. When tasks are performed at the Hardware Management Console, the commands are sent to one or more Support Elements, which then issue commands to their CPCs. Commands can be sent to as many as all of the CPCs defined to the Hardware Management Console.

On z196 models, CPCs configured to a Hardware Management Console are those CPCs whose internal Support Elements are:

- Attached by local area network (LAN) to the Hardware Management Console
- Defined to have the same domain name and domain password as the Hardware Management Console
- Defined in the defined CPCs group at the Hardware Management Console.

The internal Support Elements for each CPC allows the Hardware Management Console to monitor the CPC by providing status information. Each internal Support Element provides the Hardware Management Console with operator controls for its associated CPC so you can target operations:

- In parallel to multiple or all CPCs
- To a single CPC.

When managing an ensemble, two Hardware Management Consoles are required. The primary Hardware Management Console manages the CPCs (nodes) in an ensemble. (A single CPC, including any optional attached zBX, is called a node.) A primary Hardware Management Console can also manage CPCs that are not members of an ensemble. The alternate Hardware Management Console is used as backup. If the primary fails, the alternate Hardware Management Console takes over has the primary Hardware Management Console. An Hardware Management Console, other than the primary Hardware Management Console or the alternate Hardware Management Console, can also manage CPCs that are in an ensemble. (See Figure 16 on page 116.)

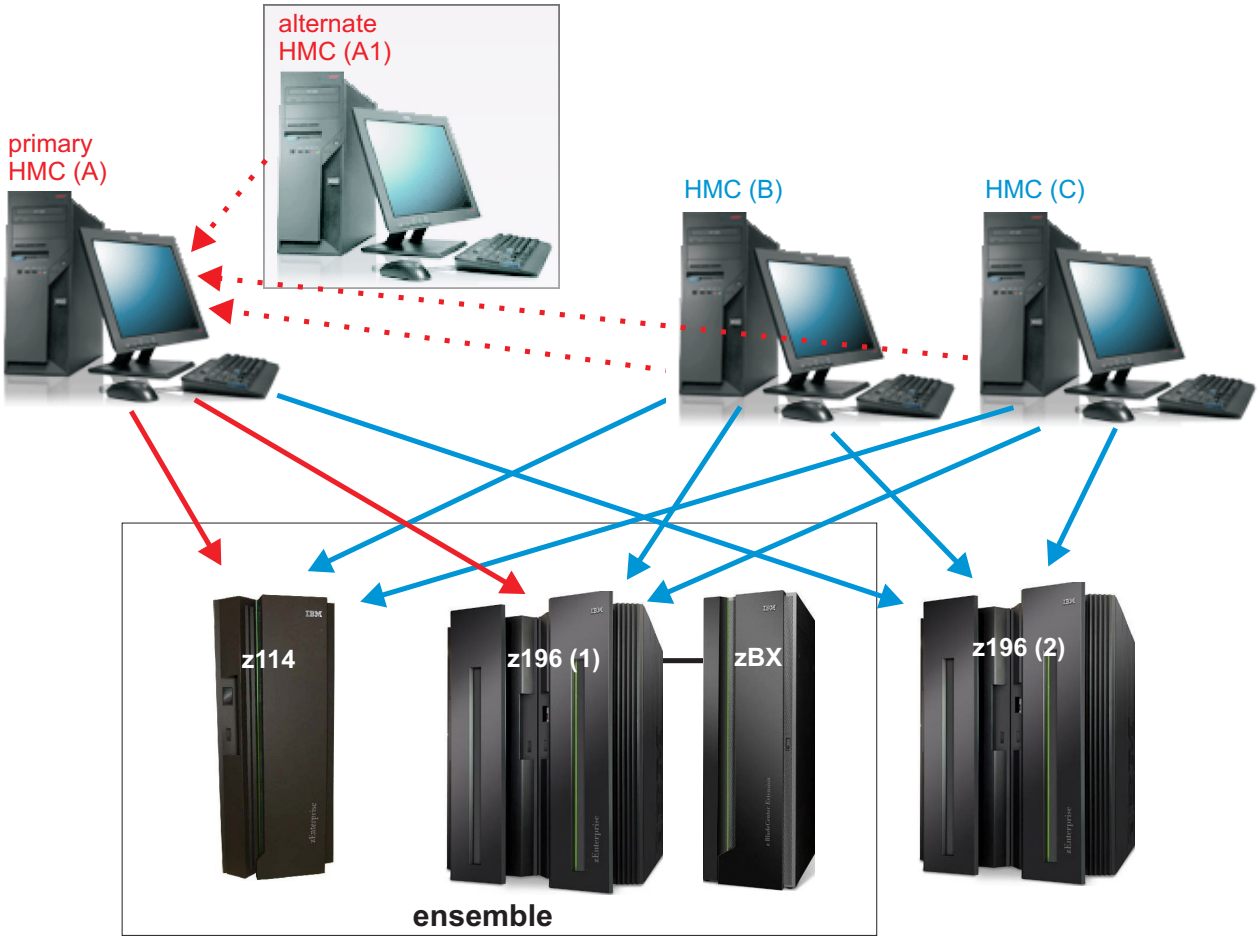


Figure 16. Hardware Management Console configuration

In Figure 16, HMC (A), the primary Hardware Management Console, can directly control the three CPCs (z114 and z196 (1) in the ensemble, and z196 (2) outside of the ensemble). HMC (A), because it is defined as the primary Hardware Management Console, will include ensemble-specific functions, but it can only perform ensemble-specific functions for z114 and z196 (1), not z196 (2).

HMC (A1), the alternate Hardware Management Console, cannot directly control any CPC at this time. If HMC (A) fails, HMC (A1) will become the primary Hardware Management Console and will manage z114, z196 (1), and z196 (2) with the same capabilities as HMC (A).

HMC (B) and HMC (C) can directly control z114, z196 (1), and z196 (2). However, because HMC (B) and HMC (C) are not defined as a primary Hardware Management Console, they cannot perform ensemble-specific functions for z114 and z196 (1).

HMC (A1), HMC (B), and HMC (C) can use the **Remote Hardware Management Console** task, which provides a remote console session to HMC (A), to perform ensemble-specific through HMC (A).

The Hardware Management Console can manage up to 100 CPCs. However, only eight of these CPCs can be a member of an ensemble managed by that Hardware Management Console. The other CPCs can be members of an ensemble managed by other Hardware Management Consoles. A CPC, that is not a member of an ensemble, can be managed by up to 32 Hardware Management Consoles. A single node can be a member of only one ensemble.

Hardware Management Console Application (HWMCA)

The Hardware Management Console Application (HWMCA) is a licensed software application installed on the Hardware Management Console. The application provides an easy-to-use object-oriented Graphical User Interface (GUI) you use to monitor and operate your CPCs. Starting the application makes the user interface available. You can directly manipulate objects displayed in the Hardware Management Console or Support Element workplace using a mouse or key combinations. The application begins whenever the console is powered-on or rebooted. For more detail about the Hardware Management Console and Support Element workplace, refer to the *System z Hardware Management Console Operations Guide* and to the *zEnterprise System Support Element Operations Guide*. You can also take an education course on Resource Link on *How to use the Hardware Management Console*.

Hardware Management Console and Support Element enhancements for z196

This section highlights the significant changes and enhancements for the Hardware Management Console and Support Element Version Code 2.11.1. For more detailed information on these enhancements, see the *System z Hardware Management Console Operations Guide* and *zEnterprise System Support Element Operations Guide*. For more information on the zBX and ensemble Hardware Management Console tasks, see the *zEnterprise System Hardware Management Console Operations Guide for Ensembles*. For more information on the zBX and ensemble Support Element tasks, see the *zEnterprise System Support Element Operations Guide*.

You can use the **What's New** wizard to review the new features available on the Hardware Management Console for each release.

Server/Application State Protocol (SASP) support for load balancing

Workloads Report – Use the **Workloads Report** task to view the Load Balancing Report. The Load Balancing Report lists the load balancing groups and group members to which external routers are distributing incoming work requests. This report also provides details about the recommended weights for each load balancing group member.

Ensemble Details – Use the **Ensemble Details** task to enable SASP support for load balancing and to configure the ensemble for SASP load balancing support. On the **Ensemble Details** window, select the **Performance Management** tab, select **Enable load balancing**, provide a single port number through which external routers establish the SASP connection with the Unified Resource Manager, and provide an IP address of each external router that is allowed to communicate with the Unified Resource Manager. This action is performed by a performance management administrator on the primary HMC.

Displaying and monitoring network resources associated with the IEDN

Use the **Network Monitors Dashboard** task to display the following network metrics and network resources associated with the IEDN:

- Display network metrics for each virtual network interface
- Display network metrics for each physical network interface
- Display VLANs defined in the ensemble
- Display network metrics for each top-of-rack (TOR) or ESM switch and its ports.

You can also use this task to view performance of the IEDN resources to validate the flow of traffic.

Discovery of storage resources

Use the **Discover Storage Resources** action on the **Manage Storage Resources** task to detect the storage resources available to one or more hypervisors.

HMC and Support Element network connection

A local Hardware Management Console must be connected to its Support Elements through a Local Area Network (LAN). z196 models use a dual ethernet (FC 0070) for the LAN wiring between the Hardware Management Console and the Support Elements. The necessary LAN adapters for the Support Elements and the Hardware Management Console may be specified as features on the system order.

The Ethernet switch (FC 0089) can be carried forward on upgrades to z196.

HMC and Support Element features and functions

Customization of the Hardware Management Console or Support Element

You can use the Hardware Management Console workplace or Support Element workplace **User Settings** task to customize the presentation characteristics of the Hardware Management Console or Support Element. These customized settings can be saved to a diskette and used on other Hardware Management Consoles or Support Elements if desired. The **User Settings** task allows you to:

- Modify the default colors or use grey patterns instead of colors.
- Associate a color or pattern with any unacceptable status value you define to distinguish between types of exceptions.
- Change the background color of the Views area used to indicate exception and non-exception situations.
- Modify the default color associated with pending hardware or operating system messages.
- Enter the Hardware Management Console or Support Element TCP/IP address and domain name.

Status reporting

Each internal Support Element monitors the operation of its associated CPC and any CPC images running on the CPC and sends status to the Hardware Management Console for consolidation and exception processing.

Exception processing surfaces only those hardware status items you define as unacceptable to the running of your systems. You can define the conditions for exception processing at the Hardware Management Console or Support Element using the Details panel associated with each managed object.

- In the tree style user interface, the exceptions icon displays in the status bar if any managed object is in an unacceptable state. The exceptions icon also displays in the status column in the work pane next to the managed object that is in an unacceptable state.
- In the classic style user interface, the Hardware Management Console and Support Element display hardware status by using color (or grey patterns) to indicate acceptable or unacceptable status for objects. The default color change is from green (acceptable status) to red (unacceptable status). You can customize these colors (or patterns) using the **Users Setting** task.

  Unacceptable status results in an exception situation that causes the color to change for the:

  – Entire Views Area background.
  – Object background in the Work Area for the object with the unacceptable status.
  – Group object background in the Work Area for any group that the object with the unacceptable status is part of.

  The default color change is from green (acceptable status) to red (unacceptable status). You can customize these colors (or patterns) using the **User Settings** task.

Service Required state

The Service Required state indicates that the spare hardware shipped with the CPC has been depleted. When a part fails causing the use of the last redundant parts of that type, you now have just the required number of parts to keep the CPC running. This message is a reminder to you and the service representative that repairs should be made at the earliest possible time before addition.

The following conditions can cause a Service Required state:
- System is in N Mode Power
- Primary Support Element loss of communications with the alternate Support Element
- Memory sparing threshold is reached
- Oscillator card is defective
- Service network is in N Mode
- Alternate Support Element is fenced
- FSP/STP card is defective
- MCM is defective
- IO Domain is in N Mode
- RAIM memory is degraded.

Degrade indicator

The text “Degraded” indicates that, although the CPC is still operating, some hardware is not working. It displays on an object in the CPC group on the remote console, the Hardware Management Console, and on the Support Elements when:
- Loss of channels due to CPC hardware failure
- Loss of memory
- One or more books are no longer functioning
- The ring connecting the books is open
- Capacity BackUp (CBU) resources have expired
- Processor cycle time reduced due to temperature problem
- CPC was IMLed during cycle time reduction.

Hardware messages

The Hardware Management Console allows you to monitor the hardware messages from any CPC, CPC images, or any group of CPCs or CPC images configured to it. The Support Element allows you to monitor the hardware messages from its CPC or any CPC images configured to it.

Hardware messages present information about problems that are detected, suggest actions where possible, and aid in requesting service when appropriate. When a message has been handled, it is deleted from all the Hardware Management Console(s) and Support Element(s).

When hardware messages are pending for a hardware object or group of hardware objects:
- In the tree style user interface, the hardware message icon displays in the status bar if any managed object received a hardware message. The hardware message icon also displays in the status column in the work pane next to the specific managed object or objects that received a hardware message.
- In classic style user interface, the background of the object and its associated group turns blue (the default) and the Hardware Messages icon turns blue and flashes.

Operating system messages

Local operating systems and Coupling Facility Control Code (CFCC) running in a coupling facility partition can use the console integration facility of the hardware to send operator messages to be displayed by the Hardware Management Console or Support Element. The Hardware Management Console and Support Element allow you to monitor and respond to the operating system messages from any CPC image, coupling facility, or any group of CPC images configured to it.

For a coupling facility partition, Coupling Facility Control Code (CFCC) uses the console integration facility to display coupling facility messages and to accept Coupling Facility Control Code (CFCC) commands. The console integration facility, through the **Operating System Messages** task, provides the only interface for entering commands to an operating coupling facility partition.

When important operating system messages are pending for a hardware object or group of hardware objects:

- In the tree style user interface, the operating system message icon displays in the status bar if any managed object received an operating system message. The operating system message icon also displays in the status column in the work pane next to the specific managed object or objects that received an operating system message.
- In classic style user interface, the background of the object and its associated group turns cyan (the default) and the Operating System Messages icon turns cyan and flashes.

Problem analysis and reporting

Each primary Support Element monitors and analyzes problems detected on the associated CPC. For problems that are isolated to a single CPC, the results are reported to the Hardware Management Console as a hardware message. For those problems that potentially involve multiple CPCs, that problem data is sent to the Hardware Management Console, where data from multiple CPCs is analyzed and reported. The Hardware Management Console configured as a problem analysis focal point can perform:

- Problem analysis for FICON channel link errors of attached Support Elements.
- Problem analysis for ESCON, coupling facility, and Sysplex Timer link faults encountered by the CPCs configured to it.
- Service calls for all CPCs configured to it. Enabling the Hardware Management Console as a call home server identifies the Hardware Management Console as having a modem or LAN/Internet connection that all CPCs configured to it can use for placing service calls.

Enablement and disablement of DEA key and AES key functions

Using the **Customize Activation Profile** task on the Support Element, you can enable the encrypt DEA key and encrypt AES key functions of the CPACF to import a clear key, then disable the encrypt DEA key and encrypt AES key functions to protect the CPACF from further imports. The CPACF feature must be installed to use the DEA key and AES key functions on the Support Element.

Virtual RETAIN

The Virtual RETAIN® function provides a way to capture problem data and place it in a temporary staging area on the Support Element hard disk for a problem that is to be called into IBM service. To ensure security and protection of the scan ring data, any hardware dump collected is encrypted before it is sent to RETAIN.

If RETAIN is not available, a hardware message is displayed for the Hardware Management Console, Support Element, and/or remote console user to instruct the customer to contact IBM Service to gather this staged problem data.

Licensed Internal Code (LIC)

Each Hardware Management Console and each Support Element has Licensed Internal Code (LIC) and is subject to periodic updates from IBM.

On systems with multiple Hardware Management Consoles, one of the Hardware Management Consoles should be configured as a LIC change management focal point. The Hardware Management Console configured can:

- Retrieve and distribute Licensed Internal Code updates for the Hardware Management Consoles remotely from IBM.

- Retrieve and distribute Support Element LIC updates to all the Support Elements of all the CPCs configured to the Hardware Management Console.

Remote I/O configuration and IOCDS management

Each CPC requires a definition of the I/O attached to it. The Hardware Configuration Definition (HCD) is a z/OS application that aids in the definition of all the I/O and aids in the distribution of the appropriate I/O definitions to the appropriate CPCs.

The Hardware Management Console configured as a change management focal point assists HCD in finding the names of all defined CPCs. A single HCD then distributes the appropriate IOCDS and IPL parameters to the various Support Elements of the CPCs defined to the same Hardware Management Console with change management capability.

Scheduled operations

The Hardware Management Console and Support Element provide support for scheduling the times and dates for automatic Licensed Internal Code (LIC) updates and backup of critical hard disk data for the Hardware Management Console, the CPCs configured to the Hardware Management Console, or the Support Element. You can accomplish this by using the **Customize Scheduled Operations** task.

For the Hardware Management Console, the **Customize Scheduled Operations** task, available from the HMC Management work pane (tree style view) or the Console Actions Work Area (classic style view), allows you to schedule the following operations:

- Accept internal code changes
- Audit and log management
- Backup critical hard disk information
- Install concurrent code changes / Activate
- Remove internal code changes / Activate
- Retrieve internal code changes
- Single step code changes retrieve and apply
- Transmit system availability data.

For the CPCs configured to the Hardware Management Console, the **Customize Scheduled Operations** task, available from the Operational Customization task list, allows you to schedule the following operations:

- Accept internal code changes
- Access external time source
- Activate selected CPC
- Audit and log management
- Backup critical hard disk information
- Change LPAR weights
- Deactivate (Power off) selected CPC
- Install concurrent code changes / Activate
- Remove concurrent code changes / Activate
- Retrieve internal code changes
- Single step code changes retrieve and apply
- Transmit system availability data.

For the Support Element, the **Customize Scheduled Operations** task, available from the CPC Operational Customization task list, allows you to schedule the following operations:

- Access external time source
- Accept internal code changes
- Activate or deactivate processor resources in an On/Off CoD record.
- Activate selected CPC
- Audit and log management
- Change LPAR weights

- Deactivate (Power off) selected CPC
- Install concurrent code changes / Activate
- Remove concurrent code changes / Activate
- Retrieve internal code changes
- Transmit system availability data.

Remote Support Facility (RSF)

The Hardware Management Console provides Remote Support Facility (RSF) to aid in the service and maintenance of your system. RSF provides:

- Automatic or customer initiated call for service
- Automatic or customer downloading of the latest LIC change levels
- Automatic downloading of the latest phone list
- Support for records staged by Customer Initiated Upgrade (CIU)
- Support to enable Electronic Service Agent™ (Service Directory) to process operating system I/O error and software inventory data.

Remote Support Facility communicates with the IBM Service Support System using secure TCP/IP protocols. (Both IPv4 and IPv6 protocols are supported.) The communication may be through the enterprise LAN to the Internet (either directly or via a customer supplied SSL Proxy) or through the HMC-provided modem and customer supplied phone line. In both cases, the connection uses high grade SSL encryption, and is an outbound only connection.

Automation and API support

Application Programming Interfaces (APIs) on the Hardware Management Console and Support Element provide an end-user with the ability to view and manipulate managed objects.

The Hardware Management Console supports Common Information Model (CIM), Simple Network Management Protocol (SNMP), and Web Services as systems management APIs.

These APIs contain the ability to get or set the Hardware Management Console or Support Elements managed object's attributes, issue commands to be performed on a managed object from a local or remote application, and receive asynchronous event notifications. These APIs provide a mechanism to IBM, independent system management vendors, and an enterprise, to integrate with the Hardware Management Console Application (HWMCA).

For detailed information on the SNMP APIs, refer to *System z Application Programming Interfaces*. For detailed information on the CIM APIs, refer to *System z Common Information Model (CIM) Management Interface*. For detailed information on the Web Services APIs, refer to *System z Hardware Management Console Web Services API* .

CPC activation

Activating a CPC is an intelligent (LIC controlled) process that takes the CPC from its current state to a fully operational state. The activation may involve a power-on, Power on Reset (POR), and IPL, as necessary.

To activate a CPC, you must specify system activation information for each CPC configured to the Hardware Management Console.

You specify the CPC system activation information using activation profiles. Activation profiles allow you to define multiple power-on reset configurations for the same CPC. These profiles allow you to tailor the CPC resources (central processors, storage, and channels) to meet varying business and workload needs.

You use activation profiles to define PR/SM LPAR mode configurations. Activation profiles are stored on the Support Element hard disk so that they are available for future activate requests.

You can modify default activation profiles that are provided with the CPC and use them as templates to produce customized profiles to fit your processing needs.

There are four types of activation profiles you can produce:
- Reset - Used during power-on reset processing
- Load - Used to load an operating system
- Image - Used to define an logical partition.
- Group - Used to specify the capacity of a group of LPARs.

For PR/SM LPAR mode, you must define a reset profile and one Image profile for each logical partition.

NTP client/server support on the Hardware Management Console

When the Hardware Management Console has the NTP client installed and running, the Hardware Management Console time can be continuously synchronized to an NTP server instead of synchronizing to a Support Element.

Also, when the Hardware Management Console has the NTP client installed and running, the Hardware Management Console can be configured to be used as an NTP server. This provides the ability for the Preferred Timer Server and Backup Time Server in an STP-only CTN to configure the external time source to use NTP with the Hardware Management Console as the NTP server.

z/VM integrated systems management

z/VM integrated systems management from the Hardware Management Console provides out-of-the-box integrated GUI-based basic management of z/VM guests. The Hardware Management Console detects the z/VM images. The z/VM integrated systems management function includes disk, network adaptor and memory management, guest activation and deactivation, and the display of guest status.

From the Hardware Management Console, you can also:
- Dynamically determine if a directory manager is installed. If it is installed, the Hardware Management Console allows any guests to be selected for management, whether it is running or not. It also allows for the defining, altering, and deleting of z/VM guests.

  Enhanced z/VM systems management from the Hardware Management Console allows selected virtual resources to be defined and managed, such as z/VM profiles, z/VM prototypes, z/VM virtual machines, and z/VM volume space.
- View and alter the Virtual Machine Resource Manager (VMRM) configuration and view the current VMRM measurement statistics.

Installation support for z/VM using the Hardware Management Console

z196 allows the installation of Linux on System z in a z/VM 5.4 or later virtual machine using the Hardware Management Console (HMC) DVD drive and the z/VM FTP server. This function does not require an external network connection between z/VM and the Hardware Management Console. Instead, it utilizes the existing internal communication path between the Support Element and the Hardware Management Console, and the content is available via the z/VM FTP server. Installing z/VM from the HMC DVD drive using the legacy support and the z/VM support, z/VM can be installed in an LPAR and both z/VM and Linux on System z installation can be started in a virtual machine from the HMC DVD drive without requiring any external network setup or a connection between an LPAR and the Hardware Management Console.

Network traffic analyzer authorization

z196 allows you to trace OSA-Express network traffic and HiperSockets network traffic. You can use the **Network Traffic Analyzer Authorization** task on the Support Element to display the channels that are currently authorized to trace OSA-Express network traffic and the NTA logical partitions that are currently authorized to trace HiperSockets network traffic. You can also use the Network Traffic Analyzer

Authorization to select or change the level of authorization of the OSA-Express Host Network Traffic Analyzer or the Hipersockets Network Traffic Analyzer.

The **Network Traffic Analyzer Authorization** task is accessible by a user with the role of access administrator (default user ID, ACSADMIN).

User authentication

You can configure the Hardware Management Console to use a LDAP server to perform user ID and password authentication at logon time. The Hardware Management Console still defines the user ID and the roles given to that user ID, but an enterprise directory (LDAP) server is used to authenticate the user. This eliminates the need to store the user ID's password locally.

This function allows the use of the current user ID/password policy for Hardware Management Console user ID/passwords, and provides one centralized user ID/password control mechanism to help meet the user's corporate security guidelines.

Network protocols

The Hardware Management Console for z196 uses a single network protocol, TCP/IP, when communicating with the Support Elements (Support Elements). This network protocol is used for both discovery and normal communications purposes.

The Hardware Management Console supports IPv6 and IPv4 protocols within any customer network (for example, for remote access to the Hardware Management Console user interface or for communication between Hardware Management Consoles and Support Elements). It can also perform electronic remote support requests to IBM service over an IPv6 or IPv4 network.

Customizable console date and time

The **Customize Console Date and Time** task uses more traditional time zone definitions rather than specific offsets from GMT. This allows for the automatic handling of special time zone characteristics such as daylight savings time.

System I/O configuration analyzer (SIOA)

The System I/O configuration analyzer allows the system hardware administrator access to the system's I/O configuration information from one place instead of obtaining it from many separate applications. The analyzer makes it easier to manage I/O configurations, especially across multiple CPCs.

Network analysis tool for Support Element communications

A network analysis tool is available that allows you to verify that all required TCP/IP ports are supported from the Hardware Management Console to the Support Element.

Instant messaging facility

An instant messaging facility is available that allows basic messaging capabilities between users of the Hardware Management Console and the Support Element. It also allows messaging between local users and remote users using existing the Hardware Management Console and Support Element interconnection protocols. The messaging capabilities include:

- Interactive chats between two partners using plain text
- Plain text broadcast message to all sessions on a selected console.

Screen capture function

The Hardware Management Console allows you to capture full screens or specific windows of the Hardware Management Console and save them as PNG, JPG, or GIF files. These files can then be viewed, copied to removable media, or deleted.

Call-home servers selection

You can select which Hardware Management Consoles can be designated as call-home servers for your Hardware Management Console.

User interface

The Hardware Management Console and Support Element allow you to choose the interface style in which you prefer to work:

- Tree style user interface
- Classic style user interface.

The tree style user interface is the default for Operator, Advanced Operator, Access Administrator, and System Programmer user roles. The classic user interface is the default for the Service Representative user role.

The **tree style** interface provides hierarchical views of system resources and tasks using drill-down and launch-in-context techniques to enable direct access to hardware resources and task management capabilities.

The **classic style** interface is the original user interface and has an object-oriented design. You can directly manipulate the objects (such as CPCs) that are defined and be aware of changes to hardware status as they are detected. There are several techniques for manipulating objects and tasks. One way to do this is to left-click an object to select it and double-click the task. An alternate method is the drag and drop technique, which involves using the mouse to pick up one or more objects, dragging them to a task, and then dropping them.

You can change from the tree style interface to the classic style using the **User Settings** task on the Hardware Management Console.

Tree style user interface features

The following items are available when using the tree style user interface:

- **Tasks Index** node is available in the navigation pane. When selected, all tasks and their descriptions are listed in the work pane either in alphabetical order or by task category.
- **Expand all** and **collapse all** icons are available in the navigation pane and the task pad. The expand icon displays all the nodes in the navigation pane or all the tasks under each task group in the task pad. The collapse icon display only the main nodes in the navigation pane or the task groups in the task pad.
- **View** drop-down in the work pane table allows you to create a customized view in the work pane.

Classic style user interface features

The Console Actions Work Area is available is three different layouts:

- Classic - displays the console tasks and the console task groups as they traditional were displayed. This is the default setting.
- List - displays the console tasks in a list. The console task groups are not displayed in this list, but the console tasks within each console task groups are included in the list.
- Groups - displays the console task groups and a few console tasks. The console task groups are:
  - Console Internal Code
  - HMC Configuration and Customization
  - Logs
  - Management
  - Security
  - Service Management Configuration
  - Shutdown or Restart

  The **Logoff** task will display in all the console task groups.

You can choose the layout of the Console Actions Work Area of the classic style user interface using the new Classic Style tab is located on both the User Settings window and the Console Default User Settings window.

Password prompt for disruptive actions

Using the **User Profiles** task, you can control whether you want to be prompted for a password for disruptive actions.

The **User Profiles** task is accessible by a user with the role of access administrator (default user ID, ACSADMIN).

User authority

User authority for the Hardware Management Console is determined by a user role that is associated with tasks. Each user role is actually a collection of authorizations. The Hardware Management Console allows additional user roles to be defined to meet your needs. Refer to the *System z Hardware Management Console Operations Guide* for a list of predefined roles and details on how to define and customer user roles.

Controlling user access to the Hardware Management Console

z196 provides you with the ability to manage the users that have access to the Hardware Management Console without having to define all the user IDs on the Hardware Management Console. Access control is set up using the **User Patterns** task and **User Templates** task.

View only access to selected Hardware Management Console and Support Element tasks

Using the **Customize User Controls** task, you can create user roles having view-only access to a select group of tasks. These view-only tasks include:

- Hardware Messages
- Operating System Messages
- Customize/Delete Activation Profiles
- OSA Advanced Facilities (Hardware Management Console only)
- Advanced Facilities (Support Element only)
- Configure Channel Path On/Off (Hardware Management Console only)
- Configure On/Off (Support Element only).

Removable writable media

The DVD-RAM in the Hardware Management Console has been replaced with a removable writeable media – the USB flash memory drive.

LPAR controls

The following LPAR control functions are available:

- The **Change LPAR Controls** task provides the ability to export LPAR control data to an Excel (.csv) file. This enables you to keep a record of the data for auditing purposes and to perform analysis on the data. You can only export this data when you are connected to the Hardware Management Console remotely through a web browser.
- You can specify a partition capping value for the Change LPAR weights scheduled operation.
- SNMP, CIM, and Web Services APIs can dynamically change LPAR group members and the LPAR group capacity setting.

Auditability support

To provide you the ability to easily view system information for audit purposes, you can offload the following HMC and Support Element log files:

- Audit Log
- Console Event Log
- Console Service History
- Task Performed Log
- Security Log

To generate, view, save, and offload this audit information, the following HMC and Support Element tasks have been added or modified:

- Use the new **Audit & Log Management** task to manually generate, view, save, and offload audit reports. The log files are offloaded as plain files (HTML and XML readable versions).
- Use the **Customize Scheduled Operations** task to schedule when you want to generate, save, and offload audit report.
- Use the **Monitor Events** task to allow for security logs to result in email notifications using the same type of filters and rules that are used for both hardware and operating system messages.
- Use the **Password Profiles** task to allow for the removal of predefined password rules by the access administrator.
- Use the SNMP, CIM, and Web Services APIs to allow user ID audit reports to be generated and retrieved.

You can offload this information to removable media as well as to remote locations via FTP.

Unified Resource Manager

The Unified Resource Manager is a Licensed Internal Code (LIC) that is part of the Hardware Management Console. The Unified Resource Manager performs hardware management and platform management functions for the physical and logical resources of a given ensemble. These functions include:

- Hypervisor management functions:
  - Management of ISO images
  - Creation of virtual networks
  - Management and control of communication between virtual server operating system and the hypervisor.
- Operation controls:
  - Autodiscovery and configuration support for new resources
  - Cross platform hardware problem detection, reporting, and call home
  - Physical hardware configuration, backup, and restore
  - Delivery of system activity using new user.
- Network management functions:
  - Private, secure, and physically isolated data and service networks
  - Management of virtual networks, including access control.
- Energy management functions:
  - Monitoring and trend reporting of CPU energy efficiency
  - Static power savings
  - Ability to query maximum potential power.
- Platform performance management functions:
  - Wizard driven management of resources in accordance with specified business service level objectives
  - A single consolidated and consistent view of resources provides by the Hardware Management Console
- Virtual server lifecycle management and workload context:
  - Single view for virtualization across platforms

- Ability to deploy multiple, cross-platform virtual servers within minutes
- Monitor resource use within the context of a business workload.

The Unified Resource Manager capabilities improve the ability to integrate, monitor, and dynamically manage heterogeneous server resources as a single logical virtualized environment.

Security considerations

Because multiple Hardware Management Consoles and internal Support Elements require connection through a LAN, it is important to understand the use and capabilities enabled for each Hardware Management Console.

Hardware Management Consoles operate as peers with equal access to the CPCs configured to them. The Support Element for each CPC serializes command requests from Hardware Management Console Applications on a first come, first served basis. There is no guarantee of exclusive control across a sequence of commands sent from a single Hardware Management Console.

You should consider these security recommendations:

- Following installation of the CPC(s), Hardware Management Console(s), and Support Element(s) in your configuration, the access administrator should change the default logon passwords at the Hardware Management Console(s) and Support Element(s).
- Create a private LAN to interconnect the Hardware Management Consoles with the controlled Support Elements.

  Using a private LAN for your configuration offers several security, availability, and performance advantages as follows:
  - Direct access to the LAN is limited to the Hardware Management Console(s), Support Element(s), CPC(s), and control unit(s) attached to it. Outsiders cannot connect to it.
  - Traffic disruption due to temporary outages on the LAN is reduced, including disruptions caused by plugging in and powering on new devices on the LAN (minor) to LAN adapters being run at the wrong speed (catastrophic).
  - LAN traffic is minimized reducing the possibility of delays at the Hardware Management Console/Support Element user interface.
- Connect the Hardware Management Consoles to the enterprise LAN using the second LAN adapter in the Hardware Management Console.
- Assign a unique domain name that includes all the CPCs controlled from one or more Hardware Management Consoles.
- Install one or more Hardware Management Consoles that have all of the CPCs you want to control defined to it.

  Place at least one of these Hardware Management Consoles in the machine room near the CPCs that form its domain.

  Use the following enable/disable controls to help you control access and provide focal point capabilities:
  - Licensed Internal Code (LIC) update (change management focal point)
  - Remote service support
  - Remote customer access
  - Remote service access
  - Auto-answer of the modem.
- Physically secure the Hardware Management Console (keep it in a locked room).
- If a remote console is used for remote operations access, assign a secure logon password.
- Log off each Hardware Management Console when it is not in use. The Hardware Management Console provides a status bar capable of displaying status colors (or grey patterns) to alert you when operator activity is needed, even when no one is logged on.

- Establish a limited list of objects and actions available to the operator.

Change management considerations

All Hardware Management Consoles are shipped with change management enabled. If you want to limit the number of Hardware Management Consoles that have change management capability such as LIC update control, I/O definition and remote IOCDS management capability using HCD, enable only those Hardware Management Consoles to be used as change management consoles. A least one Hardware Management Console in the domain must be enabled for change management.

Remote operations and remote access

Remote operations provides the ability to monitor or control a system, or group of systems, from a central or remote location. Remote capability creates a powerful tool for problem determination and diagnosis and operations assistance. Remote operations can save time and money while increasing the productivity of support staff. Technical expertise can be centralized, reducing the need for highly skilled personnel at remote locations.

Remote operations become increasingly important as:

- Data center operations and staff consolidate, with operations centers separate from those data centers
- Companies and their DP staffs merge
- World-wide operations become more common.

When considering remote operation of your z196, there are two options available. You can choose one or both, based on your needs and configuration.

The first set of options deal with manual interaction and provide various methods of allowing a person to interact with the user interface. Manual control allows an operator to monitor and control the hardware components of the system using a hardware management console or a web browser.

A second set of options deal with machine interaction and provide methods of allowing a computer to interact with the consoles through an Application Program Interface (API). These automated interfaces allow a program to monitor and control the hardware components of the system. The automated interfaces are used by various automated products, including those from IBM and other vendors of other System Management products.

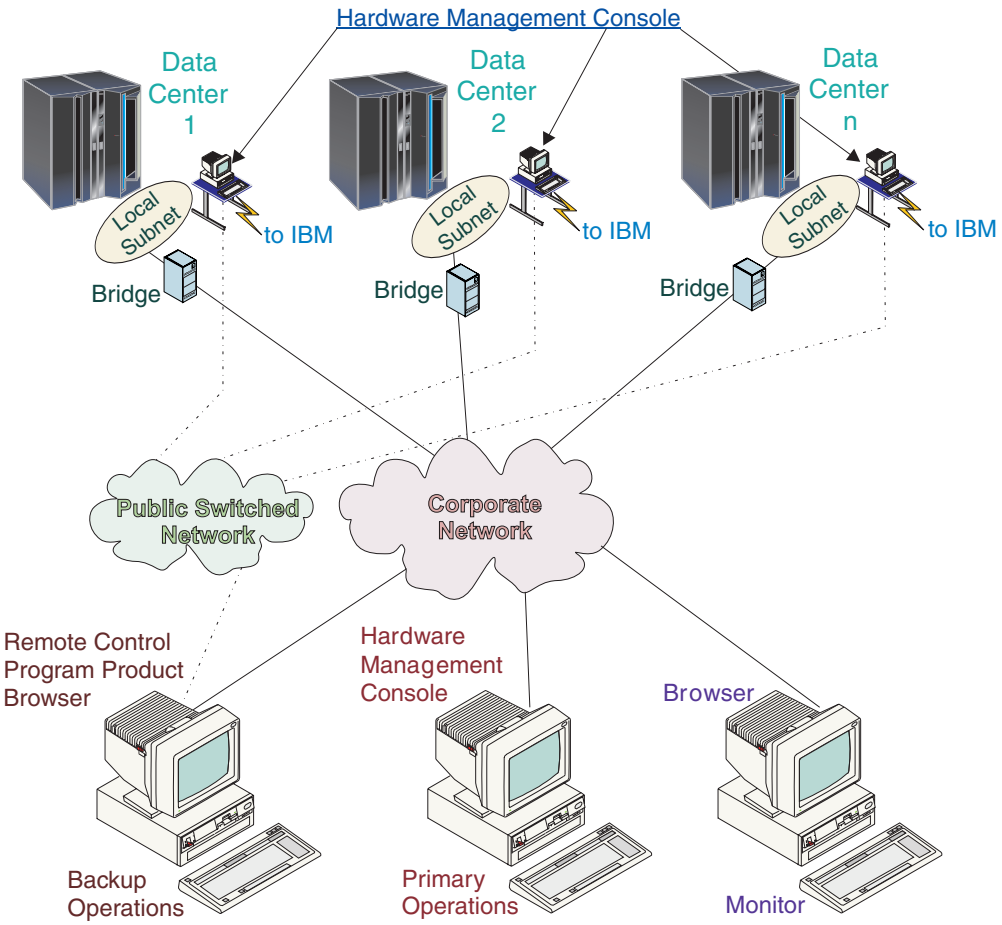


Figure 17. Remote operation example configuration

Remote manual operations

Remote manual operations use the same Graphical User Interface (GUI) used by a local Hardware Management Console operator. There are two ways to perform remote manual operations:

- Using a remote Hardware Management Console
- Using a web browser to connect to a local Hardware Management Console.

The choice between a remote Hardware Management Console and a web browser connected to a local Hardware Management Console is determined by the scope of control needed. A remote Hardware Management Console defines a specific set of managed objects that will be directly controlled by the remote Hardware Management Console, while a web browser to a local Hardware Management Console controls the same set of managed objects as the local Hardware Management Console. An additional consideration is communications connectivity and speed. LAN connectivity provides acceptable communications for either a remote Hardware Management Console or web browser control of a local Hardware Management Console but dialup connectivity is only acceptable for occasional web browser control.

Using a Hardware Management Console

A remote Hardware Management Console gives the most complete set of functions because it is a complete Hardware Management Console – only the connection configuration is different from a local Hardware Management Console. As a complete Hardware Management Console, it requires the same setup and maintenance as other Hardware Management Consoles. A remote Hardware Management

Console needs LAN TCP/IP connectivity to each Support Element to be managed. Therefore, any existing customer installed firewall between the remote Hardware Management Console and its managed objects must permit communications between the Hardware Management Console and Support Element. The remote Hardware Management Console also requires connectivity to IBM or another Hardware Management Console with connectivity to IBM for service and support.

Using a web browser

Each Hardware Management Console contains a web server that can be configured to allow remote access for a specified set of users. When properly configured, an Hardware Management Console can provide a remote user with access to all the functions of a local Hardware Management Console except those that require physical access to a DVD-RAM, CD-ROM, or USB flash memory drive. The user interface on the remote Hardware Management Console is the same as the local Hardware Management Console and has the same constraints as the local Hardware Management Console.

The web browser can be connected to the local Hardware Management Console using either a LAN TCP/IP connection or a switched, dial, network PPP TCP/IP connection. Both types connections can only use encrypted (HTTPS) protocols, as configured in the local Hardware Management Console. If a PPP connection is used, the PPP password must be configured in the local Hardware Management Console and in the remote browser system. Logon security for a web browser is provided by the local Hardware Management Console user logon procedures. Certificates for secure communications are provided, and can be changed by the user.

Browser level is the responsibility of the customer and browser service or support and maintenance does not require connectivity to IBM.

Refer to the *System z Hardware Management Console Operations Guide* for web browser requirements and information on getting ready to configure and use the web server and things to consider during your web session.

Remote automated operations

As an alternative to manual operations, z196 allows a computer to interact with the consoles through a programmable interface, or API. The automated interface allows a program to monitor and control the hardware components of the system in the same way a human can monitor and control the system. The SNMP, CIM, and Web Services APIs provide monitoring and control functions through TCP/IP to an Hardware Management Console. These APIs provide the ability to get and set a managed object's attributes, issue commands, receive asynchronous notifications, and generate traps. For additional information about APIs, see the *System z Application Programming Interfaces*, *System z Common Information Model (CIM) Management Interface*, or the *System z Hardware Management Console Web Services API* document.

The automated interfaces are used by various automation products, including *Tivoli System Automation for z/OS - Processor Operations*.

Chapter 10. Reliability, Availability, and Serviceability (RAS)

z196 reduces downtime by using standard features that provide high levels of Reliability, Availability, and Serviceability (RAS). This section lists the RAS functions available with the z196.

Reliability

The features that provide a high level of reliability include the following:

- High-reliability technology components
- Parts integration to reduce the number of parts in the machine
- To ensure data security, transmission of MCL files, restoration of backup files, and delivery of code loads via AROMs are digitally signed. In addition, any hardware dump collected from the Virtual RETAIN function is encrypted before being transmitted to RETAIN.

Availability

The functions that provide a high level of availability include the following:

Asynchronous delivery of data

The HiperSockets completion queue function allows both synchronous and asynchronous transfer of data between logical partitions. With the asynchronous support, during high-volume situations, data can be temporarily held until the receiver has buffers available in its inbound queue. This provides end-to-end performance improvement for LPAR to LPAR communications.

The HiperSockets completion queue function is available for HiperSockets on Linux on System z.

Alternate HMC preload function

The **Manage Alternate HMC** task allows you to reload internal code changes onto the alternate HMC to minimize HMC downtime during an upgrade to a new firmware level.

Server/Application State Protocol (SASP) support for load balancing

The Unified Resource Manager can provide load balancing recommendations to the configured external routers. These recommendations enable the external routers to distribute incoming work more effectively across the virtual servers in a load balancing group. For each external router, one or more load balancing groups are defined identifying the virtual servers to which the external router can send work requests. In addition, each virtual server participating in load balancing must have a guest platform management provider installed and active. Using the information received from the guest platform management provider, the Unified Resource Manager can associate one group member to a particular virtual server in an ensemble.

You can use the **Workloads Report** task to view the Load Balancing Report. The Load Balancing Report lists the load balancing groups and group members to which external routers are distributing incoming work requests. This report also provides details about the recommended weights for each load balancing group member. If a guest platform management provider is not running or installed on a virtual server, the Load Balancing Report will display "Matching IP address not found" in the **Status** column for the virtual server.

Access to Unified Resource Management capabilities using APIs

You can use application programming interfaces (APIs) to access Unified Resource Manager capabilities for inventory, provisioning, configuration, operational control, monitoring, and workload optimization of

the physical and logical resources of a zEnterprise environment. The capabilities available using the APIs are consistent with the function provided by the Hardware Management Console (HMC) user interface.

Alternate HMC preload function

The **Manage Alternate HMC** task allows you to reload internal code changes onto the alternate HMC to minimize HMC downtime during an upgrade to a new firmware level.

Redundant zBX configurations

zBX provides the following:

- Redundant configuration within a BladeCenter to provide the capacity to concurrently repair the BladeCenter components
- Redundant PDU power connections to the main power source, the management TOR switches, and the BladeCenters
- Redundant management TOR switch connections to the BPH port on z196.

RAIM

For improved availability in the memory subsystem, RAIM technology provides protection at the dynamic random access memory (DRAM), dual inline memory module (DIMM), and memory channel level.

Concurrent book add (CBA)

The concurrent book add function provides the capability of concurrently upgrading the server model by adding the second, third, or fourth processor book, which increases physical processors and memory. The concurrent book add (CBA) function allows new physical hardware to be integrated into the system without affecting ongoing customer workload.

Enhanced book availability

z196 is designed to allow a single book, in a multibook server, to be concurrently removed from the server and reinstalled during an upgrade or repair action. Enhanced book availability or concurrent book replacement is an extension of the support for concurrent book add (CBA).

Redundant I/O interconnect

Redundant I/O interconnect helps maintain critical connections to devices in the event of a HCA fanout card, InfiniBand cable, PCIe cable, or drawer failure by providing connectivity to the server I/O resources using a second path when available.

In the event of an outage, the HCA2-C fanout card and the PCIe fanout card, used for I/O, can be concurrently repaired using redundant I/O interconnect.

Flexible memory

Flexible memory provides the additional resources to maintain a constant level of memory in the event of a single book failure or during an enhanced book availability action. Flexible memory is available on Models M32, M49, M66, and M80 only. Depending on the model, the additional resources are offered in:

- 32 GB increments from 32 GB to 256 GB
- 64 GB increments from 256 GB to 512 GB
- 96 GB increments from 512 GB to 896 GB
- 112 GB increment from 896 GB to 1008 GB
- 128 GB increment from 1008 GB to 1520 GB
- 256 GB increment from 1520 GB to 3056 GB.

Contact your IBM representative to help you determine the appropriate configuration.

Plan ahead features

The Plan Ahead features allow you to order hardware and memory that your current configuration will need in the future. Ordering ahead will avoid a disruptive hardware install in the future. The Plan Ahead features include: Plan Ahead Memory (FC 1996), Plan Ahead for Line Cords feature (FC 2000), Plan Ahead for Balanced Power (FC 3001), and Plan Ahead for I/O Expansion (FC 1999)

Plan Ahead Memory (FC 1996)

The Plan Ahead Memory feature (FC 1996) adds the necessary physical memory required to support target memory sizes. Therefore, it gives you the flexibility to activate memory to any logical size offered between the starting logical memory size and the target logical memory size. You can preplan future upgrades to be nondisruptive.

The Plan Ahead Memory feature is offered in 16 GB increments.

This feature is supported by z/OS and z/VM V5.4 or later.

Plan Ahead for Line Cords (FC 2000)

The Plan Ahead for Line Cords feature (FC 2000) allows you to add a second set of line cords so you can add a book, I/O cages, I/O drawers, or PCIe I/O drawers in the future without having to modify the power infrastructure (circuit breakers, power feeds, line cords, etc.)

This feature is a corequisite for the Plan Ahead of Balanced Power feature (FC 3001).

Plan Ahead for Balanced Power (FC 3001)

The Plan Ahead for Balanced Power feature (FC 3001) allows you to order the maximum number of bulk power regulators (BPRs) on any configuration. This feature helps to ensure that your configuration will be in a balanced power environment if you intend to add processors, I/O cages, or I/O drawers to your server in the future. Regardless of your configuration, all six BPR pairs will be shipped, installed, and activated.

The Plan Ahead for Line Cords feature (FC 2000) is a corequisite. Therefore, if the Plan Ahead for Line Cord feature was not previously selected, it will be added to the order.

Plan Ahead for I/O Expansion (FC 1999)

The Plan Ahead I/O Expansion feature (FC 1999) allows you to select the number of I/O drawers or I/O cages above the minimum number assigned by the configurator. This feature is intended for customers who plan to increase their I/O capability in the future and want to avoid the outage associated with an I/O drawer or I/O cage upgrade.

Additional Memory Capacity Increments (FC 1901)

To activate plan ahead memory, order additional 16 GB increments of memory using FC 1901. For each additional FC 1901 ordered, one feature of plan ahead memory will be activated.

Enhanced driver maintenance

Licensed Internal Code (LIC) updates contribute to downtime during planned outages. The z196 can be configured to permit planned LIC updates to the server at specified driver sync points; that is, points in the maintenance process when LIC may be applied concurrently. A limited number of sync points exist throughout the life of the LIC level and once a sync point has passed, the LIC can no longer be applied concurrently. Availability is improved by taking advantage of these sync points.

Dynamic oscillator card switchover

z196 is designed with two oscillator cards, a primary and a backup. In the event the primary card fails, the backup is designed to detect the failure, switch over, and provide the clock signal to the server transparently.

Program directed re-IPL

Program directed re-IPL is designed to allow Linux on System z to re-IPL without operator intervention. Linux on System z can identify how the IPL was performed from the load device. Program directed re-IPL uses LIC to request a reload, using the same load parameters, from the same load device. Program directed re-IPL allows a Linux on System z running natively in an LPAR to execute a re-IPL.

z/OS V1R10 or later supports the program-directed IPL capability. The z/OS AutoIPL function allows the installation to specify IPL parameters to either IPL Stand-Alone Dump (SADMP), re-IPL z/OS, or both when the z/OS system requires a nonrestartable wait state to be loaded. z/OS also supports specification of IPL volumes and load parameters for IPLs that are to be performed.

Processor unit (PU) sparing

Two spare PUs are provided per server to maintain performance levels should an active CP, ICF, IFL, zAAP, zIIP, or System Assist Processor (SAP) fail on a z196 model.

- Cross-book PU sparing.
  Transparent sparing for failed processors is supported for z196. There are two spare PUs per system and sparing is supported across the books in the unlikely event that the book with the failure has no spares available.
- Transparent CP/ICF/IFL/zAAP/zIIP sparing.
  CP/ICF/IFL/zAAP/zIIP sparing is transparent in all modes of operation and requires no operator or operating system intervention to invoke a spare PU. It is effective on all models including uniprocessor models. With transparent sparing, the application that was running on the failed PU will be preserved and will continue processing on a new PU with no customer intervention required. Refer to "Application preservation" on page 137 for situations where no spare processors are available.
- Dynamic SAP sparing / reassignment.
  Dynamic recovery is provided for failure of the System Assist Processor (SAP). In the event of a SAP failure, if a spare processor unit (PU) is available, in most cases the spare PU will be dynamically activated as a new SAP. If there is no spare PU available, and the CPC has more than one Central Processor (CP), an active CP will be reassigned as a SAP. In either case, there is no customer intervention required. This capability eliminates an unplanned outage and permits a service action, if necessary, to be deferred to a more convenient time.

Processor design

Each processor unit (PU) contains dual instruction and execution units that are used to verify internal operation and that operate simultaneously. Results are compared, and in the event of a miscompare, Instruction Retry is invoked. This design simplifies checking and virtually eliminates PU failures due to soft errors.

Support Elements

z196 has two Support Elements. In the event the primary Support Element fails, switchover to the alternate is usually handled automatically.

Hardware Management Console

One Hardware Management Console is required for system monitoring and operation of the CPC(s) configured to it. For high availability applications, it is recommended that you have at least two

Hardware Management Consoles for your configuration to guarantee that the Hardware Management Console functions are available when needed. Two Hardware Management Consoles are required when your configuration consists of an ensemble.

The Hardware Management Console is concurrently maintainable with the operation of the CPCs configured to it. Having more than one Hardware Management Console provides continuous availability of Hardware Management Console functions, including the following:

- Hardware operator controls, hardware status monitoring, and hardware and operating system messages for all configured CPCs
- Capability to call for service
- Remote operations control
- Problem analysis.

Attaching to IBM service through the Internet

z196 provides the ability to connect to IBM service using the Internet. The SSL connection is made from the HMC through the corporate network and firewall to IBM service using the Internet. This is an outbound connection only.

Hardware Management Console monitor system events

The Hardware Management Console monitor system events is available on z196 models. The Message and State Monitor facilities of the HMC can be enabled to send e-mail notification to a specified recipient whenever a specific message is received from either the hardware subsystem or an operating system, or when a CPC (hardware object) or a CPC image (Operating system object) changes from one "state" to another "state". The state of an object represents its current condition and functional capability, so a state change represents a change in functional capability that may require attention. Hardware and operating system messages are intended to keep the operator informed of conditions that may require attention. However, not all messages and not all state changes are important; only specific ones require attention and notification of a responsible person.

SAPs

The z196 models provide the following base SAPs: Model M15 has three SAPs, Model M32 has six, Model M49 has nine, Model M66 has 12, and Model M80 has 14.

Application preservation

Application preservation is used in the case where a CP fails and there are no spares. The state of the failing CP is passed to another active CP where the operating system uses it to successfully resume the task in most cases without customer intervention.

Dynamic coupling facility dispatching

The dynamic coupling facility (CF) dispatching function helps enable continuous computing in the event of a coupling facility failure without requiring a standalone backup coupling facility. Enhanced dispatching algorithms enable you to define a backup coupling facility in a logical partition on your system. While this logical partition is in backup mode, although it is sharing resources with other logical partitions running other active workload, it uses very little processor resource. When the backup CF becomes active, only the resource necessary to provide coupling is allocated.

Error Correction Code (ECC)

Memory error checking and correction code detects and corrects single bit errors. Also, because of the memory structure design, errors due to a single memory chip failure are corrected.

Dynamic memory sparing

z196 does not contain spare memory DIMMs. Instead it has redundant memory distributed throughout its operational memory and these are used to bypass failing memory. Replacing memory cards require power-off of the drawer, which is disruptive. The extensive use of redundant elements in the operational memory minimizes the possibility of a failure that requires memory card replacement.

Memory scrubbing

Storage background scrubbing provides continuous monitoring of storage for the correction of detected faults before the storage is used.

Fixed HSA

Preplanning requirements are minimized by providing a fixed HSA (16 GB). A fixed HSA allows the maximum configuration capabilities to be exploited.

Dynamic changes to group capacity using an API

You can dynamically change the group capacity value for all logical partitions belonging to a defined group using a SNMP, CIM, or Web Services API.

Dynamic additions to a channel subsystem and LPARs

You can dynamically add LPARs, LCSSs, subchannel sets, and logical CPs to an LPAR without preplanning.

You can dynamically update LPAR image profiles to support Crypto Express3 without an outage to the LPAR. You can also dynamically delete or move Crypto Express3 features from an LPAR.

LPAR dynamic storage reconfiguration

PR/SM LPAR storage reconfigurations can occur allowing nondisruptive add or removal to any partition with a cooperating guest. This capability removes the restriction of storage reconfigurations only being possible from an adjacent and above logical partition.

LPAR dynamic PU reassignment

To better suit the logical partition configurations used on a server, the initial allocation of customer-usable PUs to physical books can change dynamically. Swapping of specialty engines and CPs with each other can occur as the system attempts to “pack” logical partition configurations into physical configurations that span the least number of books.

The effect of this swapping can be observed in dedicated and shared partitions that utilize HiperDispatch.

CICS subsystem storage protect

Subsystem storage protection and subspace group facility support, for use with CICS/ESA, prevents application software from overwriting CICS system software, control blocks, and address spaces.

Partial memory restart

In the event of a memory card failure, the system can be restarted with reduced memory capacity. Processing can be resumed until replacement memory is installed.

Dynamic I/O configuration

Dynamic I/O configuration enhances system availability without requiring a planned outage.

Dynamic I/O configuration allows you to add, delete, or modify the I/O definitions of channel paths, control units, and I/O devices in the CPC. You can also name previously reserved logical partitions and you can save the changes you made to the I/O configuration definitions and apply them to the active I/O Configuration Data Set (IOCDS).

Dynamic I/O configuration requires z/OS or z/VM. Linux on System z, z/VSE, and z/TPF do not provide dynamic I/O configuration support.

When z/VM is controlling the I/O configuration, z/VM's dynamic I/O support is designed to handle all of the elements of the multiple Channel Subsystem facility for dynamic I/O configuration changes. To dynamically change the I/O configuration, one of two techniques can be employed:
- z/VM Control Program (CP) suite of interactive dynamic I/O commands
- HCM/HCD - configuration management tools.

Note: Dynamic I/O configuration is available on a model with only IFLs because z/VM can run on IFLs and perform the function. However, dynamic I/O is not available on a model with only ICFs.

ESCON port sparing

ESCON 16-port I/O card includes one unused port dedicated for sparing in the event of a port failure on that card. Other unused ports are available for growth of ESCON channels without requiring new hardware.

FICON cascaded directors

FICON cascaded directors allow a native FICON (FC) channel or a FICON Channel-to-Channel (CTC) to connect a server to a device or other server with two native FICON directors between them. This is only for a two-switch configuration.

FCP full-fabric connectivity

The FCP full-fabric connectivity supports multiple numbers of directors/switches that can be placed between the server and the FCP/SCSI device, thereby allowing many hops through a storage network for I/O connectivity.

Maintenance/Upgrade for coupling

z196 provides concurrent maintenance for the ISC-3 adapter card. Also, ISC-3 coupling links may now be added concurrently. This eliminates the need for scheduled downtime in the demanding sysplex environment.

Redundant cage controllers

The z196 power and service control network features redundant cage controllers for logic and power control. This design enables nondisruptive service to the controllers and eliminates customer scheduled outage.

Concurrent channel upgrade

It is possible to concurrently add ESCON, FICON, ISC-3, and OSA channels to an I/O cage or I/O drawer provided there are unused channel positions in the I/O cage or I/O drawer. In addition, IFBs and their associated cables, can be added provided there are unused cable jack positions. This capability may help eliminate an outage to upgrade the channel configuration.

Redundant power feeds

The power system offers a redundant primary (AC) power supplies. These redundant power supplies are electrically isolated and each have their own line cord(s), allowing the system to survive the loss of customer power to either line cord(s). If power is interrupted to one of the power supplies, the other

power supply will pick up the entire load and the system will continue to operate without interruption. Therefore, the line cord(s) for each supply must be wired to support the entire power load of the system.

Refer to *zEnterprise 196 Installation Manual for Physical Planning* for more details about power feeds.

Redundant power and thermal subsystems

The DC power and thermal subsystems are designed with N +1 redundancy. Failure of a power or thermal component does not cause a system outage.

Dynamic FSP/STP card switchover

z196 is designed with two FSP/STP cards. An FSP/STP card failure will automatically switch to the other FSP/STP card. If the External Time Source (ETS) option, NTP server with PPS, is used, it is recommended that the PPS port on each FSP/STP card be attached to the PPS output of a different NTP server to provide effective resiliency for ETS failures.

Preferred Time Server and Backup Time Server

In an STP-only CTN configuration, it is required that at least one server is defined as the Preferred Time Server. It is also required that a Current Time Server is assigned. The Current Time Server (the Stratum 1 server) has the highest level of hierarchy in the STP-only CTN and has connectivity to the servers designated as Stratum 2 servers. If only a Preferred Time Server is defined in an STP-only CTN, it is assigned as the Current Time Server.

If there is a failure in the Preferred Time Server, synchronization fails if a backup configuration is not established. Therefore, it is highly recommended that a Backup Time Server is defined in an STP-only CTN. The Backup Time Server is normally a Stratum 2 server that has connectivity to the Preferred Time Server, as well as to all other Stratum 2 servers that are connected to the Preferred Time Server. By providing this connectivity, the Backup Server can take over as the Current Time Server if there is a failure with the Preferred Time Server or if a reconfiguration is planned. Therefore, the servers in the STP-only CTN can maintain synchronization.

Additionally, when the external time source for the STP-only CTN is configured to use NTP (with or without PPS), having the ETS configured on the Backup Time Server using different NTP server(s) provides continuous availability of NTP servers. In the event that the Preferred Time Server cannot access its configured NTP server(s), adjustments can be made using information from the Backup Time Server. This is achieved without reconfiguring the Backup Time Server as the Current Time Server.

Concurrent hardware maintenance

Concurrent maintenance enables the replacement of failed units concurrently with system operation. This enhances the processor availability by eliminating the need for system outage to effect the repair. Concurrent maintenance capability exists for the following elements:

- Power
- Thermal
- ESCON cards
- FICON Express8S cards
- FICON Express8 cards
- FICON Express4 cards
- OSA-Express4S feature cards
- OSA-Express3 feature cards
- OSA-Express2 feature cards
- Crypto Express3 features
- ISC-3 feature card
- HCA2-O HCA2-O LR, HCA2-C, HCA3-O, HCA2-O LR, and PCIe fanout cards
- OSC cards
- FSP/STP cards

- Hardware Management Console
- Support Element
- I/O drawers
- PCIe I/O drawer.

Concurrent Licensed Internal Code (LIC) patch

Concurrent LIC patch allows the activation of a patch concurrent with system operation thereby increasing the availability of the processor by reducing scheduled outage for LIC maintenance. This capability exists for code for the following elements:

- CP
- SAP
- Cage controllers
- LP
- CFCC
- Power
- Thermal
- ESCON channels
- FICON channels
- FCP channels
- OSA channels
- ISC-3 links
- IFB links
- IC links
- HiperSockets
- Hardware Management Console
- Support Element.
- PU core engineering data
- BladeCenter components
- IBM Smart Analytics Optimizer
- IBM POWER7 blades
- DataPower XI50z
- IBM System x blades.

Notes:

1. OSA-Express channels always require CHPID vary off/vary on cycle to activate LIC patches.
2. Concurrent patch support is available for OSA-Express2 features configured for OSD.

Concurrent internal code change

With Crypto Express3, you can dynamically update segment 3 (add or modify Common Cryptographic Architecture (CCA) applications) without having to configure the crypto off/on, therefore eliminating any disruptions to the entire system.

Electronic Service Agent (Service Director)

Electronic Service Agent (Service Director™) will have I/O error data collected by a component integral to the operating system, forwarded from the operating system through a Hardware Management Console, and then to an eService server in IBM. Electronic Service Agent provides the analysis of the data and provides various users access to the data through a web browser interface.

Internal Battery Feature (IBF)

The Internal Battery Feature (IBF) provides backup input power. The feature is packaged internal to the machine. It can be used with a UPS to provide additional protection.

Redundant coupling links

Redundant coupling links (ISC-3s from different ISC-M cards and IFBs from different HCA fanout cards) can be configured between a processor and the coupling facility. This potentially removes a single point of failure for the processor's data sharing capability in the Parallel Sysplex environment.

Large page support

Large (1 megabyte (MB)) page support provides performance value primarily to long running memory access intensive applications.

Customer Initiated Upgrade (CIU)

Customer Initiated Upgrade (CIU) allows you to permanently increase processor or memory capacity. You can request these orders through the web using IBM Resource Link.

You can perform permanent upgrades while temporary capacity is active. This allows for quick conversion of temporary capacity to permanent capacity.

Capacity Upgrade on Demand (CUoD)

Capacity Upgrade on Demand provides the capability to permanently add CPs, ICFs, IFLs, zAAPs, zIIPs, SAPs, memory, and channels nondisruptively, eliminating the need for a scheduled outage. Installations who take advantage of the CUoD option may invoke the additional capacity nondisruptively.

On/Off Capacity on Demand (On/Off CoD)

When your business needs short term additional capacity, On/Off Capacity on Demand (On/Off CoD) is designed to deliver it. On/Off CoD is designed to temporarily turn on CPs, IFLs, ICFs, zAAPs, and SAPs.

Up to eight temporary records (CBU, CPE, and On/Off CoD) can be installed and activated at any given time.. You also have the flexibility of activating some of the resources on a given record. You do not have to activate the entire record. You also have the ability to add capacity and engines and extend the duration of the temporary upgrade concurrently, therefore eliminating the need for constant ordering of new temporary records for different customer scenarios.

You can order an On/Off CoD upgrade record using Resource Link.

Capacity Backup (CBU)

The Capacity BackUp capability (temporary upgrade) enables enterprises to provide flexible, cost-effective Disaster Recovery on z196 models. You can order a CBU upgrade record using Resource Link.

Capacity for Planned Events (CPE)

Capacity for Planned Events (CPE) is designed to replace lost capacity within a customer's enterprise for planned down time events, such as system migration or relocation (for a data center move). This temporary upgrade is available for three days. You can order a CPE upgrade record using Resource Link.

Capacity provisioning

Capacity provisioning allows you to set up rules defining the circumstances under which additional capacity should be provisioned in order to fulfill a specific business need. The rules are based on criteria, such as: a specific application, the maximum additional capacity that should be activated, time and workload conditions.

This support provides a fast response to capacity changes and ensures sufficient processing power will be available with the least possible delay even if workloads fluctuate.

For more information, refer to the *z/OS MVS Capacity Provisioning Manager User's Guide*.

System-managed CF structure duplexing (CF duplexing)

A set of architectural extensions to the Parallel Sysplex is provided for the support of system managed CF structure duplexing (CF duplexing) of coupling facility structures for high availability.

Installing this software and microcode, and enabling this function is designed to:

- Provide the necessary base for highly available coupling facility structure data through the redundancy of duplexing.
- Enhance Parallel Sysplex ease of use by reducing the complexity of CF structure recover.
- Enable some installations to eliminate the requirement for standalone CFs in their Parallel Sysplex configuration.

CBU smart reactivation

CBU smart reactivation reduces outage duration, during a disaster event, if a processor book (containing the MCM) needs to be changed (while CBU is activated).

GDPS

GDPS is a collection of several offerings, each addressing a different set of Information Technology resiliency goals, that can be tailored to meet the Recovery Point Objective (RPO) and Recovery Time Objective (RTO) for your business. Each offering leverages a combination of server and storage hardware or software-based replication as well as automation and clustering software technologies. A summary of these offerings is described in more detail in "GDPS" on page 98.

Concurrent undo CBU

A prerequisite to executing this feature is that the customer or z/OS application must configure offline the processors that are being removed. So the best rule to follow is, "Configure offline the same logical processors that were configured online following the CBU activation." The concurrent undo CBU will require the following actions to configure the processors offline based on how it will be invoked:

- GDPS invoked Hardware Management Console/Support Element API
- Customer program invoked Hardware Management Console/Support Element API.

Notes:

1. As the user (or z/OS automation) deconfigures logical CPs, there is no guarantee that the logical CPs will remain in sequential numbering.
2. The Support Element panel will give no directions as to which CPs, ICFs, or IFLs to configure offline.

Fiber optic cabling

To serve the cabling needs of System z customers, IBM Site and Facilities Services has fiber optic cabling services available whether the requirements are product-level or enterprise-level. These services take into consideration the requirements for all of the protocols and media types supported on the z196 (for example, ESCON, FICON, coupling links, OSA) whether the focus is the data center, the Storage Area Network (SAN), the Local Area Network (LAN), or the end-to-end enterprise.

CHPID Mapping Tool

This tool provides a convenient interface to map hardware I/O ports on order to your CHPID definitions. An availability option automatically assigns PCHIDs to your CHPID definitions to minimize connecting critical paths to a single points of failure. This is recommended for all new z196 hardware builds or for upgrades from a z9 EC processor to a z196, as well as for making changes to an already installed machine after hardware upgrades that change or increase the number of channels.

Multipath initial program load

z/OS on z196 allows the system to attempt an IPL on alternate paths, if available, if I/O errors occur during the IPL. The system will attempt the IPL on an alternate path until all paths have been attempted or until the IPL is successful. This function increases the availability by allowing an IPL to complete using alternate paths and eliminates the need for manual problem determination when attempting an IPL.

This function is applicable for all FICON features with CHPID type FC and all ESCON features with CHPID type CNC.

Point-to-point SMP network

For z196, point-to-point SMP network provides growth paths up to a 80 engine system, where each of the 80 PUs have full access to all system resources, specially memory and I/O. A point-to-point SMP network design provides greater bandwidth and more interconnect concurrency between resources.

System-initiated CHPID reconfiguration

This function allows you to submit one request to all operating systems to configure offline or configure online all the CSS.CHPIDs associated with a partciular CHPID. It reduces the duration of a repair action when an ESCON or FICON channel; an OSA port; or an ISC-3 or IFB link is shared across logical partitions (LPARs).

Link aggregation support

Link aggregation (trunking) is designed to allow you to combine multiple physical OSA ports of the same type into a single logical link. You can have up to eight OSA ports in one virtual switch. This increases bandwidth and permits nondisruptive failover in the event that a port becomes unavailable. This function dedicates an OSA port to the z/VM 5.4 or later operating system for link aggregation under z/VM Virtual Switch-controlled link aggregation.

This support also provides dynamic add/remove of OSA ports and full-duplex mode (send and receive).

This support applies to OSA-Express2, OSA-Express3, and OSA-Express4S.

System power on/off cycle tracking

To help analyze power issues, z196 has the ability to track system power on/off cycles and transmit this data to IBM using the Transmit System Availability Data (TSAD) function.

Network Traffic Analyzer Trace facility

The Network Traffic Analyzer Trace facility is a diagnostic tool used to capture data as it enters or leaves an OSA adapter or Hipersockets channel for an attached host.

For OSA adapters, this facility is controlled and formatted by the z/OS Communications Server; however, the traced data is collected in the OSA at the network port.

For HiperSockets channels, the Support Element sets up authorization to allow tracing on selected HiperSockets. Traced data can be collected in a Linux partition; then tcpdump tools can be used to format and analyze the data.

For OSA adapters and HiperSockets Layer 2 devices, because the data is collected at the Ethernet frame level, you can trace the MAC headers for packets. For OSA adapters and HiperSockets channels, you can trace ARP packets, SNA packets, and packets being sent to and from other users sharing the OSA adapter or Hipersockets channel, including other TCP/IP stacks, Linux on System z users, and z/VM guest exploitation.

The Network Traffic Analyzer Trace facility supports OSA-Express2, OSA-Express3, OSA-Express4S, and Hipersockets.

QDIO diagnostic synchronization

Queued Direct I/O (QDIO) diagnostic synchronization provides the ability to coordinate and simultaneously capture software (z/OS) and hardware (OSA) traces. This function allows the host operating system to signal the OSA feature to stop traces and allows the operator to capture both the hardware and software traces at the same time. You can specify an optional filter that alters what type of diagnostic data is collected by the OSA-Express adapter. This filtering reduces the overall amount of diagnostic data collected and therefore decreases the likelihood that pertinent data is lost.

This support applies to OSA-Express2, OSA-Express3, and OSA-Express4S.

FICON purge path extended

The FICON purge path error-recovery function is used in FICON problem determination. The FICON purge path error-recovery function can transfer error-related data and statistics between the channel and entry switch, and from the control unit and its entry switch to the host operating system.

FICON Express8S, FICON Express8, and FICON Express4 pluggable optics for individual servicing

The FICON Express8S, FICON Express8, and FICON Express4 features have small form factor (SFF) pluggable optics to permit each channel to be individually serviced in the event of a fiber optic module failure. The traffic on the other channels on the same feature can continue to flow if a channel requires servicing.

CICS subspace group facility

z196 provides support for the subspace group facility that can enhance the data integrity and reliability of application server subsystems, such as Customer Information Control System Transaction Server (CICS TS), reducing application failures, service outages, and incorrect modification of critical business data.

Dynamic channel path management

Dynamic channel path management (DCM) enables the system to respond to changing channel requirements by moving channels from lesser used control units to more heavily used control units as needed. DCM can manage control units connected to ESCON channels.

When used with z/OS Workload Manager (z/OS WLM) in goal mode, z/OS WLM can direct Dynamic Channel Path Management to move channels to help business critical work achieve its goals. This also helps reduce the requirement for greater than 256 channels.

Serviceability

The features that provide a high level of serviceability include the following:

- Automatic error detection and fault isolation concurrent with system operation.
- Automatic remote support capability.
- High degree of concurrent maintenance capability in hardware and code.
- Multiple Channel Swap - an enhancement for channel problem determination allowing up to 16 channels to be swapped.
- Status Panel showing status of N+1 power system.

Appendix A. zEnterprise 196 Version 2.11.1 purpose and description

This appendix contains detailed information about Version 2.11.1 licensed internal code.

Preventative Service Planning (PSP) bucket considerations

Use IBM Service Link or contact your IBM representative to obtain a current copy of the 2817DEVICE bucket applicable to your environment. The PSP bucket contains corequisite software and hardware planning information that applies to various operating system environments. This includes Authorized Program Analysis Reports (APARS), Program Temporary Fixes (PTFs), and Licensed Internal Code (LIC) product patches.

Software corequisites

See the appropriate 2817DEVICE Preventative Service Planning (PSP) buckets subset ID for APAR and PTF information for the z196 models.

Table 20. Software corequisites

Software	PSP bucket subset ID
z/OS	2817/ZOS
z/VM	2817/ZVM
z/VSE	2817/ZVSE

Engineering change (EC) considerations

Version Code 2.11.1 for z196 includes the following Support Element and Hardware Management Console (HMC) Licensed Internal Code (LIC), engineering change (EC) and Microcode Load (MCL) levels:

- Support Element level: EC N48168 + MCLs
- HMC Level: EC N48180 + MCLs

To verify that the enhancements described in this document apply to your system, display the LIC EC levels running on the Support Element and the HMC.

Support Element EC N48168 + MCLs

From the HMC using the tree view, you can display the LIC EC and MCL level of the system's CPC as follows:

1. From the navigation pane, select **Tasks Index**.
2. Scroll down the Tasks Index work pane and select **System Information**. The Target Object Selection window displays.
3. Select the object and click **OK**. The System Information window displays
4. Verify that the EC level is in this list.

HMC EC N48180 + MCLs

From the HMC using the tree view, you can display the LIC EC and MCL level of the system's HMC as follows:

1. From the navigation pane, select **Tasks Index**.

2. Scroll down the Tasks Index work pane and select **View Console Information**. The View Console Information window displays.
3. Verify that the EC level is in this list.

Miscellaneous lower level ECs included in Version 2.11.1

The following table provides miscellaneous changes included in Support Element system code EC N48168 with Hardware Management Console system code EC N48180.

Table 21. ECs included in Version 2.11.1

EC number	Name
N26907	Backup-UFD New Build & MES Upgrade
N48187	SUL-UFD Driver 93
N26909	Security-Log UFD
N26910	TKE Backup-UFD New Build & MES Upgrade
N26911	SE Upgrade Data UFD MES Only
N10984	HMC DIAGS CDR
N26912	SE DIAGS CDR T510 ThinkPad
N48188	SE DIAGS CDR T520 ThinkPad
N26913	HWMCA/TKE Upgrade Data DVD MES Only
N29821	Add UFD Support for SE HDD Restore CDR
N48189	Ensemble/zBX Base SUL-UFD Driver 93
N48190	PBlade_OS SUL-UFD Driver 93

Appendix B. Resource Link

Resource Link is a customized web-based solution that provides everything you need to plan for, install, and maintain zEnterprise, System z10, System z9, eServer™ zSeries and S/390 servers and associated software.

You can access Resource Link at *http://www.ibm.com/servers/resourcelink*.

Resource Link content areas include:

- **Planning**
  Interactive planning provides a streamlined plan for the installation of a system using online planning information tailored for your system.
- **Education**
  A web-based multimedia education provides product courses that can be used for training or refreshing skills.
- **Library**
  Product documentation that can be viewed, printed, or downloaded.
- **Fixes**
  Interactive tools allow you to request, receive, and install system upgrades.
- **Problem Solving**
  Resources to help you investigate and solve specific hardware and software problems.
- **Services**
  Support for services such as installation, migration, networking, planning and relocating servers, Fiber Cabling, System z Application Programming Interfaces (APIs), and links to IBM software support.
- **Tools**
  Information about tools such as machine information, CHPID mapping, coupling facility structure sizer, power estimator, and links to software tools.
- **Customer Initiated Upgrade (CIU)**
  A web-based application that allows you to download licensed internal code (LIC) to permanently upgrade processors and memory. You can also temporarily add processor capacity using the On/Off Capacity on Demand (On/Off CoD), Capacity for Planned Events (CPE), and Capacity Backup (CBU) features.

Resource Link functions

Resource Link contains the following additional functions:

- **Customized planning aids** - Prepares you for the arrival and installation of your z196. To use Customized Planning Aids you need a valid order number and a Configuration Control Number (CCN), both available from your IBM Sales Representative.
- **CHPID Mapping Tool** - Downloadable from Resource Link, this tool allows you to map hardware I/O ports on order to your IOCP CHPID definitions. An availability option automatically maps hardware I/O ports to CHPIDs minimizing single points of failure.
- **Machine information** - Provides detailed information about your zEnterprise, System z10, System z9, or zSeries machine including information about the last time your machine called home, service states, and hardware status.
- **Power Estimation Tool** - Allows you to estimate the power consumption of a specific zEnterprise, System z10, or System z9 machine model and its associated configuration.
- **WWPN Tool** - Assists you in preplanning and setting up your Storage Area Networks (SANs) environment prior to the installation of your zEnterprise or System z10 server. Therefore, you can be

up and running much faster after the server is installed. This tool applies to all FICON channels defined as CHPID type FCP (for communication with SCSI devices).

- **Large Systems Performance Reference for IBM System z** - The IBM Large System Performance Reference (LSPR) ratios represent IBM's assessment of relative processor capacity in an unconstrained environment for the specific benchmark workloads and system control programs specified in the tables.

Appendix C. Capacity upgrades

z196 is designed to support concurrent upgrades that provide additional capacity with no server outage. The Capacity on Demand offerings provide permanent and temporary upgrades. All the upgrades are delivered by Licensed Internal Code Configuration Control (LICCC).

Licensed Internal Code Configuration Control (LICCC) provides for processor or memory upgrade with no hardware changes by enabling the activation of additional installed capacity. Concurrent upgrades using LICCC can be done for:
- CPs, SAPs, ICFs, IFLs, zIIPs, and zAAPs - requires available unused PUs in the installed drawer
- Memory - requires available capacity on installed memory cards
- Channel cards - requires available ports on channel cards.

You can order permanent upgrades using the Customer Initiated Upgrade (CIU) application through Resource Link or calling your IBM sales representative.

There are three type of temporary upgrades available on z196. The offerings providing these upgrades are: On/Off Capacity on Demand (On/Off CoD), Capacity Backup (CBU), or Capacity for Planned Events (CPE). You can order a CPE and CBU temporary upgrade using the CIU application through Resource Link or calling your IBM sales representative. You can order an On/Off CoD temporary upgrade **only** using the CIU application through Resource Link.

Each Capacity on Demand offering is available through an IBM contract. You must order the Online CoD Buying feature (FC 9900) to enable using Resource Link to order capacity upgrades. Refer to the *zEnterprise System Capacity on Demand User's Guide* for details.

Permanent upgrades

When using the CIU application through Resource Link to order a permanent upgrade, you can:
- Increase total and active model capacity
- Add specialty engines (ICFs, IFLs, zAAPs, zIIPs, and SAPs)
- Add memory
- Increase total model capacity and IFLs without changing the active model capacity and IFLs
- Activate unassigned model capacity or IFLs
- Deactivate activated model capacity or IFLs.

You can perform permanent upgrades while temporary capacity is active. This allows for quick conversion of temporary capacity to permanent capacity.

When calling your IBM sales representative to order a permanent upgrade (referred to as Capacity Upgrade on Demand (CUoD)), you can increase model capacity, add specialty engines (ICFs, IFLs, zAAPs, zIIPs, and SAPs), add memory, activate unassigned model capacity or IFLs, deactivate activated model capacity or IFLs, activate channels, activate crypto, and perform recharacterization.

Refer to the *zEnterprise System Capacity on Demand User's Guide* for more information.

Temporary upgrades

Using On/Off CoD, CBU, or CPE, you can increase model capacity and add specialty engines (ICFs, IFLs, zAAPs, zIIPs, and SAPs).

Characteristics of temporary upgrades include:

- **Permanent upgrade while temporary capacity is active** - You can add permanent processor or memory capacity while temporary On/Off CoD, CBU, or CPE records are active. This allows for quick conversion of temporary capacity to permanent capacity.

  Note: With active On/Off CoD records, On/Off CoD engines of the same type are converted to permanent engines. With active CBU and CPE records, active capacity is not replaced by permanent capacity – engines and capacity levels are added to the permanent engines and capacity levels, if sufficient engines and capacity levels are available.
- **Multiple records can be simultaneously active** - Up to eight records (On/Off CoD, CBU, and CPE) can be active at any given time. However, only one On/Off CoD record can be active at any given time.
- **Store LICCC records in an unactivated state** - Up to 200 records (On/Off CoD, CBU, and CPE) can be staged on the Support Element at any given time. This provides greater flexibility to quickly enable needed temporary capacity.
- **Automatic deactivation** - When a record expires, the resource is automatically deactivated. However, the record will not be deactivated if it means removing a dedicated engine or the last of that engine type.
- **Partial activation** - You do not have to activate the entire record. You can choose partial activation of resources up to the maximum you ordered.

On/Off Capacity on Demand (On/Off CoD)

On/Off Capacity on Demand (On/Off CoD) is designed to satisfy your need for short term additional capacity. On/Off CoD allows you to temporarily add any available unused resource (CPs, IFLs, ICFs, zIIPs, zAAPs, and SAPs) up to two times the purchased capacity. You can order this upgrade only using the CIU application through Resource Link.

The upgrade record is downloaded, staged, installed, and activated on your z196 server through its Support Element. The On/Off CoD record **is not** automatically activated when it is downloaded. It is placed in a "staged" area on the Support Element waiting to be installed and activated.

If you need the increased capacity for a longer period of time or you want to increase the capacity beyond the current record limit, you can “replenish” the record. Using Resource Link, you place an order for a replenishment record to extend the expiration date, increase the capacity limits, or add additional tokens to an existing upgrade record. Replenishment allows you to update an existing record without having to place a completely new order and to update an existing record while capacity is active for that record. Under certain situations, you have the ability to set an automatic renewal option that will automatically extend the expiration date without having to manually replenish the record at the end of 180 days.

When you order an On/Off CoD record, you can either prepay or post-pay for the upgrades. The payment method is based on the type of On/Off CoD upgrade you select:

- When you order a post-paid On/Off CoD record without spending limits, you select your upgrade configuration. There is no cost incurred when you order or install this type of record. You pay for what you activate during the activation time. You are charged on a 24-hour basis. For each month (starting with the month you activated the record), a report is generated. In the following month, you are billed for hardware and software charges.
- When you order a prepaid On/Off CoD record, you select your upgrade configuration and identify the duration of the configuration. Resource Link calculates the number of tokens you will need to activate

your selected upgrade configurations. When the order is downloaded, you are billed for the total hardware cost. As resources are used, the corresponding number of tokens are decremented. Tokens are tracked on a 24-hour basis. For each month resources are used, a report is generated. In the following month, you are billed for software charges.

- When you order a post-paid On/Off CoD record with spending limits, you select your upgrade configuration and identify your spending limit for each upgrade. Resource Link calculates the maximum number of tokens you may need to activate upgrade configurations without exceeding your spending limit. Tokens are tracked on a 24-hour basis. You will be notified when you are reaching the limit you set on your order. For each month (starting with the month you downloaded the record), a report is generated. In the following month, you are billed for hardware charges. Software charges are separate.

There are limits to the number of temporary zIIPs, zAAPs, IFLs, ICFs, and SAPs you can purchase. Refer to the *zEnterprise System Capacity on Demand User's Guide* for details.

The On/Off CoD characteristics include:

- **Reusable On/Off CoD records** - Using a single On/Off CoD upgrade record, z196 supports the moving from one capacity setting to another, either decreasing or increasing the amount of active temporary capacity. Multiple activations are possible within the same On/Off CoD upgrade record. The record remains valid for 180 days unless it is manually replenished or the automatic renewal option was set to extend the expiration date.
- **API used to activate** - z196 allows activation of On/Off CoD using SNMP, CIM , or Web Services APIs.
- **No-charge test** - The On/Off CoD test can be used to validate the process of ordering, downloading, activating, and deactivating On/Off CoD capacity nondisruptively. Activating an On/Off CoD test record actually upgrades your configuration, which differs from the administrative On/Off test. With each On/Off CoD enabled machine, you are entitled to one no-charge test. The test may run for a maximum duration of 24 hours beginning with the activation of the test record.
- **Multiple records simultaneously active** - An On/Off CoD record, CBU record, and CPE record can be active at the same time.
- **Administrative test** - A no-charge administrative On/Off CoD test record is available that allows you to validate the process of ordering, downloading, activating, and deactivating On/Off records without actually setting real capacity. This test record can be used to train personnel or test applications.

Refer to the *zEnterprise System Capacity on Demand User's Guide* for more information.

Capacity Backup (CBU)

Capacity Backup (CBU) is designed to replace lost capacity due to an emergency or disaster recovery situation. CBU increases capacity nondisruptively by allowing you to add specialty engines (IFLs, ICFs, zAAPs, zIIPs, SAPs) or add capacity by feature codes.

A combination of up to four CPE and CBU temporary records that are ordered are automatically installed during the manufacturing process. If more than four records are ordered, the records are staged on the Support Element and you can manually select which records to install.

Each CBU record is allowed one 90-day "real" activation and a number of free 10-day test activations. The number of free test activations equates to the number of years that are purchased with the CBU record. (For example, a three year CBU record has three tests activations, a one year CBU record has one test activation.) Additional test activations beyond the free tests may be purchased in **single** increments up to a maximum of 15 CBU tests per record. This maximum of 15 tests per record cannot be exceeded and includes any free activations plus additional paid test activations.

The CBU characteristics include:

- **No password is required at time of activation**.

- **Specialty engines are managed by quantities** - Added capacity is dictated by processor types. You must indicate the number of engines that can be added to the permanent configuration.
- **CP capacity is managed by feature codes** - Feature codes either adds engines or increase the capacity to a permanent engine.
- **Choice in the length of contract** - Expiration date of a contract is 1 to 5 years. You have the capability to extend your CBU record up to the maximum 5 year limit. One test activation is provided for each additional CBU year added to the CBU record.
- **Limit on the number of zIIPs or zAAPs you can order** - This number cannot exceed the total number of permanents plus temporary CPs.

Refer to the *zEnterprise System Capacity on Demand User's Guide* for more information.

Capacity for Planned Events (CPE)

Capacity for Planned Events (CPE) is designed to replace lost capacity for planned down time events, such as system migration or relocation (for a data center move). CPE increases capacity by allowing you to add model capacity or specialty engines (IFLs, ICFs, zAAPs, zIIPs, SAPs). Pricing is based on the model capacity and the type and quantity of the engines selected.

A combination of up to four CPE and CBU temporary records that are ordered are automatically installed during the manufacturing process. If more than four records are ordered, the records are staged on the Support Element and you can manually select which records to install.

Each CPE order includes 1 activation for 3 days.

Refer to the *zEnterprise System Capacity on Demand User's Guide* for more information.

Concurrent PU conversions

The z196 supports concurrent conversion of different processor unit (PU) types. This capability is extended to CPs, IFLs, ICFs, zIIPs, and zAAPs. This capability provides flexibility in configuring a z196 to meet the changing business environments.

Note: Concurrent PU conversion is not supported by CIU.

Reserved CP support in LPAR mode

With reserved CP support in LPAR mode, an LPAR may be defined with the number of logical CPs greater than the number of physical CPs. Additional CPs can be specified for the LPAR definition beyond the number of physical CPs currently installed on the model. Therefore, an enterprise planning to do a nondisruptive upgrade (with an LPAR defined of logical CPs equal to the number of physical CPs available on the installed hardware) does not need to deactivate, redefine, then reactivate in order to take advantage of the new CPs that have been activated. The enterprise simply needs to have defined additional CPs for the LPAR in advance. This ensures that any planned LPAR can be as large as the possible physical machine configuration. With the logical processor add function, the logical partition profile definition can be changed dynamically to add more logical processors to an active logical partition, nondisruptively. For more information, refer to *zEnterprise System Processor Resource/Systems Manager Planning Guide*.

Nondisruptive upgrades

The z196 Plan-Ahead process links the use of Capacity Upgrade on Demand with planning performed between IBM's account team and IBM's customer. Planning ahead enables customers to determine a future server configuration. IBM will also support its customers planning effort via capacity planning tools, IBM's order processing configurative and team sessions, with the objective of nondisruptive growth to satisfy essential capacity demand.

Processor capacity downgrades

You are allowed to downgrade your machine using CIU, CUoD, or MES. The primary benefit to downgrading is a reduction in software charges based on a lower reported machine capacity.

Some additional considerations should be noted when downgrading:

- Downgrades are done by “unassigning” either CPs or IFLs.
- There may be a charge to unassign and then reactivate engines.
- Unassigned engines are still owned by the customer.
- Unassigning unused engines can reduce software charges since many software products are priced based on the number of active engines.
- Unassigned engines can be reactivated by CIU, CUoD, or MES.
- Unassigned engines may be temporarily activated using On/Off CoD or CPE. When used as a temporary engine, unassigned engines can be used as any of the supported engine types (thus an unassigned IFL can be activated as a CP). Reduced hardware usage charges are available when using unassigned engines as the same type.
- Unassigning of engines and later reactivation is concurrent.

Appendix D. Notices

This information was developed for products and services offered in the USA.

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
Armonk, NY 10504-1785 USA

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 states 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 supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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.

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

All IBM prices shown are IBM's suggested retail prices, are current and are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

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 the names and addresses used by an actual business enterprise is entirely coincidental.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

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.

Adobe is a registered trademark of Adobe Systems Incorporated in the United States, and/or other countries.

InfiniBand is a registered trademark of the InfiniBand Trade Association.

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

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

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Other product and service names might be trademarks of IBM or other companies.

Electronic emission notices

The following 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 instructions contained in the installation 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, by installation or use of this equipment other than as specified in

the installation manual, or by any other 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.

Canadian Department of Communications Compliance Statement

This Class A digital apparatus complies with Canadian ICES-003.

Avis de conformlté aux normes du ministère des Communications du Canada

Cet appareil numérique de la classe A est conform à la norme NMB-003 du Canada.

European Union (EU) Electromagnetic Compatibility Directive

This product is in conformity with the protection requirements of EU Council Directive 2004/108/EC 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 55022. The limits for Class equipment were derived for commercial and industrial environments to provide reasonable protection against interference with licensed communication equipment.

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.

European Community contact:
IBM Deutschland GmbH
Technical Regulations, Department M372
IBM-Allee 1, 71139 Ehningen, Germany
Telephone: 0049 (0) 7032 15-2941
email: lugi@de.ibm.com

EC Declaration of Conformity (In German)

Deutschsprachiger EU Hinweis: Hinweis für Geräte der Klasse A EU-Richtlinie zur Elektromagnetischen Verträglichkeit

Dieses Produkt entspricht den Schutzanforderungen der EU-Richtlinie 89/336/EWG zur Angleichung der Rechtsvorschriften über die elektromagnetische Verträglichkeit in den EU-Mitgliedsstaaten und hält die Grenzwerte der EN 55022 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 der IBM verändert bzw. wenn Erweiterungskomponenten von Fremdherstellern ohne Empfehlung der IBM gesteckt/eingebaut werden.

EN 55022 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 89/336/EWG in der Bundesrepublik Deutschland.

Zulassungsbescheinigung laut dem Deutschen Gesetz über die elektromagnetische Verträglichkeit von Geräten (EMVG) vom 18. September 1998 (bzw. der EMC EG Richtlinie 89/336) für Geräte der Klasse A.

Dieses Gerät ist berechtigt, in Übereinstimmung mit dem Deutschen EMVG das EG-Konformitätszeichen - CE - zu führen.

Verantwortlich für die Konformitätserklärung nach Paragraf 5 des EMVG ist die IBM Deutschland GmbH, 70548 Stuttgart.

Informationen in Hinsicht EMVG Paragraf 4 Abs. (1) 4:

Das Gerät erfüllt die Schutzanforderungen nach EN 55024 und EN 55022 Klasse A.

update: 2004/12/07

People's Republic of China Class A Compliance Statement

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 actions.

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。

Japan Class A Compliance Statement

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.

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

Korean Class A Compliance Statement

이 기기는 업무용(A급)으로 전자파적합등록을 한 기기이오니
판매자 또는 사용자는 이 점을 주의하시기 바라며, 가정외의
지역에서 사용하는 것을 목적으로 합니다.

Taiwan Class A Compliance Statement

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.

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。

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

Glossary

A.

abend. See abnormal end of task.

abnormal end of task. Ending a task before its completion because of an error condition that cannot be resolved by recovery facilities while the task is being executed.

activate logical partition. An operator-initiated procedure that performs a system reset to an LPAR and assigns the previously defined hardware to that partition. It causes an automatic IPL of the system control program to occur in the partition unless the operator performs the IPL manually.

active subchannel. A subchannel that is locked and either busy or has a pending interrupt, and is indicated by subchannel status word (SCSW) bit 24 equals 1. The control information resides in the channel subsystem because it is necessary for the current operation.

Note: An active subchannel can also reside in the local working storage of an IOP or channel.

active window. The window with which users are currently interacting. This is the window that receives keyboard input.

advanced management module (AMM). A hardware unit that provides system-management functions for all the blade servers in a BladeCenter chassis.

alert. A unit of information, usually indicating the loss of a system resource, passed from one machine or program to a host to signal an error.

allocate. To assign a resource, such as a disk or a diskette file to perform a task.

alternate HMC. A System z Hardware Management Console (HMC) that is paired with the primary HMC to provide redundancy.

See also primary HMC.

American National Standard Code for Information Interchange (ASCII). The standard code, using a coded character set consisting of 7-bit coded characters (8 bits including parity), used for information interchange among data processing systems, data communication systems, and associated equipment. The ASCII set consists of control characters and graphics characters.

Note: IBM has defined an extension to ASCII code (characters 128 - 255).

AMM. See advanced management module.

ANSI. American National Standards Institute

APAR. Authorized program analysis report

API. Application programming interface

application. A program that is specific to the solution of an application problem.

A program written for or by a user that applies to the user's work, such as a program that does inventory control or payroll.

A program used to connect and communicate with stations in a network, enabling users to perform application-oriented activities.

Application Assist Processor (AAP). A special processor configured for running Java applications on zEnterprise, z10, z9, z990 and z890 class machines.

application environment. The environment that includes the software and the server or network infrastructure that supports it.

ASCII. American National Standard Code for Information Interchange

asynchronous. Pertaining to two or more processes that do not depend upon the occurrence of specific events such as common timing signals.

Without regular time relationship; unexpected or unpredictable with respect to the execution of program instructions.

authorized program analysis report (APAR). A request for correction of a problem caused by a defect in a current release of a program unaltered by the user.

auto-answer. In data communication, the ability of a station to respond automatically to a call that it receives over a switched line.

auto-call. In data communication, the ability of a station to initiate a call automatically over a switched line.

Automate suite (Automate). The second of two suites of functionality associated with the IBM zEnterprise Unified Resource Manager. The Automate suite includes goal-oriented monitoring and management of resources and energy management.

See also Manage suite.

B.

basic mode. A central processor mode that does not use logical partitioning.

batch. An accumulation of data to be processed.

A group of records or data processing jobs brought together for processing or transmission.

Pertaining to activity involving little or no user action.

blade. A hardware unit that provides application-specific services and components. The consistent size and shape (or form factor) of each blade allows it to fit in a BladeCenter chassis.

BladeCenter chassis. A modular chassis that can contain multiple blades, allowing the individual blades to share resources such as the management, switch, power, and blower modules.

block multiplexer channel. A multiplexer channel that interleaves blocks of data.

BPA. Bulk power assembly

buffer. A routine or storage used to compensate for a difference in rate of flow of data, or time of occurrence of events, when transferring data from one device to another.

To allocate and schedule the use of buffers.

A portion of storage used to hold input or output data temporarily.

bus. A facility for transferring data between several devices located between two end points, only one device being able to transmit at a given moment.

A network configuration in which nodes are interconnected through a bidirectional transmission medium.

One or more conductors used for transmitting signals or power.

byte. A string that consists of a particular number of bits, usually eight, that is treated as a unit, and that represents a character.

byte multiplexer channel. A multiplexer channel that interleaves bytes of data.

C.

cache. A special purpose buffer storage, smaller and faster than main storage, used to hold a copy of the instructions and data obtained from main storage and likely to be needed next by the processor. (T)

A buffer storage that contains frequently accessed instructions and data; it is used to reduce access time.

CAW. Channel Address Word

CBU. Capacity Backup

CCC. Channel control check

CCW. Channel command word

CDC. Channel data check

central processor (CP). The part of the computer that contains the sequencing and processing facilities for instruction execution, initial program load, and other machine operations.

central processor complex (CPC). A physical collection of hardware that consists of main storage, one or more central processors, timers, and channels. In the zEnterprise environment, the CPC consists of a zEnterprise mainframe and any attached IBM zEnterprise BladeCenter Extension (zBX).

See also node and zCPC.

central storage. Storage that is an integral part of the processor and includes both main storage and the hardware system area.

CF. Coupling facility

channel. A path along which signals can be sent, for example, input/output channel.

The system element that controls one channel path, whose mode of operation depends on the type of hardware to which it is attached.

channel adapter. A communication controller hardware unit used to attach the controller to a data channel.

Hardware that attaches a group of channels to the secondary data stager and prioritizes and stages data between the channels and the channel control element.

channel address. In S/370 mode, the 8 leftmost bits of an input/output address that identify the channel.

channel address word (CAW). An area in storage that specifies the location in main storage at which a channel program begins.

channel command word (CCW). A doubleword at the location in main storage specified by the channel address word. One or more CCWs make up the channel program that directs data channel operations.

channel command word (CCW). A doubleword at the location in main storage specified by the channel address word. One or more CCWs make up the channel program that directs data channel operations.

channel data check. A category of I/O errors, indicating a machine error in transferring data to or from storage and sensed by the channel to which a device is attached.

channel data rate. The rate at which a channel can move data between a transmission link and processor storage during the data transfer portion of an I/O operation.

channel Licensed Internal Code. That part of the channel subsystem Licensed Internal Code used to start, maintain, and end all operations on the I/O interface.

channel path. A single interface between a central processor and one or more control units along which signals and data can be sent to perform I/O requests.

channel path identifier (CHPID). The channel subsystem communicates with I/O devices by means of a channel path between the channel subsystem and devices. A CHPID is a value assigned to each channel path of the System z that uniquely identifies that path. Up to 256 CHPIDs are supported for each channel subsystem.

channel status word (CSW) . An area in storage that provides information about the termination of input/output operations.

channel subsystem (CSS). A collection of subchannels that directs the flow of information between I/O devices and main storage, relieves the processor of communication tasks, and performs path management functions.

channel subsystem (CSS) Licensed Internal Code. Code that consists of the IOP Licensed Internal Code and the channel Licensed Internal Code.

channel-to-channel (CTC). Communication (transfer of data) between programs on opposite sides of a channel-to-channel adapter (CTCA).

channel-to-channel adapter (CTCA). An input/output device that is used by a program in one system to communicate with a program in another system.

check stop. The state that occurs when an error makes it impossible or undesirable to continue the operation in progress.

CHPID. Channel path identifier

CIB. Coupling using InfiniBand

CICS. Customer Information Control System

CICS/ESA. Customer Information Control System/Enterprise Systems Architecture

CIU. Customer Initiated Upgrade

CKD. Count key data

CLIST (command list). A data set in which commands and possibly subcommands and data are stored for subsequent execution.

CNC. Mnemonic for an ESCON channel attached to an ESCON-capable device.

command chaining. The fetching of a new channel command word (CCW) immediately following the completion of the previous CCW.

command retry. A channel and control unit procedure that causes a command to be retried without requiring an I/O interrupt.

communication control unit. A communication device that controls transmission of data over lines in a network.

communication controller. A device that directs the transmission of data over the data links of a network; its operation can be controlled by a program executed in a processor to which the controller is connected or it may be controlled by a program executed within the device.

A type of communication control unit whose operations are controlled by one or more programs stored and executed in the unit. It manages the details of line control and the routing of data through a network.

concurrent maintenance. Hardware maintenance actions performed by a service representative while normal operations continue without interruption.

connectivity. A term used to describe the physical interconnections of multiple devices/computers/ networks employing similar or different technology or architecture together to accomplish effective communication between and among connected members involving data exchange or resource sharing.

console. A logical device used for communication between the user and the system.

console integration (CI). The hardware and software facilities used to bring operating systems management and hardware systems management under a single control point.

control program. A computer program designed to schedule and to supervise the execution of programs of a computer system.

control unit. A hardware unit that controls the reading, writing, or displaying of data at one or more input/output units.

control unit data rate. The rate at which a control unit can move data between itself and a transmission link during the data transfer portion of an I/O operation.

controller. A unit that controls input/output operations for one or more devices.

Conversational Monitor System (CMS). A virtual machine operating system that provides general interactive time sharing, problem solving, and program development capabilities, and operates only under the z/VM control program.

Coordinated Server Time (CST). Represents the time in a CTN. Timekeeping messages carried over the coupling links determine the CST at each server.

Coordinated Timing Network (CTN). A collection of servers that are time synchronized to Coordinated Server Time (CST). All STP-configured servers in a CTN must have the same CTN ID.

coupling facility. A special partition that provides high-speed caching, list processing, and locking functions in a Parallel Sysplex.

coupling facility channel. A high bandwidth fiber optic channel that provides the high-speed connectivity required for data sharing between a coupling facility and the central processor complexes directly attached to it.

CP. Control program

Central processor

CPC. See central processor complex.

CPC image. Set of CPC resources that support a single control program.

CPU. Central processor unit

CPUID. CPU identifier

CSS. Channel subsystem

CST. Coordinated Server Time

CSW. Channel status word

CTC. Channel-to-channel

Mnemonic for an ESCON or FICON channel attached to another ESCON or FICON channel, respectively.

CTCA. Channel-to-channel adapter

CTN. Coordinated Timing Network

CU. Control unit

CUA. Control unit address

CUod. Capacity Upgrade on Demand

Customer Information Control System (CICS). An IBM licensed program that enables transactions entered at remote terminals to be processed concurrently by user-written application programs. It includes facilities for building, using, and maintaining data bases.

CVC. Mnemonic for an ESCON channel attached to a 9034.

D.

DataPower XI50z. See IBM WebSphere DataPower Integration Appliance XI50 for zEnterprise.

DASD. Direct access storage device

DASD subsystem. A storage control and its attached direct access storage devices.

DAT. Dynamic address translation

data processing (DP). The systematic performance of operations upon data; for example, arithmetic or logic operations upon data, merging or sorting of data, assembling or compiling of programs.

data sharing. The ability of concurrent subsystems (such as DB2 or IMS DB) or application programs to directly access and change the same data while maintaining data integrity.

data streaming. In an I/O interface, a mode of operation that provides a method of data transfer at up to 4.5 MB per second. Data streaming is not interlocked between the sender and the receiver. Once data transfer begins, the sender does not wait for acknowledgment from the receiver before sending the next byte. The control unit determines the data transfer rate.

data transfer mode. The method of information exchange used on an I/O interface.

DB2. DATABASE 2

DCA. Distributed Converter Assembly

DDR. Double Data Rate

deactive logical partition. An operator-initiated procedure that releases the hardware assigned to a LPAR, making it available to other partitions.

Note: The operator should first deactivate the system control program, if possible or necessary, and then reactivate the partition, which could provide a reset to that partition, if required.

deallocate. To release a resource assigned to a task.

DES. Data Encryption Standard

DFSMS. Data Facility Storage Management Subsystem

direct access storage. A storage device that provides direct access to data.

direct access storage device (DASD). (1) A storage device in which the location of each data record can be directly addressed. (2) A device in which the access time is effectively independent of the location of the data. (Restriction: Does not refer to diskette drive.)

DP. Data processing

dual inline memory module (DIMM). A small circuit board with memory-integrated circuits containing signal and power pins on both sides of the board.

dynamic address translation (DAT). In virtual storage systems, the change of a virtual storage address to a real storage address during execution of an instruction.

dynamic reconfiguration management. In MVS, the ability to modify the I/O configuration definition without needing to perform a Power on Reset (POR) of the hardware or an initial program load (IPL).

dynamic storage reconfiguration. A PR/SM LPAR function that allows central or expanded storage to be added or removed from an LPAR without disrupting the system control program operating in the LPAR.

E.

EC. Engineering change

ECC. Error checking and correction

ECKD. Extended count key data

EIA. Electronics Industries Association. One EIA unit is 1.75 inches or 44.45 mm.

ensemble. A collection of one or more zEnterprise nodes (including any attached zBX) that are managed as a single logical virtualized system by the Unified Resource Manager, through the Hardware Management Console.

ensemble member. A zEnterprise node that has been added to an ensemble.

See also node.

Enterprise Systems Connections (ESCON). A set of products and services that provides a dynamically connected environment using optical cables as a transmission medium.

EPO. Emergency power off

error checking and correction (ECC). In a processor, the detection and correction of all single-bit errors, plus the detection of double-bit and some multiple-bit errors

ESA. Enterprise System Architecture

ESA/370. Enterprise System Architecture/370

ESA/390. Enterprise System Architecture/390

ESCON. Enterprise Systems Connection

ESCON channel. A channel having an Enterprise Systems Connection channel-to-control-unit I/O interface that uses optical cables as a transmission medium.

Ethernet definition. A communication network (USA, Xerox 1975).

ETR. External Time Reference

expanded storage. Optional high-speed storage that transfers 4 KB pages to and from central storage.

F.

firmware. Licensed Internal Code (LIC) that is shipped with hardware. Firmware is considered an integral part of the system and is loaded and run at power on. Firmware is not open for customer configuration and is expected to run without any customer setup.

G.

Gb. Gigabit

GB. Gigabyte

GbE. Gigabit Ethernet

gigabit (Gb). A unit of measure for storage size. One gigabit equals one billion bits.

Gigabit Ethernet. An OSA channel (CHPID type OSD)

gigabyte (GB). A unit of measure for storage size. One gigabyte equals 1,073,741,824 bytes. Loosely, one billion bytes.

GMT. Greenwich Mean Time

GPMP. See guest platform management provider.

guest platform management provider (GPMP). An optional suite of applications that is installed in specific z/OS, Linux, and AIX operating system images to support platform management functions. For example, the guest platform management provider collects and aggregates performance data for virtual servers and workloads.

H.

Hardware Management Console (HMC). A user interface through which data center personnel configure, control, monitor, and manage System z hardware and software resources. The HMC communicates with each central processor complex (CPC) through the Support Element. On an IBM zEnterprise 196 (z196) or IBM zEnterprise 114 (z114), using the Unified Resource Manager on the HMCs or Support Elements, personnel can also create and manage an ensemble.

See also primary HMC and alternate HMC.

Hardware system area (HSA). A logical area of central storage, not addressable by application programs, used to store Licensed Internal Code and control information.

HCA. Host Channel Adapter

HCA1-O fanout. The HCA1-O (optical) fanout card is used for coupling using an InfiniBand connection on a

z9. The HCA1-O fanout is designed to support a two-port 12x IB-SDR coupling link operating at a link data rate of 3 GBps.

HCA2-C fanout. The HCA2-C (copper) fanout card has InfiniBand connections used for internal I/O on a z10. The HCA2-C fanout is designed to support a two-port 12x IB-DDR copper link operating at a link data rate of 6 GBps.

HCA2-O fanout. The HCA2-O (optical) fanout card is used for coupling using an InfiniBand connection on a z10 or zEnterprise. The HCA2-O fanout is designed to support a two-port 12x IB-DDR coupling link operating at a link data rate of 3 GBps (if attached to a z9) and 6 GBps (if attached to a z10 or zEnterprise)

HCA2-O LR fanout. The HCA2-O LR fanout is designed to support a two-port 1x IFB coupling link with a link data rate of either 5.0 Gbps or 2.5 Gbps and a maximum unrepeated distance of 10 kilometers (6.2 miles) and a maximum repeated distance of 100 kilometers (62 miles).

HCA3-O fanout. The HCA3-O (optical) fanout card is used for coupling using an InfiniBand connection on a zEnterprise. The HCA3-O fanout is designed to support a two-port 12x IB-DDR coupling link operating at a link data rate of 6 GBps. The HCA3-O fanout also supports the 12x IFB3 protocol if four or less CHPIDs are defined per port. The 12x IFB3 protocol provides improved service times.

HCA3-O LR fanout. The HCA3-O LR fanout card is used to support four-port 1x IFB coupling link with a link data rate of 5.0 Gbps and a maximum unrepeated distance of 10 kilometers (6.2 miles) or a maximum repeated distance of 100 kilometers (62 miles). With DWDM, the HCA3-O LR fanout supports a four-port 1x IFB coupling link with a link data rate of either 2.5 or 5 Gbps. An HCA3-O LR fanout can communicate with a HCA2-O LR fanout on zEnterprise or System z10.

HCD. Hardware configuration definition

HiperSockets network traffic analyzer. Trace HiperSockets network traffic to help simplify problem isolation and resolution. Supported on zEnterprise and System z10.

HMC. See Hardware Management Console (HMC).

HMCA. Hardware Management Console Application

HSA. Hardware system area

hypervisor. A program that allows multiple instances of operating systems or virtual servers to run simultaneously on the same hardware device. A hypervisor can run directly on the hardware, can run within an operating system, or can be imbedded in platform firmware. Examples of hypervisors include PR/SM, z/VM, and PowerVM® Enterprise Edition.

I.

IBF. Internal Battery Feature

IBM blade. A customer-acquired, customer-installed select blade to be managed by IBM zEnterprise Unified Resource Manager. One example of an IBM blade is a POWER7 blade.

IBM DB2 Analytics Accelerator for z/OS. A workload-optimized, LAN-attached appliance based on Netezza technology.

IBM Smart Analytics Optimizer for DB2 for z/OS. An optimizer that processes certain types of data warehouse queries for DB2 for z/OS.

IBM System z Application Assist Processor (zAAP). A specialized processor that provides a Java execution environment, which enables Java-based web applications to be integrated with core z/OS business applications and backend database systems.

IBM System z Integrated Information Processor (zIIP). A specialized processor that provides computing capacity for selected data and transaction processing workloads and for selected network encryption workloads.

IBM WebSphere DataPower Integration Appliance XI50 for zEnterprise (DataPower XI50z). A purpose-built appliance that simplifies, helps secure, and optimizes XML and Web services processing.

IBM zEnterprise BladeCenter Extension (zBX). A heterogeneous hardware infrastructure that consists of a BladeCenter chassis attached to an IBM zEnterprise 196 (z196) or IBM zEnterprise 114 (z114). A BladeCenter chassis can contain IBM blades or optimizers.

IBM zEnterprise BladeCenter Extension (zBX) blade. Generic name for all blade types supported in an IBM zEnterprise BladeCenter Extension (zBX). This term includes IBM blades and optimizers.

IBM zEnterprise Unified Resource Manager. Licensed Internal Code (LIC), also known as firmware, that is part of the Hardware Management Console. The Unified Resource Manager provides energy monitoring and management, goal-oriented policy management, increased security, virtual networking, and data management for the physical and logical resources of a given ensemble.

IC. Internal Coupling link

ICB. Integrated Cluster Bus link

ICF. Internal Coupling Facility

ICSF. Integrated Cryptographic Service Facility

IEDN. See intraensemble data network (IEDN).

IEDN TOR switch. See intraensemble data network (IEDN) TOR switch.

IFB. InfiniBand

IFB-MP (InfiniBand Multiplexer) card. The IFB-MP card can only be used in the I/O cage or I/O drawer. The IFB-MP cards provide the intraconnection from the I/O cage or I/O drawer to the HCA2-C fanout card in a book or processor drawer.

IFCC. Interface control check

IFL. Integrated Facility for Linux

IML. Initial machine load

IMS. Information Management System

initial machine load (IML). A procedure that prepares a device for use.

initial program load (IPL). The initialization procedure that causes an operating system to commence operation.

The process by which a configuration image is loaded into storage at the beginning of a work day or after a system malfunction.

The process of loading system programs and preparing a system to run jobs.

initialization. The operations required for setting a device to a starting state, before the use of a data medium, or before implementation of a process.

Preparation of a system, device, or program for operation.

To set counters, switches, addresses, latches, or storage contents to zero or to other starting values at the beginning of, or at the prescribed points in, a computer program or process.

INMN. See intranode management network (INMN).

input/output (I/O). Pertaining to a device whose parts can perform an input process and an output process at the same time.

Pertaining to a functional unit or channel involved in an input process, output process, or both, concurrently or not, and to the data involved in such a process.

input/output configuration. The collection of channel paths, control units, and I/O devices that attach to the processor complex.

input/output configuration data set (IOCDS). The data set that contains an I/O configuration definition built by the I/O configuration program (IOCP).

input/output configuration program (IOCP). A program that defines to a system all the available I/O devices and the channel paths.

Integrated Facility for Applications (IFA). A general purpose assist processor for running specific types of applications.

interrupt. A suspension of a process, such as execution of a computer program caused by an external event, and performed in such a way that the process can be resumed.

intraensemble data network (IEDN). A private high-speed network for application data communications within an ensemble. Data communications for workloads can flow over the IEDN within and between nodes of an ensemble. The Unified Resource Manager configures, provisions, and manages all of the physical and logical resources of the IEDN.

intraensemble data network (IEDN) TOR switch. A top-of-rack switch that provides connectivity to the intraensemble data network (IEDN), supporting application data within an ensemble.

intranode management network (INMN). A private service network that the Unified Resource Manager uses to manage the resources within a single zEnterprise node. The INMN connects the Support Element to the IBM zEnterprise 196 (z196) or IBM zEnterprise 114 (z114) and to any attached IBM zEnterprise BladeCenter extension (zBX).

I/O. Input/output

IOCDS. I/O configuration data set

IOCP. I/O configuration program

IPL. Initial program load

IPv6. Internet Protocol Version 6

ISC. InterSystem Channel

K.

KB. Kilobyte

kilobyte (KB). A unit of measure for storage size. Loosely, one thousand bytes.

km. Kilometer

L.

LAN. Local area network

laser. A device that produces optical radiation using a population inversion to provide light amplification by stimulated emission of radiation and (generally) an

optical resonant cavity to provide positive feedback. Laser radiation can be highly coherent temporally, or spatially, or both.

LCSS. Logical channel subsystem

LED. Light-emitting diode

LIC. Licensed Internal Code

Licensed Internal Code (LIC). Software provided for use on specific IBM machines and licensed to customers under the terms of IBM's Customer Agreement.

light-emitting diode (LED). A semiconductor chip that gives off visible or infrared light when activated.

local area network (LAN). A computer network located on a user's premises within a limited geographical area. Communication within a local area network is not subject to external regulations; however, communication across the LAN boundary can be subject to some form of regulation.

logical address. The address found in the instruction address portion of the program status word (PSW). If translation is off, the logical address is the real address. If translation is on, the logical address is the virtual address.

logical control unit. A group of contiguous words in the hardware system area that provides all of the information necessary to control I/O operations through a group of paths that are defined in the IOCDS. Logical control units represent to the channel subsystem a set of control units that attach common I/O devices.

logical partition (LPAR). A subset of the processor hardware that is defined to support the operation of a system control program (SCP).

logical processor. In LPAR mode, central processor resources defined to operate in an LPAR like a physical central processor.

logical unit (LU). In SNA, a port to the network through which an end user accesses the SNA network and the functions provided by system services control points (SSCPs). An LU can support at least two sessions - one with an SSCP and one with another LU – and may be capable of supporting many sessions with other LUs.

logically partitioned (LPAR) mode. A central processor complex (CPC) power-on reset mode that enables use of the PR/SM feature and allows an operator to allocate CPC hardware resources (including central processors, central storage, expanded storage, and channel paths) among LPARs.

LU. Logical unit

M.

MAC. Message Authentication Code

main storage. Program-addressable storage from which instructions and other data can be loaded directly into registers for subsequent processing.

maintenance change level (MCL). A change to correct a single licensed internal code design defect. Higher quality than a patch, and intended for broad distribution. Considered functionally equivalent to a software PTF.

Manage suite (Manage). The first suite of functionality associated with the IBM zEnterprise Unified Resource Manager. The Manage suite includes core operational controls, installation, and configuration management, and energy monitoring.

management TOR switch. A top-of-rack switch that provides a private network connection between an IBM zEnterprise 196 (z196) or IBM zEnterprise 114 (z114) Support Element and an IBM zEnterprise BladeCenter Extension (zBX).

Mb. Megabit

MB. Megabyte

MBA. Memory bus adapter

MCL. Maintenance Change Level

megabit (Mb). A unit of measure for storage size. One megabit equals 1,000,000 bits.

megabyte (MB). A unit of measure for storage size. One megabyte equals 1,048,576 bytes. Loosely, one million bytes.

menu bar. The area at the top of the primary window that contains keywords that give users access to actions available in that window. After users select a choice in the action bar, a pulldown menu appears from the action bar.

MIDAW. Modified Data Indirect Address Word

MIF. Multiple Image Facility

modem. A device that converts digital data from a computer to an analog signal that can be transmitted on a telecommunication line, and converts the analog signal received to data for the computer.

Multiple Image Facility (MIF). A facility that allows channels to be shared among PR/SM LPARs in an ESCON or FICON environment.

multichip module (MCM). The fundamental processor building block for System z. Each System z "book" is comprised of a glass ceramic multichip

module of processor units (PUs) and memory cards, including multilevel cache memory.

multiplexer channel. A channel designed to operate with a number of I/O devices simultaneously. Several I/O devices can transfer records at the same time by interleaving items of data.

MVS™. Multiple Virtual Storage

MVS image. A single occurrence of the MVS/ESA operating system that has the ability to process work.

MVS system. An MVS image together with its associated hardware, which collectively are often referred to simply as a system, or MVS system.

N.

NetBIOS. Local area network basic input/output system.

network. An arrangement of nodes and connecting branches.

A configuration of data processing devices and software connected for information exchange.

node. A single IBM zEnterprise 196 (z196) or IBM zEnterprise 114 (z114) and any optionally attached IBM zEnterprise BladeCenter Extensionion (zBX). A node can be a member of only one ensemble.

See also central processor complex.

O.

On/Off Capacity on Demand (On/Off CoD). Used to temporarily turn on CPs, IFLs, ICFs, zIIPs, zAAPs, and SAPs.

operating system (OS). Software that controls the execution of programs and that may provide services such as resource allocation, scheduling, input/output control, and data management. Although operating systems are predominantly software, partial hardware implementations are possible.

optical cable. A fiber, multiple fibers, or a fiber bundle in a structure built to meet optical, mechanical, and environmental specifications.

optical fiber. Any filament made of dielectric materials that guides light, regardless of its ability to send signals.

optimizer. A special-purpose hardware component or appliance that can perform a limited set of specific functions with optimized performance when compared to a general-purpose processor. Because of its limited set of functions, an optimizer is an integrated part of a processing environment, rather than a standalone unit.

One example of an optimizer is the IBM Smart Analytics Optimizer for DB2 for z/OS.

OSA. Open Systems Adapter (OSA-Express4S, OSA-Express3, and OSA-Express2). The OSA is an integrated hardware feature that provides direct connection to clients on local area networks (LANs).

OSA/SF. Open Systems Adapter/Support Facility

P.

parallel channel. A channel having a S/360 and S/370 channel-to-control-unit I/O interface that uses bus-and-tag cables as a transmission medium.

A data path along which a group of signals representing a character or any other entity of data can be sent simultaneously.

Parallel Sysplex. A set of z/OS systems communicating and cooperating with each other through certain multisystem hardware components and software services to process customer workloads.

PIN. Personal identification number

PCI-IN card. The PCI-IN card can only be used in the PCIe I/O drawer. The PCI-IN cards provide the intraconnection from the PCIe I/O drawer to the PCIe fanout card in a book or processor drawer.

PKA. Public-key-algorithm

platform management. The subset of systems management focused on hardware and virtualization management.

point-to-point channel path configuration. In an I/O interface, a configuration that consists of a single link between a channel and one control unit.

point-to-point connection. A connection established between two data stations for data transmission.

Note: The connection may include switching facilities.

POR. Power-on reset

power-on reset (POR). A function that reinitializes all the hardware in the system and loads the internal code that enables the machine to load and run an operating system. This function is intended as a recovery function.

power-on reset state. The condition after a machine power-on sequence and before an IPL of the control program.

PowerVM. See PowerVM Enterprise Edition.

PowerVM Enterprise Edition (PowerVM). A hypervisor that provides a set of comprehensive systems technologies and services designed to enable aggregation and management of IBM POWER blade resources through a consolidated, logical view.

primary HMC. The System z Hardware Management Console (HMC) through which data personnel create and manage an ensemble. This HMC owns configuration and policy information that the Unified Resource Manager uses to monitor, manage, and adjust resources for all members of this ensemble.

See also alternate HMC.

processor. In a computer, a functional unit that interprets and executes instructions. A processor consists of at least an instruction control unit and an arithmetic and logic unit.

The boundaries of a system, exclusive of I/O control units and devices, that can be controlled by a single operating system. A processor consists of main storage, one or more central processors, time-of-day clocks, and channels, which are, or can be, placed in a single configuration. A processor also includes channel subsystems, and expanded storage where installed.

processor complex. A system configuration that consists of all the machines required for operation; for example, a processor unit, a processor controller, a system display, a service support display, and a power and coolant distribution unit.

Processor Resource/Systems Manager (PR/SM). The feature that allows the processor to use several system control programs (SCPs) simultaneously, provides logical partitioning capability for the real machine, and provides support for multiple preferred guests.

processor unit (PU). A PU can be defined as a CP, ICF, IFL, zIIP, zAAP or spare SAP.

program. Sequence of instructions for a computer. A program interacts and relies on either the hardware or other programs.

program status word (PSW). An area in storage used to indicate the sequence in which instructions are executed, and to hold and indicate the status of the computer system.

program temporary fix (PTF). A temporary solution or bypass of a problem diagnosed by IBM as resulting from a defect in a current, unaltered release of the program.

PR/SM. Processor Resource/Systems Manager

PSC. Power sequence controller

PSP. Preventive service planning

PSW. Program status word

PTF. Program temporary fix

processor unit (PU). A PU can be defined as a CP, ICF, IFL, zIIP, zAAP or spare SAP.

R.

rack. A free-standing structure or frame that can hold multiple servers and expansion units, such as BladeCenter blades.

RAS. Reliability, availability, serviceability

reconfiguration. A change made to a given configuration in a computer system; for example, isolating and bypassing a defective functional unit or connecting two functional units by an alternative path. Reconfiguration is effected automatically or manually and can be used to maintain system integrity.

The process of placing a processor unit, main storage, and channels offline for maintenance, and adding or removing components.

recovery. To maintain or regain system operation after a failure occurs. Generally, to recover from a failure is to identify the failed hardware, to deconfigure the failed hardware, and to continue or restart processing.

Remote Service Facility (RSF). A control program plus associated communication equipment that allows local personnel to connect to an IBM service center, and allows remote personnel to operate the remote system or send new internal code fixes to it, if properly authorized.

A system facility invoked by Licensed Internal Code that provides procedures for problem determination and error detection.

Remote Technical Assistance and Information Network (RETAIN). A database, accessible to service representatives, of information relating to IBM-installed products.

RETAIN. Remote Technical Assistance and Information Network

REXX. Restructured Extended Executor language

ring network. A network configuration in which devices are connected by unidirectional transmission links to form a closed path.

Note: A ring of an IBM token-ring network is referred to as a LAN segment or as a token-ring network segment.

RMF. Resource Measure Facility

RPQ. Request for price quotation

RPS. Rotational positional sensing/sensor

RSA. Rivest-Shamir-Adelman

RSF. Remote Support Facility

S.

SAD. System Activity Display

SAP. System Assist Processor

SCSI. Small Computer System Interface

SDR. Single data rate

Server Time Protocol (STP). A message based protocol designed to enable multiple servers to maintain time synchronization with each other. The timekeeping information is passed over data links (externally defined coupling links) between servers. It provides time synchronization for the z196, z114, z10 EC, z10 BC, z9 EC, z9 BC, z990, and z890 servers and CFs without requiring the Sysplex Timer.

service representative. A person who performs maintenance services for IBM hardware products or systems.

SIE. Start Interpretive Execution

single point of control. The characteristic a Parallel Sysplex displays when you can accomplish a given set of tasks from a single workstation, even if you need multiple IBM and vendor products to accomplish that particular set of tasks.

single system image. The characteristic a product displays when multiple images of the product can be viewed and managed as one image.

SNA. Systems Network Architecture

SNA network. The part of a user application network that conforms to the formats and protocols of Systems Network Architecture. It enables reliable transfer of data among end-users and provides protocols for controlling the resources of various network configurations. The SNA network consists of network addressable units (NAUs), boundary function components, and the path control network.

SNMP. Simple network management protocol

STI. Self-Timed Interconnect

STI-MP (Self-Timed Interconnect Multiplexer). For System z9®, used for an I/O cage intraconnection.

STP. Server Time Protocol

storage. A functional unit into which data can be placed, in which they can be retained, and from which they can be retrieved.

subchannel. In 370-XA, ESA/390 modes, and z/Architecture modes, the facility that provides all of the information necessary to start, control, and complete an I/O operation.

subchannel number. A system-unique 16-bit value used to address a subchannel.

subsystem. A secondary or subordinate system, or programming support, usually capable of operating independently of or asynchronously with a controlling system.

subsystem storage. See cache.

Support Element. An internal control element of a process operational functions.

A hardware unit that provides communications, monitoring, and diagnostic functions to a central processor complex (CPC).

Sysplex Timer. An IBM unit that synchronizes the time-of-day (TOD) clocks in multiple processors or processor sides. External Time Reference (ETR) is the MVS generic name for the IBM Sysplex Timer.

system. Comprises the processor complex and all attached and configured I/O and communication devices.

system area. A logical area of central storage used to store Licensed Internal Code and control information (not addressable by application programs).

Systems Network Architecture (SNA). The description of the logical structure, formats, protocols, and operational sequences for transmitting information units through, and controlling the configuration and operation of, networks.

S/370. IBM System/370

S/390. IBM System/390®

T.

target processor. The processor that controls execution during a program restart, instruction trace, standalone dump, or IPL, and whose ID is identified by highlighting on the status line.

TCP/IP. Transmission Control Protocol/Internet Protocol

TDES. Triple Data Encryption Standard

time-of-day (TOD) clock. A system hardware feature that is incremented once every microsecond, and provides a consistent measure of elapsed time suitable for indicating date and time. The TOD clock runs regardless of whether the processor is in a running, wait, or stopped state.

timing-only links. Coupling links that allow two servers to be synchronized using STP messages when a coupling facility does not exist at either end of the coupling link.

TKE. Trusted Key Entry

TOD. Time of day

token. A sequence of bits passed from one device to another on the token-ring network that signifies permission to transmit over the network. It consists of a starting delimiter, an access control field, and an end delimiter. The access control field contains a bit that indicates to a receiving device that the token is ready to accept information. If a device has data to send along the network, it appends the data to the token. When data is appended, the token then becomes a frame.

token-ring network. A ring network that allows unidirectional data transmission between data stations, by a token passing procedure, such that the transmitted data return to the transmitting station.

A network that uses ring topology, in which tokens are passed in a circuit from node to node. A node that is ready to send can capture the token and insert data for transmission.

Note: The IBM token-ring network is a baseband LAN with a star-wired ring topology that passes tokens from network adapter to network adapter.

top-of-rack (TOR) switch. A network switch that is located in the first rack of an IBM zEnterprise BladeCenter Extension (zBX).

TOR switch. See intraensemble data network (IEDN) TOR switch and management TOR switch.

TPF. Transaction processing facility

transaction. A unit of processing consisting of one or more application programs, affecting one or more objects, that is initiated by a single request.

transaction processing. In batch or remote batch processing, the processing of a job or job step. In interactive processing, an exchange between a terminal and another device that does a particular action; for example, the entry of a customer's deposit and the updating of the customer's balance.

transaction. A unit of processing consisting of one or more application programs, affecting one or more objects, that is initiated by a single request.

U.

Unified Resource Manager. See IBM zEnterprise Unified Resource Manager.

user interface. Hardware, software, or both that allows a user to interact with and perform operations on a system, program, or device.

V.

VLAN. Virtual Local Area Network

VSE. Virtual Storage Extended

W.

workload. A collection of virtual servers that perform a customer-defined collective purpose. A workload generally can be viewed as a multi-tiered application. Each workload is associated with a set of policies that define performance goals.

workstation. A terminal or microcomputer, usually one that is connected to a mainframe or network, at which a user can perform applications.

Z.

z/OS discovery and autoconfiguration (zDAC). z/OS function for FICON channels designed to detect a new disk or tape device and propose configuration changes for the I/O definition file (IODF). This applies to all FICON channels supported on that are configured as CHPID type FC.

zAAP. See IBM System z Application Assist Processor.

zBX. See IBM zEnterprise BladeCenter Extension (zBX).

zBX blade. See IBM zEnterprise BladeCenter Extension (zBX) blade.

zCPC. The physical collection of main storage, central processors, timers, and channels within a zEnterprise mainframe. Although this collection of hardware resources is part of the larger zEnterprise central processor complex, you can apply energy management policies to the zCPC that are different from those that you apply to any attached IBM zEnterprise BladeCenter Extension (zBX) or blades.

See also central processor complex.

zIIP. See IBM System z Integrated Information Processor.

z10 BC. IBM System z10 Business Class

z10 EC. IBM System z10 Enterprise Class

z114. IBM zEnterprise 114

z196. IBM zEnterprise 196

z800. IBM eServer zSeries 800

z890. IBM eServer zSeries 890

z900. IBM eServer zSeries 900

z990. IBM eServer zSeries 990

z9 BC. IBM System z9 Business Class

z9 EC. IBM System z9 Enterprise Class

Index

Special characters

Numerics

A

B

C

D

E

F

G

H

I

T

U

V

W

Z

Printed in USA

SA22-1086-03