

IBM Mainframe Bits: Understanding the Platform Hardware

Keith Winnard

Luiz Fadel

Rob Hunt

Jo Johnston

Alvaro Salla

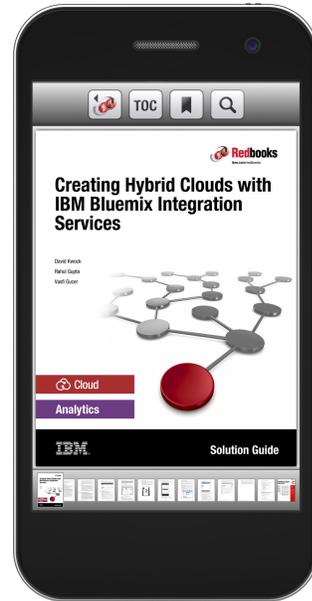


z Systems

Find and read thousands of IBM Redbooks publications

- ▶ Search, bookmark, save and organize favorites
- ▶ Get personalized notifications of new content
- ▶ Link to the latest Redbooks blogs and videos

Get the latest version of the Redbooks Mobile App



Promote your business in an IBM Redbooks publication

Place a Sponsorship Promotion in an IBM® Redbooks® publication, featuring your business or solution with a link to your web site.

Qualified IBM Business Partners may place a full page promotion in the most popular Redbooks publications. Imagine the power of being seen by users who download millions of Redbooks publications each year!



ibm.com/Redbooks
About Redbooks → Business Partner Programs

THIS PAGE INTENTIONALLY LEFT BLANK

IBM Mainframe Bits: Understanding the Platform Hardware

This IBM® Redpaper™ publication describes the hardware components and three high-level scenarios of how the IBM z Systems™ platform is used within organizations. The components and the configurations are designed and built to conform to the z/Architecture.

This paper includes the following topics:

- ▶ Introduction
- ▶ Hardware goals
- ▶ Brief tour of the CPC
- ▶ Input/output operations
- ▶ Configuration scenarios

Introduction

The IBM z Systems hardware is a physical implementation of the z/Architecture. The hardware components that are shown in Figure 1 are associated with the IBM z Systems platform.

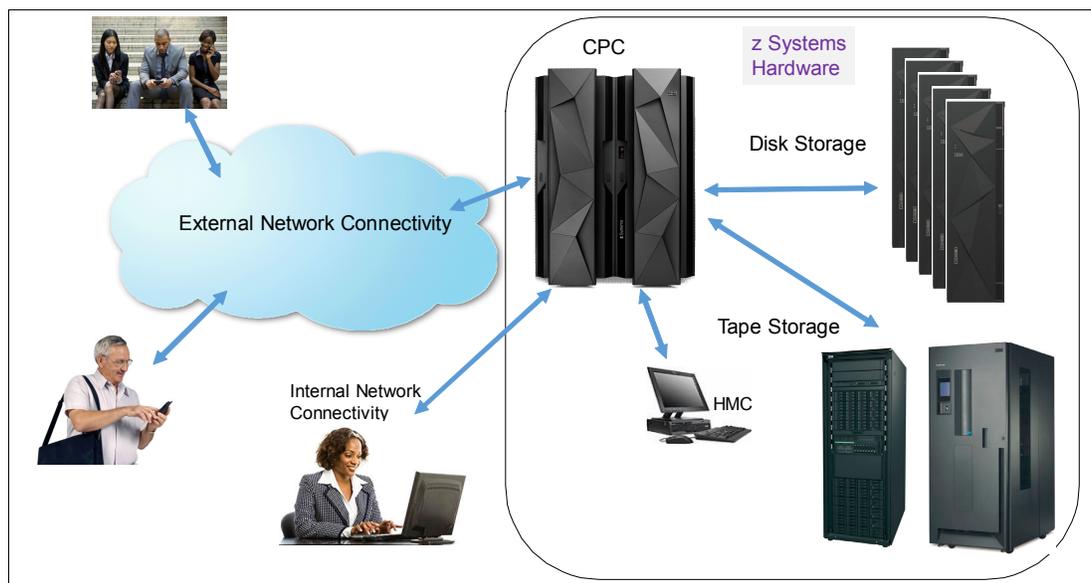


Figure 1 Hardware overview

The following key elements are featured:

- ▶ Connectivity

Enables requests to be sent or submitted and processed. If appropriate, a response to that request can be returned to the requester or sent elsewhere.

- ▶ Compute

The request typically requires some form of application processing that occurs within the processing unit (PU).

- ▶ Storage

Data that supports the organization's business processes and is related to requests and responses is stored on peripheral devices, such as disks or tapes.

Most computer systems provide connections to a network, both local area networks (LANs) and wide area networks (WANs). Connectivity allows users to access the system via different types of network connectivity, such as the internet. The users can connect from a terminal, notebook, mobile device, tablet, wearable technology, or other forms of recognized devices that can process the request and the response, such as text, video, and sound.

A Central Processor Complex (CPC) is shown in Figure 1 on page 1. It contains the electronic circuitry that carries out the instructions of a computer program by performing the arithmetic, logical, and control operations specified by the program's instructions.

Each request to the CPC activates at least one computer program to process the request. The user does not need to be aware of which computer programs are run and is only interested in the result; that is, the request that is received, processed, and the wanted result that is achieved.

A *computer program* is a sequence of instructions that is written to perform a specific task within a computer. The instruction within a program is a coded command to the processing unit to perform a specified function as defined by the processing unit instruction set.

Computer programs are written in programming languages that contain instructions in mnemonic form (for example, Assembler), higher-level languages (for example, COBOL), or script-like languages, which are much closer to English.

Programs go through compilation or translation software that converts the programming languages commands into valid instructions, which are within the scope of the z/Architecture instruction set.

The data that is needed to perform the processing task is retrieved from its location and brought into the memory so that the processing unit can access it. The result of the computation is then created and directed to its destination.

Memory is the electronic component of the CPC that hosts instructions and data that is necessary to perform tasks that are required by the program.

A *storage device* is used to temporarily and permanently store data and information. Two types of storage devices are shown in Figure 1 on page 1.

Hardware goals

The hardware components conform to the z/Architecture so they can work in unison to meet the following goals:

- ▶ Integrity of data

If the data that is stored and retrieved is inaccurately or incorrectly modified during its processing, the resulting inconsistency means that the data lost its integrity and cannot be trusted as being valid. Therefore, the data loses its value and jeopardizes the business or organization.

A key hardware function is to validate the data it is reading or writing and ensure that it maintains integrity throughout the process. If a physical component fails, the hardware might be able to reconstruct the data. Here, we are referring to the physical aspects of the reading, writing, and movement of the data between hardware components.

Note: The validation step is to ensure that the hardware reads and writes the data that it is given. It cannot validate the data from a business perspective; that process is the responsibility of the computer program.

- ▶ Availability of resources

The growth of the internet and personal digital devices increased the demand on computer systems. Hardware is designed to have resilience and provide availability levels, depending on customer requirements and configuration options.

- ▶ Performance

Hardware must perform its duties in a period that meets the expectations of its requesters. To meet the requester's expectations, the hardware must do the following tasks:

- Understand which resources are needed
- Understand the priority of the request
- Allocate resources to requests
- Monitor the progress to complete the request
- Reallocate resources as and when necessary to higher priority requests

- ▶ Security

Protection from intrusion is paramount. Mechanisms must be in place to detect tampering, provide encryption, and safeguard against unauthorized access. The hardware is securable to various levels of protection.

- ▶ Communication

The hardware must communicate with other systems by using industry standard protocols that enable the exchange of information or data.

- ▶ Flexible configuration

As needs change, so might the computer system configuration. The option to deploy specialized processors, add new storage devices, connect to other systems, and reconfigure dynamically to meet the business demands is essential. All of these processes must be done with minimal or no disruption to services. This activity often is accomplished by using the Support Element (SE) or the Hardware Management Console (HMC).

- ▶ **Scaling to meet demand**

Longer term trends that cause increases in demand for computing resources can be predicted. However, spikes in demand do occur, which cause the need to dynamically allocate further computing resources. The dynamic allocation of these extra resources to solve a spike in demand must be balanced such that the spike does not move to another resource type. Provisioning can be achieved by adding specific resources or through a combination of resources that might provide a cloud solution.

- ▶ **Virtualization of resources**

The IBM z Systems™ platform features unparalleled levels of virtualization at the hardware and software levels. Virtualization provides resources optimization and introduces the granularity that is required by organizations to create and separate environments as needed to perform different functions.

- ▶ **Compatibility**

New hardware that adheres to the z/Architecture includes component compatibility and is designed to offer enrichment through new capabilities to applications.

- ▶ **Control**

Control points to monitor, configure, maintain, and add components must be available to record and take actions where appropriate on the hardware.

- ▶ **Operating system integration**

IBM z Systems hardware integrates with several operating systems and must service requests from each operating system and allocate resources according to each operating system's needs.

The goals are achieved in many different ways and the hardware components combine within the architecture to provide a physical implementation of the architecture.

Brief tour of the CPC

The CPC is positioned to process high volumes of online transactions, enable business processes, perform real-time analytics, and encryption. Operational flexibility allows you to dynamically manage resources and allocate them to where they are needed to satisfy business application needs.

Inside the CPC

The model that is shown in Figure 2 on page 5 is an air-cooled CPC. A water-cooled option is available that features a slightly different look. When ordering your CPC, you choose the options that you want to include that are based on your business needs and environment.

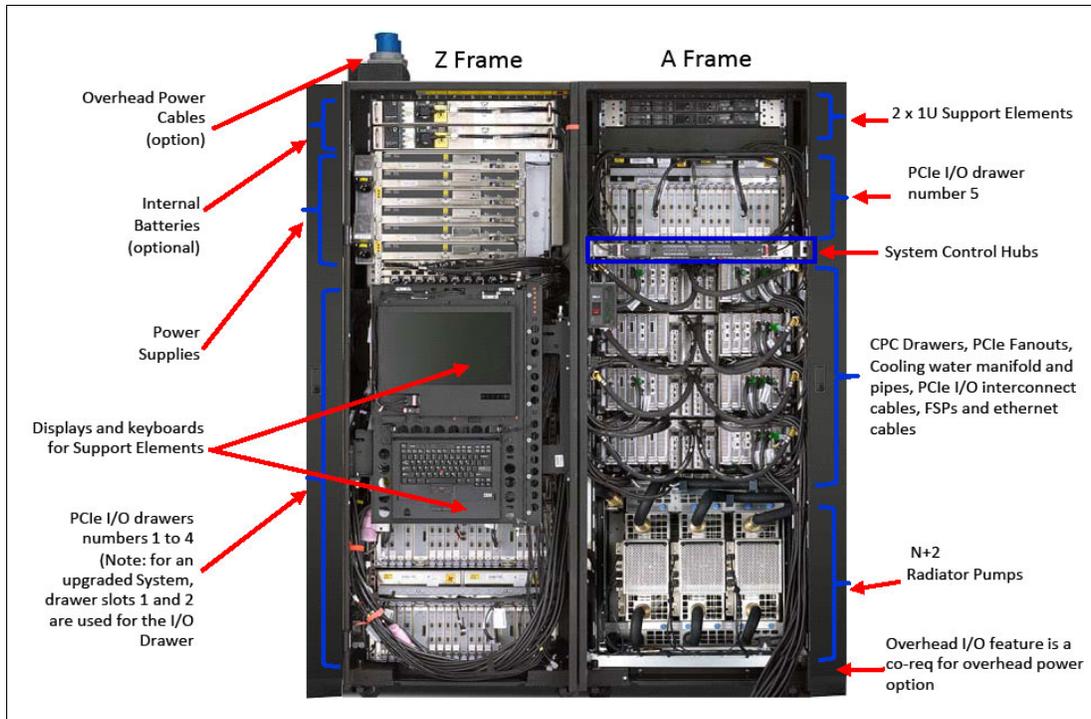


Figure 2 CPC components

Figure 2 shows the platform with the doors open, which makes it possible to see the following main components of the internal structure:

- ▶ Overhead power cables

These cables are optional. Alternatively in your environment, you can run your power cables under your machine room floor.
- ▶ Internal batteries

If there is a loss of power (for example, power to computer room was lost), the internal battery feature (IBF) provides temporary power to the processor to preserve the data. If the power loss is for an extended time, the IBF provides enough power for an orderly shutdown to be completed.
- ▶ Power supplies

The platform operates with power supplies that provide redundancy.
- ▶ Notebook that is connected to the SE

By using the displays and keyboards, you can communicate with the SE.
- ▶ PCIe drawers 1 - 4

These drawers include cards with which the CPC can connect to devices, such as disk and tapes. For more information about the input/output (I/O), see "Input/output operations" on page 12.
- ▶ PCIe I/O drawer 5

This drawer is an extra drawer to PCIe drawers 1 - 4.
- ▶ SEs

The SEs often are connected to the HMC. Together, the SEs and HMC provide hardware management, such as defining how memory is used or processors are assigned. There are two SEs for redundancy.

- ▶ System Control Hubs
The System Control Hubs control the elements for the platform and provide redundancy.
- ▶ CPC drawers
The CPC drawer contains the processors. For more information, see Figure 3 on page 7.
- ▶ Radiator Pumps
The Radiator Pumps are part of the cooling system.

Processing Units

All IBM z Systems CPCs are assigned a unique machine type. Each machine type is available in different models. Each model has a different number of PUs, memory, and other resources.

Understanding PUs

The model that you choose depends on how many PUs you require to run your workloads. There are different types of PUs. You must also define the type of PU you want (they do not have to be all the same type). You can have a combination of different types of PU.

The PUs essentially are physically the same, but differentiated by their characteristics. The PUs can be characterized in advance or dynamically. Certain PU characterizations are better-suited to specific types of tasks than others. The following PU characterizations used:

- ▶ Central processor (CP)
- ▶ Integrated Facility for Linux (IFL) processor
- ▶ z Integrated Information Processor (zIIP)
- ▶ Internal Coupling Facility (ICF)
- ▶ System assist processor (SAP)
- ▶ Integrated Firmware Processor (IFP)

Note: A PU that is not characterized cannot be used.

The following PUs are part of the base system; that is, they are not part of customer purchasable PUs and are characterized by default:

- ▶ System assist processor (SAP)
SAPs offload and manage I/O operations. Several SAPs are standard with the CPC. More SAPs can be configured if increased I/O processing capacity is needed.
- ▶ IFP
The IFP is a single PU that is dedicated to supporting native PCIe features (10 GbE RoCE Express and zEDC Express).
- ▶ Spare PUs
Spare PUs can transparently assume any characterization if a characterized PU permanently fails.

In contrast, customer purchasable PUs can assume any of the following characterizations:

- ▶ CP
This standard processor is for use with any supported operating system and user applications.

- ▶ IFL

This PU is exclusively used with Linux on z Systems and for running the IBM z/VM® hypervisor in support of Linux. z/VM is often used to host multiple Linux virtual machines (known as *guests*).
- ▶ IBM System z® Integrated Information Processor

zIIP is used under IBM z/OS® for designated workloads, which include IBM Java virtual machine (JVM), various XML System Services, IPsec offload, certain parts of IBM DB2® DRDA®, star schema, IBM HiperSockets™ for large messages, and the IBM GBS Scalable Architecture for Financial Reporting.
- ▶ ICF

ICFs are used for z/OS clustering. They are dedicated to this function and exclusively run the Coupling Facility Control Code (CFCC).
- ▶ An extra SAP can be used by the channel subsystem.

CP Assist for Cryptographic Function

The platform offers the cryptographic assist implementation, which is known as CP Assist for Cryptographic Function (CPACF). CPACF offers the full complement of the Advanced Encryption Standard (AES) algorithm and Secure Hash Algorithm (SHA) along with the Data Encryption Standard (DES) algorithm. CPACF must be explicitly enabled by using a no-charge enablement feature. The exception is the SHAs, which are shipped enabled with each server.

CPC processor drawer

The CPC that is shown in Figure 2 on page 5 features processor drawers that are fully interconnected. Each drawer contains two nodes. Each node contains Single Chip Modules (SCMs) for processor and storage control, memory, connectors to the PCIe I/O drawers, I/O drawers, and coupling link connectors. The processor drawer is shown in Figure 3.

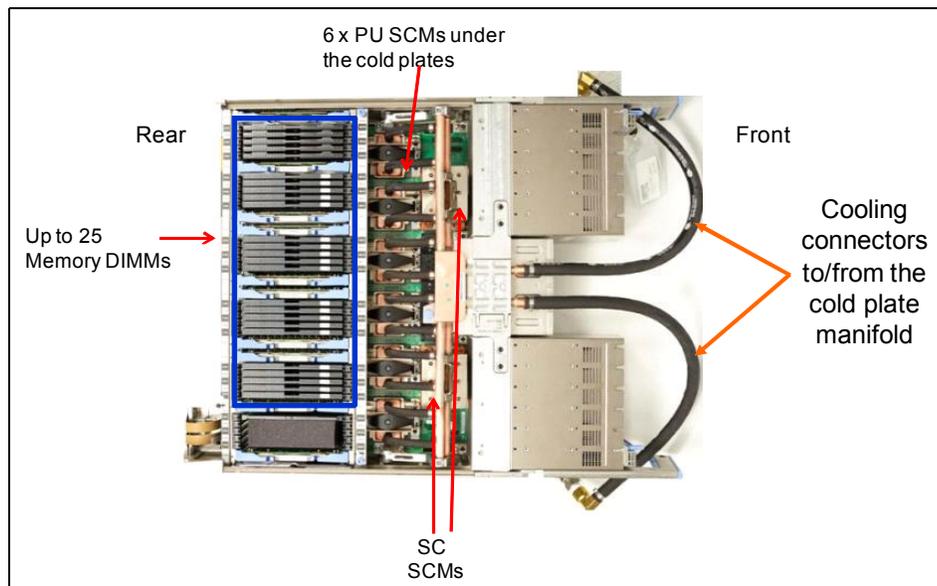


Figure 3 CPC processor drawer

This CPC processor drawer contains eight SCMs (six hold the PU chips). Each PU chip has six, seven, or eight active cores that hold Level 1, Level 2, and Level 3 caches. The remaining two SCMs are the storage control (SC) that hold the Level 4 caches.

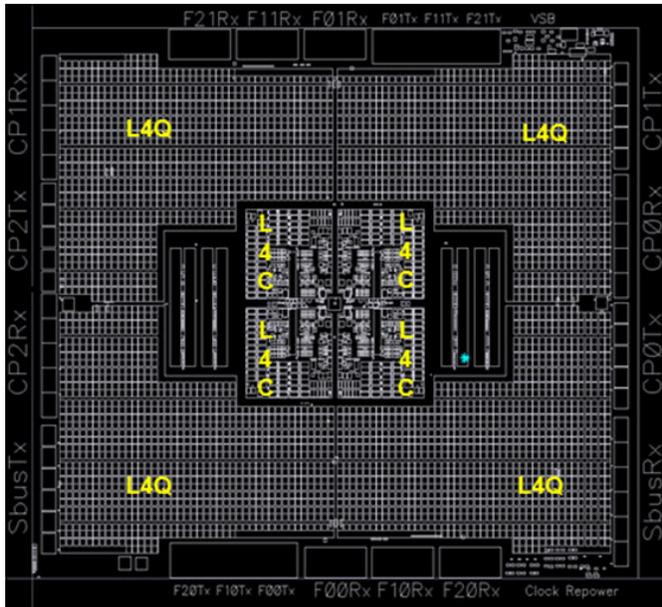


Figure 5 Storage Control Chip

A fixed hardware system area (HSA) is managed separately from client purchased memory. It is a reserved part of the system memory, contains the I/O configuration, and is used by the SAPs to manage the I/O.

Microcode

Other processors and microcode perform many functions, such as controlling and synchronizing the interaction between the hardware components. More information about these processors and microcode is not available because they form part of the internal circuitry and processing of the CPC.

Specialty processors

This section describes two types of special processors.

Cryptographic processors

Transactions and sensitive data require higher securable options. The CPC has multiple cryptographic levels. The first level is the cryptographic assist implementation, which is known as *CPACF*. It delivers cryptographic and hashing functions in support of clear-key operations.

It is physically implemented in the chip by the compression and cryptography accelerators. Each core has one dedicated coprocessor (CoP) that integrates the CPACF and the compression unit. The CPACF offers the full complement of the AES algorithm, SHA, and DES algorithm.

CPACF must be explicitly enabled by using a no-charge enablement feature. This requirement means that although the hardware has this feature, you must request its activation or it cannot be used. The CPACF is available for every PU that is characterized as CP, IFL, or zIIP.

Increased cryptography is available with the Crypto Express5S feature, and is a card that is installed in the PCIe I/O drawer. The Crypto Express5S offers a state-of-the-art, tamper-resistant cryptographic coprocessor for secure-key operations along with new hardware assists to encrypt data quickly. This capability allows data transfer across the internet to support public and private cloud and mobile workloads.

zEDC compression

zEDC compression is a mechanism for representing the same amount of information in a smaller number of bytes. This mechanism is an optional feature and provides hardware-based acceleration for data compression and decompression. It helps to improve cross platform data exchange, reduce CPU consumption, and save disk space. Similar to the Crypto Express 5S card, it is also installed in the PCIe drawers.

For resilience and to maintain availability, it is recommended that a minimum of two zEDC Express features be installed.

Simultaneous multithreading

Simultaneous multithreading (SMT) is a technique for improving the overall efficiency of certain PUs by implementing hardware multithreading. This capability is available on the PUs that were characterized as zIP or IFL. SMT permits multiple independent execution threads in a single PU to better use the resources.

Single instruction, multiple data

The CPC provides specialized hardware to improve the performance of complex mathematical models and analytic workloads through vector processing and new complex instructions. Vector processing and these complex instructions can process multiple data items by using only a single instruction. Single instruction, multiple data (SIMD) is ideal for processing data for analytical purposes and processes data faster than the single instruction, single data (SISD) because it can handle more data per instruction.

Figure 6 shows how SISD processing is more restricted than SIMD processing.

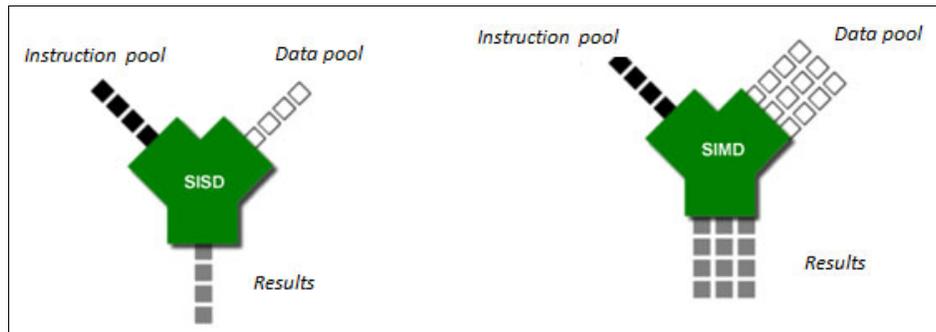


Figure 6 Single data versus multiple data

Virtualization

The ability to share resources is one of the major strengths of the mainframe. z/Architecture contains many facilities in its hardware and software that facilitate resource virtualization.

Overutilizing the mainframe environment involves creating virtual systems through logical partitions (LPARs) and assigning virtual resources, such as memory and I/O channels, to those systems. Depending on configuration options, resources can be dynamically added to or removed from these logical partitions.

Processor Resource/Systems Manager™ (PR/SM™) is a hypervisor that is integrated with all mainframe elements. PR/SM maps physical resources into virtual resources so that many logical partitions can share the physical resources.

PR/SM and logical partitions

PR/SM is a feature that enables logical partitioning of the CPC and provides isolation between partitions, which makes it possible to separate users into distinct processing images or to restrict user access to certain workloads (see Figure 7).

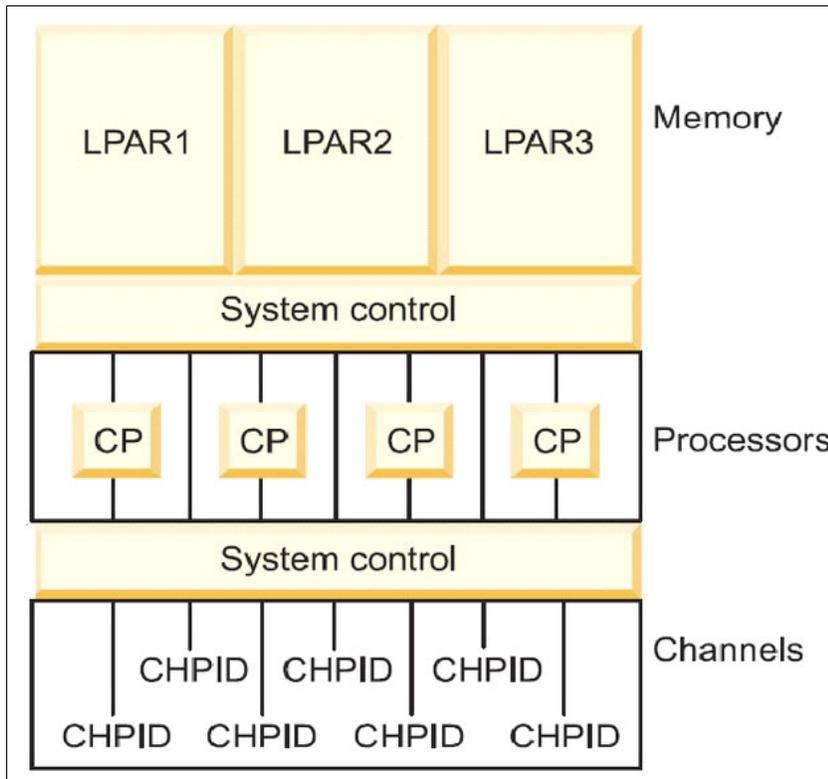


Figure 7 LPARs

In contrast to isolation, PR/SM also allows resource sharing between LPARs. Resources are defined and assigned a shareable attribute. The LPARs that can share the resources also are identified.

An LPAR (see Figure 7) features the following characteristics:

- ▶ A set of physical resources (processor, storage, and channels).
- ▶ LPARs that can be defined depends on the CPC model.

- ▶ Individual physical PUs that can be shared between multiple LPARs or dedicated for use by a single LPAR. There are considerations for mixing the PU types; not all combinations are possible. A CP or zIIP cannot be shared with an IFL because they are designed for different workloads.
- ▶ Channels that can be dedicated, reconfigured (dedicated to one LPAR but can be switched manually between LPARs), or shared.
- ▶ The CPC storage that used by an LP is dedicated, but can be reconfigured from one LP to another with planning and without stopping both logical partitions.

Input/output operations

Requests from the network and data that is stored on disks and tapes must be transported into memory before being selected for processing and after they are processed, they continue to their destination. This movement is known as an *Input/output (I/O) operation*.

Channel subsystem

The mechanisms that are used to control I/O operations are collectively called the *channel subsystem (CSS)*. The CSS directs the flow of information between I/O devices and main storage. A CSS relieves the PUs of the tasks of communicating directly with I/O devices and permits data processing to proceed on the PUs while other data is transferred concurrently to and from the I/O devices.

The CSS uses one or more channel paths by using a Channel Path ID (CHPID) as the communication link to manage the flow of information to and from I/O devices. I/O connectivity is shown in Figure 8 on page 13. Channels connect the CSS and Control Unit (CU) and the CU connects to the device.

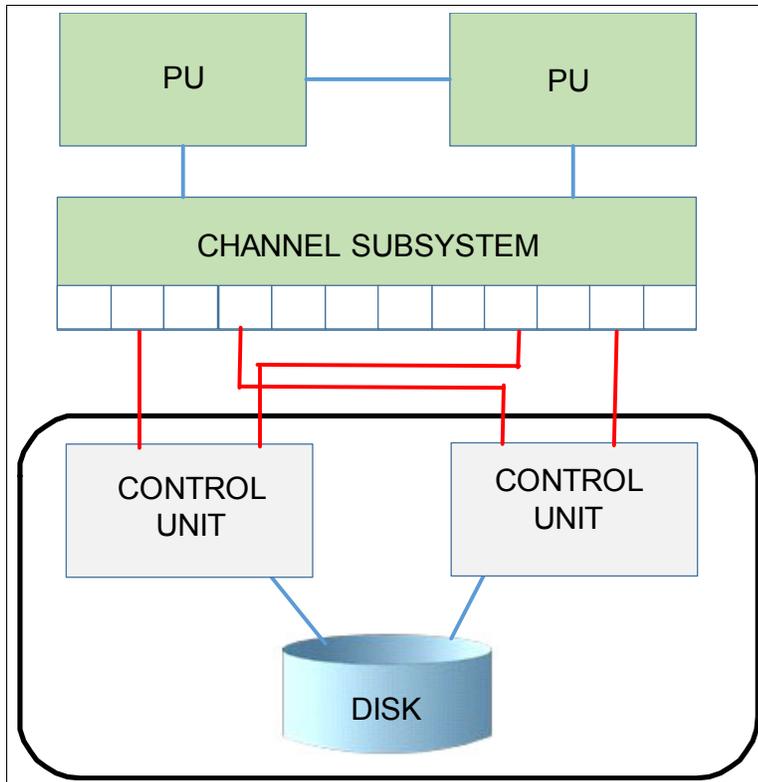


Figure 8 Path from CPC to I/O device

The channel subsystem is part of I/O processing. The channel subsystem performs a path-management operation by testing for channel path availability, choosing an available channel path, and starting the I/O operation. The following elements relate to the I/O journey:

- ▶ Channel Path: A channel path is a single interface between a system and one or more control units. A CSS has up to 256 channels attached to it.
- ▶ Subchannels: A subchannel provides the logical representation of an I/O device to the program and contains the necessary information to sustain an I/O operation to the device it represents.
- ▶ CHPID: The CSS communicates with I/O devices through channel paths between the CSS and the CUs. Each channel path is assigned an identifier that uniquely identifies that path.
- ▶ Physical channel ID (PCHID): Reflects the physical location of a channel type interface. The PCHID number is based on the following factors:
 - The PCIe I/O drawer, I/O drawer, or I/O cage location.
 - The slot number within the drawer.
 - The port number of the channel feature (such as, FICON and OSA). It is the port number on the I/O card.
- ▶ Control Unit: A CU provides the logical capabilities that are required to operate and control an I/O device and communicates as appropriate with the I/O device that is attached to it.
- ▶ I/O devices: The I/O device provides the external storage (such as disk) or a means of communication (such as a terminal) between data processing systems or its environment.

The performance of a CSS depends on its use and on the system model in which it is implemented. Channel paths are provided with different data transfer capabilities and an I/O device (such as a magnetic-tape unit or a disk storage) that is designed to transfer data at a specific rate.

Peripheral Component Interconnect Express

The Peripheral Component Interconnect Express (PCIe) Generation 3 (PCIe Gen3) provides connectivity. Figure 2 on page 5 shows the five PCIe drawers. I/O cards are placed into a slot within the I/O drawer. The I/O cards have ports; one end of the channel cable plugs into an I/O card port, the other end of the cable is typically plugged into the control unit of the device. You now have a full connection.

Direct access storage device

For IBM z Systems platforms, a disk is also known as a direct access storage device (DASD). Both terms are used in this paper and are interchangeable. A disk or DASD device can be also referred to as a volume.

Disk storage hardware is available in various models. These models include the following typical characteristics:

- ▶ Various types of devices available
- ▶ Storage tiers available
- ▶ Application to help define tiers
- ▶ Copy services
- ▶ Support Extended Address Volumes
- ▶ HMC to configure the storage
- ▶ High performance
- ▶ Multi-site business continuity
- ▶ PCIe Gen. 3 support

Devices and tiers

Often there are five types of devices that can be managed in the following tiers:

- ▶ Tier 1: Flash cards and flash drives
- ▶ Tier 2: SAS 10 k or 15 k RPM disk drives
- ▶ Tier 3: Nearline 7.2 k RPM disk drives

The drives within a tier must be homogeneous. If 10-K and 15-K RPM disk drives are in the same extent pool, they are managed as a single tier. Smart tiers can combine tiers, as shown in Figure 9. The combination of tiers and migration policies aim to achieve the optimum storage for data through its management cycle.

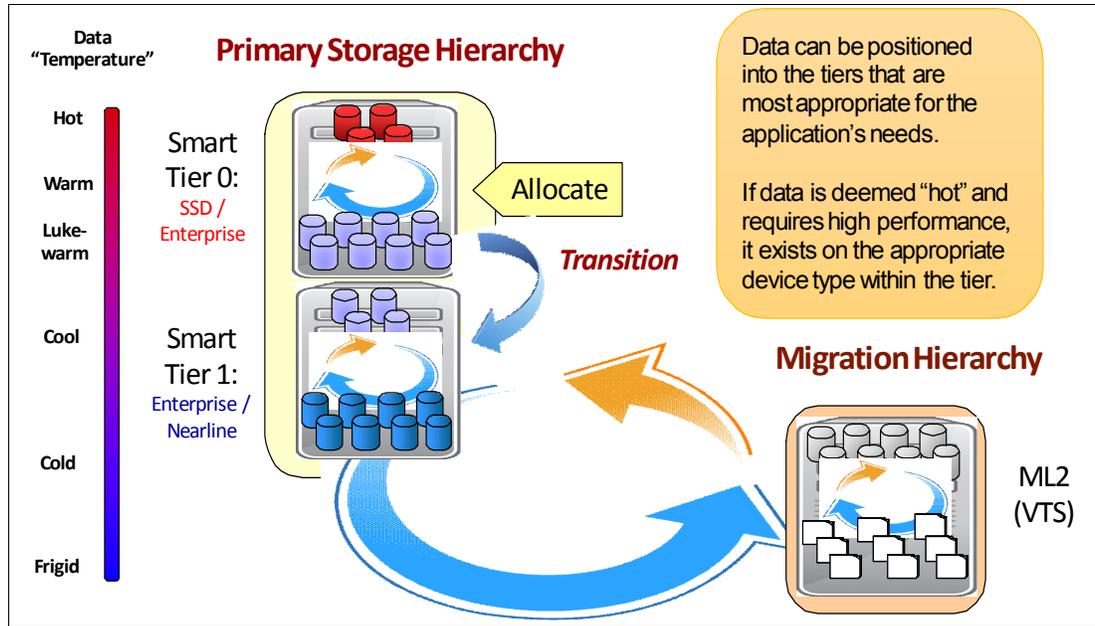


Figure 9 Data performance needs

The tiers can be set up to make the most effective use of the devices within the tier to help maintain the correct performance levels to access the data.

FICON

Storage connectivity is provided on the CPC by Fibre Channel (FC) connection features. IBM Fibre Channel connection (FICON) features follow the established FC standards to support data storage and access requirements. The FICON Express features support the following protocols:

- ▶ Native FICON

This enhanced protocol (over FC) provides for communication across channels, channel-to-channel (CTC) connectivity, and with FICON devices, such as disks, tapes, and printers. It is used in z/OS, z/VM, IBM z/VSE®, z/TPF, and Linux on z Systems environments.

- ▶ Fibre Channel Protocol (FCP)

This protocol is the standard for communicating with disk and tape devices through FC switches and directors. The FCP channel can connect to FCP SAN fabrics and access FCP/SCSI devices. FCP is used by z/VM, KVM for IBM z Systems, z/VSE, and Linux on z Systems environments.

IBM FICON Express features are implemented by using PCIe cards and offers better port granularity and improved capabilities over the previous FICON Express features. It also is the preferred technology for new systems.

Copy services

The copy services can have functionality to help you in the following ways:

- ▶ Copy data for backup and subsequent recovery, if required
- ▶ Migrate data from one tier to another
- ▶ Copy data at high speed to provide business continuity if there is an outage
- ▶ Duplicate data for testing purposes
- ▶ Copy data to a target area for reporting access

Note: Some of the copy functions that are described in this section are optional features.

The following types of copy services are available. There are specific attributes within each type of copy service:

- ▶ The IBM FlashCopy® function

You can make point-in-time copies of your data by using the FlashCopy function. You can choose to copy at the volume level or at the file (also known as a data set in z/OS) level. You can access the copy immediately to read data or update the data.

The function works by taking a source volume (the volume you want to copy from) and forming a relationship with a target volume (the volume to which you want to write the data). By using FlashCopy, the copy can occur quickly, which creates a *point-in-time* copy of the source volume (or data set) on the target volume. The target volume is immediately available for backup to tape, which might be stored off-site in a vault. You can perform a reverse restore that copies the target volume back to the source volume. There are other options available that are not described in this publication and are more advanced. The effect of an application accessing the source volume during FlashCopy is shown in Figure 10.

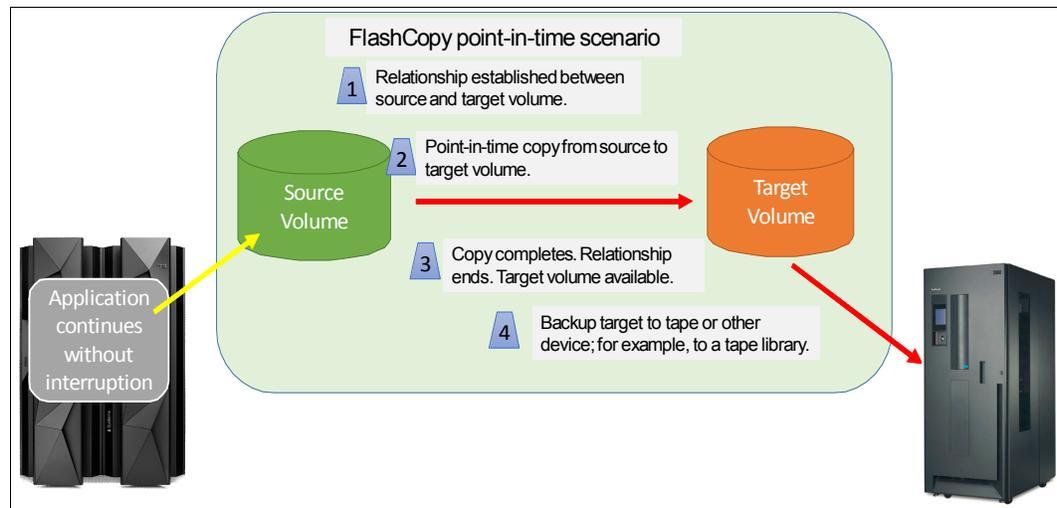


Figure 10 FlashCopy point-in-time

► Remote mirror and copy

Remote mirror and copy allows you to replicate data between a source volume and a target volume on 1- or 2-disk storage systems. The remote mirror and copy features the following functions:

– Metro mirror

Metro Mirror provides real-time mirroring of logical volumes between two disk storage systems that can be up to 300 km apart. This method is known as a *synchronous copy solution*, meaning that the data write operations are completed on both copies (local and remote) before they are considered to be completed, as shown in Figure 11. This method means that there are two copies of the data in separate locations.

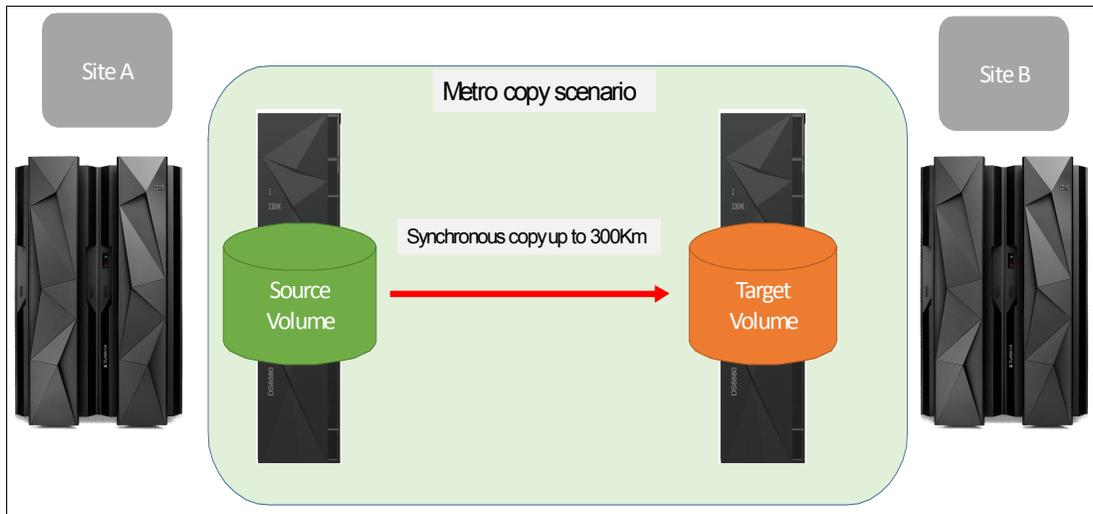


Figure 11 Simple Metro Mirror scenario

– Global copy

Global copy is an asynchronous copy over a larger distance than the metro copy function. By using global copy, the source volume periodically sends updates to the target volume. Although this method has less effect on the network bandwidth because it is using fewer sends, the source and target volumes do not always have the same data on them.

– Global Mirror

Global Mirror combines the global copy and FlashCopy functions. It mirrors data between volume pairs over longer distances without affecting overall performance. It copies to the set of target volumes every few seconds so that the data at the remote site is a point-in-time consistent copy of the data at the local site.

Protection

The disk subsystem can have the following securable functions:

- Self-encrypted disk drives
- Full Disk Encryption (FDE) that includes sophisticated key management software
- Tamper proof audit logging

Tapes

Tapes are magnetic storage devices that allow data to be stored and accessed sequentially. Data is typically stored on tape for the following reasons:

- ▶ Backup

Copies of volumes or data sets are written to tape and remain onsite or are moved to an off-site vaulting facility.

- ▶ Archive

The data might be kept for a time because of mandatory regulations. The data is no longer part of the day-to-day cycles of the business applications, so the access requirements to it are minimal or non-existent.

- ▶ Migration

Infrequently accessed data can be stored on tape instead of occupying disks. Therefore, the data might be accessed only monthly and does not require a high-performance medium.

The tapes are contained in special cartridges. These cartridges are placed in a tape drive and data is written or read from the tape. Multiple data sets can occupy the same tape. A software tape management system helps to manage the contents of each tape.

The tapes can be manually loaded and unloaded on the tape drive for use and stored in a manual tape library. The manual approach is impractical for larger numbers of tapes because of the manual overheads and increased chances of errors affecting the accuracy of the tape location.

The other option is to have an automatic tape library. The library includes designated shelving for storing the tapes. The tape drives are inside the library and robotic arms transport the tapes to and from their storage location on the shelves and the tape drives.

Some tapes are stored off-site, which is often referred to as *vaulting*. The tapes that are contained at the external location are often backups of disks and can be used for disaster recovery purposes.

The use of an automated tape library in preference to the manual method helps to provide a more consistent mount time of tapes and reduces the chances of losing tapes through misplacing them.

Virtual Tape Server

The Virtual Tape Server (VTS) is a means of storing data in tape format on disk within the VTS. The use of VTS over real tapes can be the preferred method for the following reasons:

- ▶ For certain types of data sets the VTS can optimize resource consumption
- ▶ Reduce contention on real tape devices
- ▶ Reduce wastage of real tape capacity

Data sharing

Data sharing is the capability that enables users to access the same data simultaneously. Different applications within a system often must access the same information. Data can be shared between LPARs and between CPCs. Sharing mechanisms are available to protect against any conflicts that result from sharing data.

Various sharing options are available for the channels within the same CSS and across more than one CSS. The CUs and I/O devices also must be declared as shared.

Without data sharing within your systems, you might need several copies of the data. Managing those copies might prove to be a complicated process. For example, if the system fails, you need a way to ensure that the most recently changed data is the master for all instances of that data.

Network connectivity

The network connectivity covers external connectivity and specialized internal connections for intra-system communication.

The IBM Open Systems Adapter-Express (OSA-Express) is a card that can be installed into the PCIe drawers. It can provide LAN connectivity. In addition, the OSA-Express can assume some of the PUs workload for several functions of the TCP/IP stack, which results in significant performance benefits by offloading processing from the operating system.

HiperSockets

HiperSockets is an integrated function of the IBM z Systems platforms that supplies attachments to virtual local area networks with minimal system and network overhead.

HiperSockets provides LAN connectivity across multiple system images on the same z Systems platform by performing memory-to-memory data transfers in a secure way. The HiperSockets function eliminates the use of I/O subsystem operations and the use of an external network connection to communicate between LPARs in the same CPC.

IBM 10 GbE RoCE Express

The 10 Gigabit Ethernet (10 GbE) RoCE Express feature uses Remote Direct Memory Access (RDMA) over Converged Ethernet (RoCE) to provide fast memory-to-memory communications between two CPCs or within the same CPC.

The feature is designed to help reduce consumption of CPU resources for applications that use the TCP/IP stack. It can also help reduce network latency with memory-to-memory transfers by using Shared Memory Communications - Remote Direct Memory Access (SMC-R) in z/OS V2R1 or later.

With SMC-R, you can transfer huge amounts of data quickly, at low latency. SMC-R is not apparent to the application and requires no code changes, which enables rapid time to value.

Shared Memory Communications - Direct Memory Access (SMC-D)

The platform features a communications protocol that is called Shared Memory Communications - Direct Memory Access (SMC-D). SMC-D is similar to SMC-R, but SMC-D is intended for communications *within* the same CPC, which optimizes operating systems communications in a way that is not apparent to socket applications. It reduces the CPU cost of TCP/IP processing in the data path, which enables highly efficient and application-transparent communications.

Configuration scenarios

Samples of three configurations are provided in this section to demonstrate how the z Systems platform might be deployed.

In the three scenarios that are described in this section, the IBM z Systems platform is designed to provide the hardware goals that are described in “Hardware goals” on page 3 and the system software layers for applications to provide the functionality and high-quality service to a business or organization’s customers. Without these fundamental goals, the functionality or services fail. The IBM z Systems platform is positioned and proven in meeting the challenges to provide the functionality and high-quality services.

Payroll scenario

In Figure 12, there are two applications running: payroll and the Management Information Reporting.

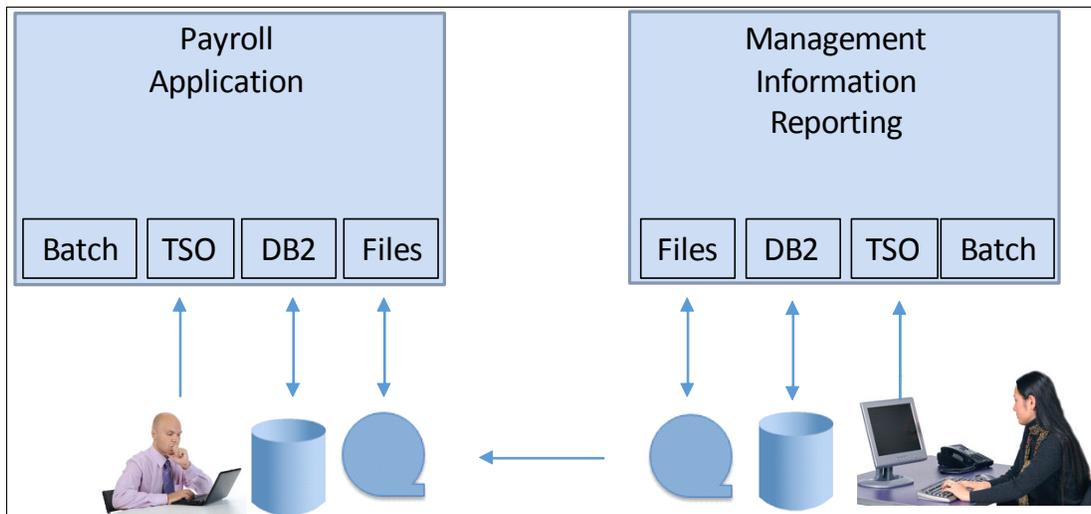


Figure 12 Payroll scenario

The batch (collection of programs) process data to calculate the base salary, overtime, and the benefits for employees. After those three processes are complete, the payroll data is backed up. After the backups are complete, a job is run to transfer the money to the bank.

Management Information Reports (MIR) detail all overtime that is worked by the employees and provides the input for the overtime reports. The overtime data is output to tape so that it can be transferred to the other LPAR. After this tape is available, the payroll can be run.

You might question why tape is used as a transfer media for the data and that the disks containing the DB2 databases with all the information in are not shared between the two LPARs. This issue might be because of regulatory restrictions.

Another use of tape for both applications that are shown in Figure 12 is to back up the data at certain points to provide a recovery point if there is a failure.

Retail scenario

This scenario is part of a retail processing system. The first mainframe is dedicated to receiving orders and queries from notebooks, tablets, mobile phones, and other internet sources. It might handle the queries and send the response. However, for more complicated queries and order placements, the first mainframe passes the request to the second mainframe. The second mainframe handles orders and queries that are passed to it via the first mainframe, and it also processes orders and queries that are coming from call centers and orders that are received via the mail. The retail scenario is shown in Figure 13.

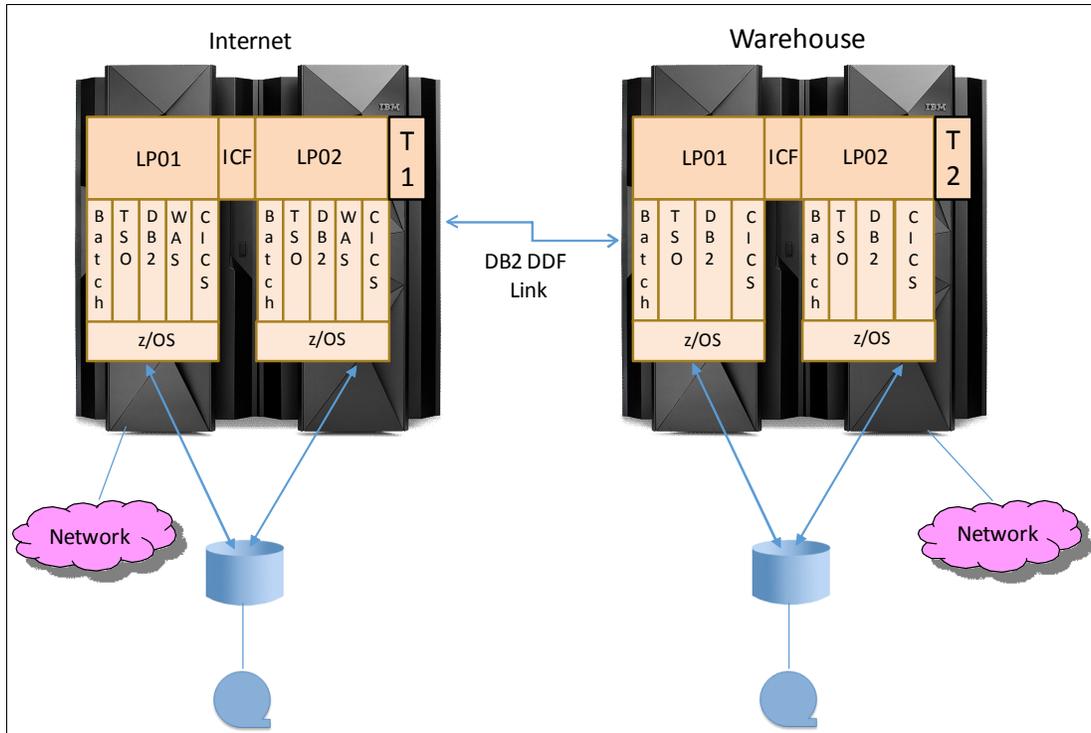


Figure 13 Retail scenario

A customer who is using the internet or a mobile phone accesses the Retailer's website to place an order for an item. This scenario assumes that the customer is going to order an item only. The time that it takes the customer to select and buy the item and receive confirmation that the order was processed must be minimal. The process also includes an SLA with an external service provider that checks the payment details and whether this transaction is fraudulent.

There are two LPARs per mainframe that include various elements that relate to system software, databases, and applications that are running on them. Both mainframes have different hardware configurations, so each can meet its specific needs. However, they work together to provide the following services to the customer:

- ▶ An online catalog that is constantly updated throughout the day with new and changed information about products, offers, and other information. The catalog consists of the following elements:
 - Thousands of graphic images that provide the option to zoom in on those images.
 - Videos to show the following elements:
 - The products in use
 - Instructional information
 - Help information for new and existing customers
 - Product descriptions for each product
- ▶ Daily order processing application, which takes orders from the internet, call centers, or mail with a common process that is sensitive to any special requirements, such as incentives to order via the internet.
- ▶ Order and customer vetting to assess the customer's account limits, and a series of anti-fraudulent processes that occur during the actual ordering transaction and after the actual order is placed.
- ▶ Delivery application, which includes address vetting, type of delivery, priority deliveries, and deliver to alternative addresses.
- ▶ Stock inventory and replenishment, which ensures that the customer's order is in stock and automatically orders new stock if a low stock threshold is set for an ordered product. The stock levels must be synchronized across the online catalog, call centers, and orders that are received through the mail. The applications interface with a high bay system that uses robotic cranes to retrieve the ordered product from its stock location.
- ▶ Delivery optimization schedules to ensure that the logistics are efficient so that trucks can leave the warehouse on time. The schedules also ensure that the contents are packed efficiently for delivery to subdepots where the content is moved to smaller delivery vehicles to deliver to more localized areas.

All of these activities are done by using workflows that minimize the time to process orders and provide a pleasant shopping experience for the customer.

Many other applications also are part of this process but are not described in this publication.

Finance scenario

This scenario is part of a bank processing system. The complexity increased further, as shown in Figure 14.

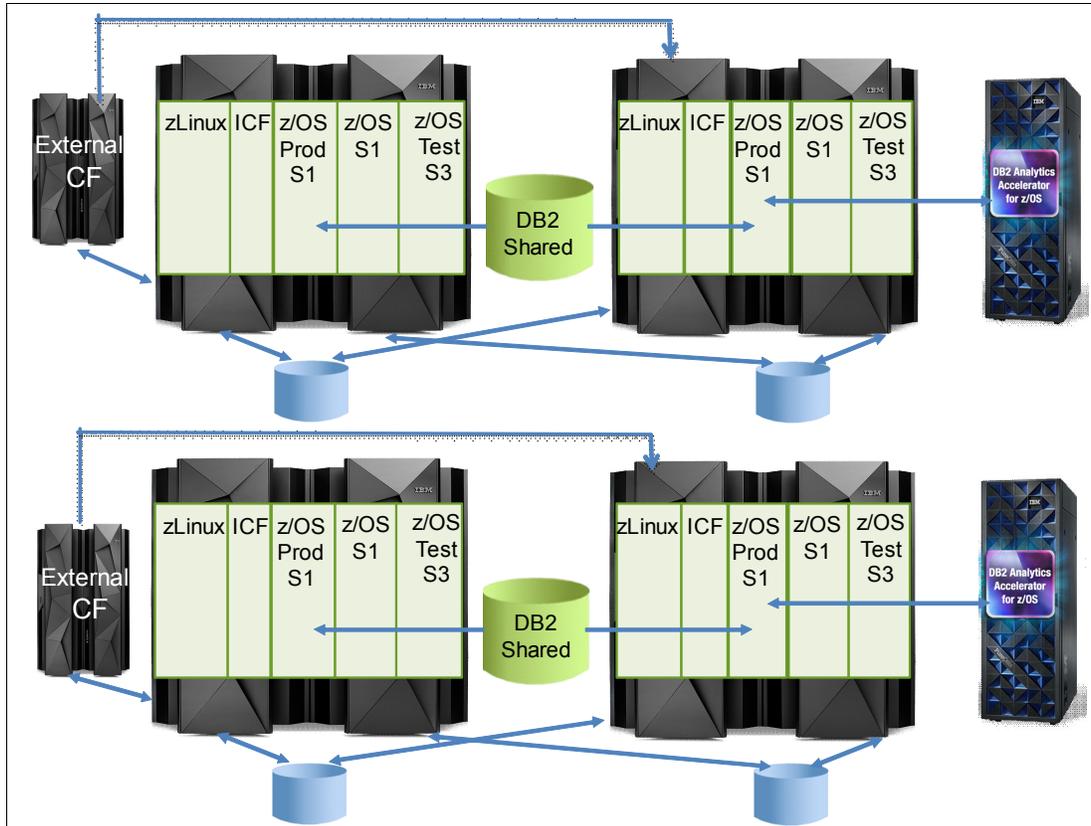


Figure 14 Finance scenario

There are multiple mainframes. There also are two groups and each group has two mainframes that communicate through an external coupling facility (CF). The mainframes in each group are not only sharing their configurations, but also the databases and other application elements. They are a system complex (sysplex). A sysplex is a group of LPARs on the same or other mainframes that share the responsibility for application performance and resilience. If a mainframe is temporarily withdrawn from the sysplex, the other mainframe continues to process the work. The external CF provides the connection for the two mainframes to synchronize with each other. LPARs that are sysplexed on one mainframe can use an internal CF to synchronize information.

A sysplex can exist with the mainframes in different locations; this condition is known as a geographically dispersed parallel sysplex (GDPS). GDPS allows systems to keep running if one of the locations becomes unavailable.

In the finance scenario, a customer might pay money into an account via the internet, mobile phone, or use the internet or a mobile phone to transfer money from their account at a different bank into a friend's account in this bank. The customer expects the money to be transferred almost instantaneously. This process is critical to the business because money transfers are a major part of the bank functions. If there are delays in transferring money, the financial institution, such as a bank, can be fined by the financial regulator and must recompense the customers for any charges that are incurred. External service providers that are used during the process might also have to pay penalties if there are issues.

Analytics are an important part of a bank's processes. In this scenario, we included the DB2 analytics accelerator for z/OS. This accelerator can work well with the SIMD capabilities to enhance the analytics, or they can be separate.

Finance demands the following high volumes of transactions:

- ▶ Money transfers
- ▶ Purchases
- ▶ Ensuring business and organization's payrolls reach their employees bank accounts
- ▶ Clearing inter-bank payments
- ▶ Investment-related transactions
- ▶ Loan applications
- ▶ Interest calculations
- ▶ Stocks and shares-related transactions
- ▶ Online statements
- ▶ Account queries

All these transactions must be performed quickly and the programs and data must be protected by securable methods, such as encryption and other protection layers.

Authors

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

Keith Winnard is the IBM Redbooks® Publications Project Leader for z/OS and related topics at the International Technical Support Organization (ITSO), Poughkeepsie Center. He joined IT in 1977 and has worked for various clients and Business Partners. He is experienced in blending traditional z/OS environments and applications with web middleware and applications, and has presented on many mainframe-related topics.

Luiz Fadel is an IBM Distinguished Engineer responsible for supporting System z for the Latin America region, which is part of the Growth Markets Unit. He joined IBM in 1969 and has supported Large Systems ever since, including working on two assignments with the ITSO. Luiz is a member of the Latin America IBM zEnterprise System Technical Introduction Advanced Technical Support team, which is responsible for handling Client Critical Situation and client claims within System z, Early Support Programs, new product installations, internal product announcements, and second-level client support, and managing complex Proof of Concepts (POC). He is a member of the zChampions team and the co-author of several IBM Redbooks® publications.

Rob Hunt is an Accredited IT Specialist with the IBM z Systems GTS Strategic Outsourcing Team in the United Kingdom. Rob has over 27 years of experience with IBM in storage, security, and systems management, supporting IBM MVS™, IBM z/VM®, and IBM z/OS environments. He has provided IBM mainframe storage and virtual machine support in the government, financial, retail, and insurance sectors.

Jo Johnston is a Certified IT Specialist and Chartered Engineer who works in the IBM z Systems Strategic Outsourcing Team in the United Kingdom. She has worked on IBM mainframe systems as a systems programmer supporting z/VM, z/OS, MVS, IBM CICS®, IBM DB2®, IBM WebSphere® Application Server, and IBM IMS™ for more than 30 years.

Alvaro Salla is a Senior IT Consultant for the ITSO. He has more than 40 years teaching and developing educational material that covers the z/OS mainframe platform and consulting services focusing on performance. Alvaro has co-authored Redbooks publications about IBM DS8000®, DFSMS, WLM, SYSPLEX, ABCs, IBM RMF™, and disaster recovery. He started with IBM in 1969.

Thanks to LindaMay Patterson, International Technical Support Organization, Rochester Center, for her contributions to this project.

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Stay connected to IBM Redbooks

- ▶ Find us on Facebook:
<http://www.facebook.com/IBMRedbooks>
- ▶ Follow us on Twitter:
<http://twitter.com/ibmredbooks>
- ▶ Look for us on LinkedIn:
<http://www.linkedin.com/groups?home=&gid=2130806>
- ▶ Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:
<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>
- ▶ Stay current on recent Redbooks publications with RSS Feeds:
<http://www.redbooks.ibm.com/rss.html>

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

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

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

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

DB2®	IBM z Systems™	WebSphere®
DRDA®	IBM z13™	z Systems™
FICON®	Redbooks®	z/OS®
FlashCopy®	Redpaper™	z/VM®
HiperSockets™	Redbooks (logo)  ®	z/VSE®
IBM®	System z®	z13™

The following terms are trademarks of other companies:

Linux is a trademark of Linus Torvalds 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 company, product, or service names may be trademarks or service marks of others.



REDP-5346-00

ISBN 0738455350

Printed in U.S.A.

Get connected

