

IBM z13

Technical Introduction

Octavian Lascu
Erik Bakker
Hans-Peter Eckam
Parwez Hamid
Rakesh Krishnakumar
Steven LaFalce
Lourenço Luitgards Moura Neto
Maurcio Andozia Nogueira
Frank Packheiser
Ewerson Palacio
Giancarlo Rodolfi
Andre Spahni

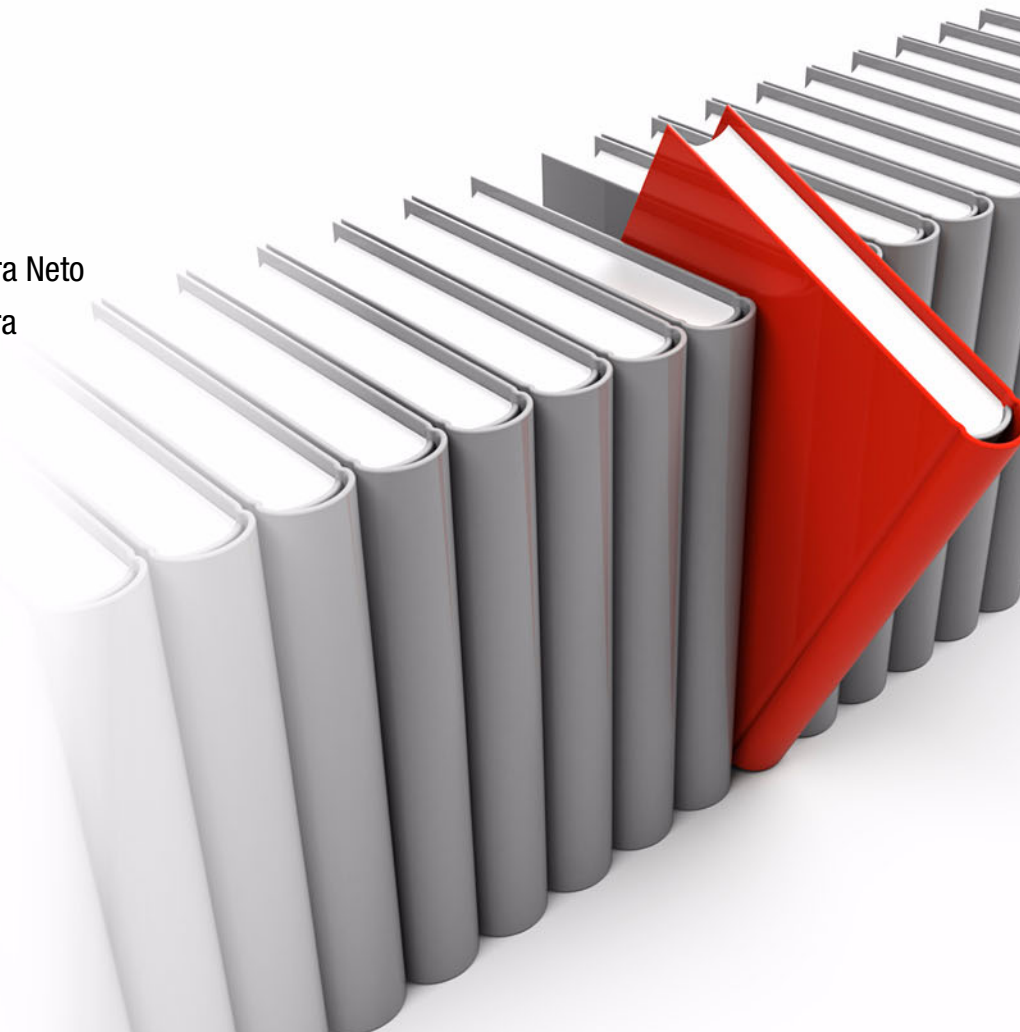


IBM Academy of Technology



Learn

IBM z Systems





International Technical Support Organization

IBM z13 Technical Introduction

January 2015

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

First Edition (January 2015)

This edition applies to the following IBM z Systems™ : IBM z13™, IBM zEnterprise EC12 (zEC12), IBM zEnterprise BC12 (zBC12), IBM zEnterprise 196 (z196), and IBM zEnterprise 114 (z114).

This document was created or updated on February 6, 2015.

Contents

Notices	vii
Trademarks	viii
Preface	ix
Authors	ix
Now you can become a published author, too!	xi
Comments welcome	xi
Stay connected to IBM Redbooks	xii
Chapter 1. Infrastructure matters	1
1.1 IBM z13 technical description	4
1.1.1 CPC drawers	6
1.1.2 Processor unit characterization	7
1.1.3 I/O subsystem	8
1.1.4 Hardware Management Console and Support Elements	15
1.1.5 Capacity on demand and performance	16
1.2 IBM z BladeCenter Extension Model 004	16
1.3 IBM z Unified Resource Manager	18
1.4 Software	19
1.5 Reliability, availability, and serviceability (RAS)	20
1.6 IBM z Advanced Workload Analysis Reporter (IBM zAware)	21
1.7 Simultaneous multithreading (SMT)	21
1.8 Single Instruction Multiple Data	22
1.9 Dynamic memory reassignment	24
1.10 Shared RoCE support	24
Chapter 2. Hardware overview	25
2.1 z13 highlights, models and upgrades	26
2.1.1 z13 highlights	26
2.1.2 IBM z13 models	27
2.1.3 z13 upgrades	28
2.2 Frames	29
2.2.1 z13 frames	29
2.3 z13 CPC drawers, and single chip modules	30
2.4 Processor chip	34
2.5 Processor unit	35
2.6 Memory	36
2.6.1 z13 memory	36
2.6.2 Concurrent memory upgrade	38
2.6.3 Redundant array of independent memory	38
2.6.4 Hardware system area	38
2.7 I/O system structure	39
2.8 I/O features	41
2.8.1 FICON Express16S	42
2.8.2 FICON Express8S	42
2.8.3 FICON Express8 (carry forward only)	43
2.8.4 OSA-Express5S	44
2.8.5 OSA-Express4S (carry forward only)	44
2.8.6 Flash Express	45

2.8.7 zEDC Express	46
2.8.8 10 Gigabit Ethernet RoCE Express	46
2.8.9 Coupling links and timing-only links	47
2.8.10 Cryptographic features	47
2.9 Coupling and clustering	48
2.9.1 Integrated Coupling Adapter (ICA SR)	49
2.9.2 InfiniBand coupling links	49
2.9.3 Internal Coupling (IC)	50
2.9.4 Coupling Facility Control Code (CFCC)	50
2.10 Time functions	51
2.10.1 Server Time Protocol (STP)	51
2.10.2 Network Time Protocol support	51
2.10.3 Time coordination for zBX components	52
2.11 Hardware Management Console and Support Element	52
2.12 Power and cooling	53
2.13 IBM z BladeCenter Extension	55
Chapter 3. Key functions and capabilities of IBM z13™	59
3.1 Virtualization	60
3.1.1 z13 hardware virtualization	60
3.1.2 IBM z Systems software virtualization	64
3.1.3 zBX virtualized environments	65
3.1.4 z Systems -based clouds	66
3.1.5 GDPS Virtual Appliance	66
3.2 z13 technology improvements	67
3.2.1 Microprocessor	67
3.2.2 Memory	72
3.2.3 Native PCIe features and integrated firmware processor	74
3.2.4 Flash Express	74
3.2.5 zEDC Express	75
3.2.6 10GbE RoCE Express	77
3.2.7 I/O capabilities	78
3.2.8 Cryptography	87
3.3 Capacity and performance	89
3.3.1 Capacity settings	90
3.3.2 z13 Capacity on Demand (CoD)	91
3.3.3 z13 performance	93
3.4 zSystems common time functions	96
3.4.1 Server Time Protocol (STP)	96
3.4.2 Network Time Protocol (NTP) client support	97
3.5 Hardware Management Console (HMC) functions	98
3.5.1 HMC key enhancements for z13	98
3.6 z13 power and cooling functions	99
3.6.1 High voltage DC power	99
3.6.2 Integrated battery feature (IBF)	100
3.6.3 Power capping and power saving	100
3.6.4 Power estimation tool	100
3.6.5 IBM Systems Director Active Energy Manager	100
3.6.6 Top Exit Power	101
3.7 IBM z BladeCenter Extension (zBX) Model 004	101
3.7.1 IBM blades	101
3.8 Reliability, availability, and serviceability (RAS)	102
3.8.1 IBM z Advanced Workload Analysis Reporter (IBM zAware)	104

3.8.2 RAS capability for the SE	104
3.8.3 RAS capability for the HMC	104
3.8.4 RAS capability for zBX	105
3.9 High availability	105
3.9.1 High availability for z Systems with Parallel Sysplex	105
3.9.2 PowerHA in zBX environment	108
3.10 IBM z Systems and emerging paradigms	110
Chapter 4. Operating system support	113
4.1 Software support summary	114
4.2 Support by operating system	117
4.2.1 z/OS	117
4.2.2 z/VM	121
4.2.3 z/VSE	124
4.2.4 z/TPF	125
4.2.5 Linux on z Systems	127
4.2.6 References	129
4.3 Software support for zBX Model 004	130
4.4 z/OS considerations	130
4.5 Coupling facility and CFCC considerations	134
4.6 Input/output configuration program (IOCP) considerations	135
4.7 GDPS Virtual Appliance	136
Chapter 5. A system of insight for digital business	137
5.1 Integrated hybrid infrastructures	138
5.1.1 A cornerstone of a smart IT infrastructure	138
5.2 Benefiting from a smart infrastructure	139
5.3 Cloud, analytics, mobile, social, and security	144
5.3.1 Cloud computing	145
5.3.2 Analytics on IBM z13	151
5.3.3 IBM z13 serving mobile and social applications	153
5.3.4 Security on IBM z13	156
Appendix A. Software licensing	161
Software licensing considerations	161
Monthly License Charge (MLC) pricing metrics	162
z Systems International Program License Agreement (IPLA)	165
zBX licensed software	165
IBM z Unified Resource Manager	166
Appendix B. Channel options	167
Abbreviations and acronyms	171
Related publications	173
IBM Redbooks	173
Other publications	173
Online resources	173
Help from IBM	174

Notices

This information was developed for products and services offered in the U.S.A.

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 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: 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.

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.


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.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®	Lotus®	Smarter Planet®
BigInsights™	OMEGAMON®	SPSS®
CICS®	Parallel Sysplex®	System p®
Cognos®	Passport Advantage®	System Storage®
DataPower®	POWER®	System z®
DB2®	Power Systems™	System z10®
DB2 Connect™	POWER7®	System z9®
Domino®	PowerHA®	SystemMirror®
DRDA®	PowerVM®	Tealeaf®
DS8000®	PR/SM™	Tivoli®
FICON®	Processor Resource/Systems Manager™	WebSphere®
GDPS®	QRadar®	Worklight®
HACMP™	RACF®	z/Architecture®
HiperSockets™	Rational®	z/OS®
HyperSwap®	Rational Rose®	z/VM®
IBM®	Redbooks®	z/VSE®
IBM Systems Director Active Energy Manager™	Redpaper™	z10™
IMS™	Redbooks (logo)  ®	z9®
InfoSphere®	Resource Link®	zEnterprise®
Insights™	RMF™	zSecure™
Language Environment®	Sametime®	

The following terms are trademarks of other companies:

Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo 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.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

Preface

This IBM® Redbooks® publication introduces the new IBM z13. IBM z13 delivers a data and transaction system reinvented as a system of insight for digital business. IBM z Systems leadership is extended with:

- ▶ Improved ability to meet Service Level Agreements with new processor chip technology that includes simultaneous multithreading, analytical vector processing, redesigned and larger cache, and enhanced accelerators for hardware compression and cryptography
- ▶ Better availability and more efficient use of critical data with up to 10 TB available RAIM memory
- ▶ Validation of transactions, management, and assignment of business priority for SAN devices through updates to the I/O subsystem
- ▶ Continued management of heterogeneous workloads with IBM z BladeCenter Extension (zBX) Model 004 and IBM z Unified Resource Manager

This IBM Redbooks publication will help you become familiar with the new z Systems platform, understand how it can help integrate data, transactions and insight for faster and more accurate business decisions.

This IBM Redbooks publication explains how, with its innovations and traditional strengths, IBM z13 it can play an essential role in today's IT environments, satisfy the demands for cloud deployments, analytics, mobile, and social applications in a trustful, reliable, and secure environment with operations that lessen business risk.

Authors

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

Octavian Lascu is a Senior IT Consultant for IBM Romania with over 20 years of experience. He specializes in designing and supporting complex IT infrastructure environments (systems, storage, and networking), including high availability and disaster recovery solutions and high-performance computing deployments. He has developed and taught over 50 workshops for technical audiences around the world. He has authored several Redbooks and Redpapers.

Ewerson Palacio is an IBM Distinguished Engineer and a Certified Consulting IT Specialist for Large Systems in Brazil. He has more than 40 years of experience in IBM large systems. Ewerson holds a Computer Science degree from Sao Paulo University. His areas of expertise include z Systems client technical support, mainframe architecture, infrastructure implementation, and design. He is an ITSO z Systems hardware official speaker, who has presented technical ITSO seminars, workshops, and private sessions to IBM clients, IBM IT Architects, IBM IT Specialists, and IBM Business Partners around the globe. He has also been a z Systems Hardware Top Gun training designer, developer, and instructor for the last generations of the IBM high-end servers. Ewerson leads the Mainframe Specialty Services Area (MF-SSA), which is part of GTS Delivery, Technology and Engineering (DT&E). He is a member of the IBM Academy of Technology.

Hans-Peter Eckam is a Support Center Specialist in the IBM Hardware Support Center in Mainz, Germany. Hans-Peter joined IBM in 1986 and has over 25 years of experience

supporting z Systems clients. His areas of expertise include z Systems, Sysplex Timer, Server Time Protocol, and connectivity such as ESCON, FICON, and InfiniBand.

Steven LaFalce started his career at IBM in 2002 working on design, testing, and qualification of z Systems memory subsystems. He later joined z Systems hardware Product Engineering working on the central electronic complex subsystem. In his current role, he is working in z Systems Field Technical Enablement on hardware innovation projects.

Erik Bakker is a Certified Senior IT Specialist working for IBM Server and Technology Group in the Netherlands. During the past 28 years he has worked in various roles within IBM and with a large number of mainframe customers. For many years he worked for Global Technology Services as a systems programmer providing implementation and consultancy services at many customer sites. He currently provides pre-sales z Systems technical consultancy in support of large and small z Systems customers. His areas of expertise include Parallel Sysplex®, z/OS®, and z Systems™.

Rakesh Krishnakumar is an IBM certified IT Specialist and PMP® certified project manager working with IBM India for the past six years. He has served multiple roles including z Systems software services delivery manager and lab manager for Linux on z Systems development and certification. Currently he is performing the role of architect/customer advocate for Linux on z Systems customers in Asia Pacific region. Rakesh has over 18 years of z Systems experience.

Andre Spahni is a Senior Support Center Representative working for IBM Global Technology Services in Switzerland. He has 13 years of experience working with and supporting z Systems clients. André has worked for the Technical Support Competence Center (TSCC) Hardware FE z Systems for Switzerland, Germany, and Austria since 2008. His areas of expertise include z Systems hardware, Parallel Sysplex, and connectivity.

Frank Packheiser is a Senior IT Specialist at IBM's Field Technical Sales Support office in Germany. He has 25 years of experience in IBM z Systems, zEnterprise, IBM zSeries, and predecessor mainframe servers. He has worked for 10 years for the IBM education center in Germany, developing and providing professional training. He also provides professional services to z Systems and mainframe clients. In 2008 and 2009, he supported clients in Middle East/North Africa (MENA) as a z Systems IT Architect. Besides co-authoring several Redbooks® publications since 1999, he has been an official ITSO presenter at ITSO workshops for the last four years.

Giancarlo Rodolfi is an IBM z Systems Technical Sales Specialist in Brazil. He has over 25 years of experience in the z Systems field. He has written extensively about the z/OS Communication Server, security and z/OS.

Mauricio Andozia Nogueira is a System Programmer at Banco do Brasil, a government owned bank in Brazil. He has seven years of experience in mainframe systems, including SAN, data center connectivity, and hardware configuration. He holds a degree in Computer Science from Marília University (Unimar). His areas of expertise include mainframe hardware configuration and storage director architecture and implementation.

Lourenço Luitgards Moura Neto is z/OS System Programmer. He manages the System Programmers team at Banco do Brasil, a government bank in Brazil. He has more than 12 years of experience in mainframe systems. He holds an engineering degree in automation and control from Brasília University. His areas of expertise include IBM Parallel Sysplex, z Systems hardware, z/OS availability and z/OS performance.

Special thanks to **Parwez Hamid** for guiding the team through the massive amount of information that was required to put this book together.

Parwez Hamid has been an Executive IT Consultant with the IBM Server and Technology Group and a Technical Staff member of the IBM UK Technical Council. During the past 39 years, he has worked in various IT roles within IBM. Since 1988, he has worked with many IBM mainframe clients, mainly introducing new technology. Currently, he works as a Consultant for System z in Poughkeepsie and provides technical support for the IBM System z hardware product portfolio. Parwez continues to co-author ITSO Redbooks publications and he prepares technical material for the worldwide announcement of System z servers. Parwez works closely with System z product development in Poughkeepsie, and provides input and feedback for future product plans. Parwez teaches and presents at numerous IBM user group and internal conferences and teaches at ITSO workshops.

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

| William G. White
International Technical Support Organization, Poughkeepsie Center

| Ellen Carbarnes, Diana Henderson, John Eells, Gregory Daynes, Daniel Jacobs, Garth Godfrey, Darelle Gent, Garth Godfrey, Bill Ogden
IBM Poughkeepsie

| Christine Smith, Amanda Stapels, Bill Bitner, Leslie Geer III, Romney White
IBM Endicott

Amy Tobenkin
IBM White Plains

| Monika Zimmermann, Uwe Zumpe, Utz Bacher, Dr. Reinhard Buendgen
IBM Germany

Dwayne Moore
IBM Canada

| Seth Adutwum
IBM Burlingame

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

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- ▶ Use the online **Contact us** review Redbooks form found at:

ibm.com/redbooks

- ▶ Send your comments in an email to:

redbooks@us.ibm.com

- ▶ Mail your comments to:

IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400

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>



1

Infrastructure matters

The digital era is witness to the transformation of technology and its emphasis on data mining and complex analysis of the huge amount (increasing at a very fast rate) of data that is available to cater to the needs of users by providing the insight that facilitates taking the best business decisions in a timely manner. The users interact with the systems through a wide range of devices and interfaces, from mobile devices (smartphones and tablets) to notebooks and desktops, requiring customized content delivered to fit the devices.

Data has become a new natural resource which an industry can use to develop existing business and create new business opportunities. To achieve the required business outcome, sophisticated analytics must be applied on a huge amounts of structured, semi-structured, and unstructured (raw) data. Based on the business requirements, new applications must be built to integrate front-end to back-end systems, such as banking applications, human resources applications, mobile applications, and so on.

Abundant bandwidth, emergence of new standards, and demand for consumability are transforming both *information technology (IT)* and business through the adoption of cloud based services approach. Successful leaders see possibilities by embracing digital business. This means supporting existing clients with new services and offerings, while providing an environment that helps previously underserved businesses and individuals gain access to products, services, and societal benefits.

A successful journey through IT transformation requires an underlying infrastructure that is efficient, secure, adaptive, and integrated. Today's IT infrastructure must be designed to handle the explosive growth of increasingly mobile users (clients and employees), to be able to leverage new and vast amounts of data, and to provide deeper, real-time insight at the point where the business impact is most significant (analytics). The IT transformation also needs to achieve these goals with a flexible, secure and resilient infrastructure (cloud-based).

This chapter introduces IBM z13 and explains its one-of-a-kind capabilities and traditional strengths for data processing and management, in the context of meeting the IT demands for cloud, analytics, mobile, and social, underpinned by security.

z13 description

IBM z13 is the mainframe that is optimized for the digital era. Unmatched virtualization capabilities, support for Enterprise grade Linux, open standards, enhanced sharing, and focus on business continuity ensure that z13 delivers state-of-the-art cloud support

The z13 has more cores, more memory, new cache design, improved I/O bandwidth, and compression than its predecessors that help to serve more data for analytics.

z13 with its performance, scale, intelligent I/O, and security enhancements supports transaction volume growth in the mobile world.

This chapter introduces the IBM z13 server and explains how, with its innovations and traditional strengths, it can play an essential role in today's IT environments, satisfy the demands for cloud deployments, analytics, mobile, and social applications. Chapter 2, "Hardware overview" on page 25, provides technical details.

Terminology: In the remainder of the book, we use the designation *CPC* to refer to the *central processor complex*.

Integrating data, transactions and insight

IBM is introducing the new generation of IBM z Systems. The new IBM z13 is designed to deliver new levels of performance and capacity for large-scale consolidation and growth. From the microprocessor to the software that exploits it, the z13 is designed for digital business. The new system features first-of-a-kind technologies that demonstrate IBM's ongoing commitment to meet changing computing demands. It is the world's fastest, most scalable and secure enterprise system with the ability to integrate resources for trusted resiliency, efficiency and data.

In 2012, The IBM Center for Applied Insights™ released a study that highlighted the emerging trends of big data and analytics, cloud, mobile, and social technologies. Recent studies show that 70% of the enterprises have adopted these technologies as opposed to only 30-50% a few years ago.

Companies using an outsourcing model for managing their IT are attracted to cloud capabilities, which provide them with lower prospects of cost while giving more flexibility, scalability, and speed in adapting to business requirements. At the same time, with evolving security requirements, business stakeholders are not willing to risk undermining the compliance and security requirements or raising IT costs to be prepared to respond to ever changing security threats.

Big data and analytics are inspiring organizations worldwide to craft new tactics for doing business. Healthcare providers are taking advantage of analytics technology to gain a more complete view of patients health (status and evolution), and applying insights to care coordination, health management, and patient management. Retail industries also are harnessing a wealth of data from a wide range of sources, such as transactional, social, and even third-party entities, to deliver a smarter and more tailored customer (shopping) experience.

Systems of record manage vast and rich amounts of data providing core transactional systems that are essential for daily business operations. These systems meet the demands of regulatory compliance with high demand on quality of service with regard to performance and continuous availability.

Systems of engagement interface with the systems of record to meet the need of social and mobile workloads. These systems analyze data using sophisticated methodologies for rapid development and delivery of new business ideas.

Systems of insight also interface the systems of record and systems of engagement to provide support and automation (where possible) for business decisions, by integrating data from all available sources, structured and unstructured. Decision support systems integrate transactions and analytics for gaining insight that can provide the favorable advantage to organizations that exploit the ever growing amount of data in the fastest and most effective way.

The z13 is optimized for the digital era requirements, focusing on these three aspects:

- ▶ Efficient and trusted cloud services to transform the economics of IT
- ▶ Integrated transaction and analytics for insight at the point of impact (data landing or source)
- ▶ Data and transaction serving for the mobile generation of applications that also facilitate and enhance social interaction.

Figure 1-1 shows the IBM z13 CPC.



Figure 1-1 IBM z13 CPC

The IBM z13 is designed to perform and scale out with intelligent and resilient I/O features. These include FICON® Express features, which improve the I/O latency, reducing the elapsed time for data access.

The IBM z13 also supports a larger memory pool for faster analysis of large volumes of data. The simultaneous multithreading (SMT) feature, introduced with IBM z13, delivers more throughput for efficient and competitive operational analytics on z Systems™, and by boosting the performance of Java and Linux workloads, and zIIP eligible workloads too.

The on-chip hardware compression acceleration facility, introduced with IBM z13, helps to compress more data, saving disk space, and reduce data transfer time, while reducing I/O bandwidth requirements.

The z13 brings a new approach for Enterprise grade Linux with offerings and capabilities for availability, virtualization with z/VM® and a focus on open standards and architecture with new support of a kernel-based virtual machine (KVM) on the mainframe (see the following Statement of Direction). Best of all, many of our clients have IBM zEnterprise EC12 and IBM zEnterprise 196 servers, which can be upgraded to the z13 for the greatest levels of investment protection.

Statement of Direction^a (KVM offering for z Systems): In addition to the continued investment in z/VM, IBM intends to support a kernel-based virtual machine (KVM) offering for z Systems that will host Linux on z Systems guest virtual machines. The KVM offering will be a software distribution that can coexist with z/VM virtualization environments, z/OS®, z/VSE®, and z/TPF. The KVM offering will be optimized for z Systems architecture and will provide standard Linux and KVM interfaces for operational control of the environment. In addition, KVM will integrate with standard OpenStack virtualization management, enabling enterprises to easily integrate Linux servers into their existing infrastructure and cloud offerings.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

1.1 IBM z13 technical description

The IBM z13 uses several leading-edge technologies, including silicon-on-insulator 22nm (CMOS 14S0 process), storage-class memory, PCIe Generation 3 (PCIe Gen3), InfiniBand, and Ethernet. The z13, when compared to its predecessor, the IBM zEC12, offers improvements in several areas such as a faster, more efficient, and redesigned high-frequency chip, more granularity options, better availability, and enhanced on-demand options.

The z13 is the first system to offer simultaneous multithreading (SMT) and single-instruction, multiple-data (SIMD) technologies. IBM z13 also offers *Transactional Execution Facility*, first introduced in the zEC12, known in the industry as *hardware transactional memory*, which can also increase transaction rates by reducing resource collision. Available with the previous zEnterprise family, the following features are also available in the z13:

- ▶ The IBM Flash Express for Storage Class Memory, a solid-state drive-based offering contributes to system availability.
- ▶ The innovative IBM zAware, an analytical and statistical appliance offering, further contribute to enhancing the system RAS by providing a near real-time, continuous learning, diagnostic, and monitoring capability.

In addition, several features are introduced, namely in the connectivity and data encryption areas.

z13 is a highly scalable *symmetric multiprocessor (SMP)* system that can be configured with up to 141 configurable processors. Five models are offered: N30, N63, N96, NC9, and NE1.

The model name represents the maximum number of processors that can be configured in the model (“C9” stands for 129, and “E1” stands for 141).

The z13 system architecture ensures continuity and upgradability from the zEC12 and z196 EC systems. The IBM commitment to the z13 and its sustained investment in the system and its predecessors is portrayed in Figure 1-2. The figure compares the z13 with previous Enterprise Class z Systems, regarding four major attributes:

- ▶ Single engine processing capacity (based on the *processor capacity index (PCI)*¹)
- ▶ Number of engines
- ▶ Memory
- ▶ I/O bandwidth (servers use a subset of their designed I/O capability)

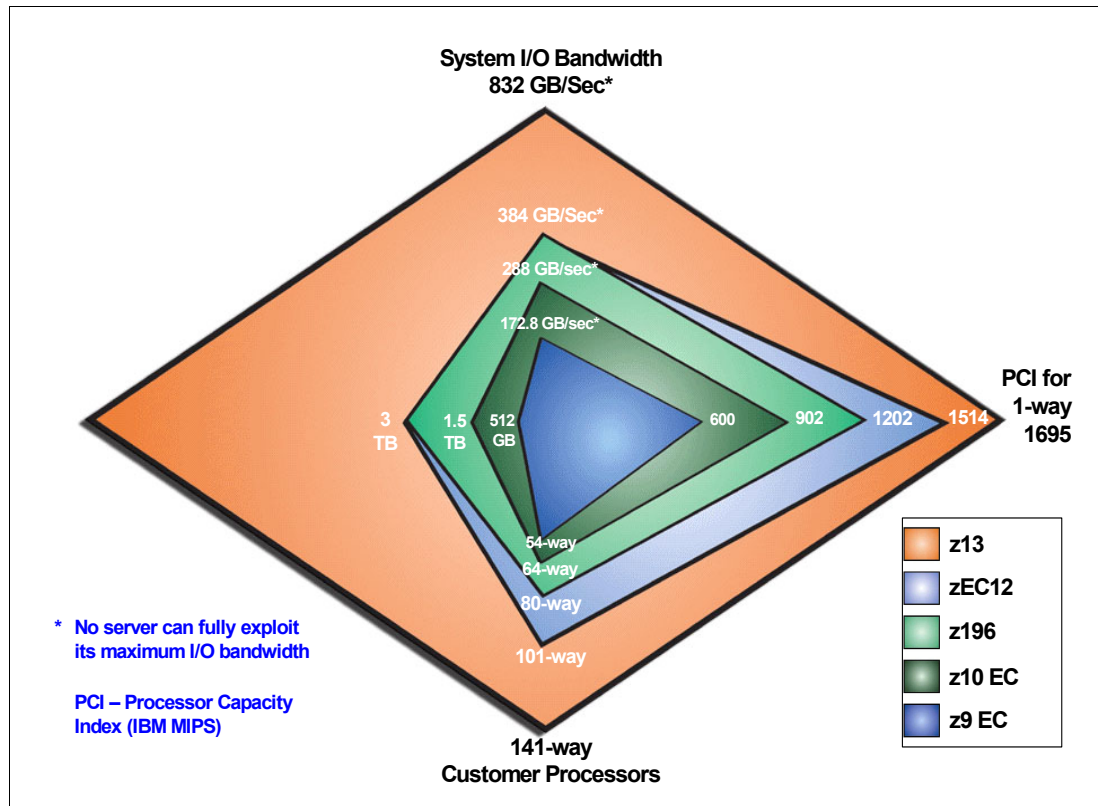


Figure 1-2 High End z Systems platform design comparison

The z13 has a machine type (M/T) designation of 2964 and is a two-frame system. The frames are known as the *A frame* and the *Z frame*.

The A frame contains the following elements:

- ▶ Up to four processor drawers, which hold the processors, memory, storage controllers, and other essential electronic components
- ▶ Two redundant Support Elements (SE)
- ▶ Two redundant System Control Hubs (SCH)
- ▶ Modular cooling units (different for water and air cooling systems)
- ▶ PCIe I/O drawers, I/O drawers, and their I/O features, available in several combinations
- ▶ Optional integrated battery features (IBF)

¹ PCI values can be obtained from *Large Systems Performance Reference*, SC28-1187: <https://www-304.ibm.com/servers/resourceLink/lib03060.nsf/pages/lisprindex>

The Z frame contains the following elements:

- ▶ Two SE displays, two SE console keyboards (with pointing device) mounted in a removable tray
- ▶ PCIe I/O drawers, I/O drawers, and their I/O features, available in several combinations
- ▶ Power supplies
- ▶ Optional IBFs

For better energy efficiency, the z13 offers, in addition to a radiator air-cooled option, a water-cooled option. To further increase energy efficiency a high-voltage DC (HVDC) power option is available for z13 (for both air- and water-cooled systems). These features mean potentially lower costs, without significantly changing the system physical footprint (the water cooling option adds a few inches of depth to the back of both system frames).

The z13 offers top-exit cabling options for power and I/O, as an alternative to having all the cables exit at the bottom of the CPC to under the raised floor. A non-raised floor installation of the z13 air-cooled systems is also possible, although water-cooled systems must be installed on a raised floor. These options are offered on new build and miscellaneous equipment specification (MES) orders. The increased flexibility allows choosing the options that best meet the data center requirements.

1.1.1 CPC drawers

The z13 inherits several technologies from the zEC12 but also incorporates new ones. Notably, the z13 central processor complex (CPC) is redesigned to package processors in *drawers*, rather than *books* (as in zEC12 systems), with node design to improve the affinity between processors, caches, and memory. The z13 processor chip and storage control chips are also redesigned to increase the numbers of processors, the size of the memory, and the size of the four levels of cache.

z13 CPC drawer

On a z13, the CPC houses from one to four CPC drawers that are fully interconnected. Each CPC 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. Despite the multi-CPC drawer design, the system is a symmetric multiprocessor, scalable up to 168 active cores (141 characterizable for customer use).

The z13 is the first mainframe to implement a high-speed eight-core design. To enable the increased number of cores sharing larger caches with quicker access time and improved capacity and performance, the z13 has a cycle time of 0.2 nanoseconds (5.0 GHz). Each core is known as a *processor unit (PU)*.

z13 is built on the proven superscalar microprocessor z/architecture. The PU chip has several innovations, notably in the superscalar pipeline design, multithreading, and on-chip caches. Improvements have been made in error checking and correcting (namely in the memory design) and specialized circuitry (for instance, to support multithreaded execution). z13 architecture has a set of instructions to improve the vector processing, through parallel processing of several operands with a single instruction.

In each node of a CPC drawer, there is one storage control chip (SC SCM) and three processor unit chips (PU SCMs). Therefore, the CPC drawer has eight single chip modules: six PU chips and two SC² chips. Each PU chip has enabled either six, seven, or eight cores.

² SC Storage Controller

z13 extracts the generated heat from the PU SCMs with an internal, closed-loop water system. This internal water system can be cooled in two ways:

- ▶ *Radiator units* with air-cooling
- ▶ *Water cooling units*, which are connected to a building's (data center) chilled water systems together with the back door heat-exchange units

One of the two cooling options must be selected at ordering time because they are factory installed. Converting between the cooling options (water, air) in the field is not possible.

Up to 10 TB of memory are orderable, with support for up to 10 TB configurable per logical partition depending on operating system level and other configuration factors. For enhanced availability, memory is implemented as a Redundant Array of Independent Memory (RAIM). In each CPC drawer, of the 3200 GB that can be installed, part is redundant, so that up to 2560 GB of usable memory can be configured. In addition, 96 GB are part of the base and reserved for the *hardware system area (HSA)*; as a result, the maximum amount of purchasable memory is 10144 GB, just short of 10 TB (with redundancy, a total of 12.8 TB are installed). *Plan-ahead memory*, a capability whereby memory can be installed but not enabled for use until needed, further enhances system availability for continuous operations.

1.1.2 Processor unit characterization

At system initialization time, each installed and enabled PU is characterized as one of various types. Also possible is to dynamically characterize PUs. A PU that is not characterized cannot be used.

Some PUs are part of the base system, that is, they are *not* part of customer purchaseable PUs and so are characterized by default:

- ▶ System assist processor (SAP):
SAPs offload and manage I/O operations. Several are standard with the z13. Additional SAPs can be configured if increased I/O processing capacity is needed.
- ▶ Integrated firmware processor (IFP):
The integrated firmware processor is a single PU dedicated to the support of the native PCIe features (10GbE RoCE Express and zEDC Express).
- ▶ Spare PUs
These can transparently assume any characterization, in the case of the permanent failure of a characterized PU.

Customer purchaseable PUs can assume any of the following characterizations:

- ▶ Central processor (CP):
This standard processor is for use with any supported operating system and user applications.
- ▶ Integrated Facility for Linux (IFL):
This is exclusively used with Linux on z Systems and for running the z/VM hypervisor in support of Linux. z/VM is often used to host multiple Linux virtual machines (called *guests*).
- ▶ IBM System z Integrated Information Processor (zIIP):
zIIP is used under 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.

zIIP^a:

- ▶ Work that is dispatched on zIIP does not incur any IBM software charges. zIIPs contribute to a lower cost of computing by taking some of the z/OS load that otherwise run on CPs.
- ▶ The z/OS system measurement facility (SMF), when properly configured, is able to project the potential utilization of zIIPs by the workloads running on the system, without actually requiring the purchase and installation of these processors. This estimate might help size the number of required zIIPs, and justify their acquisition.
 - a. IBM System z Application Assist Processor is not available on z13. zAAP qualified workload can now run on a zIIP processor (zAAP on zIIP feature), thus reducing the complexity of the z/Architecture.

- ▶ Internal Coupling Facility (ICF):

ICFs are used for z/OS clustering. They are dedicated to this function and exclusively run the *Coupling Facility Control Code (CFCC)*.

- ▶ An additional *system assist processor (SAP)* can be used by the channel subsystem.

CP Assist for Cryptographic Function

The z13 continues to offer the cryptographic assist implementation, first deployed in 2003, known as *CP Assist for Cryptographic Function (CPACF)*.

CPACF is physically implemented in the eight-core chip by the compression and cryptography accelerators. Each core has one dedicated coprocessor (CoP) integrating the CPACF and the compression unit, with improved performance and more capability when compared to zEC12.

The 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, using a no-charge enablement feature, except for the SHAs, which are shipped enabled with each server.

The CP Assist for Cryptographic Function supports the following functions:

- ▶ DES
- ▶ AES
- ▶ SHA
- ▶ Message authentication code (MAC):
- ▶ Pseudo Random Number Generation (PRNG) for cryptographic key generation
- ▶ Protected key capabilities

Keys: The keys can be provided only in clear form.

1.1.3 I/O subsystem

The z/Architecture defines an *I/O subsystem* to which I/O processing is offloaded. This is a significant contributor to the performance and availability of the system, and it strongly contrasts with the architectures of other servers.

The z/Architecture also specifies that *peripheral devices* are managed by *control units* and are reached through *channels* from the CPC. A control unit provides controlling function for a device or set of devices, and may be physically implemented with the device or in an independent unit.

The z13 I/O subsystem is fully virtualized and allows full sharing of its elements across multiple logical partitions.

As with its predecessors, the z13 implements the z/Architecture I/O subsystem through a dedicated subsystem, known as the *channel subsystem (CSS)*. z13 has six channel subsystems, each composed of the following elements:

- ▶ System assist processor (SAP)

SAP is a specialized processor that uses the installed PUs³. Its role is to offload I/O operations and manage channels and the I/O operations queues. It relieves the other PUs of all I/O tasks, allowing them to be dedicated to application logic. Enough SAP processors are automatically defined, depending on the model of the machine. The SAPs are part of the base configuration of the system.

- ▶ Hardware system area (HSA)

HSA is a reserved part of the system memory and contains the I/O configuration. It is used by SAPs. On the z13, a fixed amount of 96 GB is reserved. The HSA is not part of the client-purchased memory. This amount provides for greater configuration flexibility and higher availability by eliminating planned and preplanned outages.

- ▶ Channels

Channels are dedicated processors that communicate with the I/O control units (CU). They manage the data transfer between memory and the external devices. Channels are contained in the I/O card features.

- ▶ Channel path

Channel paths are the means by which the channel subsystem communicates with the I/O devices. Because of I/O virtualization, multiple independent channel paths can be established on a single channel, allowing sharing of the channel between multiple logical partitions, with each partition having a unique channel path. The function that allows sharing I/O paths across logical partitions is known as the multiple image facility (MIF). The z13 continues to transform the Channel subsystem enhancement for I/O resilience, providing improved throughput and reduced I/O service times. This was introduced with the IBM zEnterprise EC12 and the IBM zEnterprise zBC12

- ▶ Subchannels

Subchannels are displayed to a program as a logical device (programs do not directly communicate with the devices) and contain the information that is required to perform an I/O operation. One subchannel exists for each I/O device addressable by the channel subsystem. The z13 has four subchannel sets⁴.

In addition to the channel subsystem, the z13 also implements a *queued direct I/O (QDIO)* infrastructure, present also on predecessor systems. QDIO is a highly efficient data transfer mechanism that is designed to dramatically reduce system overhead and improve throughput by using system memory queues and a signaling protocol. Data is directly exchanged between the I/O features and the network software.

QDIO is used by Open Systems Adapter (OSA) features, HiperSockets using the QDIO Accelerator function, and the FICON channels when operating in FCP mode.

The I/O subsystem direction of the z13 is evolutionary, expanding on developments from the zEC12 and zBC12, and includes Peripheral Component Interconnect Express (PCIe), Integrated Coupling Adapter (ICA SR), InfiniBand, enhanced cards, and protocols (High

³ Each z13 PU can be characterized as one of seven types. For more information, see 1.1.2, "Processor unit characterization" on page 7.

⁴ Subchannel set 0 can have up to 63.75 K devices (256 devices are reserved). Subchannels 1,2, and 3 can each have up to 64 K minus one.

Performance FICON for z Systems (zHPF)). It can provide significant performance improvements over the I/O platforms of previous systems both by reducing overhead and latency and providing increased data throughput.

The I/O infrastructure includes I/O infrastructure elements, the PCIe I/O drawer, and the I/O drawer⁵.

Peripheral Component Interconnect Express (PCIe)

PCIe is a standard for computer add-on cards. It is designed as a serial bus following industry standard specifications and is used by a large variety of computer platforms. The bus operates at 16 GBps for Gen3, and 8 GBps for Gen2.

The PCI Special Interest Group is responsible for developing and maintaining format specifications.

PCIe in the z13 provides an internal I/O infrastructure that positions the system for continued support of the industry's direction for high-performance I/O.

The PCIe I/O bus connects the processor drawer to the PCIe I/O drawers. PCIe I/O drawers house PCIe features.

The IBM *integrated coupling adapter* (ICA SR), introduced on the z13 platform, is a two-port fanout used for short distance coupling Parallel Sysplex® connectivity and uses PCIe technology. The ICA SR is designed to drive distances up to 150m and support a link data rate of 8 GBps. The ICA SR fanout is housed in the PCIe I/O fanout slot on the z13 processor drawer. ICA SR supports z13 to z13 connectivity. InfiniBand (IFB, see next section) and ICA SR links completely replace the InterSystem Channel-3 (ISC-3) and ICB-4 offerings available on previous systems.

InfiniBand

InfiniBand is an industry-standard specification that defines a first-order interconnection technology, which is used to interconnect servers, communications infrastructure equipment, storage, and embedded systems. InfiniBand is a fabric architecture that uses switched, point-to-point channels with data transfers of up to 120 Gbps, both in chassis backplane applications and through copper and optical fiber connections.

A single connection can carry several types of traffic, such as communications, management, clustering, and storage. Additional characteristics include low processing overhead, low latency, and high bandwidth. Thus, it can become pervasive.

InfiniBand is used by the z13. Within the system, the InfiniBand I/O bus connects the processor drawer to I/O drawers. I/O drawers house legacy I/O features. For external usage, InfiniBand (IFB) links are available to interconnect z Systems CPCs in a Parallel Sysplex (z/OS cluster). IFB and ICA (see previous section) links completely replace the InterSystem Channel-3 (ISC-3) and ICB-4 offerings available on previous systems.

I/O drawer

Up to two I/O drawers can be carried forward on an upgrade to a z13. The I/O drawers can be installed, on the z13, both on the Z frame. I/O drawers were first offered with the z10™ BC, and each can accommodate up to eight FICON Express8⁵ features in z13.

⁵ No new orders. Carry forward only.

Important: The z13 supports real storage configurations up to 10 TB in a single LPAR. However, the installation of an I/O drawer limits the LPAR storage size to 1 TB only.

IBM PR/SM™ supports full system memory per LPAR. Operating system support might differ.

PCIe I/O drawer

The *PCIe I/O drawer* was introduced with the z196. This drawer provides for a higher number of cards (four times as much as the I/O drawer) and finer port granularity. The PCIe I/O drawers can be concurrently installed and repaired in the field. Each drawer can accommodate up to 32 PCIe I/O features in any combination. Up to five PCIe I/O drawers can be installed on the z13.

PCIe I/O features

The z13 supports the following PCIe features, which can be installed only in the PCIe I/O drawers:

- ▶ FICON Express16S
- ▶ FICON Express8S
- ▶ OSA-Express5S
- ▶ OSA-Express4S
- ▶ RoCE Express
- ▶ Crypto Express4S
- ▶ Crypto Express5S
- ▶ Flash Express
- ▶ zEDC Express

I/O features

When carried forward on an upgrade, z13 also supports the FICON Express8 features, which can be installed only in the I/O drawers.

In addition, ICA SR links and IFB coupling links, which attach directly to the processor drawers, are supported.

For a description of each I/O feature that is supported by the z13, see 2.8, “I/O features” on page 41.

Native PCIe features and integrated firmware processor (IFP)

From a hardware perspective, all features that reside in the PCIe I/O drawers and I/O drawers are supported from the Support Elements. This includes installing and updating the Licensed Internal Code of the features, and other operational tasks.

Many features have an application-specific integrated circuit (ASIC) that handles the adaptation layer functions required to present, in a uniform manner, the necessary features to the rest of the system. Thus, all operating systems have the same interface into the I/O subsystem.

The z13 also supports two features that were introduced with zEC12 and zBC12 (the 10GbE RoCE Express and zEDC Express), with industry-standard PCIe adapters, termed native PCIe features.

For native PCIe features, there is no adaptation layer, but the device driver is presented in the operating system. The adapter management functions (such as diagnostics and firmware updates) are provided by Resource Groups.

Resource Groups (there are two, for reliability, availability, and serviceability) use the integrated firmware processor (IFP), which is part of the system's base configuration and transparent to customers' operations.

Storage connectivity

Storage connectivity is provided on the z13, by Fibre Channel Connection (FICON) features, which support several protocols.

IBM Fibre Connection (FICON) channels

FICON channels follow the Fibre Channel (FC) standards. They support data storage and access requirements, and the latest FC technology in storage and access devices. FICON channels 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 includes the zHigh Performance FICON (zHPF) protocol and is used in z/OS, z/VM, IBM z/VSE (no zHPF), z/TPF, and Linux on z Systems environments.

- ▶ Fibre Channel Protocol (FCP)

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

Some rules must be followed when combining FICON Express16S, FICON Express8S, and FICON Express8 features.

Depending on the feature, auto-negotiated link data rates of 2, 4, 8, or 16 Gbps are supported (2, 4, and 8 for FICON Express8 and FICON Express8S; 4, 8, and 16 for FICON Express16S).

FICON Express16S is the most recent feature and provides significant improvements in throughput and response time for performance critical middleware and to shrink the batch window required to accommodate I/O bound batch work. FICON Express16S is implemented as PCIe cards and offer better port granularity and improved capabilities. FICON Express16S supports a link data rate of 16 Gbps (4,8, or 16 Gbps auto-negotiate) and is the preferred technology for new systems.

Networking connectivity

The z13 is a fully virtualized system, able to support many system images. Network connectivity, thus, covers not only those connections between the system and its outside, but also specialized internal connections for inter-system communication.

Open Systems Adapter (OSA)

The *Open Systems Adapter (OSA)* features provide local networking (LAN) connectivity and comply with IEEE standards. In addition, OSA features assume several functions of the TCP/IP stack that normally are performed by the processor. These functions can provide significant performance benefits.

The OSA-Express3 features are not supported, fulfilling the IBM Statement of General Direction.

HiperSockets

The *HiperSockets* function is an integrated function of the CPC that provides users with attachments to up to 32 high-speed virtual local area networks with minimal system and network overhead.

HiperSockets is a function of the virtualization Licensed Internal Code (LIC) and provides LAN connectivity across multiple system images on the same CPC by performing memory-to-memory data transfers in a secure way. The HiperSockets function eliminates having to use I/O subsystem operations and having to traverse an external network connection to communicate between logical partitions in the same CPC. Therefore, HiperSockets offers significant value in server consolidation by connecting virtual servers and simplifying the Enterprise network.

Improved functions of HiperSockets (also available on zEC12, zBC12, z196, and z114) include the ability to integrate in the *intraensemble data network (IEDN)*, and support for bridging to z/VM virtual switches.

10GbE RoCE Express

The 10 Gigabit Ethernet (10GbE) RoCE Express feature uses Remote Direct Access Memory over Converged Ethernet, and can provide fast memory-to-memory communications between two CPCs. It is transparent to applications.

Use of the 10GbE RoCE Express feature might help reduce consumption of CPU resources for applications that use the TCP/IP stack (such as IBM WebSphere® accessing a DB2 database). It might 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.

Coupling and Server Time Protocol (STP) connectivity

Clustering of z Systems systems under a single logical system image is possible under z/OS and uses specialized hardware.

Note: The z/VM Single System Image has a different design and does not require the features described in this section.

Coupling links

Coupling links are used when clustering z13 and z Systems running the z/OS operating system on the same or different CPCs. A clustered configuration is known as a *Parallel Sysplex* and can have up to 32 member nodes. The links provide high-speed bidirectional communication between members of a sysplex. The z13 supports the following links:

- ▶ Integrated Coupling Adapter (ICA SR) for distances up to 150 meters
- ▶ Internal coupling (IC) links for memory-to-memory transfers between LPARs on the same CPC
- ▶ Optical links: 12x InfiniBand optical links for distances up to 150 meters (492 feet)
- ▶ Optical links: 1x Infiniband optical (LR) links for unrepeated distances up to 10 Km (6.2 miles)

Important: The IBM z13 does not support InterSystem Channel-3 (ISC-3) and Host Channel Adapter 2 - Optical (HCA2-O).

Special purpose features

This section has an overview of several features that, although installed in the PCIe I/O drawer or in the I/O drawer, provide specialized functions without actually performing I/O operations, that is, no data is moved between the CPC and externally attached devices.

Cryptography

The tamper-sensing and tamper-responding Crypto Express5S and Crypto Express4S features provide high-performance cryptographic operations. Each Crypto Express5S and each Crypto Express4S feature has one PCI Express adapter. Each of the adapters can be configured in one of these modes:

- ▶ Secure IBM CCA coprocessor: For secure key encrypted transactions using CCA callable services (default).
- ▶ Accelerator: For public key and private key cryptographic operations that are used with Secure Sockets Layer/Transport Layer Security (SSL/TLS) acceleration.
- ▶ Secure IBM Enterprise PKCS #11 (EP11) coprocessor: Implements industry standardized set of services that adhere to the PKCS#11.

These features have specialized hardware to perform DES, TDES, AES, RSA, SHA-1, and SHA-2 cryptographic operations. The tamper-resistant hardware security module (HSM), which is contained in the Crypto Express5S and Crypto Express4S features, is designed to meet the FIPS 140-2 Level 4 security requirements for hardware security modules.

The configurable Crypto Express features are supported by z/OS, z/VM, z/VSE, z/TPF (accelerator mode only), and Linux on z Systems.

A paper, which is written by *Atsec information security corporation*, on Payment Card Industry compliance, recognizes the inherent qualities of the mainframe and the simplification in the infrastructure that it can provide. For more information, see this website:

http://www.atsec.com/downloads/white-papers/PCI_Compliance_for_LCS.pdf

Flash Express

Flash Express is an optional feature, introduced with the zEC12. It is also available on zBC12 and z13. When used under z/OS V1R13 and later, Flash Express can help improve availability and handling of paging workload spikes and act as an overflow area for large coupling facility structures for MQ. Using Flash Express can help availability by eliminating slow downs that can occur at the start of the workday. It can also help to eliminate delays that might occur when collecting diagnostic data during failures. Flash Express can, therefore, help organizations meet their most demanding service level agreements. Flash express can also be exploited on Linux on z Systems (SLES 11 SP3 and RHEL 6.4).

Flash Express is easy to configure, requires no special skills, and provides rapid time to value. Additional usage of Flash Express is expected to be supported in the future.

More details about the Flash Express feature are provided in *IBM z13 Technical Guide*, SG24-8251.

IBM zEnterprise Data Compression capability and zEDC Express

The growth of data that needs to be captured, transferred, and stored for large periods of time is not relenting. On the contrary, software-implemented compression algorithms are costly in terms of processor resources, and storage costs are not negligible either.

The zEDC Express, an optional feature exclusive to z13, zEC12 and zBC12, addresses that problem by providing hardware-based acceleration for data compression and decompression.

The z/OS V2R1 *IBM zEnterprise Data Compression* acceleration capability uses the *zEDC Express* feature. It delivers an integrated solution to help reduce CPU consumption, optimize performance of compression-related tasks, and enable more efficient use of storage resources, while providing a lower cost of computing and also helping to optimize the cross-platform exchange of data.

GDPS Virtual Appliance

Statement of Direction^a: In the first half of 2015, IBM intends to deliver a GDPS/Peer to Peer Remote Copy (GDPS/PPRC) multiplatform resiliency capability for customers who do not run the z/OS operating system in their environment. This solution is intended to provide IBM z Systems customers who run z/VM and their associated guests, for instance, Linux on z Systems, with similar high availability and disaster recovery benefits to those who run on z/OS. This solution will be applicable for any IBM z Systems announced after and including the zBC12 and zEC12.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

To reduce IT costs and complexity, many enterprises are consolidating independent servers into Linux images (guests) running on z Systems platform. Linux on z Systems can be implemented either as guests running under z/VM or native Linux LPARs on z Systems. Workloads with an application server running on Linux on z Systems and a database server running on z/OS are common. Two examples are as follows:

- ▶ WebSphere Application Server running on Linux and CICS®, DB2 running under z/OS
- ▶ SAP application servers running on Linux and database servers running on z/OS

With a multitiered architecture, providing a coordinated near-continuous availability and disaster recovery solution for both z/OS and Linux on z Systems is unnecessary.

GDPS Virtual Appliance is a fully integrated continuous availability and disaster recovery solution for Linux on z Systems customers and consists of these elements:

- ▶ An operating system image
- ▶ The application components
- ▶ An appliance management layer which makes the image self-containing
- ▶ APIs and UIs for customization, administration, and operation tailored for the appliance function.

The GDPS Appliance can improve both consumability and time-to-value for customers. For more information, see *IBM z13 Technical Guide*, SG24-8251.

1.1.4 Hardware Management Console and Support Elements

The Hardware Management Console (HMC) and Support Elements (SE) are appliances that together provide hardware platform management for z Systems. Hardware platform management covers a complex set of setup, configuration, operation, monitoring, and service management tasks and services that are essential to the use of the System z hardware and software.

For an example of the HMC and SE connectivity, see Figure 2-11 on page 53.

1.1.5 Capacity on demand and performance

In the same footprint, the z13 141-way system can deliver up to 40% more capacity than the largest 101-way zEC12. The z13 1-way system has approximately 10% more capacity than the zEC12 1-way. Numerous improvements in the processor chip design, including new instructions, multithreading, and redesigned and larger caches, contribute to the additional capacity. Taking advantage of some of the functionality is available only by using the most recent levels of compilers and JVMs, and operating system support.

The z13 has the capability to execute two threads of certain workloads simultaneously in a single processor, increasing the capacity. This feature is known as simultaneous multithreading (SMT) and described later in 1.7, “Simultaneous multithreading (SMT)” on page 21. Note that variations on the observed increased performance are dependent on the workload type.

Except for IBM System z Application Assist processor (zAAP), the z13 continues to offer all the specialty engines that are available on previous z Systems. zAAP qualified workload can now run on a zIIP processor (zAAP on zIIP feature), thus reducing the complexity of the z/Architecture.

On the z13, performing just-in-time deployment of processor capacity resources is possible. The Capacity on Demand (CoD) function provides flexibility, granularity, and responsiveness to business requirements changes by allowing the user to dynamically change the available system capacity.

1.2 IBM z BladeCenter Extension Model 004

The *IBM z BladeCenter Extension Model 004 (zBX)* is available as an MES upgrade from a previous zBX Model 002 or a zBX Model 003. The zBX Model 004 provides the same improvements as those that are added to zBX Model 003. It also includes a set of two 1U rack-mounted internal Support Elements (SEs), which perform the monitoring and management functions within that zBX. After it is upgraded, the zBX Model 004 becomes a stand-alone box, not associated or owned by any z Systems CPC.

The zBX Model 004 can be added to an ensemble HMC as a member (or node) of an HMC created ensemble. The zBX addition to the ensemble HMC is done in the same fashion as other z Systems CPCs, using the *Add Object Definition* task.

zBX Model 004 supports IBM AIX® on IBM POWER7®, Linux on System x, Microsoft Windows on System x, and IBM WebSphere DataPower® Integration Appliance XI50 for zEnterprise on a blade form factor. They are connected to the z Systems CPCs through virtual LANs that are supported on a high-speed private network.

The zBX is managed through its own internal Support Elements and by using the Unified Resource Manager functions. In bringing together multiple platforms, IBM has created a scalable solution that simplifies hardware and firmware management and support, and the definition and management of a network of virtualized servers.

The zBX Model 004 consists of the following components:

- ▶ Up to four IBM Enterprise racks
- ▶ Two internal 1U rack mounted SEs with display and keyboard.
- ▶ Up to eight BladeCenter chassis (two per rack), with up to 14 blades each
- ▶ Select IBM blades, up to 112

- ▶ Two Top of Rack (TOR) 1000BASE-T switches for the *intranode management network (INMN)*
The INMN provides connectivity for management purposes.
- ▶ Two TOR 10 GbE switches for the IEDN⁴
- ▶ The IEDN is used for data paths between the z Systems CPCs and the zBX, and the other ensemble members. 8 Gbps Fibre Channel switch modules for connectivity to an SAN
- ▶ Power distribution units (PDU) and cooling fans

The zBX is configured with redundant components to provide qualities of service similar to that of z13, such as firmware management and the capability for concurrent upgrades and repairs.

The zBX components are configured, managed, and serviced using its internal SEs. Although the zBX processors are not z/Architecture processors and run specific software, including hypervisors, the software intrinsic to the zBX components does not require any extra administration effort or tuning by the user. In fact, it is handled as z Systems Licensed Internal Code. The zBX hardware features are part of the mainframe, not add-ons.

GDPS/PPRC and GDPS/GM support zBX hardware components, providing workload failover for automated multi-site recovery. These capabilities can help facilitate the management of planned and unplanned outages across the z13 and other z Systems platforms.

Statement of Direction:^a IBM intends to deliver new functionality with IBM Systems Director offerings to support the IBM zBX. Such planned new capabilities will be designed to provide virtual image management and enhanced energy management functions for IBM Power Systems™ and System x blades.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

The zBX Model 004 can be controlled only from an ensemble HMC and its internal SEs. The zBX Model 003 can be controlled only from a zEC12 or zBC12, and the zBX Model 002 can be controlled only from a z196 or z114. However, z13, zEC12, zBC12, z196 and z114 can use the zBX Model 004 as well as the other zBX Models 002 and 003, by connecting the CPCs to the IEDN TOR switch or by using the zBX to zBX IEDN switch ports.

A zBX Model 002 that is controlled by a z196 or a z114, and a zBX Model 003 that is controlled by a zEC12 or zBC12, must be upgraded to Model 004 at the time of upgrading the controlling CEC to a z13. During the upgrade, the virtualization and configuration data is preserved; however, the process is disruptive and requires a planned outage.

IBM blades

When ordering a zBX Model 004 MES upgrade, a new entitlement record may be acquired by the client. This new entitlement record will allow IBM System x blades or IBM POWER7 PS701 to be ordered and added to the zBX, up to the limit of available empty (not used) slots in the existing blade centers.

Addition of new racks or new blade centers are not supported. Also the addition of the *IBM WebSphere DataPower Integration Appliance XI50 for zEnterprise* is not supported.

The POWER7 blades offer a virtualized environment through the IBM PowerVM® Enterprise Edition hypervisor. The virtual servers run the AIX operating system. The System x blades have an integrated hypervisor using *kernel-based virtual machines (KVM)*, which provide a virtualized environment for running the Linux and Windows operating systems.

IBM WebSphere DataPower Integration Appliance X150 for zEnterprise

The *IBM WebSphere DataPower Integration Appliance X150 for zEnterprise (DataPower XI50z)* is integrated into the z Systems infrastructure. DataPower XI50z is a multifunctional appliance that can help provide multiple levels of XML optimization. It can also 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. This appliance can help to simplify, govern, and enhance the network security for XML and web services.

For more details about the IBM DataPower X150z integration appliance, see this website:

<http://www-01.ibm.com/software/integration/datapower/xi50z>

1.3 IBM z Unified Resource Manager

The z13 system perfectly fits in smart and cloud infrastructures, continuing IBM high-end system leadership and being both the next step in the evolution of mainframes and a premier solution for centrally managed enterprise environments including clouds. They integrate a true hybrid computing system that is composed of virtualized heterogeneous resources that are integrated and managed as a single system by the IBM z Unified Resource Manager.

The IBM z Unified Resource Manager is an integral part of the z13 systems. It provides end-to-end management of CPCs and zBX resources and of virtualized environments, with the ability to align those resources according to individual workload requirements.

Through virtualization, the physical resources can be shared among multiple workloads. Most likely, the workloads have varying policies with different objectives. The goal of the Unified Resource Manager is to fulfill the objectives of the workload policies in the most optimal and efficient way.

The Unified Resource Manager provides energy monitoring and management, goal-oriented policy management, increased security, virtual networking, and data management, consolidated in a single interface that can be tied to business requirements.

The functions that pertain to an ensemble are provided by the Hardware Management Console and Support Elements in the CPCs and in the zBX Model 004. For more information, see 3.5, "Hardware Management Console (HMC) functions" on page 98.

The Unified Resource Manager resource management functions are delivered in tiers, by two operational suites. Within the Unified Resource Manager, several roles are defined. This configuration promotes security through task isolation and authorization.

Resource management suites

The functions that are delivered by the Unified Resource Manager are accessed through the Hardware Management Console (HMC) and provide the following capabilities:

- ▶ Integrated hardware management across all elements of the system, the CPC, the zBX, and the integrated networks.
- ▶ Fully automatic and coherent integrated resource discovery and inventory for all elements of the system without requiring user configuration, deployment of libraries or sensors, or user scheduling.
- ▶ Hypervisors for the zBX components are shipped, serviced, and deployed as Licensed Internal code (LIC). They are booted automatically at power-on reset, and managed through the isolated intranode management network (INMN).

- ▶ Virtual server lifecycle management, enabling uniform directed and dynamic virtual server provisioning across all hypervisors of zBX elements from a single point of control.
- ▶ Representation of the physical and virtual resources that are used in the context of a deployed business function as a named workload.
- ▶ Monitoring and trend reporting of CPU energy efficiency, which can be helpful in managing the costs of deployed workloads.
- ▶ Delivery of system activity through a user interface, the Monitors Dashboard (which augments the existing System Activity Display), enabling a broader and more granular view of system resources consumption.

The Unified Resource Manager offers the ability to optimize technology deployment according to individual workload requirements. To achieve this optimization, the Unified Resource Manager is delivered in two suites of tiered functionality:

- ▶ Manage
- ▶ Automate/Advanced Management

See the following IBM Redbooks publications to read more about Unified Resource Manager functions and capabilities:

- ▶ *Building an Ensemble Using IBM zEnterprise Unified Resource Manager*, SG24-7921
- ▶ *IBM z13 Technical Guide*, SG24-8251

1.4 Software

The IBM z13 supports a wide range of IBM and independent software vendor (ISV) software solutions. The range varies from traditional batch and online transaction processing (OLTP) type of environments, which include IBM Customer Information Control System (CICS), IBM Information Management System (IMS™), IBM Database 2 (DB2), and these web services, among others: Java platform, Linux and open standards, IBM WebSphere, and Worklight® for mobile application development.

The following operating systems are supported by IBM z13:

- ▶ z/OS Version 2 Release 1 with PTFs (exploitation)
- ▶ z/OS Version 1 Release 13 with PTFs (limited exploitation)
- ▶ z/OS Version 1 Release 12 with PTFs (Limited exploitation)⁶
- ▶ z/VM Version 6 Release 3 with PTFs (exploitation)
 - compatibility support
 - Crypto Express5S and Crypto Express4S with up to 85 domains support
- ▶ z/VM Version 6 Release 2 with PTFs :
 - compatibility support
 - Crypto Express5S and Crypto Express4S with up to 85 domains support
- ▶ z/VSE Version 5 release 1 with PTFs (compatibility plus crypto Express5S support)
- ▶ z/VSE Version 5 Release 2 with PTFs (compatibility plus Crypto Express5S support)
- ▶ z/TPF Version 1 Release 1 (compatibility)
- ▶ Linux on z Systems

⁶ z/OS V1.12 End of Service support-9/30/2014. Service extension support is provided for up to three years beginning 1 October 2014

- SUSE: SUSE Linux Enterprise Server (SLES) 11 and SLES 12
- Red Hat: Red Hat Enterprise Linux (RHEL) 5, RHEL 6, and RHEL7

More information about the software support is in Chapter 4, “Operating system support” on page 113.

IBM Compilers

Compilers are built with knowledge about the system’s architecture, which is used during code generation. Using the latest compilers is therefore essential to extract the maximum benefit of a server’s new capabilities. Empower your business applications with IBM compilers.

With IBM Enterprise COBOL for z/OS and IBM Enterprise PL/I for z/OS, decades of IBM experience in application development can be used to integrate COBOL and PL/I with web services, XML, and Java. Such interoperability enables capitalizing on existing IT investments while smoothly incorporating new, web-based applications into the infrastructure of your organization.

z/OS XL C/C++ and XL C/C++ for Linux on z Systems help with creating and maintaining critical business applications that are written in C or C++ to maximize application performance and improve developer productivity. These compilers transform C or C++ source code to executable code that fully leverages the z Systems architecture. This function is possible through hardware-tailored optimizations, built-in functions, performance-tuned libraries, and language constructs that simplify system programming and boost application runtime performance.

1.5 Reliability, availability, and serviceability (RAS)

The z13 continues to offer the high quality of service and *reliability, availability, and serviceability (RAS)* that is traditional in IBM mainframes. The RAS strategy employs a building-block approach. developed to meet the client's stringent requirements for achieving continuous reliable operation. These are the building blocks:

- ▶ Error prevention
- ▶ Error detection
- ▶ Recovery
- ▶ Problem determination
- ▶ Service structure
- ▶ Change management
- ▶ Measurement
- ▶ Analysis

The RAS design objective is to manage change by learning from previous generations and investing in new RAS functionality to eliminate or minimize all sources of outages. Enhancements to z Systems RAS designs were implemented on the z13 system through the introduction of new technology, structure, and requirements. Continuous improvements in RAS are associated with new features and functions to ensure z Systems platform delivers exceptional value to our clients.

The IBM mainframe systems have gone through decades of intense engineering development. The introduction of z13 adds, once again, new, carefully engineered RAS features, providing the highest possible level of RAS. One example is an expanded capability of IBM zAware, which now allows Linux on z Systems logs to be analyzed.

For more details about RAS features, see the corresponding chapter in the *IBM z13 Technical Guide*, SG24-8251.

1.6 IBM z Advanced Workload Analysis Reporter (IBM zAware)

Introduced with the zEC12, the *IBM zAware* feature is an integrated expert solution that uses sophisticated analytics to help clients identify potential problems and improve overall service levels.

IBM zAware runs analytics in a special purpose firmware partition, isolated from production. It can monitor z13 or other z Systems CPCs running z/OS V1.13 (with supporting PTFs) or later levels of z/OS. With the introduction, IBM z13 IBM zAware is enhanced to support enterprise Linux operating systems running natively on z Systems. This chargeable feature scans z/OS operator logs and Linux system logs using optimized sorting and filtering algorithms to provide accurate predictions which helps diagnose problems while they are occurring in near real time.

More details about IBM zAware are in these locations:

- ▶ 3.8.1, “IBM z Advanced Workload Analysis Reporter (IBM zAware)” on page 104
- ▶ 4.4, “z/OS considerations” on page 130
- ▶ IBM z13 *Technical Guide*, SG24-8251

1.7 Simultaneous multithreading (SMT)

The z13, aligned with multithreading industry directions, can process up two simultaneous threads in a single core. This capability is known in z Systems as simultaneous multithreading (SMT) and offers full z/Architecture capability for each thread.

SMT is supported only by zIIP and IFL speciality engines on z13. An operating system with SMT support can be configured to dispatch work to a thread on a zIIP (for eligible workloads in z/OS) or an IFL (for z/VM on behalf of Linux on z System guest or native Linux on z Systems) core in single thread or SMT mode.

The use of SMT provides a better and more efficient use of the processor resources and helps address memory latency, resulting in overall throughput gains. SMT is designed to deliver better overall throughput for many workloads, but performance in some cases may be superior using single threading.

Figure 1-3 illustrates the capacity gain and a possible decrease in the performance of a single thread in SMT mode.

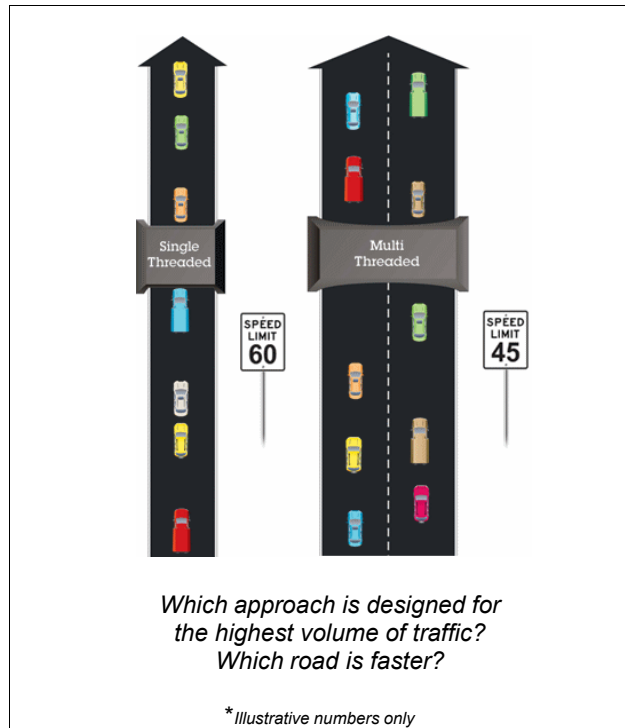


Figure 1-3 Capacity and performance in SMT mode

The SMT is functionally transparent to middleware and applications and no changes is required to run in SMT partition.

1.8 Single Instruction Multiple Data

The z13 architecture includes a subset of 139 new instructions, known as single-instruction, multiple-data (SIMD), added to improve efficiency of complex mathematical models and vector processing. These instructions allow a larger number of operands to be processed with a single instruction. The SIMD instructions use the superscalar core to process operands in parallel, enabling more interactions.

The set of SIMD instructions are a type of data parallel computing and vector processing that can decrease amount of code and accelerate code with integer, string, character and floating point data types.

Figure 1-4 shows a schematic comparison between a scalar instruction and a SIMD instructions.

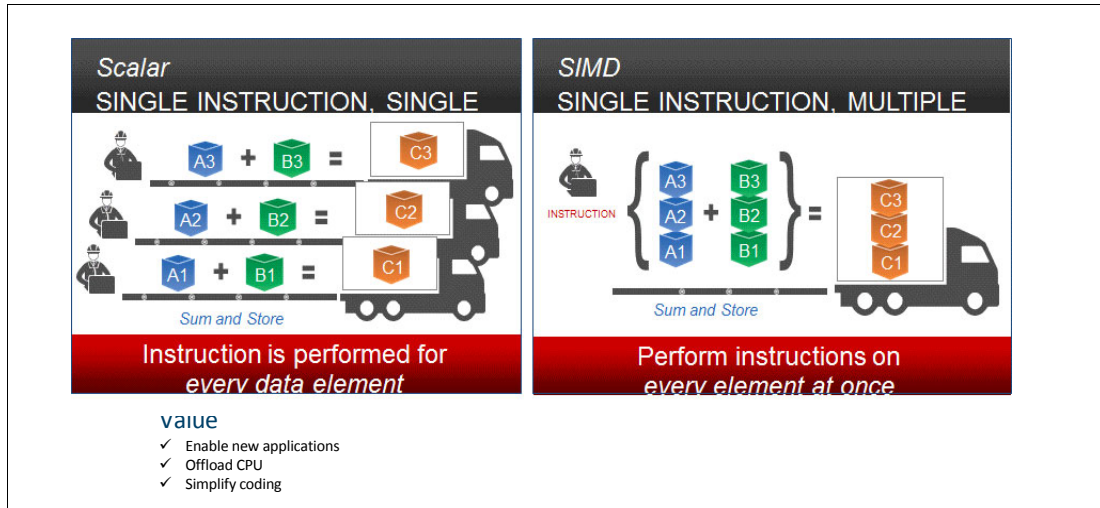


Figure 1-4 Schematic comparison between a scalar instruction and a SIMD instructions

Codes will have to be developed to take advantage of the SIMD capability. The following compilers have support and built-in functions for SIMD:

- ▶ IBM java
- ▶ XL C/C++
- ▶ Enterprise cobol
- ▶ Enterprise PL/I

Figure 1-5 shows how you can improve code generation, with significantly smaller amount of code, and analytics application performance with SIMD usage.

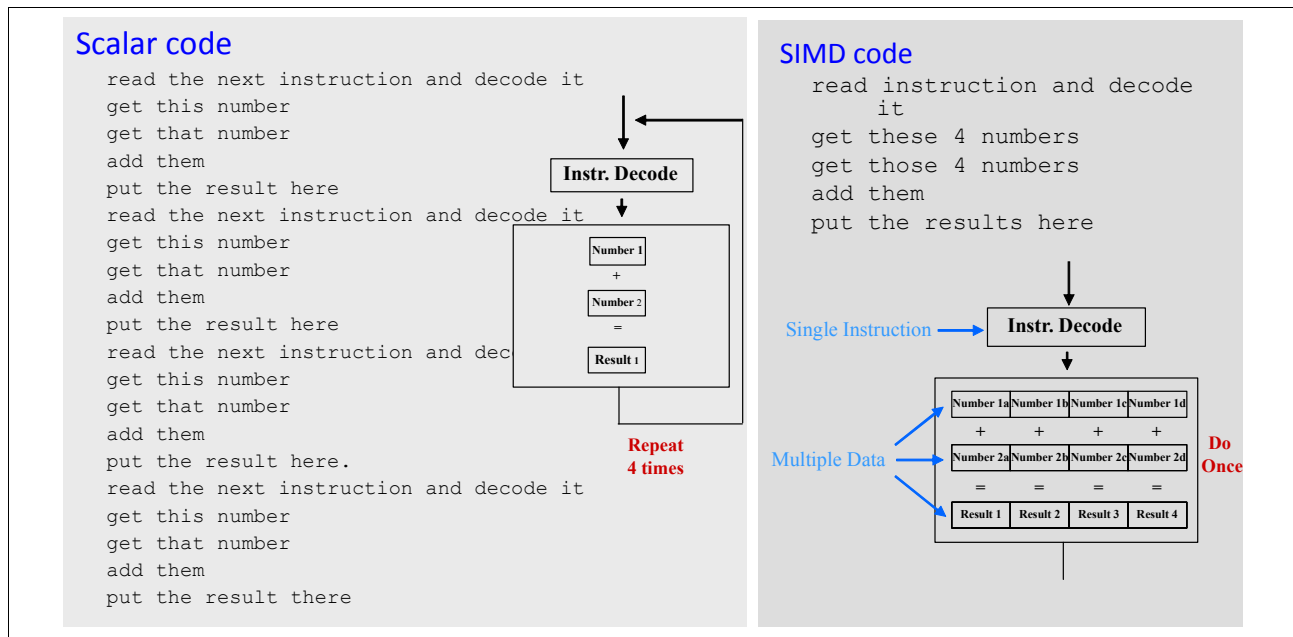


Figure 1-5 Comparison scalar code and SIMD code

The SIMD instructions improves performance of complex mathematical models and can permit integrate business transactions and analytic workloads on systems z.

Statement of direction - z/VM support for single-instruction, multiple-data (SIMD):
In a future deliverable, IBM intends to deliver support to enable z/VM guests to use the Vector Facility for z/Architecture (SIMD).

1.9 Dynamic memory reassignment

The logical partition (LPAR) memory assignment is done at image activation time. Since z10, the defined partition memory was striped across the installed books to use the performance and maximize the usage of the existing memory controllers (MC) on each of the books.

On z13 the memory allocation algorithm has changed. PR/SM will try to allocate memory in to a single processor drawer, striped between the two nodes. Basically, the PR/SM memory and processor resources allocation goal is to place all partition resources on a single processor drawer, if possible. The resources, memory and processors, are assigned to the partitions when they are activated. Later, when all partitions are activated, PR/SM can move memory between processor drawers to benefit performance, without operating system knowledge.

1.10 Shared RoCE support

z/OS V2.1 with PTF supports the new sharing capability for the Remote Direct Memory Access (RDMA) over Converged Ethernet (RoCE Express) features on z13 processors. This new enhancement will allow to utilize both ports in the RoCE adapter and to share the adapters across up to 31 z/OS images on a z13 processor. Also, the Communication Server is enhanced to support the selecting between TCP/IP and RoCE transport layer protocols automatically based on traffic characteristics. This feature is supported on z13, zEC12, and zBC12, and is installed in the PCIe I/O drawer. A maximum of 16 features can be installed. In z13 we have both ports enabled to be used and each feature, two ports, can be shared up to 31 logical partitions. Previously on zEC12 and zBC12, only one port could be used and the port had to be dedicated to a logical partition.

For details on RoCE, see *IBM z13 Technical Guide*, SG24-8251.



Hardware overview

The IBM z13 is the next step in the evolution of the mainframe family. It continues the evolution by introducing several innovations and expanding existing functions, building upon the z/Architecture.

The z13 is designed to deliver new levels of performance and capacity for large-scale consolidation and growth, and in support of cloud infrastructures, and also support for the next generation of digital signature security, cutting edge pattern recognition analytics for smart monitoring of system health, and enhanced environmental capabilities.

This chapter expands upon the overview of key hardware elements of the z13 provided in “z13 description” on page 2, and compares it with previous IBM z Systems, where relevant.

This chapter describes the following topics:

- ▶ z13 highlights, models and upgrades
- ▶ Frames
- ▶ z13 CPC drawers, and single chip modules
- ▶ Processor chip
- ▶ Processor unit
- ▶ Memory
- ▶ I/O system structure
- ▶ I/O features and Cryptographic features
- ▶ Coupling and clustering
- ▶ Time functions
- ▶ Hardware Management Console and Support Element
- ▶ Power and cooling
- ▶ IBM z BladeCenter Extension

2.1 z13 highlights, models and upgrades

The z13 models, and the improvements and upgrades over their predecessors, are discussed in this section.

2.1.1 z13 highlights

Major enhancements of IBM z13 over its predecessors include the following features:

- ▶ Increased total system capacity: 168-way system (with 141 characterizable PUs) and more subcapacity settings, offering increased levels of performance and scalability to help enable new business growth.
- ▶ Six eight-core 5.0 GHz processor chips that can help improve the execution of processor-intensive workloads.
- ▶ Up to 10 TB of addressable real memory per system (with up to 10 TB real memory per logical partition) ensuring high availability in the memory subsystem by the proven technology of redundant array of independent memory (RAIM).
- ▶ A 96 GB fixed *hardware system area (HSA)* that is managed separately from client-purchased memory (three times the size of the IBM zEnterprise EC12).
- ▶ Fourth-generation high frequency, second-generation out-of-order design, with single-instruction, multiple-data (SIMD) processor, which increases parallelism to enable analytics processing, Simultaneous multithreading (SMT-2) increases processing efficiency and throughput, and the number of instructions in flight increases by 42.8%.
- ▶ Processor cache structure improvements and larger cache sizes that can benefit most production workloads: Compared to zEC12, first level of cache (L1 private) is increased from 64 KB to 96 KB for instructions and from 96 KB to 128 KB for data. The size of second level of cache (L2) doubles from 1 MB to 2 MB for data and instructions, increases third level of cache (L3) from 48 MB to 64 MB, and fourth level of cache (L4) from 384 MB to 480 MB along with the addition of 224 MB for non-Data Integrated Coherent (NIC) Directory for L3.
- ▶ Improved cryptographic functionality and performance are achieved by having one dedicated cryptographic coprocessor per core.
- ▶ Channel subsystem enhancement for I/O resilience: The z13 channel subsystem increases the number of logical channel subsystems (LCSSs) from four to six, augmenting the number of logical partitions from 60 to 85. A fourth subchannel set is implemented to improve addressability. The number of I/O devices supported by all FICON channel features is augmented from 24 K to 32 K devices.

IBM z13 also introduces several new features and functions:

- ▶ LPAR resources allocation algorithms for PUs and memory.
- ▶ Generation of flash solid-state drives (SSD) for Flash Express feature. These SSDs are mounted in Flash Express feature cards and can be used to handle paging workload spikes and improve availability.
- ▶ Crypto Express5S, with enhanced support of cryptographic functions and 85 domains.
- ▶ PCIe Gen3 I/O fanout adapters with 16 GBps bus.
- ▶ RoCE Express feature has new functionality: Second port is now available for use, and also the ability to share adapters between LPARs.
- ▶ Integrated Coupling Adapter, ICA-SR for coupling links.

- ▶ The 1U Support Elements (SEs) replace the SE notebooks;
- ▶ Redundant System Control Hubs (SCHs) replace the Bulk Power Hubs (BPHs)
- ▶ Optional rack mountable 1U Hardware Management Console (HMC) for installation on a customer supplied rack.
- ▶ Air cooled systems: cooling radiators with N+2 redundant design.
- ▶ IBM z Advanced Workload Analysis Reporter (IBM zAware) software appliance, which provides a smart solution for detecting and diagnosing anomalies in z/OS systems is now also available for Linux on z Systems.

In all, these enhancements provide options for continued growth, continuity, and ability to upgrade.

For an in-depth discussion of the IBM z13 functions and features, see the *IBM z13 Technical Guide*, SG24-8251.

2.1.2 IBM z13 models

The z13 has an assigned machine type (M/T) of 2964, which uniquely identifies the central processor complex (CPC). The z13 is offered in five models:

- ▶ z13 N30: One CPC drawer and a maximum of 30 customizable processor units (PUs)
- ▶ z13 N63: Two CPC drawers and a maximum of 63 customizable PUs
- ▶ z13 N96: Three CPC drawers and a maximum of 96 customizable PUs
- ▶ z13 NC9: Four CPC drawers and a maximum of 129 customizable PUs
- ▶ z13 NE1: Four CPC drawers and a maximum of 141 customizable PUs

The model determines the maximum number of processor units that are available for characterization. PUs are delivered in single-engine (core) increments. The first four models use 39 PU cores on six PU SCMs in one CPC drawer, the fifth model, NE1, uses 42 PU on six SCMs, in each of the four CPC drawers to provide up to 141 configurable PUs.

Spare PUs, system assist processors (SAPs), and one integrated firmware processor (IFP) are integral to the system. Table 2-1 provides a model summary that includes SAPs and spare PUs for the various models. For an explanation of PU characterization, see “Processor unit characterization” on page 36.

Table 2-1 z13 Model summary

Model	Drawers/PUs	CPs	Standard SAPs	Spares	Integrated firmware processor
N30	1/39	0–30	6	2	1
N63	2/78	0–63	12	2	1
N96	3/117	0–96	18	2	1
NC9	4/156	0–129	24	2	1
NE1	4/168	0–141	24	2	1

The z13 offers 231 capacity levels. There are 141 capacity levels that are given by the number of physically used *central processors* (CP), plus the possibility of 90 subcapacity models for the first 30 CPs. There is also one model for all *Integrated Facility for Linux* (IFL) or all *Internal Coupling Facility* (ICF) configurations. This topic is described in more detail in “Processor unit characterization” on page 36.

2.1.3 z13 upgrades

Figure 2-1 summarizes the upgrade paths to the z13.

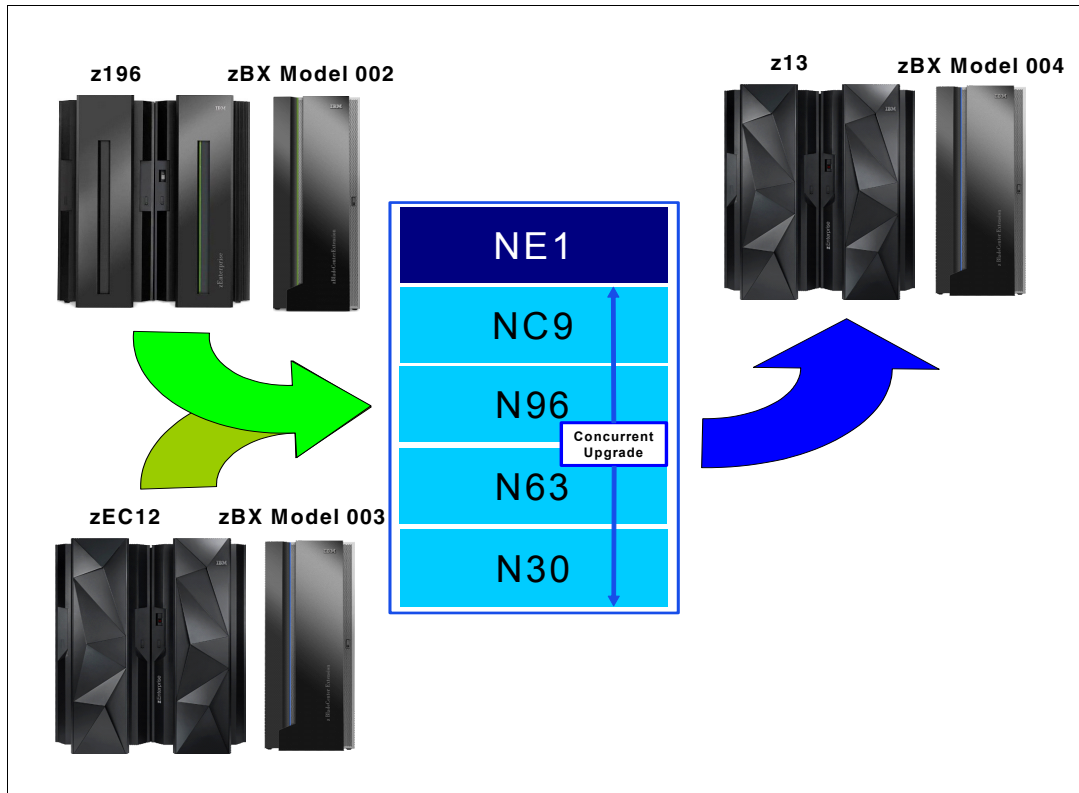


Figure 2-1 z13 upgrade paths

Concurrent upgrades of CPs, IFLs, ICFs, zAAPs, zIIPs, or SAPs are available for the z13. However, concurrent PU upgrades require that more PUs are physically installed (at a previous time), but not activated.

If an upgrade request cannot be accomplished within the specified configuration, a hardware upgrade is required. The upgrade enables the addition of one or more CPC drawers to accommodate the desired capacity. Additional CPC drawers can be installed concurrently. Upgrades from any z13 (model N30, N63, N96, HC9) to a model NE1, are disruptive because this upgrade requires the replacement of all installed CPC drawers.

Spare PUs are used to replace defective PUs. There are always two spare PUs on a z13. In the rare event of a PU failure, a spare PU is concurrently and transparently activated, and assigned the characteristics of the failing PU.

When a z196 with a zBX Model 002 is upgraded to z13, the zBX is converted to a Model 004. When a zEC12 with a zBX Model 003 is upgraded to z13, the zBX is converted to a Model 004. The virtualization and configuration data is preserved, however the upgrade process requires downtime.

2.2 Frames

The frames of the z13 are described in this section.

2.2.1 z13 frames

The z13 is always a two-frame system: the *A Frame* and the *Z Frame*. The z13 can be delivered as an air-cooled system or as a water-cooled system.

Figure 2-2 shows an internal front view of the two frames for an air-cooled system. The two frames form the Central Processing Complex (CPC). The number and type of drawer that holds I/O features can vary and depends on the number of I/O features. For a new build system, a combination of up to five PCIe I/O drawers can be installed. A miscellaneous equipment specification (MES) can carry forward up to two I/O drawers (8 slot I/O drawer). This configuration provides for a maximum number of 16 non-PCIe feature cards that can be carried forward.

Figure 2-2 shows the system with the maximum of five PCIe I/O drawers.

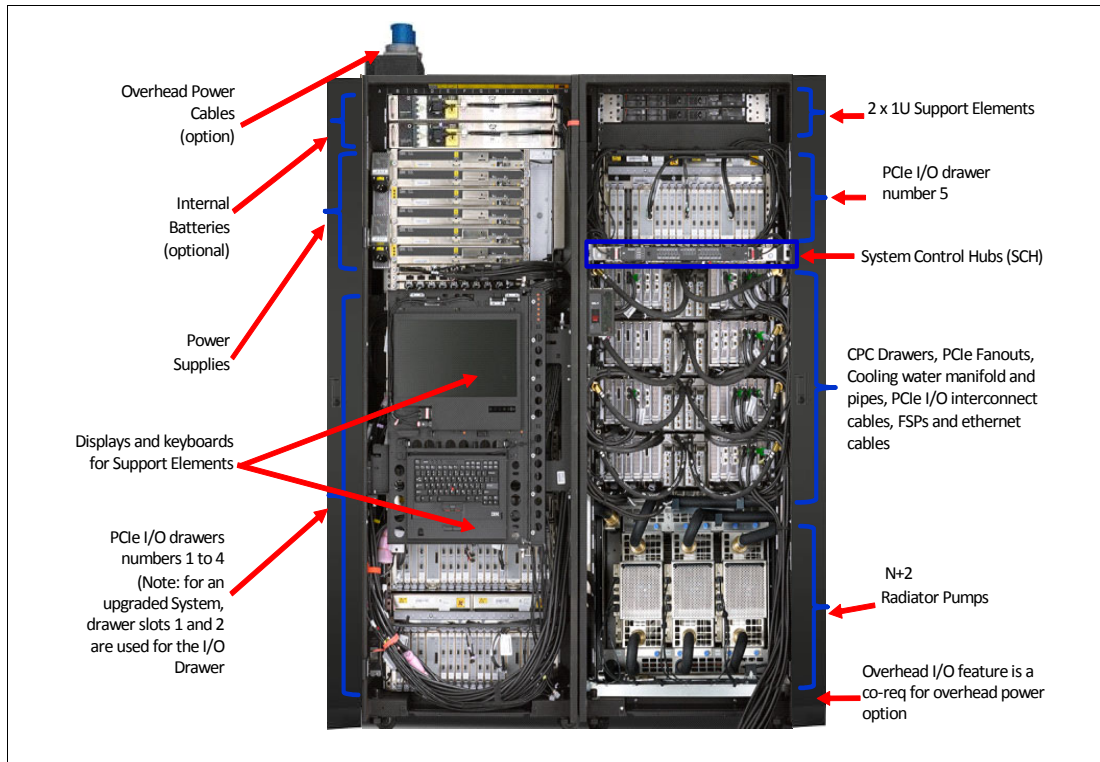


Figure 2-2 z13 internal front view: air-cooled CPC

Figure 2-3 shows an internal front view of the two frames of a water-cooled CPC without the top exit I/O cable and top exit power cable option.

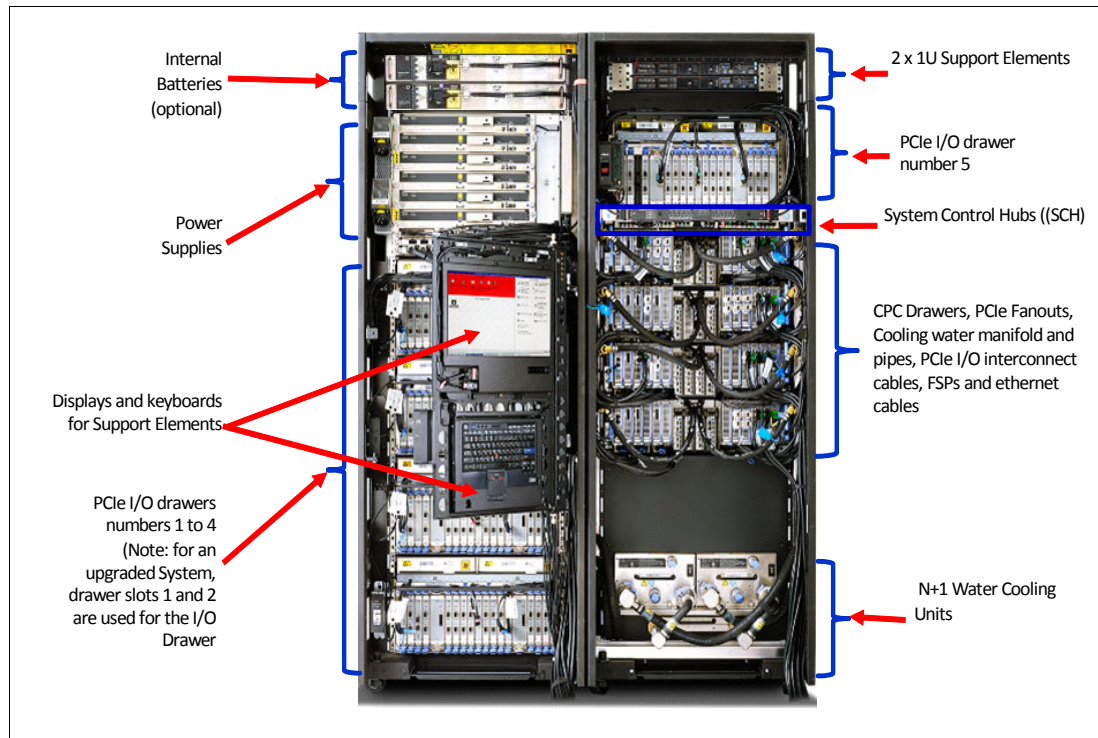


Figure 2-3 z13 internal front view: water-cooled CPC

Top exit I/O and power cabling

z13 and its predecessor, the zEC12, have the option of ordering the infrastructure to support the top exit of fiber optic cables (FICON, OSA, 12x InfiniBand, 1x InfiniBand, ICA SR and RoCE) and copper cables for the 1000BASE-T Ethernet features.

On the z13, the top exit capability is designed to provide an additional option, the overhead power cabling option. Figure 2-2 on page 29 shows this overhead power cable feature, present on the Z Frame. Instead of all the cables exiting under the CPC to under the raised floor, there is now the flexibility to choose the options that best meet the data center requirements. A non-raised floor installation of the z13 air-cooled systems is also possible. Top exit cabling can also help to increase air flow. These options are offered on new build and MES orders.

2.3 z13 CPC drawers, and single chip modules

The z13 system is a multiple CPC drawer system. Up to four CPC drawers can be installed in Frame A of a CPC. A CPC drawer slides into one of the four slots of the A Frame of the CPC. Figure 2-4 on page 31 shows the top view, no-cover CPC drawer. The location and structure of the CPC drawers are the same for both air-cooled and water-cooled systems.

Each CPC drawer contains the following elements:

- ▶ Eight single chip modules (SCM)

There are six PU SCMs and two storage controller (SC) SCMs. Each PU SCM includes eight PU cores.

SCMs are described in “z13 single chip modules” on page 33. See Table 2-1 on page 27 for the model summary and the relation between the number of CPC drawers and number of available PUs.

► Memory

A minimum of 256 GB and a maximum of 2.5 TB memory are available for client use. See Table 2-2 on page 37 for details.

For 25 available slots, 20 or 25 memory DIMMs are plugged into these slots, providing up to 2.5 TB of customer-available memory installed in a CPC drawer.

► Fanouts

A combination of up to four InfiniBand Host Channel Adapter (HCA3-Optical, HCA2-Copper) fanouts and up to 10 PCIe third generation (PCIe Gen3) fanouts.

Each fanout has one, two, or four ports, thereby up to 40 connections are supported:

- One-port PCIe 16 GBps I/O fanout, each supporting one domain in 32-slot PCIe I/O drawers)
- ICA SR two-port fanout, for coupling links (two links, 8 GBps each),
- HCA3-O 12x InfiniBand fanout, for coupling links (two ports at 6GBps each)
- HCA3-O LR 1x InfiniBand fanout, for coupling links (four ports, 5Gbps each).
- HCA2-Copper fanouts(two ports per fanout): supported only for I/O drawers that are carried forward, for a maximum of two features.

► Two Distributed Converter Assemblies (DCA) that provide power to the CPC drawer.

Loss of one of the DCAs leaves enough power to satisfy the power requirements of the entire drawer. The DCAs can be concurrently maintained.

Figure 2-4 shows a view of a z13 CPC drawer (without the containing box).

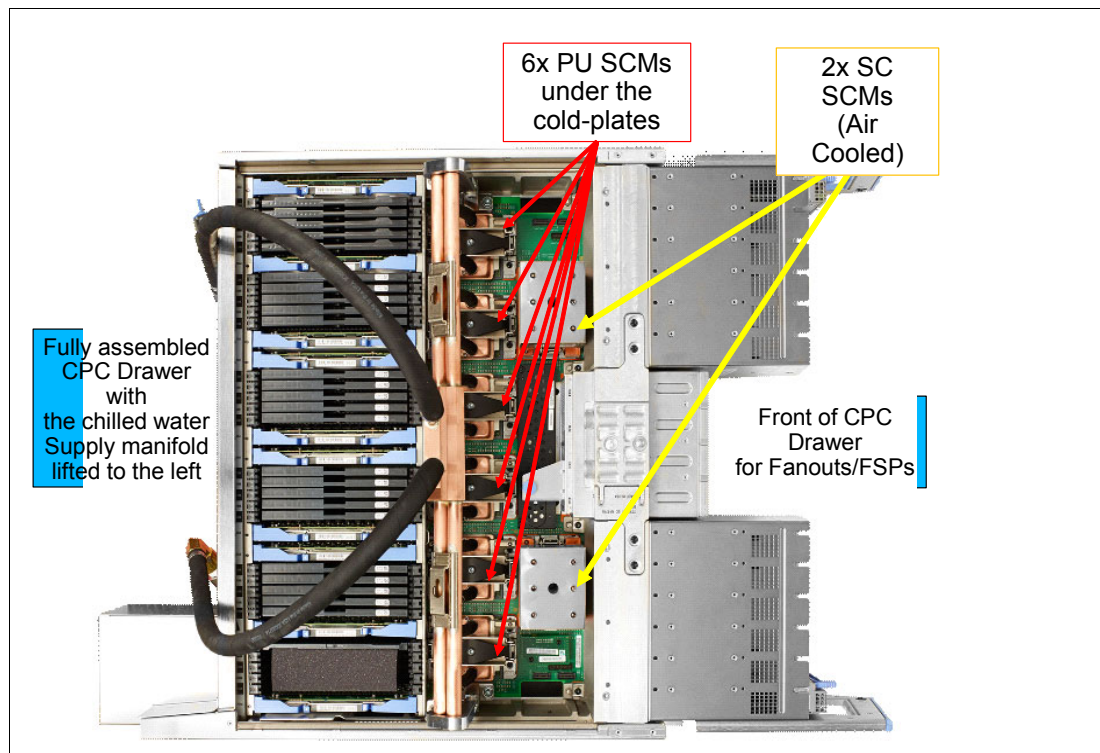


Figure 2-4 z13 CPC drawer structure and components

The z13 continues the evolution in system scalability, providing opportunity for server consolidation in a multi CPC drawer system structure. As shown in Figure 2-5 and Figure 2-6 on page 32, all CPC drawers are interconnected with high-speed communications links through the L4 shared caches. The z13 has 960 MB of L4 cache per CPC drawer, which is 2.5 times larger than its predecessor.

This design used for PU and SC interconnect allows the system to be operated and controlled by the PR/SM facility as a symmetrical, memory-coherent multiprocessor system (SMP).

Figure 2-5 shows the z13 inter-CPC drawer communication structure.

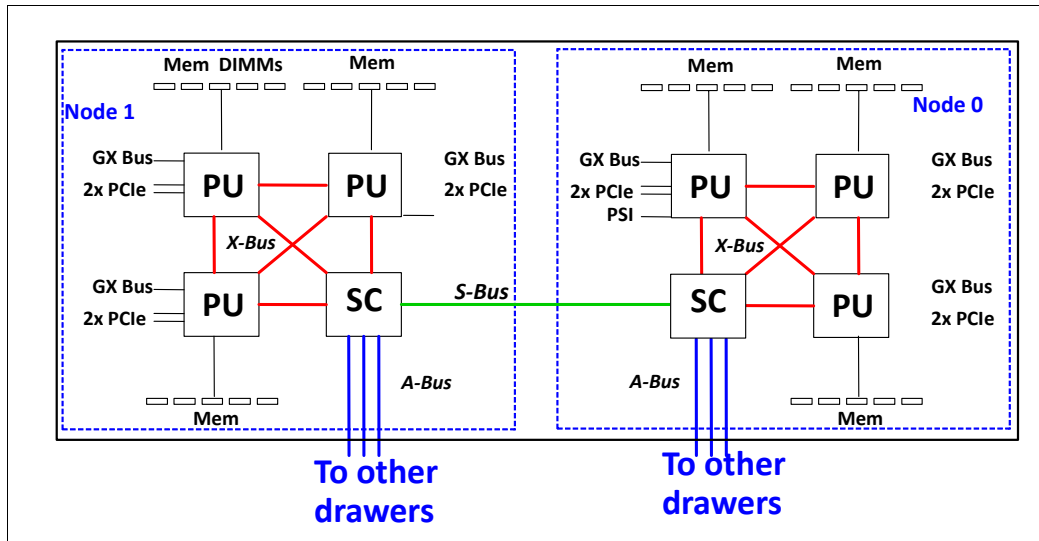


Figure 2-5 z13 CPC drawer communication topology

The point-to-point connection topology shown Figure 2-6 on page 32 allows direct communication between all CPC drawers.

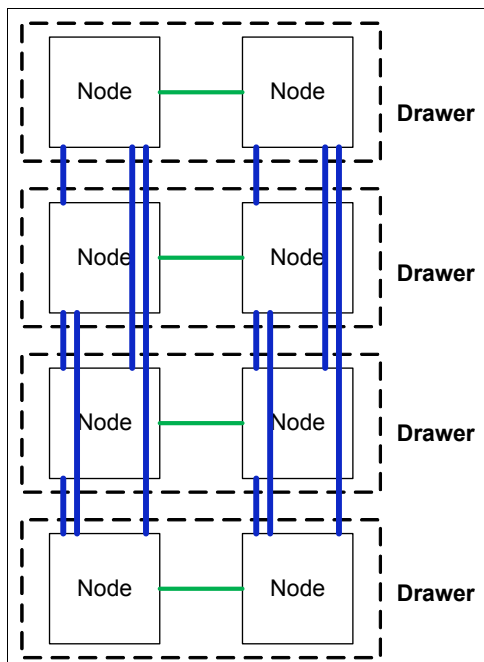


Figure 2-6 z13 CPC drawer interconnect topology (four drawers)

Although the z13 potentially includes several hundred processor chips, only the z/Architecture® processor chips are described in the section that follows.

z13 single chip modules

The *single chip modules (SCM)*, shown in Figure 2-7, are high-performance, glass-ceramic chips, providing the highest level of processing integration in the industry. The SCMs are the heart of the system. The figure shows the size compared to the currency of one US dime.

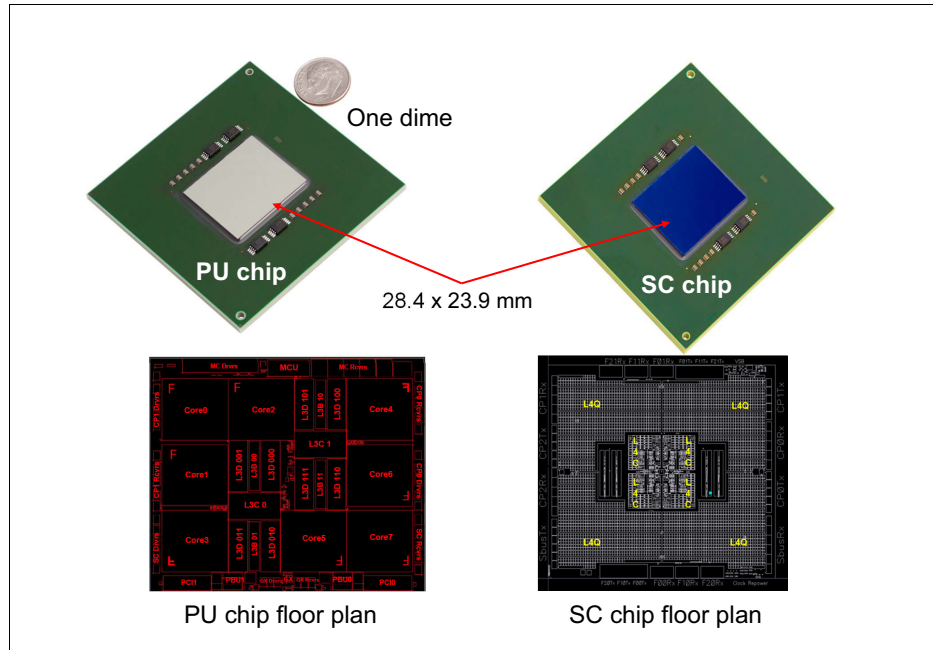


Figure 2-7 z13 PU and SC chip SCMs

Each z13 processor unit (PU) chip has eight cores. All chips use complementary metal-oxide of silicon (CMOS¹) 14S0 technology. CMOS 14S0 is a state-of-the-art microprocessor technology that is based on 17-layer copper interconnections and Silicon-On-Insulator (SOI) technologies. The chip lithography line width is 0.022 μm (22 nm). The processor unit chip contains close to 3,990,000,000 transistors in a 678.8 mm² die.

There are eight Single Chip Modules (SCMs) per CPC drawer. Six of these SCMs hold the processor chips (PU chips) and two SCMs hold storage control (SC) chips. Each PU chip has six, seven, or eight active cores, L1, L2, and L3 caches; the two SC chips holds L4 caches, as shown in Figure 2-8.

Two CPC drawer configurations are offered with 39 or 42 PUs. All the models employ CPC drawers with 39 PUs except for the model NE1, which has four CPC drawers with 42 active PUs, for a total of 168 PUs.

¹ Complementary metal-oxide-semiconductor

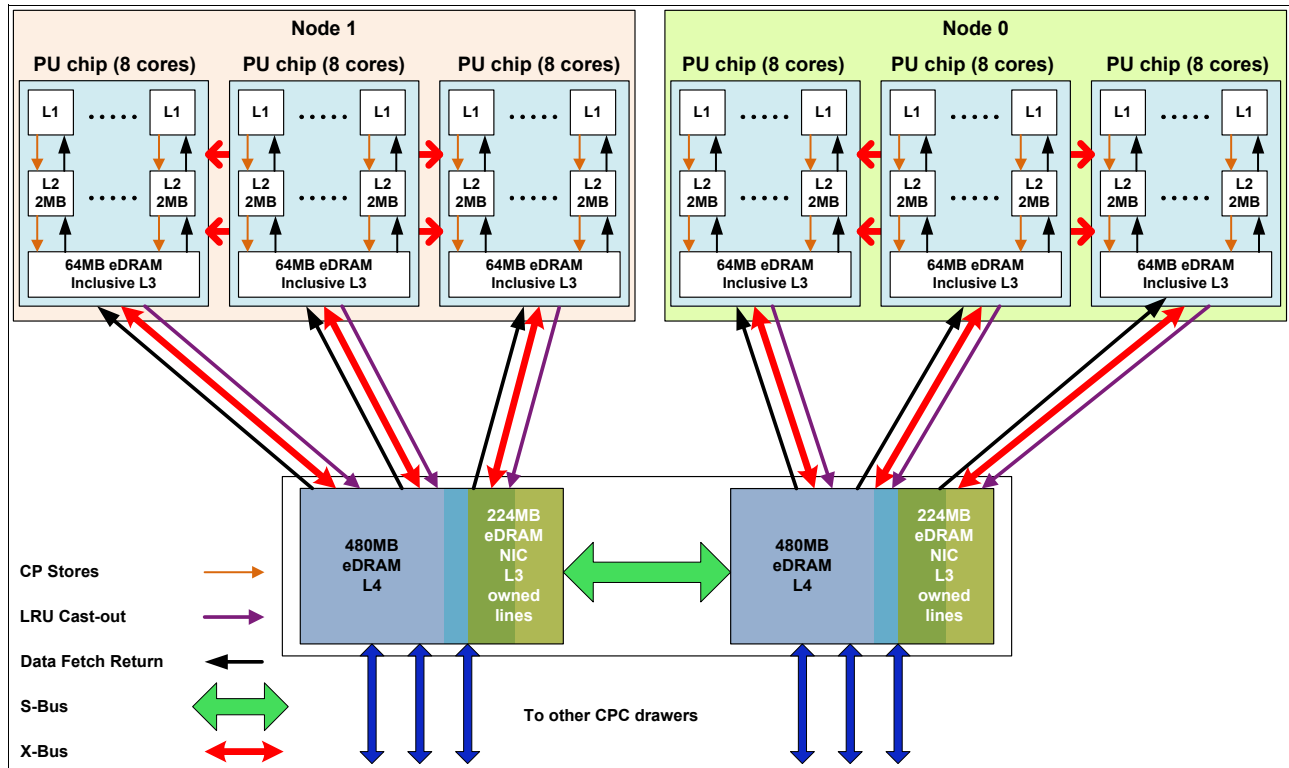


Figure 2-8 z13 CPC drawer chips and cache structure

Each SC chip packs 480 MB of eDRAM cache, interface logic for three PU chips each, and SMP fabric logic, into 678.78 mm². The two SC chips are configured to provide a single 960 MB L4 cache that is shared by all cores in the CPC drawer. This amount of cache gives a total of 3.8 GB of cache if all four CPC drawers are implemented, yielding outstanding SMP scalability on real-world workloads.

2.4 Processor chip

The z13 features a high-frequency, eight-core processor chip (PU chip), with an advanced microprocessor design, a robust cache hierarchy, that incorporates SMT² and SIMD³ architectures along with an SMP⁴ design that is optimized for enterprise database and transaction processing workloads, and for workloads such as Java and Linux.

The PU chip uses leading-edge technology and circuit design techniques while building on the rich heritage of mainframe system design, including industry-leading reliability, availability, and serviceability. Functions and features that are introduced with the z13 enable increased software efficiency and scalability, while maintaining full compatibility with existing software. Further details are given in 3.2.1, “Microprocessor” on page 67.

² Simultaneous multithreading (SMT): Two simultaneous threads running concomitantly on the same core

³ Single-instruction, multiple-data (SIMD): Vector processing unit (one instruction applied to a vector of data)

⁴ Symmetric multiprocessing (SMP)

2.5 Processor unit

Processor unit (PU) is the generic term for a z/Architecture processor. A PU is embedded in a z Systems chip core. Each PU is a superscalar processor with the following attributes:

- ▶ Up to six instructions can be decoded per clock cycle.
- ▶ Up to ten instructions can be in execution per clock cycle.
- ▶ Instructions can be issued out-of-order. A high-frequency, low-latency pipeline, providing robust performance across a wide range of workloads, is used.
- ▶ Memory accesses might not be in the same instruction order (out-of-order operand fetching).
- ▶ Most instructions flow through a pipeline with varying numbers of steps for various types of instructions. Several instructions can be in progress at any moment, subject to the maximum number of decodes and completions per cycle.

PU chip cache

The on-chip cache for the PU chip is as follows:

- ▶ Each PU (core) has an L1 cache (private) that is divided into a 96 KB cache for instructions and a 128 KB cache for data.
- ▶ Each PU (core) also has a private L2 cache, with 2 MB D-cache (D is data) and 2 MB I-cache (I is instruction).
- ▶ In addition, each PU chip contains an 64MB L3 cache, which is shared by all eight PUs (cores) on the chip. The shared L3 cache uses eDRAM.

The cache structure of z13 is shown in Figure 2-8 on page 34. This implementation optimizes performance of the system for high-frequency processors.

Each L1 cache has an associated *translation lookaside buffer (TLB)* of 512 entries. In addition, a secondary TLB is used to further enhance performance. This structure supports large working sets, multiple address spaces, and a two-level virtualization architecture.

CPU sparing

Hardware fault detection is embedded throughout the design and combined with comprehensive instruction-level retry and dynamic CPU sparing. This provides the reliability and availability that is required for true mainframe quality.

On-chip cryptographic hardware

Dedicated on-chip cryptographic hardware includes extended key and hash sizes for the Advanced Encryption Standard (AES), Secure Hash Algorithm (SHA) algorithms, and UTF8 to UTF16 conversion support. It is available to any processor type (CP, zIIP, IFL).

Software support

The z13 processor provides full compatibility with existing software for ESA/390 and z/Architecture, while extending the Instruction Set Architecture (ISA) to enable enhanced function and performance. Several hardware instructions that support more efficient code generation and execution are introduced. Examples are as follows:

- ▶ Hardware Decimal Floating-Point
- ▶ Transactional Execution Facility
- ▶ Runtime Instrumentation Facility
- ▶ Single-instruction, multiple-data (SIMD)

These features are further described in Chapter 3, “Key functions and capabilities of IBM z13™” on page 59.

Processor unit characterization

Processor units (PU) are ordered in single increments. The internal system functions, which are based on the configuration that is ordered, characterize each processor unit (core) into one of various types during initialization of the system, which is often called a *power-on reset (POR)* operation. Characterizing PUs dynamically without a POR is possible using a process called *Dynamic PU Reassignment*. A processor unit that is not characterized cannot be used.

Each PU (core) can be characterized as follows:

- ▶ 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)

At least one CP must be purchased with, or before, a zIIP can be purchased. Clients can purchase up to two zIIPs, for each purchased CP (assigned or unassigned) on the system. However, a logical partition definition can go behind the 1:2 ratio. For example, on a system with two CPs, a maximum of four zIIPs can be installed. A logical partition definition for that system can contain up to two logical CPs and four logical zIIPs. Another possible configuration would be one logical CP and three logical zIIPs.

Converting a PU from one type to any other type is possible again, using the *Dynamic PU Reassignment* process. These conversions happen concurrently with the operation of the system.

zIIP processors: The addition of ICFs, IFLs, zIIPs, and system assist processors (SAP) to a CPC does not change the system capacity setting or its MSU rating (only CPs do). IBM does not impose any software charges on work that is dispatched on zIIP processors.

zAAP processors: The zEC12 and zBC12 were the last z Systems servers to offer support for zAAP specialty engine processors. IBM supports running zAAP workloads on zIIP processors (“zAAP on zIIP”). This change is intended to help simplify capacity planning and performance management, while still supporting all the currently eligible workloads. IBM has provided a PTF for APAR OA38829 on z/OS V1R12 and V1R13. This PTF removes the restriction that prevents zAAP-eligible workloads from running on zIIP processors when a zAAP is installed on the server.

2.6 Memory

This section discusses memory subsystem for the z13.

2.6.1 z13 memory

Maximum physical memory size is directly related to the number of CPC drawers in the system. Because part of the physically installed memory is used to implement the RAIM design, a z13 system has more memory installed than ordered. This configuration results in up to 2.5 TB of available memory per CPC drawer and up to 10 TB for a four-drawer system.

The HSA memory has a fixed amount of 96 GB and is managed separately from client memory. Therefore, theoretically, up to 2.5 TB on the one CPC drawer model and up to 10 TB on the four CPC drawer models can be ordered. Because of some dependencies on the memory granularity, the maximum number of orderable memory can vary from the theoretical number.

Table 2-2 lists the maximum and minimum memory sizes for each z13 model.

Table 2-2 z13 models memory sizes

Model	Number of CPC drawers	Client memory (GB)
N30	1	256 GB to 2560 GB
N63	2	512 GB to 5120 GB
N96	3	768 GB to 7669 GB
NC9	4	1024 GB to 10 TB
NE1	4	1024 GB to 10 TB

On z13 systems, the granularity for memory orders varies from 32 GB to 512 GB. Table 2-3 shows the memory granularity depending on installed client memory.

Table 2-3 z13 memory granularity

Memory Increment (GB)	Client (offered) memory (GB)
32	64 to 192
64	256 to 448
96	544 to 928
128	1056 to 1440
256	1696 to 6048
512	6560 to 10144

Physically, memory is organized in the following ways:

- ▶ A CPC drawer always contains a minimum of 320 GB physically installed memory where 256 GB are usable memory.
- ▶ A CPC drawer can have more installed memory than enabled. The excess amount of memory can be enabled by a Licensed Internal Code load.
- ▶ Memory upgrades are satisfied from already installed unused memory capacity until exhausted. When no more unused memory is available from the installed memory cards, either the cards must be upgraded to a higher capacity or the addition of a CPC drawer with more memory is necessary.

When activated, PR/SM tries to allocate a logical partition's memory in a single CPC drawer and, if not possible, it can use memory resources located in any CPC drawer. No matter which CPC drawer the memory is in, a logical partition has access to that memory if allocated. Despite the CPC drawer structure, the z13 is still a *symmetric multiprocessor (SMP)* system.

A memory upgrade is concurrent when it requires no change of the physical memory cards. A memory card change is disruptive when no use is made of Enhanced Drawer Availability

(EDA). See *IBM z13 Technical Guide*, SG24-8251, for a description of *Enhanced Drawer Availability*.

For a model upgrade that results in the addition of a CPC drawer, the minimum memory increment is added to the system. Remember that the minimum usable memory size in a CPC drawer is 256 GB. During a model upgrade, the addition of a CPC drawer is a concurrent operation. The addition of the physical memory that is in the added drawer is also concurrent.

2.6.2 Concurrent memory upgrade

Memory can be upgraded concurrently by using *Licensed Internal Code Configuration Control (LICCC)* if physical memory is available, as described in the previous section. The *plan ahead memory function* that is available with the z13 enables nondisruptive memory upgrades by having in the system pre-plugged memory, which is based on a target configuration. Pre-plugged memory is enabled through an LICCC order that is placed by the client.

2.6.3 Redundant array of independent memory

The z196 introduced the *redundant array of independent memory (RAIM)* to z Systems, making the memory subsystem essentially a fully fault-tolerant N+1 design. RAIM design automatically detects and recovers from failures of dynamic random access memory (DRAM), sockets, memory channels, or dual inline memory module (DIMM).

The RAIM design is fully integrated in the z13 and is enhanced to include one Memory Controller Unit (MCU) per processor chip with five memory channels and one DIMM per channel. A fifth channel in each MCU enables memory to be implemented as a RAIM. This technology has significant RAS capabilities in error correction. Bit, lane, DRAM, DIMM, socket, and complete memory channel failures, including many types of multiple failures, can be detected and corrected.

Patented error correction technology in the memory subsystem provides the most robust error correction from IBM to date. Two full DRAM failures per rank can be spared and a third full DRAM failure corrected. DIMM level failures, including components such as the memory controller application-specific integrated circuit (ASIC), the power regulators, the clocks, and the system board can be corrected.

Memory channel failures, such as signal lines, control lines, and drivers/receivers on the SCM, can be corrected. Upstream and downstream data signals can be spared by using two spare wires on both the upstream and downstream paths. One of these signals can be used to spare a clock signal line (one upstream and one downstream).

2.6.4 Hardware system area

The *hardware system area (HSA)* is a fixed-size reserved area of memory, separate from the client-purchased memory. The HSA is used for several internal functions, but the bulk is used by channel subsystem functions. The HSA has grown with each successive mainframe generation. On older systems, model upgrades and also new logical partition definitions or changes required preplanning and were sometimes disruptive.

The fixed size 96 GB HSA of the z13 is large enough to accommodate any logical partition (LPAR) definitions or changes, thus eliminating outage situations. In addition, planning needs are eliminated.

A fixed large HSA allows the dynamic I/O capability to be enabled by default. It also enables the dynamic addition and removal, without planning, of the following features:

- ▶ Logical partition to new or existing channel subsystem (CSS)
- ▶ CSS (up to six can be defined on z13)
- ▶ Subchannel set (up to four can be defined on z13)
- ▶ Devices, up to the maximum that is permitted, in each subchannel set
- ▶ Logical processors by type
- ▶ Cryptographic adapters

2.7 I/O system structure

The z13 supports two types of internal I/O infrastructure:

- ▶ New generation 3 PCIe-based infrastructure for PCIe I/O drawers (PCIe Gen 3),
- ▶ InfiniBand-based infrastructure and I/O drawers (carry forward on an MES only)

The PCIe I/O infrastructure consists of the following features:

- ▶ New PCIe Gen3 fanouts in the z13 CPC drawer which support 16 GBps connectivity to the PCIe I/O drawer (zEC12 used PCIe Gen2 fanouts at 8 GBps)
- ▶ Up to five 7U PCIe I/O drawers each with 32 slots (eight slots per I/O domain) for PCIe I/O features

The InfiniBand I/O infrastructure (carry forward only) consists of the following features:

- ▶ InfiniBand fanouts in the z13 CPC drawer, which support the 6 GBps InfiniBand I/O interconnect
- ▶ InfiniBand I/O card domain multiplexers with redundant I/O interconnect in the following configuration: up to two, 5U, 8-slot, 2-domain I/O drawer (carry forward only).
- ▶ Selected non-PCIe I/O features (carry forward only)

Ordering of I/O features: Ordering of I/O feature types determines the appropriate mix of PCIe I/O drawers and I/O drawers (order dependent).

Figure 2-9 shows a high-level view of the I/O system structure for the z13

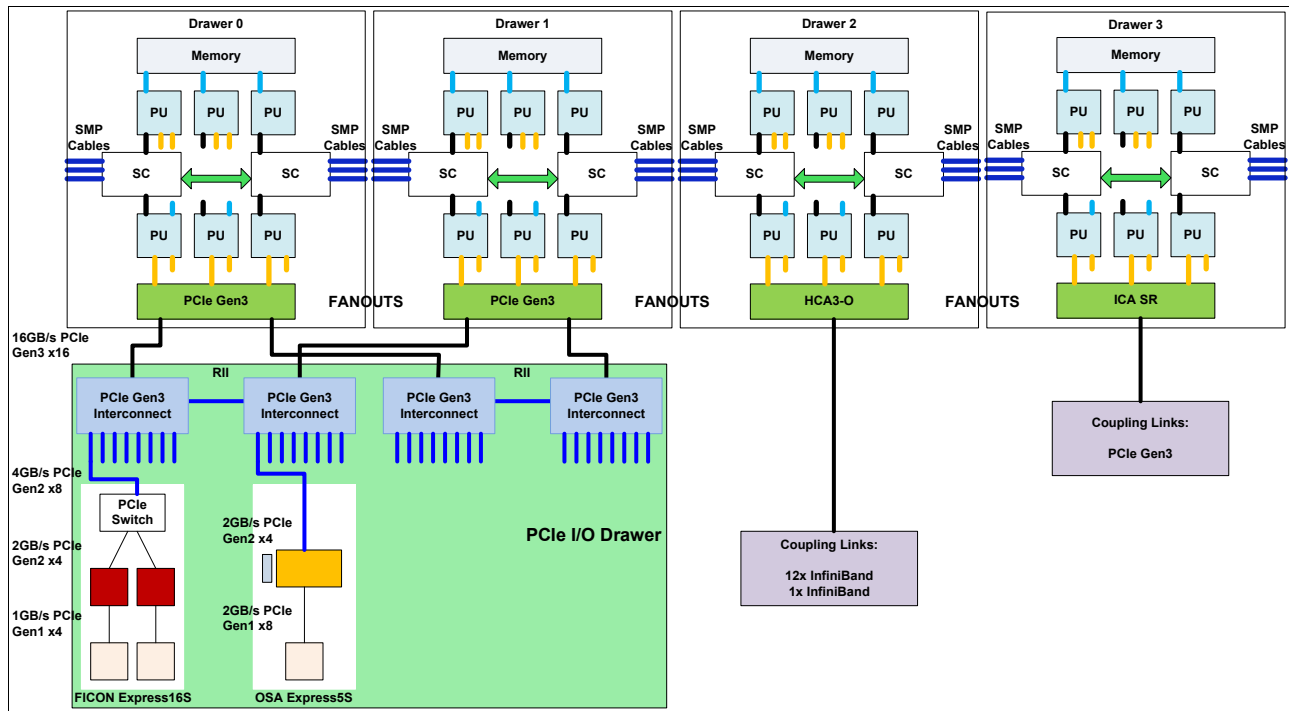


Figure 2-9 z13 I/O system structure

The z13 supports five fanout types (for fanout location see also Figure 2-10 on page 41), which are located at the front on the z13 CPC drawer:

- ▶ ICA SR
- ▶ HCA3-O
- ▶ HCA3-O LR
- ▶ PCIe Gen3
- ▶ HCA2-C (carry forward only)

The HCA3-O LR fanout includes four ports, the PCIe Gen3 fanout with one, and all the other fanouts have two ports.

The following types of internal I/O connectivity support the PCIe I/O drawer and I/O drawer:

- ▶ PCIe connections to the PCIe I/O drawers from the PCIe fanouts via copper cables. The I/O features supported through the one port on the fanouts are listed in 2.8, "I/O features" on page 41.
- ▶ InfiniBand (IFB) connections to the legacy I/O drawers from the Host Channel Adapter (HCA2-C) fanouts via copper cables in a carry forward MES⁵ only. The two ports in the fanout are dedicated to connect to an InfiniBand multiplexer (IFB-MP) card in the I/O drawer.

For coupling link connectivity (Parallel Sysplex or STP configuration), the z13 supports the following fanouts:

- ▶ ICA SR
- ▶ HCA3-O
- ▶ HCA3-O LR

⁵ MES - Miscellaneous Equipment Specification (upgrade or change)

The z13 CPC drawer (Figure 2-10) can have up to ten 1-port PCIe Gen3 fanouts (numbered LG02 to LG06 and LG11 to LG15) and up to four 2-port or 4-port InfiniBand fanouts for each CPC drawer, (numbered LG07 to LG10), which are used to connect to I/O drawers, PCIe I/O drawers, or for Parallel Sysplex InfiniBand and PCIe connectivity. In a system that is configured for maximum availability, alternate paths maintain access to critical I/O devices, such as disks and networks.

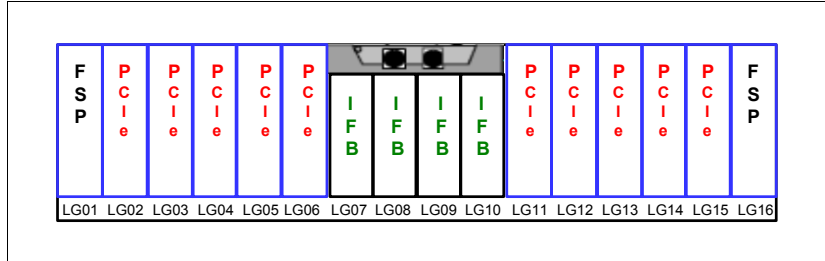


Figure 2-10 z13 CPC drawer front view

The PCIe I/O drawer is a two-sided drawer (I/O features on both sides) that is 7U high. The drawer contains 32 slots, four switch cards (two in the front and two in the rear) to support four I/O domains, each containing eight features of any type (FICON Express16S, FICON Express8S, OSA-Express5S, OSA-Express4S, Crypto Express5S, Crypto Express4S, Flash Express, zEDC Express, and 10GbE RoCE Express). Two DCAs to provide the redundant power, and two air moving devices (AMD) provide redundant cooling to the PCIe I/O Drawer.

Each (legacy) I/O drawer supports two I/O domains (A and B) for a total of eight I/O slots. Each I/O domain uses an IFB-MP card in the I/O drawer and a copper cable to connect to a Host Channel Adapter (HCA) fanout in the CPC drawer.

All features in the I/O drawer are installed horizontally. The two DCAs distribute power to the I/O drawer.

The IFB-MP cards are installed at location 09 at the rear side of the I/O drawer. The I/O features are installed from the front and rear sides of the I/O drawer. Two I/O domains are supported. Each I/O domain has up to four FICON Express8 features. The FICON Express 8 I/O features are connected to the IFB-MP card through the backplane board.

2.8 I/O features

The z13 supports two internal I/O infrastructure types:

- ▶ A PCIe-based infrastructure for PCIe I/O drawers, which supports the following I/O features:
 - FICON Express16S
 - FICON Express8S
 - OSA-Express5S
 - OSA-Express4S
 - 10GbE RoCE Express
 - Crypto Express5S
 - Crypto Express4S
 - Flash Express
 - zEDC Express
- ▶ An InfiniBand-based infrastructure for I/O drawers, which supports the following I/O features:

– FICON Express8

All migration offerings and exchange programs offer FICON Express8S, OSA-Express5S, OSA-Express4S, Flash Express, zEDC Express, 10GbE RoCE Express and Crypto Express4S features as carry forward only. The following features are no longer orderable for z13:

- ▶ ESCON
- ▶ FICON Express4
- ▶ OSA-Express3
- ▶ ISC-3
- ▶ Crypto Express3

See Table B-1 on page 168 for more details.

Features available only when carried forward: FICON Express8, OSA-Express4S, Crypto Express4S, features are available only when carried forward on upgrades.

2.8.1 FICON Express16S

Two types of transceivers for *FICON Express16S* are supported on a new build z13, one long wavelength (LX) laser version, and one short wavelength (SX) laser version:

- ▶ FICON Express16S LX feature
- ▶ FICON Express16S SX feature

Each port supports attachment to the following elements:

- ▶ FICON/FCP switches and directors that support 4 Gbps, 8 Gbps, or 16 Gbps
- ▶ Control units (Storage subsystems) that support 4 Gbps, 8 Gbps, or 16 Gbps

FICON Express16S LX feature

The *FICON Express16S LX* feature occupies one I/O slot in the PCIe I/O drawer. It has two ports, each supporting an LC duplex connector and auto-negotiated link speeds of 4 Gbps, 8 Gbps, and 16 Gbps up to an unrepeated maximum distance of 10 km.

FICON Express16S SX feature

The *FICON Express16S SX* feature occupies one I/O slot in the PCIe I/O drawer. It has two ports, each supporting an LC duplex connector and auto-negotiated link speeds of 4 Gbps, 8 Gbps, and 16 Gbps up to an unrepeated maximum distance⁶ of up to 380 meters at 4 Gbps, 150 meters at 8 Gbps, or 100 meters at 16 Gbps.

2.8.2 FICON Express8S

Statement of Direction^a: The IBM z13 server will be the last z Systems server to offer ordering of FICON Express8S channel features. Enterprises that have 2 Gb device connectivity requirements must carry forward these channels.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

⁶ Distances valid for OM3 cabling

Two types of transceivers for *FICON Express8S* are supported on a new build⁷ z13, one long wavelength (LX) laser version, and one short wavelength (SX) laser version:

- ▶ FICON Express8S 10KM LX feature
- ▶ FICON Express8S SX feature

Each port supports attachment to the following elements:

- ▶ FICON/FCP switches and directors that support 2 Gbps, 4 Gbps, or 8 Gbps
- ▶ Control units that support 2 Gbps, 4 Gbps, or 8 Gbps

FICON Express8S 10KM LX feature

The *FICON Express8 10KM LX* feature occupies one I/O slot in the I/O drawer. It has four ports, each supporting an LC duplex connector, and auto-negotiated link speeds of 2 Gbps, 4 Gbps, and 8 Gbps up to an unrepeated maximum distance of 10 km.

FICON Express8S SX feature

The *FICON Express8S SX* feature occupies one I/O slot in the I/O drawer. This feature has four ports, each supporting an LC duplex connector, and auto-negotiated link speeds of 2 Gbps, 4 Gbps, and 8 Gbps up to an unrepeated maximum distance⁸ of up to 500 meters at 2 Gbps, 380 meters at 4 Gbps, or 150 meters at 8 Gbps.

2.8.3 FICON Express8 (carry forward only)

The IBM z13 will be the last z Systems server to support FICON Express8 channels: IBM z13 will be the last high-end server to support FICON Express8. Enterprises should begin migrating from FICON Express8 channel features (#3325, #3326) to FICON Express16S channel features (#0418, #0419). FICON Express8 will not be supported on future high-end z Systems servers as carry forward on an upgrade.

The *FICON Express8* features are available only when carried forward on upgrades. Two types of transceivers for FICON Express8 are supported on z13:

- ▶ FICON Express8 10KM LX feature
- ▶ FICON Express8 SX feature

FICON Express8 10KM LX feature

The *FICON Express8 10KM LX* feature occupies one I/O slot in the I/O drawer. It has four ports, each supporting an LC duplex connector, and auto-negotiated link speeds of 2 Gbps, 4 Gbps, and 8 Gbps up to an unrepeated maximum distance of 10 km.

FICON Express8 SX feature

The *FICON Express8 SX* feature occupies one I/O slot in the I/O drawer. This feature has four ports, each supporting an LC duplex connector, and auto-negotiated link speeds of 2 Gbps, 4 Gbps, and 8 Gbps up to an unrepeated maximum distance⁸ of up to 500 meters at 2 Gbps, 380 meters at 4 Gbps, or 150 meters at 8 Gbps.

⁷ FICON Express8S is offered on new build to support point to point 2 Gbps attachments

⁸ Distances valid for OM3 cabling

2.8.4 OSA-Express5S

This section describes the connectivity options that are offered by the *OSA-Express5S* features. The following OSA-Express5S features can be installed on z13:

- ▶ OSA-Express5S 10 Gigabit Ethernet (GbE) Long Reach (LR)
- ▶ OSA-Express5S 10 Gigabit Ethernet (GbE) Short Reach (SR)
- ▶ OSA-Express5S Gigabit Ethernet Long Wavelength (GbE LX)
- ▶ OSA-Express5S Gigabit Ethernet Short Wavelength (GbE SX)
- ▶ OSA-Express5S 1000BASE-T Ethernet

OSA-Express5S 10 GbE LR feature

The *OSA-Express5S 10 GbE LR* feature occupies one slot in a PCIe I/O drawer. It has one port that connects to a 10 Gbps Ethernet LAN through a 9 μ m single mode fiber optic cable that is terminated with an LC Duplex connector. The feature supports an unrepeated maximum distance of 10 km.

OSA-Express5S 10 GbE SR feature

The *OSA-Express5S 10 GbE SR* feature occupies one slot in the PCIe I/O drawer. This feature has one port that connects to a 10 Gbps Ethernet LAN through a 62.5 μ m or 50 μ m multimode fiber optic cable that is terminated with an LC Duplex connector. The maximum supported unrepeated distance is 33 m on a 62.5 μ m multimode fiber optic cable, and 300 m on a 50 μ m multimode fiber optic cable.

OSA-Express5S GbE LX feature

The *OSA-Express5S GbE LX* occupies one slot in the PCIe I/O drawer. This feature has two ports, representing one *channel path identifier (CHPID)*, that connect to a 1 Gbps Ethernet LAN through a 9 μ m single mode fiber optic cable. This cable is terminated with an LC Duplex connector, supporting an unrepeated maximum distance of 5 km. A multimode (62.5 or 50 μ m) fiber optic cable can be used with this feature. The use of these multimode cable types requires a *mode conditioning patch (MCP)* cable at each end of the fiber optic link. Use of the single mode to multimode MCP cables reduces the supported distance of the link to a maximum of 550 meters.

OSA-Express5S GbE SX feature

The *OSA-Express5S GbE SX* occupies one slot in the PCIe I/O drawer. This feature has two ports, representing one CHPID, that connect to a 1 Gbps Ethernet LAN through 50 or 62.5 μ m multimode fiber optic cable. This cable is terminated with an LC Duplex connector over an unrepeated distance of 550 meters (for 50 μ m fiber) or 220 meters (for 62.5 μ m fiber).

OSA-Express5S 1000BASE-T feature

The *OSA-Express5S 1000BASE-T* occupies one slot in the PCIe I/O drawer. It has two ports, representing one CHPID, that connect to a 1000 Mbps (1 Gbps), 100 Mbps, or 10 Mbps Ethernet LAN. Each port has an RJ-45 receptacle for UTP Cat5 or Cat6 cabling, which supports a maximum distance of 100 meters.

2.8.5 OSA-Express4S (carry forward only)

This section describes the connectivity options that are offered by the *OSA-Express4S* features. The following OSA-Express4S features can be installed on z13:

- ▶ OSA-Express4S 10 Gigabit Ethernet (GbE) Long Reach (LR)
- ▶ OSA-Express4S 10 Gigabit Ethernet Short Reach (SR)
- ▶ OSA-Express4S Gigabit Ethernet long wavelength (GbE LX)

- ▶ OSA-Express4S Gigabit Ethernet short wavelength (GbE SX)
- ▶ OSA-Express4S 1000BASE-T Ethernet

OSA-Express4S 10 GbE LR feature

The *OSA-Express4S 10 GbE LR* feature occupies one slot in a PCIe I/O drawer. It has one port that connects to a 10 Gbps Ethernet LAN through a 9 μm single mode fiber optic cable that is terminated with an LC Duplex connector. The feature supports an unrepeated maximum distance of 10 km.

OSA-Express4S 10 GbE SR feature

The *OSA-Express4S 10 GbE SR* feature occupies one slot in the PCIe I/O drawer. This feature has one port that connects to a 10 Gbps Ethernet LAN through a 62.5 μm or 50 μm multimode fiber optic cable that is terminated with an LC Duplex connector. The maximum supported unrepeated distance is 33 m on a 62.5 μm multimode fiber optic cable, and 300 m on a 50 μm multimode fiber optic cable.

OSA-Express4S GbE LX feature

The *OSA-Express4S GbE LX* occupies one slot in the PCIe I/O drawer. This feature has two ports, representing one *channel path identifier (CHPID)*, that connect to a 1 Gbps Ethernet LAN through a 9 μm single mode fiber optic cable. This cable is terminated with an LC Duplex connector, supporting an unrepeated maximum distance of 5 km. A multimode (62.5 or 50 μm) fiber optic cable can be used with this feature. The use of these multimode cable types requires a *mode conditioning patch (MCP)* cable at each end of the fiber optic link. Use of the single mode to multimode MCP cables reduces the supported distance of the link to a maximum of 550 meters.

OSA-Express4S GbE SX feature

The *OSA-Express4S GbE SX* occupies one slot in the PCIe I/O drawer. This feature has two ports, representing one CHPID, that connect to a 1 Gbps Ethernet LAN through 50 or 62.5 μm multimode fiber optic cable. This cable is terminated with an LC Duplex connector over an unrepeated distance of 550 meters (for 50 μm fiber) or 220 meters (for 62.5 μm fiber).

OSA-Express4S 1000BASE-T feature

The *OSA-Express4S 1000BASE-T* occupies one slot in the PCIe I/O drawer. It has two ports, representing one CHPID, that connect to a 1000 Mbps (1 Gbps), 100 Mbps, or 10 Mbps Ethernet LAN. Each port has an RJ-45 receptacle for UTP Cat5 or Cat6 cabling, which supports a maximum distance of 100 meters.

2.8.6 Flash Express

The *Flash Express* is an innovative optional feature introduced with the zEC12 and also available on the z13. It is intended to provide performance improvements and better availability for critical business workloads that cannot afford any hits to service levels. Flash Express is easy to configure, requires no special skills, and provides rapid time to value.

Flash Express implements storage-class memory (SCM) through internal NAND Flash solid-state drive (SSD), in a PCIe card form factor. The Flash Express feature is designed to allow each logical partition to be configured with its own SCM address space.

For availability, this feature is available in pairs of cards. Each feature offers a capacity of 1.4 TB of usable storage per pair of cards. A maximum of four pairs of cards can be installed on a z13, providing a maximum capacity of 5.6 TB of storage.

2.8.7 zEDC Express

The *zEDC Express* is an optional feature, available to the z13, zEC12 and zBC12. It is designed to provide hardware-based acceleration for data compression and decompression for the enterprise, helping to improve cross platform data exchange, reduce CPU consumption, and save disk space.

A minimum of one feature can be ordered and a maximum of eight can be installed on the system, in the PCIe I/O drawer. Up to two zEDC Express features per domain can be installed. There is one PCIe adapter/compression coprocessor per feature which implements compression as defined by RFC1951 (DEFLATE)⁹. A zEDC Express feature can be shared between up to 15 LPARs.

The zEDC Express is a native PCI feature; the management functions are provided by Resource Groups (RG) running on the integrated firmware processor (IFP). See 3.2.3, “Native PCIe features and integrated firmware processor” on page 74 for more details about RGs and IFP.

For resilience, there are always two independent RGs on the system, sharing the IFP. Thus, it is recommended that a minimum of two zEDC Express features be installed, one per RG.

Consider also the total data throughput required and that, in the case of one feature becoming unavailable, the others should be able to absorb the load. Thus, for best data throughput and availability, the suggestion is that at least two features per RG are installed.

2.8.8 10 Gigabit Ethernet RoCE Express

The 10 Gigabit Ethernet (10GbE) RoCE Express feature is designed to help reduce consumption of CPU resources for applications utilizing the TCP/IP stack and may also help to reduce network latency with memory-to-memory transfers using Shared Memory Communications - Remote Direct Memory Access (SMC-R) in z/OS V2R1. It is transparent to applications and can be used for LPAR-to-LPAR communication on a single system or server-to-server communication in a multiple CPC environment.

This feature resides exclusively in the PCIe I/O drawer and is available to the z13, zEC12 and zBC12. The 10GbE RoCE Express feature has one PCIe adapter with two ports.

The 10GbE RoCE Express feature utilizes a short reach (SR) laser as the optical transceiver, and supports use of a multi-mode fiber optic cable terminated with an LC Duplex connector. Both point to point connection and switched connection with an enterprise-class 10 GbE switch are supported. Switch used by 10GbE RoCE Express feature must have *Pause frame* enabled as defined by the IEEE 802.3x standard.

The 10GbE RoCE Express feature does not use a CHPID. It is defined using the Input/Output Configuration Program (IOCP) **FUNCTION** statement or in the Hardware Configuration Definition (HCD).

A maximum of 16 features can be installed per system. With z13, the 10GbE RoCE adapters can be shared by up to 31 LPARs. Also, both adapter ports are now supported by z/OS, when running on z13.

⁹ See <http://www.ietf.org/rfc/rfc1951.txt>

2.8.9 Coupling links and timing-only links

Coupling links provide for communication in a Parallel Sysplex environment. They are further discussed in 2.9, “Coupling and clustering” on page 48.

2.8.10 Cryptographic features

Cryptographic coprocessor and accelerator functions can be provided in addition to the synchronous functions of the internal cryptographic coprocessor by the PCIe cryptographic adapters. z13 supports the Crypto Express5S feature and Crypto Express4S¹⁰.

CPACF on z13

One cryptographic coprocessor is part of each PU core of the SCM. SHA-1, and SHA-2 support for SHA-224, SHA-256, SHA-384, and SHA-512 are shipped enabled on all servers running synchronously on this coprocessors. All other cryptographic synchronous functions and the support of extra installed PCIe cryptographic features are provided by the CPACF. The CPACF must be explicitly enabled by using an enablement feature (feature code 3863) that is available for no extra fee. CPACF is available for every PU characterized as CP, IFL, or zIIP.

Crypto Express5S

The *Crypto Express5S* features provide the following capabilities:

- ▶ The Crypto Express5S feature occupies one I/O slot in a z13 PCIe I/O drawer.
- ▶ The Crypto Express5S feature has one PCIe adapter with one PCHID assigned to it according to its physical location in the PCIe I/O drawer.
- ▶ There is no need to define a CHPID for the Crypto Express5S feature in the HCD/IOCP. Care must be taken to avoid the use of Crypto Express5S' associated PCHID by another device in the HCD/IOCP definition.
- ▶ On z13 each Crypto Express5S PCI Express adapter can be configured as one of the following environments:
 - Coprocessor with the following characteristics:
 - *Secure IBM Common Cryptographic Architecture (CCA) coprocessor* for Federal Information Processing Standard (FIPS) 140-2 Level 4 certification.
 - *Secure IBM Enterprise PKCS #11 (EP11) coprocessor* implements industry standardized set of services that adhere to the PKCS #11 specification v2.20. A Trusted Key Entry (TKE) Workstation with a smart card reader feature is required to support the administration of the Crypto Express5S, when configured as an Enterprise PKCS #11 coprocessor.
 - *Accelerator* for public key and private key cryptographic operations that are used with Secure Sockets Layer and Transport Layer Security (SSL/TLS) processing.

These modes can be configured by the Support Element, and the PCIe adapter must be configured offline to change the mode.

- ▶ Up to 16 Crypto Express4S features are supported (16 PCI Express adapters per z13 system).

¹⁰ Crypto Express4S: Carry forward only

Crypto Express4S (carry forward only)

Crypto Express4S features provide the following capabilities:

- ▶ The Crypto Express4S feature occupies one I/O slot in a z13 I/O drawer.
- ▶ The Crypto Express4S feature has one PCIe adapter with one PCHID assigned to it according to its physical location in the PCIe I/O drawer.
- ▶ There is no need to define a CHPID for the Crypto Express4S feature in the HCD/IOCP. Care must be taken to avoid the use of Crypto Express4S associated PCHID by another device in the HCD/IOCP definition.
- ▶ On z13, each Crypto Express4S PCI Express adapter can be configured as one of the following environments:
 - Coprocessor with the following characteristics:
 - *Secure IBM Common Cryptographic Architecture (CCA) coprocessor* for Federal Information Processing Standard (FIPS) 140-2 Level 4 certification.
 - *Secure IBM Enterprise PKCS #11 (EP11) coprocessor* implements industry standardized set of services that adhere to the PKCS #11 specification v2.20. A Trusted Key Entry (TKE) workstation with a smart card reader feature is required in order to support the administration of the Crypto Express4S, when configured as an Enterprise PKCS #11 coprocessor.
 - *Accelerator* for public key and private key cryptographic operations that are used with Secure Sockets Layer / Transport Layer Security (SSL/TLS) processing.

These modes can be configured by the Support Element, and the PCIe adapter must be configured offline to change the mode.

Crypto Express4S PCIe configuration: When the Crypto Express4S PCIe adapter is configured as a secure IBM CCA coprocessor, it still provides accelerator functions. However, up to three times better performance for those functions can be achieved if the Crypto Express4S PCIe adapter is configured as an accelerator.

- ▶ Up to 16 Crypto Express4S features are supported (16 PCI Express adapters per z13).

2.9 Coupling and clustering

In the past, Parallel Sysplex communications support has been provided over several types of connection, such as InterSystem Coupling (ISC), Integrated Cluster Bus (ICB), and Internal Coupling (IC), each of which (except IC) involves unique development effort for the support code and for the hardware.

Coupling connectivity for Parallel Sysplex on z13 uses new Integrated Coupling Adapter (ICA SR) and InfiniBand (IFB) technology. The ICA SR is designed to drive distances up to 150 m. The HCA3-O LR fanout supports longer distances between systems using the IFB technology. The older versions of InfiniBand adapters HCA2-O and HCA2-O LR are not supported on z13.

ICA SR and InfiniBand technologies allow moving all of the Parallel Sysplex connectivity support to interfaces that provides high-speed interconnection at short distances and longer distance fiber optic interconnection.

See the *Coupling Facility Configuration Options* white paper for a more specific explanation regarding the coupling links technologies. The white paper is available at the following web page:

<http://www.ibm.com/systems/z/advantages/ps0/whitepaper.html>

For details about all InfiniBand features, see either of these resources:

- ▶ *IBM z Systems Connectivity Handbook*, SG24-5444
- ▶ *Implementing and Managing InfiniBand Coupling Links on System z*, SG24-7539

2.9.1 Integrated Coupling Adapter (ICA SR)

The IBM Integrated Coupling Adapter (ICA SR), introduced on the z13 platform, is a two-port fanout used for short distance coupling connectivity and utilizes a new coupling channel type: CS5. The ICA utilizes PCIe Gen3 technology, with x16 lanes that are bifurcated into x8 lanes for coupling. No performance degradation is expected compared to Coupling over InfiniBand 12X IFB3 protocol.

The ICA SR is designed to support cable length of up to 150m and support a link data rate of 8 GBps. It is also designed to support up to 4 CHPIDs per port and 7 subchannels (devices) per CHPID. The coupling links can be defined as shared between images within a CSS. They can also be spanned across multiple CSSs in a CPC. Unlike the HCA3-O 12x InfiniBand links, the ICA SR cannot define more than four CHPIDs per port.

2.9.2 InfiniBand coupling links

Two types of *Host Channel Adapter (HCA)* fanouts are used for IFB coupling links on the z13:

- ▶ HCA3-O fanout, which supports 12x InfiniBand (12x IFB)
- ▶ HCA3-O Long Reach (LR) fanout, which supports 1x InfiniBand (1x IFB)

HCA3s are the latest generation of InfiniBand host channel adapters for coupling. The HCA3-O fanout for 12x InfiniBand (12x IFB) is designed for improved service times and is available for z Systems CPCs using the 12x InfiniBand3 (12x IFB3) protocol. The HCA3-O LR fanout for 1x InfiniBand (1x IFB) provides four ports and optional additional subchannels for extended-distance solutions.

InfiniBand coupling link data rate: The InfiniBand coupling link data rate (6 GBps, 3 GBps, 5 Gbps, or 2.5 Gbps) does not represent the performance of the link. The actual performance depends on many factors, including latency through the adapters, cable lengths, and the type of workload.

The 12x InfiniBand coupling links support double data rate (DDR) at 6 GBps for a z13 to z13 or to z Systems CPCs:

- ▶ InfiniBand (HCA3-O) coupling links (12x IFB), used for z/OS-to-CF communication, CF-to-CF traffic, or STP messaging at distances up to 150 meters (492 feet) by using industry standard OM3 50 µm fiber optic cables.

When no more than four CHPIDs are defined per port, and an HCA3-O to HCA3-O connection is set up, the IFB3 protocol is used. When using the IFB3 protocol, synchronous service times are 40% faster than when using the IFB protocol.

An HCA3-O to HCA2-O connection is supported, but the standard IFB protocol is used.

HCA3-O LR 1x InfiniBand coupling links support up to 32 subchannels (devices) per CHPID, versus the current default value of seven devices per CIB type CHPID:

- ▶ InfiniBand (HCA3-O LR) coupling links (1x IFB) for z/OS-to-CF communication at unrepeated distances up to 10 km (6.2 miles) using 9 μ m single mode fiber optic cables and repeated distances up to 100 km (62 miles) using IBM z Systems qualified DWDM equipment. (Connectivity to HCA2-O LR is supported).

The HCA3-O LR has four ports, the number of supported CHPIDs remains at 16 for the fanout card.

HCA2-O and HCA2-O LR features: HCA2-O and HCA2-O LR features are not available on z13 but may be present in an existing parallel sysplex environment. These adapters were supported in a carry forward MES migrating to the zEC12 and/or zBC12 families.

Time source for Server Time Protocol traffic

IFB and ICA SR links can be used to carry Server Time Protocol (STP) timekeeping information.

For details about all InfiniBand features, see either of these resources:

- ▶ *IBM z Systems Connectivity Handbook*, SG24-5444
- ▶ *Implementing and Managing InfiniBand Coupling Links on System z*, SG24-7539

2.9.3 Internal Coupling (IC)

Internal Coupling (IC) links are used for internal communication between LPARs on the same system running coupling facilities (CF) and z/OS images. The connection is emulated in Licensed Internal Code (LIC) and provides for very fast and secure memory-to-memory communications between LPARs within a single system. No physical cabling is required.

2.9.4 Coupling Facility Control Code (CFCC)

This section discusses the various levels of Coupling Facility Control Code (CFCC).

CFCC Level 20

CFCC Level 20 is available on the z13 machines with the driver level 22 and includes the following enhancements:

- ▶ Support for up to 141 ICF processors
 - The maximum number of logical processors in a Coupling Facility Partition remains at 16
- ▶ Large memory support
 - Improve availability for larger CF cache structures and data sharing performance with larger DB2 Group Buffer Pools (GBP).
 - This support removes inhibitors to using large CF structures, enabling use of Large Memory to appropriately scale to larger DB2 Local Buffer Pools (LBP) and Group Buffer Pools (GBP) in data sharing environments.
 - CF structure size remains at a maximum of 1 TB
- ▶ Support for new IBM Integrated Coupling Adapter (ICA).

Statement of Direction^a (Coupling CHPIDs support): IBM plans to support up to 256 coupling link CHPIDs in z13 short, (that is twice the 128 coupling link CHPIDs supported on zEC12). Each CF image will continue to support a maximum of 128 coupling link CHPIDs.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

CF structure size changes are expected to grow when going from CFCC Level 19 (or earlier) to CFCC Level 20. We suggest reviewing the CF LPAR size by using the following tools:

- ▶ *CFSizer* tool, is web-based and is most useful when changing an existing workload or introducing a new one. CFSizer tool is available at the following web page:

<http://www.ibm.com/systems/support/z/cfsizer/>

- ▶ *Sizer Utility*, an authorized z/OS program download, most useful when upgrading a CF. The Sizer utility is available at the following web page:

<http://www.ibm.com/systems/support/z/cfsizer/altsize.html>

2.10 Time functions

Each server must have an accurate time source to maintain a time-of-day value.

Time functions are used to provide an accurate time-of-day value and to ensure that the time-of-day value is properly coordinated among all of the systems in a complex. These functions are critical for Parallel Sysplex operation.

IBM z Systems support attachment to an External Time Source (ETS) for clock information, support the Server Time Protocol, and can participate in a coordinated timing network (CTN).

2.10.1 Server Time Protocol (STP)

Server Time Protocol (STP) is a system-wide facility that is implemented in the Licensed Internal Code. The STP presents a single view of time to PR/SM and provides the capability for multiple CPCs and CFs to maintain time synchronization with each other. The enablement for using this protocol with the z13 is ensured by activating the optional STP feature.

More details about implementation of STP on z13 are in 3.4.1, "Server Time Protocol (STP)" on page 96.

2.10.2 Network Time Protocol support

Network Time Protocol (NTP) support is available on z13 and IBM z Systems. This implementation answers the need for a single time source across the heterogeneous platforms in the enterprise. With this implementation, the STP is synchronized by the use of a NTP server as time source.

Pulse Per Second

The z13 provides a dual-path interface for *Pulse Per Second (PPS)* support. STP tracks the highly stable accurate PPS signal from the NTP server. PPS maintains accuracy of 10 μ s as measured at the PPS input of the z13 CPC. If STP uses an NTP server without PPS, a time accuracy of 100 ms to the ETS is maintained. A cable connection from the PPS port to the

PPS output of an NTP server is required when the z13 is configured for using NTP with PPS as the external time source for time synchronization.

2.10.3 Time coordination for zBX components

NTP clients, running on blades in the zBX, can synchronize their time every hour with the NTP server provided by the zBX internal Support Element (SE). This synchronization provides the capability for the components in the zBX to maintain an approximate time accuracy of 100 milliseconds to an NTP server.

The battery operated clock (BOC) of the CPC SE synchronizes to the time-of-day (TOD) clock of the system, every hour.

2.11 Hardware Management Console and Support Element

The *Hardware Management Console (HMC)* and *Support Element (SE)* appliances together provide hardware platform management for IBM z Systems. Hardware platform management covers a complex set of configuration, operation, monitoring, service management tasks, and other services that are essential to the use of the hardware platform product.

With z13, the HMC can be a stand-alone desktop computer or an optional 1U rack-mounted computer. See 3.5, “Hardware Management Console (HMC) functions” on page 98 for more information about HMC capabilities.

The z13 is supplied with a pair of new integrated 1U SEs. One, the primary SE, is always active while the other is an alternate. Power for the SEs is supplied by the CPCs power supply, and there are no additional power requirements. The SEs are connected to the System Control Hubs (SCH) for network connectivity with the CPC and the HMCs.

The SCH replaces the Bulk Power Hub (BPH) which existed in the previous z Systems families.

The SEs and HMCs are closed systems, and no other applications can be installed on them.

The HMCs and SEs of the system are attached to a LAN. An HMC communicates with one or more z Systems and with the optional zBX model 004 own internal SEs, as shown in Figure 2-11. When tasks are performed on the HMC, the commands are sent to one or more SEs, which then issue commands to their CPCs and optional zBXs.

The HMC Remote Support Facility (RSF) provides communication with the IBM support network for hardware problem reporting and service.

Figure 2-11 shows an example of the HMC and SE connectivity.

RSF connection: RSF connection through a modem *is not supported* on the z13 HMC. An internet connection to IBM is required to have hardware problem reporting and service.

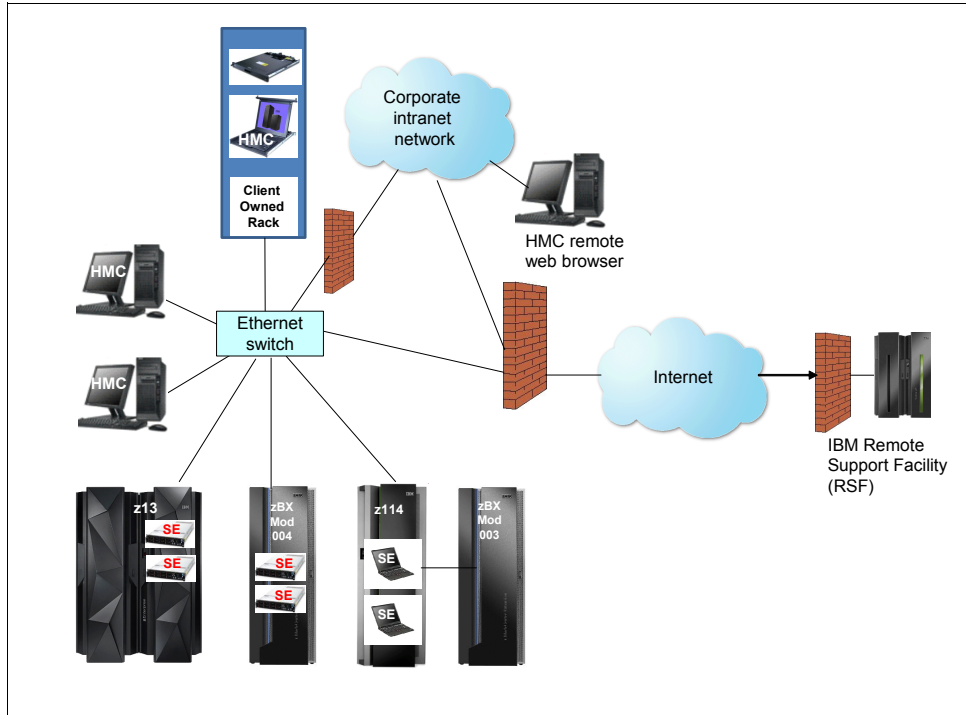


Figure 2-11 HMC and SE connectivity

2.12 Power and cooling

The power and cooling system of the z13 builds upon its predecessor, the zEC12, with the expansion of some significant newly developed technologies. The power service specifications of the z13 is almost same as those of their respective predecessors, and the total power consumption with the maximum configuration has increased only by approximately 5% as compared to the previous generation.

In the z13 a closed internal water cooling loop is used for PU SCMs cooling. Extracting the heat from the internal water loop can be done either with a radiator-cooled or a water-cooled option.

z13 radiator cooling option

The cooling system in z13 is redesigned for better availability and lower cooling power consumption. The radiator design is a closed-loop water cooling pump system for the PU SCMs. It is designed with N+2 pumps, blowers, controls, and sensors. The radiator unit is cooled by air.

z13 water cooling option

The z13 continues to offer the choice of using the building's chilled water to cool the system, by employing the water cooling unit (WCU) technology. The SCMs in the CPC drawer are cooled by an internal, closed, water cooling loop. In the internal closed loop, water exchanges heat with building chilled water through a cold plate. The source of building chilled water is provided by the client.

In addition to the SCMs, the internal water loop also circulates through two heat exchangers that are in the path of the exhaust air in the rear of the frames. These heat exchangers remove approximately 60 - 65% of the residual heat from the I/O drawers, the air-cooled logic

in the CPC drawers and the heat that is generated within the power enclosures. Almost two thirds of the total heat that is generated is removed from the system by the chilled water.

The z13 operates with two fully redundant water cooling units (WCU). One water cooling unit can support the entire load and the replacement of WCU is fully concurrent. If there is a total loss of building-chilled water or if both water cooling units fail, the rear door heat exchangers will cool the internal water cooling loop.

Conversion between air and water cooling systems is not available.

High Voltage Direct Current power feature

With the optional High Voltage Direct Current (HV DC) power feature, the z13 can directly connect to DC power input and improve data center energy efficiency by removing the need for an additional DC to AC inversion step. In addition to the data center UPS and power distribution energy savings, a zEnterprise CPC running on HV DC power draws 1 - 3% less input power.

Power considerations

The z13 operates with two sets of redundant power supplies. Each set of the power supplies has its individual power cords or pair of power cords, depending on the number of Bulk Power Regulator (BPR) pairs installed. Power cords attach a 3-phase, 50/60 Hz, 200 - 480 V AC power, or 380 - 520 V DC power. The total loss of one power supply has no effect on system operation.

The optional Balanced Power Plan Ahead feature is available for future growth, also assuring adequate and balanced power for all possible configurations. With this feature, downtime for upgrading a system is eliminated by including with the initial installation the maximum power requirements in terms of Bulk Power Regulators (BPR) and power cords. The Balance Power Plan Ahead feature is not available with DC and 1-phase line cords.

For ancillary equipment such as the Hardware Management Console, and its display, more single-phase outlets are required (customer provided).

The power requirements depend on the cooling facility that is installed, number of CPC drawers and the number and type of I/O units installed. Maximum power consumption tables for the various configurations and environments can be found in *IBM z13 Installation Manual for Physical Planning*, GC28-6938. See also the power and weight estimation tool available in IBM Resource Link®.

Top Exit Power

The z13 supports the (optional) *Top Exit Power* feature, which is combined with the *Top Exit I/O Cabling* feature, providing more flexibility to planning the computer room cabling. The radiator-cooled z13 models support installation on raised floor and non-raised floor environments. For water-cooled models, only the raised floor option is available.

2.13 IBM z BladeCenter Extension

The IBM z BladeCenter Extension (zBX) Model 004 only exists as an upgrade MES from a previous existing zBX model 003 or a zBX model 002. The zBX model 004 extends the z Systems qualities of service and management to integrate heterogeneous systems with high redundancy.

The zBX Model 004 (2458-004) connects to the z13 to become a *node* on its own, as part of an ensemble. The new zBX standalone node, in turn, creates an integrated multi-platform system with advanced virtualization management (through the Unified Resource Manager) that supports diverse workloads.

The zBX is configured with the following key components:

- ▶ Model 004 incorporates its own 1U Support Elements installed in the first zBX rack - Frame B.
- ▶ One to four standard 19-inch 42U IBM z Systems racks with required network and power infrastructure
- ▶ One to eight BladeCenter chassis with a combination of up to 112 blades¹¹
- ▶ Redundant infrastructure for fault tolerance and higher availability
- ▶ Management support through the z13 HMC and zBX Mod 004 SEs

The first rack (rack B) in the zBX is the primary rack where one or two BladeCenter chassis are located. Two pairs of Top of Rack (TOR) switches are included in rack B, one pair for the *intranode management network (INMN)* and another pair for the *intraensemble data network (IEDN)* connectivity as well as the new 1U SEs with their respective keyboard and display. The other three racks (C, D, and E) are expansion racks with one or two BladeCenter chassis each.

The zBX is managed from the primary ensemble HMC, connected to its internal primary SE via the existing client provided HMC/SE switch, as an ensemble node. The ensemble HMC will discover the zBX SE when the physical connectivity is established as it does for the CPCs today.

The IEDN provides private and secure 10 GbE high-speed data paths between all elements of an ensemble and the zBX node through the IEDN TOR switches. The IEDN connections use MAC addresses, not IP addresses (Layer 2 connection). The OSA-Express for zBX (OSX) CHPID type supports connectivity and access control from the z13 (or other z Systems) to the zBX node.

¹¹ Current maximum number of blades supported: 112 POWER7, 28 for DataPower XI50z, 56 for System x HX5.

Figure 2-12 shows the CEC and the zBX node connections through the OSA-Express5S 10 GbE or OSA-Express4S 10 GbE features (CHPID type OSX) in the z13.

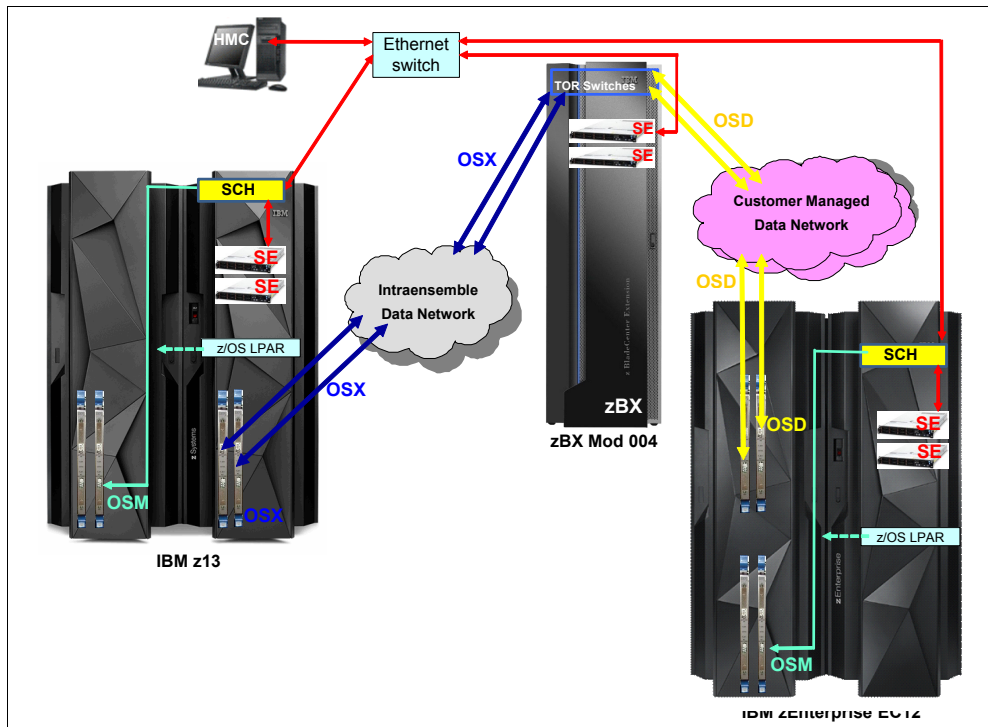


Figure 2-12 IEDN, and client-managed data networks in an ensemble

Optionally, as part of the ensemble, any OSA-Express5S, OSA-Express4S, or OSA-Express3 features (with CHPID type OSD) in z Systems can connect to the client-managed data network. The client-managed network can also be connected to the IEDN TOR switches in the zBX node.

In addition, each BladeCenter chassis in the zBX has two Fibre Channel (FC) switch modules that connect to FC disk storage a SAN switch. Each FC switch supports up to six external FC links to connect to SAN switches.

Figure 2-13 shows front and rear views of a zBX rack (Rack B).

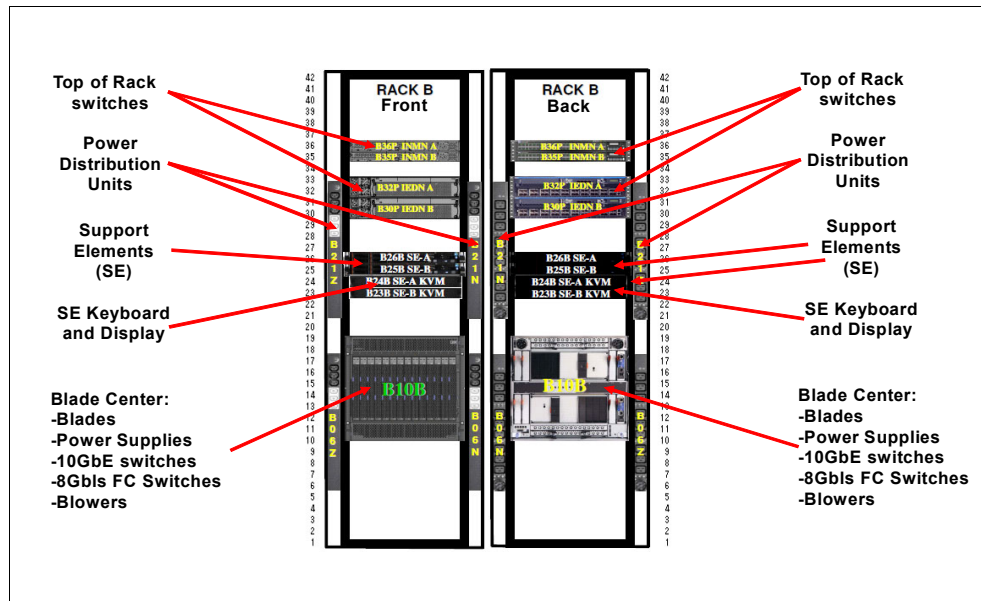


Figure 2-13 zBX rack B front and back configuration

The zBX racks include the following features:

- ▶ Two 1U internal Support Elements
- ▶ Two KVM (Keyboard-Video-Mouse) SE components
- ▶ Two TOR 1000BASE-T switches (rack B only) for the INMN
- ▶ Two TOR 10 GbE switches (rack B only) for the IEDN
- ▶ Up to two BladeCenter chassis in each rack

Each BladeCenter consists of the following features:

- Up to 14 blade slots
- Two Advanced Management Modules (AMM)
- Two Ethernet Switch Modules (ESM)
- Two 10 GbE high speed switch (HSS) modules
- Two 8 Gbps Fibre Channel switch modules
- Two blower modules
- ▶ Power Distribution Units (PDU)

The following blade types are supported in zBX:

- ▶ IBM POWER7 PS701 Express blades
- ▶ IBM System x blades (HX5 7873 dual-socket 16-core)
- ▶ IBM WebSphere DataPower XI50 for zEnterprise blades (double-width)

PowerVM Enterprise Edition is the hypervisor on the POWER7 blades, and the supported operating system is AIX. Linux on System x and Windows on System x are the supported operating systems for select System x blades, using the zBX integrated hypervisor for IBM System x blades (using a Kernel-based virtual machine). Both hypervisors are shipped, serviced, and deployed as z Systems LIC, booted automatically at power-on reset, and are isolated on the internal platform management network.

Client-supplied external disks are required with the zBX. Supported FC disk types and vendors with IBM blades are listed at the following web page:

http://www-03.ibm.com/systems/support/storage/config/ssic/displayesssearchwithoutjs.wss?start_over=yes

See the IBM *IBM z13 Technical Guide*, SG24-8251 for more information about the number of blades that are supported and configuration options.



Key functions and capabilities of IBM z13™

IBM z13 is the follow-on to the IBM zEnterprise zEC12 and the flagship of the IBM Systems portfolio. Like its predecessor, the z13 offers five hardware models, but has a more powerful processor, more processor units, and new functions and features.

The superscalar design allows the z13 to deliver a record level of capacity over the prior IBM z Systems servers. It is powered by 141 of the world's most powerful microprocessors that run at 5.0 GHz. The extreme scalability of z13 provides up to 40% more total capacity than its predecessor, the zEC12. The z13 is the industry's premier enterprise infrastructure choice for large-scale consolidation, secure data serving, transaction processing, and analytics capabilities.

For existing users of the zEnterprise BladeCenter Extension (zBX) an upgrade is available with z13 to a stand-alone *IBM z BladeCenter Extension (zBX) Model 004*.

In this chapter, we highlight some of the z13 functions and capabilities and point out solution areas where they can be of special value:

- ▶ Virtualization
- ▶ z13 technology improvements
- ▶ Capacity and performance
- ▶ zSystems common time functions
- ▶ Hardware Management Console (HMC) functions
- ▶ z13 power and cooling functions
- ▶ IBM z BladeCenter Extension (zBX) Model 004
- ▶ Reliability, availability, and serviceability (RAS)
- ▶ High availability

3.1 Virtualization

The z13 servers are highly virtualized, with the goal of maximizing the utilization of computing resources, lowering the total amount of resources that are needed for defined workloads and their cost. Virtualization is a key strength of the IBM z Systems family. Virtualization is embedded in the architecture and built into the hardware, firmware, and operating systems.

Virtualization requires a *hypervisor*, which is the control code that manages resources required for multiple independent operating system images. Hypervisors can be implemented in software or hardware, and the z13 has both. The hardware hypervisor for the z13 is known as *IBM Processor Resource/Systems Manager™ (PR/SM)*. PR/SM is implemented in firmware as part of the base system, fully virtualizes the system resources, and does not require any additional software to run. The software hypervisor is implemented by the z/VM operating system. z/VM uses some PR/SM functions.

Statement of Direction^a - KVM offering for IBM z Systems: In addition to the continued investment in z/VM, IBM intends to support a Kernel-based Virtual Machine (KVM) offering for z Systems that will host Linux on z Systems guest virtual machines. The KVM offering will be software that can be installed on z Systems processors like an operating system and can co-exist with z/VM virtualization environments, z/OS, Linux on z Systems, z/VSE and z/TPF. The KVM offering will be optimized for z Systems architecture and will provide standard Linux and KVM interfaces for operational control of the environment, as well as supporting OpenStack interfaces for virtualization management, enabling enterprises to easily integrate Linux servers into their existing infrastructure and cloud offerings.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

In the zBX, PowerVM Enterprise Edition is the hypervisor that offers a virtualization solution for any IBM Power Systems workload that runs on AIX. It allows use of the POWER7 processor-based PS 701 blades and other physical resources, providing better scalability and reduction in resource costs. IBM System x blades have a kernel-based virtual machine (KVM), integrated hypervisor with identical objectives.

Virtualization is key to establishing flexible infrastructures, with automated management and monitoring, such as those underpinning cloud offerings, including infrastructure as a service (IaaS), platform as a service (PaaS), and Software as a service (SaaS). We will return to this subject after describing the hardware and software virtualization capabilities of the z13.

3.1.1 z13 hardware virtualization

Processor Resource/Systems Manager (PR/SM) was first implemented in the mainframe in the late 1980s. It allows defining and managing subsets of the server resources that are known as logical partitions (LPAR). PR/SM virtualizes processors, memory, and I/O features. Certain features are purely virtualized implementations. For example, HiperSockets work like a LAN but do not use any I/O hardware.

PR/SM is always active on the system and has been enhanced to provide more performance and platform management benefits. PR/SM technology on zEC12 has received Common Criteria EAL5+ security certification.

Each LPAR can run any of the following supported operating systems:

- ▶ z/OS

- ▶ z/VM
- ▶ z/VSE
- ▶ z/TPF
- ▶ Linux on z Systems

The LPAR definition includes a number of logical PUs (LPUs), memory, and I/O devices. The *z/Architecture* (inherent in the z13 and its predecessors) is designed to meet those stringent requirements with low overhead and has achieved the highest security certification of the industry: Common Criteria EAL5+ with a Specific Target of Evaluation (Logical Partitions). This design has been proven in many client installations over several decades.

On z13, up to 85 LPARs can be defined, and hundreds or even thousands of virtual servers can be run under z/VM. Therefore, a high rate of context switching is to be expected, and accesses to the memory, caches, and virtual I/O devices must be kept isolated.

Logical processors

Logical processors are defined and managed by PR/SM and are perceived by the operating systems as real processors. These processors can be characterized as follows:

- ▶ Central processors (CP)
- ▶ IBM z Integrated Information Processors (zIIP)¹
- ▶ Integrated Facility for Linux (IFL)
- ▶ Internal Coupling Facility (ICF)

In addition, pre-characterized processors that are part of the system base configuration are always present: system assist processors (SAP) and integrated firmware processors (IFP). They provide support for all LPARs but are never part of an LPAR configuration.

PR/SM is responsible for accepting requests for work on logical processors by dispatching logical processors on physical processors. Physical processors can be shared across LPARs, but can also be dedicated to an LPAR. However, an LPAR must have its logical processors either all shared or all dedicated.

The sum of logical processors (LPU) defined in all of the LPARs activated in a central processor complex (CPC) might be well over the number of physical processors (PPU). The maximum number of LPUs that can be defined in a single LPAR cannot exceed the total number physical PUs that are available in the CPC. To achieve optimal ITR performance in sharing LPUs, it is suggested to keep the total number of online LPUs to a minimum. This action reduces both software and hardware overhead.

PR/SM ensures that, when switching a physical processor from one logical processor to another, the processor state is properly saved and restored, including all the registers. Data isolation, integrity, and coherence inside the system are strictly enforced at all times.

Logical processors can be dynamically added to and removed from LPARs. Operating system support is required to take advantage of this capability. Starting with z/OS V1R10, z/VM V5R4², and z/VSE V4R3³ the ability to dynamically define and change the number and type of reserved PUs in an LPAR profile can be used for that purpose. No pre-planning is required. The new resources are immediately available to the operating systems and, in the case of z/VM, to its guests. Linux on z Systems provides the Standby CPU activation/deactivation function, which is implemented in SLES 11 and RHEL 6.

¹ The zEC12 and zBC12 were the last z Systems servers to offer support for zAAP specialty engine processors. IBM supports running zAAP workloads on zIIP processors (“zAAP on zIIP”).

² z/VM V5R4 is not supported on z13

³ z/VSE V4R3 not supported on z13

z/VM mode partitions

The z/VM mode *logical partition (LPAR)*, first supported on IBM System z10®, is exclusively for running multiple workloads under z/VM. This LPAR mode provides increased flexibility and simplifies systems management by allowing z/VM to manage guests to perform the following tasks in the same z/VM LPAR:

- ▶ Operate Linux on z Systems on IFLs or CPs.
- ▶ Operate z/OS, z/VSE, and z/TPF on CPs.
- ▶ Operate z/OS while fully allowing *zIIP* usage by workloads (such as WebSphere and DB2) for an improved economics environment.
- ▶ Operate a complete Sysplex with ICF usage. This setup is especially valuable for testing and operations training; however, it is not recommended for production environments.

The z/VM-mode partitions require z/VM V5R4² or later and allows z/VM to use a wider variety of specialty processors in a single LPAR. The following processor types can be configured to a z/VM-mode partition:

- ▶ CPs
- ▶ IFLs
- ▶ zIIPs
- ▶ ICFs

If only Linux on z Systems is to be run under z/VM, then a z/VM mode LPAR is not required, and we suggest that a Linux-only LPAR be used instead.

Memory

To ensure security and data integrity, memory cannot be concurrently shared by active LPARs. In fact, a strict isolation is maintained.

Using the plan-ahead facility, memory can be physically installed without being enabled. It can then be enabled when it is necessary. z/OS and z/VM support dynamically increasing the memory size of the LPAR.

A logical partition can be defined with both an initial and a reserved amount of memory. At activation time, the initial amount is made available to the partition and the reserved amount can later be added, partially or totally. Those two memory zones do not have to be contiguous in real memory, but are displayed as logically contiguous to the operating system that runs in the LPAR.

z/OS is able to take advantage of this support by non-disruptively acquiring and releasing memory from the reserved area. z/VM V6R2 and later versions are able to acquire memory non-disruptively and quickly make it available to guests. z/VM virtualizes this support to its guests, which can also increase their memory non-disruptively. Releasing memory is still a disruptive operation.

LPAR memory is said to be virtualized in the sense that, in each LPAR, memory addresses are contiguous and start at address zero. LPAR memory addresses are different from the absolute memory addresses of the system, which are contiguous and have a single address of zero. Do not confuse this capability with the operating system virtualizing its LPARs memory, which is done through the creation and management of multiple address spaces.

The z/Architecture has a robust virtual storage architecture that allows, per LPAR, the definition of an unlimited number of address spaces and the simultaneous use by each program of up to 1023 of those address spaces. Each address space can be up to 16 EB (1 exabyte = 2⁶⁰ bytes). Thus, the architecture has no real limits. Practical limits are determined by the available hardware resources, including disk storage for paging.

Isolation of the address spaces is strictly enforced by the Dynamic Address Translation hardware mechanism. The validation of a program's right to read or write in each page frame is accomplished by comparing the page key with the key of the program that is requesting access. This mechanism has been in use since the System 370. Memory keys were part of, and used by, the original System 360 systems. Definition and management of the address spaces is under operating system control. Three addressing modes, 24-bit, 31-bit, and 64-bit, are simultaneously supported. This provides compatibility with earlier versions and investment protection.

z13 supports 2 GB pages, introduced with EC12 and zBC12, in addition to the 4 KB and 1 MB pages, and an extension to the z/Architecture: the Enhanced Dynamic Address Translation-2 (EDAT-2). With additional hardware, 1 MB pages can be pageable⁴.

Operating systems can allow sharing of address spaces, or parts of them, across multiple processes. For instance, under z/VM, a single copy of the read-only part of a kernel can be shared by all virtual machines that use that operating system, resulting in large savings of real memory and improvements in performance.

I/O virtualization

The z13 supports six Logical Channel Subsystems (LCSS) each with 256 channels, for a total of 1536 channels. In addition to the dedicated use of channels and I/O devices by an LPAR, I/O virtualization allows concurrent sharing of channels. This architecture also allows sharing the I/O devices that are accessed through these channels, by several active LPARs. This function is known as *multiple image facility (MIF)*. The shared channels can belong to different channel subsystems, in which case they are known as *spanned channels*.

Data streams for the sharing LPARs are carried on the same physical path with total isolation and integrity. For each active LPAR that has the channel configured online, PR/SM establishes one logical channel path. For availability reasons, multiple logical channel paths should exist for critical devices (for instance, disks that contain vital data sets).

When more isolation is required, configuration rules allow restricting the access of each logical partition to particular channel paths and specific I/O devices on those channel paths.

Many installations use the *parallel access volume (PAV)* function, which allows accessing a device by several addresses (normally one base address and an average of three aliases). This feature increases the throughput of the device by using more device addresses. *HyperPAV* takes the technology a step further by allowing the I/O Supervisor (IOS) in z/OS (and the equivalent function in the Control Program of z/VM) to create PAV structures dynamically. The structures are created depending on the current I/O demand in the system, lowering the need for manually tuning the system for PAV use.

In large installations, the total number of device addresses can be high. Thus, the concept of *channel sets* was introduced with the IBM System z9@. On the z13, up to four sets of approximately 64 K device addresses are available. This availability allows the base addresses to be defined on set 0 (IBM reserves 256 subchannels on set 0) and the aliases on set 1, set 2, and set 3. In total, 261,885 subchannel addresses are available per channel subsystem. Channel sets are used by the *Metro Mirror* (also referred as synchronous Peer-to-Peer Remote Copy (PPRC)) function by the ability to have the Metro Mirror primary devices defined in channel set 0. Secondary devices can be defined in channel set 1, 2, and 3, providing more connectivity through channel set 0.

To reduce the complexity of managing large I/O configurations further, starting with z/OS V1R10, z Systems introduced *extended address volumes (EAV)*. EAV is designed to provide

⁴ z/OS support only

large disk volumes. In addition to z/OS, both z/VM (starting with V5R4 with APARs) and Linux on z Systems support EAV.

By extending the disk volume size, potentially fewer volumes can be required to hold the same amount of data, making systems and data management less complex. EAV is supported by the IBM DS8000® series. Devices from other vendors should be checked for EAV compatibility.

The health checker function in z/OS V1R10 introduced a health check in the I/O Supervisor that can help system administrators identify single points of failure in the I/O configuration.

The dynamic I/O configuration function is supported by z/OS and z/VM. It provides the capability of concurrently changing the currently active I/O configuration. Changes can be made to channel paths, control units, and devices. The existence of a fixed HSA area in the z13 greatly eases the planning requirements and enhances the flexibility and availability of these reconfigurations.

3.1.2 IBM z Systems software virtualization

Software virtualization is provided by the IBM z/VM product. Strictly speaking, it is a function of the Control Program component of z/VM. Starting in 1972⁵, IBM has continuously provided commercial software virtualization in its mainframe servers.

z/VM uses the resources of the LPAR in which it is running to create functional equivalents of real z Systems CPCs, which are known as *virtual machines (VM)* or *guests*. A z/VM virtual machine is the functional equivalent of a real server. In addition, z/VM can emulate I/O peripheral devices (for instance, printers) by using spooling and other techniques, and LAN switches and disks by using memory.

z/VM allows fine-grained dynamic allocation of resources. As an example, in the case of processor sharing, the minimum allocation is approximately 1/10,000 of a processor. As another example, disks can be subdivided into independent areas, which are known as *minidisks*, each of which is exploited by its users as a real disk, only smaller. Minidisks are shareable, and can be used for all types of data and also for temporary space in a pool of on-demand storage.

Under z/VM, virtual processors, virtual memory, and all the virtual I/O devices of the VMs are dynamically definable (provisionable). z/VM supports the concurrent addition (but not the deletion) of memory to its LPAR and immediately makes it available to guests. Guests themselves can support the dynamic addition of memory. All other changes are concurrent. To make these concurrent definitions occur non disruptively requires support by the operating system that is running in the guest.

Although z/VM imposes no limits on the number of defined VMs, the number of active VMs is limited by the available resources. On a z13, thousands of VMs can be activated.

In addition to server consolidation and image reduction by vertical growth, z/VM provides a highly sophisticated environment for application integration and co-residence with data, especially for mission-critical applications.

Virtualization provides hardware-enabled resource sharing, and can also be used for the following functions:

- ▶ Isolate production, test, training, and development environments.
- ▶ Support previous applications.

⁵ Commercial name was VM/370

- ▶ Test new hardware configurations without actually buying the hardware.
- ▶ Enable parallel migration to new system or application levels, and provide easy back-out capabilities.

z/VM V6R2 introduced a new feature, *single system image (SSI)*. SSI enables improved availability, better management of planned outages, and capacity growth by creating clusters of z/VM systems with simplified management.

With SSI, is it possible to cluster up to four z/VM images in a single logical image. Highlights of SSI features include these:

- ▶ Live Guest Relocation (LGR) for Linux, the ability to move executing virtual servers without disruption from one z/VM system to another in the SSI.
- ▶ Management of resources with multi-system virtualization, to allow up to four z/VM instances to be clustered as a single system image.
- ▶ Horizontal scalability with up to four systems, even on mixed hardware generations.
- ▶ Availability, through non-disruptively moving work to available system resources and nondisruptively moving system resources to work.
- ▶ An SSI cluster can contain both 6.2 and 6.3 members, and a member can be upgraded from 6.2 to 6.3 using a new *upgrade in place* installation feature.

For more information about SSI, see the following resources:

- ▶ *An introduction to z/VM Single System Image (SSI) and Live Guest Relocation (LGR)*, SG24-8006
- ▶ *Using z/VM v 6.2 Single System Image (SSI) and Live Guest Relocation (LGR)*, SG24-8039

In light of the IBM cloud strategy and adoption of OpenStack, the management of z/VM environments in zManager is now stabilized and will not be further enhanced. Accordingly, zManager will not provide systems management support for z/VM V6R3 and later releases. However, zManager will continue to play a distinct and strategic role in the management of virtualized environments created by integrated firmware hypervisors (PR/SM, PowerVM, and System x hypervisor based on KVM) of z Systems.

The IBM z Unified Resource Manager (zManager)⁶ uses the management *application programming interface (API)* of z/VM to provide a set of resource management functions for the z/VM V6R2 environment.

Providing a more detailed description of z/VM or other highlights of its capabilities is beyond the scope of this book. For a deeper discussion of z/VM, see *Introduction to the New Mainframe: z/VM Basics*, SG24-7316-00.

3.1.3 zBX virtualized environments

On the zBX model 004, available as an upgrade of zBX model 002 or 003, the IBM POWER7 processor-based PS701 blades run PowerVM Enterprise Edition to create a virtualized environment that is similar to the one found in IBM Power Systems servers. The POWER7 processor-based LPARs run the AIX operating system.

PowerVM is EAL4+ certified and is isolated on the intranode management network, providing intrusion prevention, integrity, and secure virtual switches with integrated consolidation.

⁶ zManager does not support systems management of z/VM on z13

The IBM System x blades are also virtualized. The integrated System x hypervisor uses kernel-based virtual machines (KVMs). Support is provided for Linux and Microsoft Windows.

PowerVM, as well as the integrated hypervisor for the System x blades, is managed by the IBM z Unified Resource Manager, so it is shipped, deployed, monitored, and serviced at a single point.

Management of the zBX environment is done as a single logical virtualized environment by the Unified Resource Manager.

3.1.4 z Systems -based clouds

Cloud computing is a paradigm for providing IT services. It capitalizes on the ability to rapidly and securely deliver standardized offerings, while retaining the capacity for customizing the environment. Elasticity, allowing to accompany the ebbs and flows of demand, and exploiting just-in-time provisioning is another requirement. We make no distinction here between private and public clouds, as they are both well addressed by z Systems.

Virtualization is critical to the economic and financial viability of those offerings, because it allows to minimize over provisioning of resources. and reutilizing them at the end of the virtual server life-cycle.

Because of the extreme integration in the hardware, virtualization on z13 is highly efficient (the best in the industry) and encompasses computing as well as I/O resources, including the definition of *internal virtual networks with switches*. These are all characteristics of *Software Defined Environments*, and allow supporting on a single real server, dense sets of virtual servers and server networks, with up to 100% sustained resource utilization and the highest levels of isolation and security. Therefore, the cloud solution costs, whether hardware, software, or management, are minimized.

Cloud elasticity requirements are covered by the z13 granularity offerings, including capacity levels and Capacity on Demand. These and other technologic leadership characteristics that make the z Systems CPCs the server golden standard, are discussed in the remainder of this chapter.

Z/VM V6R3 Integrated xCAT support: If you want to get started with cloud computing, the Extreme Cloud Administration Toolkit (xCAT), a scalable open source tool developed by IBM, can be used to provision, manage, and monitor physical and virtual machines. Because xCAT is integrated into z/VM V6R3, it no longer needs to be separately downloaded, installed, and configured. You can quickly deploy xCAT with a small amount of tailoring. While xCAT provides rudimentary cloud management, z/VM V6R3 also provides OpenStack enablement for more sophisticated and complete solutions, like IBM Cloud Manager.

3.1.5 GDPS Virtual Appliance

Statement of Direction^a: In the first half of 2015, IBM intends to deliver a GDPS/Peer to Peer Remote Copy (GDPS/PPRC) multiplatform resiliency capability for customers who do not run the z/OS operating system in their environment. This solution is intended to provide IBM z Systems customers who run z/VM and their associated guests, for instance, Linux on z Systems, with similar high availability and disaster recovery benefits to those who run on z/OS. This solution will be applicable for any IBM z Systems announced after and including the zBC12 and zEC12.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

To reduce IT costs and complexity, many enterprises are consolidating independent servers into Linux images (guests) running on z Systems platform. Linux on z Systems can be implemented either as guests running under z/VM or native Linux LPARs on z Systems. Workloads with an application server running on Linux on z Systems and a database server running on z/OS are common. Two examples are as follows:

- ▶ WebSphere Application Server running on Linux and CICS, DB2 running under z/OS
- ▶ SAP application servers running on Linux and database servers running on z/OS

With a multitiered architecture, there is a need to provide a coordinated near-continuous availability and disaster recovery solution for both z/OS and Linux on z Systems.

GDPS Virtual Appliance is a fully integrated continuous availability and disaster recovery solution for Linux on z Systems customers which consists of:

- ▶ An operating system image
- ▶ The application components
- ▶ An appliance management layer which makes the image self-containing
- ▶ API and UI for customization, administration, and operation tailored for the appliance function.

It is designed to improve both consumability and time-to-value for customers.

For more information, see *IBM z13 Technical Guide*, SG24-8251.

3.2 z13 technology improvements

z13 provides technology improvements in these areas:

- ▶ Microprocessor
- ▶ Memory
- ▶ Capacity and performance
- ▶ Flash Express feature
- ▶ 10GbE RoCE Express feature
- ▶ zEDC Express feature
- ▶ Cryptography
- ▶ I/O capabilities

These features are intended to provide a more scalable, flexible, manageable, and secure consolidation and integration to the platform, which contributes to a lower total cost of ownership.

3.2.1 Microprocessor

The z13 has a newly developed microprocessor chip and a newly developed storage control chip. The chips use CMOS 14S (22nm) technology and represent a major step forward in technology use for the IBM z Systems products, resulting in increased packaging density.

The microprocessor chip and the storage control chip for the z13 are packaged each one in a single-chip module (SCM). The SCM can contain one microprocessor (PU) chip or one storage control (SC) chip. The SCMs are installed inside a CPC drawer, and the z13 can contain from one to four CPC drawers. Each CPC drawer has two nodes and each node has three microprocessor (PU) chips and one storage control chip. The CPC drawer also contains the memory arrays, I/O connectivity infrastructure, and various other mechanical and power controls.

The CPC drawer is connected to the PCI Express (PCIe) I/O drawers and I/O drawers, through one or more cables.

Standard PCIe and InfiniBand protocols are used for fast transfer of large volumes of data between the memory in the CPC drawer and the I/O cards housed in the PCIe I/O drawers and I/O drawers.

z13 processor chip

The z13 processor chips provide more functions per chip (eight cores on a single chip) because of technology improvements that allow designing and manufacturing more transistors per unit of area. Each processor chip can have six, seven or eight active cores. This configuration translates into using fewer chips to implement the needed functions, which helps enhance system availability.

The z Systems microprocessor development followed the same basic design set since the 9672-G4 (announced in 1997) until the z9. That basic design was stretched to its maximum, so a fundamental change was necessary. The z10 chip introduced a high-frequency design, which was improved with the z196 and enhanced again with the hex-core zEC12 and with eight core microprocessor chip for z13.

To allow an increased number of processors sharing larger caches with faster access time and improved capacity and performance, the z13 has adjusted the cycle time to 0.2 nanoseconds (5.0 GHz).

Each core of the processor chip (Figure 3-1) includes one co-processor for hardware acceleration of data compression and cryptography, I/O bus and memory controllers, and an interface to a separate storage controller/cache chip.

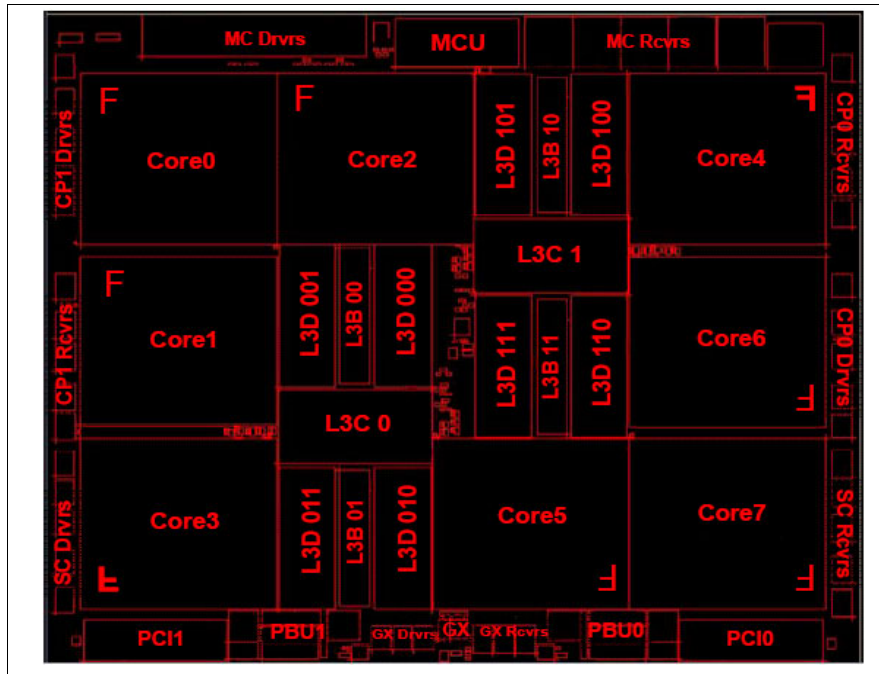


Figure 3-1 z13 eight-core microprocessor chip

On-chip cryptographic hardware includes the full complement of the *Advanced Encryption Standard (AES)* algorithm, *Secure Hash Algorithm (SHA)*, and the *Data Encryption Standard (DES)* algorithm, and also the protected key implementation.

z13 processor design highlights

The z/Architecture offers a rich *complex instruction set computer (CISC) Instruction Set Architecture (ISA)* supporting multiple arithmetic formats.

The z196 introduced 110 new instructions and offered a total of 984, out of which 762 were implemented entirely in hardware. The zEC12 also introduced new instructions, notably for the Transactional Execution and the EDAT-2 facilities. The z13 introduces 139 new instructions for analytics vector processing.

Compared to zEC12, the z13 processor design improvements and architectural extensions include the following features:

- ▶ Balanced performance growth:
 - 40% more system capacity:
 - 33% more cores in a central processor chip (6 -> 8).
 - Maximum number of cores increased from 120, on the zEC12, to 168 on the z13.
 - Maximum number of configurable cores increased from 101, on the zEC12, to 141 on the z13.
 - Fourth Generation High Frequency processor:
 - Although the frequency has been lowered from 5.5 GHz, on the zEC12, to 5.0 GHz on z13, its uniprocessor performance is up to 10% faster as compared to zEC12.
- ▶ Innovative core-cache design (L1 and L2), processors chip-cache design (L3) and node-cache design (L4) optimized by HiperDispatch, with focus on keeping more data

closer the processor, increasing the cache sizes and decreasing the latency to access the next levels of cache: :

- Total L1 per core is 40% larger.
 - Total L2 per core is 100% larger.
 - Total on-chip shared L3 is 33% larger.
 - Total shared L4 is 266% larger, with non-data inclusive coherent (NIC) directory.
 - Unique private L2 cache (2 MB for instructions and 2MB for data) design reduces L1 miss latency.
- ▶ Re-optimized pipeline depth for power and performance:
 - Increased instructions pipeline width per core
 - Number of instructions in flight is increased from seven to ten.
 - Greater integer execution bandwidth, with four fixed-point arithmetic execution units.
 - Improved fixed point and floating point divide.
 - Greater floating point execution bandwidth, with two binary and two decimal floating-point arithmetic execution units.
 - ▶ Improved Instruction Fetching Unit, with new branch prediction and instruction fetch front end to support multithreading and to improve branch prediction throughput
 - ▶ Wider instruction decode, dispatch and completion bandwidth increased to six instructions per cycle compared to three on zEC12.
 - ▶ Dedicated co-processor for each core with improved performance and more capability:
 - The Central Processor Assist for Cryptographic Function (CPACF) has been optimized to provide up to 2x faster encryption functions.
 - TDES with twice more throughput of zEC12 CPACF for large blocks
 - AES with twice more throughput of zEC12 CPACF for large blocks
 - SHA four times more throughput of zEC12 CPACF for large blocks
 - Hashing functions in CPACF are up to 3.5x faster. more performance as compared to zEC12
 - ▶ Multiple innovative architectural extensions for software usage:
 - Single-instruction, multiple-data (SIMD). Set of instructions that allows optimization of code to complex mathematical models and business analytics vector processing.
 - Simultaneous multithreading (SMT). Allows up to two active threads per core sharing the IFL or ZIIP cores resources.
 - ▶ Increased instruction issue, execution, and completion throughput:
 - Improved instruction dispatch and grouping efficiency
 - Millicode handling
 - Next Instruction Access Intent
 - Load and Trap instructions
 - Branch Prediction Preload
 - Data prefetch

Hardware decimal floating point function

Hardware decimal floating point (HDFP) support was introduced with the z9 EC and enhanced with a new decimal floating point accelerator feature in IBM zEnterprise 196. zEC12 and zBC12 facilitate better performance on traditional zoned-decimal operations with a broader usage of Decimal Floating Point facility by COBOL and PL/I programs. The z13 includes a decimal floating point *packed conversion facility* with z/OS V2R1 and z/VM V6R3 support.

This facility is designed to speed up such calculations and provide the necessary precision demanded mainly by the financial institutions sector. The decimal floating point hardware fully implements the new IEEE 754r standard.

Industry support for decimal floating point is growing, with IBM leading the open standard definition. Examples of support for the draft standard IEEE 754r include: Java BigDecimal, C#, XML, C/C++, GCC, COBOL, and other key software vendors such as Microsoft and SAP.

Support and usage of HDFP varies with operating system and release. For a detailed description, see *IBM z13 Technical Guide*, SG24-8251. See also “z/OS XL C/C++ considerations” on page 133.

Simultaneous multithreading (SMT)

The simultaneous multithreading (SMT) allows more than one thread execute simultaneous in the same core, sharing all its resources. This functionality is available in the z13 IFL and zIIP processor cores and allows up to two threads executing in the same processor, thereby providing better utilization of the cores and an increased capacity of processing.

When a program accesses a memory location that is not in the cache, it is called a *cache miss*. Because the processor then must wait for the data to be fetched from the next cache level, or from main memory, before it can continue to execute, cache misses directly influence the performance and capacity of the core to execute instructions. With simultaneous multithreading exploitation, when one thread in the core is waiting, for example, for data to be fetched from the next cache levels or from main memory, the second thread in core can utilize the shared resources rather than remain idle.

Exploitation support for SMT functionality is provided in z/OS V2R1 for zIIPs and z/VM V6R3 for IFLs.

Single Instruction Multiple Data (SIMD)

z13 architecture has designed with a set of instructions to improve the performance of complex mathematical models and analytics workloads, through vector processing and new complex instructions able to process a large number of data with only a single instruction.

This new set of instructions, known as Single Instruction Multiple Data (SIMD), which is designed for parallel computing can accelerate code with integer, string, character, and floating point data types. enabling better consolidation of analytics workloads and business transactions on the z Systems platform. SIMD provides the next phase of enhancements on z Systems analytics capability.

Transactional Execution (TX) facility

This capability, which is known in the industry as *hardware transactional memory*, allows issuing a group of instructions atomically, that is, either *all the results* of the instructions in the group are committed *or none* are, in true transactional way. The execution is optimistic: The instructions are issued, but previous state values are saved in a “transactional memory.” If the transaction succeeds, the saved values are discarded, otherwise they are used to restore the original values. Software can test the execution’s success and re-drive the code, if needed, using the same or a different path.

The TX facility provides several instructions, including declaring the beginning and end of a transaction, and to cancel the transaction. TX is expected to provide significant performance benefits and scalability to workloads by being able to avoid most of the locks. This ability is especially important for heavily threaded applications, such as Java.

Runtime Instrumentation Facility

This facility provides managed runtimes and just-in-time compilers with enhanced feedback about application behavior, allowing dynamic optimization on code generation as it is being executed.

3.2.2 Memory

The z13 can have up to 10 TB of usable memory installed. This is significantly more than its predecessor zEC12 which had 3 TB maximum.

In addition, the z13 has increased the size of the hardware system area (HSA) to 96 GB, when compared with its predecessor, zEC12 which has 32 GB of HSA. The HSA is not included in the memory which the client orders, and has a fixed size of 96 GB.

z/Architecture addressing modes: The z/Architecture simultaneously supports 24-bit, 31-bit, and 64-bit addressing modes. These modes provide compatibility with earlier versions and investment protection.

Support of large memory varies with the operating system, as follows:

- ▶ z/OS V1R12 and later support up to 4 TB.
- ▶ z/VM V6R3 supports up to 1 TB.
- ▶ z/VM V6R2 support up to 256 GB.
- ▶ z/VSE V5R1 and later support up to 32 GB.
- ▶ z/TPF V1R1 supports up to 4 TB.
- ▶ SLES 11 supports 4 TB and RHEL 6 supports 3 TB.

The maximum memory size per logical partition has changed with z13. Up to 10TB can now be defined to a logical partition in the image profile. Each operating system will be able to allocate central storage according to their individual maximum memory amount supported, as shown above.

Dynamic memory reallocation

On z13 the memory allocation algorithm has changed. PR/SM will try to allocate memory in to a single processor drawer, striped between the two nodes. Basically, the PR/SM memory and processor resources allocation goal is to place all partition resources on a single processor drawer, if possible. The resources, memory, and processors, are assigned to the partitions when they are activated. Later, when all partitions have been activated, PR/SM can move memory between processor drawers to benefit performance, without operating system knowledge.

Plan-ahead memory

When a client can anticipate the requirements for future increases of the installed memory, the initial system order can contain both a starting and additional memory sizes. The additional memory is referred to as *plan-ahead memory*. A specific memory pricing model is available in support of this capability.

The starting memory size is activated at system installation time and the rest remains inactive. When more physical memory is required, it is fulfilled by activating the appropriate number of plan-ahead memory features. This activation is concurrent and can be nondisruptive to the applications depending on the operating system support. z/OS and z/VM support this function.

Do not confuse plan-ahead memory with flexible memory support. Plan-ahead memory is for a permanent increase of installed memory, whereas flexible memory provides a temporary replacement of a part of memory that becomes unavailable.

Flexible memory

Flexible memory was first introduced on the z9 EC as part of the design changes and offerings to support enhanced book availability (EBA). Flexible memory was used to temporarily replace the memory that becomes unavailable when performing maintenance on a book.

On z13, the additional resources that are required for the flexible memory configurations are provided through the purchase of planned memory features, along with the purchase of memory entitlement. Flexible memory configurations are available only on multi-CPC drawers (models N63, N96, NC9, and NE1) and range from 256 GB to 2.5 TB, depending on the model.

Contact your IBM representative to help determine the appropriate configuration for your business.

Large page support

The size of pages and page frames has remained at 4 KB for a long time. Starting with the IBM System z10, z Systems platforms are capable of having large pages with the size of 1 MB, in addition to supporting pages of 4 KB. This capability is a performance item which addresses particular workloads and relates to large main storage usage. Both page frame sizes can be simultaneously used.

Large pages enable the *translation lookaside buffer (TLB)* to better represent the working set and suffer fewer misses by allowing a single TLB entry to cover more address translations. Users of large pages are better represented in the TLB and are expected to perform better.

This support benefits long-running applications that are memory access intensive. Large pages are not recommended for general use. Short-lived processes with small working sets are normally not good candidates for large pages and see little to no improvement. The use of large pages must be decided based on knowledge that is obtained from measurement of memory usage and page translation overhead for a specific workload.

The large page support function is not enabled without the required software support. Without the large page support, page frames are allocated at the current 4 KB size. At the time they were introduced, large pages were treated as fixed pages and were never paged out. Under z/OS, they are available only for 64-bit virtual private storage, such as virtual memory located above 2 GB. With the availability of the hardware feature Flash Express (see 3.2.4, “Flash Express” on page 74), large pages can become pageable.

Support for 2 GB large page

z13 uses 2 GB page frames, introduced with zEC12, as an architectural extension. This is to increase efficiency for DB2 buffer pools, Java heap, and other large structures. Usage of 2 GB pages increases TLB coverage without proportionally enlarging the TLB size:

- ▶ A 2 GB memory page has the following characteristics:
 - It is 2048 times larger than a large page of 1 MB size.
 - It is 524,288 times larger than an ordinary base page with a size of 4 KB.
- ▶ A 2 GB page allows for a single TLB entry to fulfill many more address translations than either a large page or ordinary base page.
- ▶ A 2 GB page provides users with much better TLB coverage, and therefore provides better performance in the following ways:
 - By decreasing the number of TLB misses that an application incurs
 - By spending less time on converting virtual addresses into physical addresses
 - By using less real storage to maintain DAT structures

2 GB Large Page exploitation: Exploitation of 2 GB pages is provided for the IBM 31-bit SDK for z/OS, Java Technology Edition, V7.0.0 (5655-W43) and SDK7 IBM 64-bit SDK for z/OS, Java Technology Edition, V7.0.0 (5655-W44).

3.2.3 Native PCIe features and integrated firmware processor

zEC12 introduced feature card types, known as native PCIe features, which require a different management design. The following native PCIe features are available:

- ▶ 10GbE RoCE Express
- ▶ zEDC Express

These features are plugged exclusively into a PCIe I/O drawer, where they coexist with the other, non-native PCIe, I/O adapters and features, but they are managed in a different way from those other I/O adapters and features. The native PCIe feature cards have a PCHID assigned according to the physical location in the PCIe I/O drawer.

For non-native PCIe features, which are plugged into a PCIe I/O drawer, and on the z13 supported carry-forward I/O drawer, all adaptation layer functions are integrated into the adapter hardware.

For the native PCIe features introduced by zEC12 and supported by z13, there are drivers included in the operating system, and the adaptation layer is not needed. The adapter management functions (such as diagnostics and firmware updates) are provided by Resource Groups partitions running on the integrated firmware processor (IFP).

The integrated firmware processor (IFP) is used to manage native PCIe adapters installed in a PCIe I/O drawer. The IFP is allocated from a pool of PUs available for the whole system. Because the IFP is exclusively used to manage native PCIe adapters, it is not taken from the pool of PUs that can be characterized for customer usage.

If a native PCIe feature is present in the system, the IFP is initialized and allocated during the system POR phase. Although the IFP is allocated to one of the physical PUs, it is not visible to the customer. In case of error or failover scenarios, the IFP will act like any other PU (that is, sparing is invoked).

3.2.4 Flash Express

The *Flash Express* feature helps to improve system and application availability and performance to compete more effectively in today's service focused market. Flash Express capabilities enable the following features:

- ▶ Improved z/OS recovery and diagnostic times
- ▶ Handling of workload shifts and coping with dynamic environments more smoothly
- ▶ Use of pageable large pages (1 MB) yielding CPU performance benefits
- ▶ Offloading GBps of random I/O from the I/O Fabric
- ▶ Predictive paging
- ▶ Overflow areas for certain Coupling Facility list structures for WebSphere MQ

Flash Express is easy to configure, requires no special skills, and provides rapid time to value. This feature is designed to allow each logical partition to be configured with its own storage-class memory (SCM) address space, to be used for paging and act as an overflow area for CF structures. Pages that are 1 MB become pageable (z/OS only). Support is provided to configure SCM increments offline through a z/OS operator command, and allows the PLPA and COMMON paging data sets to be optional.

Flash Express data privacy

For Flash Express, the data privacy relies on a symmetric key that encrypts the data that is temporarily stored on the SSD. By using a smart card and an integrated smart card reader on the Support Element (SE), the encryption key is generated within the secure environment of the smart card. The key is tightly coupled to the SE serial number, which ensures that no other SE is able to share the key or the smart card that is associated with a specific SE. The generated key is replicated in a secure way to the alternate Support Element smart card. The key is transferred from the SE to the Flash Express adapter under the protection of a private and public key pair that is generated by the firmware that manages the Flash Express adapter.

CFCC exploitation of Flash Express

With CFCC Level 19 and later, the Flash Express feature can be exploited to help handle the overflow of WebSphere MQ shared queue structures. This is designed to allow structure data to be migrated to Flash Express memory as needed and migrated back to real memory to be processed. This requires WebSphere MQ for z/OS V7, z/OS V2R1 or V1R13, with additional service.

Software and operating system support

Exploitation of pageable 1 MB pages for z/OS includes these items:

- ▶ z/OS V1R13 Language Environment® when used with a run-time option.
- ▶ Java, with the IBM 31-bit SDK for z/OS, Java technology Edition, V7.1.0 and IBM 64-bit SDK for z/OS, Java Technology Edition, V7.1.0.
- ▶ DB2 10⁷ and DB2 11 exploits pageable 1 MB frames for buffer pool and executable code.
- ▶ Pageable large pages are used by IMS⁸ Common queue server (CQS) interface buffers and selected database storage pools on a zEC12 or z13.

Flash Express is an optional feature. It is supported by z/OS V1R13 and later, with the z/OS V1R13 RSM Enablement Offering web deliverable installed. It is fully supported by z/OS V2R1.

3.2.5 zEDC Express

zEDC Express is an optional feature, available on z13, zEC12, and zBC12 systems, and is designed to help to improve cross-platform data exchange, reduce CPU consumption, and save disk space by providing hardware-based acceleration for data compression and decompression for the enterprise. It provides data compression with lower CPU consumption than compression technology previously available on z Systems.

This capability is of special interest, for instance, to clients experiencing significant year-to-year growth in storage. Savings can be realized initially by making more efficient use of existing capacity, allowing more data to be kept active and online at lower cost, and longer term by elongating time frames for acquisitions for additional storage.

z/OS V2R1 zEnterprise Data Compression

Exploitation support of zEDC Express functionality is provided *exclusively* by z/OS V2R1 zEnterprise Data Compression for both data compression and decompression.

Support for data recovery (decompression) in the case that zEDC Express is not installed, or installed but not available, on the system, is provided via software on z/OS V2R1, and on

⁷ Service is required.

⁸ IMS 12 needs PTF for CQS interface buffers and IMS 13 includes this support also.

V1R13 and V1R12 with appropriate PTFs. Software decompression is slow and uses considerable processor resources, thus it is not recommended for production environments.

z/OS guests running under z/VM V6R3 may exploit the zEDC Express feature. IBM zEnterprise Data Compression (zEDC) for z/OS V2R1 and the zEDC Express feature are designed to support a new data compression function to help provide high-performance, low-latency compression without significant CPU overhead. This may help to reduce disk usage, provide optimized cross-platform exchange of data, and provide higher write rates for System Management Facility (SMF) data.

z/OS V2R1 can use zEDC to compress SMF records. zEDC with z/OS SMF Logger alleviates SMF constraints across the entire life cycle of a record using compression technology while storing data in System Logger and reducing Logger CPU usage.

Clients who have large sequential data that uses BSAM/QSAM⁹ extended format, can use zEDC to help reduce disk space usage and improve effective bandwidth without significant CPU overhead. zEDC will also be used by DFSMSdss and DFSMSHsm to deliver efficient compression when backing up and restoring data.

IBM 31-bit and 64-bit SDK7 for z/OS Java Technology Edition, Version 7 provides transparent exploitation of the zEDC Express feature.

Similarly, the IBM Encryption Facility for z/OS (IBM EF) uses zEDC Express when used with z/OS Java Version 7. For IBM EF users not already using compression, compression with zEDC can provide IBM EF users a reduction of elapsed time and CPU times. This complements the software compression that exists today with Encryption Facility OpenPGP support.

IBM Sterling Connect: Direct for z/OS V5R2 can automatically leverage zEDC Express Accelerator for file compression and decompression as files are transferred cross-platform. Usage may provide reduction in elapsed time to transfer a file from z/OS to z/OS with minimal CPU increase.

Table 3-1 compares compression technologies.

Table 3-1 Compression technologies for z Systems

Type	Optimized for	Performance overhead	Supported data	Frequency of access post compression
CMPSC compression on z Systems processor chip	DB2 and select DFSMS files	On chip, relatively little CPU overhead and less I/O, fast	DB2 - optimized for row-wise access to data is required DFSMS files - for VSAM and non-VSAM extended format data sets	Often
Other software compression (zlib, or similar)	Most compression uses industry std today. Used by many file types	Higher CPU - software instructions executed. Note: if Java then eligible for zIIP (or zAAP)	Any. De facto standard for almost any type of data.	Often

⁹ BSAM/QSAM: basic sequential access method and queued sequential access method

Type	Optimized for	Performance overhead	Supported data	Frequency of access post compression
Tape HW compression	Tape compression –optimized for use with large files, archival purposes	Performed by tape subsystem	Any	Often/Rare (application dependent)
Archival / Backup	Archive data and data backup/copy	CPU overhead, longer wall clock time	DFSMSHsm, DFSMSdss	Often/Rare (application dependent)
Real time compression	IBM NAS storage	No performance degradation	SVC	Designed for active primary data.
zEDC Express	Active, for cross-platform data exchange. Enables compression of active and inactive data	Processing on zEDC Express: expect minimal CPU overhead, low I/O latency	SMF though logger zlib compatible Java BSAM/QSAM Extended format SOD DFSMSHsm / DFSMSdss Encryption facility	Frequent access required. Useful for files that previously used software compression as well

IBM z Batch Network Analyzer

The IBM z Batch Network Analyzer (zBNA) is a no-cost, “as-is” tool. It is available to clients, IBM Business Partners, and IBM employees.

zBNA replaces the BWATOOL. It is Windows based, provides graphical and text reports, including Gantt charts, and support for alternate processors.

zBNA can be used to analyze customer provided SMF records, in order to identify jobs and data sets which are candidates for zEDC compression, across a specified time window, typically a batch window. zBNA is able to generate lists of data sets by job:

- ▶ Those that already do hardware compression and might be candidates for zEDC
- ▶ Those that might be zEDC candidates but are not in extended format

Thus, zBNA can help estimate utilization of zEDC features and help size number of features needed.

IBM Clients can obtain zBNA and other CPS tools at this site:

<http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5132>

IBM Business Partners can obtain zBNA and other CPS tools at this site:

https://www.ibm.com/partnerworld/wps/servlet/mem/ContentHandler/tech_PRS5133

IBM Employees can obtain zBNA and other CPS tools at this site:

<http://w3.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5126>

3.2.6 10GbE RoCE Express

The 10 Gigabit Ethernet (10GbE) RoCE Express feature is designed to help reduce consumption of CPU resources for applications utilizing the TCP/IP stack (such as WebSphere accessing a DB2 database).

Use of the 10GbE RoCE Express feature may also help to reduce network latency with memory-to-memory transfers utilizing Shared Memory Communications - Remote Direct Memory Access (SMC-R) in z/OS V2R1. It is transparent to applications and can be used for LPAR-to-LPAR communication on a single CPC or server-to-server communication in a multiple CPC environment.

z/OS V2.1 with PTF supports the new sharing capability available for the Remote Direct Memory Access (RDMA) over Converged Ethernet (RoCE Express) features on z13 processors. This enhancement allows full use of the ports in the RoCE adapter and to share the adapters across up to 31 z/OS images on a z13 processor.

z/VM V6R3 supports also guest exploitation of RoCE

In addition, Communication Server has been enhanced to support selecting between TCP/IP and RoCE transport layer protocols automatically based on traffic characteristics. This feature is supported on z13, zEC12, and zBC12¹⁰ and is installed in the PCIe I/O drawer. A maximum of 16 features can be installed. In z13 we have both ports enabled to be used and each feature, two ports, can be shared up to 31 logical partitions. On zEC12 and zBC12 only one port can be used and the port had to be dedicated to a logical partition.

3.2.7 I/O capabilities

The z13 has many I/O capabilities for supporting high-speed connectivity to resources inside and outside the system. The connectivity of the z13 is designed to maximize application performance and satisfy clustering, security, storage area network (SAN), and local area network (LAN) requirements.

Multiple subchannel sets

Multiple subchannel sets (MSS) provide greater I/O device configuration capabilities for large enterprises. Up to four subchannel sets for z13 can be defined to each channel subsystem (CSS). Up to six channel subsystems can be defined on the z13.

For each additional subchannel set, the amount of addressable storage capacity is 64 K minus one subchannel, which enables a larger number of storage devices. This increase complements other functions (such as large or extended address volumes) and HyperPAV. This can also help facilitate consistent device address definitions, simplifying addressing schemes for congruous devices.

The first subchannel set (SS0) allows the definition of any type of device (such as bases, aliases, secondaries, and those devices other than disks that do not implement the concept of associated aliases or secondaries). The second, third, and fourth subchannel sets (SS1, SS2, and SS3) can be designated for use for disk alias devices (of both primary and secondary devices) and Metro Mirror secondary devices only.

Initial program load from an alternate subchannel set

The z13 support the *initial program load (IPL)* from subchannel set 1 (SS1), subchannel set 2 (SS2), or subchannel set 3 (SS3). Devices that are used early during IPL processing can be accessed by using subchannel set 1, subchannel set 2, or subchannel set 3. This flexibility allows the users of Metro Mirror (PPRC) secondary devices that are defined using the same device number and a new device type in an alternate subchannel set to be used for IPL, input/output definition file (IODF), and stand-alone dump volumes, when needed.

¹⁰ Feature sharing and dual port are not supported on zEC12 and zBC12.

Channel subsystem enhancement for I/O resilience

The z13 channel subsystem incorporates an improved load balancing algorithm that is designed to provide improved throughput and reduced I/O service times, even when abnormal conditions occur. For example, degraded throughput and response times can be caused by multi-system workload spikes. This reduction can also be caused by resource contention in storage area networks (SAN) or across control unit ports, SAN congestion, suboptimal SAN configurations, problems with initializing optics, dynamic fabric routing changes, and destination port congestion.

When such events occur, the channel subsystem is designed to dynamically select channels to optimize performance. The subsystem also minimizes imbalances in I/O performance characteristics (such as response time and throughput) across the set of channel paths to each control unit. This function is done by using the in-band I/O instrumentation and metrics of the z Systems FICON and zHPF protocols.

This channel subsystem enhancement is available to z13, zEC12, and zBC12 and is supported on all FICON channels when configured as CHPID type FC. In support of this new function, z/OS V1R12 and V1R13 with a *program temporary fix (PTF)* also provide an updated health check based on an I/O rate-based metric, rather than on initial control unit command response time.

This enhancement is transparent to operating systems. However, this feature requires an updated health check based on an I/O rate-based metric, rather than on initial control unit command response time, provided by z/OS V1R12 and V1R13 with a PTF and later.

FICON connectivity

The Fibre Connection (FICON) features in the z13 can provide connectivity to servers, FC switches, and various devices (control units, disk, tape, printers) in a SAN environment. FICON improves upon the Fibre Channel Protocol (FCP) and continues to evolve, delivering improved throughput, reliability, availability, and serviceability.

High Performance FICON for z Systems

High Performance FICON for z Systems (zHPF), first provided on System z10, is a FICON architecture for protocol simplification and efficiency, reducing the number of information units (IU) processed. Enhancements have been made to the z/Architecture and the FICON interface architecture to provide optimizations for online transaction processing (OLTP) workloads.

When used by the FICON channel, the z/OS operating system, and the control unit (appropriate levels of Licensed Internal Code are required), the FICON channel overhead can be reduced and performance can be improved. Additionally, the changes to the architecture provide end-to-end system enhancements to improve reliability, availability, and serviceability (RAS). The zHPF channel programs can be used, for instance, by z/OS OLTP I/O workloads, DB2, VSAM, PDSE, and zFS. zHPF requires matching support by the DS8000 series or similar devices from other vendors.

The zHPF is exclusive to z Systems. The FICON Express16S, FICON Express8S, and FICON Express8 (*channel path identifier (CHPID)* type FC) concurrently support both the existing FICON protocol and the zHPF protocol in the server Licensed Internal Code.

High Performance FICON for z Systems (zHPF) has been enhanced to allow all large write operations (>64KB) at distances up to 100km to be executed in a single round trip to the control unit thereby not elongating the I/O service for these write operations at extended distances. This is especially advantageous for GDPS HyperSwap® configurations.

For more information about FICON channel performance, see the technical papers on the z Systems I/O connectivity website:

http://www.ibm.com/systems/z/hardware/connectivity/ficon_performance.html

Modified Indirect Data Address Word (MIDAW) facility

The MIDAW facility is a system architecture and software usage that is designed to improve FICON performance. This facility was introduced with z9 servers and is used by the media manager in z/OS.

The MIDAW facility provides a more efficient structure for certain categories of data-chaining I/O operations resulting in improved FICON performance and I/O response times, in particular for extended format data-sets (DB2 is a major user). For more information about FICON, FICON channel performance, and MIDAW, see the following resources:

- ▶ I/O Connectivity web page:
 - <http://www.ibm.com/systems/z/connectivity/>
- ▶ These IBM Redbooks publications
 - *How does the MIDAW Facility Improve the Performance of FICON Channels Using DB2 and other workloads?*, REDP-4201
 - *DS8000 Performance Monitoring and Tuning*, SG24-7146

Extended distance FICON

Using an enhancement to the industry standard FICON architecture (FC-SB-3) can help avoid degradation of performance at extended distances by implementing a protocol for persistent information unit (IU) pacing. Control units that use the enhancement to the architecture can increase the pacing count (the number of IUs allowed to be in flight from channel to control unit). Extended distance FICON allows the channel to remember the last pacing update for use on subsequent operations to help avoid degradation of performance at the start of each new operation.

Improved IU pacing can optimize the use of the link (for example, helps to keep a 4 Gbps link that is fully used at 50 km) and allows channel extenders to work at any distance, with performance results similar to those experienced when using emulation.

The requirements for channel extension equipment are simplified with the increased number of commands in flight. This can benefit z/OS Global Mirror (also referred as Extended Remote Copy, XRC) applications, as the channel extension kit is no longer required to simulate specific channel commands. Simplifying the channel extension requirements can help reduce the total cost of ownership of end-to-end solutions.

Extended Distance FICON is transparent to operating systems and applies to all the FICON Express16S, FICON Express8S, and FICON Express8 features carrying basic FICON traffic (CHPID type FC). For usage, the control unit must support the new IU pacing protocol.

Usage of extended distance FICON is supported by the IBM System Storage® DS8000 series with an appropriate level of Licensed Machine Code (LMC).

z/OS discovery and autoconfiguration

z/OS discovery and autoconfiguration for FICON channels (zDAC) automatically performs a number of I/O configuration definition tasks for new and changed disk and tape controllers that are connected to an FC switch, when attached to a FICON channel.

Users can define a policy, by using the hardware configuration definition (HCD) dialog. Then, when new controllers are added to an I/O configuration or changes are made to existing

controllers, the system is designed to discover them and propose configuration changes that are based on that policy. This policy can include preferences for availability and bandwidth, which includes PAV definitions, control unit numbers, and device number ranges.

zDAC is designed to perform discovery for all systems in a sysplex that support the function. 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. zDAC is designed to simplify I/O configuration on z13 running z/OS and reduce complexity and setup time. zDAC applies to all FICON features supported on z Systems when configured as CHPID type FC.

FICON name server registration

The FICON channel provides the same information to the fabric as is commonly provided by open systems, registering with the name server in the attached FICON directors. This enables a quick and efficient management of storage area network (SAN) and performance of problem determination and analysis.

Platform registration is a standard service that is defined in the Fibre Channel - Generic Services 3 (FC-GS-3) standard (INCITS (ANSI) T11.3 group). It allows a platform (storage subsystem, host, and so on) to register information about itself with the fabric (directors).

This z13 function is transparent to operating systems and applicable to all FICON FICON Express16S, Express8S, and FICON Express8 features (CHPID type FC). For more information, see *IBM System z Connectivity Handbook*, SG24-5444.

FCP connectivity

Fibre Channel Protocol is fully supported on the z13. It is commonly used with Linux on z Systems and supported by the z/VM, z/VSE, and Linux on z Systems operating systems.

Fibre Channel Protocol enhancements for small block sizes

The *Fibre Channel Protocol (FCP)* Licensed Internal Code was modified to help provide increased I/O operations per second for small block sizes. This FCP performance improvement is transparent to operating systems and applies to all the FICON Express 16S, FICON Express8S, and FICON Express8 features, when configured as CHPID type FCP, communicating with SCSI devices.

For more information about FCP channel performance, see the performance technical papers on the z Systems I/O Connectivity web page:

http://www.ibm.com/systems/z/hardware/connectivity/fcp_performance.html

FCP channels to support T10-DIF for enhanced reliability

Recognizing that high reliability is important for maintaining the availability of business-critical applications, the z Systems Fibre Channel Protocol (FCP) has implemented support of the American National Standards Institute's (ANSI) T10 Data Integrity Field (DIF) standard. Data integrity protection fields are generated by the operating system and propagated through the storage area network (SAN). IBM z Systems help 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 layer through the host bus adapter (HBA), where cyclical redundancy checking (CRC) protection is implemented.

T10-DIF support by the FICON Express16S, FICON Express8S, and FICON Express8 features, when defined as CHPID type FCP, is available to z13 and to zEnterprise CPCs. Usage of the T10-DIF standard requires support by the operating system and the storage device.

N_Port ID Virtualization (NPIV)

NPIV is designed to allow the sharing of a single physical FCP channel among operating system images, whether in logical partitions or as z/VM guests. This is achieved by assigning a unique *worldwide port name (WWPN)* for each operating system that is connected to the FCP channel. In turn, each operating system appears to have its own distinct WWPN in the SAN environment, hence enabling separation of the associated FCP traffic on the channel.

Access controls that are based on the assigned WWPN can be applied in the SAN environment. This function can be done by using standard mechanisms such as zoning in SAN switches and logical unit number (LUN) masking in the storage controllers.

WWPN tool

A part of the installation of your z13 server is the planning of the SAN environment (if applicable). IBM has made a stand-alone tool available to assist with this planning before the installation. The tool, which is known as the *WWPN tool*, assigns WWPNs to each virtual Fibre Channel Protocol (FCP) channel/port. This function is done by using the same WWPN assignment algorithms that a system uses when assigning WWPNs for channels using NPIV. Thus, the SAN can be set up in advance, allowing operations to proceed much faster after the server is installed.

The WWPN tool takes a CSV (.csv) file that contains the FCP-specific I/O device definitions and creates the WWPN assignments that are required to set up the SAN. A binary configuration file that can be imported later by the system is also created. The .csv file can either be created manually or exported from the Hardware Configuration Definition/Hardware Configuration Manager (HCD/HCM).

The WWPN tool is available for download from the IBM Resource Link and is applicable to all FICON channels defined as CHPID type FCP (for communication with SCSI devices) on z13.

<http://www.ibm.com/servers/resourceLink/>

LAN connectivity

The z13 offer a wide range of functions that can help consolidate or simplify the LAN environment with the supported OSA-Express features, though also satisfying the demand for more throughput. Improved throughput (mixed inbound/outbound) is achieved by the data router function that was introduced in the OSA-Express3 and enhanced in OSA-Express5S, and OSA-Express4S features.

With the data router, the store and forward technique in DMA is no longer used. The data router enables a direct host memory-to-LAN flow. This function avoids a hop and is designed to reduce latency and to increase throughput for standard frames (1492 bytes) and jumbo frames (8992 bytes).

Queued direct I/O (QDIO) optimized latency mode

QDIO optimized latency mode can help improve performance for applications that have a critical requirement to minimize response times for inbound and outbound data. It optimizes the interrupt processing as noted in the following configurations:

- ▶ For inbound processing, the TCP/IP stack looks more frequently for available data to process, ensuring that any new data is read from the OSA-Express5S, or OSA-Express4S without requiring more program controlled interrupts (PCI).
- ▶ For outbound processing, the OSA-Express5S, or OSA-Express4S looks more frequently for available data to process from the TCP/IP stack, thus not requiring a Signal Adapter (SIGA) instruction to determine whether more data is available.

Inbound workload queuing (IWQ)

IWQ is designed to help reduce overhead and latency for inbound z/OS network data traffic and implement an efficient way for initiating parallel processing. This is achieved by using an OSA-Express5S, or OSA-Express4S feature in QDIO mode (CHPID types OSD and OSX) with multiple input queues and by processing network data traffic that is based on workload types. The data from a specific workload type is placed in one of four input queues (per device), and a process is created and scheduled to run on one of multiple processors, independent from the other three queues. This improves performance because IWQ can use the symmetric multiprocessor (SMP) architecture of the z13.

Virtual local area network (VLAN) support

VLAN is a function of OSA-Express features that takes advantage of the IEEE 802.1q standard for virtual bridged LANs. VLANs allow easier administration of logical groups of stations that communicate as though they were on the same LAN. In the virtualized environment of System z, TCP/IP stacks can exist, potentially sharing OSA-Express features. VLAN provides a greater degree of isolation by allowing contact with a server from only the set of stations that comprise the VLAN.

Virtual MAC (VMAC) support

When sharing OSA port addresses across LPARs, VMAC support enables each operating system instance to have a unique virtual MAC (VMAC) address. All IP addresses associated with a TCP/IP stack are accessible by using their own VMAC address, instead of sharing the MAC address of the OSA port. Advantages include a simplified configuration setup and improvements to IP workload load balancing and outbound routing.

This support is available for Layer 3 mode and is used by z/OS and supported by z/VM for guest usage.

QDIO data connection isolation for the z/VM environment

New workloads increasingly require multitier security zones. In a virtualized environment, an essential requirement is to protect workloads from intrusion or exposure of data and processes from other workloads.

The QDIO data connection isolation enables the following elements:

- ▶ Adherence to security and HIPAA-security guidelines and regulations for network isolation between the instances that share physical network connectivity.
- ▶ Establishment of security zone boundaries that are defined by the network administrators.
- ▶ A mechanism to isolate a QDIO data connection (on an OSA port) by forcing traffic to flow to the external network. This feature ensures that all communication flows only between an operating system and the external network.

Internal routing can be disabled on a per-QDIO connection basis. This support does not affect the ability to share an OSA port. Sharing occurs as it does today, but the ability to communicate between sharing QDIO data connections can be restricted through this support.

QDIO data connection isolation (also known as VSWITCH port isolation) applies to the z/VM environment when using the Virtual Switch (VSWITCH) function and to all of the OSA-Express5S, and OSA-Express4S features (CHPID type OSD) on z13. z/OS supports a similar capability.

QDIO interface isolation for z/OS

Some environments require strict controls for routing data traffic between servers or nodes. In certain cases, the LPAR-to-LPAR capability of a shared OSA port can prevent such controls from being enforced. With interface isolation, internal routing can be controlled on an LPAR basis. When interface isolation is enabled, the OSA discards any packets that are destined for a z/OS LPAR that is registered in the OAT as isolated.

QDIO interface isolation is supported by Communications Server for z/OS V1R11 and later and for all OSA-Express5S, and OSA-Express4S features on z13.

Open Systems Adapter for NCP (OSN)

The OSN support is able to provide channel connectivity from z Systems Operating Systems to IBM Communication Controller for Linux on z Systems (CCL). This function is done by using the Open Systems Adapter for the Network Control Program (OSA for NCP) supporting the Channel Data Link Control (CDLC) protocol.

When SNA solutions that require NCP functions are needed, CCL can be considered as a migration strategy to replace IBM Communications Controllers (374x). The CDLC connectivity option enables z/TPF environments to use CCL.

OSN: The OSN CHPID type is not supported on OSA-Express 5S GbE or OSA-Express4S GbE features.

Network management: Query and display OSA configuration

As more complex functions are added to OSA, the ability for the system administrator to display, monitor, and verify the specific current OSA configuration unique to each operating system is becoming more complex. OSA-Express5S, and OSA-Express4S have the capability for the operating system to query and display the current OSA configuration information (similar to OSA/SF) directly. z/OS uses this OSA capability by providing the TCP/IP operator command, **Display OSAINFO**, which allows the operator to monitor and verify the current OSA configuration, helping to improve the overall management, serviceability, and usability of OSA-Express5S, and OSA-Express4S features.

The **Display OSAINFO** command is exclusive to OSA-Express5S, and OSA-Express4S (CHPID types OSD, OSM, and OSX), the z/OS operating system, and is supported on z/VM for guest usage.

z Systems ensemble connectivity

With the IBM zEnterprise Systems, two CHPID types have been introduced to support the z Systems ensemble:

- ▶ OSA-Express for Unified Resource Manager (OSM) for the *intranode management network (INMN)*
- ▶ OSA-Express for zBX (OSX) for the *intraensemble data network (IEDN)*

The INMN is one of the ensemble's two private and secure internal networks. INMN is used by the Unified Resource Manager functions in the primary HMC.

The OSM connections are through the System Control Hub (SCH) in the z13. The INMN requires two OSA-Express5S 1000BASE-T, or OSA-Express4S 1000BASE-T ports from separate features.

The IEDN is the ensemble's other private and secure internal network. IEDN is used for communications across the virtualized images (LPARs and virtual machines). The IEDN connections use MAC addresses, not IP addresses (Layer 2 connection).

The OSX connections are from the z Systems CPC to the IEDN TOR switches in zBX. The IEDN requires two OSA-Express5S 10 GbE, or OSA-Express4S 10 GbE ports from separate features.

HiperSockets

HiperSockets feature has been referred to as the “network in a box.” HiperSockets simulates LANs entirely in the hardware. The data transfer is from LPAR memory to LPAR memory, mediated by microcode. The z13 support up to 32 HiperSockets. One HiperSockets network can be shared by up to 85 LPARs on a z13. Up to 4096 communication paths support a total of 12,288 IP addresses across all 32 HiperSockets.

HiperSockets Layer 2 support

The HiperSockets internal networks can support two transport modes:

- ▶ Layer 2 (link layer)
- ▶ Layer 3 (network or IP layer)

Traffic can be Internet Protocol (IP) Version 4 or Version 6 (IPv4, IPv6) or non-IP (such as AppleTalk, DECnet, IPX, NetBIOS, SNA, or others). HiperSockets devices are independent of protocol and Layer 3. Each HiperSockets device has its own Layer 2 Media Access Control (MAC) address, which is designed to allow the use of applications that depend on the existence of Layer 2 addresses such as Dynamic Host Configuration Protocol (DHCP) servers and firewalls.

Layer 2 support can help facilitate server consolidation. Complexity can be reduced, network configuration is simplified and intuitive, and LAN administrators can configure and maintain the mainframe environment the same way as they do for a non-mainframe environment. HiperSockets Layer 2 support is provided by Linux on System z, and by z/VM for guest usage.

HiperSockets Multiple Write Facility

HiperSockets performance is enhanced to allow for the streaming of bulk data over a HiperSockets link between LPARs. The receiving LPAR can now process a much larger amount of data per I/O interrupt. This enhancement is transparent to the operating system in the receiving LPAR. HiperSockets Multiple Write Facility, with fewer I/O interrupts, reduces CPU use of the sending and receiving LPAR.

The HiperSockets Multiple Write Facility is supported in the z/OS environment.

zIIP-Assisted HiperSockets for large messages

In z/OS, HiperSockets are enhanced for zIIP usage. Specifically, the z/OS Communications Server allows the HiperSockets Multiple Write Facility processing for outbound large messages that originate from z/OS to be performed on a zIIP.

zIIP-Assisted HiperSockets can help make highly secure and available HiperSockets networking an even more attractive option. z/OS application workloads that are based on XML, HTTP, SOAP, Java, and traditional file transfer can benefit from zIIP enablement by lowering general-purpose processor use for such TCP/IP traffic.

When the workload is eligible, the TCP/IP HiperSockets device driver layer (write) processing is redirected to a zIIP, which unblocks the sending application.

zIIP Assisted HiperSockets for large messages is available on z13 with z/OS V1R12 and later releases.

HiperSockets Network Traffic Analyzer (NTA)

HiperSockets NTA is a function that is available in the LIC of the z13. It can simplify problem isolation and resolution by allowing Layer 2 and Layer 3 tracing of HiperSockets network traffic.

HiperSockets NTA allows Linux on z Systems to control tracing of the internal virtual LAN. It captures records into host memory and storage (file systems) that can be analyzed by system programmers and network administrators, using Linux on z Systems tools to format, edit, and process the trace records.

A customized HiperSockets NTA rule enables authorizing an LPAR to trace messages only from LPARs that are eligible to be traced by the NTA on the selected IQD channel.

HiperSockets Completion Queue

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 communication and can be especially helpful in burst situations.

HiperSockets Completion Queue function is supported on the z13 running z/OS V1R13, z/VM V6R2 (with maintenance) and later, z/VSE V5R1 (with maintenance), Red Hat Enterprise Linux (RHEL) 6.2, or SUSE Linux Enterprise Server (SLES) 11 SP2 (with maintenance), and later.

HiperSockets integration with the intraensemble data network

The z13 servers provide the capability to integrate HiperSockets connectivity with the *intraensemble data network (IEDN)*. Thus the reach of the HiperSockets network is extended to outside the CPC to the entire ensemble, which is displayed as a single, Layer 2 network. Because HiperSockets and IEDN are both internal z Systems networks, the combination allows System z virtual servers to use an optimal path for communications.

The support of HiperSockets integration with the IEDN function is available starting with z/OS Communication Server V1R13.

HiperSockets Virtual Switch Bridge Support

The z/VM virtual switch is enhanced to transparently bridge a guest virtual machine network connection on a HiperSockets LAN segment. This bridge allows a single HiperSockets guest virtual machine network connection to also directly communicate with the following systems:

- ▶ Other guest virtual machines on the virtual switch
- ▶ External network hosts through the virtual switch OSA UPLINK port

z/VM V6R2 and later, TCP/IP, and Performance Toolkit APARs are required for this support.

A HiperSockets channel by itself is only capable of providing intra-CPC communications. The HiperSockets Bridge Port allows a virtual switch to connect z/VM guests by using real HiperSockets devices, the ability to communicate with hosts that reside externally to the CPC. The virtual switch HiperSockets Bridge Port eliminates the need to configure a separate next hop router on the HiperSockets channel to provide connectivity to destinations that are outside of a HiperSockets channel.

z/VSE fast path to Linux support

Linux Fast Path (LFP) allows z/VSE TCP/IP applications to communicate with the TCP/IP stack on Linux without using a TCP/IP stack on z/VSE. LFP for use in a z/VM guest environment is supported since z/VSE V5R1. When LFP is used in an LPAR environment, it requires the HiperSockets Completion Queue function available on z Systems CPCs. LFP in an LPAR environment is supported since z/VSE V5R1.

Coupling and Server Time Protocol connectivity

Coupling connectivity in support of Parallel Sysplex environments is provided on the z13 by the following features:

- ▶ New PCIe Gen3, Integrated Coupling Adapter (ICA SR), which allows two ports coupling links connectivity for a distance of up to 150 m (492 feet) at 8GBps each.
- ▶ HCA3-O, 12x InfiniBand coupling links offering up to 6 GBps of bandwidth between z13, zBC12, z196 and z114 systems, for a distance of up to 150 m (492 feet).
- ▶ HCA3-O LR, 1x InfiniBand up to 5 Gbps connection bandwidth between z13, zEC12, zBC12, z196 and z114 for a distance of up to 10 km (6.2 miles).
- ▶ Internal Coupling Channels (ICs), operating at memory speed.

All coupling link types can be used to carry Server Time Protocol (STP) messages. The z13 does not support ISC-3 connectivity. Also, HCA2-O 12x and HCA2-O LR 1x InfiniBand features are not supported in z13.

3.2.8 Cryptography

z13 provides cryptographic functions that, from an application program perspective, can be grouped into the following:

- ▶ Synchronous cryptographic functions, provided by the CP Assist for Cryptographic Function (CPACF)
- ▶ Asynchronous cryptographic functions, provided by the Crypto Express features

CP Assist for Cryptographic Function (CPACF)

CPACF offers a set of symmetric cryptographic functions for high performance encryption and decryption with clear key operations for SSL/TLS, VPN, and data-storing applications that do not require FIPS¹¹ 140-2 level 4 security. The CPACF is integrated with the compression unit in the coprocessor (CoP) in the z13 microprocessor core.

The CPACF protected key is a function that is designed to facilitate the continued privacy of cryptographic key material while keeping the wanted high performance. CPACF ensures that key material is not visible to applications or operating systems during encryption operations. CPACF protected key provides substantial throughput improvements for large-volume data encryption and low latency for encryption of small blocks of data.

The cryptographic assist includes support for the following functions:

- ▶ Data Encryption Standard (DES) data encrypting and decrypting:
 - DES supports the following key types:
 - Single-length key DES
 - Double-length key DES
 - Triple-length key DES (T-DES)
- ▶ Advanced Encryption Standard (AES) for 128-bit, 192-bit, and 256-bit keys

¹¹ Federal Information Processing Standards (FIPS) 140-2 Security Requirements for Cryptographic Modules

- ▶ Pseudo random number generation (PRNG)
- ▶ Message Authentication Code (MAC)
- ▶ Hashing algorithms: SHA-1 and SHA-2 support for SHA-224, SHA-256, SHA-384, and SHA-512

SHA-1 and SHA-2 support for SHA-224, SHA-256, SHA-384, and SHA-512 are shipped enabled on all servers and do not require the CPACF enablement feature. The CPACF functions are supported by z/OS, z/VM, z/VSE, z/TPF, and Linux on z Systems.

Crypto Express5S

The Crypto Express5S represents the newest generation of the *Peripheral Component Interconnect Express (PCIe)* cryptographic coprocessors. It is an optional feature exclusive to the z13. This feature provides a secure programming and hardware environment wherein crypto processes are performed. Each cryptographic coprocessor includes a general-purpose processor, non-volatile storage, and specialized cryptographic electronics.

The Crypto Express5S has one PCIe adapter per feature. For availability reasons, a minimum of two features is required. Up to 16 Crypto Express5S features are supported (16 PCI Express adapters per z13). The Crypto Express5S feature occupies one I/O slot in a z13 PCIe I/O drawer.

Each adapter can be configured as a Secure IBM CCA coprocessor, a Secure IBM Enterprise PKCS #11 (EP11) coprocessor or as an accelerator.

Crypto Express5S has been enhanced to provide domain support for up to 85 logical partitions on IBM z13.

The accelerator function is designed for maximum-speed Secure Sockets Layer and Transport Layer Security (SSL/TLS) acceleration, rather than for specialized financial applications for secure, long-term storage of keys or secrets. The Crypto Express5S can also be configured as one of the following configurations:

- ▶ *Secure IBM CCA coprocessor* for Federal Information Processing Standard (FIPS) 140-2 Level 4 certification. This standard includes secure key functions and is optionally programmable to deploy more functions and algorithms using User Defined Extension (UDX).
- ▶ *Secure IBM Enterprise PKCS #11 (EP11) coprocessor*, implementing an industry standardized set of services that adheres to the PKCS #11 specification v2.20 and more recent amendments. It was designed for extended FIPS and Common Criteria evaluations to meet industry requirements.

This new cryptographic coprocessor mode introduced the PKCS #11 secure key function.

TKE feature: The Trusted Key Entry (TKE) Workstation feature is required for supporting the administration of the Crypto Express5S when configured as an Enterprise PKCS #11 coprocessor.

When the Crypto Express5S PCI Express adapter is configured as a secure IBM CCA coprocessor, it still provides accelerator functions. However, up to three times better performance for those functions can be achieved if the Crypto Express5S PCI Express adapter is configured as an accelerator.

Crypto Express4S

The *Crypto Express4S* can be carried forward to z13. It is an optional feature available to zEC12, zBC12 or z13. This feature provides a secure programming and hardware

environment wherein crypto processes are performed. Each cryptographic coprocessor includes a general-purpose processor, non-volatile storage, and specialized cryptographic electronics.

The Crypto Express4S has one PCIe adapter per feature. For availability reasons, a minimum of two features is required. Up to 16 Crypto Express4S features are supported (16 PCI Express adapters per z13). The Crypto Express4S feature occupies one I/O slot in a z13 PCIe I/O drawer.

Each adapter can be configured as a Secure IBM CCA coprocessor, a Secure IBM Enterprise PKCS #11 (EP11) coprocessor or as an accelerator.

With the new LIC, Crypto Express4S has been enhanced to provide domain support for up to 85 logical partitions on IBM z13.

The accelerator function is designed for maximum-speed SSL/TLS acceleration, rather than for specialized financial applications for secure, long-term storage of keys or secrets. The Crypto Express4S can also be configured as one of the following configurations:

- ▶ *Secure IBM CCA coprocessor* for Federal Information Processing Standard (FIPS) 140-2 Level 4 certification. This standard includes secure key functions and is optionally programmable to deploy more functions and algorithms using User Defined Extension (UDX).
- ▶ *Secure IBM Enterprise PKCS #11 (EP11) coprocessor*; implementing an industry standardized set of services that adheres to the PKCS #11 specification v2.20 and more recent amendments. It was designed for extended FIPS and Common Criteria evaluations to meet industry requirements.

This new cryptographic coprocessor mode introduced the PKCS #11 secure key function.

TKE feature: The Trusted Key Entry (TKE) Workstation feature is required for supporting the administration of the Crypto Express4S when configured as an Enterprise PKCS #11 coprocessor.

- ▶ When the Crypto Express4S PCI Express adapter is configured as a secure IBM CCA coprocessor, it still provides accelerator functions. However, up to three times better performance for those functions can be achieved if the Crypto Express4S PCI Express adapter is configured as an accelerator

Web deliverables

For z/OS downloads, see the z/OS website:

<http://www.ibm.com/systems/z/os/zos/downloads/>

3.3 Capacity and performance

The z13 offers significant increases in capacity and performance over its predecessor, zEC12. Many factors contribute to this effect, including the larger number of processors, individual processor performance, memory caches, simultaneous multithreading-2 (SMT-2) and machine instructions, including the new single-instruction, multiple-data (SIMD). Subcapacity settings continue to be offered.

3.3.1 Capacity settings

The z13 expands the offer on subcapacity settings. Finer granularity in capacity levels allows the growth of installed capacity to more closely follow the enterprise growth, for a smoother, pay-as-you-go investment profile. There are many performance and monitoring tools that are available on z Systems environments that are coupled with the flexibility of the capacity on-demand options (see 3.3.2, “z13 Capacity on Demand (CoD)” on page 91). These features help to manage growth by making capacity available when needed.

Regardless of the installed model, the z13 offers four distinct capacity levels for the first 30 central processors (CP): full capacity and three subcapacities. These processors deliver the scalability and granularity to meet the needs of medium-sized enterprises, while also satisfying the requirements of large enterprises that have large-scale, mission-critical transaction and data-processing requirements.

A capacity level is a setting of each CP to a subcapacity of the full CP capacity. The clock frequency of those processors remains unchanged. The capacity adjustment is achieved through other means.

Full capacity CPs are identified as CP7. On the z13 server, 141 CPs can be configured as CP7. The three subcapacity levels are identified by CP6, CP5, and CP4, respectively, and are displayed in hardware descriptions as feature codes on the CPs.

If more than 30 CPs are configured to the system, then all must be full capacity because all CPs must be on the same capacity level. Granular capacity adds 90 subcapacity settings to the 141 capacity settings that are available with full capacity CPs (CP7). The 231 distinct capacity settings in the system, provide for a range of over 1:320 in processing power.

A processor that is characterized as anything other than a CP, such as a zIIP, an IFL, or an ICF, is always set at full capacity. There is, correspondingly, a separate pricing model for non-CP processors regarding purchase and maintenance prices, and various offerings for software licensing.

On z13, the CP subcapacity levels are a fraction of full capacity, as follows:

- ▶ Model 7xx = 100%
- ▶ Model 6xx = 63%
- ▶ Model 5xx = 44%
- ▶ Model 4xx = 15%

For administrative purposes, systems that have only ICF or IFL processors, are now given a capacity setting of 400. For either of these systems, it is possible to have up to 141 ICFs or IFLs, which always run at full capacity.

Figure 3-2 gives more details about z13 full capacity and subcapacity offerings.

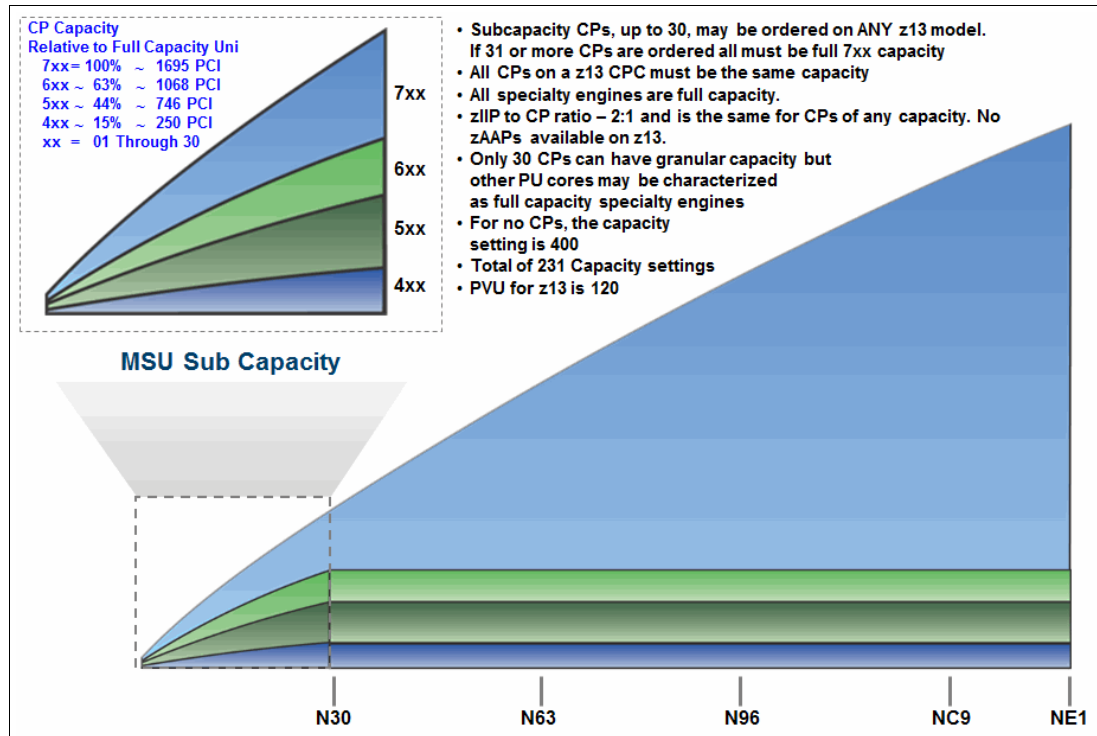


Figure 3-2 z13 full and subcapacity CP offerings¹²

To help size a z Systems platform to fit clients' requirements, IBM provides a no-cost tool that reflects the latest IBM LSPR measurements, called the *IBM Processor Capacity Reference for z Systems (zPCR)*. The tool can be downloaded from the following web page:

<http://www.ibm.com/support/techdocs/atmastr.nsf/WebIndex/PRS1381>

Also, see 3.3.3, "z13 performance" on page 93 for more information about LSPR measurements.

3.3.2 z13 Capacity on Demand (CoD)

The z13 continues to offer on-demand offerings. The offerings provide flexibility and control to the client, ease the administrative burden in the handling of the offerings, and give the client finer control over resources that are needed to meet the resource requirements in various situations.

The z13 can perform concurrent upgrades, providing an increase of processor capacity with no server outage. In most cases, with operating system support, a concurrent upgrade can also be non disruptive to the operating system. It is important to consider that these upgrades are based on the enablement of resources already physically present in the z13.

Capacity upgrades cover both permanent and temporary changes to the installed capacity. The changes can be done using the Customer Initiated Upgrade (CIU) facility, without requiring IBM service personnel involvement. Such upgrades are initiated through the web by using IBM Resource Link. Use of the CIU facility requires a special contract between the client and IBM, through which terms and conditions for online capacity on demand (CoD)

¹² PCI - Processor Capacity Index

buying of upgrades and other types of CoD upgrades are accepted. For more information, consult the IBM Resource Link.

For more information about the CoD offerings, see *IBM z13 Technical Guide*, SG24-8251.

Permanent upgrades

Permanent upgrades of processors (CP, IFL, ICF, zIIP, and SAP) and memory, or changes to a server's Model-Capacity Identifier, up to the limits of the installed processor capacity on an existing z13, can be performed by the client through the IBM Online Permanent Upgrade offering by using the CIU facility.

Temporary upgrades

Temporary upgrades of a z13 can be done by On/Off CoD, Capacity Backup (CBU), or Capacity for Planned Event (CPE) ordered from the CIU facility.

On/Off CoD function

On/Off CoD is a function that is available on the z13 that enables concurrent and temporary capacity growth of the CPC. On/Off CoD can be used for client peak workload requirements, for any length of time, has a daily hardware charge and can have an associated software charge. On/Off CoD offerings can be pre-paid or post-paid. Capacity tokens are available on z13. Capacity tokens are always present in prepaid offerings and can be present in post-paid if the client wants that. In both cases capacity tokens are being used to control the maximum resource and financial consumption.

When using the On/Off CoD function, the client can concurrently add processors (CP, IFL, ICF, zIIP, and SAP), increase the CP capacity level, or both.

Capacity Backup (CBU) function

CBU allows the client to perform a concurrent and temporary activation of additional CP, ICF, IFL, zIIP, and SAP, an increase of the CP capacity level, or both. This function can be used in the event of an unforeseen loss of z Systems capacity within the client's enterprise, or to perform a test of the client's disaster recovery procedures. The capacity of a CBU upgrade cannot be used for peak workload management.

CBU features are optional and require unused capacity to be available on CPC drawers of the backup system, either as unused PUs or as a possibility to increase the CP capacity level on a subcapacity system, or both. A CBU contract must be in place before the LIC-CC code that enables this capability can be loaded on the system. An initial CBU record provides for one test for each CBU year (each up to 10 days in duration) and one disaster activation (up to 90 days in duration). The record can be configured to be valid for up to five years. Client can also order additional tests for a CBU record if needed, in quantities of five tests up to a maximum of 15.

Proper use of the CBU capability does not incur any additional software charges from IBM.

Capacity for Planned Event (CPE) function

CPE allows the client to perform a concurrent and temporary activation of additional CPs, ICFs, IFLs, zIIPs, and SAPs, an increase of the CP capacity level, or both. This function can be used in the event of a planned outage of z Systems capacity within the client's enterprise (for example, data center changes, system or power maintenance). CPE cannot be used for peak workload management and can be active for a maximum of three days.

The CPE feature is optional and requires unused capacity to be available on CPC drawers of the back-up system, either as unused PUs or as a possibility to increase the CP capacity level

on a subcapacity system, or both. A CPE contract must be in place before the LIC-CC that enables this capability can be loaded on the system.

z/OS capacity provisioning

Capacity provisioning helps clients manage the CP and zIIP capacity of z13 that is running one or more instances of the z/OS operating system. By using z/OS Capacity Provisioning Manager (CPM) component, On/Off CoD temporary capacity can be activated and deactivated under control of a defined policy. Combined with functions in z/OS, the z13 provisioning capability gives the client a flexible, automated process to control the configuration and activation of On/Off CoD offerings.

3.3.3 z13 performance

The z Systems microprocessor chip of the z13 has a high-frequency design that uses IBM leading technology and offers more cache per core than other chips. In addition, an enhanced instruction execution sequence, along with processing technologies such as SMT delivers world-class per-thread performance. z/Architecture is enhanced by providing more instructions, including SIMD, that are intended to deliver improved CPU-centric performance and analytics. For CPU-intensive workloads, more gains can be achieved by multiple compiler-level improvements. Improved performance of the z13 is a result of the enhancements that we described in Chapter 2, “Hardware overview” on page 25 and 3.2, “z13 technology improvements” on page 67.

The z13 Model NE1 is designed to offer up to 40% more capacity than the largest zEC12 system. Uniprocessor performance has also increased significantly. A z13 Model 701 offers, based on an average workload, performance improvements of up to 10% over the zEC12 Model 701.

However, variations on the observed performance increase are dependent upon the workload type.

LSPR workload suite: z13 changes

To help in better understanding workload variations, IBM provides a no-cost tool that is called the *IBM Processor Capacity Reference for z Systems (zPCR)*:

<http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS1381>

IBM continues to measure performance of the systems by using various workloads and publishes the results in the Large Systems Performance Reference (LSPR) report. The LSPR is available at the following web page:

<https://www-304.ibm.com/servers/resourceLink/lib03060.nsf/pages/lsprindex?>

The MSU ratings are available at the following web page:

<http://www-03.ibm.com/systems/z/resources/swprice/reference/exhibits/hardware.html>

Historically, LSPR capacity tables, including pure workloads and mixes, have been identified with application names or a software characteristic. Examples are as follows:

- ▶ CICS
- ▶ IMS
- ▶ OLTP-T: Traditional online transaction processing workload (formerly known as IMS)
- ▶ CB-L: Commercial batch with long-running jobs
- ▶ LoIO-mix: Low I/O Content Mix Workload
- ▶ TI-mix: Transaction Intensive Mix Workload

However, capacity performance is more closely associated with how a workload uses and interacts with a particular processor *hardware design*. Workload capacity performance is sensitive to three major factors:

- ▶ Instruction path length
- ▶ Instruction complexity
- ▶ Memory hierarchy

With the availability of the CPU measurement facility (MF) data, the ability to gain insight into the interaction of workload and *hardware design* in production workloads has arrived. CPU MF data helps LSPR to adjust workload capacity curves that are based on the underlying hardware sensitivities, in particular the processor access to caches and memory. This is known as *nest activity intensity*. With the IBM zEnterprise System, the LSPR introduced three workload capacity categories that replace all prior primitives and mixes:

- ▶ LOW (relative nest intensity):

A workload category that represents light use of the memory hierarchy. This category is similar to past high scaling primitives.

- ▶ AVERAGE (relative nest intensity):

A workload category that represents average use of the memory hierarchy. This category is similar to the past LoIO-mix workload and is expected to represent most of the production workloads.

- ▶ HIGH (relative nest intensity):

A workload category that represents heavy use of the memory hierarchy. This category is similar to the past TI-mix workload.

These categories are based on the relative nest intensity, which is influenced by many variables such as application type, I/O rate, application mix, CPU usage, data reference patterns, LPAR configuration, and the software configuration that is running, among others. CPU MF data can be collected by z/OS System Measurement Facility on SMF 113 records.

Guidance in converting LSPR previous categories to the new ones is provided, and built-in support is added to the IBM zPCR tool.

In addition to low, average, and high categories, the latest zPCR provides the low-average and average-high mixed categories, which allow better granularity for workload characterization.

The LSPR tables continue to rate all z/Architecture processors running in LPAR mode and 64-bit mode. The single-number values are based on a combination of the default mixed workload ratios, typical multi-LPAR configurations, and expected early-program migration scenarios. In addition to z/OS workloads used to set the single-number values, the LSPR tables contain information that pertains to Linux and z/VM environments.

The LSPR contains the internal throughput rate ratios (ITRR) for the z13 and the previous generations of processors that are based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user might experience varies depending on factors such as the amount of multiprogramming in the user's job stream, the I/O configuration, and the workload processed.

Experience demonstrates that System z servers can be run at up to 100% utilization levels, sustained, although most clients prefer to leave a bit of white space and run at 90% or slightly under. For any capacity comparison, using "one number" such as the MIPS or MSU metrics is not a valid method. That is why, while doing capacity planning, we suggest using zPCR and

involving IBM technical support. For more information about z13 performance, see the *IBM z13 Technical Guide*, SG24-8251

Throughput optimization with z13

The z990 was the first server to use the concept of books. The memory and cache structure implementation in the z13 CPC drawers have been enhanced from the z990, through successive system generations to the z13 to provide sustained throughput and performance improvements. Despite the fact that the memory is distributed throughout the CPC drawers and the CPC drawers having individual levels of caches private to the cores and shared by the cores, all processors have access to the highest level of caches and all of the memory. Thus, the system is managed as a memory coherent symmetric multiprocessor (SMP).

Processors within the z13 CPC drawer structure have different distance-to-memory attributes. As described in 2.3, “z13 CPC drawers, and single chip modules” on page 30, CPC drawers are connected in a star configuration to minimize the distance. Other non-negligible effects result from data latency when grouping and dispatching work on a set of available logical processors. To minimize latency, the system attempts to dispatch and later re-dispatch work to a group of physical CPUs that share the same cache levels.

PR/SM manages the use of physical processors by logical partitions by dispatching the logical processors on the physical processors. But PR/SM is not aware of which workloads are being dispatched by the operating system in what logical processors. The Workload Manager (WLM) component of z/OS has the information at the task level, but is unaware of physical processors. This disconnect is solved by enhancements that allow PR/SM and WLM to work more closely together. They can cooperate to create an affinity between task and physical processor rather than between logical partition and physical processor. This is known as HiperDispatch.

HiperDispatch

HiperDispatch, introduced with the z10 Enterprise Class, and evolved in z196, and zEC12 is further enhanced in z13. It combines two functional enhancements, one in the z/OS dispatcher and one in PR/SM. This function is intended to improve efficiency both in the hardware and in z/OS. z/VM HiperDispatch is introduced by z/VM V6R3.

In general, the PR/SM dispatcher assigns work to the minimum number of logical processors that are needed for the priority (weight) of the LPAR. On z13, PR/SM attempts to group the logical processors into the same node (see Figure 2-5 on page 32) or in the neighbor node in the same CPC drawer and, if possible, in the same chip. This results in reducing the multi-processor effects, maximizing use of shared cache, and lowering the interference across multiple partitions.

The z/OS dispatcher is enhanced to operate with multiple dispatching queues, and tasks are distributed among these queues. Specific z/OS tasks can be dispatched to a small subset of logical processors. PR/SM ties these logical processors to the same physical processors, thus improving the hardware cache reuse and locality of reference characteristics, such as reducing the rate of cross communication.

To use the correct logical processors, the z/OS dispatcher obtains the necessary information from PR/SM through interfaces that are implemented on the z13. The entire z13 stack (hardware, firmware, and software) now tightly collaborates to obtain the full potential of the hardware. z/VM HiperDispatch provides support similar to the z/OS one.

The HiperDispatch function is enhanced on the z13 to use the new eight-core chip and improve computing efficiency. It is possible to dynamically switch on and off HiperDispatch without requiring an initial program load (IPL).

Note: HiperDispatch is required if SMT is enabled.

3.4 zSystems common time functions

Each server must have an accurate time source to maintain a time-of-day value. Logical partitions use their system's time. When system images participate in a Sysplex, coordinating the time across all system images in the Sysplex is critical to its operation.

The z13 supports the Server Time Protocol and can participate in a STP-only coordinated timing network (CTN).

3.4.1 Server Time Protocol (STP)

Server Time Protocol (STP) is a message-based protocol in which timekeeping information is passed over data links between servers. The timekeeping information is transmitted over externally defined coupling links. The STP feature is the supported method for maintaining time synchronization between the z13 and coupling facilities (CF) in Sysplex environments.

The STP design uses a concept called *Coordinated Timing Network (CTN)*. A CTN is a collection of CPCs that are time synchronized to a time value called *Coordinated Server Time (CST)*. Each CPC to be configured in a CTN must be STP-enabled. STP is intended for CPCs that are configured to participate in a Parallel Sysplex or CPCs that are not in a Parallel Sysplex, but must be time synchronized.

STP is implemented in LIC as a system-wide facility of the z13 and other z Systems CPCs. STP presents a single view of time to PR/SM and provides the capability for multiple CPCs to maintain time synchronization with each other. The z13 server is enabled for STP by installing the STP feature code. Additional configuration is required for a z13 to become a member of a CTN.

Important: The IBM z13 *cannot* join a CTN that includes a z10 or earlier system as a member. Because the z10 was the last server that supported the IBM Sysplex Timer (9037) connectivity, the z13 cannot be configured as a member of a mixed CTN. The z13 can join only an STP-only CTN.

STP provides the following additional value over the former used-time synchronization method by a Sysplex Timer:

- ▶ STP supports a multi-site timing network of up to 100 km (62 miles) over fiber optic cabling, without requiring an intermediate site. This protocol allows a Parallel Sysplex to span these distances and reduces the cross-site connectivity that is required for a multi-site Parallel Sysplex.
- ▶ The STP design allows more stringent synchronization between CPCs and CFs by using communication links that are already used for the sysplex connectivity. With the z13, STP supports coupling links over InfiniBand or Integrated Coupling Adapter links.
- ▶ STP helps eliminate infrastructure requirements, such as power and space, needed to support the Sysplex Timers and helps eliminate maintenance costs that are associated with the Sysplex Timers.
- ▶ STP can reduce the fiber optic infrastructure requirements in a multi-site configuration because it can use the coupling links that are already in use.

Server Time Protocol (STP) recovery enhancement

When HCA3-O, HCA3-O LR, or ICA SR coupling links are used, an unambiguous “*going away signal*” is sent when the server on which the HCA3 or ICA is running is about to enter a failed state. When the *going away signal* that is sent by the Current Time Server (CTS) in an STP-only CTN is received by the Backup Time Server (BTS), the BTS can safely take over as the CTS. The takeover can occur 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.

The previously available STP recovery design is still available for the cases when a going away signal is not received or for other failures different from a system failure.

3.4.2 Network Time Protocol (NTP) client support

The use of NTP servers as an external time source (ETS) usually fulfills a requirement for a time source or common time reference across heterogeneous platforms and for providing a higher time accuracy.

NTP client support is available in the Support Element (SE) code of the z13. The code interfaces with the NTP servers. This interaction allows an NTP server to become the single time source for z13 and for other servers that have NTP clients. NTP can be used only for an STP-only CTN environment.

Pulse per second (PPS) support

Two oscillator cards (OSC), included as a standard feature of the z13, provide a dual-path interface for the PPS signal. The cards contain a BNC connector for PPS attachment at the rear side of the CPC frame A. The redundant design allows continuous operation, in case of failure of one card, and concurrent card maintenance.

STP tracks the highly stable accurate PPS signal from the NTP server. PPS maintains accuracy of 10 μ s as measured at the PPS input of the z13 CPC.

If STP uses an NTP server without PPS, a time accuracy of 100 ms to the ETS is maintained. A cable connection from the PPS port to the PPS output of an NTP server is required when the z13 is configured for using NTP with PPS as the external time source for time synchronization.

NTP server on HMC with security enhancements

The NTP server capability on the HMC addresses the potential security concerns that users can have for attaching NTP servers directly to the HMC/SE LAN. When using the HMC as the NTP server, the pulse per second capability is not available.

HMC NTP broadband authentication support for z13

The HMC NTP authentication capability is provided by the HMC Level 2.12.0 and later. SE NTP support stays unchanged. To use this option for STP, configure the HMC as the NTP server for the SE.

The authentication support of the HMC NTP server can be set up in either of two ways:

- ▶ NTP requests are UDP socket packets and cannot pass through the proxy. If a proxy to access outside corporate data center is used, then this proxy must be configured as an NTP server to get to target servers on the web. Authentication can be set up on the client's proxy to communicate to the target time sources.
- ▶ If a firewall is used, HMC NTP requests must pass through the firewall. Clients in this configuration should use the HMC authentication to ensure untampered time stamps.

For more details about STP, see the following books:

- ▶ *Server Time Protocol Planning Guide*, SG24-7280
- ▶ *Server Time Protocol Implementation Guide*, SG24-7281

3.5 Hardware Management Console (HMC) functions

The HMC and SE are appliances that provide hardware platform management for z Systems. Hardware platform management covers a complex set of setup, configuration, operation, monitoring, and service management tasks and services that are essential to the use of the z Systems hardware platform product.

When tasks are performed on the HMC, the commands are sent to one or more SEs, which issue commands to their CPCs and zBXs.

HMC/SE Version 2.13.0 is the current version available for the z13. See *IBM z13 Technical Guide*, SG24-8251, for more information about these HMC functions and capabilities, and also zBX mod 004.

3.5.1 HMC key enhancements for z13

The HMC application has several enhancements:

- ▶ Tasks and panels are updated to support configuring and managing Flash Express, IBM zAware, zEDC Express, and 10GbE RoCE Express features.
- ▶ The Backup for HMC and SE can be saved additional to an FTP server for z13.
- ▶ OSA/SF is available on the HMC for specific OSA-Express features.
- ▶ For STP NTP broadband security, authentication is added to the HMC's NTP communication with NTP time servers and panels to configure STP are redesigned.
- ▶ Modem support is removed from HMC. The Remote Support Facility (RSF) for IBM support, service, and configuration update is only possible through an Ethernet broadband connection.
- ▶ The Monitors Dashboard on the HMC and SE is enhanced with an adapter table. The Crypto Utilization percentage is displayed on the Monitors Dashboard according to the PCHID number. The adapter table also displays Flash Express. You can now display the activity for an Logical Partition (LPAR) by processor type and the Monitors Dashboard is enhanced with showing simultaneous multithreading (SMT) usage.
- ▶ The Environmental Efficiency Statistic Task provides historical power consumption and thermal information for z13 on the HMC. This task provides similar data along with a historical summary of processor and channel use. The initial chart display shows the 24 hours that precede the current time so that a full 24 hours of recent data is displayed. The data is presented in table form, graphical (histogram) form, and it can also be exported to a .csv formatted file so that it can be imported into a spreadsheet.
- ▶ The microcode update to a specific bundle is possible.

Statements of Direction^a:

Removal of support for Classic Style User Interface on the Hardware Management Console and Support Element: The IBM z13 will be the last z Systems server to support Classic Style User Interface. In the future, user interface enhancements will be focused on the Tree Style User Interface.

Removal of support for the Hardware Management Console Common Infrastructure Model (CIM) Management Interface: IBM z13 will be the last z Systems server to support the Hardware Console Common Infrastructure module (CIM) Management Interface. The Hardware Management Console Simple Network Management Protocol (SNMP), and Web Services Application Programming Interfaces (APIs) will continue to be supported.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

For more information about the key capabilities and enhancements of the HMC, see *IBM z13 Technical Guide*, SG24-8251.

3.6 z13 power and cooling functions

As environmental concerns raise the focus on energy consumption, z13 offer a holistic focus on the environment. New efficiencies and functions, such as an improved integrated cooling system and static power save mode enable a reduction of energy usage. The new design of the rear door covers addresses past data center issues regarding airflow challenges. You have the possibility that the covers are in a *vectored down or up* orientation for the outgoing air.

3.6.1 High voltage DC power

In today's data centers, many businesses are paying increasing electric bills and are also running out of available power.

This feature allows CPCs to directly use the high voltage DC distribution in new, green data centers. A direct HVDC¹³ data center power design can improve data center energy efficiency by removing the need for a DC-to-AC and AC-to-DC inversion/conversion steps. The z13 bulk power supplies are able to support HVDC, so the only difference in shipped HW to implement the option is the DC power cords.

Because HVDC is a new technology, there are multiple proposed standards. The z13 supports both ground-referenced and dual-polarity (differential) HVDC supplies, such as +/-190V or +/-260V, or +380V. Beyond the data center uninterruptible power supply (UPS) and power distribution energy savings, a z13 running on HVDC power draws 1 - 3% less input power by eliminating the AC to DC internal conversion. HVDC does not change the number of power cords that a system requires.

¹³ HVDC - High Voltage Direct Current

3.6.2 Integrated battery feature (IBF)

IBF is an optional feature on the z13. See Figure 2-2 on page 29 or Figure 2-3 on page 30 for a pictorial view of the location of IBF. IBF provides the function of a local uninterrupted power source.

The IBF further enhances the robustness of the power design, increasing power line disturbance immunity. The feature provides battery power to preserve processor data if there is a total loss of power from the utility company. The IBF can hold power briefly during a brownout, or for orderly shutdown in a longer outage.

3.6.3 Power capping and power saving

Power capping limits the maximum power consumption and reduces the cooling requirements especially with the zBX. A z13 server does not support power capping.

A static power-saving mode is also available for the z13 when the Unified Resource Manager Automate Firmware Suite feature is installed. It uses frequency and voltage reduction to reduce energy consumption and can be set up ad hoc or as a scheduled operation. It means, for example, in periods of low utilization or on CBU systems, that clients can set the system in a static power-saving mode. Power Saving functions are also provided for the blades in the zBX.

3.6.4 Power estimation tool

The power estimation tool for z13 is a web-based tool that is available to registered users of IBM Resource Link. The tool allows entering the exact server configuration to produce an estimate of power consumption.

Log in to IBM Resource Link and go to **Planning Tools Power Estimation Tools**. Specify the quantity for the features that are installed in the machine. The tool estimates the power consumption for the specified configuration. The tool does not verify whether the specified configuration can be physically built.

Power consumption: The exact power consumption for a machine will vary. The objective of the tool is to produce an estimation of the power requirements to aid in planning for machine installation. Actual power consumption after installation can be confirmed with the HMC monitoring tools.

3.6.5 IBM Systems Director Active Energy Manager

IBM Systems Director Active Energy Manager[™] (AEM) is an energy management solution building block that returns true control of energy costs to the client. This feature enables management of the actual power consumption and resulting thermal loads that IBM servers place on the data center. It is an industry-leading cornerstone of the IBM energy management framework. In tandem with chip vendors Intel and AMD, and consortium such as Green Grid, AEM advances the IBM initiative to deliver price performance per unit of area.

AEM runs on Windows, Linux on System x, AIX, Linux on IBM System p®, and Linux on z Systems. For more information, see its documentation:

<http://www.ibm.com/systems/software/director/resources.html>

How AEM works

The following list is a brief overview of how AEM works:

- ▶ Hardware, firmware, and systems management software in servers and blades, can take inventory of components.
- ▶ AEM adds power draw-up for each server or blade and tracks that usage over time.
- ▶ When power is constrained, AEM allows power to be allocated on a server-by-server basis. Consider the following information:
 - Be careful that limiting power consumption does not affect performance.
 - Sensors and alerts can warn the user if limiting power to this server could affect performance.
- ▶ Certain data can be gathered from the SNMP API on the HMC:
 - System name, machine type, model, serial number, firmware level
 - Ambient and exhaust temperature
 - Average and peak power (over a 1-minute period)
 - Other limited status and configuration information

3.6.6 Top Exit Power

The z13 support the optional *Top Exit Power* feature. This feature enables installing a radiator (air) cooled z13 on a non-raised floor, when the optional top exit I/O cabling feature is also installed. Water-cooled z13 models cannot be installed on a non-raised floor as top exit support for water cooling systems is not available. On a raised floor, either radiator or water cooling is supported.

3.7 IBM z BladeCenter Extension (zBX) Model 004

z13 introduces the IBM z BladeCenter Extension (zBX) Model 004. The IBM z BladeCenter Extension (zBX) Model 004 continues to support workload optimization and integration. As a stand-alone *node* of an existing ensemble the zBX can house multiple environments that include AIX, Linux on System x, and Windows, supporting a “fit for purpose” application deployment.

The zBX is tested and packaged together at the IBM manufacturing site and shipped as one unit, relieving complex configuration and set up requirements. With a focus on availability, the zBX has hardware redundancy that is built in at various levels: the power infrastructure, rack-mounted network switches, power and switch units in the BladeCenter chassis, and redundant cabling for support and data connections. The IBM z BladeCenter Extension (zBX) Model 004 components are configured, managed, and serviced using a pair of internal 1U rack mounted Support Elements as a node of an ensemble, defined to the ensemble HMC as any other ensemble member.

Although the zBX processors are not z/Architecture PUs, the zBX is handled by z Systems firmware called *IBM z Unified Resource Manager*.

GDPS/PPRC and GDPC/GM support zBX hardware components, providing workload failover for automated multi-site recovery. These capabilities can help facilitate the management of planned and unplanned outages across IBM z13.

3.7.1 IBM blades

zBX Model 004 supports IBM AIX on IBM POWER7, Linux on System x, Microsoft Windows on System x and IBM WebSphere DataPower Integration Appliance X150 for **zEnterprise** on a

blade form factor, which are connected to the z Systems CPCs through virtual LANs supported on a high-speed private network.

IBM BladeCenter PS701 Express blades virtualized by PowerVM Enterprise Edition. The virtual servers in PowerVM run the AIX operating system. PowerVM handles all the access to the hardware resources, providing a Virtual I/O Server (VIOS) function and the ability to create logical partitions. The logical partitions can be either dedicated processor LPARs, which require a minimum of one core per partition, or shared processor LPARs (micro-partitions), which in turn can be as small as 0.1 core per partition.

A select set of IBM BladeCenter HX5 (7873) blades can be used by the zBX. These blades have an integrated hypervisor, and their virtual machines run Linux on System x and Windows Server 2012.

When ordering a zBX Model 004 MES upgrade, a new entitlement record can be acquired by the client. This new entitlement record allows IBM System x blades or IBM POWER7 PS701 to be ordered and added to the zBX, up to the limit of available empty (not used) slots in the zBX existing blade centers.

Unsupported: The addition of new racks or new blade centers cannot be done and are not supported. Also the addition of the *IBM WebSphere DataPower Integration Appliance XI50 for zEnterprise* is not supported.

3.8 Reliability, availability, and serviceability (RAS)

The IBM z Systems family presents numerous enhancements in the *RAS* areas. Focus was given to reducing the planning requirements, while continuing to reduce planned, scheduled, and unscheduled outages. One of the contributors to scheduled outages are LIC Driver updates that are performed in support of new features and functions. Enhanced driver maintenance (EDM) can help reduce the necessity and eventual duration of a scheduled outage. When properly configured, the z13 can concurrently activate a new LIC Driver level. Concurrent activation of the select new LIC Driver level is supported at specifically released synchronization points. However, for certain LIC updates, a concurrent update or upgrade is not possible.

The effects of drawer repair and upgrade actions are minimized on the z13 with enhanced drawer availability (EDA). In a multiple drawer system, a single drawer can be concurrently removed and reinstalled for an upgrade or repair. To ensure that the z13 configuration supports removal of a drawer with minimal affect to the workload, consider the Flexible Memory option (see “Flexible memory” on page 73).

The z13 provides a method to increase memory availability, called redundant array of independent memory¹⁴ (RAIM), where a fully redundant memory system can identify and correct memory errors without stopping. The implementation is similar to the RAID concept used in storage systems for a number of years. See *IBM z13 Technical Guide*, SG24-8251 for a detailed description of the RAS features.

To help prevent outages, improvements in several components of z13 are introduced. These enhancements include changes to the:

¹⁴ Meaney, P.J.; Lastras-Montano, L.A.; Papazova, V.K.; Stephens, E.; Johnson, J.S.; Alves, L.C.; O'Connor, J.A.; Clarke, W.J., “IBM zEnterprise redundant array of independent memory subsystem,” *IBM Journal of Research and Development*, vol.56, no.1.2, pp.4:1,4:11, Jan.-Feb. 2012
doi: 10.1147/JRD.2011.2177106

- physical packaging
- bus structures
- processor cores
- memory and cache hierarchy
- power subsystem
- thermal subsystem
- service subsystem
- integrated sparing.

The z13 central processor complex (CPC) subsystem consists of a horizontal drawers designed as a field replaceable unit (FRU). Connections among the drawers are established using symmetric multi processing (SMP) cables. Each drawer is comprised of two nodes, and each node contains three Processing unit (PU) chips, one system cache (SC) chip, and 10 or 15 DDR3 DIMM slots. With a two-node drawer structure, the z13 design supports system activation with partial-drawer resources in a degraded mode, if necessary. The PU and SC chips are designed as single chip modules (SCMs) and FRUs.

A redundant pair of distributed converter assemblies (DCAs) step down the bulk power and connect to 10 point of load (POL) cards, which provide power conversion and regulation. Two redundant oscillators are connected to the drawers through an isolated backplane. Time domain reflectometry (TDR) techniques are applied to isolate failures on the SMP cables, between chips (PU-PU, PU-SC, and SC-SC), and between the PU chips and DIMMs.

Additional redundancy is designed into new N+1 system control hubs (SCHs) and associated power supplies, and 1U service elements (SEs). Improvements to the z13 I/O infrastructure reliability include better recovery of FICON channels facilitated through forward error correction code (FEC) technology.

An air-cooled configuration features a fully-redundant N+2 radiator pump design that cools the PU chips through a water manifold FRU.

Further RAS enhancements include integrated sparing, error detection and recovery improvements in caches and memory, refreshes to IBM zAware, Flash Express, RoCE, and PCIe coupling, Fibre Channel Protocol support for T10-DIF, a fixed HSA with its size increased to 96 GB on the z13, OSA firmware changes to increase the capability of concurrent maintenance change level (MCL) updates, a new radiator cooling system with N+2 redundancy, new CFCC level, and RMF™ reporting.

z13 continues to support concurrent addition of resources, such as processors or I/O cards to an LPAR to achieve better serviceability. If an additional system assist processor (SAP) is required on a z13 (for example, as a result of a disaster recovery situation), the SAPs can be concurrently added to the CPC configuration.

Concurrently adding CP, zIIP, IFL, and ICF processors to an LPAR is possible. This function is supported by z/VM V5R4¹⁵ and later, and also (with appropriate PTFs) by z/OS and z/VSE V4R3 and later. Previously, proper planning was required to add CP, zAAP, and zIIP to a z/OS LPAR concurrently. Concurrently adding memory to an LPAR is possible. This is supported by z/OS and z/VM.

z13 supports adding Crypto Express features to an LPAR dynamically by changing the cryptographic information in the image profiles. Users can also dynamically delete or move Crypto Express features. This enhancement is supported by z/OS, z/VM, and Linux on z Systems.

¹⁵ z/VM V5R4 is not supported on z13

3.8.1 IBM z Advanced Workload Analysis Reporter (IBM zAware)

Introduced with the zEC12 and also available with the zBC12, the IBM zAware feature is an integrated expert solution that uses sophisticated analytics to help clients identify potential problems and improve overall service levels.

IBM zAware runs analytics in a dedicated Logical Partition (LPAR) and intelligently examines z/OS message logs for potential deviations, or inconsistencies, or variations from the norm, providing out-of-band monitoring and machine learning of operating system health.

IBM zAware can accurately identify system anomalies in minutes. This feature analyzes massive amounts of processor data to identify problematic messages and provides information that can feed other processes or tools. The IBM zAware virtual appliance monitors the z/OS operations log (OPERLOG), which contains all messages that are written to the z/OS console, including application-generated messages. IBM zAware provides a graphical user interface (GUI) for easy drill-down into message anomalies, which can lead to faster problem resolution.

IBM zAware has been enhanced to support Linux on z Systems images running natively or as guests in z/VM, identifying the unusual system behavior by analyzing the syslog.

Statement of Direction:^a IBM intends to deliver IBM z Advanced Workload Analysis Reporter (IBM zAware) support for z/VM. This future release of IBM zAware is intended to help identify unusual behaviors of workloads running on z/VM in order to accelerate problem determination and improve service levels

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

More detail about IBM zAware can be found in *IBM z13 Technical Guide, SG24-8251, IBM zAware Concept Guide*, SG24-8070, and *Advanced Workload Analysis Reporter (IBM zAware)*, SC27-2623.

3.8.2 RAS capability for the SE

Enhancements are made to the Support Element (SE) design for z13. Notebooks that were used on prior generations of zSystem servers have been replaced with rack-mounted 1U System x servers in a redundant configuration on z13. The new, more powerful 1U SEs offer RAS improvements such as ECC memory, redundant physical networks for SE networking requirements, redundant power modules, and better thermal characteristics.

3.8.3 RAS capability for the HMC

Enhancements are made to the HMC designs for z13 as well. New for z13 is an option to order 1U System x servers for traditional and ensemble HMC configurations. This new 1U HMC offers the same RAS improvements as those of the 1U SE. The 1U HMC option is a customer-supplied rack and power consolidation solution that can save space in data centers. The MiniTower design used prior to z13 will still be available.

The Unified Resource Manager is an active part of the ensemble infrastructure. Thus, the HMC has a stateful environment that needs high-availability features to ensure survival of the system in case of an HMC failure.

Each ensemble requires two HMC workstations: a primary, and a backup called alternate. The contents and activities of the primary are updated on the alternate HMC synchronously so that the alternate can take over the activities of the primary should the primary fail. Although the primary HMC can perform the classic HMC activities in addition to the Unified Resource Manager activities, the alternate HMC can only be a backup.

3.8.4 RAS capability for zBX

The zBX was built following the traditional z Systems hardware quality of service (QoS) to include RAS capabilities. The zBX Model 004 provides extended service capability as a member of the ensemble infrastructure. Flexibility and scale-out are improved by eliminating the management coupling between a controlling CPC and the zBX Mod 004. CPC upgrade complexities are also reduced. With the zBX Model 004 configured as an independent node in the ensemble, serviceability on the zBX Model 004 does not affect other CPCs in the ensemble and vice versa. The ensemble HMC provides management and control functions for the zBX Model 004 solution.

Independent of the number of zBX racks installed, the zBX Mod 004 is configured to provide N+1 redundancy. Two SEs with HMC network connectivity are included in the upgrade. Installed only on zBX Model 004's first rack are four Top of Rack (TOR) switches, two per each network (INMN and IEDN). These switches provide N + 1 connectivity for the data network between a CPC and the zBX Model 004, and for the management network used for monitoring and controlling the zBX Model 004 components. The zBX components can be replaced concurrently.

zBX firmware

The testing, delivery, installation, and management of the zBX Model 004 firmware is handled the same way as for the z13. The same z13 processes and controls are used. Any fixes to the zBX Model 004 are downloaded and applied independent of the CPCs in the ensemble. The zBX Model 004 SE will contain only the applicable MCL streams for the zBX configurations. The z13 will not contain any of the blade MCL streams. Most MCLs for the zBX Model 004 are concurrent and their status can be viewed at the ensemble's HMCs.

These and additional features are described in *IBM z13 Technical Guide*, SG24-8251.

3.9 High availability

z Systems platform is renowned for its reliability, availability, and serviceability capabilities, of which Parallel Sysplex is an exponent. Extended availability technology with IBM PowerHA® for AIX is available for blades in the zBX. We describe the z Systems Parallel Sysplex technology and the PowerHA technology.

3.9.1 High availability for z Systems with Parallel Sysplex

Parallel Sysplex technology is a clustering technology for logical and physical servers, allowing highly reliable, redundant, and robust z Systems technology to achieve availability that is near-continuous. Both hardware and software tightly cooperate to achieve this result.

A Parallel Sysplex has the following minimum components:

► Coupling facility (CF)

This is the cluster center. It can be implemented either as an LPAR of a stand-alone z Systems CPC or as an additional LPAR of a z Systems CPC where other loads are running. Processor units that are characterized as either CPs or ICFs can be configured to this LPAR. ICFs are often used because they do not incur any software license charges. Two CFs are recommended for availability.

► Coupling Facility Control Code (CFCC)

This IBM Licensed Internal Code is both the operating system and the application that runs in the CF. No other code runs in the CF. The code is used to create and maintain the structures which are exploited under z/OS by software components such as z/OS itself, DB2 for z/OS, WebSphere MQ, among others.

CFCC can also run in a z/VM virtual machine (as a z/VM guest system). In fact, a complete Sysplex can be set up under z/VM, allowing, for instance, testing and operations training. This setup is not recommended for production environments.

► Coupling links

These are high-speed links that connect the several system images (each running in its own logical partition) that participate in the Parallel Sysplex. At least two connections between each physical server and the CF must exist. When all of the system images belong to the same physical server, internal coupling links are used.

On the software side, the z/OS operating system uses the hardware components to create a Parallel Sysplex. One example of z/OS and CF collaboration is the System-managed CF structure duplexing, which provides a general-purpose, hardware-assisted, easy-to-exploit mechanism for duplexing structure data held in CFs. This function provides a robust recovery mechanism for failures (such as loss of a single structure on CF or loss of connectivity to a single CF). The recovery is done through rapid failover to the other structure instance of the duplex pair.

Clients that are interested in deploying system-managed CF structure duplexing can read the technical paper *System-Managed CF Structure Duplexing*, ZSW01975USEN, which can be accessed by clicking **Learn more** on the Parallel Sysplex website:

<http://www.ibm.com/systems/z/ps0/index.html>

z/TPF: z/TPF can also use the CF hardware components. However, the term *Sysplex* exclusively applies to z/OS usage of the CF.

Normally, two or more z/OS images are clustered to create a Parallel Sysplex, although it is possible to have a configuration setting with a single image, called a *monoplex*. Multiple clusters can span several z Systems CPCs, although a specific image (logical partition) can belong to only one Parallel Sysplex.

A z/OS Parallel Sysplex implements shared-all access to data. This is facilitated by System z I/O virtualization capabilities such as the *multiple image facility (MIF)*. MIF allows several logical partitions to share I/O paths in a secure way, maximizing use and greatly simplifying the configuration and connectivity.

In short, a Parallel Sysplex comprises one or more z/OS operating system images that are coupled through one or more coupling facilities. A properly configured Parallel Sysplex cluster is designed to maximize availability *at the application level*. Rather than a quick recovery of a failure, the Parallel Sysplex design objective is *zero failure*.

The major characteristics of a Parallel Sysplex include the following features:

▶ Data sharing with integrity

The CF is key to the implementation of a share-all access to data. Every z/OS system image has access to all the data. Subsystems in z/OS declare resources to the CF. The CF accepts and manages lock and unlock requests on those resources, guaranteeing data integrity. A duplicate CF further enhances the availability. Key users of the data sharing capability are DB2, WebSphere MQ, WebSphere ESB, IMS, and CICS. Because these are major infrastructure components, applications that use them inherently benefit from sysplex characteristics. For instance, many large SAP implementations have the database component on DB2 for z/OS, in a Parallel Sysplex.

▶ Continuous (application) availability

Changes, such as software upgrades and patches, can be introduced one image at a time, while the remaining images continue to process work. For more details, see *Improving z/OS Application Availability by Managing Planned Outages*, SG24-8178.

▶ High capacity

Parallel Sysplex scales from two to 32 images. Remember that each image can have from one to 128 (z/OS V2R1) processor units. CF scalability is near-linear. This structure contrasts with other forms of clustering that employ n-to-n messaging, which leads to rapidly degrading performance with a growing number of nodes.

▶ Dynamic workload balancing

Viewed as a single logical resource, work can be directed to any of the Parallel Sysplex cluster operating system images where capacity is available.

▶ Systems management

This architecture provides the infrastructure to satisfy a client requirement for continuous availability, while enabling techniques for achieving simplified systems management consistent with this requirement.

▶ Resource sharing

A number of base z/OS components use CF shared storage. This usage enables the sharing of physical resources with significant improvements in cost, performance, and simplified systems management.

▶ Single system image

The collection of system images in the Parallel Sysplex is displayed as a single entity to the operator, user, database administrator, and so on. A single system image ensures reduced complexity from both operational and definition perspectives.

▶ N-2 support

Multiple hardware generations (normally three, which are the current and the two previous ones) are supported in the same Parallel Sysplex. This configuration provides for a gradual evolution of the systems in the Sysplex, without forcing changing all simultaneously. Similarly, software support for multiple releases or versions is supported.

Figure 3-3 illustrates the components of a Parallel Sysplex as implemented within the System z architecture. The diagram shows one of many possible Parallel Sysplex configurations.

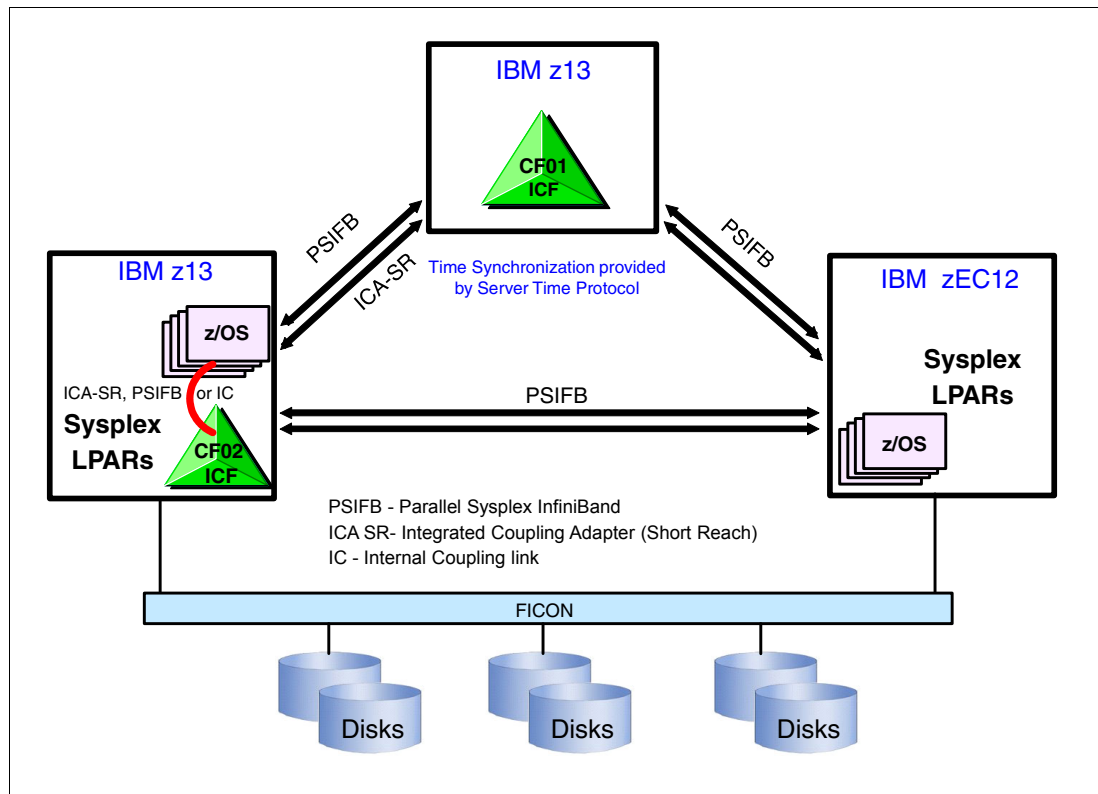


Figure 3-3 Sysplex hardware overview

Figure 3-3 shows a z13 system that contains multiple z/OS sysplex partitions and an internal coupling facility (CF02), a z13 server containing a stand-alone CF (CF01), and a zEC12 containing multiple z/OS sysplex partitions. STP over coupling links provides time synchronization to all servers. Appropriate CF link technology (1x IFB, 12x IFB, or ICA-SR) selection depends on server configuration and how distant they are physically located. ICA-SR links can only be used from z13 to z13, within a short distance.

3.9.2 PowerHA in zBX environment

High availability for applications running on AIX is provided by the IBM PowerHA SystemMirror® for AIX (formerly known as IBM HACMP™¹⁶). PowerHA is easy to configure (menu-driven) and helps define and manage resources (required by applications) running on AIX by providing service/application continuity through platform resources and application monitoring, and automated actions (start/manage/monitor/restart/move/stop).

Terminology: Resource movement and application restart on the second server is known as *failover*.

Automating the failover process speeds up recovery and allows for unattended operations, thus providing improved application availability.

A PowerHA configuration or *cluster* consists of two or more servers¹⁷ (up to 32) that have their resources managed by PowerHA cluster services to provide automated service recovery

for the applications managed. Servers can have physical or virtual I/O resources, or a combination of both.

PowerHA performs the following functions at the cluster level:

- ▶ Manage and monitor operating systems and hardware resources.
- ▶ Manage and monitor application processes.
- ▶ Manage and monitor network resources.
- ▶ Automate applications (start, stop, restart, move).

The virtual servers that are defined and managed in zBX use only virtual I/O resources. PowerHA can manage both physical and virtual I/O resources (virtual storage and virtual network interface cards).

PowerHA can be configured to perform automated service recovery for the applications that run in virtual servers that are deployed in zBX. PowerHA automates application failover from one virtual server in an IBM POWER® processor-based blade to another virtual server in a different POWER processor-based blade that has a similar configuration.

Failover protects service (masks service interruption) in case of unplanned or planned (scheduled) service interruption. During failover, users might experience a short service interruption while resources are configured by PowerHA on the new virtual server.

The PowerHA configuration for the zBX environment is similar to standard POWER environments, with the particularity that it uses only virtual I/O resources. Currently, PowerHA for zBX support is limited to failover inside the same zBX.

¹⁶ High Availability Cluster Multi-Processing

¹⁷ Servers can be also virtual servers. One server equals one instance of the AIX Operating System.

Figure 3-4 shows a typical PowerHA cluster.

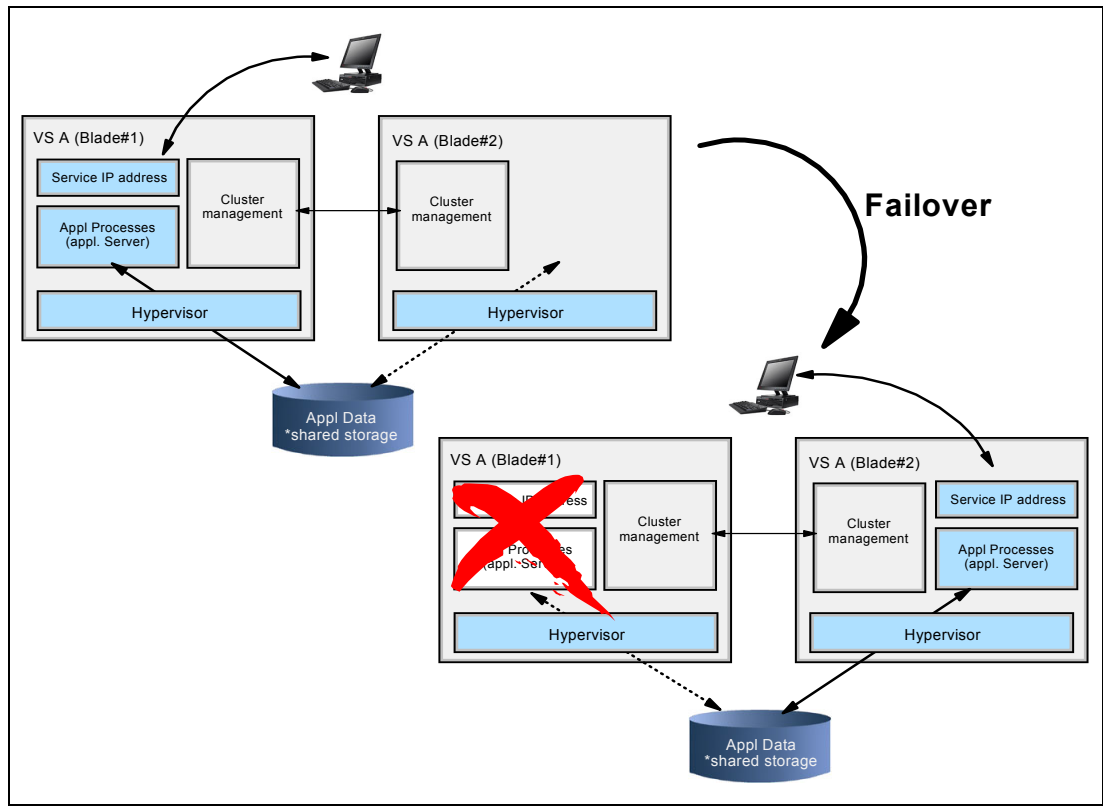


Figure 3-4 Typical PowerHA cluster diagram

For more information about IBM PowerHA SystemMirror, see the following web page:

<http://www.ibm.com/support/knowledgecenter/SSPHQG/welcome>

3.10 IBM z Systems and emerging paradigms

Having reviewed the most recent and important characteristics of z Systems, we would like to conclude this discussion with some observations on the role z13 can play in today's leading IT initiatives.

We are witnessing a transformation of the interaction between users and systems, increasingly based on mobile devices, and instrumented devices ("the Internet of things"). This front office transformation requires highly responsive and dynamic transaction systems, and demands very high security. As discussed in 3.1.4, "z Systems -based clouds" on page 66, and evidenced by the descriptions of z13, these systems can answer the infrastructure hardware requirements, whether I/O bandwidth, computing dynamic scalability, or security.

In addition, software requirements are covered. We note that several software licensing offerings are available on z Systems to cater to various environments and workloads. In particular, Linux on z Systems closely follows the distributed paradigm; see Appendix A, "Software licensing" on page 161.

Several transactional servers are available on the Linux on z and z/OS environments, which can be used by mobile applications such as those developed with the state-of-the-art IBM

Worklight software. Those applications can benefit from the unmatched reliability, availability and, serviceability (RAS) features offered by a z Systems.

IBM DB2 Analytics Accelerator, which can transparently, that is without application modification, benefit from the radical acceleration of very complex queries such as those used by Business Intelligence and Data Analytics, enabling their insertion into on-line applications.

Finally, but no less important, the need for a coherent security landscape across the enterprise is increasingly being recognized. The z13 have specialized offerings such as the Enterprise Key Management Foundation enabling their security features, such as the Crypto Express and secure key management, to be used by the larger enterprise.

For further information about how to benefit from IBM z13, see Chapter 5, “A system of insight for digital business” on page 137.



4

Operating system support

This chapter contains operating system requirements and support considerations for the IBM z13 and its features.

This chapter describes the following topics:

- ▶ Software support summary
- ▶ Support by operating system
- ▶ References
- ▶ z/OS considerations
- ▶ Coupling facility and CFCC considerations
- ▶ Input/output configuration program (IOCP) considerations
- ▶ GDPS Virtual Appliance

Support of the IBM z13 functions depends on the operating system version and release. This information is subject to change. Therefore, for the most current information, see the *Preventive Service Planning (PSP)* bucket for 2964DEVICE.

4.1 Software support summary

The software portfolio for the IBM z13 includes a many various operating systems and middleware that support the most recent and significant technologies. Continuing the mainframe-rich tradition, five major operating systems are supported:

- ▶ z/OS
- ▶ z/VM
- ▶ z/VSE
- ▶ z/TPF
- ▶ Linux on z Systems

For software supported on the zBX Model 004, see 4.3, “Software support for zBX Model 004” on page 130.

Operating systems summary

Table 4-1 summarizes the current and minimum operating system levels that are required to support the z13. Operating system levels that are no longer in service are not covered in this publication. These older levels can provide support for certain features.

Table 4-1 z13 operating system requirements

Operating system	ESA/390 (31-bit mode)	z/Architecture (64-bit mode)	End of service	Notes
z/OS V2R1	No	Yes	September 2018 ^a	See the z/OS, z/VM, z/VSE, and z/TPF subsets of the 2964DEVICE Preventive Service Planning (PSP) buckets before installing z13.
z/OS V1R13	No	Yes	September 2016 ^a	
z/OS V1R12	No	Yes	September 2014 ^b	
z/VM V6R3 ^c	No	Yes	April 2017 ^a	
z/VM V6R2 ^c	No	Yes	December 2016 ^a	
z/VSE V5R2 ^d	No	Yes	Not announced	
z/VSE V5R1 ^d	No	Yes	June 2016 ^a	
z/TPF V1R1	Yes	Yes	Not announced	
Linux on z Systems	No ^e	See Table 4-6 on page 127	Support information is available for SUSE and Red Hat ^f	

a. Planned date. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these Statements of Direction is at the relying party's sole risk and will not create liability or obligation for IBM.

b. z/OS V1R12 with required maintenance (compatibility support only) and extended support agreement

c. z/VM V6R2 and V6R3 require an architectural level set exclusive to z10 and successors.

d. z/VSE V5 requires an architectural level set exclusive to z9 and successors.

e. 64-bit distributions include a 31-bit emulation layer to run 31-bit software products.

f. SUSE: <http://www.suse.com/support/>

Red Hat: <http://www.redhat.com/security/updates/errata/>

PTFs and PSP buckets: The use of several features depends on a particular operating system. In all cases, program temporary fixes (PTF) might be necessary with the operating system level indicated.

Preventive Service Planning (PSP) buckets are continuously updated and are reviewed regularly when planning for installation of a new system. They contain the latest information about installation, hardware and software service levels, service recommendations, and cross-product dependencies.

For Linux on z Systems distributions, consult the distributor's support information.

Middleware

The following middleware offerings are for the IBM z13 environments:

- ▶ Transaction processing
 - WebSphere Application Server and WebSphere Extended Deployment
 - CICS Transaction Server
 - CICS Transaction Gateway
 - IMS DB and IMS DC
 - IMS Connect
- ▶ Application integration and connectivity
 - WebSphere Message Broker
 - WebSphere MQ
 - WebSphere ESB
- ▶ Process integration
 - WebSphere Process Server
 - WebSphere MQ Workflow
 - WebSphere Business Integration Server
- ▶ Database
 - IBM DB2 for z/OS
 - IBM DB2 for Linux
 - IBM DB2 Connect™

Service Management

The IBM Tivoli® brand has a large product set that includes the following offerings:

- ▶ IBM Tivoli Service Management Center
- ▶ IBM Tivoli Information Management for z/OS
- ▶ IBM Tivoli Workload Scheduler
- ▶ IBM Tivoli OMEGAMON® XE
- ▶ IBM Tivoli System Automation

Security

A highly secure z Systems environment can be implemented at various levels by using the following products:

- ▶ IBM Security zSecure™ suite
- ▶ The Security Server component of z/OS and z/VM; includes Resource Access Control Facility (IBM RACF®)
- ▶ IBM Tivoli Directory Server for z/OS
- ▶ z/OS Communications Server and Policy Agent; for policy-based network security

- ▶ The z/OS Cryptographic Services component of z/OS; includes the Integrated Cryptographic Service Facility (ICSF)

Application development and languages

Several programming languages are available for the IBM z13 environment. Because the Linux environment is similar to Linux on other servers, we focus on the z/OS environment.

In addition to the traditional COBOL, PL/I, Fortran, and Assembler languages, z Systems support C, C++, and Java (including Java Platform, Enterprise Edition and batch environments) programming languages as well.

Development can be conducted by using the latest software engineering technologies and advanced *integrated development environments (IDE)*. The extensive tool set uses a workstation environment for development and testing, with final testing and deployment performed on z/OS. Application development tools, many of which have components that are based on the Eclipse platform, are provided through the following offerings:

- ▶ IBM Rational® Application Developer for WebSphere
- ▶ IBM Rational Developer for z Systems
- ▶ IBM WebSphere Developer for z Systems
- ▶ IBM Rational Rose® product line
- ▶ IBM Rational Software Architect and Software Modeler

For more information about software for z Systems, see this web page:

<http://www.ibm.com/software/products/us/en/category/sysz?pgel=1nav>

We cannot emphasize enough the importance of using the most recent versions of the compilers. The compilers enable the use of the latest technologies that are implemented on the system and take advantage of the performance benefits that are introduced. Examples of benefits include new cache structures, new machine instructions, and instruction execution enhancements.

For example, z13 processor introduces *single-instruction, multiple-data (SIMD) instruction set*, which uses the enhanced superscalar z13 core to process a large number of operands (vector) through a single instruction, allowing the development of smaller and optimized codes to improve efficiency of complex mathematical models and vector processing. This feature will be fully used by z/OS V2R1 operating system and several compilers have built-in functions for SIMD.

Statement of Direction^a (SIMD Support - z/VM support SIMD): In a future deliverable, IBM intends to deliver support to enable z/VM guests to used the Vector Facility for z/Architecture (SIMD).

- All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

IBM compilers

Each new version of IBM z/OS compilers (Enterprise COBOL, Enterprise PL/I, XL C/C++) underscores the continuing IBM commitment to the COBOL, PL/I, and C/C++ programming languages on the z/OS platform.

Enterprise COBOL

The latest version of Enterprise COBOL uses the latest z/Architecture and performance optimization, enhanced XML parsing support, and capability of programming with Unicode, and supports Java 7 SDKs for Java interoperability.

Enterprise PL/I

The latest version of Enterprise PL/I provides web interoperability, which includes web services, XML parsers, and Java Platform, Enterprise Edition (Java EE). The compiler also includes the expanded support for UTF-16.

z/OS XL C/C++

The z/OS XL C/C++ uses the latest z/Architecture, including the latest z13 servers. It enables developing high performance oriented applications, through the services provided by Language Environment and Runtime Library extension base elements, and works in concert with z/OS problem determination tools.

Application development tools

IBM Enterprise COBOL and Enterprise PL/I support are strategic components (separately orderable products) for IBM Rational Developer for IBM z Systems software. These features provide a robust, integrated development environment (IDE) for COBOL and PL/I and connecting web services, Java Platform, Enterprise Edition (Java EE) applications, and traditional business processes.

z/OS XL C/C++ programmers can also tap into Rational Developer for z Systems to boost their productivity by easily editing, compiling, and debugging z/OS XL C and XL C++ applications right from their workstation.

4.2 Support by operating system

In this section, we list the support by in-service operating systems of selected functions of the z13. For a detailed description of z13 and its features, see the *IBM z13 Technical Guide*, SG24-8251. For an in-depth description of all I/O features, see the *IBM System z Connectivity Handbook*, SG24-5444.

4.2.1 z/OS

z/OS Version 1 Release 12 is the earliest in-service release that supports the z13, with required maintenance (compatibility support only) and extended support agreement.

Table 4-2 summarizes the support requirements of selected z13 functions for the currently supported z/OS releases. In the table, Y (yes) means the function is supported, N (no) means the function is not supported.

Table 4-2 z/OS support summary

Function	V2R1	V1R13	V1R12
z13	Y	Y ^a	Y ^a
Support of 141 PUs by a single system image	Y	N ^b	N ^b
Support of IBM zAware	Y	Y	N
IBM z <i>Integrated Information Processor</i> (zIIP)	Y ^c	Y ^c	Y ^c
4 TB memory per LPAR exploitation	Y ^d	N	N
Pageable 1MB Large page support	Y	Y ^e	N
Decimal Floating point support for packed decimal conversions	Y	Y	N
2 GB Large Page Support	Y	Y	N
Support up to 85 LPARs	Y ^f	Y ^g	N
Support for six logical channel sub systems (CSS)	Y	Y	N
Support for 4 sub channel sets per CSS	Y	Y	N
Single-instruction, multiple-data (SIMD) support	Y	N	N
Two way simultaneous multi threading support (SMT-2)	Y ^g	N	N
HiperDispatch	Y	Y	Y
EP 11 cryptography support	Y	Y ^h	Y ^f
Common Cryptographic architecture >16 Domain Support	Y	Y ⁱ	N
CPACF	Y	Y	Y
CPACF AES-128, AES-192, and AES-256	Y	Y	Y
CPACF SHA-1, SHA-224, SHA-256, SHA-384, SHA-512	Y	Y	Y
CPACF protected key	Y	Y	Y
Crypto Express5S	Y ^j	Y ^h	Y ^h
Crypto Express4S	Y	Y	Y
Secure IBM Enterprise PKCS #11 (EP11) coprocessor mode	Y	Y	Y
Elliptic Curve Cryptography (ECC)	Y	Y	Y
Flash Express	Y	Y	N
zEDC Express	Y	N ^k	N ^k
Shared (SR-IOV) 10GbE RoCE Express	Y	N	N
zHPF (High Performance FICON) Extended Distance II	Y	Y	N
FICON Dynamic Routing	Y	Y	N
Fabric Priority for an I/O request	Y	Y	N
FICON Express 16S	Y	Y	Y
FICON Express8S	Y	Y	Y
FICON Express8 ^l	Y	Y	Y

Function	V2R1	V1R13	V1R12
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSD	Y	Y	Y
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSX	Y	Y	Y
OSA-Express5S Gigabit Ethernet LX and SX CHPID type OSD (using two ports per CHPID)	Y	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSC (using two ports per CHPID)	Y	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSD (using two ports per CHPID)	Y	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSE(using one port per CHPID)	Y	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSM	Y	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSN (using one port per CHPID)	Y	Y	Y
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSD	Y	Y	Y
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSX	Y	Y	Y
OSA-Express4S Gigabit Ethernet LX and SX CHPID type OSD (using two ports)	Y	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSC (using two ports per CHPID)	Y	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSD (using two ports per CHPID)	Y	Y	Y

Function	V2R1	V1R13	V1R12
OSA-Express4S 1000BASE-T CHPID type OSE (using two ports per CHPID)	Y	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSM (using two ports per CHPID)	Y	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSN (using two ports per CHPID)	Y	Y	Y
Coupling using InfiniBand CHPID type CIB	Y	Y	Y
InfiniBand coupling links (12x IFB-SDR or 12x IFB-DDR) at a distance of 150 m	Y	Y	Y
InfiniBand coupling links (1x IFB-SDR or 1x IFB-DDR) at an unrepeated distance of 10 km	Y	Y	Y
IBM Integrated Coupling Adapter support (ICA)	Y	Y	Y
Server Time Protocol	Y	Y	Y
CFCC Level 20	Y	Y	Y
Assembler instruction mnemonics	Y	Y	Y
C/C++ exploitation of hardware instructions	Y ^m	N	N
CPU measurement facility	Y	Y	Y

- a. Service is required for support of z13.
- b. z/OS 1.13 and z/OS 1.12 support 100 cores per LPAR
- c. On an upgrade from z196 or zEC12, installed zAAPs are converted to zIIPs by default. zAAPs are no more supported on z13.
- d. 4TB Support on z/oS LPAR with z/OS V2R1 is SoD
- e. A web deliverable is required for Pageable 1M Large Page Support.
- f. This supported only if the LPARs are running z/OS V2R1. IF z/OS V1R12 is running in any of the LPARs then only 60 LPARs are supported
- g. Only for zIIP eligible workload
- h. With Cryptographic support for z/OS V1R12-V1R13 web deliverable
- i. With PTF and RSM enablement offering
- j. Enhanced cryptographic support for z/OS V1R13-z/OS V2R1 web deliverable
- k. Software decompression only.
- l. Carry forward from zEC12/z196 only
- m. z13 Exploitation Support for V2R1 XL C/C++ web deliverable

4.2.2 z/VM

At general availability, z/VM V6R3 and z/VM V6R2 provide compatibility support with use of some new z13 functions.

Statements of Direction:^a

- ▶ *Removal of support for Expanded Storage (XSTORE):* z/VM 6.3 is the last z/VM release to support XSTORE for either host or guest usage. The IBM z13 will be the last high-end server to support XSTORE.
- ▶ *Stabilization of z/VM V6.2 support:* The IBM z13 server family is planned to be the last z Systems server supported by z/VM V6.2 and the last z Systems server that will be supported where z/VM V6.2 is running as a guest (second level). This is in conjunction with the statement of direction that the IBM z13 server family will be the last to support ESA/390 architecture mode, which z/VM V6.2 requires. z/VM V6.2 will continue to be supported until December 31, 2016, as announced in announcement letter # 914-012.
- ▶ *Product Delivery of z/VM on DVD/Electronic only:* z/VM 6.3 will be the last release of z/VM that will be available on tape. Subsequent releases will be available on DVD or electronically
- ▶ *Enhanced RACF password encryption algorithm for z/VM:* In a future deliverable an enhanced RACF/VM password encryption algorithm is planned. This support will be designed to provide improved cryptographic strength using AES-based encryption in RACF/VM password algorithm processing. This planned design is intended to provide better protection for encrypted RACF password data in the event that a copy of RACF database becomes inadvertently accessible
- ▶ *z/VM support for Single Instruction Multiple Data (SIMD):* In a future deliverable IBM intends to deliver support to enable z/VM guests to exploit the Vector Facility for z/Architecture (SIMD)

a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

Table 4-3 summarizes the support requirements of selected functions for the currently supported z/VM releases. In the table, Y (yes) means the function is supported, N (no) means the function is not supported.

Important: Any new functions listed as Y for z/VM 6.2 and 6.3 will require service (for both compatibility and exploitation).

Table 4-3 z/VM support summary

Function	V6R3	V6R2
z13	Y	Y
Support of up to 64 PUs for single thread	Y ^a	N ^b
zIIP ^c	Y	Y
1 TB real memory support	Y	N ^d
Large page support	N	N
Decimal floating point support for packed decimal numbers	Y	Y

Function	V6R3	V6R2
Two way simultaneous multithreading (SMT) Support	Y	N
Single-instruction, multiple-data (SIMD) support	N	N
CPU measurement facility counter	Y	Y
HiperDispatch	Y	N
CPACF	Y	Y
CPACF AES-128, AES-192, and AES-256	Y	Y
CPACF SHA-1, SHA-224, SHA-256, SHA-384, SHA-512	Y	Y
CPACF protected key	Y	Y
Crypto Express4S	Y	Y
Crypto Express5S	Y	Y
Secure IBM Enterprise PKCS #11 (EP11) coprocessor mode	Y	Y ^e
Elliptic Curve Cryptography (ECC)	Y	Y ^e
Common Cryptographic Architecture (CCA) > 16 Domain Support	Y	Y
EP 11 Support	Y	Y
Flash Express	N	N
zEDC Express ^f	Y ^g	N
Shared 10GbE RoCE (SR-IOV) support ^h	Y	N
High Performance FICON (zHPF) Extended distance ⁱ	N	N
FICON Express8S	Y	Y
FICON Express8	Y	Y
FICON Express16S	Y	Y
OSA-Express QDIO data connection isolation for z/VM environments	Y	Y
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSD	Y	Y
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSX	N ^j	N ^j
OSA-Express5S Gigabit Ethernet LX and SX CHPID type OSD (using two ports)	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSC (using two ports per CHPID)	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSD (using two ports per CHPID)	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSE ^k (using one or two ports per CHPID)	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSM	N ^j	N ^j

Function	V6R3	V6R2
OSA-Express5S 1000BASE-T CHPID type OSN (using two ports per CHPID)	Y	Y
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSD	Y	Y
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSX	Y	Y
OSA-Express4S Gigabit Ethernet LX and SX CHPID type OSD (using two ports)	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSC (using two ports per CHPID)	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSD (using two ports per CHPID)	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSE (using two ports per CHPID)	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSM	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSN (using two ports per CHPID)	Y	Y
Multi-Vswitch Link Aggregation	Y	N
Dynamic I/O support for InfiniBand CHPIDs	Y	Y
InfiniBand coupling links (12x IFB-SDR or 12x IFB-DDR) at a distance of 150 m	Y	Y
InfiniBand coupling links (1x IFB-SDR or 1x IFB-DDR) at an unrepeated distance of 10 km	Y	Y
IBM Integrated Coupling Adapter support (ICA)	Y	Y
Dynamic I/O support for ICA CHPIDs	Y	Y
CFCC Level 20	Y	Y

- a. 64 cores support without multi-threading and 32 cores with multithreading-2
- b. Supports only 32 PU in compatibility mode in z13
- c. zAAPs not available on z13
- d. Real memory limit is 256 GB
- e. Service is required.
- f. Minimum firmware bundle level 21
- g. For z/OS guest support available via APAR
- h. z/OS APAR required for exploiting SR-IOV for running z/OS as a guest in z/VM
- i. zHPF(High Performance FICON) support available. But no support for extended distance.
>64KB distance over 100 KM distance normally used by HyperSwap functionalities
- j. Only dynamic I/O when z/VM is the controlling LPAR.
- k. A CHPID Type OSE supports SNA (LLC2) and IP connectivity over Ethernet (802.3 or DIX V2).

z/VM logical partitions: IBM z13 CPs and IFLs have increased capacity over that of their predecessors. Therefore, we suggest that the capacity of z/VM logical partitions and of any guests, in terms of the *number* of IFLs and CPs (real or virtual), be reviewed and adjusted to achieve the required capacity. Virtual machine shares might also need to be adjusted.

4.2.3 z/VSE

Table 4-4 summarizes the support requirements of selected z13 functions for the currently supported z/VSE releases. In the table, Y (yes) means the function is supported, N (no) means the function is not supported.

Table 4-4 z/VSE support summary

Function	V5R2 ^a	V5R1 ^a
z13	Y	Y
Support for up to ten CPs ^b	Y	Y
Large page support for data spaces (2GB)	Y	Y
Simultaneous multithreading 2 support	N	N
Single-instruction, multiple-data (SIMD) support	N	N
CPACF	Y	Y
CPACF AES-128, AES-192, and AES-256	Y	Y
CPACF SHA-1, SHA-224, SHA-256, SHA-384, SHA-512	Y	Y
CPACF protected key	N	N
Common Cryptographic Architecture (CCA) > 16 Domain Support ^c	Y ^d	Y ^d
Crypto Express4S Toleration	Y	Y
Secure IBM Enterprise PKCS #11 (EP11) coprocessor mode	N	N
Crypto Express5S toleration	Y ^d	Y ^d
Elliptic Curve Cryptography (ECC)	N	N
FICON Express8S	Y	Y
FICON Express8	Y	Y
FICON Express16S	Y	Y
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSD	Y	Y
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSX	Y	Y
OSA-Express5S Gigabit Ethernet LX and SX CHPID type OSD (using two ports)	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSC (using two ports per CHPID)	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSD (using two ports per CHPID)	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSE ^e (using two ports per CHPID)	Y	Y
OSA-Express5S 1000BASE-T CHPID type OSM	N	N
OSA-Express5S 1000BASE-T CHPID type OSN ^f (using two ports per CHPID)	Y	Y

Function	V5R2 ^a	V5R1 ^a
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSD	Y	Y
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSX	Y	Y
OSA-Express4S Gigabit Ethernet LX and SX CHPID type OSD (using two ports)	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSC (using two ports per CHPID)	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSD (using two ports per CHPID)	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSE ^e (using two ports per CHPID)	Y	Y
OSA-Express4S 1000BASE-T CHPID type OSM	N	N
OSA-Express4S 1000BASE-T CHPID type OSN ^f (using two ports per CHPID)	Y	Y
HiperSockets Completion Queue for Linux Fast Path function in LPAR	Y	Y

- a. z/VSE V5 executes in z/Architecture mode and supports 64-bit real and 64-bit virtual memory addressing. PTF support required for exploitation and compatibility
- b. z/VSE Turbo Dispatcher supports up to four CPs and tolerates up to 10-way LPARs
- c. Coprocessor and Accelerator mode
- d. Support available with PTF
- e. A CHPID Type OSE supports both SNA (LLC2) and IP connectivity over Ethernet (802.3 or DIX V2).
- f. One port is configured for OSN. The other port is unavailable.

4.2.4 z/TPF

Table 4-5 summarizes the support requirements of selected z13 functions for the currently supported z/TPF release. In the table, Y (yes) means the function is supported, N (no) means the function is not supported.

Table 4-5 TPF and z/TPF support summary

Function	z/TPF V1R1 ^a
z13	Y
Support for up to 141 PUs	Y
4 TB Real Storage Support	Y
CPACF	Y
CPACF AES-128, AES-192, and AES-256	Y ^b
CPACF SHA-1, SHA-224, SHA-256	Y ^c
CPACF protected key	N
Crypto Express4S ^d	Y ^e
Common Cryptographic Architecture (CCA) up to 85 Domain Support	Y

Function	z/TPF V1R1 ^a
Secure IBM Enterprise PKCS #11 (EP11) coprocessor mode	N
Crypto Express5S	Y
Elliptic Curve Cryptography (ECC)	N
FICON Express16S	Y
FICON Express8S	Y
FICON Express8	Y
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSD	Y
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSX	N
OSA-Express5S Gigabit Ethernet LX and SX CHPID type OSD (using two ports)	Y
OSA-Express5S 1000BASE-T CHPID type OSC (using two ports per CHPID)	N
OSA-Express5S 1000BASE-T CHPID type OSD (using two ports per CHPID)	Y
OSA-Express5S 1000BASE-T CHPID type OSE(using two ports per CHPID)	N
OSA-Express5S 1000BASE-T CHPID type OSM	N
OSA-Express5S 1000BASE-T CHPID type OSN(using two ports per CHPID)	Y
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSD	Y
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSX	N
OSA-Express4S Gigabit Ethernet LX and SX CHPID types OSD (using two ports)	Y
OSA-Express4S 1000BASE-T CHPID type OSC (using one or two ports per CHPID)	N
OSA-Express4S 1000BASE-T CHPID type OSD (using two ports per CHPID)	Y
OSA-Express4S 1000BASE-T CHPID type OSE(using one or two ports per CHPID)	N
OSA-Express4S 1000BASE-T CHPID type OSM	N
OSA-Express4S 1000BASE-T CHPID type OSN (using one or two ports per CHPID)	Y
Coupling over InfiniBand CHPID type CIB	Y ^f

Function	z/TPF V1R1 ^a
IBM Integrated Coupling Adapter (ICA)	Y
CFCC Level 20	Y

- a. PTF support required
- b. Supports only AES-128 and AES-256.
- c. Supports only SHA-1 and SHA-256.
- d. Service is required.
- e. Accelerator mode only.
- f. Compatibility is supported.

4.2.5 Linux on z Systems

Linux on z Systems distributions are built separately for the 31-bit and 64-bit addressing modes of the z/Architecture. The newer distribution versions are built only for 64-bit. You can run 31-bit applications in the 31-bit emulation layer on a 64-bit Linux on z Systems distribution.

None of the current versions of Linux on z Systems distributions (SUSE SLES 11 and SLES 12; Red Hat RHEL 5, RHEL 6 and RHEL 7) require toleration support; therefore, any release of these distributions can run on the z13.

Table 4-6 lists the most recent service levels of the current SUSE and Red Hat supported to run on IBM z13.

Table 4-6 Current Linux on z Systems distributions, by z/Architecture mode

Linux distribution	z/Architecture (64-bit mode)
SUSE SLES 11	Yes
SUSE SLES 12	Yes
Red Hat RHEL 7	Yes
Red Hat RHEL 6	Yes

Table 4-7 lists selected z13 features, showing the minimum level of SUSE and Red Hat distributions that support each feature.

Table 4-7 Linux on z Systems support summary

Function	SUSE	Red Hat
z13	SLES 12 SLES 11	RHEL 7 RHEL 6
Maximum number of CPs or IFLs	64/256 ^a	64/256 ^a
Large page support	SLES 12 SLES 11	RHEL 7 RHEL 6
Decimal floating point for Packed decimal numbers	No	No
CPACF	SLES 12 SLES 11	RHEL 7 RHEL 6
CPACF AES-128, AES-192, and AES-256	SLES 12 SLES 11	RHEL 7 RHEL 6

Function	SUSE	Red Hat
CPACF SHA-1, SHA-224, SHA-256, SHA-384, SHA-512	SLES 12 SLES 11	RHEL 7 RHEL 6
CPACF protected key	Yes ^b	Yes ^b
Crypto Express4S	SLES 12 SLES 11 SP3	RHEL 7 RHEL 6.6
Secure IBM Enterprise PKCS #11 (EP11) coprocessor mode	SLES 11,12 ^c	RHEL 6,7 ^c
Crypto Express5S ^d	SLES 12 SLES 11 SP3	RHEL 7 RHEL 6.6
Elliptic Curve Cryptography (ECC)	Yes ^b	Yes ^b
Common Cryptographic Architecture (CCA) > 16 domain support	SLES 12 SLES 11 SP3	RHEL 7 RHEL 6.6
HiperSockets Completion Queue	SLES 12 SLES 11	RHEL 7 RHEL 6
HiperSockets Virtual Switch Bridge	SLES 12 SLES 11	RHEL 7 RHEL 6
HiperSockets Layer 2 support	SLES 12 SLES 11	RHEL 7 RHEL 6
Shared 10GbE RoCE (SR-IOV) Express	No	No
High Performance FICON (zHPF) extended distance	No	No
Simultaneous Multi-threading 2 Support (SMT-2) support	SLES 12	RHEL 7
single-instruction, multiple-data (SIMD) support	SLES 12	RHEL 7
zFlash Express	SLES 12 SLES 11 SP3	RHEL 7 RHEL 6.4
zEDC Support	No	No
IBM zAware Support	SLES 12 SLES 11	RHEL 7 RHEL 6
FICON Express8S CHPID type FC and FCP	SLES 12 SLES 11	RHEL 7 RHEL 6
FICON Express8, CHPID types FC and FCP	SLES 12 SLES 11	RHEL 7 RHEL 6
FICON Express16S CHPID type FC and FCP	SLES 12 SLES 11	RHEL 7 RHEL 6
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSD	SLES 12 SLES 11	RHEL 7 RHEL 6
OSA-Express5S 10 Gigabit Ethernet LR and SR CHPID type OSX	SLES 12 SLES 11 SP1 ^e	RHEL 7 RHEL 6
OSA-Express5S Gigabit Ethernet LX and SX CHPID type OSD (using two ports)	SLES 12 SLES 11	RHEL 7 RHEL 6
OSA-Express5S 1000BASE-T CHPID type OSC (using one or two ports per CHPID)	No	No

Function	SUSE	Red Hat
OSA-Express5S 1000BASE-T CHPID type OSD (using two ports per CHPID)	SLES 12 SLES 11	RHEL 7 RHEL 6
OSA-Express5S 1000BASE-T CHPID type OSE ^f (using one or two ports per CHPID)	No	No
OSA-Express5S 1000BASE-T CHPID type OSM	SLES 12 SLES 11 SP2	RHEL 7 RHEL 6
OSA-Express5S 1000BASE-T CHPID type OSN (using two ports per CHPID)	SLES 12 SLES 11	RHEL 7 RHEL 6
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSD	SLES 12 SLES 11	RHEL 7 RHEL 6
OSA-Express4S 10 Gigabit Ethernet LR and SR CHPID type OSX	SLES 12 SLES 11 SP1 ^c	RHEL 7 RHEL 6
OSA-Express4S Gigabit Ethernet LX and SX CHPID type OSD (using two ports)	SLES 12 SLES 11	RHEL 7 RHEL 6
OSA-Express4S 1000BASE-T ^g CHPID type OSD (using two ports per CHPID)	SLES 12 SLES 11	RHEL 7 RHEL 6
OSA-Express4S 1000BASE-T ^e CHPID type OSM	SLES 12 SLES 11 SP3	RHEL 7 RHEL 6
OSA-Express4S 1000BASE-T ^e CHPID type OSN (using two ports per CHPID)	SLES 12 SLES 11	RHEL 7 RHEL 6

- a. SLES12 and RHEL7 supports 256 PUs and SLES11 and RHEL6 support 64 PUs. SLES and RHEL support 64 PUs when running as a guest under z/VM without multithreading and 32 PUs with multithreading.
- b. Via CCA host library
- c. Requires EP11 host library
- d. CEX5S toleration mode support
- e. Maintenance update required.
- f. A CHPID Type OSE supports both SNA (LLC2) and IP connectivity over Ethernet (802.3 or DIX V2).
- g. Not available on zBC12.

IBM is working with its Linux distribution partners so that use of further IBM z13 functions are provided in future Linux on z Systems distribution releases. We suggest the following updates:

- ▶ SUSE SLES 12 or Red Hat RHEL 7 should be used in any new projects for the z13.
- ▶ Any Linux distributions should be updated to their latest service level before migration to z13.
- ▶ Adjust the capacity of any z/VM or Linux logical partitions, and of any z/VM guests, in terms of the *number* of IFLs and CPs, real or virtual, in face of the increased processor unit (PU) capacity of the z13.

4.2.6 References

Planning information for each operating system is available on the following support websites:

- ▶ z/OS:
<http://www.ibm.com/systems/z/os/zos>

- ▶ z/VM:
<http://www.ibm.com/systems/z/os/zvm>
- ▶ z/TPF:
<http://www.ibm.com/systems/z/os/tpf>
- ▶ z/VSE:
<http://www.ibm.com/systems/z/os/zvse>
- ▶ Linux on z Systems:
<http://www.ibm.com/systems/z/os/linux>

4.3 Software support for zBX Model 004

IBM z BladeCenter Extension (zBX) model 004 is available as an upgrade from an existing zBX Model 002 or Model 003. The following operating systems would be supported on zBX model 004:

- ▶ AIX(R) (on POWER7(R) blade located in IBM BladeCenter(R) Extension Mod 004):
AIX 5.3, AIX 6.1 and AIX 7.1 and subsequent releases and PowerVM Enterprise Edition
- ▶ Linux on System x (on IBM BladeCenter HX5 blade installed in zBX Mod 004):
Red Hat RHEL 5.5 and up 6.0 and up, RHEL 7.0 and up, SLES 10 (SP4) and up, 11 SP1 and up, SLES 12 and up - 64 bit only
- ▶ Microsoft Windows (on IBM BladeCenter HX5 blades installed in ZBX Mod 004)
Microsoft Windows Server 2012, Microsoft Windows Server 2012 R2, Microsoft Windows Server 2008 R2 and Microsoft Windows Server 2008, SP2 (Datacenter Edition recommended) 64 bit only

4.4 z/OS considerations

z13 base processor support is required in z/OS. With that exception, software changes do not require the new z13 functions and, equally, the new functions do not require functional software. The z/OS approach is to (where applicable) automatically decide to *enable* or *disable* a function that is based on the presence (enable) or absence (disable) of the required hardware and software.

General recommendations

The z13 introduces the latest System z technology; notable cases are Flash Express, zEDC Express, 10GbE RoCE Express support, Simultaneous Multi Threading (SMT) and Single Instruction Multiple Data (SIMD) processing. Although support for z13 is provided by z/OS starting with z/OS V1R12, the use of z13 functions is dependent on the z/OS release.

In general, we suggest the following approaches:

- ▶ Do not migrate software releases and hardware at the same time.
- ▶ Keep members of the sysplex at the same software level other than during brief migration periods.
- ▶ Migrate to a Server Time Protocol (STP) - only Coordinated Timing Network (CTN) before introducing a z13 into a Sysplex.

- ▶ Review z13 restrictions and considerations before creating an upgrade plan.

Flash Express

z13 supports the Flash Express feature, also available on the zEC12 and BC12, which can help improve resilience and performance of the z/OS system. Flash Express is designed to assist with the handling of workload spikes or increased workload demand that might occur at the opening of the business day, or in the event of a workload shift from one system to another.

z/OS is the first exploiter to use Flash Express storage as Storage Class Memory (SCM) for paging store and SVC dump. SVC dump data capture time is expected to be substantially reduced. As a paging store, Flash Express storage is suitable for workloads that can tolerate paging and does not benefit workloads that cannot afford to page. The z/OS design for Flash Express storage does not completely remove the virtual storage constraints that are created by a paging spike in the system.

Flash Express storage is allocated to logical partition, similar to main memory. The initial and maximum amount of Flash Express Storage that is available to a particular logical partition is specified at the SE or HMC via a new flash memory allocation panel.

The amount of Flash Express storage in the partition can be changed dynamically, between the initial and the maximum amount at the SE or HMC. For z/OS, this can also be done by an operator command. Each partition's Flash Express storage is isolated similarly to main storage and each partition sees only its own space in the flash memory space.

Flash express feature has been enhanced to support pageable large (1 MB) pages. Also coupling facility (CF) support for Flash Express for certain list structures (Websphere MQSeries for z/OS Vsrion 7) has been enabled with PTF support.

zEnterprise Data Compression (zEDC) Express

zEDC Express, an optional feature available for zEC12,zBC12 and z13, addresses data growth requirements by providing hardware-based acceleration for data compression and decompression. zEDC provides data compression with lower CPU consumption than previously existing compression technology on System z.

The z/OS V2R1 zEnterprise Data Compression capability exploits zEDC Express and has the following minimum requirements:

- ▶ z/OS V2R1 with PTFs and the zEDC for z/OS feature
- ▶ z/OS V1R13 with PTFs (software decompression support only, no compression)
- ▶ z/OS V1R12 with PTFs (software decompression support only, no compression)

Initial exploiters included the following use cases.

- ▶ SMF archive data compressed using zEDC provided substantial benefit to increase the amount of data kept online
- ▶ zSecure product output size of monitor and unload files reduced giving substantial media savings benefit
- ▶ z/OS SVC and stand-alone dumps could be stored in up to 5X less space

Support for DFSMS (BSAM/QSAM) has been enabled, providing a new type of policy-based compression support for non-VSAM extended format data sets.

zIIP support

Starting z13 there are no more *IBM System z Application Assist processor (zAAP)* available. *IBM System z Integrated Information Processor (zIIP)* will be able to drive zAAP eligible workload too. A maximum of two zIIPs per one CP can be installed.

Because z/VM can dispatch virtual zIIPs on real CPs¹, the z/VM partition does not require any real zIIPs defined to it. However, in general, real zIIPs should be used due to software licensing reasons.

Large page support

On z/OS, memory that is reserved for large page support needs to be defined in the IEASYSxx member of SYS1.PARMLIB. The definition cannot be dynamically changed.

HiperDispatch

Hiperdispatch exploits the new core-cache design, the new chip-cache design, and the new node and drawer-cache design of the z13 to increase the number of instructions per cycle.

On z/OS, a parameter of the IEAOPTxx member of SYS1.PARMLIB controls whether *HiperDispatch* is enabled or disabled for the z/OS image. It can be dynamically changed (without an initial program load (IPL) or any outage). The default is the hiperdispatch enabled.

To use HiperDispatch effectively, adjustment of defined *Workload Manager (WLM)* goals and policies might be required. We suggest that WLM policies and goals are reviewed and updated as necessary, mainly after a migration to a new system z server or any important change on the system.

A health check is provided to verify whether HiperDispatch is enabled on z13

Capacity provisioning

Installation of the capacity provision function on z/OS requires the following tasks to be completed:

1. Setting up and customizing z/OS RMF, including the Distributed Data Server (DDS)
2. Setting up the z/OS CIM Server (a z/OS base element with z/OS V1R9 and later)
3. Performing capacity provisioning customization as described in the *z/OS MVS Capacity Provisioning User's Guide*, SC33-8299

Use of the capacity provisioning function requires the following elements:

- ▶ TCP/IP connectivity to observed systems
- ▶ TCP/IP connectivity from the observing system to the HMC of observed systems
- ▶ IBM Resource Measurement Facility (RMF) Distributed Data Server must be active
- ▶ Common Information Model (CIM) Server must be active
- ▶ Security and CIM customization
- ▶ Capacity Provisioning Manager customization

In addition, the Capacity Provisioning Control Center must be downloaded from the host and installed on a personal computer (PC) workstation. This application is only used to define policies. It is not required to manage operations.

Customization of the capacity provisioning function is required on the operating system that observes other z/OS systems in one or multiple sysplexes. See the following resources:

¹ The z/VM system administrator can use the SET CPUAFFINITY command to influence the dispatching of virtual specialty engines on CPs or real specialty engines.

- ▶ For a description of the capacity provisioning domain, see *z/OS MVS Capacity Provisioning User's Guide*, SC33-8299.
- ▶ For more details about capacity provisioning, see *IBM System z10 Enterprise Class Capacity On Demand*, SG24-7504.

Integrated Cryptographic Service Facility (ICSF)

Integrated Cryptographic Service Facility (ICSF) is a base component of z/OS. It is designed to transparently use the available cryptographic functions, whether CPACF or Crypto Express features, to balance the workload and help address the applications' bandwidth requirements.

Despite being a z/OS base component, ICSF new functions are generally made available through a web deliverable support a couple of months after a new z/OS release is launched. Because of this fact, new functions must be related to an ICSF *function modification identifier (FMID)* instead of a z/OS version.

See the *IBM zEnterprise EC12 Technical Guide*, SG24-8049 for a table listing the ICSF FMIDs and web-deliverable codes for z/OS V1R10 through V2R1. Later FMIDs include the functions of previous ones.

The Cryptographic Support for z/OS V1R13-z/OS-V2R1 web deliverable (ICSF FMID HCR77B0) is required to use the new functions that are available with the z13. Crypto Express5S toleration is available through maintenance at ICSF FMID HCR7780 or higher.

Coupling links

Each system can use, or not use, *Internal coupling links*, *InfiniBand coupling links* or *Integrated Coupling Adapter (ICASR) coupling links* independently of what other systems are doing. z13 does not support participating in a Parallel Sysplex with System z10 and earlier systems.

Coupling connectivity is only available when other systems also support the same type of coupling. We suggest that, when planning to use the InfiniBand coupling or ICA SR coupling links technology, the *Coupling Facility Configuration Options white paper* be consulted. The white paper is available at the following website:

<http://www.ibm.com/systems/z/advantages/pso/whitepaper.html>

z/OS XL C/C++ considerations

z/OS V2R1 with program temporary fixes (PTFs) or newer is required to use the latest level (11) of the following two C/C++ compiler options:

- ▶ ARCHITECTURE. This option selects the minimum level of machine architecture on which the program runs. Certain features that are provided by the compiler require a minimum architecture level. ARCH(11) uses instructions that are available on the z13.
- ▶ TUNE. This option allows optimization of the application for a specific machine architecture, within the constraints that are imposed by the ARCHITECTURE option. The TUNE level must not be lower than the setting in the ARCHITECTURE option.

For more information about the ARCHITECTURE and TUNE compiler options, see *z/OS V2R1.0 XL C/C++ User's Guide*, SC14-7307.

C/C++: A C/C++ program that is compiled with the ARCH(11) or TUNE(11) options runs only on z13. Otherwise, an operation exception results. This is a consideration for programs that might need to run on different-level servers during development, test, production, and fallback or disaster recovery.

IBM z Advanced Workload Analysis Reporter (IBM zAware)

With IBM zEnterprise EC12 and BC12, IBM introduced a new technology, IBM zAware, employing analytic algorithms based on machine learning developed by IBM Research.

The new version of IBM zAware introduces a new generation of technology with improved analytics to reduce false positives and provide better results. The previous version required messages with message ID (for example, z/OS and other IBM software). In the improved implementation, IBM zAware can process message streams that do not have message IDs (Linux operating system).

IBM zAware uses near real-time continuous learning algorithms, providing a diagnostics capability intended to help you quickly pinpoint problems, which in turn, can help you to more rapidly address service disruptions. IBM zAware uses analytics to intelligently examine z/OS console messages and Linux system log to find unusual patterns, inconsistencies, and variations.

Large z/OS operating system environments can sometimes generate more than 25 million messages per day. This can make manual analysis time consuming and error-prone when exceptional problems occur. IBM zAware provides a simple graphical user interface (GUI) to help you find message anomalies quickly which can help speed problem identification.

IBM zAware and z/VM: IBM intends to deliver IBM z Advanced Workload Analysis Reporter (IBM zAware) support for z/VM. This future release of IBM zAware is intended to help identify unusual behaviors of workloads running on z/VM in order to accelerate problem determination and improve service levels.

4.5 Coupling facility and CFCC considerations

Coupling facility connectivity to a z13 is supported on the zEC12, zBC12, z196, z114, or another z13. The logical partition running the Coupling Facility Control Code (CFCC) can be located on any of these supported systems.

Up to 141 ICF engines can be ordered on a single z13 server across multiple coupling facility LPARs. This helps environments that use server hosting multiple coupling facilities to support multiple Parallel Sysplexes. There is still a limit of 16 ICF engines for a single coupling facility LPAR.

Because coupling link connectivity to System z10 and previous systems is not supported, this might affect the introduction of z13 into existing installations and require more planning. For more information, see the companion book *IBM z13 Technical Guide*, SG24-8251.

z13 supports CFCC Level 20. To support migration from one CFCC level to the next, different levels of CFCC can be run concurrently if the Coupling Facility logical partitions are running on different central processor complexes (CPCs). CF logical partitions running on the same CPC share the CFCC level.

CFCC exploitation Flash Express for WebSphere MQ

CFCC Level 19 and above can exploit Flash Express feature to help handle the overflow of WebSphere MQ shared queues. Overflow areas for certain Coupling Facility list structures can be provided by the Flash Express feature. This is designed to allow structure data to be migrated to Flash Express memory as needed and migrated back to real memory to be processed. This requires WebSphere MQ for z/OS V7, running on z/OS V2R1 or V1R13, with additional service.

CFCC Coupling Thin Interrupts

The Coupling Thin Interrupts enhancement, delivered with CFCC 19, improves the performance of a Coupling Facility partition and improves the dispatching of z/OS Ipars awaiting the arrival of a returned asynchronous CF requests, when used in a shared engine environment.

CFCC Large Memory support

CFCC Level 20 supports Coupling Facility Use of Large Memory with improved availability for larger CF cache structures and data sharing performance with larger DB2 group buffer pools (GBP). This support removes inhibitors to using large CF structures, enabling use of large memory to appropriately scale to larger DB2 local buffer pools (LBP) and group buffer pools (GBP) in data sharing environments.

CFCC support 256 CHPIDs

CHPID support^a:

- ▶ At the time of this writing IBM plans to support up to 256 coupling CHPIDs on z13, that is twice the 128 coupling CHPIDs supported on zEC12. This provides enhanced connectivity and scalability for a growing number of coupling channel types and facilitates consolidation of multiple sysplexes into the same set of physical servers.
- ▶ Each CF image will continue to support a maximum of 128 coupling CHPIDs.

a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

For more information about CFCC code levels, see the Parallel Sysplex website:

<http://www.ibm.com/systems/z/pso/cftable.html>

For latest recommended levels see the current exception letter published on Resource Link:

<https://www-304.ibm.com/servers/resourceLink/lib03020.nsf/pages/exceptionLetters?openDocument>

4.6 Input/output configuration program (IOCP) considerations

All z Systems require a description of their I/O configuration. This description is stored in *Input/Output Configuration Data Set (IOCDS)* files. The *Input/Output Configuration Program (IOCP)* allows creation of the IOCDS file from a source file that is known as the *Input/Output Configuration Source (IOCS)*.

The IOCS file contains detailed information for each channel and path assignment, each control unit, and each device in the configuration.

The required level of IOCP for the z13 is *V5 R1 L0* (IOCP 5.1.0) or later with PTF. For details, see *System z Input/Output Configuration Program User's Guide for ICP IOCP*, SB10-7037-12.

4.7 GDPS Virtual Appliance

Is it typical to deploy IT environments based on IBM z Systems that are running only Linux on z Systems. The GDPS Virtual Appliance is a building block of High Availability and Disaster Recover solutions for those environments that do not have nor require z/OS skills.

Statement of Direction^a: In the first half of 2015, IBM intends to deliver a GDPS/Peer to Peer Remote Copy (GDPS/PPRC) multiplatform resiliency capability for customers who do not run the z/OS operating system in their environment. This solution is intended to provide IBM z Systems customers who run z/VM and their associated guests, for instance, Linux on z Systems, with similar high availability and disaster recovery benefits to those who run on z/OS. This solution will be applicable for any IBM z Systems announced after and including the zBC12 and zEC12.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

Is it typical to deploy IT environments based on IBM z Systems that are running only Linux on z Systems. The GDPS Virtual Appliance is a building block of High Availability and Disaster Recover solutions for those environments that do not have nor require z/OS skills.

GDPS Virtual Appliance is a fully integrated continuous availability and disaster recovery solution for Linux on z Systems customers which consists of:

- ▶ An operating system image²
- ▶ The application components
- ▶ An appliance management layer which makes the image self-containing
- ▶ API and UI for customization, administration, and operation tailored for the appliance function.

It is designed to improve both consumability and time-to-value for customers.

The GDPS Virtual Appliance solution implements the GDPS/PPRC Multiplatform Resilience for System z, also known as xDR. xDR coordinates near-continuous availability and DR solution by these methods:

- ▶ Disk error detection
- ▶ Heartbeat for sanity checks
- ▶ Re-IPL in place
- ▶ Coordinated site takeover
- ▶ Coordinated HyperSwap
- ▶ Single point of control

Although the GDPS Virtual Appliance is supported with z/VM V5R4 or higher, on IBM z13 z/VM V6R2 or higher is required.

For more information, see *IBM z13 Technical Guide*, SG24-8251.

² GDPS Virtual Appliance operating system image requires 1 LPAR with one logical CP to run.



A system of insight for digital business

We started this book by reviewing current IT infrastructure deployments and the most pressing problems that they present, and the need to better align IT with business requirements. In this chapter, we further explore how the z13 can help establish an infrastructure that matters, with improved ability to rapidly respond to business needs: agile, flexible, resilient, and secure.

This chapter describes the following topics:

- ▶ Integrated hybrid infrastructures
- ▶ Benefiting from a smart infrastructure
- ▶ Cloud, analytics, mobile, social, and security
 - Cloud computing
 - Analytics on IBM z13
 - IBM z13 serving mobile and social applications
 - Security on IBM z13

5.1 Integrated hybrid infrastructures

IBM z Systems has long been an integrated diverse platform, with specialized hardware, and dedicated computing capabilities. Recall, for instance, in the mid-1980s, the IBM 3090, and its vector facility (occupying a separate frame). Or note that the cryptographic processors and all the I/O cards, which are specialized dedicated hardware, run non z Systems code on specialty processors. This design allows z Systems processor units (PU)s to concentrate on application's computational tasks, while specialized processors take care of the other tasks, thus providing more value for the money through lower cost of effective computation of application data,

All of these specialized hardware components have been seamlessly integrated within the mainframe for over a decade. IBM extended, where possible, the same integration and simplification philosophy to other servers outside of the mainframe platform, creating a logical environment of shared resources capable of handling diverse workloads.

It might seem that increased flexibility inevitably leads to increased complexity. However, it does not need to be that way. IT operational simplification greatly benefits from the z13 intrinsic autonomic characteristics and the ability to consolidate and reduce the number of system images. There are also benefits from management best practices and products that were developed and are available for the mainframe, in particular for the z/OS and Linux on z Systems environments.

5.1.1 A cornerstone of a smart IT infrastructure

An important point is that the z Systems stack consists of much more than just a system. This claim can be made because of the total systems view that guides z Systems development. The z Systems stack is built around services, systems management, software, and storage. It delivers a complete range of policy-driven functions, pioneered and most advanced in the z/OS environment. It includes the following functions:

- ▶ Access management, to authenticate and authorize who can access specific business services and associated IT resources.
- ▶ Use-management to drive maximum utilization of the system. Unlike other classes of servers, z Systems systems are designed to run at 100% utilization all the time, which is based on the varied demands of its users.
- ▶ Availability: Clustering can be used to achieve 99,999% availability *at the application level* (often referred to as the “five nines”).
- ▶ Just-in-time capacity, to deliver more processing power and capacity when needed.
- ▶ Virtualization security, to enable clients to allocate resources on demand without fear of security risks.
- ▶ Enterprise-wide operational management and automation, leading to a more autonomic environment.

IBM z13 is the result of the IBM sustained and continuous investment and development policies. Commitment to IBM systems design means that z13 brings all this innovation while helping clients to optimize their current investment in the mainframe and improve the economics of IT.

A number of enterprises are reaching the limits of available physical space and electrical power at their data centers. The extreme virtualization capabilities of the z13 enable the creation of dense and simplified infrastructures that are highly secure and can lower

operational costs. They are, therefore, the most powerful tools available to reduce cost (including cost of energy) and complexity in enterprise data centers.

Further simplification is possible by using the HiperSockets¹ and z/VM virtual switch functions. These functions can be used, at no additional cost, to replace physical routers, switches, and their cables, while eliminating security exposures and simplifying configuration and administration tasks. In actual simplification cases, measured at customer sites, cables have been reduced by 97%.

In the following sections we describe how these capabilities play a key role in cloud computing. The following capacity range and flexibility features and benefits translate into a flexible infrastructure that is based on an integrated heterogeneous environment, on which a wide range of workloads can be seamlessly deployed and managed:

- ▶ A processor equally able to handle compute-intensive and I/O-intensive workloads
- ▶ Specialty engines for improved price/performance
- ▶ Extreme virtualization
- ▶ Secure access to data (and the network)
- ▶ Additional platforms and the Unified Resource Manager

5.2 Benefiting from a smart infrastructure

When distilled to a single central theme, embracing a heterogeneous infrastructure and creating an ensemble has the single objective of providing an optimized infrastructure.

In this smart infrastructure, multiple workloads can be deployed across heterogeneous environments and managed under a common umbrella. The computing resources of different hardware platforms are managed as a single view with these characteristics

- ▶ Tuned for the task and optimized across the infrastructure
- ▶ Managed end-to-end for flexible delivery of high value services, in a cloud-like manner
- ▶ Designed for enterprise-wide real-time data modeling

Adopting this smarter infrastructure and extracting benefit from it can be done gradually, and does not require disruptive moves such as those popularized under the names *lift and shift* and *rip and replace*. In the IBM solution, protection of the client's investment in all IT aspects is maximized.

The IT infrastructures of clients are all unique, differing in areas such as virtualization, monitoring, and automation. A single approach does not fit all; there are multiple points of entry to the adoption of an integrated, heterogeneous, virtualized infrastructure.

¹ For a description of HiperSockets, see "HiperSockets" on page 85. The z/VM virtual switch is a z/VM system function that uses memory to emulate network switching hardware.

Figure 5-1 shows a simplified view of a progressive adoption process, where building blocks are added to achieve an enterprise class infrastructure for heterogeneous workload deployment. Again, one can notice a resemblance with cloud computing.

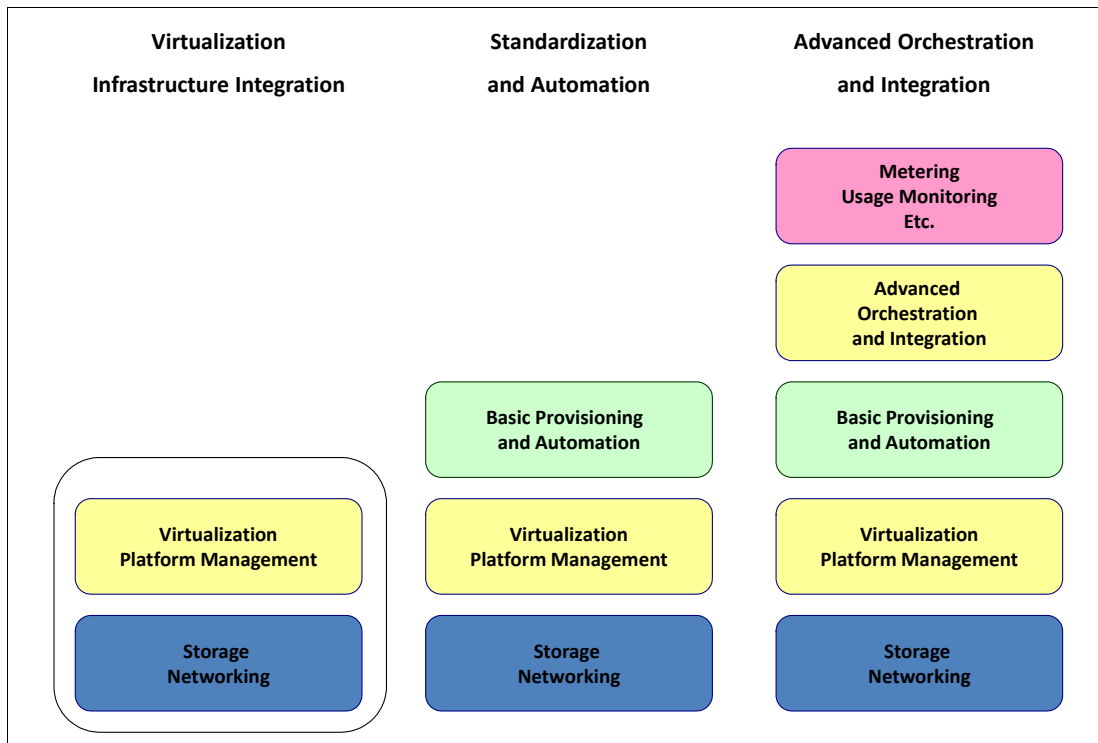


Figure 5-1 Enterprise class heterogeneous platform deployment building blocks

Workloads

Successfully deploying workloads across such an infrastructure requires not only knowledge of application requirements and behavior, but also might affect the enterprise's operational procedures and methodologies, and even the organization.

Understand the terms *application* and *workload*:

Application A computer program or a set of computer programs that are dedicated to perform a defined computational work.

Workload An application (load) runs on a computer to perform a set of functions (work), and requires computing resources to perform these functions.

Workloads require computing resources, which can be classified as the following functions:

- ▶ CPU
- ▶ Memory
- ▶ I/O (networking, storage, and other devices; for instance, graphics devices)

I/O handling also uses a certain amount of CPU and memory resources, but the functions are highly specialized for the type of actions (I/O) that are performed.

Each application function requires a certain amount of the previously mentioned resources. In principle, every type of computer, regardless of the architecture employed (for instance, z Systems, IBM POWER, or System x) provides the three types of resources and can (in principle) perform the same tasks. Thus, a workload is the work that is performed which uses a certain amount of computing resources to run the functions of an application.

Each application has a specific workload *profile* that is determined by the type resources needed computational (CPU, memory) or data movement (I/O) and their variation with time. Throughout the evolution of computing platforms, applications have been developed to use the platform characteristic features (CPU, memory, I/O). Historically, in certain cases, a platform has been changed or adapted to better suit the application needs (for example, the continuous evolution of the z Systems platform and the design of the *reduced instruction-set computer (RISC)* architecture).

The two basic types of application-required resources result from the two most important performance objectives:

- ▶ Response time: The ability to return results in a specified time
- ▶ Throughput: The amount of data that can be processed in a specified time interval

In addition, data and service reliability have driven the evolution of platforms by adding specific mechanisms and tools to achieve the desired results. These characteristics are the measurable parameters enabling the establishment of a service level agreement (SLA).

Platform design has been improved to serve application requirements more effectively and safely. Thus, specialized engines that carry out specific tasks have been developed. Such examples include (but are not limited to) the following features:

- ▶ I/O coprocessors, which are designed to offload the main processors from the I/O-related tasks
- ▶ Specialized cryptographic co-processors, which are designed to offload the computational intensive mathematical functions required to encrypt and decrypt data
- ▶ Specialized co-processors for accelerating compressing and decompressing of data
- ▶ Specialized units inside the processor core, such as the binary-coded decimal (BCD) arithmetic logic unit or vector units
- ▶ Special purpose units, such as the IBM DB2 Analytics Accelerator, for accelerating complex queries such as the ones found in Data Warehousing workloads

These improvements have been driven by the necessity of reserving the CPU execution capabilities for the core work of the application (main data computation). The diversity of the business needs and platforms also has determined two approaches for application design:

- ▶ Custom code
- ▶ Commercial off the shelf (COTS) code

Across industries, we see various applications with their specific workload profiles that run on various platforms. Historically, the choice of a platform has been determined by two major aspects:

- ▶ Platform availability (development costs always must consider this aspect)
- ▶ Platform fit-for-the-purpose

The use of programming tools, such as compilers, has greatly contributed to application portability (the ability to run on several platforms). However, because of the diversity of application workload profiles, not all platforms can run the same workload with the same efficiency.

Moreover, depending on the industry and business requirements, an application that might perform the same core functionality, might have more non-functional requirements. Examples include specific data security and availability, that cannot be obtained in a cost-effective manner on all available platforms.

Examples are as follows:

- ▶ In *banking*, there are components across retail and wholesale banking that employ several architectures to run, but the core of most banking applications relies on z Systems and z/OS.
- ▶ *Insurance* typically maintains claims processing on z Systems but reaches out to the internet for interaction with consumers, using Linux, UNIX, Power, and x86. Fraud detection can benefit from exploiting the vast amounts of data managed by z Systems, using z Systems based computing.
- ▶ The *public sector* is relying more on the web-based capability to reach out to citizens and improve the rate of return for taxes, accurate payment of social benefits, election process, and even census-based reporting. Here, too, fraud detection and tax evasion can benefit from z Systems capabilities.
- ▶ The *retail industry* can effectively use a heterogeneous infrastructure (for instance, by benefiting from z Systems large I/O capacity to implement large databases, using business intelligence to characterize their clients).
- ▶ *Airline reservation systems* are one example of extreme online transaction processing. z13 provides the z/TPF operating system and application environment specifically for this situation. Another example of z/TPF utilization is in banking with credit card and ATM processing.

The workloads tend to follow well-established technology and infrastructure patterns, such as the following applications:

- ▶ Core applications (for example, database engines)
- ▶ Multitier web serving
- ▶ Data warehouse/data mining

Attributes of workload components

Each workload pattern is typically made up of components that have distinct characteristics and requirements. The components are woven together with application programs and middleware to enable a business process to achieve business objectives.

Based on the workload attributes (characteristics), we can identify the following main types:

- ▶ Transaction processing and database (OLTP):
 - High transaction rates
 - High quality of service
 - Peak workloads
 - Resiliency and security
- ▶ Analytics and high performance:
 - Compute or I/O intensive
 - High memory bandwidth
 - Floating point and vector processing (SIMD²)
 - Scale-out capable (horizontal scalability)
- ▶ Business applications:
 - Scale up (vertical scalability)
 - High quality of service
 - Large memory footprint
 - Responsive infrastructure
- ▶ Web, collaboration, and infrastructure management:
 - Highly threaded (SMT³)
 - Throughput-oriented

² Single-instruction, multiple-data (SIMD): Same instruction that is run on a vector of data

³ Simultaneous multithreading (SMT): More than one thread runs simultaneous in the same processor

- Scale out capable
- Lower quality of service

From an architectural perspective, it is critical to deploy workload components on the server technology that offers the best fit and is most effective in satisfying their requirements. Thus, multiple platforms might be appropriate. In a more synthetic approach, the workloads might be characterized as the following types:

- ▶ Shared data and multiple work queues (OLTP, for example, or large batch jobs)
- ▶ Parallel data structures (HPC and Analytics)
- ▶ Highly threaded (for example, business applications)
- ▶ Small discrete applications

In addition, today's applications can rarely be classified as only one of the aforementioned types. See Figure 5-2. In most cases, an application consists of components with varying requirements and workload profiles. Thus, proper platform choice is key in obtaining the results (SLA) that you want.

For example, a monothread batch job requires fast movement of data through the processor, and does not, of itself, use any multi-threading capabilities that a platform might have. Thus, the higher the CPU speed, the faster the job is processed. Conversely, a highly threaded application (web services, for example) performs better on a multi-threaded capable platform.

With z13, and in addition to the advanced workload management capabilities of z/OS, the IBM z Unified Resource Manager also offers workload management capabilities of zBX environments:

- ▶ For virtual servers on POWER7 blades, through the POWER/VM hypervisor.
- ▶ For System x in the zBX, by using a function in the KVM hypervisor known as *cgroups*. This allows monitoring the availability of workload resources to satisfy a defined workload service level policy. The Unified Resource Manager will exploit *cgroups* by assigning *cgroup* to virtual servers and dynamically managing the CPU share of a virtual server based on policy goals.

Figure 5-2 shows an image of workload characterization.

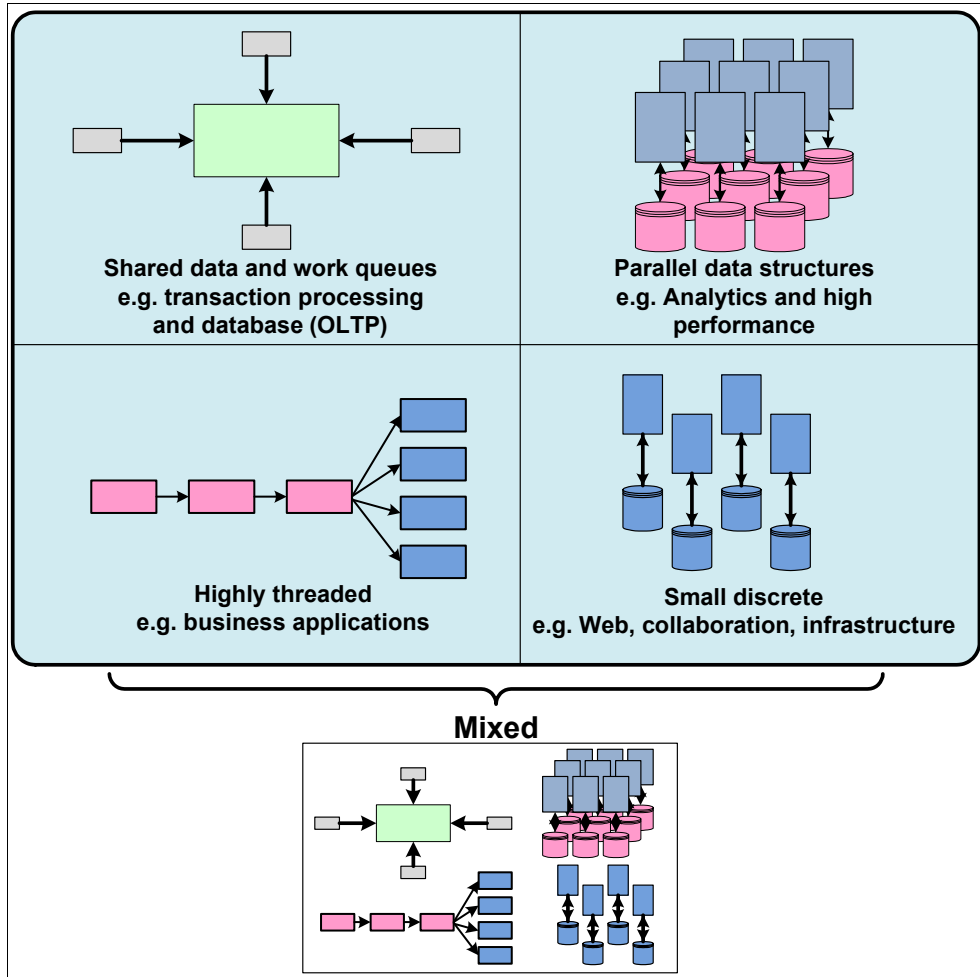


Figure 5-2 Workload characterization

The z13 is the platform of choice for the integration of the new generations of applications with existing applications and data. They truly are cornerstones of a smart IT infrastructure.

5.3 Cloud, analytics, mobile, social, and security

IBM mainframes traditionally provide an advanced combination of reliability, availability, security, scalability, and virtualization. The IBM z13 is designed to extend these capabilities into heterogeneous resources and are optimized for today's business needs. The IBM z13 continues the evolution of the mainframe, building upon the z/Architecture definitions as the mainframe system that is optimized for the digital era.

IBM z13 is an integrated, highly scalable computer system that allows many different pieces of work to be handled at the same time, sharing the same information as needed, with protection, handling very large amounts of information for many users with security, without users experiencing any failures in service.

The IBM z13 is emerging as a choice platform for cloud computing. Cloud technologies can improve the integration of people, processes, and technology to help businesses run more cost effectively while also supporting business growth and innovation. Cloud is the infrastructure model of the digital era, and z13 provides the infrastructure to support all

dimensions of cloud service deliveries. z/OS and Linux on z Systems are the foundations of the most secure and scalable private, public, and hybrid cloud infrastructures.

The z13 is the trusted enterprise platform for integrating data, transactions, and insights. Enterprises typically use traditional integer, floating point, string, and XML character-based data. The amount of data is increasing exponentially, and IT shops need to respond quickly to the diversity of data. Therefore, an aspect that is becoming more important is for customers to do computations and analytics closer to the data. The z13 supports new analytics workloads and offers new and improved capabilities for mathematical optimization of complex models for deeper analytics insights. For the last three generations, z Systems has changed its capabilities in analytics compute processing, such as superscalar, out of order, compiler improvements and precision floating point. Single instruction, multiple-data (SIMD), simultaneous multithreading (SMT), and large memory provide the next phase of enhancements to business and IT analytics transaction processing and compute-intensive competitiveness on z Systems.

The z13 continues to reinvent enterprise IT for digital business. The mobile revolution is changing how people interact and organizations are transforming their business processes to address an increasing demand for mobile capabilities in a dynamic and agile marketplace. The z13 is designed and enabled for a mobile generation. High performance, large-scale, and intelligent I/O and security enhancements on the z13 support transaction growth in the mobile world. New technologies that improve throughput and response time enable organizations to provide rapid data and transaction service to their customers.

The z13 system enables the integration of social technology throughout business processes. Organizations are incorporating social media platforms and realizing value in connecting employees, suppliers, and customers. In a global marketplace, virtual interactions can lead to more efficient and effective results. Communications can be accelerated as teams worldwide become more connected with enterprise collaboration tools, independent of time zones. Software such as IBM Connections and IBM Sametime® are components of the IBM Platform for Social Business, the industry's most comprehensive portfolio to support social technology.

The z13 offers the highest level of security certification for commercial server applications. As the world continues to become more digitized and interconnected, doors to emerging security threats and vulnerabilities continue to open. Cybersecurity attacks have become more sophisticated, with new motivations, and increasing speed and dexterity. The z13 offers the most secure and reliable computing platform through auditable protection of data, simplification of management and compliance, and capability for security analytics. The security strategy is to combine processor-based cryptography and encryption, with firmware and operating system enhancements, to ensure no attack vectors are exploited to compromise the security and integrity of the system.

5.3.1 Cloud computing

“Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.”⁴ A cloud can be public, private, or a hybrid of both. With cloud computing, the application can be running on a server anywhere in the world. That flexibility is why it is changing the way companies provide services to their clients and suppliers.

⁴ US National Institute of Standards and Technology:
<http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>

Figure 5-3 shows the IBM Cloud Computing Reference Architecture (CCRA⁵), which defines the fundamental architectural elements which constitute a cloud computing environment. It is required that all of these infrastructure components be managed from a single, central Common Cloud Management Platform with the ability to place instances of each cloud service on the corresponding infrastructure. This requirement perfectly fits the z13-based heterogeneous infrastructure, with its end-to-end management capabilities for flexible delivery of high-value services.

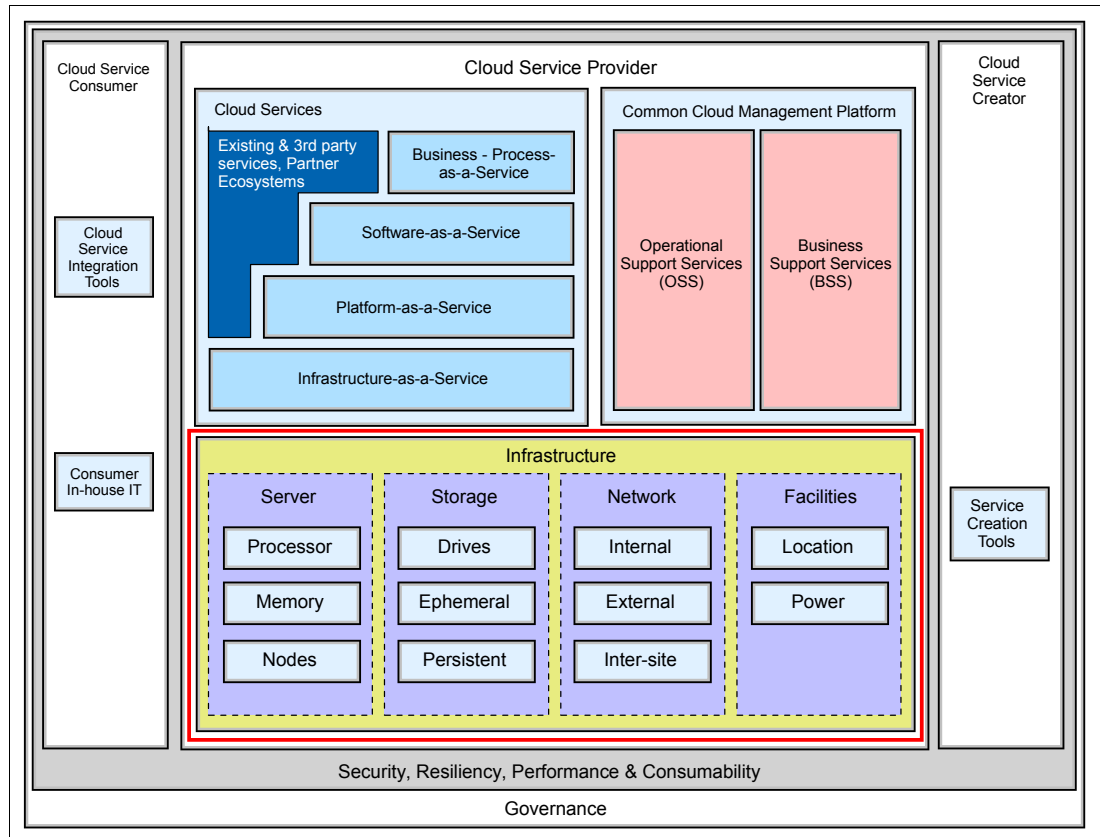


Figure 5-3 The IBM Cloud Computing Reference Architecture infrastructure details

Virtualization is the foundation for cloud, and the benefits of consolidation and virtualization are widely accepted by the IT community. Adding standardization and automation to a virtualized environment enables IT optimization for cloud computing. Workflow orchestration, monitoring, and metering for accounting are other major components of cloud computing.

⁵ https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/Wf3cce8ff09b3_49d2_8ee7_4e49c1ef5d22/page/IBM%20Cloud%20Computing%20Reference%20Architecture%203.0

Deploying a cloud infrastructure is not a simple process, but there is a defined path that can be followed. Figure 5-4 depicts the path from standard managed services to cloud.

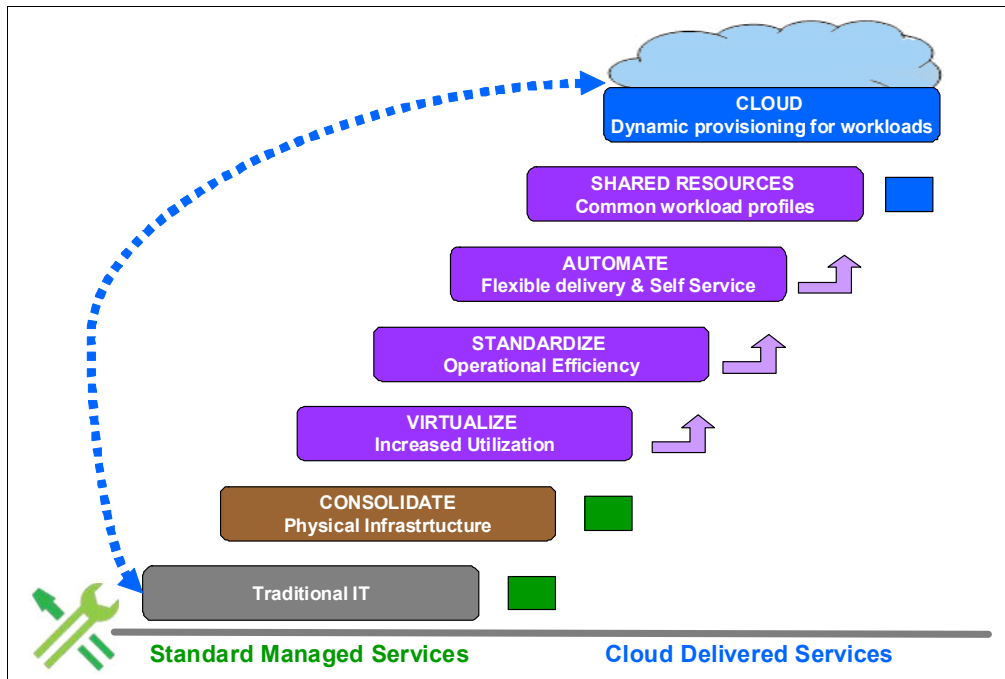


Figure 5-4 Cloud delivered services

Because z13, with its “shared everything infrastructure,” offers a fully virtualized system, it becomes easier to integrate a cloud computing deployment as part of the existing IT optimization strategy and roadmap. Table 5-1 summarizes potential benefits that are provided through cloud computing.

Table 5-1 Benefits with cloud computing

Virtualization	Standardization	Automation
Higher utilization Economy-of-scale benefits Lower capital expense Lower operating expense	Easier access Flexible pricing Reuse and sharing Easier integration	Faster cycle times Lower support costs Optimized utilization Improved compliance Optimized security Better user experience

Cloud computing on z13 builds on the industry’s leading virtualization technology that uses virtualization, standardization, and automation to free operational budget for new investments. This technology also allows you to optimize new investments for direct business benefits. z13 provides the following features:

- ▶ A highly scalable heterogeneous pool of virtualized resources that are managed in a single system
- ▶ On-demand activation, allocation, prioritization, and retiring of resources, and automation of service delivery
- ▶ Maximizing utilization of resources for improved return on investment (ROI) and lower cost of service delivery
- ▶ Increased levels of security, resiliency, and manageability to create a cloud environment that is enterprise ready

- ▶ More processing capacity, multithreading, and larger memory to consolidate different applications with less physical infrastructure.

Building upon all the previous resource management capabilities and core functionality, application-supporting middleware, such as transactional servers, are adding function to enable existing applications to become cloud accessible, without the need to rewrite and replace them. True to IBM commitment, and similarly to web enablement, applications can become naturally immersed in the cloud, while preserving the client's investment.

CICS Transaction Server support for cloud computing: CICS Transaction Server for z/OS V5.1 provides new application, platform, and policy capabilities that can help clients build private clouds from new and existing CICS applications. This capability is intended to assist CICS clients to deploy new and updated CICS applications faster, more easily, and with greater levels of confidence.

The IBM z13 (Figure 5-5 on page 149) enables the following attributes, being the IT industry's first multi-architecture cloud solution:

- ▶ Higher utilization:
 - Up to 100% CPU utilization
 - *Shared everything* architecture
 - Hosting of thousands of mixed workloads
 - Multithreading capability
- ▶ Increased productivity:
 - Single point of control for a heterogeneous infrastructure at a platform level, with the Unified Resource Manager
 - Efficient, rapid provisioning
 - Superior workload management that is enabled with Unified Resource Manager
 - Workload optimization with fit for purpose approach
 - Collocating applications where industry-leading z/OS transaction and data services run
- ▶ More efficient data center:
 - Less power and cooling
 - Less floor space
 - Fewer parts to manage
- ▶ Greater reliability and availability:
 - Built-in hardware redundancy
 - Decades of RAS innovation
 - Capacity and backup on demand
 - Decades-proven virtualization security protecting sensitive data and critical business processes in the cloud
 - Resiliency management and fewer points of failure
- ▶ Security:
 - Extending z Systems security to a private network across heterogeneous resources

Figure 5-5 on page 149 shows a z13 for IT optimization, consolidation, and cloud computing.

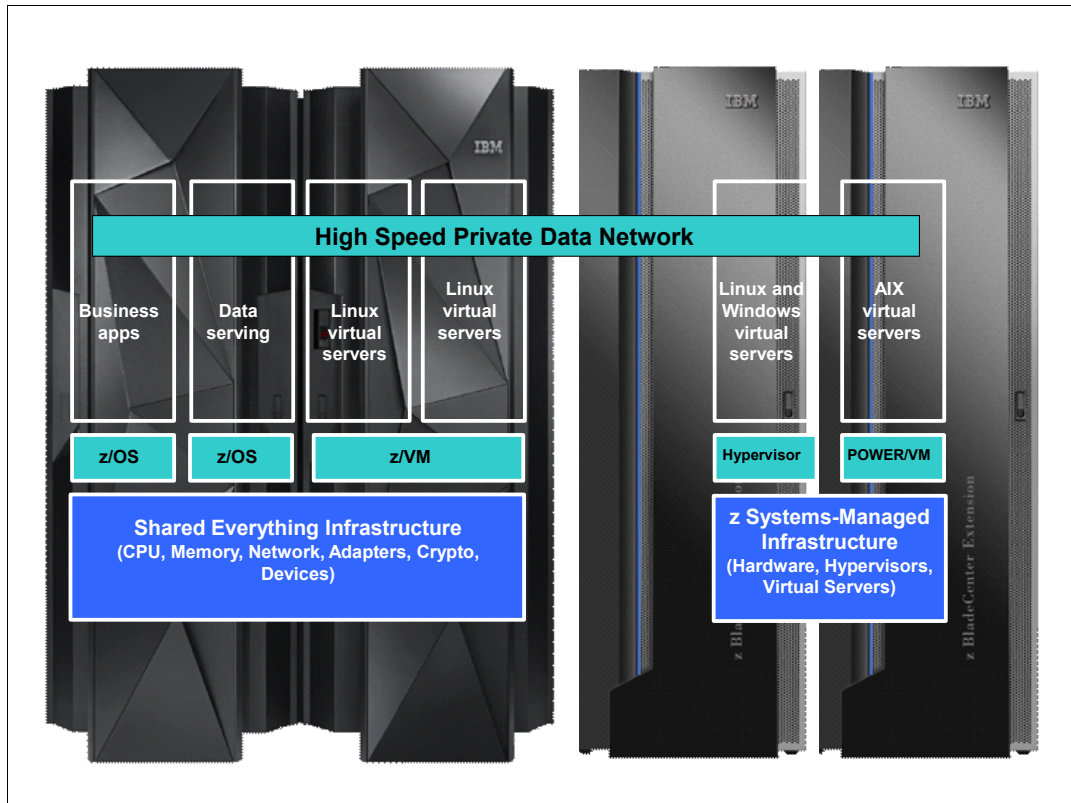


Figure 5-5 z13 for IT optimization, consolidation, and cloud computing

Quality of service improvements

In an ensemble (Figure 5-5), the qualities for which z Systems are renowned are extended to the non z Systems components of the ensemble, providing support for mission-critical workloads running on the ensemble's heterogeneous infrastructure. Compared to other heterogeneous infrastructures, the ensemble provides the following benefits:

- ▶ Potentially lower the cost of enterprise computing by implementing a single management and policy framework across web serving, transaction, database, and servers
- ▶ Simplified operations through integration of multiplatform management capabilities through extended functionality in the well-known mainframe HMC
- ▶ Improved infrastructure reliability by extending the mainframe systems management and service to the zBX environment
- ▶ Improved service through dynamic resource management of the mainframe to all devices within a multitier architecture
- ▶ Simplified and improved infrastructure management through monitoring and management of a heterogeneous solution as a single, logical virtualized solution
- ▶ Better alignment of IT with business objectives by managing the platform's resources in accordance with specified business-service-level objectives
- ▶ Improved infrastructure manageability through management of virtual servers as part of the overall deployed business workload
- ▶ Dramatically simplified infrastructure, improved application performance, and simplified management by using IEDN, the secure and managed Layer 2 network which connects the zBX blades with the CPC

Cloud computing is one of the key ways to address the challenges of today and build an IBM Smarter Planet®. As more companies embrace cloud computing, z13 becomes more obviously the perfect platform for delivering large-scale *software as a service (SaaS)* application software services.

For more information about cloud on an IBM z Systems, see *Deploying a Cloud on IBM System z*, REDP-4711.

GDPS Virtual Appliance

Statement of General Direction^a: In the first half of 2015, IBM intends to deliver a GDPS/Peer to Peer Remote Copy (GDPS/PPRC) multiplatform resiliency capability for customers who do not run the z/OS operating system in their environment. This solution is intended to provide IBM z Systems customers who run z/VM and their associated guests, for instance, Linux on z Systems, with similar high availability and disaster recovery benefits to those who run on z/OS. This solution will be applicable for any IBM z Systems announced after and including the zBC12 and zEC12.

- a. All statements regarding IBM plans, directions, and intent are subject to change or withdrawal without notice. Any reliance on these statements of general direction is at the relying party's sole risk and will not create liability or obligation for IBM.

Deploying IT environments, based on IBM z Systems that are running only Linux on z Systems is common today. The GDPS Virtual Appliance is a building block of high availability and disaster recover solutions for those environments that do not have nor require z/OS skills.

To reduce IT costs and complexity, many enterprises are consolidating independent servers into Linux images (guests) running on z Systems platform. Linux on z Systems can be implemented either as guests running under z/VM or native Linux LPARs on z Systems. Workloads with an application server running on Linux on z Systems and a database server running on z/OS are common. Two examples are as follows:

- ▶ WebSphere Application Server running on Linux and CICS, DB2 running under z/OS
- ▶ SAP application servers running on Linux and database servers running on z/OS

With a multi-tiered architecture, there is a need to provide a coordinated near-continuous availability and disaster recovery solution for both z/OS and Linux on z Systems.

GDPS Virtual Appliance is a fully integrated continuous availability and disaster recovery solution for Linux on z Systems customers and consists of these items:

- ▶ An operating system image
- ▶ The application components
- ▶ An appliance management layer, which makes the image self-containing
- ▶ APIs and UIs for customization, administration, and operation tailored for the appliance function

It is designed to improve both consumability and time-to-value for customers.

For more information, see *IBM z13 Technical Guide*, SG24-8251.

5.3.2 Analytics on IBM z13

Analytics projects mainly draw on mainframe-based transactional data. However, in the last two decades, analytics solutions have been deployed as departmental applications, based on very specific business needs of the requesting department.

Even though most the source of the data fed into the analytics applications is common (shared among various solutions), the results are often not correlated, as they reside in diverse systems and formats (silos).

What seemed initially like a small investment (at individual server and application levels) has evolved into a significant investment when viewed in the aggregate. With many organizations concerned about costs (hardware, software, energy, and labor) and real estate availability (data center), maintaining such a diverse infrastructure can become a mess.

In addition, after the applications follow their lifecycles (design, development and testing, quality assurance, production, disaster recovery), return on investment in the associated infrastructure becomes difficult to evaluate and control.

Data is the most important resource of our clients. The way our clients view and analyze the data increases their competitive advantage:

- ▶ **Real-Time View:** The operational systems that house the books of record. These data sources are critical to day-to-day business processes (systems of record).
- ▶ **Historical View:** Data warehouses, marts, and others. These data sources support reporting and predictive model creation.
- ▶ **Predictive View:** Analytics models and software. These are the tools that deliver actionable insights from data.

Business-critical analytics solutions depend on where the source data resides. IBM z Systems servers delivers a single workload-optimized system for both operations and analytics by integrating and managing real-time, historical and predictive views of data.

Real-time analytics-based decisions is a top priority for outperforming organizations, and this requires tight integration with operational data. The following problems are usually encountered when implementing analytics solutions in a silo manner:

- ▶ Significant effort is spent for copying and moving data, resulting in veracity and security issues
- ▶ Business does not have access to the most current view (data in systems of record is in constant change).
- ▶ Complicated, bifurcated infrastructure requiring multiple skill types
- ▶ No single point of management
- ▶ Business continuity concerns

With keeping data and analyzing it at the source, several advantages can be identified:

- ▶ Less movement of data, resulting in higher quality and less risk of loss
- ▶ Integration with core systems delivers most accurate view to the business
- ▶ Integrated architecture leveraging existing environment
- ▶ Single view simplifies management
- ▶ Business continuity inherited from core systems

z Systems analytics enables organizations to improve performance and lower cost by bringing the analytic processing to where the data resides (see Figure 5-6). Organizations can maximize their current IT investments while adding additional functionality, and improved price/performance with z13 over zEC12. With new z13 features, applications can gain

increased throughput for distributed operational business intelligence (operational BI), DB2 query workloads, thus saving money (hardware, software, labor).

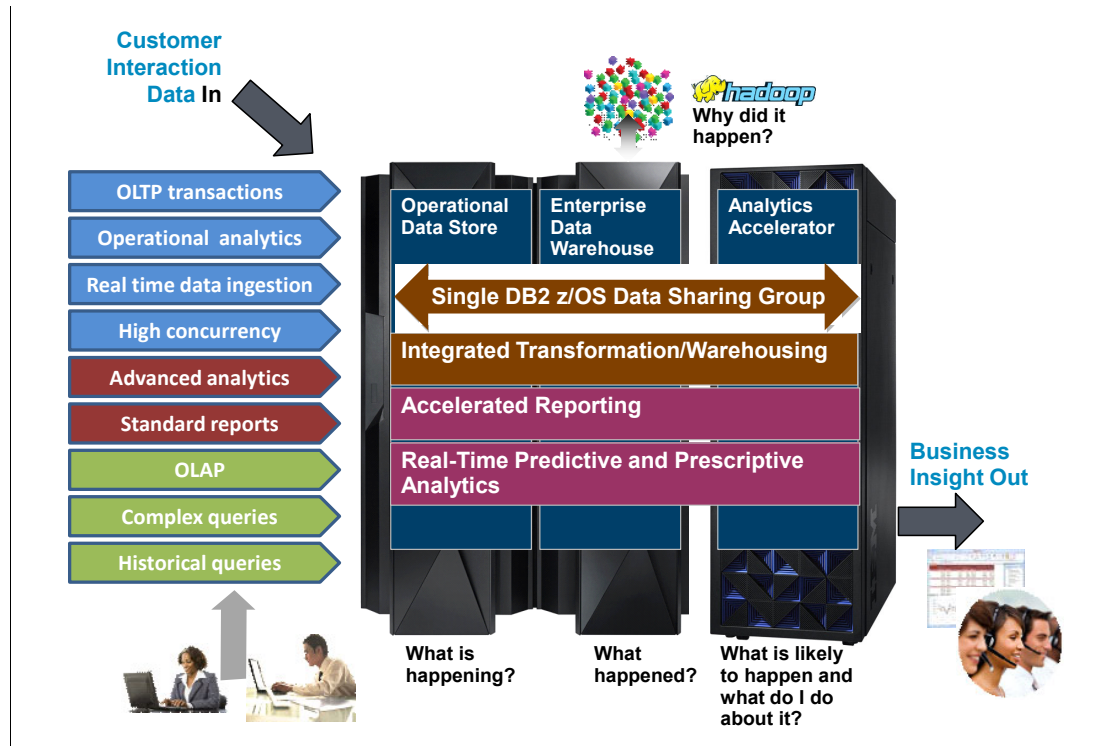


Figure 5-6 IBM z Systems transaction and analytic processing (hybrid)

By keeping (big) data and analytics applications in the z Systems ecosystem, customers can achieve better information governance while saving money to safeguard data and applications. The following aspects set z Systems apart for information governance:

- ▶ Co-location of transactional data and historical data
- ▶ Minimizes duplicate copies of the data
- ▶ Fewer security intrusion points
- ▶ Scalability to leverage a single, consistent copy of the data
- ▶ Consistent business rules across the enterprise

IBM z Systems can help achieve the best-in-class consistency of security controls across the data:

- ▶ Enterprise-wide security, auditing and compliance
- ▶ Real-time monitoring and alerting
- ▶ User identification and authentication at all levels
- ▶ Access defined down to the cell level
- ▶ End-to-end auditing
- ▶ Data is secured with cryptography, encryption

In the context of analytics, the following features and function of z13 bring benefit to workloads:

- ▶ SIMD - MASS and ATLAS libraries, for z/OS and Linux on z Systems:
 - Allow construction of richer, complex analytics models using ISV/SWG Analytics products that exploit SIMD to provide better accuracy of insight,
 - Allow Analytics workloads to be ported from Power and x86 with ease that can accelerate analytics to provide speedy business insight,

- Increase programmer productivity of ISV and customer analytics workload development leading to rapid business insight generation for competitive advantage
- ▶ SMT - process more workload in the same amount of time (zIIPs, IFLs)
- ▶ Enterprise security (built in, one-of-a-kind cryptographic and encryption functions) maintaining data secure, inside the platform
- ▶ Large memory support (z/OS, DB2)

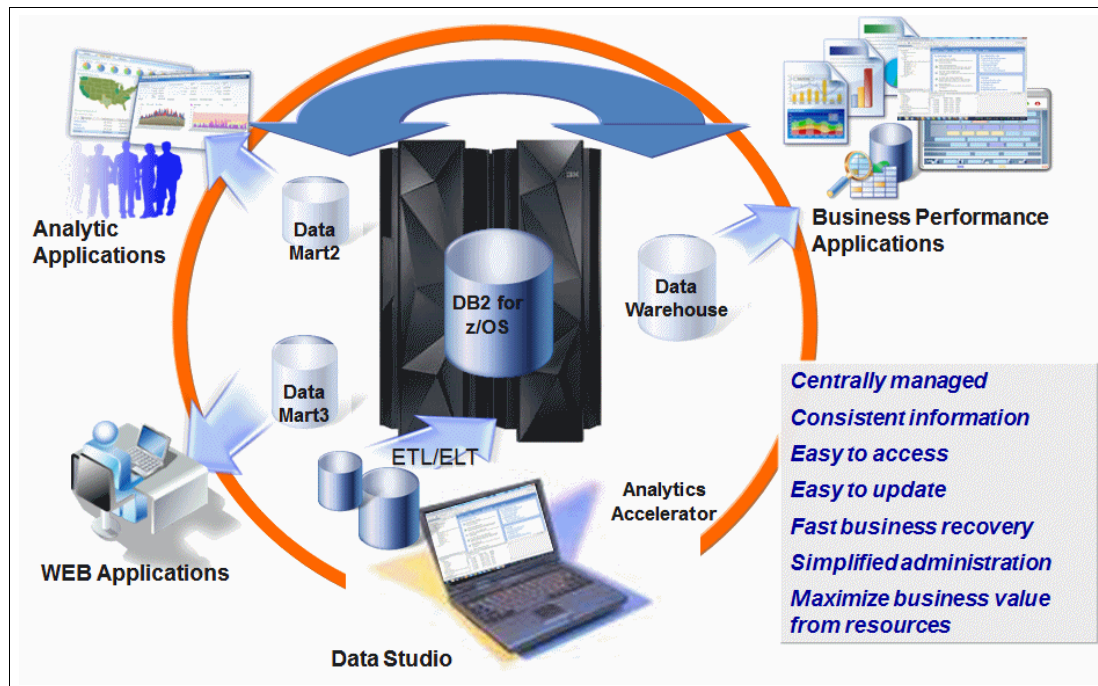


Figure 5-7 Centralized control of decision information

Conclusion

IBM z Systems platform excels as custodian of the data model, providing accurate, secure single copy of information that ensures the veracity of the data for reliable analytics (see Figure 5-7).

5.3.3 IBM z13 serving mobile and social applications

In today's mobile era, over 10 billion devices are accessing information. Enterprises are challenged with integrating new mobile services with existing organizational processes, without sacrificing the client's experience. Figure 5-8 depicts a typical environment where access to applications and interaction with the systems is achieved from mobile devices.

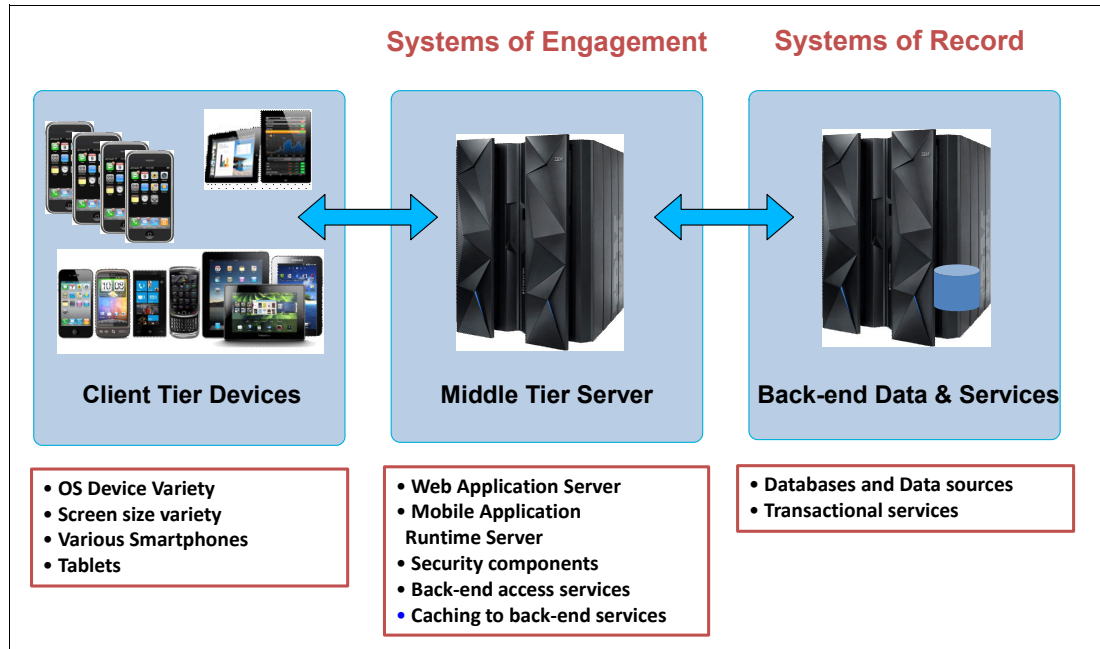


Figure 5-8 Typical mobile environment

IBM z Systems provide you with enterprise mobility solutions which can scale to handle the huge number of often unpredictable transaction rates and volumes, deliver proven mobile end-to-end integration with reliability, availability, and security, and ensure that your customer data is protected.

IBM is the leader platform in system of record (z/OS). IBM provides easily consumable mobile access to all the data and transactions in z Systems software subsystems (DB2, CICS, IMS, MQ, and others). Customers can create engaging mobile apps today using existing transactions on z Systems. IBM flagship operating system z/OS availability and scalability is crucial for mobile workloads.

IBM is also a key player in systems of engagement (Linux on z Systems). Based on its availability and scalability z Systems platform is capable to handle mobile workloads. In this context, Linux on z Systems provides an excellent environment for mobile infrastructure. The diagram in Figure 5-9 on page 155 shows a typical deployment of IBM z Systems in a mobile application environment.

The following features benefit the development and running of mobile applications on IBM z13:

- ▶ SMT: IFLs and zIIPs with software support
- ▶ SIMD: Accelerating computation for analytics
- ▶ Large memory: Providing faster response time to Java and DB2 based applications (in general, any applications that can exploit large amount of in-memory data)
- ▶ Enterprise security capabilities can help you simplify and improve a complex set of operational security processes. IBM z Systems are designed for the highest level of security for commercial grade platforms.
- ▶ High performance communications (network and I/O): SMC-R, FICON Express16s, and so on

IBM provides the tools to satisfy the lifecycle requirements for mobile application development, including, but not limited to these tools:

- ▶ IBM MobileFirst Studio (formerly IBM Worklight Studio) offers leading tools for mobile app development that help maximize code reuse and accelerate development.
- ▶ IBM MobileFirst Server⁶ (formerly IBM Worklight Server) is mobile-optimized middleware that serves as a gateway between applications, back-end systems and cloud-based services.

The IBM MobileFirst Platform provides the following benefits:

- ▶ Exploit z Systems platform security and encryption for use by mobile apps.
- ▶ Exploit co-location with z/OS data and transactions.
- ▶ Easily create new mobile development and production clouds

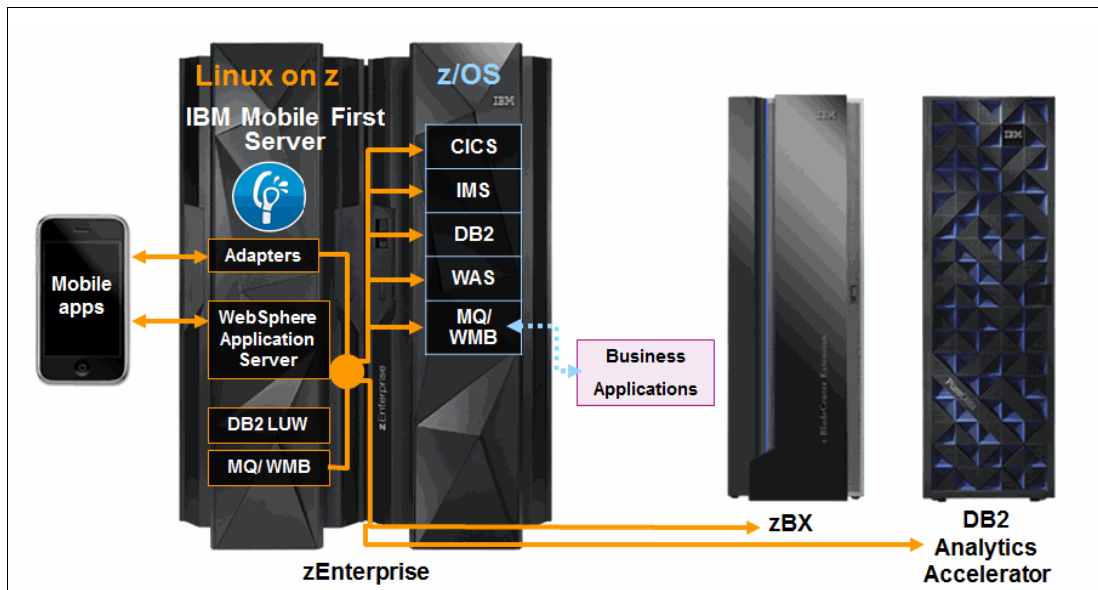


Figure 5-9 Mobile applications on IBM z Systems

IBM z Systems address key mobile development and delivery challenges, such as these:

- ▶ Fragmentation and developing for multiple mobile platforms:

A highly fragmented set of devices, platforms, languages, and tools complicates development, test, and operations.

IBM response is IBM MobileFirst Studio and IBM Rational Developer for z:

 - Seamless integration with z Systems data and transactions;
 - Device runtime provides mobile device independence.
- ▶ Accelerated time to market requirements:
 - Accelerated development demands instant provisioning of development servers.
 - Spikey mobile traffic demands highly scalable cloud-based infrastructures.

IBM response is cloud offerings:

 - Linux on z Systems enables rapid provisioning of dev servers.
 - High scalability can handle workload increases
- ▶ Connecting apps with existing enterprise systems:

⁶ <http://www-03.ibm.com/software/products/en/mobilefirstfoundation>

- Apps typically need to leverage existing enterprise services, which must be made mobile-consumable, and remain secure.
- Enterprise systems must be able to instantly provision new services and environments.

IBM response is that z/OS is mobile enabled:

- All z/OS subsystems are mobile-ready, with added consumability enhancements planned.
- DB2 tech preview – MongoDB⁷ API;
- End to end mobile security.
- High-performance access from Linux on z Systems

Social

While mobile applications enable customer engagement anytime, anywhere, the huge volume of data generated can also be fed into analytics applications that provide business decisions with the social perspective. Add real-time scoring to your OLTP workload with minimal impact on CPU consumption and response time. z Systems answer this call drastically reducing analytic latency via in-database scoring combined with new support for IBM DB2 Analytics Accelerator (IDAA)-based modelling. Real-time scoring (real-time SPSS scoring in DB2 for z/OS) allows organizations to take next best action within the transaction with minimal CPU impact. Some of the applications that involve social aspects are:

- ▶ Social media analytics
- ▶ Customer sentiment
- ▶ Customer identification
- ▶ Identification of illegal or suspicious activities
- ▶ Employee and customer experiences

Social behavior analysis provides another perspective to business, enhancing business decisions with prediction of future trends and the capability of detecting errors before they may actually occur (fraud detection is one example).

Some applications examples are (including but not limited to):

- ▶ IBM SPSS® -- Statistical Package for the Social Sciences
<http://www-01.ibm.com/software/analytics/spss/>
- ▶ IBM Employee Experience Suite for Linux on z Systems
http://www-01.ibm.com/common/ssi/rep_ca/6/897/ENUS214-166/ENUS214-166.PDF
- ▶ IBM Connections
<http://ibmconnections.com/>

Conclusion

The role of z Systems in the mobile and social environments can be resumed as follows:

- ▶ Build an agile approach to deliver applications.
- ▶ Make every transaction secure.
- ▶ Use mobile analytics to improve outcomes at every moment.

5.3.4 Security on IBM z13

Analytics, mobile, social and cloud all have one thing in common: the need for a platform that has a deeply integrated security stack. This is where IBM z Systems excel. With the threats to business in the current world, including everything from hackers looking to see what they can

⁷ <http://www.mongodb.org/>

do, to government backed attacks, IBM z Systems offer a platform with layers of defense to protect your customer data, intellectual property, and your reputation.

This infrastructure is the basis for a platform ideal to support today's cloud, mobile, analytics and social networking workloads. With the holistic approach to security, the IBM z Systems provide the security foundation for cloud, analytics, mobile, and social applications, as shown in Figure 5-10.

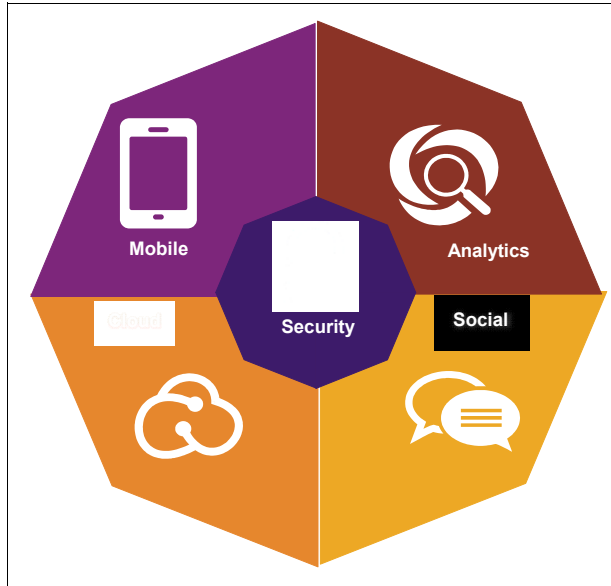


Figure 5-10 Cloud, analytics, mobile, and social, underpinned by Security

Since the IBM System 360 was first introduced, it had security built into the hardware and operating system with storage protection keys. We have continued to build upon the security with processor base encryption, LPAR isolation, operating system protections, network and secure communications, and self-encrypting tape and disk storage solutions. For more information on how to configure and secure operations on a mainframe, see *Security on the IBM Mainframe, Vol. 1*, SG24-7803-01.

IBM Security Framework

These matters of business-driven enterprise security are encapsulated in a concept known as the IBM Security Framework. The IBM Security Framework (Figure 5-11) provides a business view of the security posture of an enterprise. It is a high level view, but it incorporates all that is necessary for consideration.

IBM Security Framework can help with these items:

- ▶ Enable innovation through security-rich, end-to-end infrastructure and platforms.
- ▶ Reduce number and complexity of required security controls.
- ▶ Reduce redundant security expenses.
- ▶ Improve organizational and operational agility and resiliency.
- ▶ Leverage industry expertise to help unify policy management.
- ▶ Deliver needed visibility, control, and automation.

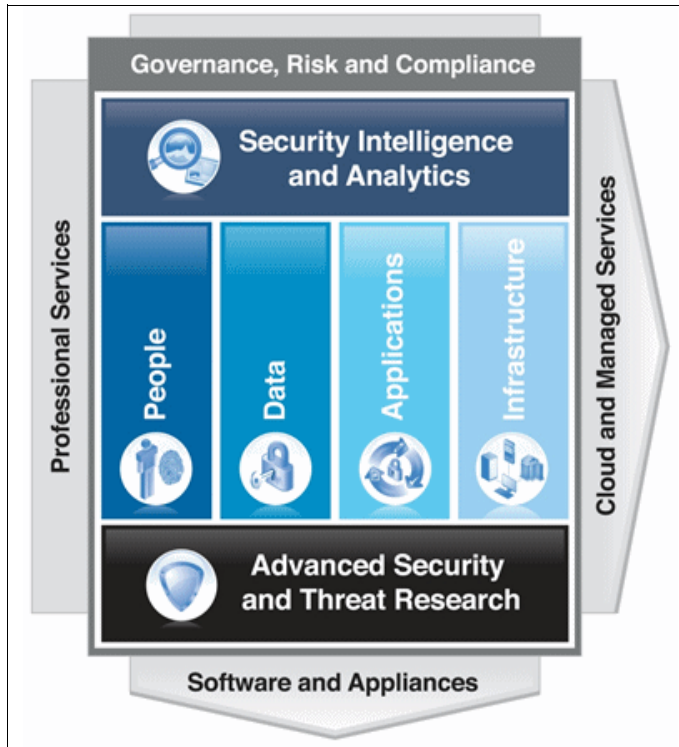


Figure 5-11 IBM Security Framework

The security landscape has grown substantially more complex. The evolution of a highly networked, data-intensive computing model has introduced significant security risks to IT infrastructures. Greater interconnectivity and system concentration exacerbates the likelihood of cascading system failures. And the increasing prevalence of social, mobile, virtualization and cloud technologies present a profound security challenge.

This increased cost and complexity make traditional data security practices obsolete. More sophisticated threats and higher risk are driving the need for more responsive, context-aware approaches to data security management. To take advantage of new marketplace opportunities while minimizing risks, companies must proactively incorporate security as a business enabler.

The average security breach can cost \$5.85M US dollars⁸ and the cost of protecting an enterprise is much more.

In addition to built-in security and cryptographic features and functions, one important aspect of securing your IT environment is to prevent and eliminate security risks. One of the important features of a secure environment is to have the capabilities to analyze security related data using powerful analytics that help early detection of security and help predict a breach on the verge of happening. Analytics is also key to modern fraud detection and prevention methods.

Security around data and multitenancy work together. Mobile security depends on access rules and endpoint protection. IBM z Systems already have encryption and secured isolation, while network security for the cloud is addressed too. Security for the Internet with SSL transactions and secure coprocessing is delivered with Crypto Express5S.

⁸ Ponemon Institute:

<http://www-935.ibm.com/services/us/en/it-services/security-services/cost-of-data-breach>

Today's social networking generation is more likely to reveal personal data to other parties in venues like social networking sites. Abusing mobility has proven to be an effective means to spread malware. Social engineering through mobile apps puts users and companies they work for in danger.

Many mobile phone makers do not rapidly deploy software patches to devices; malicious apps are often distributed through third-party app markets. Another troubling trend is that some mobile malware can collect a user's personal information for use in phishing attacks. Mobile attacks are potentially providing an entry into corporate assets.

To effectively detect and prevent security breaches, security intelligence and powerful analytics must be implemented. Such capabilities are provided by InfoSphere BigInsights and IBM Security QRadar® which also take advantage of traditional mainframe security and the capabilities of IBM z13, such as these examples:

- ▶ CPACF and Crypto Express5s
CPACF is designed to improve performance for cryptographic functions. The optional Cryptographic Coprocessor adapter (Crypto Express5) provides new virtualization capabilities and performance increase.
- ▶ SIMD allows construction of richer, complex analytics models that use SIMD to provide better accuracy of insight:
 - Allow analytics workloads to be ported from IBM Power and x86 with ease and can accelerate analytics to provide speedy business insight.
 - Increase programmer productivity of ISV and customer analytics workload development leading to rapid business insight generation for competitive advantage.
- ▶ SMT: Process more workload (throughput for IFLs).

InfoSphere BigInsights empowers enterprises of all sizes to cost effectively manage and analyze big data: the massive volume, variety, and velocity of data that consumers and businesses create every day. InfoSphere BigInsights helps increase operational efficiency by modernizing your data warehouse environment as a queryable archive, allowing you to store and analyze large volumes of multi-structured data without straining the data warehouse.

IBM QRadar Security Intelligence Platform products provide a unified architecture for integrating security information and event management (SIEM), log management, anomaly detection, incident forensics and configuration and vulnerability management. These products offer advanced threat detection, greater ease of use and lower total cost of ownership.

IBM QRadar Security Intelligence Platform products deliver these features:

- ▶ A single architecture for analyzing log, flow, vulnerability, user and asset data
- ▶ Near real-time correlation and behavioral anomaly detection to identify high-risk threats
- ▶ High-priority incident detection among billions of data points
- ▶ Full visibility into network, application and user activity
- ▶ Automated regulatory compliance with collection, correlation and reporting capabilities

Conclusion

The IBM Security Framework and the IBM Security Blue are tools to enable the security architect to understand the components and facilities needed to design a secure solution. These tools are useful for all platforms and can be applied to systems hosted on IBM z Systems and to applications which span multiple heterogeneous hardware and software platforms.

**A**

Software licensing

Software licensing options are available for the z13. These options and basic information about software licensing for the IBM z BladeCenter Extension (zBX) Model 004 environments are covered.

Software licensing considerations

The z13's IBM software portfolio includes operating system software (that is, z/OS, z/VM, z/VSE, and z/TPF) and middleware that runs on these operating systems. The portfolio also includes middleware for Linux on z Systems environments.

zBX software products are covered by the International Program License Agreement (IPLA) and additional agreements, such as the IBM International Passport Advantage® Agreement, similar to other AIX, Linux on System x, and Windows environments. POWER/VM Enterprise Edition licenses must be ordered for IBM POWER7 blades.

For the z13, two metric groups for software licensing are available from IBM, depending on the software product:

- ▶ Monthly license charge (MLC)

MLC pricing metrics have a recurring charge that applies each month. In addition to the right to use the product, the charge includes access to IBM product support during the support period. MLC metrics, in turn, include various offerings.

- ▶ International Program License Agreement (IPLA)

IPLA metrics have a single, up-front charge for an entitlement to use the product. An optional and separate annual charge, called subscription and support, entitles clients to access IBM product support during the support period. With this option, you can also receive future releases and versions at no additional charge.

For more details about software licensing, see the following resources:

- ▶ Learn about Software licensing:

http://www-01.ibm.com/software/lotus/passportadvantage/about_software_licensing.html

- ▶ Base license agreements:

<http://www-03.ibm.com/software/sla/slabd.nsf/viewbla>

- ▶ IBM z Systems Software Pricing Reference Guide:

<http://www.ibm.com/systems/z/resources/swprice/reference/index.html>

- ▶ IBM z Systems Software Pricing:

<http://www.ibm.com/systems/z/resources/swprice/index.html>

- ▶ The IBM International Passport Advantage Agreement can be downloaded from the “Learn about Software licensing” web page:

ftp://ftp.software.ibm.com/software/passportadvantage/PA_Agreements/PA_Agreement_International_English.pdf

The remainder of this appendix describes the software licensing options that are available for the z13.

Monthly License Charge (MLC) pricing metrics

MLC pricing applies to z/OS, z/VSE, or z/TPF operating systems. Any mix of z/OS, z/VM, Linux, z/VSE, and z/TPF images is allowed. Charges are based on processor capacity, which is measured in millions of service units (MSU) per hour.

Charge models

Various Workload License Charges (WLC) pricing structures support two charge models:

- ▶ Variable charges (several pricing metrics)

Variable charges apply to products such as z/OS, z/VSE, z/TPF, DB2, IMS, CICS, and WebSphere MQ. Several pricing metrics employ the following charge types:

- Full-capacity license charges

The total number of MSUs of the central processor complex (CPC) is used for charging. Full-capacity licensing is applicable when the CPC of the client is not eligible for subcapacity.

- Subcapacity license charges

Software charges are based on the utilization of the logical partitions where the product is running.

- ▶ Flat charges

Software products that are licensed under flat charges are not eligible for subcapacity pricing. There is a single charge for each CPC on the z13.

Subcapacity license charges

For eligible programs, subcapacity licensing allows software charges that are based on the measured utilization by logical partitions instead of the total number of MSUs of the CPC. Subcapacity licensing removes the dependency between the software charges and CPC (hardware) installed capacity.

The subcapacity licensed products are charged monthly based on the highest observed 4-hour rolling average utilization of the logical partitions in which the product runs (except for

products that are licensed by using the select application license charge (SALC) pricing metric). This type of charge requires measuring the utilization and reporting it to IBM.

The four-hour rolling average utilization of the logical partition can be limited by a defined capacity value on the image profile of the partition. This value activates the soft capping function of the Processor Resource/Systems Manager (PR/SM), limiting the four-hour rolling average partition utilization to the defined capacity value. Soft capping controls the maximum four-hour rolling average usage (the last four-hour average value at every five-minute interval), but does not control the maximum instantaneous partition use.

Also available is a logical partition (LPAR) group capacity limit, which sets soft capping by PR/SM for a group of logical partitions running z/OS.

Even by using the soft capping option, the use of the partition can reach up to its maximum share based on the number of logical processors and weights in the image profile. Only the four-hour rolling average utilization is tracked, allowing utilization peaks above the defined capacity value.

Some pricing metrics apply to stand-alone z Systems. Others apply to the aggregation of multiple z Systems workloads within the same Parallel Sysplex.

For more information about WLC and details of how to combine LPAR utilization, see *z/OS Planning for Workload License Charges*, SA22-7506, which you can find at this website:

http://www.ibm.com/systems/z/os/zos/bkserv/find_books.html

IBM z13

Metrics that are applicable to a stand-alone z13 include the following charges:

- ▶ Advanced Workload License Charges (AWLC)
- ▶ IBM z Systems New Application License Charges (zNALC)
- ▶ Parallel Sysplex License Charges (PSLC)

Metrics that are applicable to a z13 in an actively coupled Parallel Sysplex include the following charges:

- ▶ Advanced Workload License Charges (AWLC), when all nodes are z Systems CPCs (z13, zEC12, zBC12, z196, or z114)
- ▶ Variable Workload License Charges (VWLC), allowed only under the AWLC Transition Charges for Sysplexes when not all of the nodes are zEnterprise CPCs
- ▶ IBM z Systems New Application License Charges (zNALC)
- ▶ Parallel Sysplex License Charges (PSLC)

Advanced Workload License Charges (AWLC)

AWLC type was introduced with the IBM zEnterprise 196. These charges use the measuring and reporting mechanisms, and also the existing MSU tiers, from VWLCs, although the prices for each tier were lowered.

AWLC can be implemented in full-capacity or subcapacity mode. The AWLC applies to z/OS and z/TPF and their associated middleware products such as DB2, IMS, CICS, and IBM WebSphere MQ, and IBM Lotus® IBM Domino®.

With z13, Technology Transition Offerings are available that extend the software price and performance of the AWLC pricing metric:

- ▶ Technology Update Pricing for z13 is applicable for clients that run on a stand-alone z13 or in an aggregated Parallel Sysplex consisting exclusively of z13 servers.
- ▶ New Transition Charges for Sysplexes (TC3) are applicable when z13 and zEC12 and zBC12 are the only servers in an actively coupled Parallel Sysplex.
- ▶ Transition Charges for Sysplexes (TC2) apply when two or more servers exist in an actively coupled Parallel Sysplex consisting of one or more z13, zEC12, zBC12, z196 or z114 servers.

For more information, see the AWLC web page:

<http://www.ibm.com/systems/z/resources/swprice/mlc/awlc.html>

z Systems New Application License Charges (zNALC)

zNALC offers a reduced price for the z/OS operating system on logical partitions that run a qualified new workload application. An example includes Java language business applications that run under WebSphere Application Server for z/OS or SAP.

z/OS with zNALC provides a strategic pricing model that is available on the full range of z Systems for simplified application planning and deployment. zNALC allows for aggregation across a qualified Parallel Sysplex, which can provide a lower cost for incremental growth across new workloads that span a Parallel Sysplex.

For more information, see the zNALC web page:

<http://www.ibm.com/systems/z/resources/swprice/mlc/zna1c.html>

Midrange Workload License Charges (MWLC)

MWLC applies to z/VSE V5 and later when running on z13, zEC12, z196, System z10, and z9 servers. The exceptions are: the z10 BC and z9 BC servers at the capacity setting A01, to which zELC applies; and z114 and zBC12 where MWLC is not available.

Similar to workload license charges, MWLC can be implemented in full-capacity or subcapacity mode. An MWLC applies to z/VSE V5 and later, and several IBM middleware products for z/VSE. All other z/VSE programs continue to be priced as before.

The z/VSE pricing metric is independent of the pricing metric for other systems (for instance, z/OS) that might be running on the same server. When z/VSE is running as a guest of z/VM, z/VM V5R4 or later is required.

To report usage, the subcapacity report tool is used. One subcapacity reporting tool (SCRT) report per server is required.

For more information, see the MWLC web page:

<http://www.ibm.com/systems/z/resources/swprice/mlc/mwlc.html>

Parallel Sysplex License Charges (PSLC)

PSLC applies to a large range of mainframe servers. The list is available at this web page:

<http://www.ibm.com/systems/z/resources/swprice/reference/exhibits/hardware.html>

Although it can be applied to stand-alone CPCs, the metric provides aggregation benefits only when applied to a group of CPCs in an actively coupled Parallel Sysplex cluster according to IBM terms and conditions.

Aggregation allows charging a product that is based on the total MSU value of the systems where the product runs (as opposed to all the systems in the cluster). In an uncoupled environment, software charges are based on the MSU capacity of the machine.

For more information, see the PSLC web page:

<http://www.ibm.com/systems/z/resources/swprice/mlc/pslc.html>

z Systems International Program License Agreement (IPLA)

For z Systems, the following types of products are generally in the IPLA category:

- ▶ Data management tools
- ▶ DB2 for z/OS VUE
- ▶ CICS TS VUE V5 and CICS tools
- ▶ IMS DB VUE V12 and IMS tools
- ▶ Application development tools
- ▶ Certain WebSphere for z/OS products
- ▶ Linux middleware products
- ▶ z/VM V5 and V6

Generally, three pricing metrics apply to IPLA products for z13 and z Systems:

- ▶ Value Unit (VU):

This pricing metric applies to the IPLA products that run on z/OS. Value Unit pricing is typically based on the number of MSUs and allows for lower cost of incremental growth. Examples of eligible products are IMS Tools, CICS Tools, DB2 Tools, application development tools, and WebSphere products for z/OS.

- ▶ Engine-Based Value Unit (EBVU):

This pricing metric enables a lower cost of incremental growth with more engine-based licenses purchased. Examples of eligible products include z/VM V5 and V6, and certain z/VM middleware, which are priced based on the number of engines.

- ▶ Processor Value Unit (PVU):

This pricing metric is determined from the number of engines, under the Passport Advantage terms and conditions. Most Linux middleware is also priced based on the number of engines.

For more information, see the IPLA web page:

<http://www.ibm.com/systems/z/resources/swprice/zipla/index.html>

zBX licensed software

The software licensing for the zBX select System x and POWER7 blades and DataPower X150z follows the same rules as licensing for blades installed outside of zBX.

POWER/VM Enterprise Edition *must be* licensed for POWER7 blades at the time of ordering the blades.

The hypervisor for the select System x blades for zBX is provided as part of the IBM z Unified Resource Manager (zManager).

IBM z Unified Resource Manager

The z Unified Resource Manager is available through z13, zEC12, and zBC12 *hardware* features, either ordered with the system or later. No separate software licensing is required.

**B**

Channel options

This appendix describes all channel attributes, the required cable types, the maximum unrepeated distance, and the bit rate for the z13.

For all optical links, the connector type is LC Duplex except for the 12xIFB and the ICA connection, which are established with Multifiber Push-On (MPO) connectors. The MPO connector of the 12xIFB connection has one row of twelve fibers, while the MPO connector of the ICA connection has two rows of twelve fibers. The electrical Ethernet cable for the Open Systems Adapter (OSA) connectivity is connected through an RJ45 jack.

Table B-1 on page 168 lists the attributes of the channel options that are supported on z13.

Table B-1 z13 channel feature support

Channel feature	Feature codes	Bit rate in Gbps (or stated)	Cable type	Maximum unrepeated distance ^a	Ordering information
Fiber Connection (FICON)					
FICON Express16S 10KM LX	0418	4, 8, or 16	SM 9 μ m	10 km (6.2 miles)	New build
FICON Express16S SX	0419	4, 8, or 16	OM2, OM3	See Table B-2 on page 169.	New build
FICON Express8S 10KM LX	0409	2, 4, or 8	SM 9 μ m	10 km (6.2 miles)	New build
FICON Express8 10KM LX	3325				Carry forward
FICON Express8S SX	0410	2, 4, or 8	OM1, OM2, OM3	See Table B-2 on page 169.	New build
FICON Express8 SX	3326				Carry forward
Open Systems Adapter (OSA)					
OSA-Express5S 10 GbE LR	0415	10	SM 9 μ m	10 km (6.2 miles)	New build
OSA-Express4S 10 GbE LR	0406				Carry forward
OSA-Express5S 10 GbE SR	0416	10	MM 62.5 μ m MM 50 μ m	33 m (200) 300 m (2000) 82 m (500)	New build
OSA-Express4S 10 GbE SR	0407				Carry forward
OSA-Express5S GbE LX	0413	1	SM 9 μ m	5 km (3.1 miles)	New build
OSA-Express4S GbE LX	0404				Carry forward
OSA-Express5S GbE SX	0414	1	MM 62.5 μ m MM 50 μ m	220 m (166) 275 m (200) 550 m (500)	New build
OSA-Express4S GbE SX	0405				Carry forward
OSA-Express5S 1000BASE-T	0417	100 or 1000 Mbps	Cat 5, Cat 6 copper		New build
OSA-Express4S 1000BASE-T	0408	10, 100, or 1000 Mbps			Carry forward
10GbE Remote Direct Memory Access (RDMA) over Converged Ethernet (RoCE) Express	0411	10	OM3	300 m	New build
Parallel Sysplex					
ICA (PCIe-O SR)	0172	8 GBps	OM4	150 m	New build
			OM3	100 m	New build
HCA3-O (12x IFB)	0171	6 GBps	OM3	150 m	New build
HCA3-O LR (1x IFB)	0170	2.5 or 5 Gbps	SM 9 μ m	10 km (6.2 miles)	New build
IC	N/A		N/A	N/A	N/A

Channel feature	Feature codes	Bit rate in Gbps (or stated)	Cable type	Maximum unrepeated distance ^a	Ordering information
Cryptography					
Crypto Express5S	0890	N/A	N/A	N/A	New build
Crypto Express4S	0865	N/A	N/A	N/A	Carry forward
Flash Express	0402	N/A	N/A	N/A	New build
zEDC Express	0420	N/A	N/A	N/A	New build

a. Where applicable, the minimum fiber bandwidth distance in MHz-km for multi-mode fiber optic links is included in parentheses.

Table B-2 shows the maximum unrepeated distances for FICON SX features.

Table B-2 Maximum unrepeated distance for FICON SX features

Cable type\bit rate	1 Gbps	2 Gbps	4 Gbps	8 Gbps	16 Gbps
OM1 (62.5 μm at 200 MHz·km)	300 meters	150 meters	70 meters	21 meters	N/A
	984 feet	492 feet	230 feet	69 feet	N/A
OM2 (50 μm at 500 MHz·km)	500 meters	300 meters	150 meters	50 meters	35 meters
	1640 feet	984 feet	492 feet	164 feet	115 feet
OM3 (50 μm at 2000 MHz·km)	860 meters	500 meters	380 meters	150 meters	100 meters
	2822 feet	1640 feet	1247 feet	492 feet	328 feet
OM4 (50 μm at 4700MHz·km)	N/A ^a	N/A	400 meters	190 meters	125 meters
	N/A	N/A	1312 feet	623 feet	410 feet

a. N/A - Data not available

Abbreviations and acronyms

This abbreviations and acronyms file is optional.

Use this file by adding names and descriptions to it. Sort these names: **highlight rows** → **Table** → **Sort** → **Sort By: Column 1** → **Sort**
or
optionally add names and descriptions to the Index file instead of this file by indexing the first use of an abbreviation or acronym: **highlight text** → **Special** → **Marker** → **Index** → **New Marker**

abbreviation1	Description1
abbreviation2	Description2
IBM	International Business Machines Corporation
ITSO	International Technical Support Organization
abbreviation3	Description3
abbreviation4	Description4

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- ▶ *IBM z13 Technical Guide, SG24-8251*
- ▶ *IBM z Systems Connectivity Handbook, SG24-5444*
- ▶ *IBM z13 Configuration Setup, SG24-8260*
- ▶ *Securing your Mobile Mainframe, REDP-5176*
- ▶ *z Systems Simultaneous Multithreading Revolution, REDP-5144*
- ▶ *SIMD Business Analytics Acceleration on z Systems, REDP- 5145*
- ▶ *Enhancing Value to Existing and Future Workloads with IBM z13, REDP-5135*
- ▶ *z/OS Infrastructure Optimization using Large Memory, REDP-5146*
- ▶ *Get More Out of Your IT Infrastructure With IBM z13 I/O Enhancements, REDP-5134*
- ▶ *IBM z Systems Functional Matrix, REDP-5157*

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

- ▶ *Installation Manual for Physical Planning, GC28-6938*
- ▶ *PR/SM Planning Guide, SB10-7162*
- ▶ *Capacity on Demand User's Guide, SC28-6943*

Online resources

These websites are also relevant as further information sources:

- ▶ IBM Resource Link
<https://www-304.ibm.com/servers/resourceLink/hom03010.nsf?OpenDatabase&login>
- ▶ IBM Offering Information
http://www-01.ibm.com/common/ssi/index.wss?request_locale=en

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(->Hide:)>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine:fm still open and **File>Import>Formats** the Conditional Text Settings (ONLY!) to the book files.

Draft Document for Review February 6, 2015 12:27 pm

8250spine.fm 175



IBM z13 Technical Introduction

SG24-8250-00

ISBN



(1.5" spine)

1.5" <-> 1.998"

789 <-> 1051 pages



IBM z13 Technical Introduction

SG24-8250-00

ISBN



(1.0" spine)

0.875" <-> 1.498"

460 <-> 788 pages

Redbooks

IBM z13 Technical Introduction

SG24-8250-00

ISBN



(0.5" spine)

0.475" <-> 0.873"

250 <-> 459 pages

Redbooks

IBM z13 Technical Introduction

(0.2" spine)

0.17" <-> 0.473"

90 <-> 249 pages

(0.1" spine)
0.1" <-> 0.169"
53 <-> 89 pages

To determine the spine width of a book, you divide the paper PPI into the number of pages in the book. An example is a 250 page book using Plainfield opaque 50# smooth which has a PPI of 526. Divided 250 by 526 which equals a spine width of .4752". In this case, you would use the .5" spine. Now select the Spine width for the book and hide the others: **Special>Conditional Text>Show/Hide>SpineSize(->Hide:)>Set** . Move the changed Conditional text settings to all files in your book by opening the book file with the spine:fm still open and **File>Import>Formats the Conditional Text Settings (ONLY)** to the book files.

Draft Document for Review February 6, 2015 12:27 pm

8250spine.fm 176



IBM z13 Technical Introduction

SG24-8250-00

ISBN

(2.5" spine)
2.5" <-> mmm.n"
1315 <-> mmm pages



IBM z13 Technical Introduction

SG24-8250-00

ISBN

(2.0" spine)
2.0" <-> 2.498"
1052 <-> 1314 pages





SG24-8250-00

ISBN

Printed in U.S.A.

Get connected

