Redbooks

ibm.com/redbooks

IBM z14 Model ZR1 Technical Introduction

Bill White

Hervey Kamga

Octavian Lascu

Frank Packheiser

Martijn Raave

John Troy

IBM Z

IBM

Redbooks

International Technical Support Organization

IBM z14 Model ZR1 Technical Introduction

October 2018

SG24-8550-01

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

Second Edition (October 2018)
This edition applies to IBM Z platform IBM z14™ Model ZR1 (Machine type 3907).

Contents

Notices

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

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, MD-NC119, Armonk, NY 10504-1785, US

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

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

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

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

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

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

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

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

COPYRIGHT LICENSE:

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

Trademarks

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

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

CICS®
Concert™
Db2®
DB2®
DS8000®
FICON®
GDPS®
Geographically Dispersed Parallel Sysplex™
HiperSockets™
HyperSwap®
IBM®
IBM MobileFirst™
IBM z®
IBM z Systems®
IBM z13s™
IMS™
Language Environment®
Parallel Sysplex®
Passport Advantage®
PR/SM™
Processor Resource/Systems Manager™
Rational®
Rational Team Concert™
Redbooks®
Redbooks (logo)  ®
Resource Link®
RMF™
System Storage®
System z®
System z10®
System z9®
WebSphere®
z Systems®
z/Architecture®
z/OS®
z/VM®
z/VSE®
z10™
z13®
z13s™
z9®
zEnterprise®

The following terms are trademarks of other companies:

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

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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

Preface

This IBM® Redbooks® publication introduces the latest member of the IBM Z platform, the IBM z14 Model ZR1 (Machine Type 3907). It includes information about the Z environment and how it helps integrate data and transactions more securely, and provides insight for faster and more accurate business decisions.

The z14 ZR1 is a state-of-the-art data and transaction system that delivers advanced capabilities, which are vital to any digital transformation. The z14 ZR1 is designed for enhanced modularity, which is in an industry standard footprint. This system excels at the following tasks:

- Securing data with pervasive encryption
- Transforming a transactional platform into a data powerhouse
- Getting more out of the platform with IT Operational Analytics
- Providing resilience towards zero downtime
- Accelerating digital transformation with agile service delivery
- Revolutionizing business processes
- Mixing open source and Z technologies

This book explains how this system uses new innovations and traditional Z strengths to satisfy growing demand for cloud, analytics, and open source technologies. With the z14 ZR1 as the base, applications can run in a trusted, reliable, and secure environment that improves operations and lessens business risk.

Authors

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

Bill White is an IBM Redbooks Project Leader and Senior Networking and Connectivity Specialist at IBM Redbooks, Poughkeepsie Center.

Hervey Kamga is an IBM z® Systems Product Engineer with the EMEA I/O Connectivity Team in Montpellier, France. Before his current role, he was a Support Engineer and Engineer On Site for 13 years with Sun MicroSystems and Oracle in EMEA. Hervey's areas of expertise include Oracle Solaris (operating system and hardware products), Virtualization (VMware and virtualBox), Linux (Ubuntu), and IBM z Systems® I/O features and protocols (IBM FICON® and OSA).

Octavian Lascu is an IBM Redbooks Project Leader and a Senior IT Consultant for IBM Romania with over 20 years of experience. He specializes in designing, implementing, 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 materials for and taught over 50 workshops for technical audiences around the world. He has written several IBM publications.

Frank Packheiser is a Senior IT Specialist at the Field Technical Sales Support office in Germany. He has 26 years of experience in IBM z Systems®. Frank has worked for 10 years in 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, Frank supported clients in Middle East/North Africa (MENA) as a zIT Architect. In addition to co-authoring several IBM Redbooks publications since 1999, he has been an official ITSO presenter at ITSO workshops for the last four years.

Martijn Raave is a certified IBM Z & LinuxONE Client Architect for IBM Netherlands. Over a period of 20 years, his professional career has revolved around the mainframe platform. Before joining IBM in 2005, he worked for a large Dutch client as a systems programmer with expertise in the areas of IBM z/OS®, IBM Parallel Sysplex®, and hardware. Later, he decided to explore other aspects of the mainframe environment within IBM, and joined STG as a Client Technical Specialist, supporting several Dutch IBM System z® clients, IBM Business Partners, and IBM sales representatives. In 2016, the LinuxONE brand was added to his portfolio and he is now working as a Client Architect with a broad range of customers on different engagements. He is also a board member of Guide Share Europe (GSE) Netherlands.

John Troy is a z Systems and Storage hardware National Top Gun in the northeast area of the United States. He has 36 years of experience in the service field. His areas of expertise include technical and customer support for z Systems server and high-end storage systems. John has been a z Systems hardware technical support course designer, developer, and instructor for the last seven generations of IBM high-end servers.

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

Todd Arnold
Rick Brown
Les Geer
Michael Groetzner
Parwez Hamid
Gary King
Jeff Kubala
Barbara Sannerud
Philip Sciuto
Christine Smith
Rhonda Sundlof
Dave Surman
John Torok
Marna Walle
Charles Webb
Barara Weiler
Harry Yudenfriend
Monika Zimmermann
Uwe Zumpe

Thanks to the team that wrote the *IBM z14 Technical Introduction*, SG24-8450:

Hervey Kamga
Michal Kordyzon
Frank Packheiser
John Troy
Esra Ufacik
Bo Xu

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 two - 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:

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:

  ibm.com/redbooks

- Send your comments in an email:

  redbooks@us.ibm.com

- Mail your comments:

  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

The foundation of a trusted digital transformation

The effects of today's digital transformation cannot be ignored by any business, regardless of the size. Demand is growing for new services and individualized client experiences across all industries.

The need to efficiently build and deploy highly secure cloud-based services as part of a delivery strategy is essential. More than ever, deriving real-time insights and increasing value through data and complex analytics is the mechanism by which businesses gain competitive advantage.

Today, IT must provide an infrastructure in which reliable data and transactions are always protected, available, and delivered with speed. At a minimum, the infrastructure should meet the following needs:

- Provide consistently high performance to ensure integrity of service while operating at high volumes
- Allow for rapid application development and delivery by using open and industry standard technologies
- Deliver real-time insights within the current business transaction
- Encrypt all data to protect against insider and external threats

The latest member of the IBM Z family, the z14 Model ZR1, features a tried-and-true architecture to support your digital transformation, create a strong and reliable cloud infrastructure, and expose back-end services through secure APIs. The z14 ZR1 can also streamline your ability to integrate disparate data center systems and create a single, cohesive IT infrastructure.

The z14 ZR1 can help your business make consistently optimal decisions, gain operational data insights so you get the most value from your IT investment, and fully protect your data with encryption, while facilitating regulatory compliance.

A trusted digital transformation can be built on the z14 ZR1. It is a premier system for enabling a secure infrastructure and is designed for data that is serving in a cognitive era. The z14 ZR1 offers a high-value architecture that supports an open and connected world.

This chapter includes the following topics:

1.1 The z14 ZR1: A secure and agile platform

The z14 ZR1 offers a fast, scalable, securable, and adaptable system. Compared to its predecessor platforms, the z14 ZR1 provides more of what you need to satisfy the following growing IT demands of today:

- Compute power for increased throughput
- Large-scale memory to process data faster
- Industry-unique cache design to optimize performance
- Accelerated I/O bandwidth to process massive amounts of data
- Data compression to economically store and process information
- High-speed cryptographic operations to help secure transactions and data

From the hundreds of microprocessors to the software stack, the z14 ZR1 (as shown in Figure 1-1) is built to quickly respond to change. This evolution of the Z platform embodies a proven infrastructure that is designed from the ground up for data and transactions.

Figure 1-1 The IBM z14 ZR1

The z14 ZR1 meets the needs of the digital era by using the following techniques:

- Building and deploying highly secure cloud-based services
- Securing data with pervasive encryption
- Transforming a transactional platform into a data powerhouse
- Getting more out of the platform with IT Operational Analytics
- Providing resilience towards zero downtime
- Accelerating digital transformation with agile service delivery
- Revolutionizing business processes
- Mixing open source and Z technologies

1.1.1 Building and deploying highly secure cloud-based services

The z14 Model ZR1 provides a new industry-standard infrastructure to meet the demands of entry-level businesses. It allows for easier and quicker delivery of secure cloud-based[1] services.

The z14 ZR1 excels with integrated security features that are built into the hardware, firmware, and operating systems. The built-in features range from storage protection keys and workload isolation to encryption co-processors, and more. The z14 ZR1 is also highly virtualized, with the goal of maximizing the utilization of computing resources, while lowering the total number of resources and costs that are needed to run cloud-base services.

With up to 30 configurable cores, the z14 ZR1 features performance and scaling advantages over older generations, and double the memory capacity (8 TB) than the IBM z13s™. In addition, the new 16U Reserved rack space feature allows you to populate available 19-inch standard rack space with your choice of network and storage switches, and storage devices to create all-in-one secure solutions for building and deploying cloud-based services within a single rack.

1.1.2 Securing data with pervasive encryption

IBM Z pervasive encryption provides the comprehensive data protection that your business and customers demand. By placing the security controls on the data, the solution creates an envelope of protection around the data on Z. For example, Z pervasive encryption helps protect the *at-rest* and *in-flight* data that is on your Z infrastructure. Also, centralized, policy-based data encryption controls significantly reduce the costs that are associated with data security and regulatory compliance, including the new General Data Protection Regulations (GDPR).[2]

IBM Z pervasive encryption implements this comprehensive security with your ongoing operations in mind. Therefore, it does not require you to make any application changes. It also can be implemented by using policy-based controls with low overhead. These capabilities can slash the costs that are associated with data security and compliance.

The Central Processor Assist for Cryptographic Function (CPACF), which is standard on every core, supports pervasive encryption and provides hardware acceleration for encryption operations. CPACF encryption rates for similar modes and data sizes on z14 ZR1 are up to 6x faster than z13s.

The new Crypto Express6S also receives a performance boost on the z14 ZR1. Combined, these enhancements perform encryption more efficiently than on earlier Z platforms. Also, SSL handshake throughput on the z14 ZR1 with Crypto Express6S is up to 2X higher than SSL handshake throughput on the z13s with Crypto Express5S.

1.1.3 Transforming a transactional platform into a data powerhouse

Currently, data is one of the most valuable resources an organization possesses. Deriving insights from that data to drive optimal business decisions becomes one of the biggest challenges. To maximize the value of that resource, your enterprise might need to integrate more external data sources to extract hidden insights.

---

[1] For more information about cloud-based options, see 4.2.2, “IBM Z based clouds” on page 68.
[2] For more information about GDPR, see the Clear the Path to the GPDR website.

For decades, clients typically copied business-critical data from their mainframe transactional systems to other platforms, or even data lakes, to perform sophisticated analytics. This process was inefficient, expensive, time-consuming, and introduced risk on lower-security platforms and data latency.

By accessing your enterprise data in-place with minimal data duplication or movement, you can minimize the cost and complexity of analytics. You can also make enterprise data highly accessible to analytics applications and tools by integrating transactional and analytics processing, and protect sensitive data by keeping it within the secure Z platform.

Because of the advanced infrastructure, with double the cache density on each chip and up to 8 TB memory, the z14 ZR1 can support the following state-of-the-art cognitive solutions:

- Open Data Analytics for z/OS

  This open source, in-place analytics solution for z/OS simplifies big data analysis. Open Data Analytics for z/OS gives developers and data scientists the ability to analyze business-critical z/OS data in place, with no data movement. Open Data Analytics for z/OS can also provide a federated view by accessing and analyzing distributed and local data.

- IBM Machine Learning for z/OS

  This comprehensive solution manages the entire machine learning workflow, beginning with quick ingestion and transformation of Z data where it is stored. The solution then securely creates, deploys, and manages high-quality, self-learning behavior models to help you extract hidden insights that more accurately anticipate organizational needs.

- The IBM DB2® Analytics Accelerator for z/OS

  This high-performance appliance transforms your mainframe into a highly efficient transactional and analytics-processing environment. It supports the full lifecycle of a real-time analytics solution on a single system that integrates transactional data, historical data, and predictive analytics.

1.1.4 Getting more out of the platform with IT Operational Analytics

Today, demands for 24 x 7 high-performance operations continue to rise. At the same time, allowed service windows shrink and are much less frequent. Increasing system complexity makes planning, maintaining, and troubleshooting more difficult and time-consuming. IT operations analytics represent a possible solution to this challenge.

The z14 ZR1 provides the infrastructure to host real-time analytics tools so you can clearly see your operating environment and then, maximize operational efficiencies to help reduce costs.

IBM designed IBM Operations Analytics for z (IOAz), IBM Common Data Provider for z, and IBM z Operational Insights to ensure that your Z operates at peak performance. To get the most out of your system, Operations Analytics for z provides deep insights that are based on IBM's industry-leading expertise into your Z operational data.

Note: IBM z Advanced Workload Analysis Reporter (IBM zAware), which was delivered as a firmware appliance that was running on a dedicated LPAR, is now part of IOAz.

1.1.5 Providing resilience towards zero downtime

Every second of downtime (planned or unplanned) can mean lost revenue. It is crucial to keep critical systems running 24 x 7, and to rapidly recover from an outage and resume critical business operations.

The Z in IBM Z stands for zero downtime, and the z14 ZR1 features the same proven reliability for which all IBM Z platforms are known. As with previous generations, the z14 ZR1 provides technology and services to help identify and remove any sources of outages.

Also, with platform-level redundancy, the z14 ZR1 is designed to handle failures while maintaining user access. Components can be repaired, maintenance performed, and products migrated with minimal affect on business. Some capabilities, such as capacity-on-demand, automatically turn components on and off based on current needs.

Although the z14 ZR1 platform is highly robust and even more so in a sysplex[3] environment, implementing the IBM Geographically Dispersed Parallel Sysplex™ (IBM GDPS®) family of solutions improves resilience in cases of unplanned failures, power outages, fire, or human error. The GDPS family of solutions provides more tools to ensure Z availability, and mask or significantly reduce the effects of critical component outages or failures.

By using IBM HyperSwap® technology, I/O traffic can be seamlessly routed from disk subsystems that cannot service the I/O request to a second disk subsystem that can. Also, the GDPS/Active-Active solution can route workload from a server location that is experiencing problems to a second location that is operating well.

1.1.6 Accelerating digital transformation with agile service delivery

An effective DevOps solution breaks down development silos, unifies infrastructure platforms, and enables ongoing deliveries. z14 ZR1 provides the scalable and secure infrastructure for enterprises that must rapidly create and deliver critical applications, while meeting agreed-on levels for quality, availability, regulatory compliance, and end-use expectations.

IBM DevOps for Z solutions operate from application understanding through deployment and management. In addition, DevOps for Z solutions gives you a single, cost-effective toolset to maintain and modernize valuable applications on Z and distributed platforms.

For example, Application Discovery and Delivery Intelligence helps development teams understand application interdependencies, complexity, and quality across platforms, environments, and languages. This ability gives your teams an edge in identifying potential API candidates, and provides insight about maintainability and complexity. As a result, the candidate API quality rises, and the user experience also improves.

In addition, the z14 ZR1 provides the infrastructure to support the mission-critical workloads of cloud services. The new high-performance processors, large memory, and enhanced access to data enable the z14 ZR1 to integrate business transactions, operational data, and analytics into a single workflow.

The IBM z14 ZR1 is designed as a strategic asset to power the API economy.[4] The use of the API economy demands fortified clouds, which can be open, private, public, and hybrid. The z14 ZR1 gives you the hardware platform that is necessary to support those cloud environments.

[3] Sysplex is a system clustering technique for high availability. For more information, see “High availability with parallel sysplex” on page 77.

[4] For more information about the API economy, see the Reach new customers with the API economy website.

For Linux assets, Z platforms are optimized for open source software, enhanced scalability, and sharing, while focusing on business continuity to support cloud. For traditional z/OS-based assets, Z offerings provide intuitive tools to help developers speed Representational State Transfer (RESTful) API development.

No matter which asset class you choose, the z14 ZR1 allows developers to incorporate z/OS business-critical data and transactions into their mobile and cloud services without needing to understand z/OS subsystems.

1.1.7 Revolutionizing business processes

Blockchain is poised to revolutionize how industries do business. It is a technology for a new generation of transactional applications that establishes trust, accountability, and transparency while streamlining business processes.

In a blockchain network, members can access a distributed, shared ledger that is *cryptographically* secure, updated by consensus, and becomes an immutable, indelible record of all *transactions*. The ledger functions as a single source of “truth”. Considering that blockchain is all about increasing trust in *business transactions*, it makes perfect sense to run blockchain for business on Z.

Depending on your business or regulatory policies, you can choose an on-premises installation that is supported by IBM-certified Docker images that are running on Linux on IBM Z, or the IBM Blockchain Platform service plan. IBM Blockchain Platform is a fully managed blockchain service that is running in the IBM Cloud. It delivers a secure, isolated compute environment that is ideally suited for workloads with sensitive data.

1.1.8 Mixing open source and Z technologies

The correct blend and balance of open source technologies, ISV tools, and IT platform is key to enable businesses and organizations to deliver change at a much quicker pace. To this end, IBM created a system of clients, Business Partners, and ISVs who are engaging in an open source development community to bring the most important and sought-after foundational open source technologies to its IT platforms. In addition, IBM is a member of many open-standard organizations and software governance consortia that help to shape the future of open source software.

The combination of a robust and securable hardware platform with the power of a Linux distribution can optimize the building, testing, and deploying of modern applications. It also can accommodate scale-out clusters and scalable cloud environments.

The z14 ZR1 provides a secure, massive capacity Linux platform that can be deployed as stand-alone, or side-by-side with other Z operating systems on a single physical platform. Therefore, you can easily integrate Linux workloads on the z14 ZR1 with solutions that benefit from data and applications being tightly collocated with fast internal communication and improved availability.

With the z14 ZR1, Linux on IBM Z gives you the performance and vertical scale that you need to meet the demands of your digital enterprise while controlling server sprawl costs. Combined with the integration benefits, Linux on IBM Z on the z14 ZR1 allows you to deploy innovative new services or cognitive analytics and consolidate x86 workloads.

In addition, deploying Linux on the z14 ZR1 can benefit your bottom line. Compared to virtualized x86 alternatives and public cloud solutions, the lower costs for administration and management, software licensing, business continuity, floor space, and new power technology can reduce your total cost of ownership.

1.2 z14 ZR1 technical description

When compared to its predecessor (IBM z13s™), the IBM z14 ZR1 offers several improvements, such as faster, more efficient, and redesigned high-frequency chips, more granularity options, better availability, faster encryption, and enhanced on-demand options.

1.2.1 Technical highlights

The z14 ZR1 is a highly scalable symmetric multiprocessor (SMP) system. The architecture ensures continuity and upgradeability from the previous z13s. It features one model with four features: Max4, Max12, Max24, and Max30. The z14 ZR1 is housed in an industry-standard 19-inch rack that can be easily installed in any data center.

The main technical enhancements in the z14 ZR1 over its predecessor platforms are listed in Table 1-1.

Table 1-1 Technical highlights in the z14 ZR1

Improved single processor performance and greater total system capacity. The IBM z/Architecture® ensures continuity and upgradeability from previous models. Up to 6 CPs and 156 subcapacity settings.	Up to 30 characterizable processor units. A 10% uni-processor performance improvement over z13s.
Multi-core, single-chip modules that help improve the execution of processor-intensive workloads.	4.5 GHz (14 nm FINFET Silicon-On-Insulator (SOI))
More real memory per system, which ensures high availability in the memory subsystem through use of proven redundant array of independent memory (RAIM) technology.	Up to 8 TB of addressable real memory per system.
A large, fixed hardware system area (HSA) that is managed separately from client-purchased memory.	Fixed HSA is 64 GB.
Proven technology (fifth-generation high frequency and fourth-generation out-of-order design) with a single-instruction, multiple-data (SIMD) processor that increases parallelism to accelerate analytics processing. In addition, simultaneous multithreading (SMT) increases processing efficiency and throughput and raises the number of instructions in flight.	
Processor cache structure improvements and larger cache sizes to help with more of today's demanding production workloads. The z14 ZR1 offers the following levels of cache: ► First-level cache (L1 private): 128 KB for instructions, 128 KB for data ► Second-level cache (L2): 2 MB for instructions and 4 MB for data ► Third-level cache (L3): 128 MB ► Fourth-level cache (L4): 672 MB	
Improved cryptographic functions and performance, which is achieved by having one dedicated cryptographic co-processor per processor unit.	
IBM zHyperLink Express is a short distance, I/O adapter that is designed for up to 5x lower latency than High Performance FICON for read requests.	

The channel subsystem is built for I/O resilience. The number of logical channel subsystems (LCSS), subchannel sets, and I/O devices are consistent with its predecessor platform, as is the number of logical partitions (LPARs).	► Three LCSS ► 40 LPARs ► Three subchannel sets ► 32,000 I/O devices per channel

You can compare the z14 ZR1 to the previous two IBM Z generations by using the Compare IBM Z mainframes tool.

To ensure a balanced and highly available system, the z14 ZR1 includes the following other features and functions:

- Enhanced LPAR resource allocation algorithms for processor units and memory (8 TB per LPAR).
- IBM Virtual Flash Memory (VFM) can be used to handle paging workload spikes and can improve system availability. The VFM feature (0614) is the replacement for the Flash Express features (0402 and 0403), which were available on the zBC12 and z13s.
- Next generation Crypto Express6S feature supports up to 40 domains.
- New CMPSC with Huffman Coding compression for faster expansion algorithms and reduced overhead.
- Guarded Storage Facility for improved Java performance by reducing program pauses during Java Garbage Collection.
- The 25GbE RoCE Express2 provides a technology refresh for RDMA over Converged Ethernet (RoCE) on IBM Z. This feature provides increased networking performance to take advantage of higher speed processors.
- Coupling Express Long Reach (LR) for coupling links that must extend up to 10 km (6.21 miles).
- Dynamic I/O for a standalone Coupling Facility (CF) is a means for activating a new or changed I/O configuration without disruption.
- zHyperLink Express with IBM DS8880 for extremely low latency.
- Next generation FICON Express16S+ provides benefits for workloads through improvements in the base technology.
- OSA-Express7S 25GbE meets increased networking performance demands that are driven by high speed processors and faster network-attached devices.
- Secure Service Container[5] to build and host highly secure virtual appliances.
- Support for ASHRAE Class A3 data center operating environments.

For more information about the z14 ZR1, see Chapter 2, “IBM z14 ZR1 hardware overview” on page 17.

For more information about IBM z14 ZR1 functions and features, see *IBM z14 Model ZR1 Technical Guide,* SG24-8651.

1.2.2 Storage connectivity

Storage connectivity is provided on the z14 ZR1 by IBM Fibre Connection (FICON) and the IBM zHyperLink Express features.

[5] For more information, see “Secure Service Container” on page 70.

FICON Express

FICON Express features follow the established Fibre Channel (FC) standards to support data storage and access requirements, along with the latest FC technology in storage and access devices. FICON Express features support the following protocols:

- Native FICON

  This enhanced protocol (over FC) provides for communication across channels, channel-to-channel (CTC) connectivity, and with FICON devices, such as disks, tapes, and printers. It is used in z/OS, IBM z/VM®, IBM z/VSE®, z/TPF, and Linux on IBM Z and the KVM hypervisor[6] environments.

- Fibre Channel Protocol (FCP)

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

FICON Express16S+ features provide significant improvements in throughput and response time for performance-critical middleware, and to shrink the batch window that is required to accommodate I/O-bound batch work. FICON Express16S+ features are implemented by using PCIe cards. They offer better port granularity and improved capabilities over the previous FICON Express features by using enhanced adapter technology. FICON Express16S+ features support a link data rate of 16 Gbps (4, 8, or 16 Gbps auto-negotiate), and is the preferred technology for new systems.

zHyperLink Express

zHyperLink was created to provide fast access to data by way of extremely low latency connections between the Z platform and storage.

The zHyperLink Express feature uses a new adapter to allow you to make synchronous requests for data that is in the storage cache of the IBM DS8880. This process is done by directly connecting the zHyperLink Express port in the z14 ZR1 to an I/O Bay port of the DS8880. This short distance of up to 150 m (492.12 feet) direct connection is intended to speed up DB2 for z/OS *blocking read* requests.

zHyperLink can improve application response time by up to 50% without requiring application changes.

Important: The zHyperLink channels work with FICON channels; they do *not* replace them. Both adapters are required. zHyperLink is specifically designed for low latency reads and writes, such as with IBM Db2® logging.

For more information about the available FICON Express and zHyperLink Express features, see 3.3, “Storage connectivity” on page 36.

1.2.3 Network connectivity

The z14 ZR1 is a fully virtualized platform that can support many system images at once. Therefore, network connectivity covers the connections between the platform and external networks with Open Systems Adapter-Express (OSA-Express) and RoCE Express features. It also supports specialized internal connections for intra-system communication through IBM HiperSockets™ and Shared Memory Communications–Direct Memory Access (SMC-D).

[6] For more information, see 5.2.6, “KVM hypervisor” on page 92.

OSA-Express

The OSA-Express features provide local area network (LAN) connectivity and comply with IEEE standards. In addition, OSA-Express features assume several functions of the TCP/IP stack that normally are performed by the processor unit, which allows significant performance benefits by offloading processing from the operating system.

OSA-Express6S features that were introduced with the z14 ZR1 is a technology refresh. They continue to support 1000BASE-T Ethernet for copper environments and 10 Gigabit Ethernet (10GbE) and Gigabit Ethernet fiber optic (single-mode and multimode) environments. The OSA-Express7S 25GbE feature was added to support higher throughput by way of 25GbE network infrastructures.

HiperSockets

IBM HiperSockets is an integrated function of the Z platforms that supplies attachments to up to 32 high-speed virtual LANs with minimal system and network overhead.

HiperSockets is a function of the Licensed Internal Code (LIC). It provides LAN connectivity across multiple system images on the same Z platform by performing memory-to-memory data transfers in a secure way. The HiperSockets function eliminates the use of I/O subsystem operations and having to traverse an external network connection to communicate between logical partitions in the same Z platform. In this way, HiperSockets can help with server consolidation by connecting virtual servers and simplifying the enterprise network.

RoCE Express

The 25GbE and 10GbE RoCE Express2 features use Remote Direct Memory Access (RDMA) over Converged Ethernet (RoCE) to provide fast memory-to-memory communications between two Z platforms.

Those features are designed to help reduce consumption of CPU resources for applications that use the TCP/IP stack, such as IBM WebSphere® that accesses an IBM Db2 database. They can also help reduce network latency with memory-to-memory transfers by using Shared Memory Communications over RDMA (SMC-R) in z/OS V2R1 or later.

With SMC-R, you can transfer huge amounts of data quickly, at low latency. SMC-R is completely transparent to the application and requires no code changes, which enables rapid time to value.

SMC-D

The z14 ZR1 also uses the communications protocol called Shared Memory Communications-Direct Memory Access (SMC-D). SMC-D is similar to SMC-R, but is used for communications *within* the same Z platform. SMC-D optimizes operating systems communications in a way that is transparent to socket applications. It also reduces the CPU cost of TCP/IP processing in the data path, which enables highly efficient and application-transparent communications.

SMC-D requires no extra physical resources, such as RoCE Express features, PCIe bandwidth, ports, I/O slots, network resources, or Ethernet switches. Instead, SMC-D uses LPAR-to-LPAR communication through HiperSockets or an OSA-Express feature for establishing the initial connection.

For more information about the available network connectivity features, see 3.4, “Network connectivity” on page 40.

1.2.4 Cryptography

z14 ZR1 provides two major groups of cryptographic functions: CPACF and Crypto Express6S.

CPACF is a high-performance, low-latency co-processor that performs symmetric key encryption and calculates message digests (hashes) in hardware. Supported algorithms are AES, DES/TDES, SHA-1, SHA-2, and SHA-3. Latency of the CPACF in the z14 ZR1 is significantly less when compared to the CPACF in the z13s. The z14 ZR1 offers up to 6x better performance than the z13s.

The tamper-sensing and tamper-responding Crypto Express6S features provide acceleration for high-performance cryptographic operations and support up to 85 domains. This specialized hardware performs AES, DES/TDES, RSA, Elliptic Curve (ECC), SHA-1, and SHA-2, and other cryptographic operations. It also supports specialized high-level cryptographic APIs and functions, including those APIs and functions that are required in the banking industry. Crypto Express6S features are designed to meet the FIPS 140-2 Level 4 and PCI HSM security requirements for hardware security modules.

For more information about cryptographic features and functions, see 3.6, “Cryptographic features” on page 49.

1.2.5 Clustering connectivity

A parallel sysplex is an IBM Z clustering technology that is used to make applications that are running on logical and physical servers highly reliable and always available. The servers in a parallel sysplex are interconnected by way of coupling links.

Coupling connectivity for parallel sysplex on z14 ZR1 use Coupling Express Long Reach (CE LR) feature and the Integrated Coupling Adapter Short Reach (ICA SR) in the CPC drawer. The ICA SR supports distances up to 150 m (492.12 feet). The CE LR supports longer unrepeated distances of up to 10 km (6.21 miles) between systems. As with previous Z platforms, the z14 ZR1 does not support InfiniBand coupling links.

For information about coupling and clustering features, see 3.7, “Coupling and clustering” on page 52.

1.2.6 Special-purpose features and functions

When it comes to Z development, IBM takes a *total systems* view. The Z stack is built around digital services, agile application development, connectivity, and systems management. This configuration creates an integrated, diverse platform with specialized hardware and dedicated computing capabilities.

The z14 ZR1 delivers a range of features and functions that allow processor units to concentrate on computational tasks, while distinct, specialized features take care of the rest. The following special-purpose features and functions are offered with the z14 ZR1:

- Data compression:
  - The Compression Coprocessor (CMPSC) is a high-performance coprocessor that uses compression algorithms (such as new Huffman encoding) to help reduce disk space and memory usage.

  - The IBM zEnterprise® Data Compression (zEDC) Express feature delivers an integrated solution to help reduce CPU consumption, optimize performance of compression-related tasks, and enable more efficient use of storage resources.

- Secure Services Container

  This special-purpose firmware partition is isolated from production and enables the secure deployment of software appliances.

- GDPS Virtual Appliance

  The GDPS Virtual Appliance is a continuous availability and disaster recovery solution for enterprises that run only Linux on IBM Z on their Z platform. It can help improve availability and your recovery time objective (RTO) in planned and unplanned outage situations.

  The GDPS Virtual Appliance is installed in its own LPAR by way of the Hardware Management Console.

- Dynamic Partition Manager (DPM)

  DPM is a guided management interface that is used to define the Z hardware and virtual infrastructure, including integrated dynamic I/O management that runs Linux on IBM Z and the KVM hypervisor[7] environments.

- Guarded Storage Facility (GSF)

  Also known as pause-less garbage collection, GSF is available in the z14 ZR1 to enable enterprise scale Java applications to run with fewer and shorter pauses for garbage collection on larger heaps.

- Instruction Execution Protection Facility (IEPF)

  Instruction Execution Protection is a new hardware function that was introduced with z14 ZR1. It enables software, such as IBM Language Environment®, to mark certain memory regions (for example, a heap or stack) as non-executable to improve the security of programs that are running on Z against stack-overflow or similar attacks.

- Simultaneous multithreading (SMT)

  By using SMT, you can process up to two simultaneous threads in a single core to optimize throughput. An operating system with SMT support can be configured to dispatch work to a thread on a zIIP[8] or an IFL[9]. SAP engines also support SMT.

- Single-instruction, multiple-data (SIMD)

  SIMD is set of instructions that allows optimization of code to complex mathematical models and business analytics vector processing. The set of SIMD instructions are a type of data-parallel computing and vector processing that can decrease the amount of code and accelerate mathematical computations with integer, string, character, and floating point data types.

- Coupling and parallel sysplex

  Data sharing and serialization functions are offloaded to the coupling facility by way of a special coupling link network, which provides the infrastructure to run a single production workload that accesses a common set of data across many z/OS system images.

For more information about z14 ZR1 features, see Chapter 3, “Supported features and functions” on page 31, and Chapter 4, “Strengths of the z14 ZR1” on page 57.

[7] For more information, see 5.2.6, “KVM hypervisor” on page 92.

[8] IBM z Integrated Information Processor (zIIP) is used under z/OS for designated workloads, which include IBM Java virtual machine (JVM) and various XML System Services.

[9] An Integrated Facility for Linux (IFL) is exclusively used with Linux on IBM Z and for running the z/VM or KVM hypervisor in support of Linux.

1.2.7 Capacity on demand and performance

The z14 ZR1 enables just-in-time deployment of processor resources. The capacity on demand (CoD) function allows users to dynamically change available system capacity. This function helps companies respond to new business requirements with flexibility and precise granularity.

Also contributing to the extra capacity on the z14 ZR1 are numerous improvements in processor chip design, including new instructions, multithreading, and redesigned and larger caches.

The largest IBM z14 ZR1 configuration is expected to provide up to 13% (6-way versus 6-way) more capacity for z/OS and 60% (30-way to 20-way) for Linux than the largest z13s configuration. It also is expected to provide up to 62% (6-way versus 6-way) more capacity for z/OS and 240% (30-way to 13-way) for Linux than the largest zBC12 configuration.

For more information, see 4.3.2, “Capacity on demand” on page 71, and 4.3.3, “z14 ZR1 performance” on page 73.

1.2.8 Reliability, availability, and serviceability

The z14 ZR1 offers the same high quality of service and reliability, availability, and serviceability (RAS) that is traditional in Z platforms. The RAS strategy uses a building-block approach that is designed to meet stringent client requirements for achieving continuous, reliable operation. The RAS features the following 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 product releases and investing in new RAS functionality to eliminate or minimize all sources of outages.

For more information about RAS, see 4.4, “Reliability, availability, and serviceability” on page 75.

1.3 Software support

The z14 ZR1 supports a wide range of IBM and independent software vendor (ISV) software solutions. This range includes traditional batch and online transaction processing (OLTP) environments, such as IBM Customer Information Control System (IBM CICS®), IBM Information Management System (IBM IMS™), and IBM DB2. It also includes the following web services (in addition to other web services that are not listed):

- Java platform
- Linux and open standards applications
- WebSphere
- IBM MobileFirst™ Platform Foundation for mobile application development
- IBM z/OS Connect Enterprise Edition

The following operating systems are supported on the z14 ZR1:

- z/OS Version 2 Release 3
- z/OS Version 2 Release 2 with PTFs (exploitation)
- z/OS Version 2 Release 1 with PTFs (exploitation)
- z/OS Version 1 Release 13 with PTFs (limited exploitation, requires extended support)
- z/VM Version 7 Release 1
- z/VM Version 6 Release 4 with PTFs (compatibility and exploitation support)
- z/VSE Version 6 Release 2 with PTFs
- z/VSE Version 6 Release 1 with PTFs
- z/VSE Version 5 Release 2 with PTFs
- z/TPF Version 1 Release 1 (compatibility support)
- Linux on IBM Z:
  - SUSE Linux Enterprise Server, SLES 12 SP2 with service, and SLES 11 SP4 with service
  - Red Hat Enterprise Linux; RHEL 7.3 with service and RHEL 6.9 with service
  - Canonical Ubuntu 16.04 LTS with service
- KVM hypervisor, which is offered by some Linux distributions. See your distribution's documentation to ensure that KVM on Z is supported.

Note: KVM for IBM z Systems was withdrawn from marketing and is no longer serviced.

Detailed service levels are identified during toleration tests. For more information about recommended distribution levels, see the Linux on IBM Z page of the IBM Z website.

For more information about the z14 ZR1 software support, see Chapter 5, "Operating system support" on page 85.

1.3.1 IBM compilers

Compilers are built with specific knowledge of the system architecture, which is used during code generation. Therefore, the use of the latest compilers is essential to extract the maximum benefit of a platform's capabilities. IBM compilers use the latest architecture enhancements and new instruction sets to deliver more value.

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 makes it possible to capitalize on IT investments, while smoothly incorporating new, web-based applications into the infrastructure.

z/OS, XL C/C++, and XL C/C++ for Linux on IBM Z 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 into executable code that fully uses the Z architecture. This transformation is possible thanks to hardware-tailored optimizations, built-in functions, performance-tuned libraries, and language constructs that simplify system programming and boost application runtime performance.

Compilers, such as COBOL, PL/I, and z/OS v2.3 XL C/C++, are inherently optimized on the IBM z14 ZR1 because they use floating point registers rather than memory or fast mathematical computations. The use of compilers that use hardware enhancements is key to improving application performance, reducing CPU usage, and lowering operating costs.

The IBM z14 ZR1 also offers release-to-release improvements for Java. Combined with cryptography acceleration, the z14 ZR1 can deliver improvements in throughput per core. The new z14 ZR1 pause-less garbage collection capability also provides improvements in throughput per core. Because of shorter, more consistent Java response times, the z14 ZR1 achieves both improvements without significantly affecting overall throughput. For Java applications that must remain highly responsive, enabling the new garbage collection mode is a good option.

2

IBM z14 ZR1 hardware overview

This chapter expands on the description of the key hardware elements of the z14 ZR1 that was presented in Chapter 1, “The foundation of a trusted digital transformation” on page 1. It includes the following topics:

2.1 Models and upgrade paths

The z14 Model ZR1 features an assigned machine type (MT) of 3907, which uniquely identifies the central processor complex (CPC). All z14 ZR1 use processor units single chip modules (PU SCMs, up to four per system, feature driven) with five, six, seven, eight, or nine active processor unit cores. Spare processor units, system assist processors (SAPs), and one integrated firmware processor (IFP) are integral to the z14 ZR1 system.

The z14 ZR1 is based on four new feature codes: Max4, Max12, Max 24, and Max30. The feature codes determine the maximum number of characterizable processor units available in the CPC drawer.

The number of characterizable processor units, SAPs, and spare processor units for the various models is listed in Table 2-1. For more information about processor unit characterization, see “PU characterization” on page 23.

Table 2-1 z14 ZR1 summary (machine type 3907)

Feature name	Feature code	Characterizable processor units	Standard SAPs	Spares	Integrated firmware processor
Max4	0636	1 - 4	2	1	1
Max12	0637	1 - 12	2	1	1
Max24	0638	1 - 24	2	1	1
Max30	0639	1 - 30	2	1	1

The upgrade paths for the z14 ZR1 are shown in Figure 2-1. Because of the new design of the z14 ZR1, no supported migration path to any of the z14 models is available.

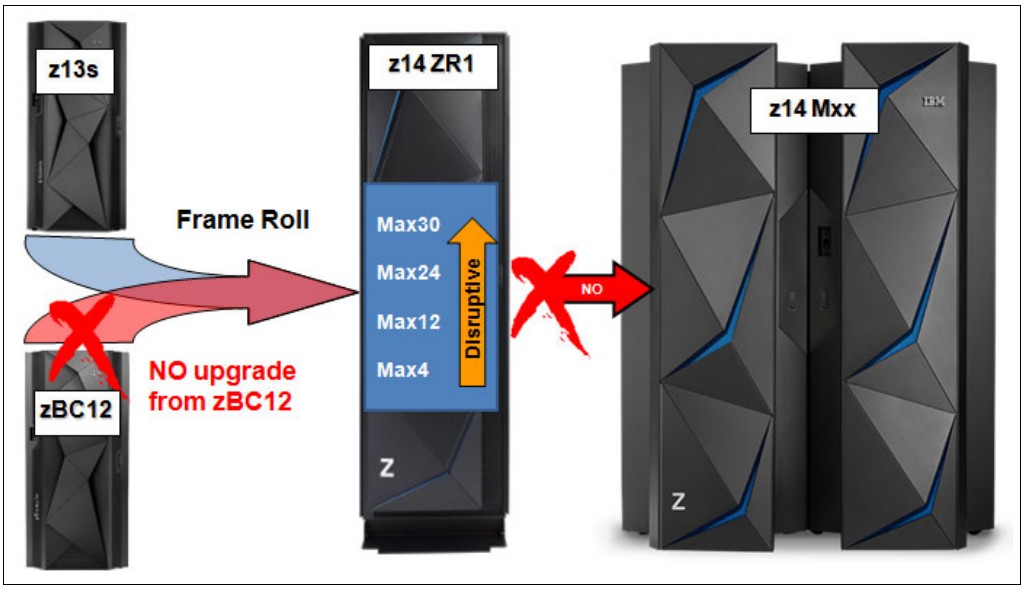

Figure 2-1 z14 ZR1 upgrade paths

On the z14 ZR1, concurrent processor upgrades (for more information, see in Chapter 4, “Strengths of the z14 ZR1” on page 57) are available for CPs, IFLs, ICFs, Z-Integrated Information Processors (zIIPs), and SAPs. However, concurrent processor unit upgrades require that more processor units are physically installed, but not activated at a previous time.

If an upgrade request cannot be accomplished in the configuration, a hardware upgrade is required in which one or more PU SCMs, memory, and CPC drawer I/O features are added to accommodate the wanted capacity. On the z14 ZR1, more PU SCMs and memory DIMMs cannot be installed concurrently because z14 ZR1 includes a single CPC drawer.

Spare processor units are used to replace defective processor units and one spare processor unit (core) always is available on a z14 ZR1. In the rare event of a processor unit failure, the spare processor unit is immediately and transparently activated and assigned the characteristics of the failing processor unit.

2.2 Rack and cabling

The z14 ZR1 is the first IBM Z that is designed to fit into an IBM supplied industry standard 19-inch rack. The z14 ZR1 is a single-rack system that is delivered as an air-cooled system.

The rack forms the z14 ZR1 CPC and contains one CPC drawer. The number of PCIe+ I/O drawers can vary based on the number of I/O features. Up to four PCIe+ I/O drawers can be installed. PCIe I/O drawers can be added concurrently[1].

In addition, the z14 ZR1 (new builds and MES orders) offers top-exit options for the fiber optic and copper cables that are used for I/O and power. These options (*Top Exit Power* and *Top Exit I/O Cabling*) provide more flexibility in planning where the system is installed. This flexibility can free you from running cables under a raised floor, which increases air flow over the system. Top and bottom cabling are supported for power and I/O.

The z14 ZR1 supports an installation on raised floor and non-raised floor environments.

Internal, front, and rear views of the z14 ZR1 system with the maximum four PCIe+ I/O drawers are shown in Figure 2-2 on page 20.

[1] The number of available PCIe fanout slots depends on the CPC drawer feature (Max4, Max12, Max24/Max30),

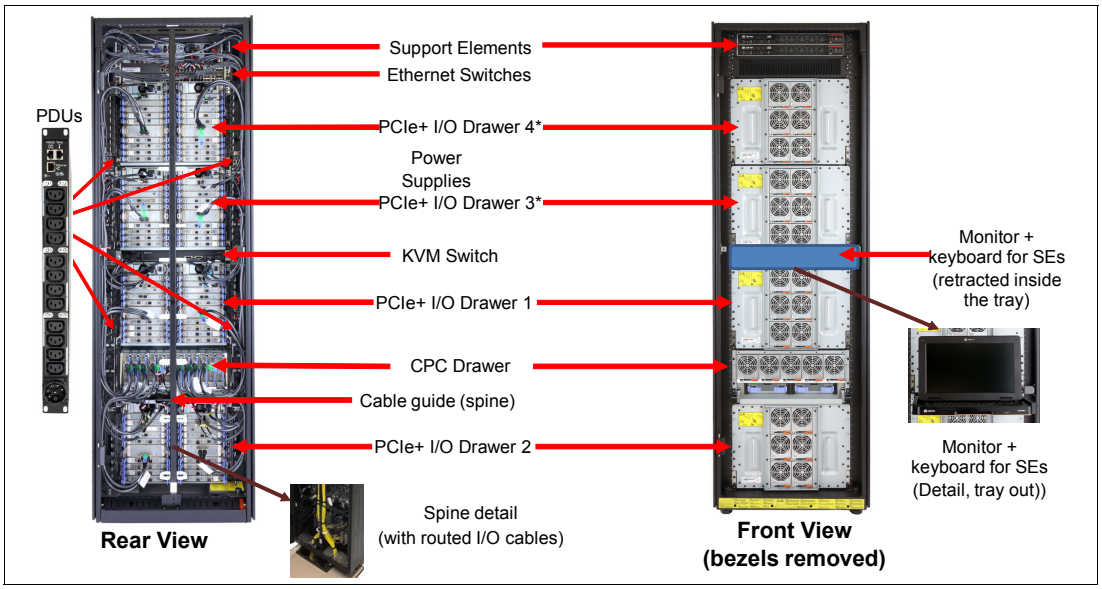

Figure 2-2 z14 ZR1 front and rear views of a configuration with four PCIe+ I/O drawers

2.3 CPC drawer

A fully populated CPC drawer (for example, a Max24 or Max30) is shown in Figure 2-3. The z14 ZR1 features 672 MB of L4 cache. The Storage Controller (SC) chip provides X-Bus connectivity for PUS SCMs in a different logical cluster.

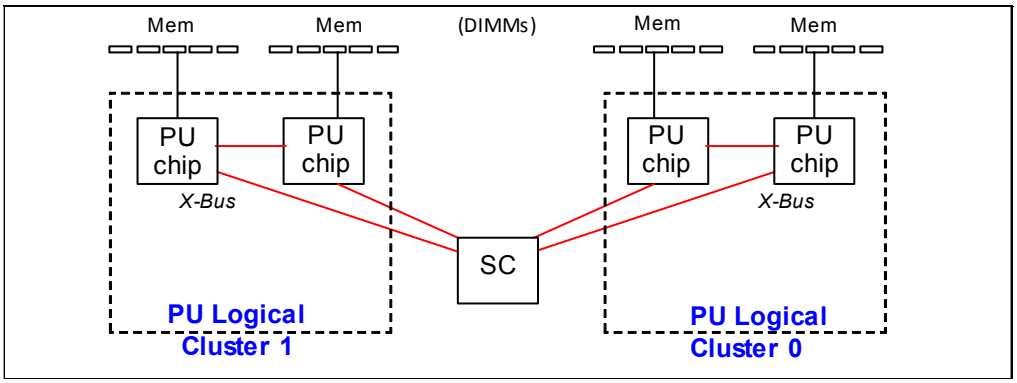

Figure 2-3 z14 ZR1 CPC drawer communication topology

The design that is used to connect the processor unit and SC allows the system to be operated and controlled by the IBM Processor Resource/Systems Manager™ (PR/SM™) facility as a memory-coherent symmetrical multiprocessor (SMP) system.

The z14 ZR1 features one CPC drawer that contains the following elements:

- Single chip modules:
  - One, two, or four PU single chip modules, each containing five, six, seven, eight, or nine processor unit cores (air-cooled). For more information, see Table 2-1 on page 18.
  - One System Controller single chip module, with a total of 672 MB L4 cache.

For more information about single chip modules, see 2.3.1, “Single chip modules” on page 21. For more information about available models and the relationship between the number of PU SCMs and number of available processor units, see Table 2-1 on page 18.

- Memory:
  - A minimum of 64 GB and a maximum of 8 TB of memory (excluding 64 GB HSA) is available for client use. For more information, see Table 2-2 on page 24.
  - A total of 5, 10, 15, or 20 memory DIMMs are plugged in a CPC drawer.
  - The number of memory DIMMs that can be plugged into the CPC drawer depends on the CPC drawer feature.
- Fanouts

  The CPC drawer provides up to eight PCIe Gen3 fanout adapters to connect to the PCIe+ I/O drawers and Integrated Coupling Adapter Short Reach (ICA SR) coupling links. The number of fanouts that can be installed depends on the CPC drawer feature (see Table 2-2 on page 24).

  Each fanout can be one of the following configurations:
  - A single port PCIe 16 GBps I/O fanout, each supporting one domain in a 16-slot PCIe+ I/O drawer.
  - A dual port ICA SR PCIe fanout for coupling links (two links, 8 GBps each).
- Two or four Power Supply Units (PSUs) provide power to the CPC drawer and are accessible from the rear (depending on the number of PU SCMs that is installed).

  Loss of one PSU leaves enough power to satisfy the power requirements of the entire drawer. The PSUs can be concurrently maintained.
- Two Flexible Support Processors (FSPs) provide redundant interfaces to the internal management network.
- Two Oscillator Cards (OSCs) provide clock synchronization to the CPC.

2.3.1 Single chip modules

The CPC drawer has up to four PU single chip modules (SCMs) and one SC SCM. Each PU SCM has five, six, seven, eight, or nine active PU cores, and L1, L2, and L3 caches. The SC SCM holds the L4 cache (672 MB).

For z14 ZR1, four CPC drawer configurations are offered with 8, 16, 28, and 34 active processor units.

The Storage Controller SCM includes 672 MB shared eDRAM cache, interface logic the four PU SCMs, and SMP fabric logic. The SC SMC is configured to provide a single 672 MB L4 cache that is shared by all PU cores in the CPC drawer.

2.3.2 Processor unit

Processor unit (PU) is the generic term for an IBM z/Architecture processor. Each PU is a superscalar processor with the following attributes:

- Up to six instructions can be decoded per clock cycle.
- Up to 10i instructions can be in execution per clock cycle.
- Instructions can be issued out of order. The PU uses a high-frequency, low-latency pipeline that provides robust performance across a wide range of workloads.

- 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 different types of instructions. Several instructions can be in execution at any moment, which are subject to the maximum number of decodes and completions per cycle.

PU cache

The on-chip cache for the PU (core) features the following design:

- Each PU core has an L1 cache (private) that is divided into a 128 KB cache for instructions and a 128 KB cache for data.
- Each PU core has a private L2 cache, with 4 MB D-cache (D for data) and 2 MB I-cache (I for instruction).
- Each PU SCM contains a 128 MB L3 cache that is shared by all PU cores in the SCM. The shared L3 cache uses eDRAM.

This on-chip cache implementation optimizes system performance for high-frequency processors, with cache improvements, new Translation/TLB2 design, pipeline optimizations, and better branch prediction.

The CPC drawer cache structure is shown in Figure 2-4.

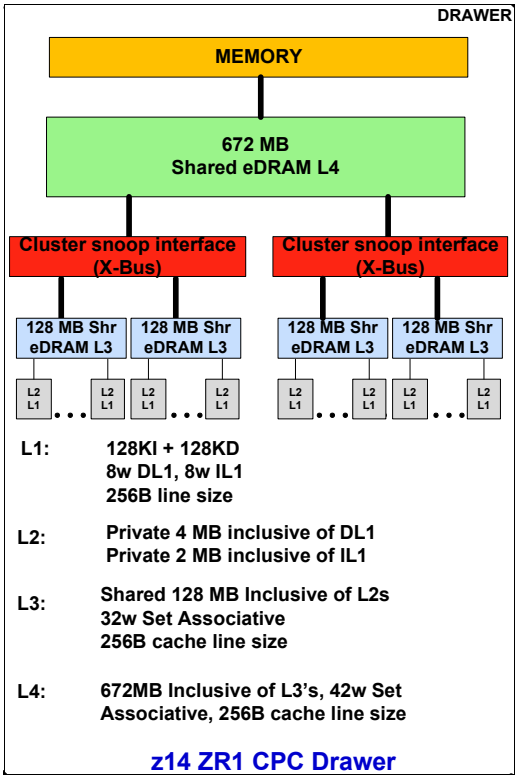

Figure 2-4 z14 ZR1 cache structure

PU sparing

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

On-chip cryptographic hardware

Dedicated on-chip cryptographic hardware for each PU core includes extended key and hash sizes for the Advanced Encryption Standard (AES) and Secure Hash Algorithm (SHA), and support for UTF8 to UTF16 conversion. This cryptographic hardware is available with any processor type; for example, CP, IBM zIIP, and IFL.

Software support

The z14 ZR1 PUs provide full compatibility with software for z/Architecture, and extend the Instruction Set Architecture (ISA) to enable enhanced functionality and performance. The following hardware instructions support more efficient code generation and execution are introduced in the z14 ZR1:

- CP Assist for Cryptographic Functions (CPACF)
- Compression call (CMPSC)
- Hardware decimal floating point (HDFP)
- Transactional Execution Facility
- Runtime Instrumentation Facility
- Single-instruction, multiple-data (SIMD)

For more information about these capabilities, see Chapter 4, “Strengths of the z14 ZR1” on page 57.

PU characterization

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

- CP
- IFL processor
- zIIP
- ICF
- SAP
- IFP

At least one CP must be purchased with a zIIP or before a zIIP can be purchased. You can purchase up to two zIIPs for each purchased CP (assigned or unassigned) on the system.

However, an LPAR definition can go beyond the 1:2 ratio. For example, on a system with two physical CPs, a maximum of four physical zIIPs can be installed. An LPAR definition for that system can contain up to two logical CPs and four logical zIIPs. Another possible configuration is one logical CP and three logical zIIPs.

Converting a PU from one type to any other type is possible by using the Dynamic Processor Unit Reassignment process. These conversions occur concurrently with the system operation.

Note: The addition of ICFs, IFLs, zIIPs, and SAP to the z14 ZR1 does not change the system capacity setting or its million service units (MSU) rating.

2.3.3 Memory

The maximum physical memory size is directly related to the number of PU SCMs that is installed in the system. With the z14 ZR1, up to 8 TB of memory can be ordered.

Important: z/OS V2R3 requires a minimum of 8 GB of memory (2 GB of memory when running under z/VM). z/OS can support up to 4 TB of memory in an LPAR.

The minimum and maximum memory sizes for each z14 ZR1 feature are listed in Table 2-2.

Table 2-2 z14 ZR1 memory per feature

Feature name	PU SCMs	Memory
Max4 (FC 0636)	1	64 - 1984 GB
Max12 (FC 0637)	2	64 - 4032 GB
Max24 (FC 0638)	4	64 - 8128 GB
Max30 (FC 0639)	4	64 - 8128 GB

The hardware system area (HSA) on the z14 ZR1 features a fixed amount of memory (64 GB) that is managed separately from orderable memory. However, the maximum amount of orderable memory can vary from the theoretical number because of dependencies on the memory granularity. On z14 ZR1 platforms, the granularity for memory are in 8 GB, 32 GB, 64 GB, 128 GB, 256 GB, and 512 GB increments.

Physically, memory is organized in the following ways:

- The CPC drawer always contains a minimum of 128 GB of installed memory, of which 64 GB is usable by the operating system.
- The CPC drawer can have more installed memory than is enabled. The excess memory can be enabled by a Licensed Internal Code load.
- Memory upgrades are first satisfied by using installed but unused memory capacity until it is exhausted. When no more unused memory is available from the installed cards, cards must be added or the cards must be upgraded to a higher capacity.

A memory upgrade is considered to be concurrent when it requires no change of the physical memory cards. A memory card change is disruptive.

Concurrent memory upgrade

If physical memory is available, memory can be upgraded concurrently by using Licensed Internal Code Configuration Control (LICCC), as wanted. The plan ahead memory function that is available with the z14 ZR1 enables nondisruptive memory upgrades by having pre-plugged memory (based on a target configuration) in the system. Pre-plugged memory is enabled through an LICCC order that is placed by the client.

Redundant array of independent memory

RAIM technology makes the memory subsystem a fully fault-tolerant N+1 design in essence. The RAIM design automatically detects and recovers from failures of dynamic random access memory (DRAM), sockets, memory channels, or dual inline memory modules (DIMMs).

The RAIM design is fully integrated in the z14 ZR1 and was 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 RAIM.

This technology has significant reliability, availability, and serviceability (RAS) capabilities in the area of error correction. Bit, lane, DRAM, DIMM, socket, and complete memory channel failures (including many types of multiple failures) can be detected and corrected.

For more information about memory design and configuration options, see *IBM z14 Model ZR1 Technical Guide*, SG24-8651.

2.3.4 Hardware system area

The HSA is a fixed-size, reserved area of memory that is separate from the purchased memory. Although the HSA is used for several internal functions, the bulk of it is used by channel subsystem functions.

The fixed-size 64 GB HSA for z14 ZR1 is large enough to accommodate any LPAR definitions or changes, which eliminates most outage situations and the need for extensive planning.

A fixed, large HSA allows the dynamic I/O capability of the z14 ZR1 to be enabled by default. It also enables the dynamic addition and removal of the following features:

- LPAR to new or existing channel subsystem (CSS)
- CSS (up to three can be defined in z14 ZR1)
- Subchannel set (up to three can be defined in z14 ZR1)
- Devices, up to the maximum number permitted, in each subchannel set
- Logical processors by type
- Cryptographic adapters

2.4 I/O system structure

The z14 ZR1 supports Generation 3 PCIe-based infrastructure for PCIe+ I/O drawers (PCIe Gen3). The number of supported PCIe+ I/O drawers is listed in Table 2-3.

Table 2-3 z14 ZR1 CPC drawer fanouts per feature

Feature name	PU SCMs	Max. PCIe fanouts	Max. PCIe+ I/O drawers[a]
Max4 (FC 0636)	1	2	1
Max12 (FC 0637)	2	4	2
Max24 (FC 0638)	4	8	4
Max30 (FC 0639)	4	8	4

a. If the 16U Reserved feature (FC 0617) is ordered, the maximum number of PCIe+ I/O drawers is two. For more information, see 2.4.1, “16U Reserved feature” on page 28.

The PCIe I/O infrastructure consists of PCIe Gen3 fanouts in the CPC drawer that support 16 GBps connectivity to the PCIe+ I/O drawer.

Note: Ordering I/O feature types and quantities determines the appropriate number of PCIe+ I/O drawers.

A high-level view of the connectivity between the CPC drawer and PCIe+ I/O drawer of the z14 ZR1 is shown in Figure 2-5.

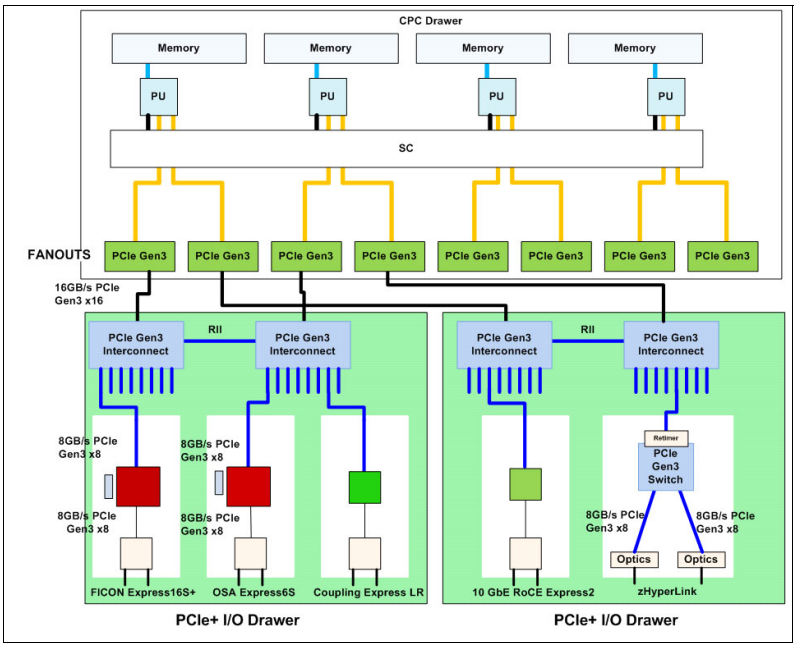

Figure 2-5 z14 ZR1 connectivity of the I/O subsystem

The z14 ZR1 supports the following fanout types (for more information about fanout location, see Figure 2-6), which are at the front of the CPC drawer:

- ICA SR
- PCIe Gen3 (for PCIe+ I/O drawer)

The PCIe Gen3 fanout has one port; ICA SR has two ports.

For coupling link connectivity (parallel sysplex and STP configuration), the z14 ZR1 supports the following link types:

- ICA SR
- Coupling Express LR

The z14 ZR1 CPC drawer (see Figure 2-6 on page 27) can have up to eight 1-port PCIe Gen3 or 2-port ICA SR PCIe coupling fanouts (numbered LG01 - LG04 and LG07 - LG10). All coupling fanouts support parallel sysplex connectivity.

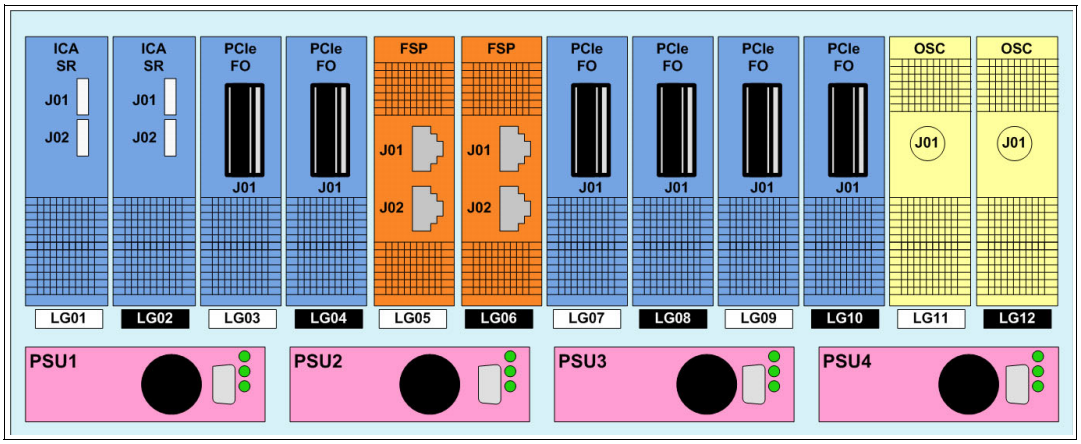

Figure 2-6 Front view of z14 ZR1 CPC drawer

The PCIe+ I/O drawer (see Figure 2-7) is a 19-inch single side drawer that is 8U high. I/O features are installed horizontally, with cooling air flow from front to rear. The drawer contains 16 slots and two switch cards. These features support two I/O domains that each contain eight features of any of the following types:

- FICON Express16S+, FICON Express16S, or FICON Express8S
- OSA-Express7S, OSA-Express6S, OSA-Express5S, or OSA-Express4S
- Crypto Express6S or Crypto Express5S
- zEDC Express
- 25GbE RoCE Express2, 10GbE RoCE Express2, or 10GbE RoCE Express
- zHyperLink Express
- Coupling Express LR

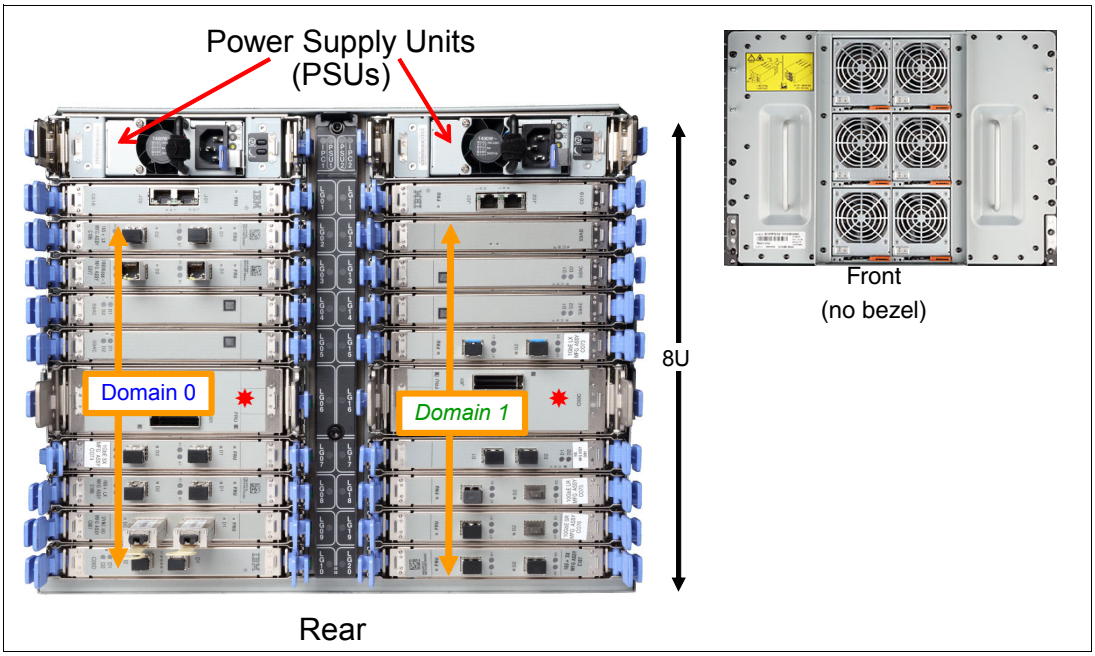

Figure 2-7 PCIe I/O drawer - front and rear views

Two PSUs provide redundant power, and six front-side fans provide redundant cooling to the PCIe+ I/O Drawer.

More information about the z14 ZR1 supported I/O features, see Chapter 3, “Supported features and functions” on page 31.

2.4.1 16U Reserved feature

IBM z14 ZR1 is a single 19-inch rack industry-standard footprint. In some cases, your installation might not require full four PCIe+ I/O drawer configuration.

If no more than two PCIe+ I/O drawers (eight EIA units each) are needed, the 19-inch rack has 16 EIA units contiguous rack space that is not populated with I/O. In addition, because no I/O is installed in those rack locations, no other power is available to the rack.

The 16U Reserved feature (FC 0617) uses the empty locations in the z14 ZR1 rack, which allows for maximizing data center space efficiency. The 16U Reserved feature tags 16U of contiguous rack space in the z14 ZR1 rack as “available” (in lieu of a third and fourth PCIe+ I/O drawer) for you to populate with your choice of equipment if it meets the documented specifications.

To stay within warranty boundaries of the z14 ZR1, the non-Z hardware content that is considered for installation must follow the following requirements (for more information, see the feature documentation):

- Fits within 19-inch rail-to-rail width, and 28 717-mm (1/4-inch) front-to-rear depth.
- Uses front to rear airflow.
- Features no more than 15.8 Kg (35 lbs.) weight per EIA location; for example, a 4U drawer can weigh no more than 65.5 Kg. (140 lbs.).
- Installed components show safety certification labels that are required for server components in the country where the system is used (such as UL or TUC).
- FC 0617 is ordered to reserve the space for planning and stay under warranty.

The maximum power that is reserved for this feature is 3400 W.

The DS8882F is a rack-mounted, all-flash storage system that can take advantage of the 16U space available in the z14 ZR1, which creates a single-footprint solution that works well for space-constrained environments.

The DS8882F offers the following features:

- Six-core processors with options of 64 GB, 128 GB, and 256 GB memory combinations
- Support for up to a maximum of four 4-port host adapters
- A variety of all-flash, high-performance drive sets, which include 400 GB, 800 GB, 1600 GB, and 3200 GB
- Two all-flash, high-capacity drive set options: 3.84 TB and 7.68 TB
- Support for one high-performance flash enclosure (HPFE) Gen 2 pair that can be populated with high-performance or high-capacity drive sets, and a maximum raw capacity of 368.64 TB
- Battery backup units that protect the write in-flight if a power loss occurs
- Same advanced functions as the rest of the DS8880F family members, except zHyperLink support

For more information about the 16U Reserved rack space, see *IBM 3907 Installation Manual for Physical Planning*, CG28-6974.

The location of the 16U Reserved feature is shown in Figure 2-8.

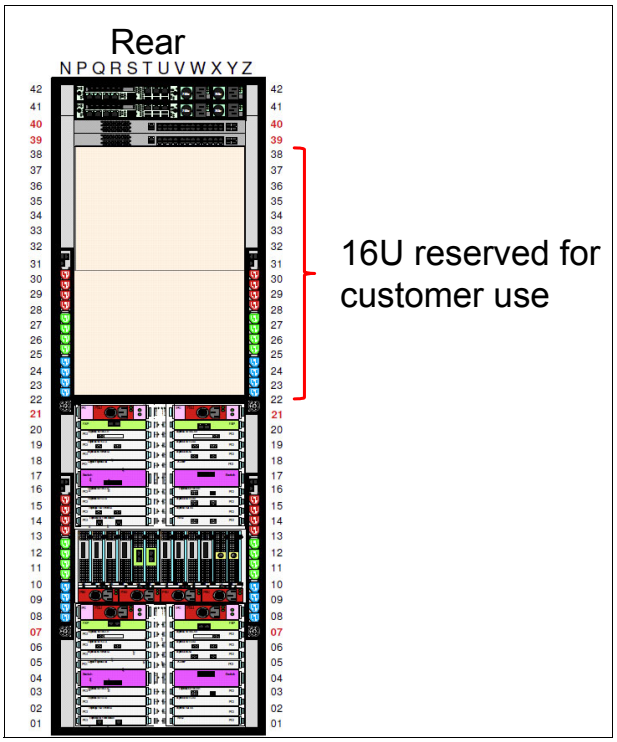

Figure 2-8 16U Reserved feature location

2.5 Power and cooling

The z14 ZR1 meets the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) Class A3 specifications. (ASHRAE is an organization that is devoted to the advancement of indoor environment control technology in the heating, ventilation, and air conditioning industry.)

The air-cooling system in the z14 ZR1 is designed with N+1 blowers/fans, controls, and sensors for the CPC and PCIe+ I/O drawers.

The power and cooling system was redesigned for the IBM z14 ZR1 and includes the following features:

- To fit in a single 42U 19-inch IBM rack (replacing the 24-inch frame).
- Intelligent Power Distribution Units (PDUs) replace the Bulk Power Assemblies (BPAs):
  - Two or four intelligent PDUs, depending on the machine configuration
  - Single phase 200 - 240 VAC (two or four power cords)
- Air cooled (CPC and I/O): Front to rear air flow (all cabling in the rear, no cabling in front)
- No High-Voltage DC (HVDC) option or Internal Battery Feature (IBF)
- No Emergency Power Off switch (EPO)
- Uses single phase power as opposed to three-phase power

2.5.1 Power considerations

The z14 ZR1 has redundant PDUs, redundant PSUs for the CPC drawer, PCIe+ I/O drawers and Support Elements. The Ethernet switches are redundant and can be replaced with no effect on system operation. The loss of one PDU or one PSU also has no effect on system operation.

If the 16U Reserved feature is ordered, the power consumption and other considerations and restrictions for the non-Z equipment that is installed in the z14 ZR1 rack that might apply are described in *IBM 3907 Installation Manual for Physical Planning*, GC28-6974.

Specific power requirements depend on the CPC drawer features and the number and type of I/O units that are installed. For more information about the maximum power consumption for the various configurations and environments, see *IBM 3907 Installation Manual for Physical Planning*, GC28-6974.

For more information about the available power and weight estimation tool, see the IBM Resource Link® website.

3

Supported features and functions

This chapter describes the I/O and other miscellaneous features and functions of the z14 ZR1. The information in this chapter expands upon the overview of the key hardware elements that was provided in Chapter 2, “IBM z14 ZR1 hardware overview” on page 17.

Only the enhanced features and functions that were introduced with the z14 ZR1 are described more in detail in this chapter. The remaining supported features from earlier generations of Z platforms are listed for convenience.

Note: Throughout the chapter, reference is made to the *IBM z14 ZR1 Technical Guide,* SG24-8651.

This chapter includes the following topics:

3.1 I/O features at a glance

The z14 ZR1 supports a PCIe-based infrastructure for PCIe+ I/O drawers to support the following I/O features:

- zHyperLink Express
- FICON Express16S+
- FICON Express16S (carry forward only)
- FICON Express8S (carry forward only)
- OSA-Express7S
- OSA-Express6S
- OSA-Express5S (carry forward only)
- OSA-Express4S (carry forward only, select features)
- 25GbE RoCE Express2
- 10GbE RoCE Express2
- 10GbE RoCE Express (carry forward only)
- Crypto Express6S
- Crypto Express5S (carry forward only)
- zEnterprise Data Compression

The following clustering and coupling links are support on the z14 ZR1:

- Integrated Coupling Adapter Short Reach (ICA SR)
- Coupling Express Long Reach (CE LR)

The following features that were part of earlier Z platforms are *not orderable* and cannot be carried forward to the z14 ZR1:

- ESCON
- FICON Express8 and older
- OSA-Express3 and older
- ISC-3
- Crypto Express4S and older
- Flash Express
- Host Channel Adapter3 - Optical Long Reach (HCA3-O LR)
- Host Channel Adapter3 - Optical (HCA3-O)

Connector type LC Duplex is used for all fiber optic cables, except those cables that are used for the zHyperLink Express, and ICA SR connections, which feature multi-fiber push-on (MPO) connectors. The MPO connector of the zHyperLink Express and the ICA SR connection includes two rows of 12 fibers and are interchangeable.

Storage connectivity options are listed in Table 3-1. For more information about zHyperLink, FICON, and FCP connectivity in relation to the z14 ZR1, see 3.3, “Storage connectivity” on page 36.

Table 3-1 Storage connectivity features

Feature	Feature codes	Bit rate in Gbps (or stated)	Cable type	Maximum unrepeated distance	Ordering information
zHyperLink Express	0431	8 GBps	OM4, OM5	150 m	New build
			OM3	100 m	
FICON Express16S+ 10KM LX	0427	4, 8, or 16	SM 9 µm	10 km (6.2 miles)	New build
FICON Express16S+ SX	0428	4, 8, or 16	OM2, OM3, and OM4	See Table 3-2	New build
FICON Express16S 10KM LX	0418	4, 8, or 16	SM 9 µm	10 km (6.2 miles)	Carry forward
FICON Express16S SX	0419	4, 8, or 16	OM2, OM3, and OM4	See Table 3-2	Carry forward
FICON Express8S 10KM LX	0409	2, 4, or 8	SM 9 µm	10 km (6.2 miles)	Carry forward
FICON Express8S SX	0410	2, 4, or 8	OM2, OM3, and OM4	See Table 3-2	Carry forward

The maximum unrepeated distances for different multimode fiber optic cable types when used with FICON SX (shortwave) features running at different bit rates are listed in Table 3-2.

Table 3-2 Unrepeated distances for different multimode fiber optic cable types

Cable type (Modal bandwidth)	2 Gbps	4 Gbps	8 Gbps	16 Gbps
OM1 (62.5 µm at 200 MHz·km)	150 meters	70 meters	21 meters	N/A
	492 feet	230 feet	69 feet	N/A
OM2 (50 µm at 500 MHz·km)	300 meters	150 meters	50 meters	35 meters
	984 feet	492 feet	164 feet	115 feet
OM3 (50 µm at 2000 MHz·km)	500 meters	380 meters	150 meters	100 meters
	1640 feet	1247 feet	492 feet	328 feet
OM4 (50 µm at 4700 MHz·km)	N/A	400 meters	190 meters	125 meters
	N/A	1312 feet	623 feet	410 feet

The network connectivity options are listed in Table 3-3. For more information about OSA-Express and RoCE Express connectivity in relation to the z14 ZR1, see 3.4, “Network connectivity” on page 40.

Table 3-3 Network connectivity features

Feature	Feature codes	Bit rate in Gbps (or stated)	Cable type	Maximum unrepeated distance[a]	Ordering information
OSA-Express7S 25GbE SR	0429	25	MM 50 µm	70 m (2000) 100 m (4700)	New build
OSA-Express6S 10GbE LR	0424	10	SM 9 µm	10 km (6.2 miles)	New build
OSA-Express5S 10GbE LR	0415				Carry forward
OSA-Express4S 10GbE LR	0406				
OSA-Expess6S 10GbE SR	0425	10	MM 62.5 µm MM 50 µm	33 m (200) 82 m (500) 300 m (2000)	New build
OSA-Expess5S 10GbE SR	0416				Carry forward
OSA-Express4S 10GbE SR	0407				
OSA-Express6S GbE LX	0422	1.25	SM 9 µm	5 km (3.1 miles)	New build
OSA-Express5S GbE LX	0413				Carry forward
OSA-Express4S GbE LX	0404				
OSA-Express6S GbE SX	0423	1.25	MM 62.5 µm MM 50 µm	275 m (200) 550 m (500)	New build
OSA-Express5S GbE SX	0414				Carry forward
OSA-Express4S GbE SX	0405				
OSA-Express6S 1000BASE-T	0426	100 or 1000 Mbps	Cat 5, Cat 6 unshielded twisted pair (UTP)	100 m	New build
OSA-Express5S 1000BASE-T	0417				Carry forward
25GbE RoCE Express2	0430	25	MM 50 µm	70 m (2000) 100 m (4700)	New build
10GbE RoCE Express2	0412	10	MM 62.5 µm MM 50 µm	33 m (200) 82 m (500) 300 m (2000)	New build
10GbE RoCE Express	0411				Carry forward

a. The minimum fiber bandwidth distance in MHz-km for multi-mode fiber optic links is included in parentheses, where applicable.

Coupling link options are listed in Table 3-4. For more information about the parallel sysplex or STP-only link connectivity in relation to the z14 ZR1, see 3.7, “Coupling and clustering” on page 52, and 3.9, “Server Time Protocol” on page 54.

Table 3-4 Coupling and clustering features[1]

Feature	Feature codes	Bit rate	Cable type	Maximum unrepeated distance	Ordering information
CE LR	0433	10 Gbps	SM 9 µm	10 km (6.2 miles)	New build or Carry forward
ICA SR	0172	8 GBps	OM4, OM5	150 m	New build or Carry forward
			OM3	100 m	
Internal Coupling (IC)	No coupling link feature or fiber optic cable is required.				

Special purpose features, such as cryptographic or compression features, and Virtual Flash Memory are listed in Table 3-5. For more information about cryptographic the feature, see 3.6, “Cryptographic features” on page 49.

Table 3-5 Special-purpose features

Feature	Feature codes	Bit rate in Gbps	Cable type	Maximum unrepeated distance	Ordering information
Crypto Express6S	0893	N/A	N/A	N/A	New build
Crypto Express5S	0890	N/A	N/A	N/A	Carry forward
zEDC Express	0420	N/A	N/A	N/A	New build or Carry forward
Virtual Flash Memory	0614	N/A	N/A	N/A	New build

3.2 Native PCIe features and integrated firmware processor

The following native PCIe features are available on the z14 ZR1:

- zHyperLink Express
- Coupling Express Long Reach (CE LR)
- 25GbE RoCE Express2
- 10GbE RoCE Express2
- 10GbE RoCE Express
- zEDC Express

PCIe+ I/O drawer for z14 ZR1

On the z14 ZR1, a PCIe+ I/O drawer (FC 4001) is used to host the PCIe features. The drawer fits in the 19-inch standard rack of the z14 ZR1. The PCIe+ I/O drawer hosts up to 16 PCIe features.

Note: PCIe I/O drawers from earlier Z platforms cannot be carried forward to z14 ZR1.

[1] The IBM z14 Model ZR1 does *not* support InfiniBand features

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. However, they are managed in a different way from the other I/O adapters and features. The native PCIe feature cards feature a PCHID that is assigned according to the physical location in the PCIe+ I/O drawer.

For the native PCIe features that are supported by z14 ZR1, drivers are 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 Group partitions that are running on the integrated firmware processor (IFP). The z14 ZR1 includes four Resource Groups.

The IFP is used to manage native PCIe adapters that are installed in a PCIe+ I/O drawer. The IFP is allocated from a pool of processor units that are available for the entire system. Because the IFP is exclusively used to manage native PCIe adapters, it is not taken from the pool of processor units 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 processor units, it is not visible. In error or failover scenarios, the IFP acts as any other processor unit (that is, sparing is started).

3.3 Storage connectivity

In the storage connectivity area, the main focus is on improving the latency for I/O transmission. With the introduction of zHyperLink Express, IBM ensures the optimization of the I/O infrastructure. The FICON Express 16S+ feature offers the same functions as its predecessor, with increased performance through new design enhancements.

For more information about FICON channel, see the Z I/O connectivity website. Technical papers about performance data are also available.

3.3.1 zHyperLink Express

IBM zHyperLink Express is a short distance Z I/O adapter with up to 5x lower latency than High-Performance FICON for read requests. This

feature is in the PCIe+ I/O drawer and is a two-port adapter that is used for short distance, direct connectivity between a z14 ZR1 and a DS8880. The zHyperLink Express is designed to support distances up to 150 meters at a link data rate of 8 GigaBytes per second (GBps).

A 24x MTP-MTP cable is required for each port of the zHyperLink Express feature. It is single 24-fiber cable with Multi-fiber Termination Push-on (MTP) connectors. Internally, the single cable houses 12 fibers for transmit and 12 fibers for receive.

Note: FICON connectivity to each storage system is still required. The FICON connection is used for zHyperLink initialization, I/O requests that are not eligible for zHyperLink communications, and as an alternative path if zHyperLink requests fail (for example, storage cache misses or busy storage device conditions).

3.3.2 FICON functions

FICON features continue to evolve and deliver improved throughput, reliability, availability, and serviceability (RAS). FICON features in the z14 ZR1 can provide connectivity to systems, Fibre Channel (FC) switches, and various devices in a SAN environment.

The FICON protocol is fully supported on the z14 ZR1. It is commonly used with IBM z/OS, IBM z/VM (and guest systems), Linux on Z, IBM z/VSE, and IBM z/TPF. The FICON enhancements are described next.

FICON multi-hops and cascaded switch support

The z14 ZR1 supports three hops (up to four FC switches) in a cascaded switch configuration. This support can help simplify the infrastructure with optimized RAS functionality. The support for a FICON multi-hop environment must also be provided by the FC switch vendor.

High-Performance FICON for z Systems

High-Performance FICON for z Systems (zHPF) simplifies and improves protocol efficiency by reducing the number of information units (IU) that are processed. Enhancements to the z/Architecture and the FICON protocol provide optimizations for online transaction processing (OLTP) workloads. zHPF can also be used by z/OS for IBM Db2, VSAM, PDSE, and zFS.

zHPF was further enhanced to allow all large write operations that are greater than 64 KB to be run in a single round trip to the control unit at distances up to 100 km (62.1 miles). This enhancement does not elongate the I/O service for these write operations at extended distances. It is especially useful for IBM GDPS HyperSwap configurations.

Additionally, the changes to the architecture provide end-to-end system enhancements to improve RAS.

zHPF requires matching support by the IBM System Storage® DS8880 series or similar devices from other vendors. FICON Express16S+, FICON Express16S, and FICON Express8S support the FICON protocol and the zHPF protocol in the server Licensed Internal Code.

FICON Forward Error Correction

Even with proper fiber optic cable cleaning discipline, errors can still occur on 16 Gbps links. Forward Error Correction (FEC) is a technique that is used for controlling errors in data transmission over lower quality communication channels. With FEC, I/O errors are decreased, which reduces the potential effect on workload performance that is caused by I/O errors.

When running at 16 Gbps, FICON Express16S+ and FICON Express16S features can use FEC when connected to devices that support FEC, such as the IBM DS8880. FEC allows channels to operate at higher speeds over longer distances and with reduced power and higher throughput, while retaining the same reliability and robustness for which FICON channels are traditionally known.

FICON Dynamic Routing

FICON Dynamic Routing (FIDR) is designed to support the dynamic routing policies that are supplied by FICON Director providers, such as Brocade's Exchange Based Routing (EBR) and Cisco's Open Exchange ID Routing (OxID).

With FIDR, you are no longer restricted to the use of static storage area network (SAN) routing policies for inter-switch links (ISLs) in a cascaded FICON Directors configuration. Performance of FICON and FCP traffic improves because of SAN dynamic routing policies that better use all of the available ISL bandwidth through higher utilization.

The IBM DS8880 also supports FIDR. Therefore, in a configuration with the z14 ZR1, capacity planning and management can be simplified and provide persistent and repeatable performance and higher resiliency.

All devices in the SAN environment must support FICON Dynamic Routing to use this feature.

The z14 ZR1s continue to provide the functions that were introduced on other Z platforms with the supported FICON features. For more information, see *IBM Z Connectivity Handbook,* SG24-5444.

3.3.3 FCP functions

Fibre Channel Protocol (FCP) is fully supported on the z14 ZR1. It is commonly used with Linux on IBM Z and supported by the z/VM and z/VSE. The key FCP functions are described next.

N_Port ID Virtualization

N_Port ID Virtualization (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 goal 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, which enables separating 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.

The following preferred and allowable operating characteristic values in the FCP protocol were increased:

- The preferred maximum number of NPIV hosts defined to any single physical FCP channel increased from 32 to 64.
- The allowable maximum number of remote N_Ports a single physical channel can communicate with increased from 512 to 1024.
- The maximum number of LUNs that is addressable by a single physical channel increased from 4096 to 8192.

In support of these increases, the FCP channels also were designed to now support 1528 concurrent I/O operations, which is an increase from the previous generation FCP channel limit of 764.

Export/import physical port WWPNs for FCP channels

IBM Z platforms automatically assign WWPNs to the physical ports of an FCP channel. This WWPN assignment changes when an FCP channel is moved to a different physical slot position in the I/O drawer. The z14 ZR1 allows for the modification of these default assignments, which permits FCP channels to keep previously assigned WWPNs. This capability eliminates the need for reconfiguration of the SAN environment when a Z platform upgrade occurs or when a FICON Express feature is replaced.

Fibre Channel Read Diagnostic Parameter

An extended link service (ELS) command called Read Diagnostic Parameter (RDP) was added to the Fibre Channel T11 standard to allow Z platforms to obtain more diagnostic data from the Small Form-factor Pluggable (SFP) optics that are throughout the SAN fabric. RDP can identify a failed or failing component without unnecessarily replacing more components in the SAN fabric (such as FICON features, optics, and cables).

FICON and FCP channels provide a means to read this extra diagnostic data for all of the ports that are accessed in the I/O configuration and make the data available to a Z LPAR. For FICON channels, z/OS displays the data with a message and display command. For Linux on IBM Z and the KVM hypervisor, z/VM, and z/VSE, this diagnostic data is made available in a window in the SAN Explorer tool on the Hardware Management Console (HMC).

3.3.4 FICON Express16S+

The following types of transceivers for FICON Express16S+ are supported on a new build system:

- FICON Express16S+ LX feature (long wavelength)
- FICON Express16S+ SX feature (short wavelength)

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

Note: Both ports of FICON Express16S+ adapter must be the same CHPID type (FC or FCP).

FICON Express16S+ LX feature

The FICON Express16S LX feature occupies one I/O slot in the PCIe+ I/O drawer. It features 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 (6.2 miles).

FICON Express16S+ SX feature

The FICON Express16S SX feature occupies one I/O slot in the PCIe+ I/O drawer. It features 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[2] of 380 meters (1246.7 feet) at 4 Gbps, 150 meters (492.1 feet) at 8 Gbps, or 100 meters (328 feet) at 16 Gbps.

3.3.5 FICON Express16S (carry forward only)

The following types of transceivers for FICON Express16S are available only when carried forward on upgrades and are supported on z14 ZR1:

- 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

[2] Distances are valid for OM3 cabling. For more information about available options, see Table 3-5 on page 35.

Note: To permit the mix of different CHPID types (FC and FCP) in the FICON Express16S features, the keyword `MIXTYPE` must be defined in the IODF to at least one port of the adapter.

FICON Express16S LX feature

The FICON Express16S LX feature occupies one I/O slot in the PCIe+ I/O drawer. It features 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 (6.2 miles).

FICON Express16S SX feature

The FICON Express16S SX feature occupies one I/O slot in the PCIe+ I/O drawer. It features 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 380 meters (1246 feet) at 4 Gbps, 150 meters (492 feet) at 8 Gbps, or 100 meters (328 feet) at 16 Gbps.

3.3.6 FICON Express8S (carry forward only)

The FICON Express8S features are available only when carried forward on upgrades. The following types of transceivers for FICON Express8 are supported on z14 ZR1:

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

FICON Express8S 10KM LX feature

The FICON Express8S 10KM LX feature occupies one I/O slot in the I/O drawer. It includes 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 (6.2 miles).

FICON Express8S SX feature

The FICON Express8S SX feature occupies one I/O slot in the I/O drawer. This feature includes 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 500 meters (1640 feet) at 2 Gbps, 380 meters (1246 feet) at 4 Gbps, or 150 meters (492 feet) at 8 Gbps.

3.4 Network connectivity

The z14 ZR1 offers a wide range of functions that can help consolidate or simplify the network environment. These functions are supported by OSA-Express, RoCE Express, and HiperSockets features.

3.4.1 OSA-Express functions

Improved throughput (mixed inbound/outbound) is achieved by the data router function that was introduced in the OSA-Express3, and enhanced in OSA-Express6S and OSA-Express5S 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 increase throughput for standard frames (1492 bytes) and jumbo frames (8992 bytes).

The most current OSA-Express functions are described next.

OSM CHPID for usage with Dynamic Partition Manager

Dynamic Partition Manager (DPM) requires that the z14 ZR1 has two OSA-Express5S 1000BASE-T Ethernet or OSA-Express6S 1000BASE-T Ethernet features that are defined as CHPID type OSM for connectivity. OSA-Express features that are defined with OSM cannot be shared with other CHPID types and must be dedicated for use by DPM. DPM supports Linux on IBM Z (running in an LPAR) under KVM hypervisor or z/VM.

DPM can be ordered along with Ensemble membership, but both cannot be enabled at the same time on the z14 ZR1.

Statement of direction[a]**:** IBM z14 ZR1 is the last platform to support Ensembles and zEnterprise Unified Resource Manager (zManager).

a. IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. The development, release, and timing of any future features or functionality that is described for our products remain at IBM's sole discretion.

OSA-ICC support for Secure Sockets Layer

When configured as an integrated console controller CHPID type (OSC) on the z14 ZR1, the Open Systems Adapter supports the configuration and enablement of secure connections by using the Transport Layer Security (TLS) protocol versions 1.0, 1.1, and 1.2. Server-side authentication is supported by using a self-signed certificate or customer supplied certificate, which can be signed by a customer-specified certificate authority.

The certificates that are used must include an RSA key length of 2048 bits and be signed by using SHA-256. This support negotiates a cipher suite of AES-128 for the session key.

Queued direct I/O optimized latency mode

Queued direct I/O (QDIO) optimized latency mode can help improve performance for applications that feature 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, which ensures that any new data is read from the OSA-Express6S or OSA-Express5S without requiring more program-controlled interrupts.
- For outbound processing, the OSA-Express6S or OSA-Express5S looks more frequently for available data to process from the TCP/IP stack; therefore, a Signal Adapter instruction is not required to determine whether more data is available.

Inbound workload queuing

Inbound workload queuing (IWQ) can help to reduce overhead and latency for inbound z/OS network data traffic and implement an efficient way for initiating parallel processing. This improvement is achieved by using OSA-Express features in QDIO mode (CHPID type OSD) 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). A process is created and scheduled to run on one of the multiple processors, independent from the other three queues. This change can improve performance because IWQ can use the symmetric multiprocessor (SMP) architecture of the Z.

Virtual local area network support

Virtual local area network (VLAN) is a function of OSA-Express features that takes advantage of the Institute of Electrical and Electronics Engineers (IEEE) 802.q 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 the Z, TCP/IP stacks can exist and potentially share 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 support

When sharing OSA port addresses across LPARs, Virtual MAC (VMAC) support enables each operating system instance to have a unique VMAC address. All IP addresses that are 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 can include a simplified configuration setup and improvements to IP workload load balancing and outbound routing.

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

z/VM multi-VSwitch link aggregation support

z/VM provides multi-VSwitch link aggregation support, which allows a port group of OSA-Express features to span multiple virtual switches within a single z/VM LPAR or between multiple z/VM LPARs. Sharing a link aggregation port group (LAG) with multiple virtual switches increases optimization and utilization of the OSA-Express when handling larger traffic loads. With this support, a port group is no longer required to be dedicated to a single z/VM virtual switch.

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 HIPPA-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.
- Use of 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 the Virtual Switch (VSWITCH) function is used, and to all supported OSA-Express features (CHPID type OSD) on Z. 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 supported OSA-Express features on Z.

3.4.2 OSA-Express7S

This section describes the OSA-Express7S 25GbE Short Reach (SR) connectivity feature that is offered on z14 ZR1.

OSA-Express7S 25GbE SR feature

The OSA-Express7S 25GbE SR feature occupies one slot in the PCIe+ I/O drawer. This feature includes one port that connects to a 25 Gbps Ethernet LAN through a 50 µm multimode fiber optic cable that ends with an LC Duplex connector.

The maximum supported unrepeated distance is 100 meters (328 feet) with a 50 µm (at 4700 MHz·km) multimode fiber optic cable and 70 meters (229.6 feet) with a 50 µm (at 2000 MHz·km) multimode fiber optic cable.

3.4.3 OSA-Express6S

This section describes the connectivity options that are offered by the OSA-Express6S features. The following OSA-Express6S features can be installed on z14 ZR1:

- OSA-Express6S 10GbE Long Reach (LR)
- OSA-Express6S 10GbE Short Reach (SR)
- OSA-Express6S Gigabit Ethernet Long Wavelength (GbE LX)
- OSA-Express6S Gigabit Ethernet Short Wavelength (GbE SX)
- OSA-Express6S 1000BASE-T Ethernet

OSA-Express6S 10GbE LR feature

The OSA-Express6S 10GbE LR feature occupies one slot in a PCIe+ I/O drawer. It includes one port that connects to a 10 Gbps Ethernet LAN through a 9 µm single mode fiber optic cable that ends with an LC Duplex connector. The feature supports an unrepeated maximum distance of 10 km (6.2 miles).

OSA-Express6S 10GbE SR feature

The OSA-Express6S 10GbE SR feature occupies one slot in the PCIe+ I/O drawer. This feature includes one port that connects to a 10 Gbps Ethernet LAN through a 62.5 µm or 50 µm multimode fiber optic cable that ends with an LC Duplex connector.

The maximum supported unrepeated distance is 33 m (108 feet) on a 62.5 µm multimode fiber optic cable, and 300 m (984 feet) on a 50 µm multimode fiber optic cable.

OSA-Express6S GbE LX feature

The OSA-Express6S GbE LX occupies one slot in the PCIe+ I/O drawer. This feature includes two ports, which represent one channel path identifier (CHPID), that connect to a 1 Gbps Ethernet LAN through a 9 µm single mode fiber optic cable. This cable ends with an LC Duplex connector, which supports an unrepeated maximum distance of 5 km (3.1 miles).

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 (1804 feet).

OSA-Express6S GbE SX feature

The OSA-Express6S GbE SX occupies one slot in the PCIe+ I/O drawer. This feature includes two ports (which represent one CHPID) that connect to a 1 Gbps Ethernet LAN through 50 or 62.5 µm multimode fiber optic cable. This cable ends with an LC Duplex connector over an unrepeated distance of 550 meters (1804 feet) for 50 µm fiber or 220 meters (721 feet) for 62.5 µm fiber.

OSA-Express6S 1000BASE-T feature

The OSA-Express6S 1000BASE-T occupies one slot in the PCIe+ I/O drawer. It features two ports (which represent one CHPID) that connect to a 1000 Mbps (1 Gbps) or 100 Mbps Ethernet LAN. Each port has an RJ-45 receptacle for UTP Cat5 or Cat6 cabling, which supports a maximum distance of 100 meters (328 feet).

Statement of Direction[a]**:** The OSA-Express6S 1000BASE-T feature is the last generation to support connections that operate at 100 Mbps link speed. Future OSA-Express 1000BASE-T features will support operation at only 1 Gbps link speed.

a. IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. The development, release, and timing of any future features or functionalit)y described for our products remain at IBM's sole discretion.

3.4.4 OSA-Express5S (carry forward only)

This section describes the connectivity options that are offered by the OSA-Express5S features. The following OSA-Express5S features can be installed on z14 ZR1:

- OSA-Express5S 10GbE Long Reach (LR)
- OSA-Express5S 10GbE 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 10GbE LR feature

The OSA-Express5S 10GbE LR feature occupies one slot in a PCIe+ I/O drawer. It features one port that connects to a 10 Gbps Ethernet LAN through a 9 µm single mode fiber optic cable that ends with an LC Duplex connector. The feature supports an unrepeated maximum distance of 10 km (6.2 miles).

OSA-Express5S 10GbE SR feature

The OSA-Express5S 10GbE SR feature occupies one slot in the PCIe+ I/O drawer. This feature includes one port that connects to a 10 Gbps Ethernet LAN through a 62.5 µm or 50 µm multimode fiber optic cable that ends with an LC Duplex connector. The maximum supported unrepeated distance is 33 m (98 feet) on a 62.5 µm multimode fiber optic cable, and 300 m (984 feet) on a 50 µm 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 includes two ports (which represent one CHPID) that connect to a 1 Gbps Ethernet LAN through a 9 µm single mode fiber optic cable. This cable ends with an LC Duplex connector, which supports an unrepeated maximum distance of 5 km (3.1 miles).

A multimode (62.5 or 50 µm) fiber optic cable can be used with this feature. The use of these multimode cable types requires an 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 (1804 feet).

OSA-Express5S GbE SX feature

The OSA-Express5S GbE SX occupies one slot in the PCIe+ I/O drawer. This feature includes two ports (which represent one CHPID) that connect to a 1 Gbps Ethernet LAN through 50 or 62.5 µm multimode fiber optic cable. This cable ends with an LC Duplex connector over an unrepeated distance of 550 meters (1804 feet) for 50 µm fiber or 220 meters (721 feet) for 62.5 µm fiber.

OSA-Express5S 1000BASE-T feature

The OSA-Express5S 1000BASE-T occupies one slot in the PCIe+ I/O drawer. It features two ports (which represent one CHPID) that connect to a 1000 Mbps (1 Gbps) or 100 Mbps Ethernet LAN. Each port includes an RJ-45 receptacle for UTP Cat5 or Cat6 cabling, which supports a maximum distance of 100 meters (328 feet).

3.4.5 OSA-Express4S (carry forward only, select features)

This section describes the connectivity options that are offered by the OSA-Express4S features. The following OSA-Express4S features can be installed on z14 ZR1:

- OSA-Express4S 10GbE Long Reach (LR)
- OSA-Express4S 10GbE Short Reach (SR)
- OSA-Express4S Gigabit Ethernet Long Wavelength (GbE LX)
- OSA-Express4S Gigabit Ethernet Short Wavelength (GbE SX)

OSA-Express4S 10GbE LR feature

The OSA-Express4S 10GbE LR feature occupies one slot in a PCIe+ I/O drawer. It features one port that connects to a 10 Gbps Ethernet LAN through a 9 µm single mode fiber optic cable that ends with an LC Duplex connector. The feature supports an unrepeated maximum distance of 10 km (6.2 miles).

OSA-Express4S 10GbE SR feature

The OSA-Express4S 10GbE SR feature occupies one slot in the PCIe+ I/O drawer. This feature includes one port that connects to a 10 Gbps Ethernet LAN through a 62.5 µm or 50 µm multimode fiber optic cable that ends with an LC Duplex connector. The maximum supported unrepeated distance is 33 m (108 feet) on a 62.5 µm multimode fiber optic cable, and 300 m (984 feet) 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 includes two ports (which represent one CHPID) that connect to a 1 Gbps Ethernet LAN through a 9 µm single mode fiber optic cable. This cable ends with an LC Duplex connector that supports an unrepeated maximum distance of 5 km (3.1 miles).

A multimode (62.5 or 50 µm) fiber optic cable can be used with this feature. The use of these multimode cable types requires an 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 (1804 feet).

OSA-Express4S GbE SX feature

The OSA-Express4S GbE SX occupies one slot in the PCIe+ I/O drawer. This feature includes two ports (which represent one CHPID) that connect to a 1 Gbps Ethernet LAN through 50 or 62.5 µm multimode fiber optic cable. This cable ends with an LC Duplex connector over an unrepeated distance of 550 meters (1804 feet) for 50 µm fiber or 220 meters (721 feet) for 62.5 µm fiber.

3.4.6 HiperSockets functions

IBM HiperSockets are referred to as the “network in a box” because it simulates LAN environments entirely within the IBM Z platform. The data transfer is from LPAR memory to LPAR memory, which is mediated by IBM Z firmware. The z14 ZR1 supports up to 32 HiperSockets. One HiperSockets network can be shared by up to 40 LPARs. Up to 4096 communication paths support a total of 12,288 IP addresses across all 32 HiperSockets.

The HiperSockets internal networks can support the following transport modes:

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

Traffic can be Internet Protocol Version 4 or Version 6 (IPv4, IPv6) or non-IP. HiperSockets devices are independent of protocol and Layer 3. Each HiperSockets device has its own Layer 2 Media Access Control (MAC) address. This address 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 IBM Z, and by z/VM for guest use.

The most current HiperSockets functions are described in the following sections.

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 eligible for zIIP processing. 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.

HiperSockets network traffic analyzer

HiperSockets network traffic analyzer (NTA) is a function that is available in the Licensed Internal Code (LIC) of IBM Z systems. It can simplify problem isolation and resolution by allowing Layer 2 and Layer 3 tracing of HiperSockets network traffic.

HiperSockets NTA allows Linux on IBM Z 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. These administrators can use Linux on IBM Z 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 can be traced by the NTA on the selected IQD channel.

HiperSockets completion queue

The HiperSockets completion queue function allows synchronous and asynchronous transfer of data between logical partitions. With the asynchronous support, data can be temporarily held until the receiver has buffers available in its inbound queue during high volume situations. This function can provide performance improvement for LPAR-to-LPAR communication, and can be especially helpful in burst situations.

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

A HiperSockets channel alone can provide intra-CPC communications only. The HiperSockets Bridge Port allows a virtual switch to connect z/VM guests by using real HiperSockets devices, which provides the ability to communicate with hosts that are external 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.

3.4.7 Shared Memory Communications functions

The Shared Memory Communication (SMC) capabilities of the z14 ZR1 optimizes the communications between applications in server-to-server (SMC-R) or LPAR-to-LPAR (SMC-D) connectivity.

SMC-R provides application transparent use of the RoCE Express features that can reduce the network overhead and latency of data transfers, which effectively offers the benefits of optimized network performance across processors.

The Internal Shared Memory (ISM) virtual PCI function takes advantage of the capabilities of SMC-D. ISM is a virtual PCI network adapter that enables direct access to shared virtual memory, which provides a highly optimized network interconnect for Z intra-system communications. Up to 32 channels for SMC-D traffic can be defined in a z14 ZR1, whereby each channel can be virtualized to a maximum of 255 Function IDs[3]. No other hardware is required for SMC-D.

[3] The 10GbE RoCE features and the ISM adapters are identified by a hexadecimal Function Identifier (FID) with a range of 00 - FF.

3.4.8 RoCE Express features

This section describes the connectivity options that are offered by the Remote Direct Memory Access (RDMA) over Converged Ethernet (RoCE) Express features. The following RoCE features can be installed in the z14 ZR1:

- 25GbE RoCE Express2
- 10GbE RoCE Express2
- 10GbE RoCE Express (carry forward only)

The RoCE Express features help reduce the use of CPU resources for applications that use the TCP/IP stack. It might also help to reduce network latency with memory-to-memory transfers that use SMC-R in z/OS V2R1 and later versions. It is transparent to applications, and can be used for server-to-server communication in a multiple Z platform environment.

These features are installed in the PCIe+ I/O drawer and use a short reach (SR) laser as the optical transceiver. Both point-to-point connections and switched connections with an Ethernet switch are supported. Ethernet switches must include the *Pause frame* enabled as defined by the IEEE 802.3x standard.

A maximum of *eight* features (any combination of 25GbE RoCE Express2, 10GbE RoCE Express2, or 10GbE RoCE Express features) can be installed in the z14 ZR1.

Note: 25GbE RoCE Express2 should not be mixed with 10GbE RoCE Express2 or 10GbE RoCE Express in the same SMC-R link group. 10GbE RoCE Express2 can be mixed with 10GbE RoCE Express in the same SMC-R link group.

25GbE RoCE Express2

The 25GbE RoCE Express2 feature occupies one slot in the PCIe+ I/O drawer. This feature includes two ports that connect to a 25 Gbps Ethernet LAN through a 50 µm multimode fiber optic cable that ends with an LC Duplex connector.

The maximum supported unrepeated distance is 100 meters (328 feet) with an OM4 multimode fiber optic cable and 70 meters (229.6 feet) with an OM3 multimode fiber optic cable. The 25GbE RoCE Express2 supports 63 Virtual Functions (VFs)[4] per physical port for a total of 126 per adapter.

10GbE RoCE Express2

The 10GbE RoCE Express2 feature occupies one slot in the PCIe+ I/O drawer. This feature includes two ports that connect to a 10 Gbps Ethernet LAN through a 62.5 µm or 50 µm multimode fiber optic cable that ends with an LC Duplex connector.

The maximum supported unrepeated distance is 300 meters (984 feet) on an OM3 multimode fiber optic cable, and can be increased to 600 meters (1,968 feet) when a switch is shared across two 10GbE RoCE Express2 features. The 10GbE RoCE Express2 supports 63 Virtual Functions (VFs) per physical port for a total of 126 per adapter.

[4] Virtual Function ID is defined when PCIe hardware or the ISM is shared between LPARs.

10GbE RoCE Express (carry forward only)

The 10GbE RoCE Express feature occupies one slot in the PCIe+ I/O drawer. This feature includes two ports that connect to a 10 Gbps Ethernet LAN through a 62.5 µm or 50 µm multimode fiber optic cable that ends with an LC Duplex connector. The maximum supported unrepeated distance is 300 m (984 feet) on an OM3 multimode fiber optic cable, and can be increased to 600 m (1968 feet) when sharing a switch across two RoCE Express features. The RoCE Express supports 31 VFs per feature.

3.5 Compression options

Two types of compression options are available with the z14 ZR1: A standard internal compression coprocessor that is tightly connected to each processor unit, and an external native PCIe feature.

3.5.1 Compression Coprocessor

The Compression Coprocessor (CMPSC) is a well-known feature that works with the processor units in the Z platform. This coprocessor works with a proprietary compression format and is used for many types of z/OS data.

3.5.2 zEnterprise Data Compression

zEnterprise Data Compression (zEDC) Express is an optional native PCIe feature that is available in the z14 ZR1. It provides hardware-based acceleration for data compression and decompression for the enterprise, which can help 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 16 can be installed on the system in the PCIe+ I/O drawer. Up to two zEDC Express features per domain can be installed. One PCIe adapter/compression processor is available per feature that implements compression as defined by RFC1951 (DEFLATE). A zEDC Express feature can be shared between up to 15 LPARs.

3.6 Cryptographic features

The z14 ZR1 provides cryptographic functions that, from an application program perspective, can be categorized in the following groups:

- Synchronous cryptographic functions, which are provided by the CP Assist for Cryptographic Function (CPACF) or the Crypto Express features when defined as an accelerator.
- Asynchronous cryptographic functions, provided by the Crypto Express features.

3.6.1 CP Assist for Cryptographic Function

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 FIPS 140-2 level 4 security[5]. The CPACF is an optional feature that is integrated with the compression unit in the coprocessor in the z14 ZR1 microprocessor core.

The CPACF protected key is a function that facilitates 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:

- Advanced Encryption Standard (AES) for 128-bit, 192-bit, and 256-bit keys
- Data Encryption Standard (DES) data encryption and decryption with single, double, or triple length keys.
- Pseudo-random number generation (PRNG)
- True-random number generator (TRNG)
- Message authentication code (MAC)
- Hashing algorithms: SHA-1, SHA-2, and SHA-3

SHA-1, SHA-2, and SHA-3 support are enabled on all Z platforms 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 IBM Z.

For information about the use of this function, see “Pervasive encryption” on page 82.

3.6.2 Crypto Express6S

The Crypto Express6S represents the newest generation of the Peripheral Component Interconnect Express (PCIe) cryptographic coprocessors, which are an optional feature that is available on the z14 ZR1. These coprocessors are Hardware Security Modules (HSMs) that provide high-security cryptographic processing as required by banking and other industries.

This feature provides a secure programming and hardware environment wherein crypto processes are performed. Each cryptographic coprocessor includes general-purpose processors, non-volatile storage, and specialized cryptographic electronics, which are all contained within a tamper-sensing and tamper-responsive enclosure that eliminates all keys and sensitive data on any attempt to tamper with the device. The security features of the HSM are designed to meet the requirements of FIPS 140-2, Level 4, which is the highest security level defined.

The Crypto Express6S includes one PCIe adapter per feature. For availability reasons, a minimum of two features is required. Up to 16 Crypto Express6S features are supported. The Crypto Express6S feature occupies one I/O slot in a 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 Express6S provides domain support for up to 40 logical partitions.

[5] Federal Information Processing Standards (FIPS) 140-2 Security Requirements for Cryptographic Modules

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 Express6S can also be configured as one of the following configurations:

- The Secure IBM CCA coprocessor includes secure key functions with emphasis on the specialized functions that are required for banking and payment card systems. It is optionally programmable to add custom functions and algorithms by using User Defined Extensions (UDX).

  A new mode, called Payment Card Industry (PCI) PIN Transaction Security (PTS) Hardware Security Module (HSM) (PCI-HSM), is available exclusively for Crypto Express6S in CCA mode. PCI-HSM mode simplifies compliance with PCI requirements for hardware security modules.

- The Secure IBM Enterprise PKCS #11 (EP11) coprocessor implements 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 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 Express6S when configured as an Enterprise PKCS #11 coprocessor or managing the CCA mode PCI-HSM.

When the Crypto Express6S PCI Express adapter is configured as a secure IBM CCA co-processor, it still provides accelerator functions. However, up to 3x better performance for those functions can be achieved if the Crypto Express6S PCI Express adapter is configured as an accelerator.

CCA enhancements include the ability to use triple-length (192-bit) Triple-DES (TDES) keys for operations, such as data encryption, PIN processing, and key wrapping to strengthen security. CCA also extended the support for the cryptographic requirements of the German Banking Industry Committee Deutsche Kreditwirtschaft (DK).

Several features that support the use of the AES algorithm in banking applications also were added to CCA. These features include the addition of AES-related key management features and the AES ISO Format 4 (ISO-4) PIN blocks as defined in the ISO 9564-1 standard. PIN block translation is supported as well as usage of AES PIN blocks in other CCA callable services. IBM continues to add enhancements as AES finance industry standards are released.

3.6.3 Crypto Express5S (carry forward only)

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. The Crypto Express5S feature occupies one I/O slot in a 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 provides domain support for up to 40 logical partitions on IBM z14 ZR1.

The Crypto Express5S feature supports all the functions of the Crypto Express6S, except the PCI-HSM standard.

3.7 Coupling and clustering

Coupling connectivity for Parallel Sysplex on z14 ZR1 use CE LR and Integrated Coupling Adapter Short Reach (ICA SR). The ICA SR is designed to support distances up to 150 m (492 feet). The CE LR supports longer distances between systems, up to 10 km (6.2 miles) unrepeated.

CE LR and ICA SR allow all of the z/OS-to-CF communication, CF-to-CF traffic, or Server Time Protocol (STP)[6] through high-speed fiber optic connections at short, which are up to 150 m (492 feet) or long, which are up to 10 km (6.2 miles) unrepeated distances.

Important: zEC12 or zBC12 can coexist in the same parallel sysplex with z14 ZR1 only if the CPC that is hosting the CFs includes coupling connectivity to the zEC12 or zBC12 and z14 ZR1. z14 ZR1 does not support direct coupling connectivity to zEC12, zBC12, or older Z platforms.

For more information about coupling links technologies, see the *Coupling Facility Configuration Options* white paper.

Note: The z14 ZR1 does *not* support InfiniBand coupling connectivity. For more information, see the Statement Of Direction from *Hardware Announcement 117-031* - fulfilled.

3.7.1 Coupling Express Long Reach

The CE LR is a two-port PCIe native adapter that is used for long-distance coupling connectivity. CE LR uses the CL5 coupling channel type. The CE LR feature also uses PCIe Gen3 technology and is hosted in a PCIe+ I/O drawer.

The feature supports communication at unrepeated distances up to 10 km (6.2 miles) by using 9 µm single-mode fiber optic cables and repeated distances up to 100 km (62 miles) by using IBM Z qualified DWDM vendor equipment. It supports up to 4 CHPIDs per port and 8 or 32 subchannels (devices) per CHPID. The coupling links can be defined as shared between images within a CSS or spanned across multiple CSSs in a Z system.

3.7.2 Integrated Coupling Adapter Short Reach

The Integrated Coupling Adapter Short Reach (ICA SR) is a two-port fanout in the CPC drawer. It is used for short distance coupling connectivity and includes the coupling channel type CS5. The ICA SR uses PCIe Gen3 technology, with x16 lanes that are bifurcated into x8 lanes for coupling.

The ICA SR supports cable length of up to 150 m (492 feet) and supports a link data rate of 8 GBps. It also supports up to four CHPIDs per port and eight 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 Z system.

[6] All coupling links can be used to carry STP timekeeping information.

3.7.3 Internal coupling

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

3.7.4 Coupling Facility Control Code Level

Various levels of Coupling Facility Control Code (CFCC) are available for the z14 ZR1. For more information, see this IBM Knowledge Center web page.

CF structure size changes can be expected when moving from one CFCC level to another. Review the CF LPAR size by using the following tools:

- The CFSizer tool is a web-based and is useful when a workload is changed or introduced
- The Sizer Utility, which is an authorized z/OS program download, is useful when upgrading a CF

3.8 Virtual Flash Memory

IBM Virtual Flash Memory (VFM) is the replacement for the Flash Express features that were available on the z13s and zBC12. On z14 ZR1, the VFM feature (0614) can be ordered in 512 GB increments up to 2 TB in total.

VFM is designed to help improve availability and handling of paging workload spikes when z/OS V2.1, V2.2, or V2.3 is run. With this support, z/OS is designed to help improve system availability and responsiveness by using VFM across transitional workload events, such as market openings, and diagnostic data collection. z/OS is also designed to help improve processor performance by supporting middleware exploitation of pageable large (1 MB) pages.

VFM can also be used in CF images to provide extended capacity and availability for workloads that use WebSphere MQ Shared Queues structures. The use of VFM can help availability by reducing latency from paging delays that can occur at the start of the workday or during other transitional periods. It is also designed to eliminate delays that can occur when diagnostic data is collected during failures.

Therefore, VFM can help meet most demanding service level agreements and compete more effectively. VFM is easy to configure and provides rapid time to value.

No application changes are required to migrate from IBM Flash Express to VFM.

3.9 Server Time Protocol

Each system 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 z14 ZR1 supports the Server Time Protocol (STP) and can participate in a STP-only Coordinated Timing Network (CTN). A CTN is a collection of Z platforms 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.

Important: If a z14 ZR1 plays a CTN role (PTS/BTS/Arbiter), the other CTN role playing Z platforms must include direct coupling connectivity to the z14 ZR1. The z14 ZR1 does not support direct coupling connectivity to zEC12, zBC12, or older Z platforms.

STP is a message-based protocol in which timekeeping information is passed over coupling 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 z14 ZR1 and CF in sysplex environments.

STP is implemented in LIC as a system-wide facility of the z14 ZR1 and other Z platforms. 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 z14 ZR1 is enabled for STP by installing the STP feature code. Extra configuration is required for a z14 ZR1 to become a member of a CTN.

For high availability purposes, non-disruptive capability were implemented in the z14 ZR1 firmware that allows two CTNs to be merged into one, or to split one CTN into two, dynamically.

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.

Network Time Protocol client support

The use of Network Time Protocol (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 z14 ZR1. The code interfaces with the NTP servers. This interaction allows an NTP server to become the single-time source for z14 ZR1 and for other servers that have NTP clients.

Pulse per second support

Two oscillator cards (OSCs), which are included as a standard feature of the z14 ZR1, provide a dual-path interface for the pulse per second (PPS) signal. The cards contain a BNC connector for PPS attachment at the rear side of the CPC drawer. The redundant design allows continuous operation during the 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 µs as measured at the PPS input of the z14 ZR1.

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 z14 ZR1 is configured for the use of NTP with PPS as the ETS for time synchronization.

For more information about STP, see the following publications:

- *Server Time Protocol Planning Guide*, SG24-7280
- *Server Time Protocol Implementation Guide*, SG24-7281
- *Server Time Protocol Recovery Guide*, SG24-7380

3.10 Hardware Management Console and Support Element

The HMC and SE are appliances that provide hardware management for IBM Z platforms. Hardware platform management covers a complex set of configuration, operation, monitoring, service management tasks, and other services that are essential to the operations of the Z platform.

With z14 ZR1, the HMC can be a stand-alone desktop computer or an optional 1U rack-mounted computer.

Important: IBM z14 ZR1 is planned to be the last platform that allows HMC support across the previous four generations of Z platforms (N - N-4). Future HMC releases are intended to be tested for support of the previous two generations (N - N-2).

The z14 ZR1 is supplied with a pair of integrated 1U SEs. The primary SE is always active; the other SE is an alternative. Power for the SEs is supplied by the rack PDUs. Each support Element features dual PSU units, 1+1 redundant.

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

The SEs are connected to Ethernet switches for network connectivity with the CPC and the HMCs. An HMC communicates with one or more Z platforms, as shown in Figure 3-1.

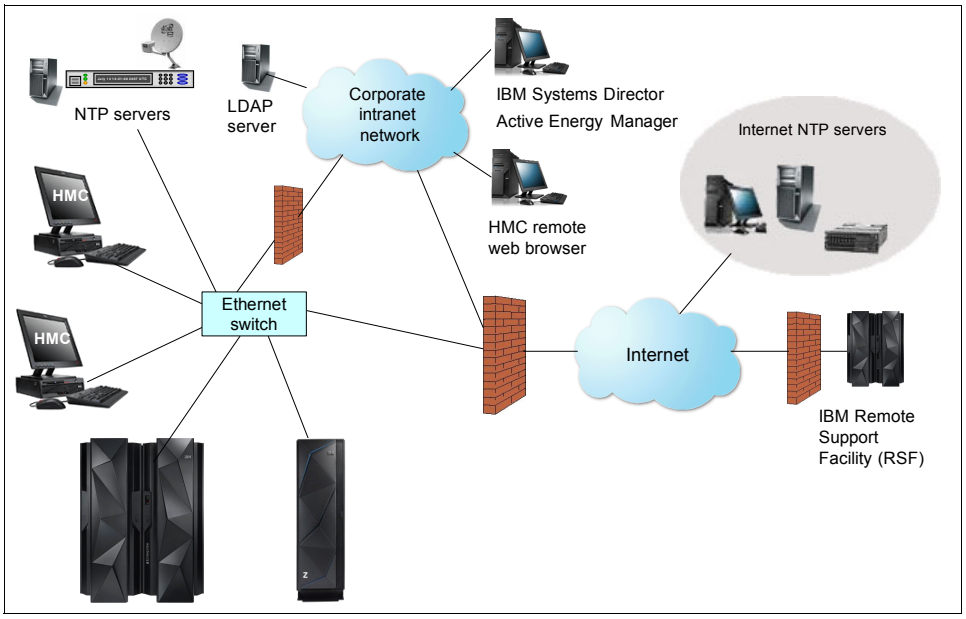

Figure 3-1 HMC and SE connectivity

When tasks are performed on the HMC, the commands are sent to one or more SEs, which then issue commands to their CPCs.

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

Note: An RSF connection through a modem is *not* supported on the z14 ZR1 HMC. An internet connection to IBM is required to enable hardware problem reporting and service.

HMC/SE Version 2.14.0 or later is required for the z14 ZR1. For information about the key capabilities and enhancements of the HMC, see *IBM z14 ZR1 Technical Guide,* SG24-8651.

4

Strengths of the z14 ZR1

Computer systems that stay relevant for more than 50 years demonstrate a forward-looking system architecture. IBM Z platforms are a prime example of this idea with the IBM z/Architecture[1].

Advanced capabilities are introduced with each new Z platform that cause it to move ahead of its predecessor platforms in terms of efficiency, flexibility, security, reliability, and much more. Whenever new capabilities are implemented, the z/Architecture is extended rather than replaced, which helps sustain the compatibility, integrity, and longevity of the Z platform. Therefore, *protection* and *compatibility with an earlier version* of existing workloads and solutions are key for the new capabilities that are introduced in the z/Architecture.

To handle new and different workloads, the scope of software and application options must be accommodated by the operating system. Also, the hardware and firmware components of the system must provide a viable option to integrate functionality into the architecture's capabilities. The Z platforms and operating systems always conform to the z/Architecture to ensure support of current and future workloads and solutions.

This chapter highlights several z14 ZR1 capabilities and strengths, and explains how they can be of value for businesses and organizations. Throughout the chapter, reference is made to the following IBM Redbooks publications:

- *IBM z14 Model ZR1 Technical Guide*, SG24-8651
- *IBM Z Connectivity Handbook*, SG24-5444

This chapter includes the following topics:

[1] IBM z/Architecture is the conceptual structure of the Z platform that determines its basic behavior. The architecture was first introduced as System/360 in 1964.

4.1 Technology improvements

The z14 ZR1 includes technology improvements that are intended to make systems integration more scalable, flexible, manageable, and securable.

The following sections provide more information about the technology improvements for the z14 ZR1.

4.1.1 Processor design highlights

The z/Architecture that underlies the z14 ZR1 offers a rich complex instruction set computer (CISC) that supports multiple arithmetic formats.

z/Architecture addressing modes: The z/Architecture simultaneously supports 24-bit, 31-bit, and 64-bit addressing modes. This feature provides compatibility with earlier versions and, with that, investment protection.

Compared to its predecessor system, the z14 ZR1 processor design includes the following improvements and architectural extensions:

- Better performance and throughput:
  - Faster processor units (4.5 GHz compared to 4.3 GHz in the z13s).
  - More capacity (up to 30 characterizable processor units versus 20 on the z13s).
  - Larger cache (and shorter path to cache) means faster uniprocessor performance
  - Innovative core-cache design (L1 and L2), processor chip-cache design (L3), and cluster design (L4). The objective is to keep more data closer to the processor by increasing the cache sizes and decreasing the latency to access the next levels of cache.
- Reoptimized pipeline depth for power and performance:
  - Improved instruction delivery
  - Faster branch wakeup
  - Reduced execution latency
  - Improved Operand Store Compare (OSC) avoidance on Dispatch Store Table (DST)
  - Optimized second-generation SMT2
- New translation design:
  - Four concurrent translations (from one in the z13s)
  - Reduced latency
  - Lookup that is integrated into L2 access pipe
  - Translation Lookaside Buffer enhancements:
    - 2x CRSTE growth
    - 1.5x PTE growth
    - New 64 entry 2 GB
- Dedicated co-processor for each processor unit (PU):
  - The Central Processor Assist for Cryptographic Function (CPACF) in the z14 ZR1 Model ZR1 is optimized to provide up to 6x faster encryption functions than the z13s. CPACF on z14 ZR1 supports new SHA-3 standard, True Random Number Generator, and 4x Advanced Encryption Standard (AES) speedup.

– On-chip Compression CMPSC on z14 ZR1 offers up to 2x expansion speed up and supports Entropy Encoding (Huffman Coding) and Order Preserving Compression for index/sort-file compression.

Transactional Execution Facility

The Transactional Execution Facility, which is known in the industry as *hardware transactional memory*, allows instructions to be issued atomically. Therefore, *all results* of the instructions in the group are committed or *no results* are committed, in a truly transactional manner. The execution is optimistic.

The instructions are issued, but previous state values are saved in transactional memory. If the transaction succeeds, the saved values are discarded. If it fails, they are used to restore the original values. Software can test the success of execution and redrive the code, if needed, by using the same path or a different path.

The Transactional Execution Facility provides several instructions, including instructions to declare the start and end of a transaction and to cancel the transaction. This capability can provide performance benefits and scalability to workloads by helping to avoid most of the locks on data. This ability is especially important for heavily threaded applications, such as Java.

Guarded Storage Facility

Also known as less-pausing garbage collection, Guarded Storage Facility is a new architecture that was introduced with z14 ZR1 to enable enterprise scale Java applications to run without periodic pause for garbage collection on larger heaps. This facility improves Java performance by reducing program pauses during Java Garbage Collection.

Instruction Execution Protection

Instruction Execution Protection (IEP) is a hardware function on the z14 ZR1 that enables software, such as Language Environment, to mark certain memory regions (for example, a heap or stack) as non-executable to improve the security of programs that are running on Z against stack-overflow or similar attacks.

Simultaneous multithreading

Simultaneous multithreading (SMT) is built into the z14 ZR1 IFLs, zIIPs, and system assist processors (SAPs), which allows more than one thread to simultaneously run in the same core and shares all of its resources. This function improves utilization of the cores and increases processing capacity.

Adjusted with the growth in the core cache and TLB2, second-generation SMT on z14 ZR1 improves thread balancing, supports multiple outstanding translations, optimizes hang avoidance mechanisms, and delivers improved virtualization performance to benefit Linux. z14 ZR1 provides economies of scale with next generation multithreading (SMT) for Linux and zIIP-eligible workloads while adding support for the I/O SAP.

Hardware decimal floating point function

The hardware decimal floating point (HDFP) function is designed to speed up calculations and provide the precision demanded by financial institutions and others. The HDFP fully implements the IEEE 754r standard.

Vector Packed Decimal Facility

Vector Packed Decimal Facility allows packed decimal operations to be performed in registers rather than memory by using new fast mathematical computations. Compilers, such as Enterprise COBOL for z/OS, V6.2, Enterprise PL/I for z/OS, V5.2, z/OS V2.3 XL C/C++, the COBOL optimizer, Automatic Binary Optimizer for z/OS, V1.3, and Java, are optimized on z14 ZR1.

Single instruction, multiple data

The z14 ZR1 includes a set of instructions called single instruction, multiple data (SIMD) that can improve the performance of complex mathematical models and analytics workloads. This improvement is accomplished through vector processing and complex instructions that can process a large volume of data with a single instruction.

SIMD is designed for parallel computing and can accelerate code that contains integer, string, character, and floating point data types. This system enables better consolidation of analytics workloads and business transactions on the Z platform.

Runtime Instrumentation Facility

The Runtime Instrumentation Facility provides managed run times and just-in-time compilers with enhanced feedback about application behavior. This capability allows dynamic optimization of code generation as it is being run.

Large page support

The size of pages and page frames remained at 4 KB for a long time. IBM Z platforms can include large pages of 1 MB, in addition to supporting pages of 4 KB. This capability relates primarily to large main storage usage. VFM supports large pages and can provide increased performance. Both page frame sizes can be used simultaneously.

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. Large pages are better represented in the TLB and are expected to perform better.

Note: Large pages can benefit long-running applications that are memory-access intensive and might not be the best fit for general use. Short-lived processes with small working sets see little to no improvement. Base the decision to use large pages on measurements of memory usage and page translation overhead for specific workloads.

Support for 2 GB large pages

z14 ZR1 uses 2 GB page frames to increase efficiency for DB2 buffer pools, Java heaps, and other large structures. The use of 2 GB pages increases TLB coverage without proportional growth in the size of the TLB. Consider the following points:

- A 2 GB memory page is 2048x larger than a 1 MB large page, and 524,288x larger than an ordinary 4 KB base page.
- A 2 GB page allows a single TLB entry to fulfill many more address translations than a large page or ordinary base page.
- A 2 GB page provides users with much better TLB coverage, which improves the following aspects of performance:
  - Decreases the number of TLB misses that an application incurs
  - Spends less time converting virtual addresses into physical addresses
  - Uses less real storage to maintain DAT structures

Central Processor Assist for Cryptographic Function

CPACF is a high-performance, low-latency co-processor that can use DES, TDES, AES-128, AES-256, SHA-1, SHA-2, and SHA-3 ciphers to perform symmetric key encryption and calculate message digests in hardware. It is well-suited for encrypting large amounts of data in real time because of its proximity to the processor unit. For the z14 ZR1 Model ZR1, CPACF encryption modes are accelerated 4 - 6x over the z13s.

Compression Coprocessor

Compression Coprocessor (CMPSC) is a high-performance coprocessor that uses compression algorithms to help reduce disk space and memory usage. Each processor unit features a dedicated CMPSC that connects to the main cache-structure for better throughput of the compression dictionaries.

In the z14 ZR1, the compression and expansion performance are improved with fewer CPU cycles. In addition, the compression ratio with Huffman coding garners more disk space and memory usage savings, even where compression is in use. Also, order-preserving compression for search trees and sort files can be used for large parts of data and DB2 indexes that were not practical to compress previously.

4.1.2 Memory

Memory is significantly greater in the new Z models. The z14 ZR1 can have up to 8 TB of usable memory installed, compared with the 4 TB maximum on the z13s.

In addition, the hardware system area (HSA) on the z14 ZR1 is expanded to 64 GB (from 40 GB on z13s). The HSA includes a fixed size and is not counted in the memory that the client orders.

The maximum memory size per logical partition (LPAR) Also changed. For example, on the z14 ZR1, up to 8 TB of memory can now be defined to an LPAR in the image profile. Each operating system can allocate main storage up to the maximum memory amount supported.

Plan-ahead memory

If you anticipate someday increasing the installed memory, the initial system order can contain starting and potential extra memory sizes. The extra memory is referred to as *plan-ahead memory*, which includes a specific memory pricing model to support it.

The starting memory size is activated when the system is installed, 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 might be nondisruptive to applications, depending on the level of operating system support. z/OS and z/VM support this function.

Note: Do not confuse *plan-ahead* and *flexible*[a] *memory* support. Consider the following points:

- Plan-ahead memory is for a permanent increase of installed memory.
- Flexible memory provides a temporary replacement of a part of memory that becomes unavailable.

a. Flexible memory option is not supported on z14 ZR1 (which includes a single CPC drawer).

IBM Virtual Flash Memory

The Virtual Flash Memory (VFM) feature is offered from the main memory capacity. For z14 ZR1, up to four VFM features can be ordered, each of 512 GB. VFM replaces the Flash Express adapters that were available on the zBC12 and z13s.

VFM provides much simpler management and better performance by eliminating the I/O of the adapters that are in the PCIe drawers. VFM does not require any application changes when moving from IBM Flash Express.

VFM can help improve availability and handling of paging workload spikes when z/OS is run. VFM can also be used in coupling facility images to provide extended capacity and availability for workloads that use WebSphere MQ Shared Queues structures.

VFM can improve availability by reducing latency from paging delays that can occur during peak workload periods. It is also designed to help eliminate delays that can occur when diagnostic data is collected during failures.

4.2 Virtualization

Virtualization is a key strength of Z platforms. It is embedded in the architecture and built into the hardware, firmware, and operating systems. For decades, Z platforms were designed based on the concept of partitioning resources (such as CPU, memory, storage, and network resources) so that each set of features can be used independently with its own operating environment.

Virtualization requires a *hypervisor*, which is the control code that manages resources that are required for multiple independent operating system images. Hypervisors can be implemented as software or hardware, and the z14 ZR1 has both.

The hardware hypervisor is IBM Processor Resource/System Manager (PR/SM). PR/SM is implemented in firmware as part of the base system, fully virtualizes the system resources, and runs without any other software.

A software hypervisor is implemented with the z/VM operating system or the KVM hypervisor[2], both of which use PR/SM functions.

The hypervisors are designed to enable simultaneous execution of multiple operating systems, which provides operating systems with virtual resources.

[2] IBM is changing how KVM for IBM Z is delivered. KVM hypervisor will now be offered through our Linux distribution partners.

Multiple software hypervisors can exist on the same Z platform (see Figure 4-1).

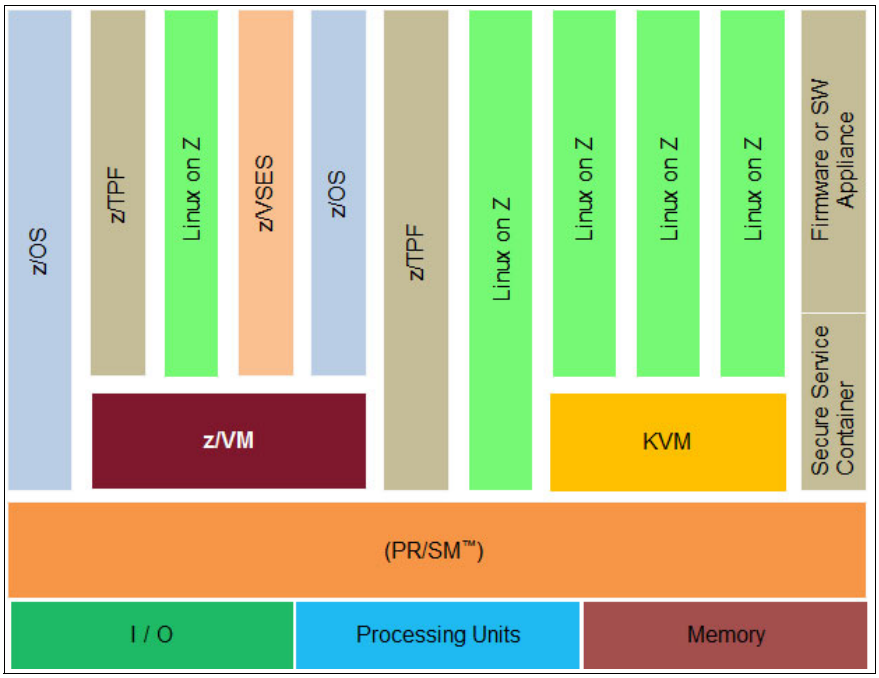

Figure 4-1 Support for coexistence of different hypervisors (in PR/SM mode)

The various virtualization options in Z platforms allow you to build flexible virtualized environments to take advantage of open source software or upgrade to new cloud service offerings, such as infrastructure as a service (IaaS) and platform as a service (PaaS)

PR/SM

Unique to Z platforms, PR/SM is a Type-1 hypervisor that runs directly on bare metal, which allows you to create multiple LPARs on the same physical server. PR/SM is a highly stable, proven, and secure, firmware-encapsulated virtualization technology that allows multiple operating systems to run on the same physical platform. Each operating system runs in its own logical partition.

PR/SM logically partitions the platform across the various LPARs to share resources, such as processor units, memory, and I/O (for networks and storage), which allows for a high degree of virtualization.

Dynamic Partition Manager

Dynamic Partition Manager (DPM) is a management infrastructure mode in the z14 ZR1. It is intended to simplify virtualization management and is easy to use, especially for users who are less experienced with Z. It does not require you to learn complex syntax or command structures.

Software hypervisors, operating systems, secure service containers, and software appliances that can exist on the Z platform in DPM mode are shown in Figure 4-2 on page 64.

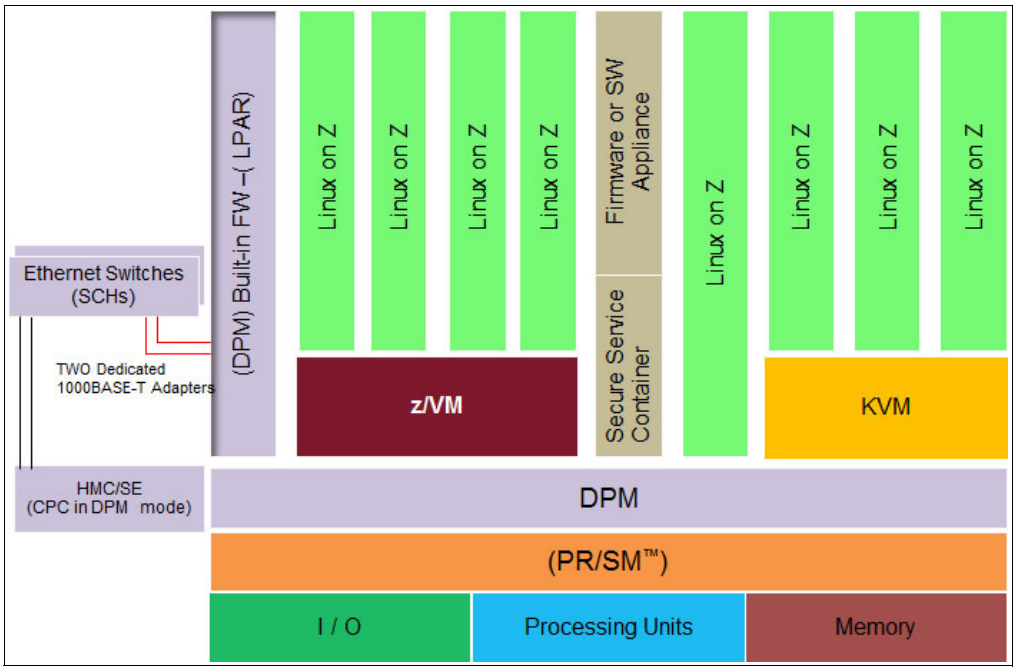

Figure 4-2 Support for coexistence of different software hypervisors (in DPM mode)

DPM provides simplified hardware and virtual infrastructure management, including partition lifecycle and integrated dynamic I/O and PCIe functions management for Linux that is running in an LPAR, under the KVM hypervisor, and under z/VM. By using DPM, an environment can be created, provisioned, and modified without disrupting running workloads. It also can be monitored for troubleshooting.

DPM provides the following capabilities through the Hardware Management Console (HMC):

- Create and provision an environment, including new partitions, assignment of processors and memory, and configuration of I/O adapters.
- Manage the environment, including the ability to modify system resources without disrupting workloads.
- Monitor and troubleshoot the environment to identify system events that might lead to degradation.

Enhancements to DPM on z14 ZR1 simplify the installation of the Linux operating system, support more hardware features, and enable base cloud provisioning through Openstack, including the following enhancements:

- Support for auto-configuration of devices to simplify Linux Operating System Installation, where Linux distribution installers use functions
- Secure FTP through HMC for starting and installing an Operating system by using FTP
- Support for OSA-Express7S, OSA-Express6S, FICON Express 16S+, Crypto Express6S, 25GbE RoCE Express, and 10GbE RoCE Express2 features

Configuration note: The z14 ZR1 can be configured in DPM mode or in PR/SM mode, but cannot be configured in both modes at the same time. DPM supports FCP and ECKD storage.

z/VM

z/VM is a Type-2 hypervisor that allows sharing the mainframe's physical resources, such as disk, memory, network adapters, and CPUs (called *CPs* and *IFLs*). These resources are managed by the z/VM hypervisor, which typically runs on an LPAR and other virtual machines (VMs) that run on top of the hypervisor. Typically, the z/VM hypervisor is used to run Linux virtual servers, but other operating systems (such as z/OS) can also run on z/VM. z/VM is a proven and well-established virtualization platform. It provides industry-leading capabilities to efficiently scale both horizontally and vertically.

KVM hypervisor

The KVM hypervisor available in recent Linux on Z distributions is a Type-2 hypervisor that provides simple, cost-effective server virtualization for Linux workloads that are running on the Z platform. It enables you to share real CPUs (called IFLs), memory, and I/O resources through platform virtualization and can coexist with z/VM virtualization environments, Linux on IBM Z, z/OS, z/VSE, and z/TPF.

The KVM hypervisor support information is provided by the Linux distribution partners. For more information, see the documentation for your distribution.

For more information about the use of KVM on Z, see the Linux on KVM page of IBM Knowledge Center.

4.2.1 Hardware virtualization

PR/SM was first implemented in the mainframe in the late 1980s. It allows you to define and manage LPARs. PR/SM virtualizes processor units, memory, and I/O features. Certain features are purely virtualized implementations.

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

The LPAR definition includes several logical processor units (LPUs), memory, and I/O devices. IBM z/Architecture is designed to meet requirements with low overhead with a Specific Target of Evaluation (Logical Partitions). This design was proven in many installations over several decades.

Up to 40 LPARs can be defined on the IBM z14 Model ZR1 and hundreds or even thousands of virtual servers can be run under z/VM or the KVM hypervisor.

Logical processors

Logical processors are defined and managed by PR/SM and are perceived by the operating systems as real processors. These processors are sorted into the following characterizations:

- Central processors (CP) are standard processors for use with any supported operating system and user applications.
- IBM System z Integrated Information Processor (zIIP) is used under z/OS for designated workloads, including the following workloads:
  - IBM Java virtual machine (JVM)
  - Various XML System Services
  - IPSec offload
  - Certain parts of IBM DB2 DRDA
  - DFSMS System Data Mover for z/OS Global Mirror
  - IBM HiperSockets for large messages

  - IBM GBS Scalable Architecture for Financial Reporting (SAFR) enterprise business intelligence reporting
- IFL is exclusively used with Linux on IBM Z, and for running the z/VM and KVM hypervisors in support of Linux VMs (also called *guests*).
- Internal Coupling Facility (ICF) is used for z/OS clustering. ICF is dedicated to this function and exclusively run the Coupling Facility Control Code (CFCC).

In addition, the following pre-characterized processors are part of the base system configuration and are always present:

- SAP that runs I/O operations
- IFP for native PCIe features

Although these processors provide support for all LPARs, they are never part of an LPAR configuration.

PR/SM accepts 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, the logical processors of an LPAR must be all shared or all dedicated.

The sum of logical processors that are defined in all active LPARs in a Z system might be higher than the number of physical processor units. The maximum number of LPUs that can be defined in a single LPAR *cannot exceed the total number physical processor units that are available in the CPC*. To achieve optimal ITR performance in sharing LPUs, the total number of online LPUs should be kept to a minimum, which reduces software and hardware overhead.

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

Logical processors can be dynamically added to and removed from LPARs. Operating system support is required to use this capability. z/OS, z/VM, and z/VSE each can dynamically define and change the number and type of reserved processor units in an LPAR profile. No pre-planning is required.

The new resources are immediately available to the operating systems and, for z/VM, to its guest images. Linux on IBM Z provides the Standby CPU activation and deactivation functions.

Memory

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

A logical partition can be defined with an initial and 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 the addressing area (for initial and reserved memory) is presented to the operating system that runs in the LPAR as contiguous.

By using the plan-ahead option, memory can be physically installed without being enabled. It can then be enabled when necessary. z/OS can use this support by nondisruptively acquiring and releasing memory from the reserved area.

z/VM can acquire memory nondisruptively and quickly make it available to guests. z/VM virtualizes this support to its guests, which can also increase their memory nondisruptively. Releasing memory is still a disruptive operation.

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

The z/Architecture features a robust virtual storage architecture that allows LPAR-by-LPAR 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 = 2^{60} 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. A program's right to read or write in each page frame is validated by comparing the page key with the key of the program that is requesting access. This mechanism was 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, which provides compatibility with earlier versions and investment protection.

z14 ZR1 supports 4 KB, 1 MB, and 2 GB pages, and an extension to the z/Architecture that is called Enhanced Dynamic Address Translation-2 (EDAT-2).

Operating systems can allow sharing of address spaces, or parts of them, across multiple processes. For example, under z/VM, a single copy of the read-only part of a kernel can be shared by all VMs that use that operating system. Known as *discontiguous shared segment* (DCSS), this shared memory exploitation for many VMs can result in large savings of real memory and improvements in performance.

I/O virtualization

The z14 ZR1 supports three logical channel subsystems (LCSSs), each with 256 channels, for a total of 768 channels. In addition to the dedicated use of channels and I/O devices by an LPAR, I/O virtualization allows concurrent sharing of channels. The z/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 includes the channel configured online, PR/SM establishes one logical channel path. For availability reasons, multiple logical channel paths should be available 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, which lowers the need for manually tuning the system for PAV use.

In large installations, the total number of device addresses can be high. Therefore, the concept of *channel sets* is part of the z/Architecture.

Subchannel sets

On the z14 ZR1, up to three sets of approximately 64,000 device addresses are available. This availability allows the base addresses[3] to be defined on set 0 (IBM reserves 256 subchannels on set 0) and the aliases on set 1, and set 2. In total 196,349 subchannel addresses are available per channel subsystem.

Subchannel sets are used by the *Metro Mirror* (also referred to as *synchronous Peer-to-Peer Remote Copy* [PPRC]) function by having the Metro Mirror primary devices that are defined in subchannel set 0. Secondary devices can be defined in subchannel sets 1, and 2, which provides more connectivity through subchannel set 0.

To reduce the complexity of managing large I/O configurations further, Z introduced extended address volumes (EAV). EAV provides large disk volumes. In addition to z/OS, z/VM and Linux on IBM Z support EAV.

By extending the disk volume size, potentially fewer volumes are required to hold the same amount of data, which simplifies systems and data management. EAV is supported by the IBM DS8000® series. For more information about EAV compatibility, see the devices from other vendors.

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. A fixed HSA area in the z14 ZR1 greatly eases the planning requirements and enhances the flexibility and availability of these reconfigurations.

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

4.2.2 IBM Z based clouds

Cloud computing capitalizes on the ability to rapidly and securely deliver standardized service offerings, while retaining the capacity for customizing the environment. Elasticity and just-in-time provisioning allow the system to deal with the ebbs and flows of demand dynamically.

Virtualization is critical to the economic and financial viability of cloud service offerings because it allows minimizing the over-provisioning of resources and reusing them at the end of the virtual server lifecycle.

Because of the extreme integration in the hardware, virtualization on z14 ZR1 is highly efficient (the best in the industry) and encompasses computing and I/O resources, including the definition of *internal virtual networks* with *virtual switches*. These characteristics are common to *software-defined environments*.

[3] Only a z/OS base device must be in subchannel set 0. Linux on IBM Z supports base devices in the other subchannels sets.

These characteristics also allow support on a single platform, 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 z14 ZR1 granularity offerings, including capacity levels and capacity on demand. These and other technologic leadership characteristics make the Z platforms the server golden standard.

In addition, managing a cloud environment requires tools that can take advantage of a pool of virtualized compute, storage, and network resources, and present them to the consumer as a service in a secure way. A cloud management system should also help with the following tasks:

- Offering open cloud management and application programming interfaces (APIs)
- Improving the usage of the infrastructure
- Lowering administrative overhead and improving operations productivity
- Reducing management costs and improving responsiveness to changing business needs
- Automating resource allocation
- Providing a self-service interface
- Tracking and metering resource usage

A cloud management system must also support the management of virtualized IT resources to support different types of cloud service models and cloud deployment models. OpenStack (offered for z/VM and the KVM hypervisor) can satisfy a wide range of cloud management demands. It integrates various components to automate IT infrastructure service provisioning.

The Z cloud architecture that provides an industrial-strength base for hybrid cloud and API economy is shown in Figure 4-3.

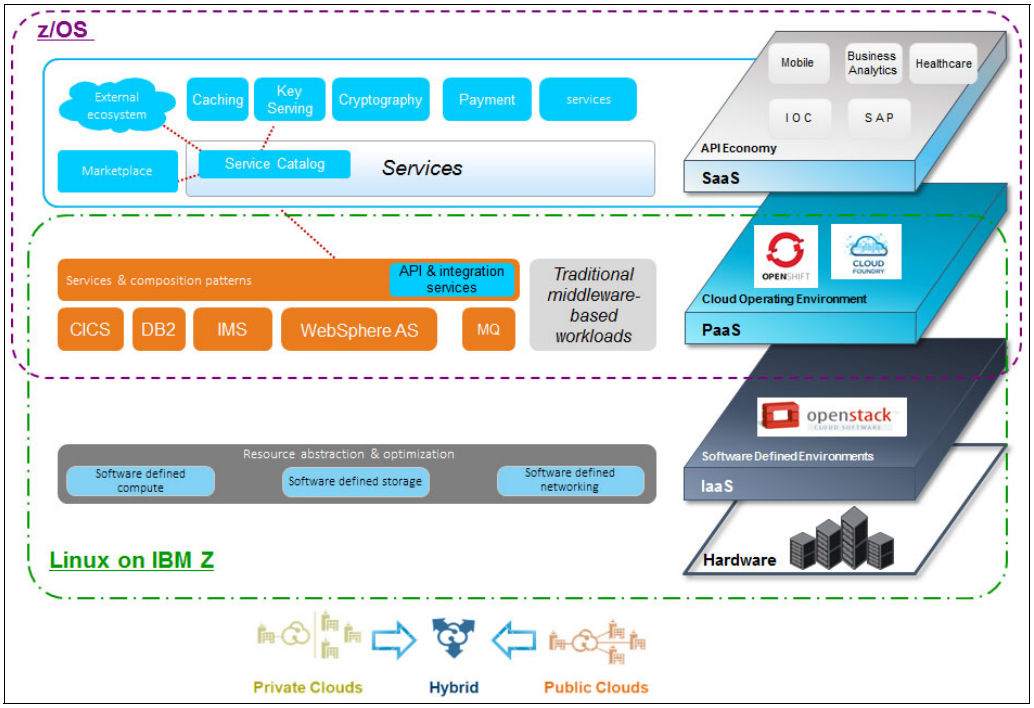

Figure 4-3 IBM Z cloud architecture

4.2.3 Secure Service Container

The IBM Secure Service Container provides the base infrastructure to create and deploy an IBM Z Appliance, which includes operating system, middleware, SDK, and firmware support. With a Secure Service Container, deploying an appliance that provides a function or a service takes minutes instead of days, while providing simplified management and maintenance. When deployed in a Secure Service Container LPAR (SSC LPAR), the workload is protected from inadvertent access from an external attacker or even from a system administrator.

IBM Z Appliance is an integration of operating system, middleware, and various software components that work autonomously to provide core infrastructure services while focusing on consumability and security. The appliance is deployed in an SSC LPAR that is running in an IBM Z platform. The SSC LPAR in the Z platform provides support for the following components:

- Encapsulated Operating Systems
- Remote APIs (RESTful) and web interfaces
- Embedded monitoring and self-healing
- Tamper-protection
- Protected IP

The platform also is tested and qualified by IBM for a specific use case and can be delivered as firmware, platform, or software.

At the time of this writing, the following IBM Z Appliances were available to be deployed in a Secure Service Container:

- z/VSE Network Appliance (VNA)
- IBM z Systems Advanced Workload Analysis Reporter (IBM zAware), now deployed as software appliance that is running in a secure service container and integrated with IBM Operations Analytics for Z

4.3 Capacity and performance

The z14 ZR1 offers significant increases in capacity and performance over its predecessor, the z13s. Several elements contribute to this effect, including the larger number of processors, individual processor performance, memory caches, and SMT and machine instructions, including the SIMD. Subcapacity settings continue to be offered.

Note: Capacity and performance ratios are based on measurements and projections that use standard IBM benchmarks in a controlled environment. Actual throughput can vary, depending on several factors, such as the job stream, I/O and storage configurations, and workload type.

4.3.1 z14 ZR1 capacity settings

The z14 ZR1 offers processor subcapacity settings. The fine 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. Many performance and monitoring tools are available on Z environments that are coupled with the flexibility of the capacity on-demand options (see 4.3.2, “Capacity on demand” on page 71). These features help to manage growth by making capacity available when needed.

Capacity levels

The z14 ZR1 offers 26 distinct capacity levels for up to six CPs in the configuration, for a total of 156 capacity settings (26 x 6). These processors deliver the scalability and granularity to meet the needs of small and medium-sized enterprises.

A Processor Unit that is characterized as anything other than a CP, such as a zIIP, IFL, or an ICF, is always set to full capacity.

The z14 ZR1 capacity settings are shown in Figure 4-4.

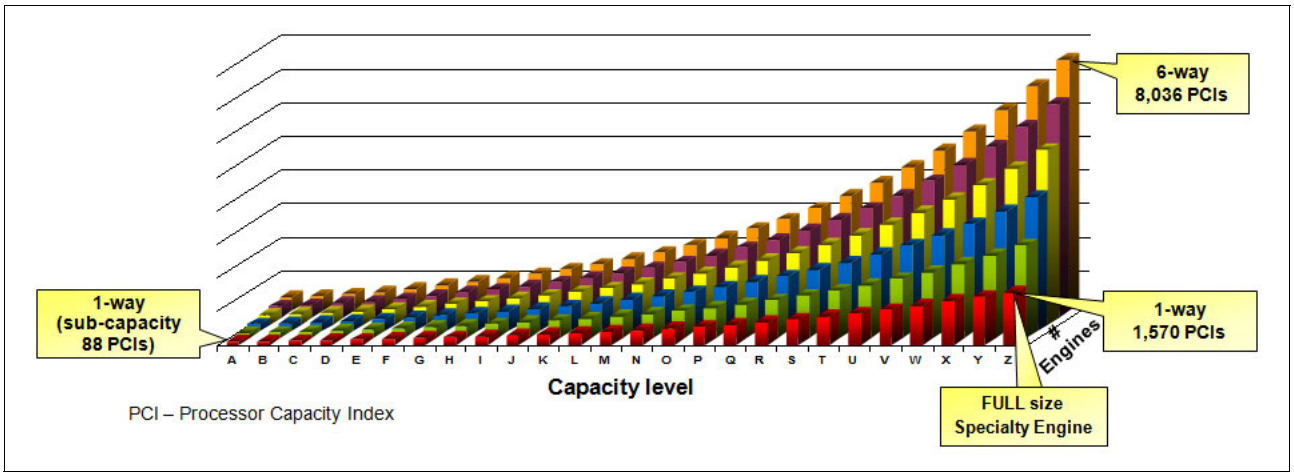

Figure 4-4 z14 ZR1 capacity settings offerings

A capacity level is a setting of each CP[4] to a subcapacity of the full CP capacity. The clock frequency of those processors remains unchanged. The capacity adjustment is achieved through other means.

To help you size a Z platform, IBM provides a no-cost tool that reflects the latest IBM LSPR measurements, called the IBM Processor Capacity Reference for Z (zPCR). You can download the tool here.

For more information about LSPR measurements, see 4.3.3, “z14 ZR1 performance” on page 73.

4.3.2 Capacity on demand

The z14 ZR1 continues to provide capacity on-demand (CoD) offerings. They 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 z14 ZR1 can perform concurrent upgrades, which provide an increase of processor capacity with no server outage. In most cases, a concurrent upgrade can also be nondisruptive to the operating system with operating system support. It is important to consider that these upgrades are based on the enablement of resources that are physically present in the z14 ZR1.

[4] The CP is the standard processor for use with any supported operating system, but is required to run z/OS.

Capacity upgrades cover permanent and temporary changes to the installed capacity. The changes can be done by using the Customer Initiated Upgrade (CIU) facility, without requiring IBM service personnel involvement. Such upgrades are started 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 CoD buying of upgrades and other types of CoD upgrades are accepted. For more information, see the IBM Resource Link.

For more information about the CoD offerings, see *IBM z14 Model ZR1 Technical Guide*, SG24-8651.

Permanent upgrades

Permanent upgrades of processors (CP, IFL, ICF, zIIP, and SAP) and memory, or changes to a platform's Model-Capacity Identifier up to the limits of the installed processor capacity on an existing z14 ZR1, can be performed by the client through the IBM Online Permanent Upgrade offering by using the CIU facility.

Temporary upgrades

Temporary upgrades of a z14 ZR1 can be done by On/Off CoD, Capacity Backup (CBU), or Capacity for Planned Event (CPE) that is ordered from the CIU facility.

On/Off CoD function

On/Off CoD is a function that is available on the z14 ZR1 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. It features a daily hardware charge and can include an associated software charge.

On/Off CoD offerings can be prepaid or post-paid. Capacity tokens are available on z14 ZR1. Capacity tokens are always present in prepaid offerings and can be present in post-paid if so wanted by the client. In both cases, capacity tokens are used to control the maximum resource and financial consumption.

When the On/Off CoD function is used, the client can concurrently add processors (CP, IFL, ICF, zIIP, and SAP), increase the CP capacity level, or both.

Capacity Backup function

CBU allows the client to perform a concurrent and temporary activation of more CP, ICF, IFL, zIIP, and SAP, an increase of the CP capacity level, or both. This function can be used during an unforeseen loss of Z 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 as unused processor units, 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 more tests for a CBU record in quantities of five tests up to a maximum of 15 tests.

Proper use of the CBU capability does not incur any extra software charges from IBM.

Capacity for Planned Event function

CPE allows the client to perform a concurrent and temporary activation of more CPs, ICFs, IFLs, zIIPs, and SAPs, an increase of the CP capacity level, or both. This function can be used during a planned outage of Z 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 backup system, as unused processor units, 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 z14 ZR1 that is running one or more instances of the z/OS operating system. By using the 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 z14 ZR1 provisioning capability gives the client a flexible, automated process to control the configuration and activation of On/Off CoD offerings.

4.3.3 z14 ZR1 performance

The Z microprocessor chip of the z14 ZR1 features 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 z14 ZR1 is a result of the enhancements that are described in Chapter 2, "IBM z14 ZR1 hardware overview" on page 17, and in 4.1, "Technology improvements" on page 58.

A fully configured z14 ZR1 (Max30, FC 0639) offers up to 54% more capacity than the largest z13s Model N20. Uniprocessor performance also increased significantly. Single processor capacity of a z14 ZR1 is approximately 10% greater than a z13s with equal *n*-way configurations. Performance varies depending on workload type and configuration.

LSPR workload suite: z14 ZR1 changes

To help you better understand workload variations, IBM provides a no-cost tool, zPCR, which is available at the IBM Presentation and Tools website.

IBM continues to measure performance of the systems by using various workloads and publishes the results in the Large Systems Performance Reference (LSPR) report.

IBM also provides a list of MSU ratings for reference.

Capacity performance is closely associated with how a workload uses and interacts with a particular processor hardware design. Workload capacity performance is sensitive to the following major factors:

- Instruction path length
- Instruction complexity
- Memory hierarchy

The CPU measurement facility (MF) data allows you to gain insight into the interaction of workload with the hardware design. 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. With the Z, the LSPR introduced the following workload capacity categories that replace all older primitives and mixes:

- LOW (relative nest intensity): Represents light use of the memory hierarchy.
- AVERAGE (relative nest intensity): Represents average use of the memory hierarchy. This category is expected to represent most production workloads.
- HIGH (relative nest intensity): Represents heavy use of the memory hierarchy.

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 or z/VM Monitor starting with z/VM V5R4.

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 that are 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 that are used to set the single-number values, the LSPR tables contain information that pertains to Linux and z/VM environments.

The LSPR includes the internal throughput rate ratios (ITRRs) for the z14 ZR1 and the previous generations of processors that are based on measurements and projections that use standard IBM benchmarks in a controlled environment. The actual throughput that any user might experience varies depending on several factors, such as the amount of multiprogramming in the user's job stream, I/O configuration, and processed workload.

Experience shows that Z platforms can be run at up to 100% utilization levels, sustained. However, most clients prefer to leave a bit of white space and run at 90% or slightly under. For any capacity comparison, the use of "one number," such as the MIPS or MSU metrics, is not a valid method. Therefore, use zPCR and include IBM technical support when you are planning for capacity. For more information about z14 ZR1 performance, see *IBM z14 Model ZR1 Technical Guide*, SG24-8651.

Throughput optimization with z14 ZR1

The memory and cache structure that is implementation in the z14 ZR1 was significantly enhanced compared to previous generations to provide sustained throughput and performance improvements. Processors within the z14 ZR1 CPC drawer feature different distance-to-memory attributes. To minimize latency, the system attempts to dispatch and later redispatch work to a group of physical CPUs that share cache levels.

PR/SM manages the use of physical processors by LPARs by dispatching the logical processors on the physical processors. However, 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 enable 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, which is known as *HiperDispatch*.

HiperDispatch

HiperDispatch combines two functional enhancements, one of which is in the z/OS dispatcher and the other in PR/SM. This function is intended to improve computing efficiency in the hardware, z/OS, and z/VM.

The PR/SM dispatcher assigns work to the minimum number of logical processors that are needed for the priority (weight) of the LPAR. On the z14 ZR1, PR/SM attempts to group the logical processors into the same logical cluster or in the neighbor logical cluster in the same CPC drawer and, if possible, in the same chip. This configuration 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, which improves 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 z14 ZR1. The entire z14 ZR1 stack (hardware, firmware, and software) tightly collaborates to obtain the full potential of the hardware. z/VM HiperDispatch provides support similar to the z/OS HiperDispatch in z/OS. It is possible to dynamically turn on and off HiperDispatch without requiring an initial program load (IPL).

Note: HiperDispatch is required if SMT is enabled.

4.4 Reliability, availability, and serviceability

The IBM Z family presents numerous enhancements in reliability, availability, and serviceability (RAS). 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 Licensed Internal Code (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 z14 ZR1 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 might not be possible.

On a z14 ZR1 concurrent repair or upgrade on the CPC drawers is *not* supported.

z14 ZR1 builds on the RAS characteristics of the z13s, with the following RAS improvements:

- z14 ZR1 Level 3 cache enhancements were made by using powerful symbol ECC to extend the reach of prior z13s cache and memory improvements for improved availability. The level 3 cache powerful symbol ECC is designed to make it resistant to more failure mechanisms. The z13s hardened the level 4 cache and the main memory was hardened with RAIM before that.
- Preemptive DRAM marking was added to the main memory to isolate and recover failures more quickly.
- Small array error handling was improved in the processor cores.
- Error thresholding was added to the processor core to isolate “sick but not dead” failure scenarios.
- The number of Resource Groups was increased to four to reduce the affect of firmware updates and failures.
- An OSA-Express6S TCP checksum was added on large sends.

z14 ZR1 also continues to feature a redundant array of independent memory[5] (RAIM) that provides a method to increase memory availability, where a fully redundant memory system can identify and correct memory errors without stopping. The implementation is similar to the RAID concept that is used in storage systems for several years. For more information about RAS features, see *IBM z14 ZR1 Technical Guide*, SG24-8651.

The z14 ZR1 consists of a single CPC drawer that is designed as a field replaceable unit (FRU). The CPC drawer uses a modular, field upgradable construction that consists of two CP clusters and contains one, two, or four processing unit (PU) single-chip modules (SCMs) and one storage controller (SC) SCM. In addition to SCMs, CPC drawer hosts memory DIMMs, connectors for I/O, oscillator interface, and Flexible Support Processors (FSPs).

The CPC drawer is equipped with two or four redundant Power Supply Units (AC to DC), depending on the CPC drawer configuration (redundant PSUs). The PSUs connect to point of Load.

A redundant pair of Distributed Converter Assemblies (DCAs) step down the bulk power and connect to six point of load (POL) cards, which provide power conversion and regulation. Two redundant oscillators are connected to the drawer.

Time domain reflectometry (TDR) techniques are applied to isolate failures between chips (PU-PU and PU-SC), and between the processor unit chips and DIMMs. More redundancy is designed into N+1 Ethernet switches, which replace the System Control Hubs (SCHs), and associated Power Distribution Units (PDUs) and 1U Support Elements (SEs).

z14 ZR1 inherits I/O infrastructure reliability improvements from z13s, including Forward Error Correction (FEC) technology that enables better recovery of FICON channels facilitated. The system is air cooled with redundant N+1 fans for the CPC drawer and the PCIe+ I/O drawer.

The following RAS enhancements also are included:

- Improved integrated sparing
- Error detection and recovery improvements in caches and memory
- Fibre Channel Protocol support for T10-DIF
- A fixed HSA with its size increased to 64 GB on the z14 ZR1

[5] 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

- OSA-Express firmware changes to increase the capability of concurrent maintenance change level (MCL) updates
- Air cooled system with redundant fans (N+1) for all major components
- New CFCC level
- Enhanced IBM RMF™ reporting

z14 ZR1 continues to support the concurrent addition of resources, such as processors or I/O cards, to an LPAR to achieve better serviceability. If another SAP is required on a z14 ZR1 (for example, as a result of a disaster recovery situation), the SAPs can be concurrently added to the CPC configuration.

CP, zIIP, IFL, and ICF processors can be added concurrently to an LPAR. This function is supported by z/VM, and by z/OS and z/VSE with appropriate PTFs. Previously, proper planning was required to concurrently add CP, zAAP, and zIIP to a z/OS LPAR. Concurrently adding memory to an LPAR also is possible. This ability is supported by z/OS and z/VM.

z14 ZR1 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 IBM Z.

4.4.1 RAS capability of the PDUs and Ethernet switches

The IBM z14 ZR1 uses a modular construction with single phase power that is provided to the rack components by way of up to four redundant (N+1) intelligent PDUs. The System Control Hubs were replaced with two redundant GbE switches that connect the internal management infrastructure (FSPs, SEs, and PDUs).

4.4.2 RAS capability for the Support Element

Enhancements are made to the SE design for z14 ZR1. Notebooks that were used on past generations of Z servers were replaced with rack-mounted 1U servers in a redundant configuration on z14 ZR1. The more powerful SEs offer RAS improvements, such as ECC memory, redundant physical networks for SE networking requirements, redundant power modules, and better thermal characteristics. Also, the SEs provide Firmware Integrity Monitoring.

4.4.3 RAS capability for the Hardware Management Console

Enhancements are also made to the HMC designs for z14 ZR1. New for z14 ZR1 is an option to order 1U servers for traditional and ensemble HMC configurations. This 1U HMC offers the same RAS improvements as the improvements in 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 before z14 ZR1 is still available.

4.5 High availability with parallel sysplex

The parallel sysplex technology is a clustering technology for logical and physical servers, which allows highly reliable, redundant, and robust Z technology to achieve near-continuous availability. Hardware and software tightly cooperate to achieve this result.

A parallel sysplex features the following minimum components:

- Coupling facility (CF)

  The CF is the cluster center. It can be implemented as an LPAR of a stand-alone Z platform, or as another LPAR of a Z platform in which other LPARs are running. Processor units that are characterized as CPs or ICFs can be configured to this LPAR. ICFs are often used because they do not require any software license charges. Two or more CFs are recommended for availability.

- Coupling Facility Control Code

  This IBM LIC is the operating system and application that runs in the CF. No other code runs in the CF. The code is used to create and maintain the structures. These structures are used under z/OS by software components, such as z/OS, DB2 for z/OS, CICS TS, and WebSphere MQ.

  CFCC can also run in a z/VM virtual machine (as a z/VM guest system). A complete sysplex can be set up under z/VM, which allows testing and operations training, for example. This setup is not recommended for production environments.

- Coupling links

  These high-speed links 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-use 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.

For more information about deploying system-managed CF structure duplexing, see the technical paper *System-Managed CF Structure Duplexing,* ZSW01975USEN, which is available by clicking **Learn more** at the Parallel Sysplex website.

Note: z/TPF can also use the CF hardware components. However, the term *sysplex* exclusively applies to z/OS use of the CF.

Normally, two or more z/OS images are clustered to create a Parallel Sysplex. Multiple clusters can span several Z platforms, 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 configuration is facilitated by Z I/O virtualization capabilities, such as MIF. MIF allows several logical partitions to share I/O paths in a secure way, which maximizes use and greatly simplifies the configuration and connectivity.

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.*

Parallel Sysplex includes the following major characteristics:

- Data sharing with integrity

  The CF is key to the implementation of share-all access to data. Every z/OS system image can access all of the data. Subsystems in z/OS declare resources to the CF. The CF accepts and manages lock and unlock requests on those resources, which helps ensure 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 components are major infrastructure components, applications that use them inherently benefit from sysplex characteristics. For example, many large SAP implementations have the database component on DB2 for z/OS in a Parallel Sysplex.

- Near-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 information, see *Improving z/OS Application Availability by Managing Planned Outages*, SG24-8178.

- High capacity

  Parallel sysplex scales 2 - 32 images. Each image can have 1 - 170[6] processor units. The scalability is near-linear as z/OS images are added to a sysplex. This structure contrasts with other forms of clustering that use n-to-n messaging, which leads to rapidly degrading performance with a growing number of nodes.

- Dynamic workload balancing

  Because the system is 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 and enables techniques for achieving simplified systems management consistent with this requirement.

- Resource sharing

  Several 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 operational and definition perspectives.

- N-1 support

  Three hardware generations (the current and the two previous generations) often are supported in the same parallel sysplex. However, the z14 ZR1 supports N/N-1 coupling and STP connectivity only. The z14 ZR1 supports the following coupling features:

  – Integrated Coupling Adapter Short Reach (ICA SR; 8x link, up to 150 m)
  – Coupling Express Long Reach (CE LR; 1x link, up to 10 km unrepeated)

  Software support for multiple releases or versions is supported.

[6] The IBM z14 ZR1 can have a maximum of 6 CPs or up to 30 ICFs. The maximum number of PUs per CF LPAR is 16

- CF encryption support

  Provides support for encrypted data while it is being transferred to and from the CF because it is in the Coupling Facility Structure:

  – z/OS Systems must have the cryptographic hardware configured and activated to perform cryptographic functions and hold AES master keys within a secure boundary. Feature 3863, CPACF DES/TDES Enablement must be installed to use the Crypto Express4 Coprocessor (CEX4C), the Crypto Express5 Coprocessor (CEX5C), or the Crypto Express6 Coprocessor (CEX6C) feature.

  – Support provided can be enabled only when all systems are z/OS 2.3 or higher. Toleration support with reduced functionality will be provided for z/OS 2.2 and z/OS 2.1.

- Dynamic activation of I/O configurations for stand-alone CFs

  Dynamic I/O configuration changes can be made to a stand-alone CFs without requiring a disruptive reset. An LPAR with a firmware-based appliance version of a HCD instance is used to apply the new I/O configuration changes. The LPAR on a z14 is driven by an updated HCD instance that is running in a z/OS LPAR on a remote z14.

  Note: This function requires new z14 firmware support for the stand-alone CF. To enable this function, an IML or POR action must be performed on the stand-alone CF after the firmware is installed. Firmware support is also required on the connected Z platform where the HCD instance resides.

- Improved performance and resiliency

  Asynchronous Cache Cross-Invalidation is a new sysplex capability for performance, scalability, and improved cross-site operation. This function allows the cache coherency messages that flow across the sysplex to maintain data integrity to be performed in an asynchronous fashion rather than synchronously, with exploitation from the data manager (Db2). Asynchronous Cache Cross-Invalidation is expected to reduce CF cache structure service times and sysplex coupling overhead, particularly in sysplex environments that involve multiple sites with significant cross-site distances. This function requires PTFs for z/OS and Db2.

The components of a Parallel Sysplex as implemented within the Z architecture are shown in Figure 4-5. The configuration is one of many possible parallel sysplex configurations.

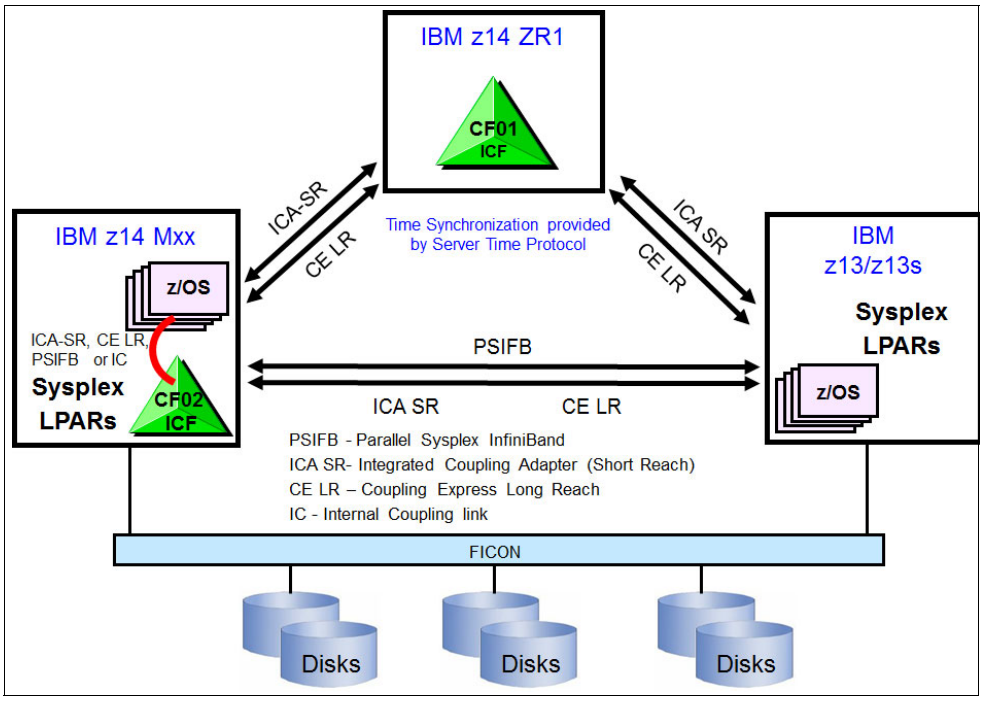

Figure 4-5 Sysplex hardware overview

A z14 ZR1 that contains multiple z/OS sysplex partitions and an internal coupling facility (CF02), a z14 ZR1 server containing a stand-alone CF (CF01), and a z13 containing multiple z/OS sysplex partitions also is shown in Figure 4-5. Server Time Protocol (STP) over coupling links provide time synchronization to all servers.

Note: The z14 ZR1 does not include InfiniBand coupling. The zEC12 or zBC12 does not include ICA SR or CE LR coupling links; therefore, they cannot connect directly.

The z14 ZR1 allows only direct connectivity back to z13 or z13s that migrated to ICA SR or CE LR coupling links.

Appropriate CF link technology (1x IFB, 12x IFB, ICA SR, or CE LR) selection depends on platform configuration and how distant they are physically located. ICA SR links can be used within a short distance only; a CE LR can support a distance up to 10 km (6.2 miles). For more information about coupling link options, see 3.7, “Coupling and clustering” on page 52.

4.6 Pervasive encryption

Data protection and security are business imperatives, and regulatory compliance is increasingly complex. Extensive use of encryption is one of the best ways to reduce the risks and financial losses of a data breach and meet complex compliance mandates. However, implementing encryption can be a complex process for organizations. The following factors must be determined:

- What data should be encrypted?
- Where should encryption occur?
- Who is responsible for encryption?

Because the data is the new perimeter, encryption policies must cover data in-flight and data at-rest, but should not require costly application changes to achieve this goal. Organizations need a transparent and consumable approach to enable extensive encryption of data in-flight and at-rest to substantially simplify and reduce the costs that are associated with protecting the data at the core of their enterprise and achieving compliance mandates.

With solutions around privileged identity management, sensitive data protection, and integrated security intelligence, Z security offers the next generation of secure, trusted transactions.

Pervasive encryption is a data-centric approach to information security that entails protecting data entering and exiting the z14 ZR1 platform. It involves encrypting data in-flight and at-rest to meet complex compliance mandates and reducing the risks and financial losses of a data breach. It is a paradigm shift from selective encryption (where only the data that is required to achieve compliance is encrypted) to pervasive encryption. Pervasive encryption with z14 ZR1 is enabled through tight platform integration that includes the following features:

- Integrated cryptographic hardware: CPACF is a co-processor on every processor unit that accelerates encryption. Crypto Express features can be used as hardware security modules (HSMs)[7].
- Data set and file encryption: You can protect Linux file systems and z/OS data sets by using policy-controlled encryption that is not apparent to applications and databases.
- Network encryption: You can protect network data traffic by using standards-based encryption from endpoint to endpoint.
- Full disk encryption: You can use disk drive encryption that protects data at rest when disk drives are retired, sent for repair, or repurposed.
- CF encryption: This encryption secures the parallel sysplex infrastructure, including the CF links and data that stored in the CF, by using policy-based encryption.
- Secure Service Container: This container secures the deployment of software appliances, including tamper protection during installation and runtime, restricted administrator access, and encryption of data and code in-flight and at-rest.

[7] An HSM is a physical computing device that safeguards and manages digital keys for strong authentication and provides crypto processing.

Data is encrypted when in-flight and at-rest, as shown in Figure 4-6. Data is decrypted only when it is processed by the operating system.

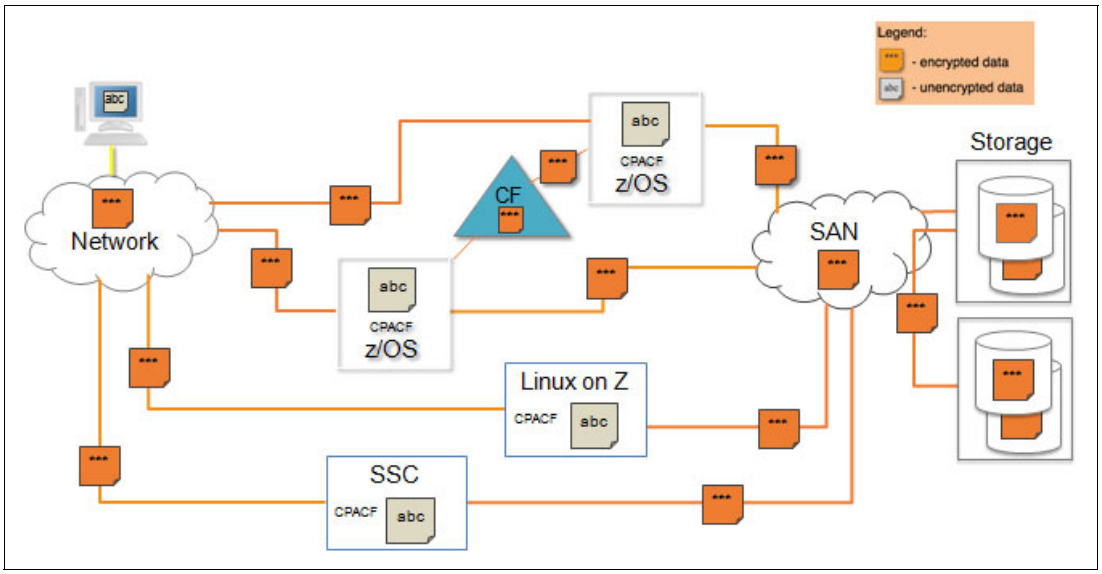

Figure 4-6 Protecting data in-flight and at-rest

Pervasive encryption includes the following advantages:

- The ability to encrypt data by policy without application change
- A simplified way to protect data at a much coarser scale with industry best performance
- Greatly simplified audit, which enables clients to pass compliance audits more easily

5

Operating system support

This chapter describes the operating system requirements and support considerations for the z14 ZR1 and their features.

Support and use of hardware functions depend on the operating system version and release. The information in this chapter is subject to change. Therefore, for the most current information, see *z14 Model ZR1: Preventive Service Planning (PSP)* bucket for 3907DEVICE

This chapter includes the following topics:

5.1 Software support summary

The software portfolio for the z14 ZR1 includes various operating systems and middleware that support the most recent and significant technologies. The following major operating systems are supported:

- z/OS
- z/VM
- z/VSE
- z/TPF
- Linux on IBM Z and the KVM hypervisor

5.1.1 Operating systems summary

The current and minimum operating system levels that are required to support the z14 ZR1 are listed in Table 5-1 on page 87. Operating system levels that are no longer in service are not covered in this publication. These older levels can support certain features.

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 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.

PTFs for z/OS, z/VM, and z/VSE can be ordered electronically from IBM Shopz.

For more information about obtaining access to download the z/TPF and z/TPFDF APAR packages, contact TPFQA@us.ibm.com.

For Linux on IBM Z distributions and the KVM hypervisor, consult the distributor's support information.

Fix packs for IBM software products that are running on Linux on IBM Z can be downloaded from IBM Fix Central.

Table 5-1 z14 ZR1 operating system requirements

Operating system[a]	End of service	Notes
z/OS V2R3	September 2022[b]	See the z/OS, z/VM, z/VSE, and z/TPF subsets of the 3907DEVICE Preventive Service Planning (PSP) buckets before installing the z14 ZR1.
z/OS V2R2	September 2020[b]	
z/OS V2R1	September 2018[b]	
z/OS V1R13[c]	September 2016	
z/VM V7R1	Not announced	
z/VM V6R4	Not announced	
z/VSE V6R2	Not announced	
z/VSE V6R1[d]	June 2019	
z/VSE V5R2[e]	October 2018[b]	
z/TPF V1R1	Not announced	
Linux on IBM Z[f]	Support information is available for SUSE[g], Red Hat[h] and Canonical[i]	
KVM hypervisor	For more information about minimal and recommended distribution levels, see the Linux distributors' websites	

a. Only z/Architecture mode is supported.
b. 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.
c. Compatibility only. The IBM Software Support Services for z/OS V1.13, offered as of October 1, 2016, provides the ability for customers to purchase extended defect support service for z/OS V1.13.
d. z/VSE V6 requires an architectural level set
e. z/VSE V5 requires an architectural level set
f. For more information, see the Linux on IBM Z page of the IBM Z website.
g. For more information, see the Support page of the SUSE website.
h. For more information, see the Red Hat Enterprise Linux Life Cycle page of the RedHat website.
i. For more information, see the Ubuntu for IBM LinuxONE and IBM Z page of the Ubuntu website.

z/Architecture mode: As announced on January 14, 2015 with Announcement letter 115-001, beginning with z14 ZR1, all systems will support only operating systems that are running in z/Architecture mode. This support applies to operating systems that are running native on PR/SM and operating systems that are running as second-level guests.

IBM operating systems that run in ESA/390 mode are no longer in service or currently available with only extended service contracts, and they are not usable on systems beginning with z14 ZR1. However, z14 ZR1 does provide ESA/390-compatibility mode, an environment supporting a subset of DAT-off ESA/390 applications in a hybrid architectural mode.

All 24-bit and 31-bit problem state application programs that were originally written to run on the ESA/390 architecture will be unaffected by this change.

5.1.2 Application development and languages

Several programming languages are available for the z14 ZR1 environments. Because the Linux environment is similar to Linux on other servers, this description focuses on the z/OS environment.

In addition to the traditional COBOL, PL/I, Fortran, and assembler languages, Z platforms support C, C++, Java (including Java Platform, Enterprise Edition, and batch environments), Go, Swift, and JavaScript.

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.

IBM Z embraces emerging concepts, such as DevOps, which is the process of bringing Development and Operations together to share processes and procedures with a goal to reduce the risk of change and improve the speed of deployment. As organizations embark on their journey with digital transformation and enter the API economy, connecting business-critical applications that are running on mainframes with mobile and cloud applications to better engage with customers becomes more essential. A key step in this evolution is to understand what assets exist in the enterprise.

IBM's current DevOps offerings, such as IBM Application Delivery Foundation for z and IBM Rational® Team Concert™, coupled with IBM Application Discovery and Delivery Intelligence's (ADDI) application discovery technology, are designed to enable developers to understand the applications, gain cognitive insights into the process, and evolve those valuable older assets at the speed of business with reduced risk to the enterprise.

The use of modern development practices is part of the transformation. IBM's Rational Team Concert™ and open source-based Git Version Control Tools for IBM z/OS (ported by Rocket Software) are modern source code managers (SCM) that run on and support z/OS.

For more information about software for Z platforms, see the Products catalog website.

Note: The use of the most recent versions of the compilers is of utmost importance. 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.

z14 ZR1 introduces new features and functions, such as Guarded Storage Facility (GSF) enabling less-pausing garbage collection for Java workloads. z14 ZR1 processors also inherit and further enhance features and functions from its predecessor generation z13, such as the single-instruction, multiple-data (SIMD), which allows the development of smaller and optimized codes to improve efficiency of complex mathematical models and vector processing.

5.1.3 IBM compilers

Each new version of IBM z/OS compilers (Enterprise COBOL, Enterprise PL/I, and 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 most recent version of Enterprise COBOL uses the most recent z/Architecture and performance optimization, enhanced XML parsing support, and capability of programming with Unicode. It also supports Java 8 SDKs for Java interoperability and delivers new COBOL statements, new and changed compiler options, and changed APIs.

- IBM Automatic Binary Optimizer for z/OS

  The Automatic Binary Optimizer for z/OS improves the performance of compiled COBOL programs. The Optimizer does not require source code, source code migration, or performance options tuning. It uses modern optimization technology to target Z platforms and accelerate the performance of compiled COBOL applications.

- 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. Enterprise PL/I for z/OS allows you to capitalize on IT investments while modernizing your infrastructure.

- z/OS XL C/C++

  The z/OS XL C/C++ enables developing high performance-oriented applications, through the services that are provided by IBM Language Environment and Runtime Library extension base elements. It also works in concert with z/OS Problem Determination Tools.

IBM Enterprise COBOL and Enterprise PL/I support are strategic components (separately orderable products) for IBM Rational Developer for IBM Z software. These features provide a robust, 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 to help boost productivity by editing, compiling, and debugging z/OS XL C and XL C++ applications from the workstation.

5.2 Support by operating system

This section lists the support by in-service operating systems of selected functions of the z14 ZR1.

For more information about the z14 ZR1 and its features, see *IBM z14 Model ZR1 Technical Guide,* SG24-8651.

For more information about all of the I/O features, see *IBM Z Connectivity Handbook*, SG24-5444.

5.2.1 z/OS

z/OS Version 2 Release 1 is the earliest in-service release that is supported on the z14 ZR1. Although service support for z/OS Version 1 Release 13 ended in September of 2016, a fee-based extension for defect support (for up to three years) is available by ordering IBM Software Support Services - Service Extension[1] for z/OS 1.13.

z14 ZR1 capabilities differ depending on the z/OS release. Toleration support is provided on z/OS V1R13. Exploitation support is provided only on z/OS V2R1 and later by way of PTFs.

For more information about all z14 ZR1 features and functions that are supported by the z/OS releases, see *IBM z14 Model ZR1 Technical Guide,* SG24-8651.

Web deliverables

For more information about z/OS downloads, see the z/OS website.

5.2.2 z/VM

z/VM V6R4[2] provides support that enables guests to use the following functions that are supported by z/VM on z14 ZR1:

- z/Architecture support
- New hardware facilities
- ESA/390-compatibility mode for guests
- Crypto Clear Key ECC operations
- Encrypted paging support
- Guest exploitation support:
  - Pause-less garbage collection for Java
  - Instruction Execution Protection Facility
- Dynamic I/O support for managing the configuration of:
  - OSA-Express7S and OSA-Express6S for OSD and OSX CHPID types
  - FICON Express16S+ for FC and FCP CHPID types
  - zHyperLink Express
  - Coupling Express LR
  - 25GbE RoCE Express2
  - 10GbE RoCE Express2
- Improved memory management support

z/VM logical partitions: z14 ZR1 CPs and IFLs feature increased capacity over the capacity of their predecessors (on z13s). 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 machines might also need adjustment.

[1] Beginning with z/OS V1.12, IBM Software Support Services replaced the IBM Lifecycle Extension for z/OS offering with a service extension for extended defect support.

[2] z/VM V6R3 is *not* supported on the z14 Model ZR1

IBM z/VM V7.1 offers a new release delivery and new function support that provides predictability for lifecycle management of z/VM systems through continuous delivery. z/VM V7.1 includes the following enhancements:

- Integration of the Single System Image (SSI) function into the base, at no extra cost
- Enhancements to the dump process to reduce the time that is required to create and process dumps
- Upgrades to a new Architecture Level Set
- More functionality will be delivered as service after general availability
- Enhancements to the dynamic configuration capabilities of a running z/VM system with Dynamic Memory Downgrade support

For more information about PTF availability, see the z/VM Continuous Delivery News web page.

For more information about all z14 ZR1 features and functions that are supported by the z/VM releases, see *IBM z14 Model ZR1 Technical Guide,* SG24-8651.

5.2.3 z/VSE

z14 Model ZR1 support is provided by z/VSE V5R2 and later. Consider the following points:

- z/VSE runs in z/Architecture mode only.
- z/VSE V5 has an architectural level set (ALS) that requires IBM System z9® or later.
- z/VSE V6.1 has an ALS that requires IBM System z10® or later.
- z/VSE V6.2 has an ALS that requires IBM zEnterprise 196 or later.
- z/VSE V6.2 supports High-Performance FICON (zHPF) and SIMD.

For more information about all z14 ZR1 features and functions that are supported by the z/VSE releases, see *IBM z14 Model ZR1 Technical Guide,* SG24-8651.

5.2.4 z/TPF

z14 Model ZR1 support is provided by z/TPF V1R1 with PTFs. For more information about all z14 ZR1 features and functions that are supported by the z/TPF, see *IBM z14 Model ZR1 Technical Guide,* SG24-8651.

5.2.5 Linux on IBM Z

The earliest service levels of SUSE, Red Hat, and Ubuntu releases that are supported on the z14 ZR1 are listed in Table 5-2.

Table 5-2 Linux on IBM Z distributions

Linux on IBM Z distribution[a]	Earliest supported version and release
SUSE Linux Enterprise Server	SLES 12 SP2 with service
SUSE Linux Enterprise Server	SLES 11 SP4 with service
Red Hat Enterprise Server	RHEL 7.3 with service
Red Hat Enterprise Server	RHEL 6.9 with service
Canonical	Ubuntu 16.04 LTS with service

a. Only z/Architecture (64-bit mode) is supported. IBM testing identifies the “minimum required level” and the “recommended levels” of the tested distributions.

For more information about supported Linux distributions on Z servers, see the Linux on IBM Z page of the IBM Z website.

For more information about all z14 ZR1 features and functions that are supported by the Linux on IBM Z distributions, see *IBM z14 Model ZR1 Technical Guide,* SG24-8651.

5.2.6 KVM hypervisor

With Announcement Letter 917-04, dated March 7, 2017, IBM announced that it is working with Linux distributions to support the KVM component on IBM Z.

For more information about KVM support for the Z platform, see your distribution’s documentation.

5.3 Software licensing

The Z 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 IBM Z environments. For the z14 ZR1, the following metric groups for software licensing are available from IBM, depending on the software product:

- Monthly license charge (MLC)

  MLC pricing metrics feature a recurring monthly charge. In addition to the permission to use the product, the charge includes access to IBM product support during the support period. MLC pricing applies to z/OS, z/VSE, and z/TPF operating systems. Charges are based on processor capacity, which is measured in millions of service units (MSU) per hour.

- IPLA

  IPLA metrics feature 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 extra charge.

Software Licensing References

For more information about software licensing, see the following resources:

- Learn about Software licensing
- Base license agreements
- IBM Z Software Pricing reference guide
- The IBM International Passport Advantage® Agreement can be downloaded from the Learn about Software licensing website

Subcapacity pricing terms for z/VM and select z/VM-based programs

Subcapacity pricing for the z/VM V6 operating environment is available to clients that are running z/VM Version 6 Release 3 or higher. Software pricing at less than full machine capacity can provide more flexibility and improved cost of computing as you manage the volatility and growth of new workloads.

For more information about subcapacity pricing terms for z/VM and z/VM-based programs, see announcement letter 217-267, dated July 17, 2017.

For more information about software licensing options that are available for z14 ZR1, see *IBM z14 Model ZR1 Technical Guide,* SG24-8651.

IBM Secure Service Container

The IBM Secure Service Container expanded to integrate with the IBM Cloud Private platform for hybrid and private cloud deployments on IBM Z. You can deploy containerized IBM Middleware applications and use common management tooling for deploying homegrown or other third-party Docker and Kubernetes-based applications.

IBM Secure Service Container supports the deployment of software container technology without requiring application changes. This support is especially useful considering the regulatory focus on protecting critical data from internal and external threats. Support features (0103 and 0104) ensure that Secure Service Container applications run smoothly, whether you use only one application type, or many types.

5.4 References

For current planning information, see the following operating system web pages:

- z/OS
- z/VM
- z/VSE
- z/TPF
- Linux on IBM Z

Redbooks

IBM z14 Model ZR1 Technical Introduction

(0.2”spine)
0.17”<->0.473”
90<->249 pages

SG24-8550-01

ISBN 0738457183

Printed in U.S.A.