Front cover

Redbooks
ibm.com/redbooks

IBM z17 Technical Introduction

Bill White

Ewerson Palacio

IBM Z

IBM

Redbooks

IBM Redbooks

IBM z17 Technical Introduction

July 2026

SG24-8580-01

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

Second Edition (July 2026)

This edition applies to IBM z17 ME1, IBM z17 ME2, and IBM z17 MER.

© Copyright International Business Machines Corporation 2025, 2026.
Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices vii
Trademarks viii

Preface ix
Authors ix
Now you can become a published author, too! x
Comments welcome x
Stay connected to IBM Redbooks x

Chapter 1. The intelligent core for modern enterprise workloads 1
1.1 Addressing enterprise challenges with IBM z17 2
1.2 IBM z17 technical highlights 4
1.3 AI, automation, and security: Exploring IBM z17 use cases 7
1.3.1 Enabling IBM z17 for AI-driven innovation 7
1.3.2 Automating operations and improving efficiency 8
1.3.3 Enhancing security and risk management 8

Chapter 2. IBM z17 ME1 hardware overview 11
2.1 IBM z17 ME1 upgrade paths 12
2.2 Frames and cabling 13
2.3 CPC drawers 16
2.3.1 Dual-chip modules 17
2.3.2 Processor unit 17
2.3.3 Memory 20
2.3.4 Hardware System Area 22
2.4 I/O system structure 22
2.5 Power and cooling 24
2.5.1 Power options 24
2.5.2 Cooling options 25

Chapter 3. IBM z17 ME2 and IBM z17 MER hardware overview 27
3.1 IBM z17 ME2 and IBM z17 MER upgrade paths 28
3.2 Frame and cabling 29
3.3 CPC drawers 31
3.3.1 Dual-chip modules 33
3.3.2 Processor unit 33
3.3.3 Memory 35
3.3.4 Hardware System Area 37
3.4 I/O system structure 37
3.5 Power and cooling 40
3.5.1 Power options 40
3.5.2 Cooling options 40
3.5.3 Power considerations 40

Chapter 4. IBM z17 supported features and functions 41
4.1 IBM z17 I/O connectivity overview 42
4.2 Introducing the IBM z17 I/O design 43
4.3 Storage connectivity 43
4.3.1 FICON Express features 43

© Copyright IBM Corp. 2025.

4.3.2 zHyperLink Express 44
4.4 Network connectivity 44
4.4.1 Network features 45
4.4.2 HiperSockets 45
4.4.3 ISM communications 46
4.5 Clustering connectivity 46
4.5.1 Dynamic I/O configuration 47
4.5.2 Quantum-safe protection for Coupling Facility Control Code 47
4.6 IBM Z Coordinated Server Time 47
4.6.1 STP implementation with IBM z17 48
4.6.2 Network Time Protocol Client support 48
4.6.3 Precision Time Protocol client support 48
4.6.4 Pulse per second support 49
4.7 Cryptography 49
4.7.1 Central Processor Assist for Cryptographic Functions 50
4.7.2 Crypto Express8S 51
4.7.3 Crypto Express7S (carry forward only) 52
4.8 Special-purpose features and functions 53
4.9 Hardware management 53

Chapter 5. IBM z17 system design strengths 55
5.1 IBM z17 technology enhancements 56
5.1.1 Processor design highlights 56
5.1.2 Architectural enhancements 56
5.1.3 Integrated accelerators 57
5.1.4 I/O subsystem improvements 57
5.1.5 Memory design 58
5.1.6 IBM Virtual Flash Memory 58
5.2 IBM Z virtualization 58
5.2.1 Hardware virtualization 59
5.2.2 Cloud environments 60
5.3 IBM z17 capacity 61
5.3.1 Capacity settings 61
5.3.2 CoD offerings 63
5.3.3 IBM Z Flexible Capacity for Cyber Resiliency 64
5.3.4 Tailored Fit Pricing for IBM Z 64
5.4 IBM z17 performance 65
5.4.1 LSPR workload suite: IBM z17 changes 65
5.4.2 IBM z17 throughput optimization 66
5.4.3 HiperDispatch 67
5.5 IBM z17 AI acceleration 68
5.5.1 AI acceleration architecture 68
5.5.2 Integrated AI Accelerator 68
5.5.3 AI workload integration 68
5.5.4 Composite AI architecture 69
5.5.5 Generative AI and AI services 69
5.5.6 AI development ecosystem 69
5.6 IBM z17 sustainability 70
5.7 IBM Z resiliency and cyber resiliency 70
5.7.1 Traditional resiliency foundations 70
5.7.2 IBM Z cyber resiliency 71
5.8 IBM Z cybersecurity 72
5.9 Automation on IBM Z 76

5.9.1 Automation in IBM Z AIOps 77
5.9.2 Application development and integration 77

Chapter 6. IBM z17 software support 79
6.1 Software support 80
6.1.1 Application development and languages 80
6.1.2 Supported IBM compilers 81
6.2 Support by operating system 83
6.2.1 z/OS 84
6.2.2 z/VM 84
6.2.3 z/TPF 84
6.2.4 Linux on IBM Z 85
6.2.5 Kernel-based Virtual Machine hypervisor 85
6.3 Software licensing 85

Abbreviations and acronyms 87

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 customer 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.

© Copyright IBM Corp. 2025.

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®
DS8000®
FICON®
GDPS®
IBM®
IBM Cloud®
IBM Concert®
IBM Security®
IBM Spyre™
IBM Telum®
IBM watsonx®
IBM Z®
IBM z16®
Language Environment®
Parallel Sysplex®
RACF®
Rational®
Redbooks®
Redbooks (logo) ®
Resource Link®
Spyre™
Terraform®
watsonx®
watsonx Assistant®
watsonx Code Assistant®
WebSphere®
z/OS®
z/VM®
z15®
z16®
z17™

The following terms are trademarks of other companies:

The registered trademark Linux® is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis.

Zowe, are trademarks of the Linux Foundation.

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

Red Hat, Ansible, OpenShift, are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

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

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

Preface

This IBM Redbooks® publication introduces the newest members of the IBM Z® family: the IBM z17™ ME1, IBM z17 ME2, and IBM z17 MER. It focuses on the architectural design strengths and software ecosystem that enable organizations to address modern enterprise requirements. It also highlights how IBM z17 supports key business outcomes, including operational resilience, security, scalability, and integration with hybrid cloud and AI-driven workloads.

This publication describes the core capabilities of IBM z17, including balanced system design across compute, memory, and I/O resources; advanced virtualization for consolidation and efficiency; integrated security and resiliency; and flexible capacity and automation models. These capabilities can help organizations optimize resource usage, simplify infrastructure management, and respond effectively to changing workload demands.

In addition, the publication outlines the supported software environment, including operating systems, development frameworks, and open-source technologies. It shows how IBM z17 enables application modernization, supports DevOps practices, and integrates AI capabilities directly into operational workflows, while maintaining compatibility with existing investments.

This publication examines the key components of IBM z17, providing insights and example scenarios that illustrate how organizations can apply these capabilities. It can help you understand how IBM z17 can support evolving infrastructure strategies and business objectives.

Authors

This book was produced by specialists from IBM Redbooks:

Bill White is an IBM Redbooks Project Leader and Senior IT Infrastructure Specialist at IBM Poughkeepsie, New York.

Ewerson Palacio is an IBM Redbooks Project Leader. He holds a bachelor's degree in math and computer science. Ewerson worked for IBM Brazil for over 40 years and retired in 2017 as an IBM Distinguished Engineer. Ewerson has co-authored many IBM Z Redbooks publications and created and presented IBM Redbooks seminars around the globe.

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

Jonathan Bradbury, Marc Coq, Anne Dames, Jacky Doll, Michael Grötzner, Dave Hollar, Richard Kisley, Nicolas Mäding, Purvi Patel, Priyal Shah, André Spahni, Dave Surman, Benedikt Wenz
IBM®

© Copyright IBM Corp. 2025.

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 IBM Redbooks residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

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

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

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

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

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, IBM Redbooks
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- 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 intelligent core for modern enterprise workloads

Organizations today face many pressures, including accelerating digital demand, evolving cyberthreats, increasing regulatory requirements, and the need to use AI at scale. Traditional approaches that are characterized by fragmented architectures, siloed data, bolt-on AI solutions, and reactive security models are insufficient to address these challenges consistently and efficiently.

As a result, many enterprises are shifting their focus from incremental adoption of AI capabilities to the integration of AI directly into core business processes. This shift reflects a growing need for real-time insight, where transactional systems record activity and analyze and act on data as it is created. In this context, both AI and generative AI are becoming integral to business workflows, supporting improved decision-making, operational efficiency, and service delivery.

IBM z17 can help support this migration by embedding AI, automation, and advanced security capabilities into the transactional core. By integrating these capabilities at the platform level, IBM z17 enables organizations to process, analyze, and protect data within a unified environment. This approach supports real-time intelligence, enhances cyber resilience, and streamlines operations, enabling enterprises to increase the value that is derived from mission-critical workloads while maintaining the reliability and scalability required for enterprise computing.

This chapter outlines the features of IBM z17 and illustrates the resulting business benefits.

This chapter describes the following topics:

- 1.1, “Addressing enterprise challenges with IBM z17” on page 2
- 1.2, “IBM z17 technical highlights” on page 4
- 1.3, “AI, automation, and security: Exploring IBM z17 use cases” on page 7

© Copyright IBM Corp. 2025, 2026.

1.1 Addressing enterprise challenges with IBM z17

Enterprises operating in modern digital environments must balance innovation with operational stability while addressing increasing demands for security, performance, and regulatory compliance. As organizations expand their use of hybrid cloud architectures and integrate AI into core workflows, these requirements become more complex and interdependent.

Transaction processing systems (TPSs) continue to play a central role in this landscape. They serve as systems of record and critical points of integration where data is generated, processed, and shared across distributed environments. As enterprises move toward architectures that require real-time insight and continuous availability, TPS platforms must evolve to support higher levels of performance, resilience, and intelligence.

IBM z17 can help address these evolving requirements by providing an integrated platform that combines advanced processing capabilities, built-in AI acceleration, automation, and comprehensive security. This integration enables organizations to support AI and hybrid cloud strategies while maintaining the reliability and scalability expected of enterprise transaction systems.

IBM z17 offers targeted solutions that can address these obstacles effectively, enhancing both performance and security while supporting sustainable operations:

- Data management and performance
  - Challenge: Efficiently managing increasing volumes of data while sustaining high transaction throughput and minimal latency.
  - Solution: IBM z17 delivers enhanced performance and scalability through the following capabilities:
    - The IBM Telum® II processor, which provides increased core density, expanded processing capacity, and architectural improvements
    - Large-scale memory that supports high-speed data access and processing
    - Optimized cache structures that improve data locality and system efficiency
    - High-bandwidth I/O capabilities to support large-scale data movement and transactional workloads

  These capabilities enable consistent performance for high-volume, latency-sensitive workloads.
- Data security, compliance, and resilience
  - Challenge: Protecting sensitive data against current and emerging threats while maintaining compliance with industry and regional regulations.
  - Solution: IBM z17 strengthens enterprise security and resilience by providing the following capabilities:
    - Advanced cryptographic capabilities, including support for quantum-safe approaches
    - Confidential computing technologies that protect data in use
    - Centralized management of cryptographic assets and secrets
    - Built-in resilience features, including rapid recovery mechanisms to minimize downtime
    - Support for geographically distributed operations and dynamic resource activation to enhance DR

These capabilities improve visibility, control, and protection of sensitive information across the enterprise.

- AI-driven innovation
  - Challenge: Integrating AI into business processes to enable real-time insight, improve decision-making, and support new application models.
  - Solution: IBM z17 enables AI adoption within transactional environments through the following capabilities:
    - On-chip AI acceleration with the IBM Telum II processor, supporting real-time inference
    - The IBM Spyre™ Accelerator, helping support scalable, on-premises generative AI workloads
    - Integration of AI into transaction processing pipelines, enabling analysis at the point of data creation
    - Efficient resource usage to support AI workloads without disproportionate infrastructure growth

  These capabilities help organizations to incorporate both predictive and generative AI into core operations without disrupting existing applications.
- Operational efficiency and automation
  - Challenge: Reducing operational complexity, improving consistency, and accelerating time to delivery.
  - Solution: IBM z17 improves operational efficiency through the following capabilities:
    - End-to-end automation of infrastructure and system management tasks
    - Support for infrastructure as code (IaC) to enable standardized and repeatable deployments
    - Integrated observability, including OpenTelemetry based monitoring and system-wide visibility
    - Reduced reliance on specialized skills through automation and tools

  This approach helps organizations to streamline operations while maintaining control and reliability.
- Sustainability
  - Challenge: Reducing environmental impact while continuing to meet performance and availability requirements.
  - Solution: IBM z17 supports sustainability objectives through the following capabilities:
    - Energy-efficient system design, reducing power consumption compared to previous generations
    - High-density computing, enabling consolidation of workloads and reduced data center footprint
    - Efficient resource usage across workloads, including AI processing

  These improvements contribute to smaller operational costs and a reduced environmental impact.

IBM z17 extends the role of traditional TPSs by integrating AI, automation, and advanced security into a unified platform. By addressing key challenges in performance, security, operational efficiency, and sustainability, it enables organizations to evolve their transaction processing environments to support modern hybrid cloud and AI-driven workloads. From an architectural perspective, IBM z17 serves as the transactional core of these hybrid environments, integrating with distributed systems and cloud services while maintaining centralized control over data processing, security, and operations. This approach preserves the strengths of high-volume transaction processing while enabling real-time, AI-driven capabilities across the enterprise.

1.2 IBM z17 technical highlights

IBM z17 is delivered as a family of systems that share a common architecture while supporting different scales and deployment models. The offerings include the following systems:

- IBM z17 ME1: A multi-frame, enterprise-scale system for maximum capacity and scalability
- IBM z17 ME2: A compact, single-frame system for smaller footprints with reduced scale
- IBM z17 MER: A rack-based system that is installed in a client-supplied 19-inch rack, offering deployment flexibility

All IBM z17 systems share a common architectural foundation, including processor units (PUs), central processor complex (CPC) drawers, memory subsystems, and a PCIe-based I/O infrastructure. The differences between these systems are primarily in scale, packaging, and supported capacity ranges.

This architecture enables scalability from a single frame up to four frames based on the specific system requirements and configurations. Figure 1-1 shows the diverse configuration possibilities.

Figure 1-1 IBM z17 configuration options

IBM z17 is designed for seamless integration, inheriting compatibilities from prior IBM Z servers. This continuity supports a range of upgrade paths through five distinct features: Max43, Max90, Max136, Max183, and Max208 with the IBM z17 ME1. IBM z17 ME2 and IBM z17 MER offer four distinct features: Max6, Max18, Max39, and Max82. Each feature's name reflects the maximum number of characterizable PUs that is available for configuration.

Table 1-1 lists some of the IBM z17 ME1, IBM z17 ME2, and IBM z17 MER technical attributes at a glance.

Table 1-1 IBM z17 technical attributes

Technical attribute	IBM z17 ME1	IBM z17 ME2 and IBM z17 MER
Frames and form factor	1 - 4 IBM Z frames	Single IBM Z frame or client rack
CPC drawers	1 - 4	1 or 2
I/O drawers[a]	Up to 12	Up to three
Processor chip IBM Telum II	► 5 nm technology ► 5.5 GHz ► 8 cores per chip ► 24.1 miles of wire per chip ► 43 billion transistors	► 5 nm technology ► 4.8 GHz ► 8 cores per chip ► 24.1 miles of wire per chip ► 43 billion transistors
Cache	► 128 KB L1 (separate instruction and data caches per core) ► 36 MB of L2 ► 360 MB of virtual L3 ► 2.88 GB of virtual L4	► 128 KB L1 (separate instruction and data caches per core) ► 36 MB of L2 ► 360 MB of virtual L3 ► 2.88 GB of virtual L4
Capacity	208 characterizable PUs	82 characterizable PUs
Total system memory	► 4U DDR5 DDIMM ► 64 TB maximum memory	► 4U DDR5 DDIMM ► 17.96 TB maximum memory
IBM Virtual Flash Memory (VFM)	Up to 6 TB in 12 increments of 0.5 TB	Up to 2 TB, in four increments of 0.5 TB
On-chip AI acceleration (Artificial Intelligence Unit (AIU))	Built for ultra-low latency, the AIU can deliver submillisecond inference under specific workloads and configurations.	Built for ultra-low latency, the AIU can deliver submillisecond inference under specific workloads and configurations.
Off-chip PCIe Accelerator card	► IBM Spyre Accelerator ► 32 Gen AI-ready cores on extended adapters ► Up to 48 adapters per system	► IBM Spyre Accelerator ► 32 Gen AI-ready cores on extended adapters ► Up to 24 adapters per system
On-chip I/O Data Processing Unit (DPU)	A 64 KB L1 cache that is dedicated to processing I/O acceleration and virtualization	A 64 KB L1 cache that is dedicated to processing I/O acceleration and virtualization.
I/O subsystem	Six LCSSs in four subchannel sets across 85 LPARs	Three logical LCSSs in three subchannel sets across 40 LPARs
Hardware System Area (HSA)	884 GB is used to manage critical firmware functions, partition definitions, and channel subsystems (CSSs) for I/O.	476 GB is used to manage critical firmware functions, partition definitions, and CSSs for I/O.

a. The number of I/O drawers is determined by the total number of features that is ordered. (Up to 16 features per I/O drawer.)

IBM z17 also provides the following capabilities:

- A single instruction multiple data (SIMD) processor increases parallelism to accelerate analytics and AI processing. Also, simultaneous multithreading (SMT) increases processing efficiency and throughput and raises the number of instructions in flight. Special coprocessors and new hardware instructions for accelerating selected workloads are also available.
- The IBM Z Sort accelerator helps reduce CPU costs, speed up the sorting process, and improve database functions.
- IBM Integrated Accelerator for IBM zEnterprise Data Compression (zEDC) improves data compression operations by having one dedicated compression coprocessor per PU (on-chip compression), and new hardware instructions.
- For improved I/O channel performance, the application-specific integrated circuit (ASIC) functions that are part of the I/O features were moved to assist processors in the PU chip, which is called the DPU. The corresponding I/O features are IBM FICON® Express32-4P (4 ports) for storage connectivity and Network Express (2 ports) for network connectivity.
- The Integrated Network Adapter 100G (dual ports in the CPC drawer) delivers higher throughput by using RDMA-based data transfers. This approach reduces data movement delays and enables smooth handling of high-volume, latency-sensitive workloads, making it ideal for environments constrained by bandwidth, latency, or CPU usage.
- Just-in-time deployment of processor resources. The CoD function can dynamically change available system capacity. This function can help respond to new business requirements with flexibility and precise granularity.
- Tailored Fit Pricing for IBM Z options are designed to deliver unmatched simplicity and predictability of hardware capacity and software pricing.
- The same high quality of service and RAS that is traditional in the IBM Z platform is provided. The RAS strategy uses a building-block approach that meets the stringent requirements for achieving continuous, reliable operation.
- Enhanced resiliency with IBM Z Flexible Capacity for Cyber Resiliency to make system resources available across locations to proactively avoid disruptions.
- Integrated quantum-safe protection through quantum-safe cryptography APIs and crypto-discovery tools. Quantum-safe secure boot technology helps protect IBM z17 firmware from quantum attacks through a built-in dual signature scheme with no configuration changes required for enablement.

For more information about IBM z17 configuration options, hardware components, and upgrade paths, see Chapter 2, “IBM z17 ME1 hardware overview” on page 11 and Chapter 3, “IBM z17 ME2 and IBM z17 MER hardware overview” on page 27.

For more information regarding cryptographic features and storage, network, and clustering connectivity options, see Chapter 4, “IBM z17 supported features and functions” on page 41.

The enterprise challenges and corresponding platform capabilities highlight the need for an integrated approach, which is further illustrated through representative use cases in the following section.

1.3 AI, automation, and security: Exploring IBM z17 use cases

To demonstrate how integrated AI, automation, and security features can be used to improve business outcomes, operational efficiency, and risk management, three use cases are highlighted. Rather than treating these capabilities independently, IBM z17 enables their combined use within transactional environments, where data is created, processed, and acted on in real time.

The following examples cover three areas:

- AI-driven innovation and growth
- Operational automation and efficiency
- Security and risk mitigation

Each use case explores the transformative potential of IBM z17, offering practical insights for businesses looking to optimize their operations.

1.3.1 Enabling IBM z17 for AI-driven innovation

As organizations integrate AI into core operations, the ability to run AI models in real time within transactional workflows becomes important. IBM z17 supports this approach by embedding AI capabilities directly into the platform. These capabilities enable AI to run as a part of transaction processing.

Key capabilities include the following ones:

- Composite AI integration

  IBM z17 supports the combined use of multiple AI techniques, including machine learning (ML), deep learning, and LLMs. This approach enables organizations to apply different models to a single problem domain, improving accuracy and enabling broader analysis across both structured and unstructured data.

- On-chip AI acceleration with IBM Telum II

  The IBM Telum II processor enables real-time inference within high-volume transaction processing. This approach enables AI models to be invoked as part of the transaction itself, supporting immediate decision-making without requiring data movement to external systems.

- Integrated AI acceleration architecture

  AI acceleration resources are accessible across processor cores, enabling efficient scaling of inference workloads. This architecture supports high-throughput AI processing for latency-sensitive use cases, improving responsiveness and maintaining transactional performance.

- Support for scalable AI workloads with IBM Spyre™ Accelerator

  IBM Spyre extends platform capabilities to support larger and more complex AI models, including generative AI. This approach enables organizations to deploy and scale AI workloads within controlled, on-premises environments, including support for emerging agentic AI systems that can autonomously run tasks and workflows.

- Secure execution of AI workloads

  AI models and data are protected through hardware-based encryption, isolated execution environments, and integrated security controls. This approach helps maintain data privacy and supports compliance requirements when applying AI to sensitive workloads.

These capabilities enable AI integration across a range of transactional use cases, including fraud detection, risk analysis, personalization, and operational optimization.

The business impact of these capabilities includes the following items:

- Real-time decision-making within transactions
- Improved prediction accuracy through combined AI models
- Reduction in the latency that is associated with external AI processing
- Support for adoption of generative AI within enterprise environments

1.3.2 Automating operations and improving efficiency

Operational complexity is a challenge in hybrid cloud environments. IBM z17 addresses this challenge by integrating automation and observability into the platform.

Key capabilities include the following ones:

- End-to-end automation

  Routine operational tasks, including provisioning, configuration, and lifecycle management, can be automated to improve consistency and reduce manual effort.

- IaC

  Standardized deployment models enable environments to be defined and managed programmatically, improving repeatability and reducing configuration variability.

- Integrated observability

  Support for OpenTelemetry provides visibility across applications, infrastructure, and system components, enabling faster identification and resolution of issues. This level of observability reduces operational blind spots and improves system-wide transparency, enabling more proactive and informed operational decision-making.

- Operational standardization

  Automation reduces reliance on specialized skills and helps enforce consistent operational practices across environments.

The business impact of these capabilities includes the following items:

- Reduced operational impact and manual intervention
- Faster deployment and time-to-market
- Improved system reliability through consistent processes
- Enhanced visibility into system behavior and performance

1.3.3 Enhancing security and risk management

Security remains a foundational requirement for enterprise systems, particularly as organizations increase data sharing and AI adoption. IBM z17 integrates advanced security capabilities to address both current and emerging risks.

Key capabilities include the following ones:

- Advanced cryptographic capabilities

  Support for modern encryption techniques, including quantum-safe approaches, helps protect data against evolving threats.

- Confidential computing

  Sensitive data can be processed in protected environments, reducing exposure during computation.

- Centralized secrets and cryptographic management

  Improved visibility and control over cryptographic assets helps reduce risk and simplify compliance. These capabilities also support digital sovereignty requirements by enabling organizations to maintain control over data location, access, and processing in accordance with regional and regulatory constraints.

- Resilience and recovery mechanisms

  Built-in capabilities support rapid recovery and continuous availability, minimizing disruption in failures or attacks.

- AI-assisted security operations

  AI techniques can be applied to identify anomalies, detect threats, and support proactive risk mitigation.

The business impact of these capabilities includes the following items:

- Improved protection of sensitive data and workloads
- Enhanced ability to meet regulatory and compliance requirements
- Reduced risk of data exposure or system disruption
- Faster detection and response to security events

These use cases illustrate how IBM z17 integrates AI, automation, and security within transactional environments to address the challenges that are outlined in this chapter. By embedding these capabilities at the platform level, IBM z17 enables organizations to evolve from traditional TPSs into intelligent, adaptive platforms that support real-time insight, operational consistency, and resilient business operations.

As enterprises continue to adopt AI-driven, hybrid architectures, platforms such as IBM z17 play a central role in bridging established transaction processing with modern requirements for intelligence, automation, and security. This integration simplifies architectures, strengthens resilience, and increases the value that is derived from mission-critical workloads. This evolution reflects a broader shift toward doing more at the core, where TPSs become integrated platforms for secure, automated, and intelligent data processing.

For more information about the IBM z17 capabilities that underpin these use cases, see Chapter 5, “IBM z17 system design strengths” on page 55.

2

IBM z17 ME1 hardware overview

This chapter introduces the IBM z17 hardware architecture as a reference implementation. It describes the key design principles, components, and functional elements that define the system. This architectural foundation serves as the basis for the alternative implementations that are described in Chapter 3, “IBM z17 ME2 and IBM z17 MER hardware overview” on page 27.

This chapter describes the following topics:

- 2.1, “IBM z17 ME1 upgrade paths” on page 12
- 2.2, “Frames and cabling” on page 13
- 2.3, “CPC drawers” on page 16
- 2.4, “I/O system structure” on page 22
- 2.5, “Power and cooling” on page 24

© Copyright IBM Corp. 2025, 2026.

2.1 IBM z17 ME1 upgrade paths

IBM z17 ME1 is built by using the IBM Telum II processor design. Each processor chip consists of eight cores. Two processor chips are packaged in the dual-chip module (DCM). Each DCM can have 9 - 11 or 10 - 15 active processor unit (PU) cores, depending on the configuration. With IBM z17 ME1, the maximum number of characterizable processors is represented by the feature names (see Table 2-1): Max43, Max90, Max136, Max183, and Max208. Spare PUs, System Assist Processors (SAPs), and two Integrated Firmware Processors (IFPs) are included in the IBM z17 ME1 configuration.

The number of characterizable PUs, SAPs, and spare PUs for the various features is listed in Table 2-1. For more information about PU characterization types, see “PU characterization” on page 19.

Table 2-1 IBM z17 ME1 processor unit configurations

Feature name	Number of CPC drawers	Feature Code	Characterizable processor units	SAPs	Spares
Max43	1	0571	0 - 43	5	2
Max90	2	0572	0 - 90	10	2
Max136	3	0573	0 - 136	16	2
Max183	4	0574	0 - 183	21	2
Max208	4	0575	0 - 208	24	2

IBM z17 ME1 helps ensure continuity and upgrades from IBM z16® A01 and IBM z15® T01. The supported upgrade paths are shown in Figure 2-1.

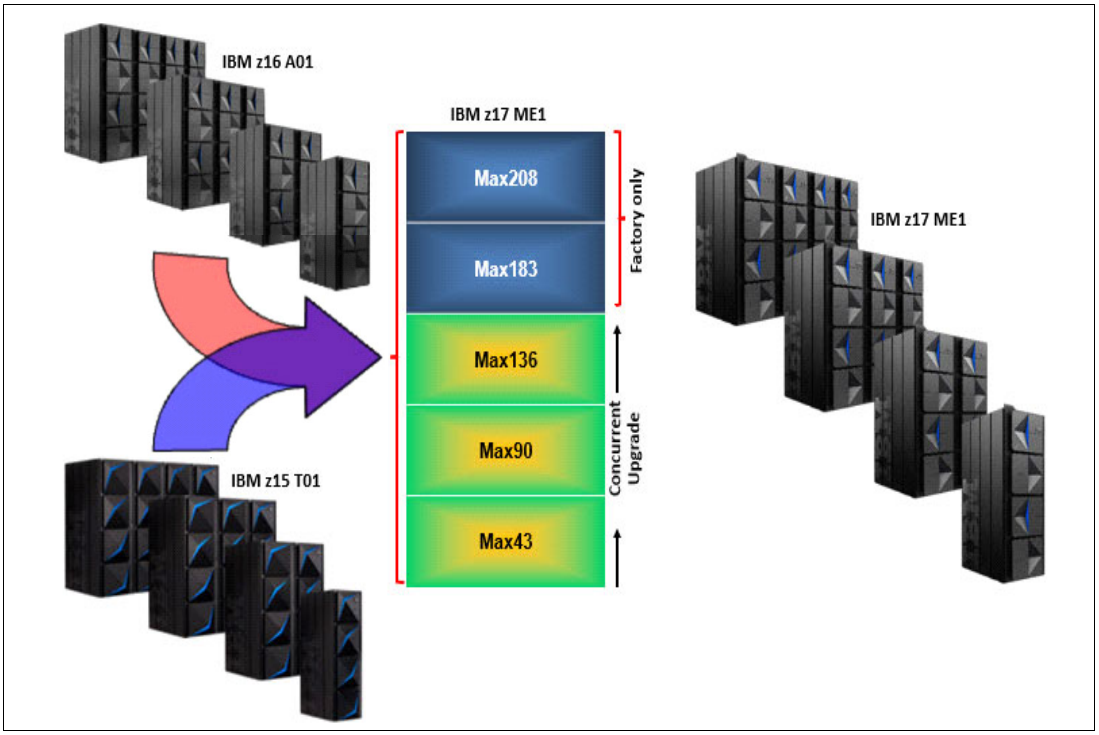

Figure 2-1 IBM z17 ME1 upgrade paths

If an upgrade request cannot be fulfilled by using the existing configuration, a hardware upgrade is required in which one or two CPC drawers are added to accommodate the necessary capacity. With IBM z17 ME1, CPC drawers can be installed concurrently from a Max43 to a Max90, and to a Max136.

Note: No field upgrade is available to a Max183 or a Max208; these two features are factory-shipped only.

With IBM z17 ME1, concurrent upgrades are available for central processors (CPs), Integrated Facilities for Linux (IFLs), Internal Coupling Facilities (ICFs), and IBM Z Integrated Information Processor (zIIP) processors. However, concurrent PU upgrades require that more PUs are physically installed but not activated.

In a PU failure, one of the spare PUs is immediately and transparently activated and assigned the characteristics of the failing PU. Two spare PUs are always available on an IBM z17 ME1.

IBM z17 ME1 offers 337 capacity levels. In all, 208 capacity levels are based on the number of physically used CPs, plus up to 129 subcapacity models for the first 43 CPs.

For more information, see 5.3.1, “Capacity settings” on page 61.

2.2 Frames and cabling

IBM z17 ME1 employs 19-inch frames and adheres to industry-standard power and hardware norms. Configurable as a 1-4 frame system, it occupies two standard 24-inch (60-centimeter) floor tiles per frame, which fits contemporary data center layouts.

The IBM z17 configuration options as compared to previous IBM Z servers are listed in Table 2-2.

Table 2-2 IBM z17 ME1 configuration options as compared to IBM z16 A01 and IBM z15 T01 configurations

System	Number of frames	Number of CPC drawers	Number of I/O drawers	I/O and power connections	Power options[a]	Cooling options
IBM z17 ME1	1 - 4	1 - 4	0 - 12	Rear only	PDU only	Radiator (air) only
IBM z16® A01	1 - 4	1 - 4	0 - 12[b]	Rear only	PDU or BPA	Radiator (air) only
IBM z15 T01	1 - 4	1 - 5	0 - 12[b]	Rear only	PDU or BPA	Radiator (air) or water-cooling unit

a. The PDU option supports the air-cooling (radiator) option. The Bulk Power Assembly (BPA) option supports both air-cooling and water-cooling options.
b. A maximum of 12 if ordered with the PDU option or a maximum of 11 if ordered with the BPA option.

The number of PCIe+ I/O drawers can vary based on the number of I/O features, and the number of CPC drawers that is installed. For a maximum configuration, up to 12 PCIe+ I/O drawers can be installed. PCIe+ I/O drawers can be added concurrently.

Also, IBM z17 ME1 supports top-exit options for the fiber-optic and copper cables that are used for I/O and power. These options give you more flexibility in planning where the system is installed, eliminate the need for cables to be run under a raised floor, and increase air flow over the system.

IBM z17 ME1 supports installation on raised floor and non-raised floor environments.

Figure 2-2 shows the front view of a fully configured IBM z17 ME1 with radiator cooling, four CPC drawers, and 12 PCIe+ I/O drawers.

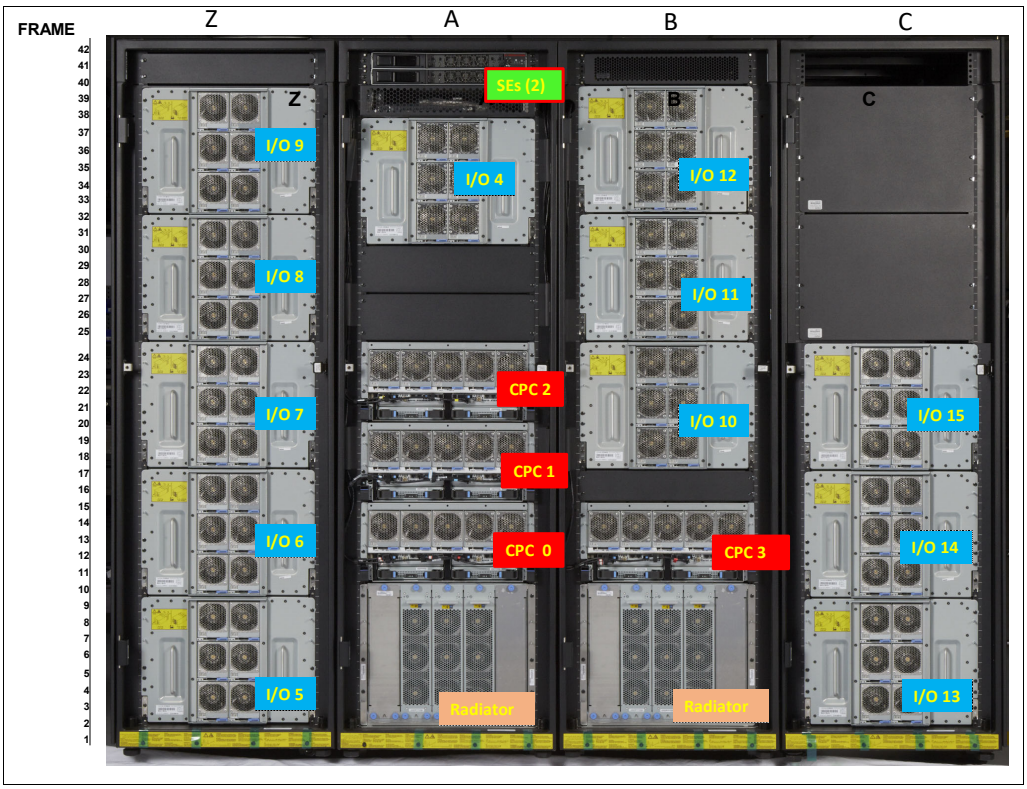

Figure 2-2 Front view of a fully configured IBM z17 ME1 with radiator cooling

Figure 2-3 shows the rear view of a fully configured IBM z17 ME1 with a total of 16 drawers (I/O and CPC combined), and two radiator cooling units.

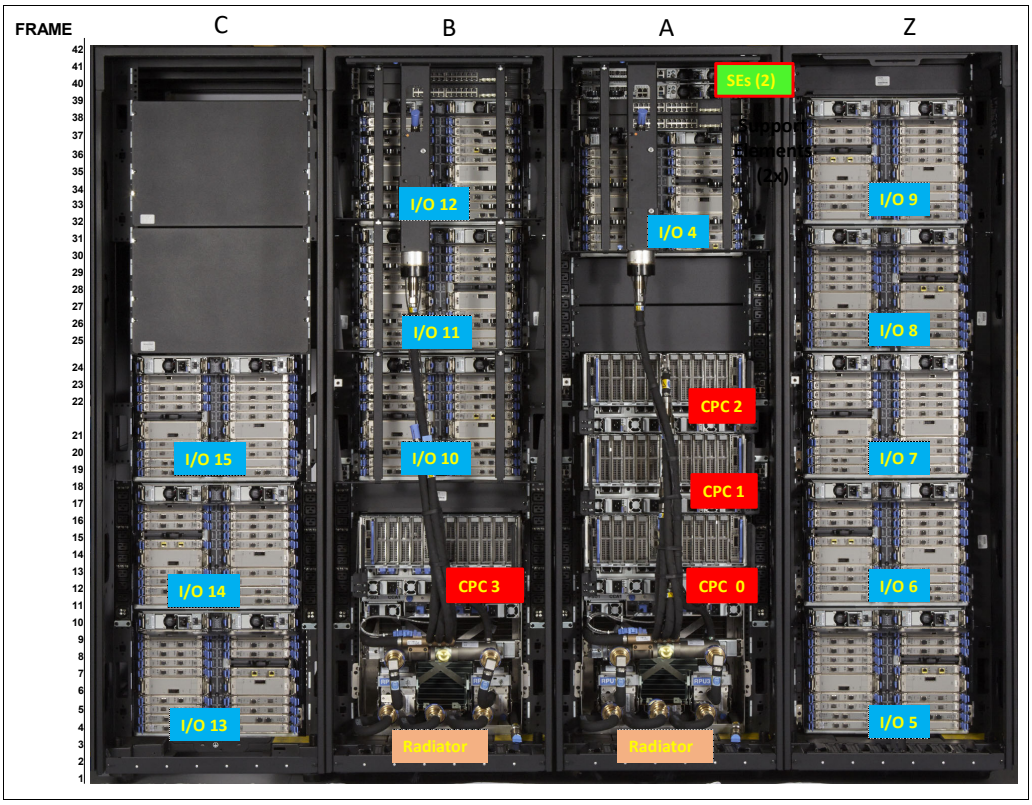

Figure 2-3 Rear view of a fully configured IBM z17 ME1

During the ordering process, the IBM configurator determines the necessary frame count and optimally positions CPC and PCIe+ I/O drawers. Factors that determine the number of frames for an IBM z17 ME1 configuration include the following features:

- Number of CPC drawers
- Plan-ahead features for more CPC drawers
- Number of I/O features (determines the number of PCIe+ I/O drawers)
- Number of PDUs

2.3 CPC drawers

IBM z17 ME1 can be configured with up to four CPC drawers (three in the A Frame and one in the B Frame). Each CPC drawer contains the following elements:

- DCMs

  Four DCMs with internal water loops providing dedicated cooling for every PU chip.

- Memory
  - The memory allocation ranges from a base 512 GB to an extended 64 TB with four CPC drawer configurations, excluding the 884 GB reserved for the Hardware System Area (HSA). For more information, see Table 2-3 on page 20.
  - A CPC drawer can accommodate up to 48 DIMMs, ranging in capacity from 32, 64, 128, 256, or 512 GB each.

- Fanouts

  Each CPC drawer supports up to 12 Gen4 PCIe+ fanout adapters to connect to the Gen4 PCIe+ I/O drawers, and contains Integrated Coupling Adapter Short Reach (ICA SR) coupling links:
  - A 2-port PCIe 32 GBps I/O fanout. Each port supports one domain in the 16-slot Gen4 PCIe+ I/O drawers.
  - ICA SR2.0 PCIe fanouts for coupling links (two links of 8 GBps each).

- Four power supply units (PSUs) provide power to the CPC drawer and are accessible from the rear. The loss of one PSU leaves enough power to satisfy the power requirements of the entire drawer. The PSUs can be concurrently maintained.
- Two dual-function Baseboard Management Controllers (BMCs) or oscillator cards (OSCs) provide redundant interfaces to the internal management network and provide clock synchronization to the IBM Z platform.
- Two dual-function Processor Power Cards (PPCs) that control voltage regulation, PSUs, and fans. The PPCs are redundant and can be concurrently maintained.
- The CPC drawer features five front-mounted fans, which help ensure efficient airflow and cooling for all components except the PU DCMs, which are internally cooled by using a water loop.

The CPC drawer communication topology is shown in Figure 2-4 on page 17. All CPC drawers are interconnected with high-speed communication link SMP-10 Assemblies (A-Bus) that are directly attached to the DCMs. The SMP-10 cables interconnect all the CPC drawers. The X-Bus provides connectivity between DCMs on the drawer, and the M-Bus connects the two PU chips on each DCM.

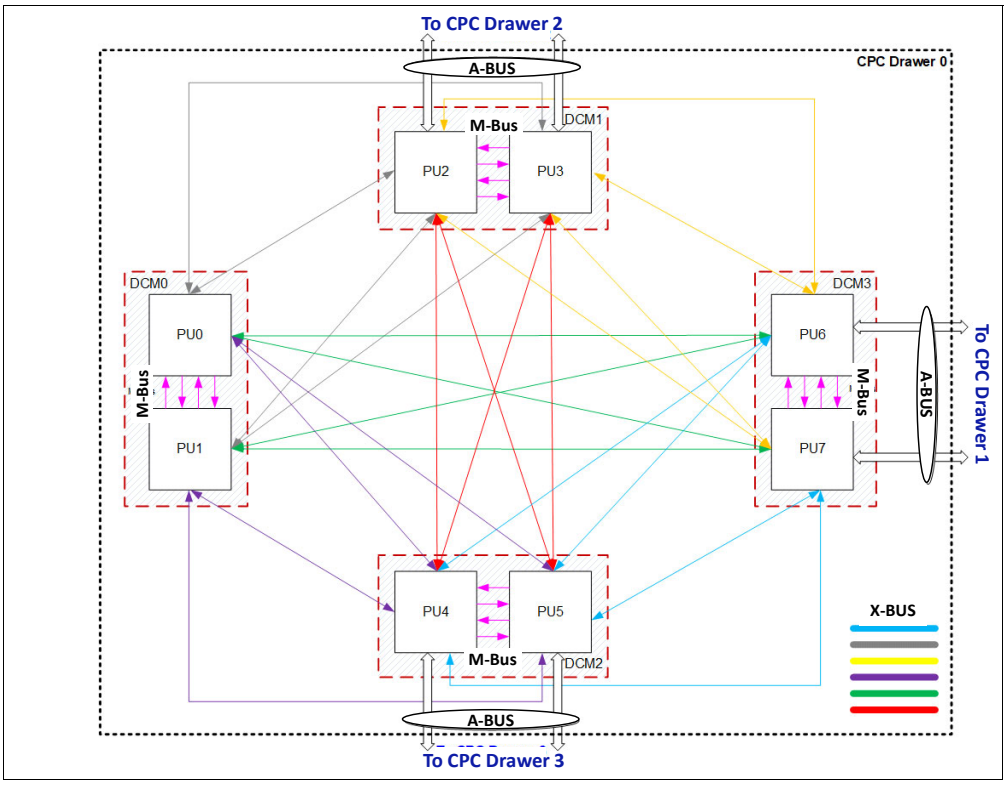

Figure 2-4 IBM z17 ME1 CPC drawer communication topology

Using a dedicated connection between the PU and storage controls, the IBM Processor Resource/Systems Manager (PR/SM) facility enables seamless operation and management of this memory-coherent symmetric multiprocessing (SMP) system.

2.3.1 Dual-chip modules

A DCM contains two PU chips. Each PU chip contains eight general purpose PUs with a 128 KB Instruction and Data L1 cache and ten 36 MB semi-private L2 caches. Each PU chip also has one Data Processing Unit (DPU) I/O Engine with access to the semi-private L2 caches for PCIe Gen5 x16 I/O operations.

2.3.2 Processor unit

Processor unit 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 10 instructions can run 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 a varying number of steps for different types of instructions. Several instructions can be running at any moment, and they are subject to the maximum number of decodes and completions per cycle.

Processor cache structure

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 semi-private L2 cache of 36 MB, which is next to the core.

Note: L1 and L2 are physical caches and implemented in dense static random access memory (SRAM).

- Each PU chip contains a 360 MB shared-victim virtual L3 cache, which is available to each PU core. The shared-victim virtual L3 cache is a logical construction that comprises all ten semi-private L2s (10 x 36 MB = 360 MB) belonging to all cores on the same chip[1].
- Each CPC drawer contains a 2.8 GB shared-victim virtual L4, consisting of all virtual L3 caches of the DCMs in the CPC drawer.

This on-chip cache implementation optimizes system performance for high-frequency processors and includes the following features:

- Cache improvements
- New translation and TLB2 design
- Pipeline optimizations
- Better branch prediction
- New accelerators and architectures
- Secure Execution support

The IBM z17 ME1 cache structure is shown Figure 2-5.

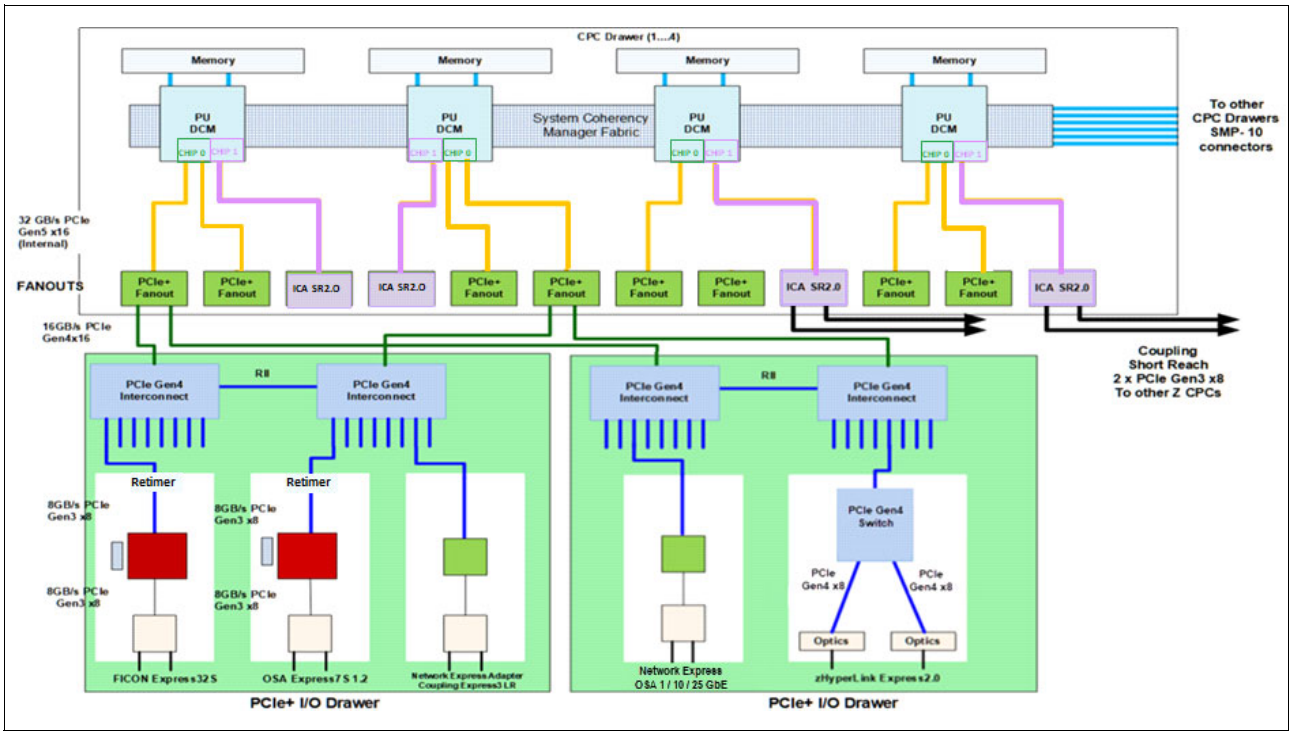

Figure 2-5 IBM z17 ME1 cache structure

[1] There are ten L2s in the pool (one per PU), plus one extra L2 to serve the DPU and one as part of the PU L2 pool.

PU sparing

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

On-core cryptographic hardware

Dedicated cryptographic hardware for each PU core includes extended key and hash sizes for AES and Secure Hash Algorithm (SHA). For more information, see 4.7, “Cryptography” on page 49. This cryptographic hardware is available with any processor type.

On-chip functions

Key on-chip dedicated coprocessors and accelerators include the following ones:

- IBM Integrated Accelerator for Artificial Intelligence is designed for high-speed, real-time inferencing at scale.
- DPU helps improve I/O channel throughput along with improved I/O density.
- Central Processor Assist for Cryptographic Functions (CPACF) encrypts large amounts of data in real time.
- IBM Integrated Accelerator for IBM Z Sort uses the sort instruction (**SORTL**) to accelerate the sorting of data.
- IBM Integrated Accelerator for IBM zEnterprise Data Compression (zEDC) is for data compression as defined by RFC195.

Software support

IBM z17 ME1 PUs help ensure complete compatibility with the z/Architecture software while expanding the ISA. These enhancements unlock improved functions and enhanced performance, introducing new Artificial Intelligence Unit (AIU) instructions on the IBM z17 ME1.

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 *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 with one of the following characterizations:

- Central processor (CP): These standard processors are used for general workloads.
- Integrated Facility for Linux (IFL): Designates processors to be used specifically for Linux natively or on top of a hypervisor like IBM z/VM® or Kernel-based Virtual Machine (KVM).
- Unassigned Integrated Facilities for Linux: Enables you to directly purchase an IFL feature that is marked as being deactivated upon installation, which avoids software charges until the IFL is brought online for use.
- zIIP: An “off-load processor” that is used under IBM z/OS® for designated types of workloads. For a list of zIIP use candidates, see “Logical processors” on page 59. zIIP is also used for the IBM System Recovery Boost (SRB) feature. For more information, see *Introducing IBM Z System Recovery Boost*, REDP-5563.
- Unassigned zIIP: A processor that is purchased for future use as a zIIP. It is offline and cannot be used until an upgrade for the zIIP is installed. It does not affect software licenses or maintenance charges.

- ICF: Designates processors to be used specifically for Coupling Facility (CF) partitions. CF partitions can run on shared or dedicated ICF processors, or on shared CPs. On IBM z17, the use of dedicated CPs for CF partitions was eliminated.
- Unassigned Internal Coupling Facility: Procures a deactivated feature during installation for later use.
- SAP: Designates processors to be used specifically for assisting I/O operations.

Before acquiring zIIPs, at least one CP must be purchased. The maximum number of zIIPs that is available is limited to one fewer than the system's maximum PU configuration; for example, IBM z17 ME1 Max208 can support up to 207 zIIPs. These guidelines apply similarly to unassigned zIIPs and IFL processors.

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

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

2.3.3 Memory

The maximum physical memory size is directly related to the number of central processor complex (CPC) drawers in the system. An IBM Z server includes more installed memory than is ordered because part of the installed memory is used to implement the redundant array of independent memory (RAIM) design. With IBM z17 ME1, up to 16 TB of memory per CPC drawer can be ordered, and up to 64 TB for a four-CPC drawer system.

Important: z/OS requires a minimum of 8 GB of memory (2 GB of memory when running under z/VM®). z/OS can support up to 16 TB of memory in an LPAR.

The minimum and maximum memory sizes for each IBM z17 ME1 feature are listed in Table 2-3.

Table 2-3 IBM z17 ME1 memory per feature

Feature name	CPC drawers	Memory
Max43 (Feature Code 0571)	1	512 GB - 16 TB
Max90 (Feature Code 0572)	2	512 GB - 32 TB
Max136 (Feature Code 0573)	3	512 GB - 48 TB
Max183 (Feature Code 0574)	4	512 GB - 64 TB
Max208 (Feature Code 0575)	4	512 GB - 64 TB

The HSA on IBM z17 ME1 has a fixed amount of memory (884 GB) that is managed separately from customer available memory.

However, the maximum amount of orderable memory can vary from the theoretical number because of dependencies on the memory granularity. On IBM z17 ME1, the granularity for memory is in 64, 128, 256, 512, 1024, and 2048 GB increments.

Physically, the memory is organized in the following ways:

- A CPC drawer always contains a minimum of 1024 GB to a maximum of 16 TB of installed memory.
- A CPC drawer can have more memory that is installed than is enabled. The excess memory can be enabled by a LIC load.
- Memory upgrades are first satisfied by activating installed but unused memory capacity until it is exhausted. When no more unused memory is available from the installed cards, the cards must be upgraded to a higher capacity (replacing existing memory cards or DIMMs in the CPC drawer). Other options include adding extra DIMMs into an existing drawer or installing an extra CPC drawer with more memory (if permitted by the configuration).

On LPAR activation, PR/SM strives to assign PUs and memory within a single CPC drawer. If this allocation proves unfeasible, PR/SM can then use memory resources from any available CPC drawer. For example, if the allocated PUs extend across multiple drawers, PR/SM attempts to allocate memory from those specific drawers, even if sufficient memory exists in only one of them.

No matter which CPC drawer the memory is installed in, an LPAR can access that memory after it is allocated. IBM z17 ME1 is an SMP system because the PUs can access all available memory.

A memory upgrade is considered to be concurrent when it requires no change of the physical memory cards. A memory card change is disruptive when the system does not use Enhanced Drawer Availability (EDA). In a multiple-CPC drawer system, a single CPC drawer can be concurrently removed and reinstalled for a repair with EDA.

For model upgrades that involve adding a CPC drawer, the minimum usable memory increment (512 GB) is added to the system. During an upgrade, adding a CPC drawer and physical memory in the new drawer are concurrent operations.

Concurrent memory upgrade

When unused capacity is available on the installed memory cards, the total system memory can be upgraded concurrently through LICCC. LICCC provides for system upgrades without hardware changes by activating extra (physically installed) unused capacity.

Redundant array of independent memory

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

The RAIM design is fully integrated in IBM z17 ME1 and enhanced to include one memory controller unit (MCU) per processor chip, with eight memory channels and one DIMM per channel. The MCU enables memory to be implemented as RAIM. This technology has 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 z17 (9175) Technical Guide*, SG24-8579.

2.3.4 Hardware System Area

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

Featuring a fixed-size 884 GB HSA, IBM z17 ME1 provides ample capacity to accommodate LPAR definitions or modifications, reducing outage occurrences and needing minimal planning.

A fixed, large HSA enables the dynamic I/O capability of IBM z17 ME1 to be enabled by default. It also enables the dynamic addition and removal of the following features:

- An LPAR to a new or existing CSS
- Devices, up to the maximum number permitted, in each subchannel set
- Logical processors by type
- Cryptographic adapters

2.4 I/O system structure

IBM z17 ME1 supports the PCIe-based infrastructure for the PCIe+ I/O drawers. The PCIe I/O infrastructure consists of the dual-port PCIe fanouts in the CPC drawers that support 32 GBps connectivity to the PCIe+ I/O drawer.

Ordering of I/O features: Ordering I/O feature types determines the number of PCIe+ I/O drawers.

Figure 2-6 shows a high-level view of the I/O system structure for IBM z17 ME1.

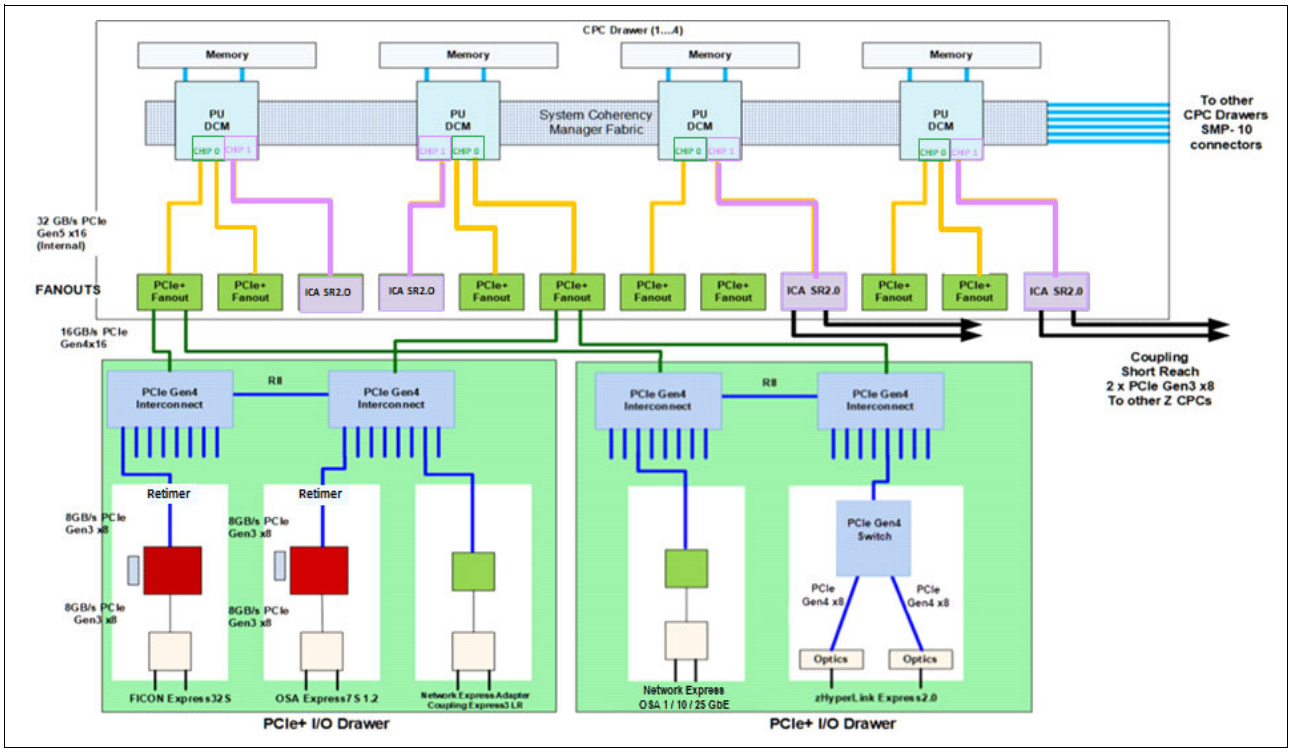

Figure 2-6 IBM z17 ME1 I/O system structure

The IBM z17 ME1 CPC drawer contains 12 fanout positions (LG01 - LG12). These slots can accommodate any of the following components:

- Dual-port PCIe+ fanouts for connecting to a PCIe+ I/O drawer
- ICA SR2.0 fanouts for coupling purposes
- Filler plates to help with airflow cooling

For coupling link connectivity (IBM Parallel Sysplex® and Server Time Protocol (STP) configuration), IBM z17 ME1 supports the following link types:

- ICA SR2.0 (to be installed in a CPC drawer)
- Coupling Express3 Long Reach (CE3 LR) (for installation within a PCIe+ I/O drawer)

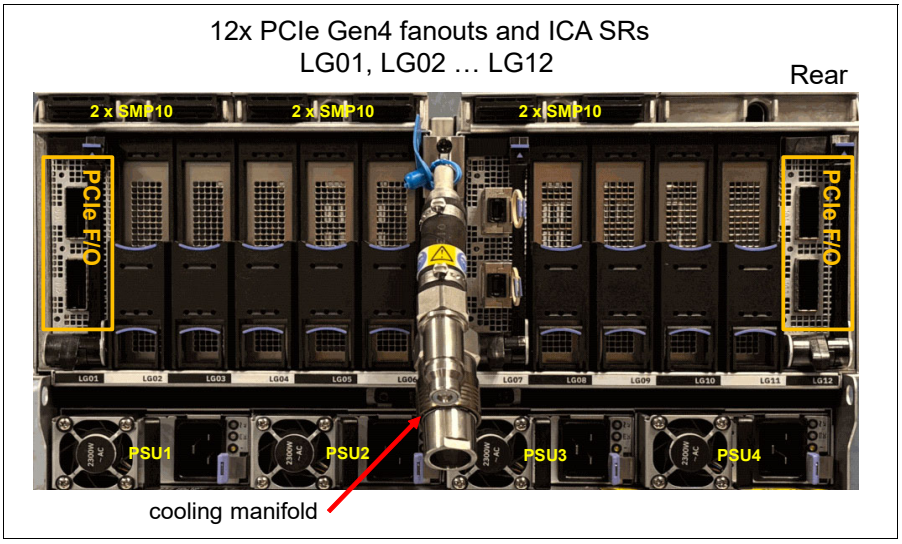

Figure 2-7 IBM z17 ME1 CPC drawer: Rear view

The PCIe+ I/O drawer (see Figure 2-8) 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 adapter slots and two slots for PCIe switch cards.

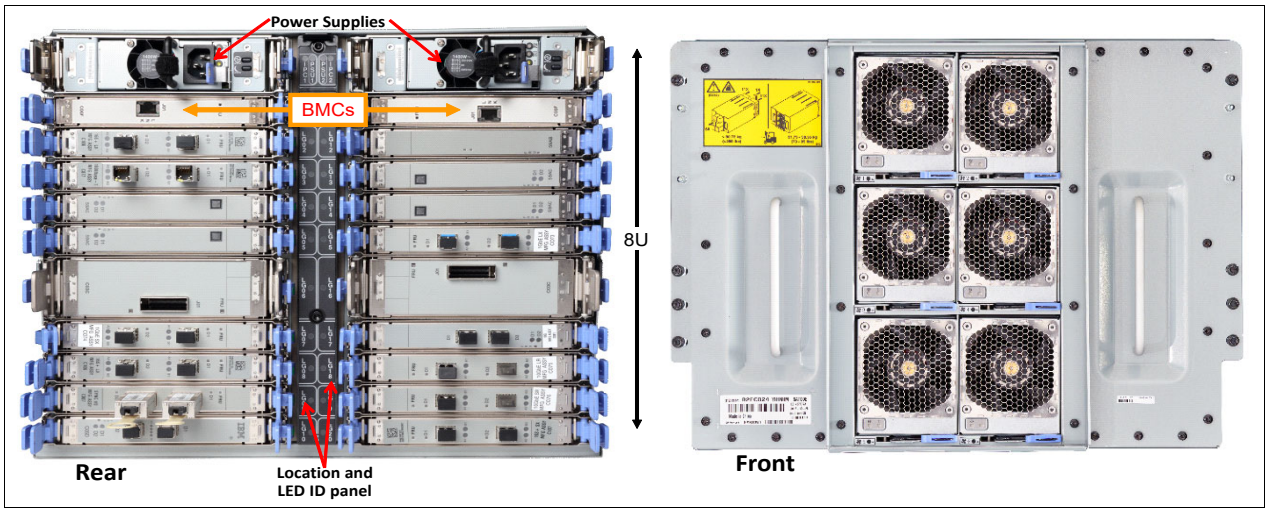

Figure 2-8 PCIe+ I/O drawer: Front and rear views

The two I/O domains per drawer each contain up to eight I/O features that support the following types:

- IBM FICON Express32-4P, IBM FICON Express32S, or IBM FICON Express16SA
- OSA-Express7S 1.2 and OSA-Express7S
- Integrated Network Adapter 100G
- Network Express
- IBM zHyperLink Express 2.0
- CE3 LR
- Crypto-Express8S and Crypto-Express7S
- IBM Spyre Accelerator

For more information about the I/O features that are available with IBM z17 ME1, see Chapter 4, “IBM z17 supported features and functions” on page 41.

2.5 Power and cooling

IBM z17 ME1 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.

2.5.1 Power options

The IBM z17 19-inch frames are available with PDU power (BPA is not supported).

Using PDU for IBM z17 ME1 can enable fewer frames, which enables extra I/O slots and improves power efficiency to reduce overall energy costs. It offers some standardization of data center installation planning. PDU-based systems support up to 12 PCIe+ I/O drawers. The loss of one power supply per set has no effect on system operation.

Depending on the configuration, IBM z17 ME1 requires 2 - 8 PDU connections for full redundancy. Power cords support attachment to either 200 - 208 V 30/60 A 3 Phase (Low Voltage) or 380 - 415 V 32 A 3 Phase Wye (High Voltage).

Consider the following points about power:

- A total of one to four 42U 19-inch IBM frames are used.
- Air flow is front to rear. All blowers are mounted on the front of the frame.
- All external power cabling is at the rear of the frames (no power cabling in front).
- Top or bottom exit power is supported.
- No Emergency Power Off (EPO) switch is used.
- PDUs are different for Low Voltage versus High Voltage (never plug a High-Voltage source to a Low Voltage PDU).

Specific power requirements depend on the number of frames, CPC drawers, and type of I/O features that are installed, which dictate the number of PDUs that are necessary.

For more information about the maximum power consumption tables for the various configurations and environments, see *IBM 9175 Installation Manual for Physical Planning*, GC28-7049.

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

2.5.2 Cooling options

The IBM z17 ME1 cooling system is available only with the radiator (air) cooling option. The previous customer water-cooling solution that was available with IBM z15 is not offered with IBM z17.

DCMs are always cooled with an internal (closed) coolant loop. The liquid in the internal system is cooled by using an internal radiator. The radiator, PCIe+ I/O drawers, power enclosures, and CPC drawers are cooled by chilled air with blowers.

The air-cooling system in IBM z17 ME1 is redesigned for better availability and reduced cooling power consumption. The radiator design is a closed-loop coolant pump system for the DCMs in the CPC drawers. It is designed with N+1 pumps, reservoir, blowers, controls, and sensors. The radiator unit is cooled by air.

The cooling loop liquid for IBM z17 ME1 is a propylene glycol + water solution instead of the previously used water solution. Systems ship prefilled and sealed from manufacturing. The previous Fill Drain Tool and BTA water jugs are eliminated.

3

IBM z17 ME2 and IBM z17 MER hardware overview

The IBM z17 hardware architecture as the reference model was established in Chapter 2, “IBM z17 ME1 hardware overview” on page 11. This chapter extends that topic by introducing two compact deployment options:

- IBM z17 ME2, which is delivered in an IBM factory-configured 19-inch frame
- IBM z17 MER, which is designed for installation in a client-supplied rack with power infrastructure

This chapter describes the following topics:

- 3.1, “IBM z17 ME2 and IBM z17 MER upgrade paths” on page 28
- 3.2, “Frame and cabling” on page 29
- 3.3, “CPC drawers” on page 31
- 3.4, “I/O system structure” on page 37
- 3.5, “Power and cooling” on page 40

© Copyright IBM Corp. 2025, 2026.

3.1 IBM z17 ME2 and IBM z17 MER upgrade paths

IBM z17 ME2 and IBM z17 MER are built by using the IBM Telum II processor design. Each processor chip consists of eight cores. Two processor chips are packaged in the dual-chip module (DCM). Each DCM can have 9 - 11 or 11 - 13 active processor unit (PU) cores (depending on the configuration). With IBM z17 ME2 and IBM z17 MER, a maximum number of characterizable processors are represented by feature names Max6, Max18, Max39, and Max82. Spare PUs, System Assist Processors (SAPs), and two Integrated Firmware Processors (IFPs) are included in both configurations.

The number of characterizable PUs, SAPs, and spare PUs for the various features are listed in Table 3-1.

Table 3-1 IBM z17 ME2 and IBM z17 MER processor unit configurations

Feature name	Feature Code	Characterizable processor units	Central processors	Standard SAPs	Spares
Max6	0676	1 - 6	0 - 6	2	2
Max18	0677	1 - 18	0 - 6	2	2
Max39	0678	1 - 39	0 - 6	5	2
Max82	0679	0 - 82	0 - 6	10	2

For more information about PU characterization types, see “PU characterization” on page 35.

IBM z17 ME2 and IBM z17 MER help ensure continuity and upgradeability from IBM z16 A02 and IBM z16 AGZ. The supported upgrade paths are shown in Table 3-2.

Table 3-2 Upgrade paths

	To an IBM z17 ME2	To an IBM z17 MER
From an IBM z16 A02	Yes	Yes
From an IBM z16 AGZ	Yes	Yes

If an upgrade request cannot be fulfilled by using the existing configuration, a hardware upgrade is required in which one extra central processor complex (CPC) drawer is added to accommodate the needed capacity. With IBM z17 ME2, the 2-CPC drawer, feature Max82, is a factory-only model. Also, a single CPC drawer Max6 can be concurrently upgraded to a Max18.

With IBM z17 ME2 and IBM z17 MER, concurrent upgrades are available for central processors (CPs), Integrated Facilities for Linux (IFLs), Internal Coupling Facilities (ICFs), zIIP processors, and SAPs. However, concurrent PU upgrades require that more PUs are physically installed but not activated previously.

If there is a PU failure, one of the spare PUs is immediately and transparently activated and assigned the characteristics of the failing PU. Two spare PUs are always available on IBM z17 ME2 and IBM z17 MER.

IBM z17 ME2 and IBM z17 MER offer 156 subcapacity levels, A01 - Z06, which use 1 - 6 available CPs with features Max18, Max39, and Max82. For a Max6 feature, the subcapacity models range from A01 - Z06 (six CPs maximum).

For more information, see 3.3.1, “Dual-chip modules” on page 33.

3.2 Frame and cabling

IBM z17 ME2 is delivered in a 19-inch frame with industry-standardized power and hardware. It is a single frame (rack) system that takes up two standard 24-inch floor tiles of space, which aligns with modern data center layouts. IBM z17 MER is delivered as a bundle that is installed in a client-supplied EIA standard 19-inch rack with PDUs. Both systems share features and functions.

The IBM z17 ME2 and IBM z17 MER configuration options, as compared to previous IBM Z servers, are listed in Table 3-3.

Table 3-3 IBM z17 ME2 and IBM z17 MER configuration options compared to IBM z16 and IBM z15

System	Number of frames	Number of CPC drawers	Number of I/O drawers	I/O and power connections	Power options[a]	Power cord options
IBM z17 ME2 IBM z17 MER	1[b]	1 - 2	0 - 3[c]	Rear only	PDU	Single- or three-phase
IBM z16 A02 IBM z16 AGZ	1[d]	1 - 2	0 - 3[e]	Rear only	PDU	Single- or three-phase
IBM z15 T02	1	1 - 2	0 - 4[f]	Rear only	PDU	Single- or three-phase

a. The PDU option is used with different power cords for different frame configurations and countries.
b. There is no frame for IBM z17 MER. A client-supplied rack with PDUs is required.
c. Only PCIe+ I/O drawers are available. New build only. No carry forward of any I/O drawer.
d. There is no frame for IBM z16 AZG. A client-supplied rack with PDUs is required.
e. Only PCIe+ I/O drawers are available. New build only. No carry forward of any I/O drawer.
f. A maximum of four if ordered with single CPC drawer or maximum of three if ordered with two CPC drawers. Only PCIe+ I/O drawers are available.

The number of PCIe+ I/O drawers can vary based on the number of I/O features and number of CPC drawers that are installed. For IBM z17 ME2 and IBM z17 MER PDUs, a maximum configuration of up to three PCIe+ I/O drawers can be installed. PCIe+ I/O drawers can be added concurrently.

In addition, IBM z17 ME2 and IBM z17 MER support top-exit options for the fiber optic and copper cables that are used for I/O and power. These options give you more flexibility in planning where the system is installed, eliminate the need for cables to be run under a raised floor, and increase air flow over the system.

IBM z17 ME2 and IBM z17 MER support installation on a raised floor and a non-raised floor.

Figure 3-1 shows a fully configured IBM z17 ME2 with two CPC drawers and three PCIe+ I/O drawers.

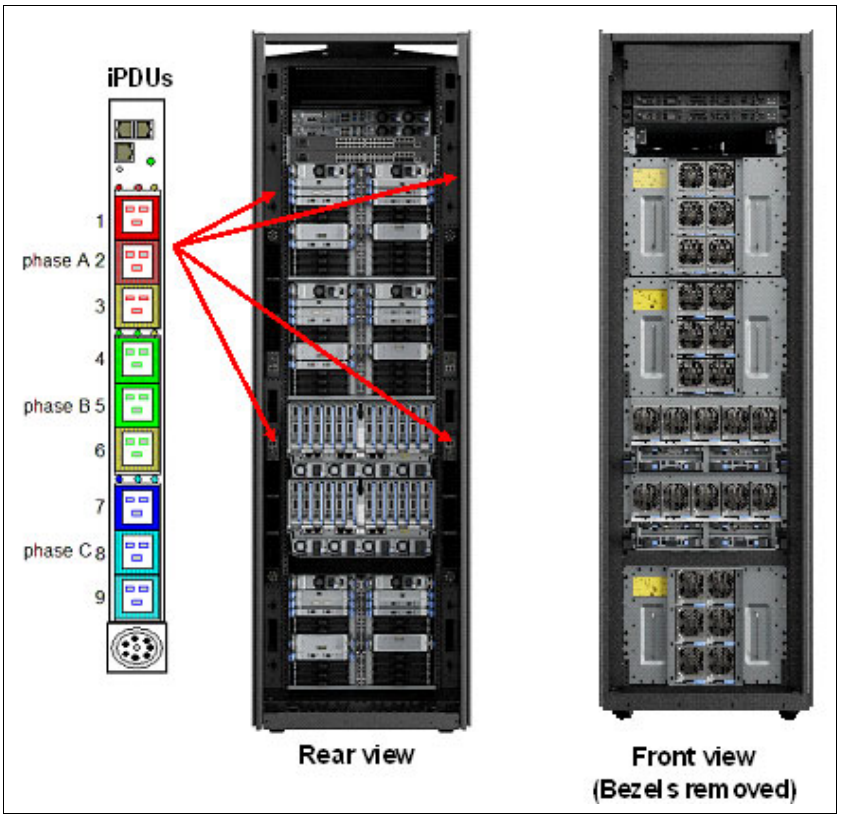

Figure 3-1 Fully configured IBM z17 ME2 with two CPC drawers (Max82) and iPDUs

Figure 3-2 on page 31 shows an IBM z17 MER system with one CPC drawer and one PCIe+ I/O drawer.

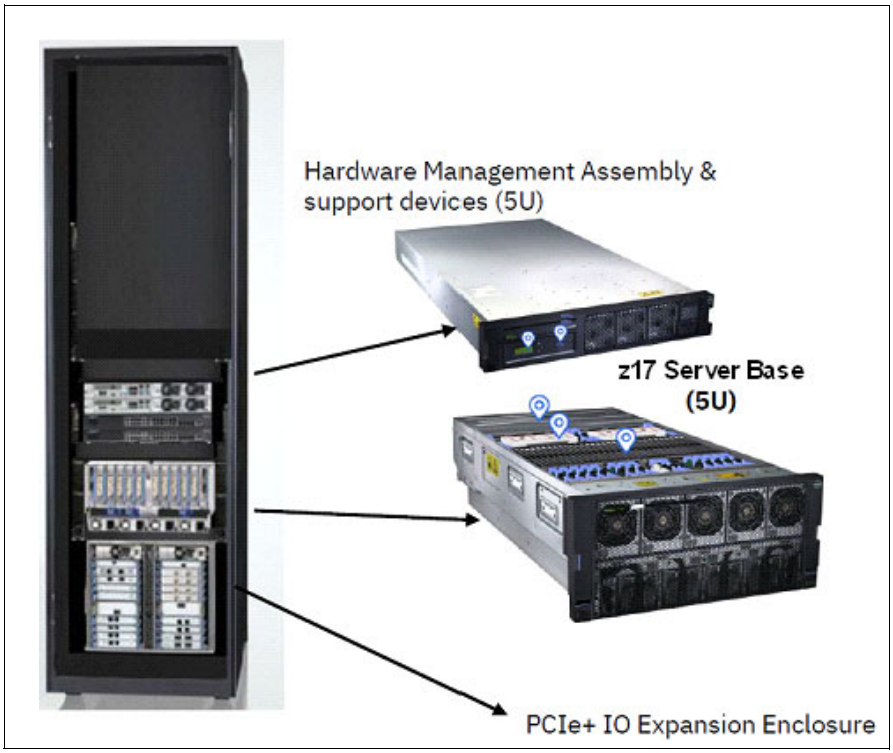

Figure 3-2 An IBM z17 MER system with one CPC drawer and one PCIe+ I/O drawer

The IBM configurator that is used during the order process calculates the placement of CPC and PCIe+ I/O drawers. Factors that determine the configuration of IBM z17 ME2 and IBM z17 MER include the following examples:

- Feature: Max6, Max18, Max39, or Max82
- Number of I/O features (Determines the number of PCIe+ I/O drawers.)
- Single- or three-phase power

3.3 CPC drawers

IBM z17 ME2 and IBM z17 MER can be configured with one or two CPC drawers. Each CPC drawer contains the following elements:

- DCMs
  - Two or four DCMs per drawer (model-dependent and air-cooled).
  - DCM supports two PU chips with an 8-core maximum.
  - Systems are configured by using 9 - 13 core DCMs.
  - 96 PUs maximum on a Max82 with up to 82 characterizable PUs.
- Memory
  - A minimum of 64 GB and a maximum of 17.96 TB of memory per system (excluding 476 GB for the HSA) is available for use. For more information, see Table 3-4 on page 35.
  - Up to 48 DIMMs that are 32, 64, 128, or 256 GB are plugged into a CPC drawer.

- Fanouts

  Each CPC drawer with four DCMs supports up to 12 PCIe+ fanout adapters to connect to the PCIe+ I/O drawers, and Integrated Coupling Adapter Short Reach (ICA SR) coupling links:

  - Two-port PCIe 16 GBps I/O fanout. Each port supports one domain in the 16-slot PCIe+ I/O drawers.
  - ICA SR2.0 PCIe fanouts for coupling links (two links of 8 GBps each).

- Two or four power supply units (PSUs), depending on the CPC drawer configuration, provide power to the CPC drawer and are accessible from the rear. The loss of one PSU leaves enough power to satisfy the power requirements of the entire drawer. The PSUs can be concurrently maintained.
- Two dual-function Baseboard Management Controllers (BMCs) or oscillator cards (OSCs) provide redundant interfaces to the internal management network and provide clock synchronization to the IBM Z platform.
- Two dual-function Processor Power Cards (PPCs) control voltage regulation, and PSU and fan control. The PPCs are redundant and can be concurrently maintained.
- Five fans are installed at the front of the drawer to provide cooling airflow for the resources that are installed in the drawer, including the PU DCMs.

The CPC drawer communication topology is shown in Figure 3-3.

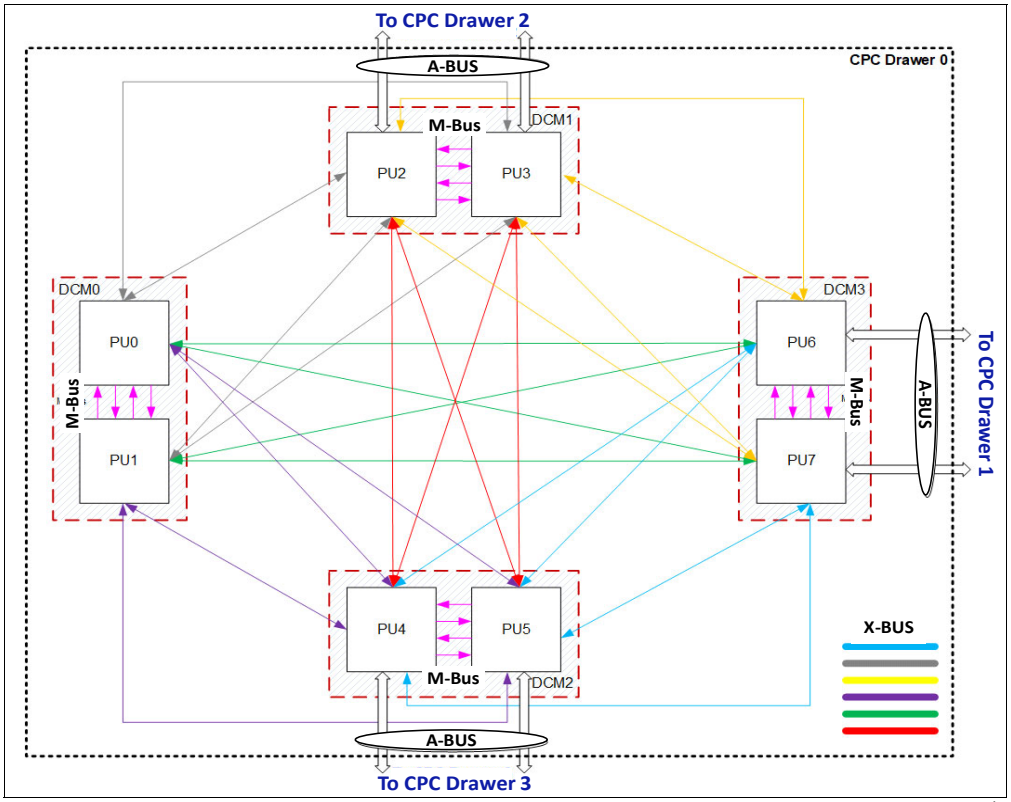

Figure 3-3 IBM z17 ME2 and IBM z17 MER CPC with two drawers communication topology[1]

[1] IBM z17 ME2 and IBM z17 MER Max82 is configured with two CPC drawers and four DCMs per drawer. Max6 and Max18 are configured with one CPC drawer with two DCMs per drawer. Max39 is configured with one CPC drawer with four DCMs.

CPC drawers are interconnected with high-speed communications links (A-Bus) through the PU chips. SMP-9 cables interconnect the CPC drawers. The X-Bus provides connectivity between DCMs within the drawer, and the M-Bus connects the two PU chips on each DCM.

The design that is used to connect the PU and storage control enables the system to be operated and controlled by the IBM Processor Resource/Systems Manager (PR/SM) facility as a memory-coherent symmetric multiprocessing (SMP) system.

3.3.1 Dual-chip modules

The CPC drawer may include two or four DCMs. Each DCM contains two PU chips. Each PU chip contains eight cores each with 128 KB instruction and data L1 cache and 36 MB semi-private L2 caches.

3.3.2 Processor unit

Processor unit 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 10 instructions can be running 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 running at any moment, and are subject to the maximum number of decodes and completions per cycle.

Processor cache structure

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 semi-private L2 cache, which is implemented as 36 MB, near the core.

Note: L1 and L2 are physical caches and are implemented in dense static random access memory (SRAM).

- Each PU core contains a 360 MB shared-victim virtual L3 cache. The shared-victim virtual L3 cache is a logical construction that comprises all eight semi-private L2s (10 x 36 MB = 360 MB) belonging to the other cores.
- The IBM z17 ME2 and IBM z17 MER Max39 and Max82 CPC drawers contain four DCMs and a maximum of 2.88 GB shared-victim virtual L4, consisting of the “remote” virtual L3 caches of the DCMs in the CPC drawer. Max6 and Max18 contain two DCMs per CPC drawer and a maximum of 1.4 GB shared-victim virtual L4.

This on-chip cache implementation optimizes system performance for high-frequency processors and includes the following features:

- Cache improvements
- New translation and TLB2 design
- Pipeline optimizations
- Better branch prediction
- New accelerators and architectures
- Secure Execution support

The IBM z17 ME2 and IBM z17 MER cache structure is shown Figure 3-4.[2]

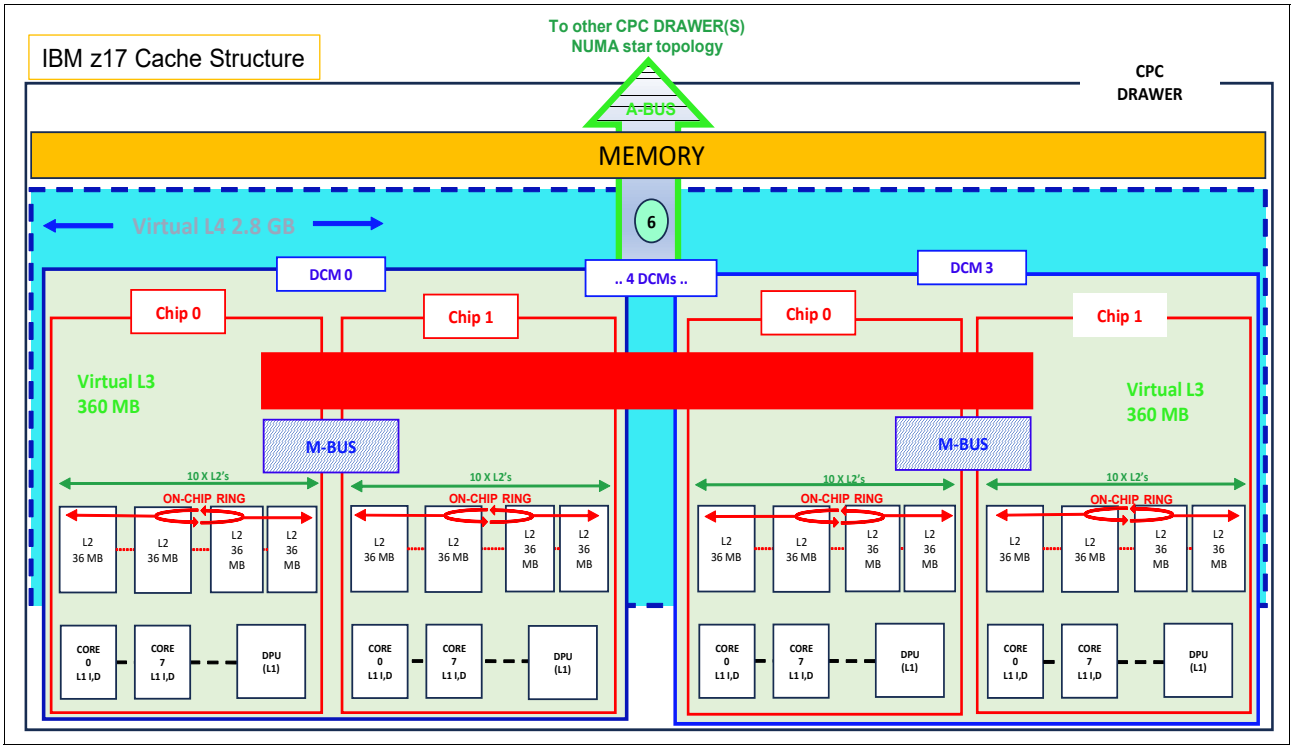

Figure 3-4 IBM z17 ME2 and IBM z17 MER cache structure

PU sparing

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

On-core cryptographic hardware

Dedicated on-chip cryptographic hardware for each PU core includes extended key and hash sizes for AES and Secure Hash Algorithm (SHA). For more information, see 4.7, “Cryptography” on page 49. This cryptographic hardware is available with any processor type, for example, CP, zIIP, and IFL.

[2] Figure 3-4 represents an IBM z17 ME2 and IBM z17 MER Max82 with two drawers and four DCMs per CPC drawer.

On-chip functions

Consider the following points:

- IBM Integrated Accelerator for IBM zEnterprise Data Compression (zEDC) replaces the IBM zEnterprise Data Compression Express PCIe feature that was on previous IBM Z servers.
- The sort accelerator uses the sort instruction (**SORTL**) to be used by **DFSORT** and the IBM Db2® Utilities for z/OS Suite to help reduce CPU usage and improve elapsed time for sorting workloads.
- IBM Integrated Accelerator for Artificial Intelligence Unit (AIU) is implemented on each PU chip and shared among all cores. It provides a matrix array for multiplication and convolution alongside the specialty engines for complex functions.

  The AIU provides a Neural Network Processing Assist (NNPA) instruction, which operates directly on Tensor data in the user space.

Software support

The IBM z17 ME2 and IBM z17 MER PUs provides full compatibility with software for z/Architecture, and extend the ISA to enhance functions and performance. New with IBM z17 ME2 and IBM z17 MER are instructions for AIU.

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 *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. For a description of the designated characterizations, see “PU characterization” on page 19.

3.3.3 Memory

The maximum physical memory size is directly related to the number of CPC drawers that are in the system. An IBM Z server includes more installed memory than is ordered because part of the installed memory is used to implement the redundant array of independent memory (RAIM) design. With IBM z17 ME2 and IBM z17 MER, up to 8 TB of memory per CPC drawer can be ordered and up to 16 TB for a two-CPC drawer system.

Important: IBM z/OS requires a minimum of 8 GB of memory (2 GB of memory when running under IBM z/VM). z/OS can support up to 16 TB of memory in an LPAR.

The minimum and maximum memory sizes for each IBM z17 ME2 and IBM z17 MER feature are listed in Table 3-4.

Table 3-4 IBM z17 ME2 and IBM z17 MER memory per feature

Feature name	CPC drawers	Memory
Max6 (Feature Code 0676)	1	64 GB - 3.62 TB
Max18 (Feature Code 0677)	1	64 GB - 3.62 TB
Max39 (Feature Code 0678)	1	64 GB - 8.74 TB
Max82 (Feature Code 0679)	2	64 GB - 17.96 TB

The HSA on IBM z17 ME2 and IBM z17 MER has a fixed amount of memory (476 GB) that is managed separately from available memory. However, the maximum amount of orderable memory can vary from the theoretical number because of dependencies on the memory granularity. On IBM z17 ME2 and IBM z17 MER, the granularity for memory is in 64, 128, 256, and 512 GB increments. Physically, memory is organized in the following ways:

- A CPC drawer always contains a minimum of 320 GB to a maximum of 8.74 TB of installed memory.
- A CPC drawer can have more memory that is installed than is enabled. The excess memory can be enabled by a LIC 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, the cards must be upgraded to a higher capacity, or a CPC drawer with more memory must be installed.

When an LPAR is activated, PR/SM attempts to allocate PUs and the memory of the LPAR in a single CPC drawer. However, if this allocation is not possible, PR/SM uses memory resources in any CPC drawer.[3] For example, if the allocated PUs span to the additional CPC drawer, PR/SM attempts to allocate memory across that same set of CPC drawers (even if all required memory is available in a single CPC drawer).

No matter which CPC drawer the memory is installed in, an LPAR can access it after it is allocated. IBM z17 ME2 and IBM z17 MER are SMP systems because the PUs have access to all the available memory.

Enhanced Drawer Availability

A memory upgrade is considered concurrent when it requires no change of the physical memory cards. A memory card change is disruptive when Enhanced Drawer Availability (EDA) is not used. In IBM z17 ME2 and IBM z17 MER systems with two CPC drawers, a single CPC drawer can be concurrently removed and reinstalled for a memory upgrade or a repair with EDA.

For upgrades that involve the addition of a CPC drawer, the minimum usable memory increment (256 GB) is added to the system. During an upgrade, adding a second CPC drawer and physical memory in the new drawer are concurrent operations.

Concurrent memory upgrade

If physical memory is available, you can upgrade memory concurrently by using LICCC.

Redundant array of independent memory

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

The RAIM design is fully integrated in IBM z17 ME2 and IBM z17 MER and enhanced to include one memory controller unit (MCU) per processor chip, with eight memory channels and one DIMM per channel. The MCU enables memory to be implemented as RAIM. This technology has RAS capabilities in the area of error detection and correction. Bit, lane, DRAM, DIMM, socket, and complete memory channel failures (including many types of multiple failures) can be detected and corrected.

[3] In IBM z17 ME2 or IBM z17 MER, the memory resources can be allocated in the second CPC drawer if it is available.

3.3.4 Hardware System Area

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

The fixed-size 476 GB HSA of IBM z17 ME2 and IBM z17 MER 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 enables the dynamic I/O capability of IBM z17 ME2 and IBM z17 MER to be enabled by default. It also enables the dynamic addition and removal of the following features:

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

3.4 I/O system structure

IBM z17 ME2 and IBM z17 MER support the PCIe-based infrastructure for the PCIe+ I/O drawers. The PCIe I/O infrastructure consists of dual-port PCIe fanouts in the CPC drawers that support 32 GBps connectivity to the PCIe+ I/O drawer.

Ordering of I/O features: The number and type of features determine the number of I/O drawers that are needed.

Figure 3-5 shows a high-level view of the I/O system structure for IBM z17 ME2 and IBM z17 MER.

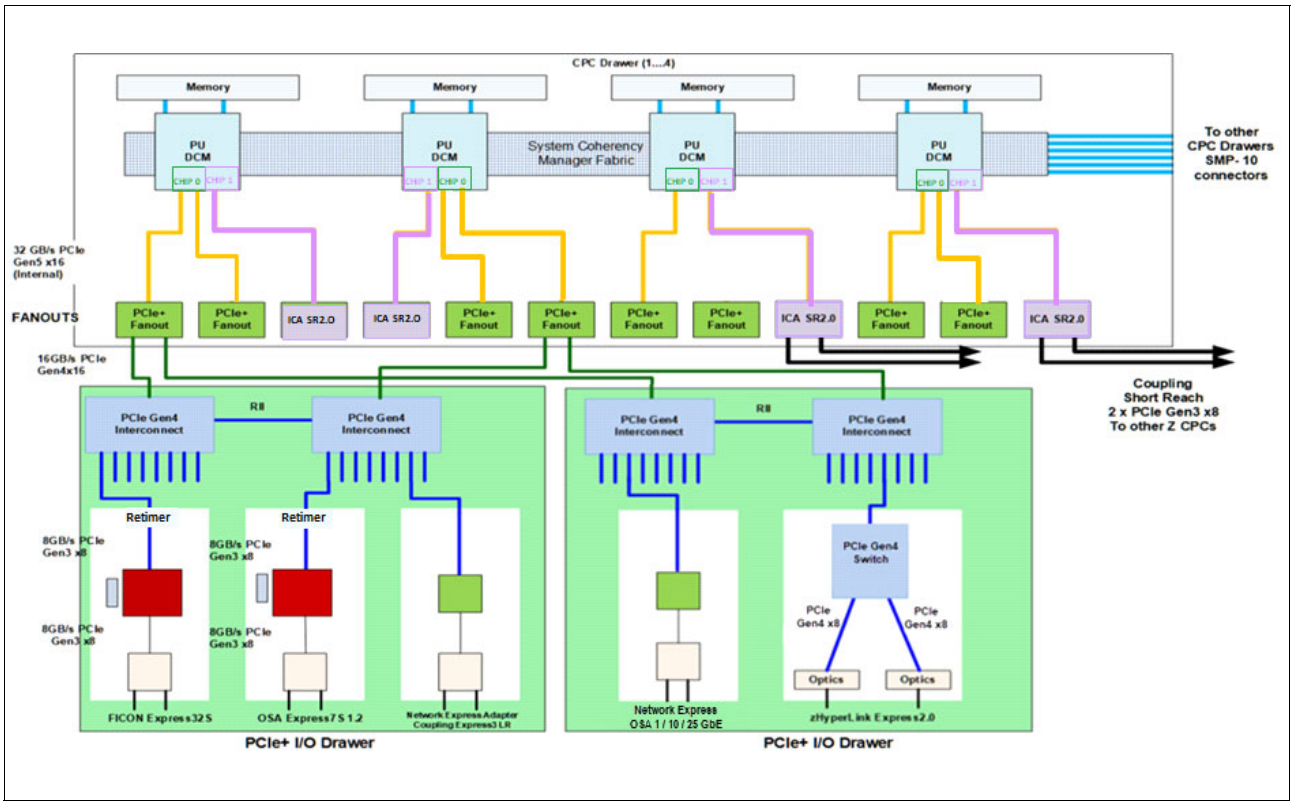

Figure 3-5 IBM z17 ME2 and IBM z17 MER I/O system structure

The IBM z17 ME2 and IBM z17 MER Max6 and Max18 CPC drawers have six fanouts, and the Max39 and Max82 have 12 fanouts per CPC drawer. Fanout locations are numbered LG01 - LG12 and can be one of the following types:

- Dual-port PCIe+ fanouts for PCIe+ I/O drawer connectivity
- ICA SR fanouts for coupling
- Filler plates to help with airflow cooling

For coupling link connectivity (IBM Parallel Sysplex and Server Time Protocol (STP) configuration), IBM z17 ME2 and IBM z17 MER support the following link types:

- ICA SR2.0 (installed in a CPC drawer)
- Coupling Express3 Long Reach (CE3 LR) (installed in a PCIe+ I/O drawer)

For systems with two CPC drawers, the locations of the PCIe+ fanouts are configured and plugged across all drawers for maximum availability. This configuration helps ensure that alternative paths maintain access to critical I/O devices, such as storage and networks (see Figure 3-6 on page 39).

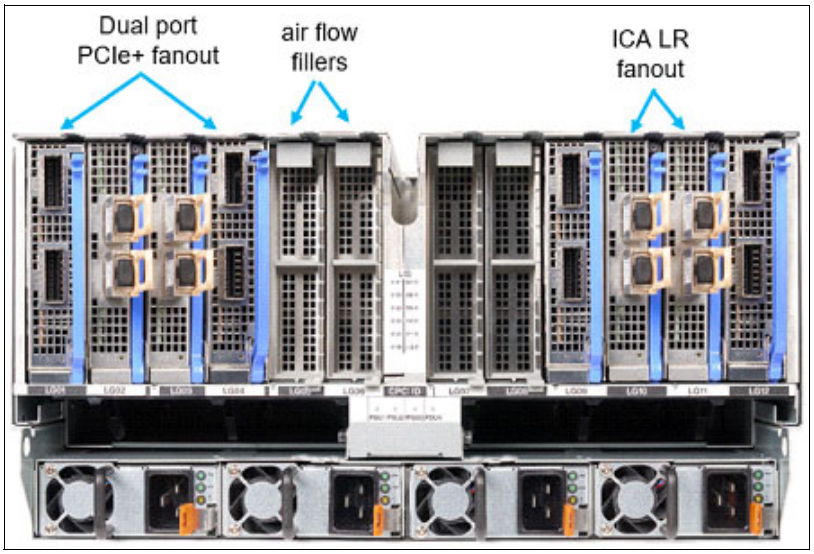

Figure 3-6 CPC drawer: Rear view

The PCIe+ I/O drawer (see Figure 3-7) 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 adapter slots and 2 slots for PCIe switch cards.

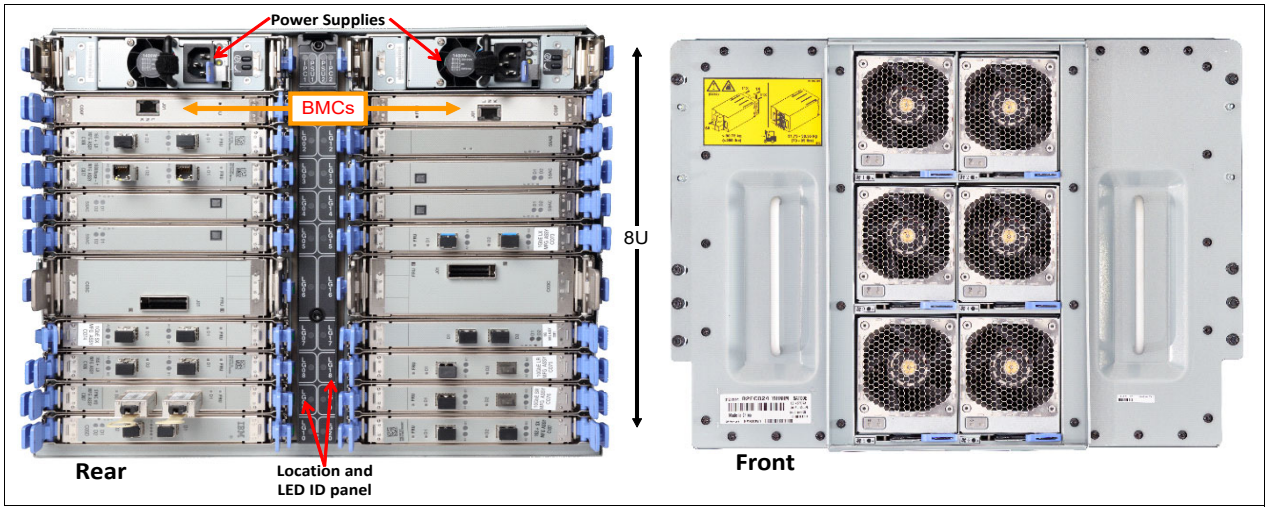

Figure 3-7 PCIe+ I/O drawer: Front and rear views

The two I/O domains per drawer each contain up to eight I/O features that support the following types:

- IBM FICON Express32-4P, IBM FICON Express32S, or IBM FICON Express16SA
- OSA-Express7S 1.2 and OSA-Express7S
- Integrated Network Adapter 100G
- Network Express
- IBM zHyperLink Express 2.0
- CE3 LR
- Crypto-Express8S and Crypto-Express7S
- IBM Spyre Accelerator

For more information about the I/O features that are available with IBM z17 ME2 and IBM z17 MER, see Chapter 4, “IBM z17 supported features and functions” on page 41.

3.5 Power and cooling

IBM z17 ME2 and IBM z17 MER meet 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.

3.5.1 Power options

The IBM z17 ME2 19-inch frame is available with the following power options:

- Single-phase PDU

  Single-phase power for IBM z17 ME2 can be used for systems with a single CPC drawer. A system with a single CPC drawer supports up to three PCIe+ I/O drawers.

- Three-phase PDU

  Both single- and dual-CPC drawer systems support three-phase power. However, systems with two CPC drawers support only three-phase power and only up to three PCIe+ I/O drawers. IBM z17 ME2 does not support plan-ahead or balanced power options.

IBM z17 ME2 operates with two or four sets of redundant PDUs. Each set has its own individual power cords or pair of power cords, depending on the configuration. Power cords attach to a three-phase, 50/60 Hz, 200 - 480 V AC power source or to a single-phase 50/60 Hz, 200 - 240 V AC power source. The loss of only one power supply per set has no effect on system operation.

3.5.2 Cooling options

The IBM z17 ME2 and IBM z17 MER cooling system is available only with one cooling option: air cooling. DCMs are cooled with forced air. PCIe+ I/O drawers and CPC drawers are cooled by chilled air with blowers.

3.5.3 Power considerations

Consider the following points about power:

- The 42U 19-inch IBM frames are used.
- Air flow is front to rear. All blowers are mounted on the front of the frame.
- All external power cabling is at the rear of the frames (no power cabling in front).
- Top or bottom exit power is supported.
- A High-Voltage DC option is not available.
- No Emergency Power Off (EPO) switch is used.

Specific power requirements depend on the number of CPC drawers, and the number and type of I/O features that are installed.

For more information about the maximum power consumption tables for the various configurations and environments, see *9176 Installation Manual for Physical Planning (Models MER/MLR)*, GC28-7075-00, and *9176 Installation Manual for Physical Planning (Models ME2/ML2)*, GC28-7080-00.

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

4

IBM z17 supported features and functions

The features that are described in this chapter apply across IBM z17 systems, with specific availability and scale depending on the system configuration (IBM z17 ME1, IBM z17 ME2, or IBM z17 MER).

The information that is provided expands on the key hardware elements that are introduced in Chapter 2, “IBM z17 ME1 hardware overview” on page 11 and Chapter 3, “IBM z17 ME2 and IBM z17 MER hardware overview” on page 27.

This chapter describes the following topics:

- 4.1, “IBM z17 I/O connectivity overview” on page 42
- 4.2, “Introducing the IBM z17 I/O design” on page 43
- 4.3, “Storage connectivity” on page 43
- 4.4, “Network connectivity” on page 44
- 4.5, “Clustering connectivity” on page 46
- 4.6, “IBM Z Coordinated Server Time” on page 47
- 4.7, “Cryptography” on page 49
- 4.8, “Special-purpose features and functions” on page 53
- 4.9, “Hardware management” on page 53

© Copyright IBM Corp. 2025, 2026.

4.1 IBM z17 I/O connectivity overview

IBM z17 provides a PCIe-based infrastructure for the PCIe+ I/O drawers to support the following features:

- Storage connectivity
  - IBM zHyperLink Express2.0 (new build only)
  - IBM FICON Express32-4P (new build only)
  - IBM FICON Express32S (carry forward only)
  - IBM FICON Express16SA (carry forward only)
- Network connectivity
  - Integrated Network Adapter 100G (new build only)
  - Network Express 25G (new build only)
  - Network Express 10G (new build only)
  - Network Express 1G (new build only)
  - OSA-Express7S 1.2 (new build or carry forward)
  - OSA Express 1000BASE-T 7S (carry forward from IBM z15 only)
- Clustering connectivity
  - Integrated Coupling Adapter2.0 (ICA SR2.0) (new build only)
  - Two types of Coupling Express3 Long Reach (CE3 LR) (new build only):
    - Coupling Express3 Long Reach 25G
    - Coupling Express3 Long Reach 10G
- Cryptographic features
  - Crypto Express8S, one or two Hardware Security Modules (HSMs) (new build or carry forward)
  - Crypto Express7S, one or two HSMs (carry forward only)

For more information about these features, see *IBM Z Functional Matrix*, REDP-5157 and *IBM Z Connectivity Handbook*, SG24-5444.

The following features that were supported on earlier IBM Z platforms are *not* orderable and *cannot* be carried forward to IBM z17:

- zHyperLink Express1.1 and zHyperLink Express
- IBM FICON Express16S+, IBM FICON Express16S, and IBM FICON Express8S
- OSA Express6S and OSA-Express5S
- RoCE Express3, RoCE Express2.1, and RoCE Express2
- Crypto Express6S and Crypto Express5S
- IBM zEnterprise Data Compression (zEDC)
- Coupling Express2 Long Reach (CE2 LR) and Coupling Express Long Reach (CE LR)
- ICA SR1.1 and ICA SR

Note: The LC Duplex connector type is used for all fiber optic cables, except the cables that are used for zHyperLink Express, and ICA SR2.0 connections, which have Multi-Fiber Termination Push-On (MTP) connectors.

4.2 Introducing the IBM z17 I/O design

The IBM Z platform is designed to meet the growing demands of modern businesses by providing advanced features and functions that enhance I/O throughput and connectivity. The IBM Z I/O stack for storage and network connectivity has advanced on application-specific integrated circuit (ASIC) chips for over 30 years. With IBM z17, a novel I/O design featuring a Data Processing Unit (DPU) is introduced to enhance performance and boost I/O rates for Fibre Channel Protocol (FCP), FICON, High Performance FICON, and OSA protocols. The DPU is an on-chip processor that integrates I/O acceleration directly onto the IBM Telum II processor.

The DPU consists of 32 cores that are organized into four clusters of eight, running firmware that uses a standard ISA tailored to address IBM Z-specific I/O requirements. A PCIe Bridge Unit offers system memory access, previously mediated by data routers in ASIC chips for prior IBM Z I/O features.

The architecture integrates extra offload or accelerator engines for both I/O and non-I/O functions alongside enhanced infrastructure management capabilities.

The IBM z17 also doubles port density for storage and network connectivity features.

Note: Storage and network connectivity features on IBM z17, excluding the ones that use the DPU, continue to use the ASIC chip design and I/O operations similar to IBM z16 and earlier systems.

4.3 Storage connectivity

Storage connectivity is achieved through various methods, including direct storage attachment to SANs. The main focus for storage connectivity with IBM Z is to continuously improve the latency for I/O transmission. The new I/O design, including the DPU and the FICON Express32-4P (4-port feature), embodies that goal, as does zHyperLink Express.

4.3.1 FICON Express features

FICON Express features in IBM z17 evolve, offering enhanced throughput and RAS. In IBM z17, these features provide connectivity to other systems, such as Fibre Channel (FC) switches and various devices in a SAN environment.

FICON Express features follow the established 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:

- FICON

  This enhanced protocol provides communication across SAN Fabric switches, CTC connectivity, and FICON devices, such as disks, tapes, and printers.

- High Performance FICON

  The FICON Channel Enhancement boosts FICON channel performance by shifting I/O management tasks from the channel subsystem (CSS) to the controller and host adapter, optimizing system efficiency.

- FCP

  This standard protocol is used for communicating with disk and tape devices through FC switches and directors. The FCP channel can connect to FCP SAN fabrics and access FCP and SCSI devices.

The FICON Express32-4P feature uses the DPU to move data to and from memory and storage devices. It supports the same protocols as its predecessor (FICON Express32S), but with higher port density (four versus two). Both features support a link data rate up to 32 Gbps and can auto-negotiate to 8, 16, or 32 Gbps.

The FICON Express32S implements PCIe cards, providing enhanced port granularity and superior capabilities compared to preceding FICON Express features.

4.3.2 zHyperLink Express

zHyperLink was created to provide fast access to data through low-latency connections between the IBM Z platform and storage.

With the zHyperLink Express2.0 feature, you can make synchronous requests for data that is in the storage cache of the IBM DS8000®. This process is done by directly connecting the zHyperLink Express2.0 port in IBM z17 to an I/O bay port of the DS8000. This short-distance (up to 150 m (492 feet)), direct connection is designed for low-latency read/write operations, such as with IBM Db2 for z/OS synchronous I/O reads and log writes.

zHyperLink can process more than 4 K data transfers in a synchronous fashion, which is new for IBM z17, enabling more software usage that was not possible with only data transfers of 4 K or less.

Working with the FICON SAN Infrastructure, zHyperLink improves application response time, which reduces I/O-sensitive workload response time by half without requiring application changes.

Note: The zHyperLink channels complement FICON channels, but they do *not* replace FICON channels. FICON remains the main data driver and is mandatory for zHyperLink usage.

4.4 Network connectivity

Network connectivity covers the connections between the platform and external networks with Integrated Network Adapter 100G (2-port feature), Network Express (2-port features), and Open Systems Adapter-Express (OSA-Express) features. Intra-system communication is provided through IBM HiperSockets and Internal Shared Memory (ISM).

The new I/O design for network connectivity, including the DPU and the Network Express features, improves data transmission latency.

Note: OSA-Express7S features are supported by IBM z17. However, OSA-Express features can be replaced by Network Express features.

IBM z17 does not support RoCE Express features; instead, it uses the Network Express features.

4.4.1 Network features

Integrated Network Adapter 100G, Network Express, and OSA-Express features provide high-performance LAN connectivity while adhering to IEEE standards. These features enhance overall efficiency by offloading selected TCP/IP stack processing functions from the processor unit (PU), reducing CPU usage and improving application performance. For Network Express, these offloaded functions are handled by an integrated DPU.

IBM z17 introduces the Integrated Network Adapter 100G, delivering 100 GbE connectivity through fiber optic MTP cables, supporting both Base-8 (8-fiber) and Base-12 (12-fiber) configurations per port. OSA-Express and Network Express features continue to support both single-mode and multimode fiber environments. With this generation, IBM signals a migration away from copper networking: IBM z17 is the final system to support copper connections, which are limited to the OSA-Express7S 1000BASE-T feature that is carried forward from IBM z15.

The Integrated Network Adapter 100G delivers high-bandwidth, low-latency connectivity by integrating dual 100 GbE ports within the CPC drawer, reducing external I/O impact and streamlining data paths. RDMA support enables data transfers that reduce CPU intervention, cache pollution, and context switching. Physical function-level access through the Network Express Adapter Direct function type enhances resource partitioning and parallelism, optimizing packet processing and throughput under load. Together, these capabilities minimize queuing delays and protocol impact, delivering deterministic performance and scalable efficiency for bandwidth-intensive, latency-sensitive enterprise workloads.

The Network Express feature supports high-speed, memory-to-memory communication between IBM Z systems by using RDMA over Converged Ethernet (RoCE). This capability enables efficient data transfer by bypassing the traditional TCP/IP stack for certain workloads.

Further performance improvements are achieved through Shared Memory Communication over Remote Direct Memory Access (SMC-R), an industry-standard protocol that delivers low-latency, high-throughput data transfer. SMC-R operates transparently to applications, requiring no code changes, which accelerates adoption and time to value.

Collectively, these capabilities reduce CPU consumption for TCP/IP-intensive workloads like IBM WebSphere® accessing Db2 while improving throughput and lowering latency, particularly for large data transfers between systems.

4.4.2 HiperSockets

Referred to as the “network in a box”, HiperSockets simulate LAN environments exclusively within IBM Z servers. It enables direct data transfer between LPAR memories, facilitated by IBM Z firmware.

A component of the LIC, HiperSockets offers LAN connectivity across various system images within an IBM Z platform through secure, memory-to-memory data transfers.

By eliminating I/O subsystem operations and bypassing the need for external network connections between LPARs on an IBM Z platform, HiperSockets supports server consolidation through seamless virtual server interconnectivity and simplified enterprise networking.

4.4.3 ISM communications

ISM, functioning as a virtual PCIe network adapter, facilitates direct access to shared virtual memory and offers an optimized interconnect for intra-communications on the IBM Z platform. It is used by Shared Memory Communication-Direct Memory Access (SMC-D).

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 PCIe bandwidth, ports, I/O slots, network resources, or Ethernet switches). Instead, SMC-D uses LPAR-to-LPAR communication through HiperSockets, an OSA-Express, or a Network Express feature for establishing the initial connection.

IBM z/OS and Linux on IBM Z platforms both support Shared Memory Communication (SMC), enabling memory-to-memory data sharing between the operating systems for enhanced efficiency and performance.

4.5 Clustering connectivity

IBM Parallel Sysplex is a clustering technology enabling multiple z/OS images to operate concurrently, appearing as one system to users. This configuration facilitates near-continuous availability. The Coupling Facility (CF), which supports parallelism, can function as a separate LPAR (internal CF) or within dedicated hardware (external CF).

What makes a group of such z/OS images into a sysplex is inter-communication, which is handled through coupling links. Coupling links enable all the z/OS to CF communication, CF-to-CF traffic, and Server Time Protocol (STP)[1].

For more information about options, see Coupling Facility Configuration Options.

There are three types of features that provide sysplex connectivity between or within IBM z17s, IBM z16s, and IBM z15s servers:

- ICA SR2.0

  Supports distances up to 150 meters. It is used for infra-site connectivity between systems.

- CE3 LR

  IBM z17 introduces CE3 LR 25G, a two-port 25 GbE-based long reach coupling adapter that carries a new coupling channel type: CL6. Also available on IBM z17 is CE3 LR 10G, a two-port 10 GbE-based long reach coupling adapter that carries the coupling channel type CL5.

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

CE3 LR supports longer (unrepeated) connectivity to systems up to 10 kilometers, so it is mainly (not exclusively) used for inter-site connectivity between systems. Longer distances, up to 100 kilometers, are possible with the use of a DWDM extender.

- Internal Coupling Peer

  Facilitates rapid and secure memory-to-memory communication between LPARs on the same system, running CF and z/OS images. This connection is emulated in LIC, eliminating the need for physical cabling.

4.5.1 Dynamic I/O configuration

You can make dynamic I/O configuration changes to stand-alone CF[2], Linux on IBM Z, and IBM z/TPF LPARs without requiring a disruptive POR by using a firmware-based activation service in IBM z17.

The firmware-based LPAR is driven by updates from an HCD instance that runs in a z/OS LPAR on a different IBM Z server that is connected to the same HMC.

A firmware LPAR is within the reserved range that is allocated by IBM, offering no attached I/O and thus consuming no configurable resources. Therefore, it remains invisible to clients.

4.5.2 Quantum-safe protection for Coupling Facility Control Code

Quantum-safe protection covers the Coupling Facility Control Code (CFCC). CFCC boot and update processes are protected with Cryptographic Suite for Algebraic Lattices (Module-Lattice-Based Digital Signature (ML-DSA)), which are standardized by NIST, for quantum-safe digital signatures. For more information about quantum-safe digital signatures, see *Transitioning to Quantum-Safe Cryptography on IBM Z*, SG24-8525.

4.6 IBM Z Coordinated Server Time

IBM Z Coordinated Server Time (CST) helps ensure that all systems within a network are synchronized to the same time. This approach is crucial for accurate logging of events, managing network traffic, running time-sensitive transactions, and coordinating actions based on precise timestamps.

STP employs message transmission over coupling links between IBM Z servers for maintaining synchronized timekeeping. IBM z17 can join an STP Coordinated Timing Network (CTN), which is a group of interconnected IBM Z servers that synchronize their time to CST.

STP is implemented in LIC as a system-wide facility of IBM z17 and is enabled by installing the STP feature. Extra configuration is required for an IBM z17 server to become a member of a CTN.

For HA purposes, a nondisruptive capability is implemented in IBM z17 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[3]. This protocol enables a Parallel Sysplex to span these distances for a multi-site Parallel Sysplex.

[2] A stand-alone CF does not have any running instances of z/OS or IBM z/VM.
[3] With DWDM.

Note: If an IBM z17 server plays a CTN role (Primary Time Server (PTS), Backup Time Server (BTS), or Arbiter), the other CTN IBM Z server must include direct coupling connectivity to IBM z15, z16, or z17.

4.6.1 STP implementation with IBM z17

The following items are the STP implementation characteristics with IBM z17:

- Network Time Protocol (NTP) v4/Chrony improves the overall resiliency of IBM Z NTP implementation for STP. It also improves the accuracy and stability capabilities of IBM Z STP by using the full suite of NTP algorithms by using Chrony.
- The Network Time Security (NTS) protocol for NTP is an approved IETF standard (RFC 8915) that provides cryptographic security for the client/server mode of NTP. NTS uses TLS and Authenticated Encryption with Associated Data to accomplish the following goals:
  - Enable NTP users to obtain time in an authenticated manner.
  - Help ensure that time data that is received is authenticated and secure.
  - Protect against attacks such as spoofing, man-in-the-middle, and replay.
- Support of Mixed Mode of NTP and Precision Time Protocol (PTP) improves the overall STP resiliency capabilities by enabling STP to use up to five External Time Sources (ETSs) (three NTP and two PTP) simultaneously. Mixed mode operation also improves STP accuracy and stability.
- Direct network connection to the IBM z17 CPC drawer of the ETS. ETS for NTP, PTP, and pulse per second (PPS) network cables now connect directly to the IBM z17 central processor complex (CPC) drawers.[4] This connection provides greater accuracy to the external time reference.
- The n-mode Power STP Imminent Disruption Signal option. A loss of the PTS on IBM Z servers has severe implications for the timing network and the entire workload execution environment within an IBM Z sysplex.

4.6.2 Network Time Protocol Client support

Employing NTP servers as an ETS typically meets the need for a time reference across diverse platforms, offering enhanced time accuracy compared to alternative methods.

NTP client support is available in the firmware partition that runs on IBM z17. The code interfaces with the NTP servers. This interaction enables an NTP server to become the single-time source for IBM z17 and for other servers that have NTP clients. The NTP Ethernet cable must plug directly into the Baseboard Management Controller (BMC) or oscillator card (OSC) ports on the IBM z17 CPC drawer. Redundant cabling and ETS must be configured.

4.6.3 Precision Time Protocol client support

IEEE 1588 PTP is implemented on IBM z17 as an ETS for a CTN. The PTP Ethernet cable must plug directly into the BMC or OSC ports on IBM z17. Redundant cabling and ETS should be configured.

[4] The ETS and PPS cables are connected to the BMC or OSC adapters and from the front of the CPC drawer.

4.6.4 Pulse per second support

Two OSCs,[5] which are included as a standard feature of IBM z17, provide a dual-path interface for the PPS signal. The cards contain a Bayonet Neill-Concelman connector for PPS attachment at the front 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 and accurate PPS signal from the external time server. PPS maintains an accuracy of 10 µs as measured at the PPS input of IBM z17.

A cable connection from the PPS port to the PPS output of an NTP server is required when IBM z17 is configured for NTP with PPS as ETS for time synchronization.

PPS is optional for PTP but might still be required for NTP to meet financial regulations regarding the accuracy of time synchronization between servers.

For more information, see *IBM Z Time Synchronization Implementation Guide*, SG24-8480.

4.7 Cryptography

The IBM Z platform offers cryptographic engines that provide high-speed cryptographic operations:

- Central Processor Assist for Cryptographic Functions (CPACF)

  Cryptographic functions are provided through a set of instructions, which are available in the hardware on every PU.

- Crypto Express features

  Cryptographic functions that are provided through high-security, tamper-sensing, and tamper-responding HSMs.

The IBM Z platform provides cryptographic functions that can be categorized in the following groups from an application program perspective:

- Symmetric cryptographic functions,[6] which are provided by the CPACF or Crypto Express features when defined as an accelerator.
- Asymmetric cryptographic functions,[7] which are provided by the Crypto Express features.

Also, cryptographic features and functions are designed to protect IBM z17 from attacks, including threats that might use quantum computers. The system includes the use of quantum-safe algorithms that are selected by NIST through the many firmware layers that are loaded during the boot process. Only authentic, IBM recognized firmware is accepted. This hardware-protected verification of the firmware uses a dual-signature scheme, consisting of a combination of both classical and quantum-safe algorithms that are selected by NIST. The protection is anchored in the IBM Z *Root of Trust*.[8]

For more information about the quantum-safe algorithms that are selected by NIST and used in IBM z17, see *Transitioning to Quantum-Safe Cryptography on IBM Z*, SG24-8525.

[5] The OSCs are combined with the BMCs.
[6] With symmetric encryption, the same cryptographic key is used for encryption and decryption of the data (the sender and receiver of the data use the same key).
[7] With asymmetric encryption, the receiver's public key is used for encryption and the receiver's private key is used for decryption.
[8] Root of Trust is a source that can always be trusted within a cryptographic system.

4.7.1 Central Processor Assist for Cryptographic Functions

Each PU chip in IBM z17 has an independent cryptographic engine (known as a *cryptographic assist*). The no-charge CPACF enablement feature offers a set of cryptographic functions for high-performance encryption and decryption with clear key operations for SSL and TLS, VPN, and data-storing applications that do not require Federal Information Processing Standards (FIPS) 140-2 Level 4 security[9].

The CPACF-protected key is a function that facilitates the continued privacy of cryptographic key material while maintaining the wanted high performance. CPACF helps ensure that key material is not visible to applications or operating systems during encryption operations. A CPACF-protected key provides substantial throughput improvements for large-volume data encryption and reduced latency for encryption of small blocks of data.

The cryptographic assist includes support for the following functions:

- AES for 128-bit, 192-bit, and 256-bit keys
- Improved performance of AES Galois/Counter Mode encryption
- DES data encryption and decryption with single, double, or triple length keys
- Pseudo-Random Number Generator (PRNG)
- Deterministic Random Number Generator (DRNG)
- True Random Number Generator (TRNG)
- Message authentication code
- Elliptic Curve Cryptography (ECC) support[10]
- Hashing algorithms: Secure Hash Algorithm (SHA)-1, SHA-2, and SHA-3, and Hash-based Message Authentication Code

IBM Z servers inherently support SHA-1, SHA-2, and SHA-3 without the CPACF feature requirement. The CPACF functions are supported by z/OS, z/VM, 21 Century Software VSE^n V6.3, z/TPF, and Linux on IBM Z.

Attention: Cryptographic methods like DES, along with hashing algorithms such as SHA-1, are deemed inadequate for current cybersecurity standards due to their vulnerability against contemporary attacks.

This risk can be mitigated by switching to stronger algorithms, such as AES, SHA-256, SHA-3, ML-DSA, and Module-Lattice-Based Key-Encapsulation Mechanism (ML-KEM).

IBM provides several tools that can aid in the management and discovery process:

- IBM z/OS Integrated Cryptographic Service Facility (ICSF)
- IBM Application Discovery and Delivery Intelligence (ADDI)
- IBM Crypto Analytics Tool (CAT)
- IBM z/OS Encryption Readiness Technology (zERT)

These tools can help you identify certificates, encryption protocols, algorithms, and key lengths that are at risk in your IBM Z environment.

[9] FIPS 140-2 Security Requirements for Cryptographic Modules.

[10] Message-Security-Assist extension 9 is a hardware feature that was introduced in IBM z15 that provides the ECC capabilities, including message authentication, key generation, and scalar multiplication.

4.7.2 Crypto Express8S

A Crypto Express8S PCIe cryptographic co-processor is an optional feature that is available on IBM z17. This co-processor is an HSM that provides high-security cryptographic processing as required by banking and other industries.

The Crypto Express8S is designed with APIs that provide access to the quantum-safe algorithms that are selected by NIST along with classical cryptographic algorithms to modernize existing applications and build new ones.

This feature provides a secure programming and hardware environment where crypto-processes are performed. Each cryptographic coprocessor includes general-purpose processors, nonvolatile storage, and specialized cryptographic electronics. All these features are contained within a tamper-sensing and tamper-responsive enclosure that wipes 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.

The Crypto Express8S (2-adapter) feature includes two PCIe Cryptographic Coprocessors (PCIeCCs), and the Crypto Express8S (1-adapter) feature includes one PCIeCC. For availability reasons, a minimum of two features is required for both features. Up to 30 Crypto Express8S (2-adapter) features are supported on IBM z17. The maximum number of the 1-adapter features is 16. The Crypto Express8S feature occupies one I/O slot in a PCIe+ I/O drawer.

The adapters can be configured as a Secure IBM Common Cryptographic Architecture (CCA) coprocessor, a Secure IBM Enterprise PKCS #11 (EP11) coprocessor, or as an accelerator.

Crypto Express8S provides domain support for up to 85 LPARs on IBM z17 ME1 and up to 40 LPARs on IBM z17 ME2 and IBM z17 MER.

This accelerator function focuses on expediting SSL and TLS processes at the fastest possible speeds, rather than providing dedicated support for securing and storing financial application keys or secrets over extended periods. The Crypto Express8S can be configured as one of the following configurations:

- The Secure IBM CCA coprocessor includes secure key functions with an 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.

  A mode that is called *Payment Card Industry HSM* (PCI HSM) is available in CCA mode. PCI HSM mode simplifies compliance with Payment Card Industry requirements for HSMs.
- The Secure IBM EP11 coprocessor implements an industry-standardized set of services that adheres to the PKCS #11 2.20 specification and more recent amendments. It was designed for extended FIPS and Common Criteria evaluations to meet industry requirements.

CCA features include the ability to use AES keys for operations, such as data encryption, ISO and IBM PIN processing, and PCI SSC-compliant key wrapping to strengthen security. CCA is also extended to support the cryptographic requirements of the German Banking Industry Committee, Deutsche Kreditwirtschaft, and the NIST-selected quantum-safe cryptographic algorithms.

The Crypto Express8S with CCA firmware adds secure key support for quantum-safe cryptographic private keys for signature and key encapsulation mechanism (KEM) purposes.

The CCA interface supports the NIST-standardized algorithm ML-DSA, including HashML-DSA to secure private keys that are usable for digital signature generation and verification. The CCA interface also supports the NIST-standardized algorithm ML-KEM. These components are also combined with Elliptic Curve Diffie-Hellman (ECDH) support to offer a quantum-safe cryptographic hybrid key exchange scheme. This hybrid scheme is designed with secure CCA private keys for all involved private keys in scenarios where data needs enhanced authentication against future quantum computing attacks on conventional cryptography.

Several features that support the usage of the AES algorithm in banking applications were also 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 conversion is supported and used in AES PIN blocks in other CCA callable services. IBM continues to add enhancements as AES finance industry standards are released.

Other enhancements for CCA on Crypto Express8S include the following ones:

- TR-31 (X9.143) native support in all operational APIs. TR-31 is the standard key block in the financial services industry. Users of the CEX8S can now create TR-31 keys under the CCA MK like any other key type and use them in CCA services.
- SHA-3 is now available for hash services and digital signature generation and verification. The CCA host library for Linux on IBM Z also now has support to send SHA-3 requests to the CPACF.

RSA updates include RSA 8192-bit key support, OAEP 2.1 encoding for RSA encryption, decryption, and import and export of RSA-AESKW wrapped keys, which enable a key exchange between CCA and PKCS#11 APIs.

Enhancements for EP11 on Crypto Express8S include the following ones:

- An extra FIPS 140-2 approved image. Customers may switch from the regular to the FIPS approved image on a per AP basis if required by regulators. Usage of the regular image is preferred because it contains the latest features and bug fixes. The FIPS image matches the version that is mentioned by the most recent CMVP certificate that is available for the module.
- A continuous update for z17 adds support for the BLS digital signature algorithm with the BLS12-381 pairing-friendly curve to EP11. BLS signatures are deterministic and non-malleable and used by the Ethereum 2 cryptocurrency system. It enables both aggregation of signatures and public keys. Also, support for the key derivation that is required by deterministic hierarchical wallets as defined in EIP2333 is added.
- A continuous update for z17 adds support for assigning Crypto Adapters in EP11 mode to guest configurations running in secure execution mode. Adapters must be bound and associated with the guest configuration to become operational. Association employs a guest secret that is known to the owner of the image only to help ensure that generated secrets can be used by this secure guest image only, even if the adapter moves to a different configuration without zeroization.

4.7.3 Crypto Express7S (carry forward only)

The Crypto Express7S feature has one or two PCIeCC (HSMs) per feature. For availability reasons, a minimum of two features is required. Up to 16 Crypto Express7S single adapter or up to 30 Crypto Express7S 2-adapter features are supported on IBM z17.

The adapters can be configured as a Secure IBM CCA coprocessor, a Secure IBM EP11 coprocessor, or as an accelerator.

On IBM z17 ME1, Crypto Express7S offers domain support for as many as 85 LPAR instances, and up to 40 LPARs on IBM z17 ME2 and IBM z17 MER.

Trusted Key Entry (TKE) feature: The TKE Workstation feature is required to support the administration of the Crypto Express features when configured as an Enterprise PKCS #11 coprocessor or managing the CCA mode PCI HSM.

Changes were made to the TKE feature to use the quantum-safe cryptographic algorithms that were selected by NIST when authenticating Crypto Express8S, verifying replies from the Crypto Express8S coprocessors, and protecting key parts in flight for the CCA. Finally, the IBM Z pervasive encryption functions are updated to use the quantum-safe algorithms that were selected by NIST for key management.

4.8 Special-purpose features and functions

IBM takes a *total systems* view regarding the design and development of the IBM Z platform. The IBM Z stack is built around digital services, agile application development, connectivity, and system management. This design approach creates an integrated, diverse platform with specialized hardware and dedicated computing capabilities.

IBM z17 delivers a range of features and functions so that CPUs can concentrate on computational tasks while distinct, specialized features take care of the rest.

The latest additions with IBM z17 are the second-generation IBM Z Integrated Accelerator for AI and the IBM Spyre Accelerator, which provide enhanced AI capabilities. The Integrated Accelerator for AI (with the IBM Telum II processor) delivers AI inference in real time, at large scale and rate, without needing to offload data from IBM z17 to perform AI inference. The Spyre Accelerator is a PCIe-attached AI accelerator that is designed to handle larger, more complex AI use cases, and can be clustered together to provide scalable and sustainable AI capabilities.

For more information about other IBM z17 on-chip features, see 5.1.1, “Processor design highlights” on page 56.

4.9 Hardware management

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

The minimum driver level for HMC and SE for IBM z17 is Driver 61. Driver 61 is equivalent to Version 2.17.0. The HMC with Driver 61 or Version 2.17.0 can manage N-2 generations of IBM Z servers and be installed on IBM z15 and IBM z16.

On IBM z17, two HMCs are delivered with the Hardware Management Appliance (HMA) feature. You can order the HMA feature later. However, only new microcode is delivered without HMC hardware.

Note: The HMC code runs on the two integrated 1U rack-mounted servers on the top of the IBM z17 A frame. Stand-alone HMCs that are outside the IBM z17 HMCs (tower or rack mount) are no longer supported and can no longer be ordered.

As closed systems, SEs, HMCs, and HMAs cannot accommodate extra software installations beyond their intended functions.

With IBM z17 and HMA, the SE and HMC codes run *virtually* on the integrated two SEs on the two integrated 1U rack-mounted servers on the top of the IBM z17 A frame. One SE is the Primary SE (active) and the other is the Alternative SE (backup).

The SEs are connected to Ethernet switches for network connectivity with IBM Z servers and the HMAs. An HMC or HMA can communicate with one or more IBM Z servers.

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

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

5

IBM z17 system design strengths

The design strengths that are described in this chapter apply to all IBM z17 servers. Although capacities and configurations vary by model, the underlying architecture and functional capabilities remain consistent across the platform.

Each generation of IBM Z extends the IBM z/Architecture with new functions that enhance performance, security, agility, and flexibility while maintaining compatibility with earlier implementations. This approach enables modern and existing workloads to coexist, protects investments, and supports long-term application continuity.

The IBM z17 design is based on four key architectural areas:

- High-performance processing through advanced cache hierarchy and integrated AI acceleration
- Hardware-based virtualization, which enables efficient consolidation of diverse workloads
- Integrated security and resiliency capabilities across system layers
- Flexible capacity and automation features that support hybrid cloud environments

Together, these characteristics enable IBM z17 to support a wide range of workloads, including transaction processing, data-intensive applications, and AI-enabled services. IBM z17 integrates processor, memory, I/O, AI acceleration, cryptographic functions, and virtualization technologies into a system design that provides scalable performance, efficient resource usage, and high levels of availability within a single platform.

This chapter covers the following topics:

- 5.1, “IBM z17 technology enhancements” on page 56
- 5.2, “IBM Z virtualization” on page 58
- 5.3, “IBM z17 capacity” on page 61
- 5.4, “IBM z17 performance” on page 65
- 5.5, “IBM z17 AI acceleration” on page 68
- 5.6, “IBM z17 sustainability” on page 70
- 5.7, “IBM Z resiliency and cyber resiliency” on page 70
- 5.8, “IBM Z cybersecurity” on page 72
- 5.9, “Automation on IBM Z” on page 76

© Copyright IBM Corp. 2025, 2026.

5.1 IBM z17 technology enhancements

The architectural enhancements of IBM z17 influence how workloads perform and scale, and how effectively the system balances processing, memory, and I/O resources in mixed-workload environments.

IBM z17 follows a holistic system design approach in which resources operate in concert. This approach integrates hardware, firmware, and software enhancements to improve execution efficiency for operations such as sorting, inferencing, compression, and encryption.

IBM z17 provides improvements in the following areas:

- Increased processing capacity and scalability across central processor complex (CPC) configurations
- Enhanced cache hierarchy and memory subsystem for improved data locality
- IBM Integrated Accelerator for AI, cryptography, sorting, and compression
- Improved I/O throughput and reduced latency through dedicated processor units (PUs)

These enhancements support a wide range of workloads, including transactional processing, hybrid cloud applications, and AI-driven services while maintaining high usage and predictable performance.

5.1.1 Processor design highlights

IBM z17 introduces the IBM Telum II processor, which extends z/Architecture and improves execution efficiency across diverse workloads.

Key characteristics include the following items:

- Superscalar design with out-of-order execution (up to six instructions per cycle)
- A high-frequency, low-latency pipeline that is optimized for mixed transactional and analytical workloads
- Support for 24-bit, 31-bit, and 64-bit addressing modes, preserving compatibility with existing applications

5.1.2 Architectural enhancements

IBM z17 includes the following architectural enhancements:

- Cross-CPC drawer (A-Bus):
  - Improved resource management across drawers
  - Reduced interface latency (greater than 10%)
  - Increased effective bandwidth (approximately 10%)
- Processor technology:
  - 5 nm silicon process for improved performance and efficiency
  - Enhanced branch prediction design by using static random access memory (SRAM)
- Cache hierarchy (expanded by approximately 40%):
  - L1 (128 KB instruction and data per core on-chip)
  - L2 (36 MB that is semi-private to the core)
  - Virtual L3 cache (360 MB shared victim cache)
  - Virtual L4 cache (2.8 GB shared victim cache across the system)

- Core-to-nest interface:
  - Increased data movement capacity between processor cores and shared resources
  - Improved throughput for memory-intensive workloads
  - Enhanced security characteristics for data movement

These improvements increase throughput, reduce latency, and improve cache usage for workloads with varying memory access patterns.

5.1.3 Integrated accelerators

Each processor core includes integrated accelerators to offload specialized processing tasks:

- AI Accelerator (Artificial Intelligence Unit (AIU)):
  - Second-generation on-chip design
  - Supports real-time inferencing at scale
  - Provides more than 24 TFLOPS of compute capacity shared across cores
  - Optimized for low-latency, in-transaction AI processing
- Central Processor Assist for Cryptographic Functions (CPACF):
  - High-speed, hardware-assisted encryption and decryption
  - Supports tracking of cryptographic usage, algorithms, and compliance metrics
- IBM Integrated Accelerator for zSort:
  - Uses the SORTL instruction to accelerate sort operations
  - Reduces CPU usage and improves database and batch performance
- IBM Integrated Accelerator for Data Compression:
  - Implements RFC195 compression
  - Reduces the impact that is associated with data storage, movement, and processing
  - Operates transparently without requiring application changes

These accelerators improve overall system efficiency by reducing the load on general-purpose processors and accelerating common workload operations.

5.1.4 I/O subsystem improvements

IBM z17 introduces a Data Processing Unit (DPU) to enhance I/O performance and scalability:

- Improves throughput for I/O and network protocols
- Reduces latency for I/O-intensive and data-driven workloads
- Includes dedicated L1 caches (32 KB instruction and 32 KB data) for I/O processing
- Supports I/O virtualization and offload processing

These enhancements improve end-to-end system throughput and support high-volume, latency-sensitive transaction processing environments.

5.1.5 Memory design

System memory is a key contributor to overall system performance and scalability. IBM z17 memory design includes the following components:

- Large-scale physical memory capacity that is based on the CPC drawer configuration
- Redundant array of independent memory (RAIM) that provides fault tolerance and memory error correction, and uses extra installed memory that is not customer-billable

Here are some example system capacities:

- Up to 64 TB (IBM z17 ME1 with four CPC drawers)
- Up to 17.96 TB (IBM z17 ME2 and IBM z17 MER with two CPC drawers)

The Hardware System Area (HSA) is increased to support system control functions and is not included in application-accessible memory.

5.1.6 IBM Virtual Flash Memory

IBM Virtual Flash Memory (VFM) replaces earlier Flash Express features and provides the following features:

- Reduced paging latency during workload transitions (for example, workload ramp-up periods)
- Improved responsiveness during diagnostic data collection
- Support for meeting service-level objectives in peak or recovery scenarios

VFM can be configured in 512 GB increments as follows:

- Up to 6 TB (IBM z17 ME1)
- Up to 2 TB (IBM z17 ME2 and IBM z17 MER)

These capabilities help improve system responsiveness and reduce the impact of transient workload spikes.

The architectural improvements that are introduced in IBM z17 enhance the foundation for its advanced virtualization capabilities.

5.2 IBM Z virtualization

Virtualization determines how efficiently workloads can be consolidated, isolated, and managed, directly affecting system usage, flexibility, and operational cost.

Virtualization has been a key strength of the IBM Z platform for more than five decades and plays a central role in addressing business challenges such as scalability, resource efficiency, and the secure consolidation of diverse workloads. It is embedded in the hardware, firmware, and operating systems. All computing resources (such as CPU, memory, and I/O) are virtualized, and each set of resources can be used independently within separate operating environments (known as *guest systems*).

The IBM Z platform is designed to concurrently run multiple virtual guest systems, enabling efficient workload consolidation while providing dynamic sharing of resources with minimal cost and performance impact. This capability directly supports the high usage and operational efficiency requirements.

Virtualization management, referred to as a *hypervisor*, operates at both the hardware and software levels of the IBM Z platform. The hardware hypervisor (type 1) is IBM Processor Resource/Systems Manager (PR/SM), which is integrated into the firmware. PR/SM manages hardware resources and builds LPARs that run operating systems, middleware, and applications. PR/SM has evolved over the decades into a secure and fundamental IBM Z technology, with each generation delivering improvements that are aimed at increasing performance, scalability, and efficiency.

The supported software hypervisors (type 2) are IBM z/VM and Kernel-based Virtual Machine (KVM):

- z/VM supports the simultaneous running of multiple virtual guest systems within an LPAR, including nested (multi-level) virtualization. It provides extensive flexibility by emulating hardware devices for guest systems.
- KVM offers an open-source-based virtualization option, enabling the sharing of CPUs (Integrated Facility for Linux (IFL) PUs), memory, and I/O resources, and supporting modern Linux-based workloads.

Together, z/VM and KVM interoperate with PR/SM, using its capabilities to provide a flexible and layered virtualization model. This approach enables the IBM Z platform to support both traditional enterprise applications and modern cloud-native workloads, aligning with the hybrid IT and application modernization use cases.

Multiple hypervisors can coexist and run simultaneously on the IBM Z platform, enabling a wide range of virtualized environments with high performance and integrated security. IBM Z continues to advance virtualization capabilities, delivering highly scalable and dynamic platforms that support both traditional and modern workloads within a single system footprint.

5.2.1 Hardware virtualization

The IBM Z platform is known for its advanced virtualization capabilities, enabling the deployment of diverse workloads with high performance and minimal impact. These capabilities help address key enterprise requirements such as workload isolation, reliability, and efficient resource usage.

Workload separation is a critical design principle. PR/SM in IBM z17 meets the Common Criteria EAL5+ security level, which helps ensure strong isolation between workloads and prevent cross-contamination between applications. This level of isolation is essential for supporting secure multi-tenant and mixed-workload environments.

Logical processors

All physical processors are virtualized as logical processors and can be defined as any of the following PUs:

- CPs for general-purpose workloads
- IBM Z Integrated Information Processor (zIIP) processors for specialized IBM z/OS workloads (such as JVM, IBM Db2, and analytics processing)
- IFL for Linux workloads and associated hypervisors
- Internal Coupling Facilities (ICFs) for IBM Parallel Sysplex clustering

This workload-specific processor classification improves efficiency and optimizes performance for different workload types, supporting the diverse application landscape.

Other system processors include System Assist Processors (SAPs) (for I/O operations) and Integrated Firmware Processors (IFPs) (for firmware and PCIe functions).

PR/SM dynamically dispatches logical processors across physical processors, enabling shared or dedicated configurations. Logical processors can be added or removed dynamically, enabling systems to respond quickly to changing workload requirements, an important capability for elastic and cloud-based environments.

Memory

Memory isolation is strictly enforced to help ensure data integrity and security. Each LPAR contains dedicated memory regions, preventing concurrent access across partitions while presenting a contiguous memory space to the operating system.

Dynamic memory management enables memory to be added nondisruptively, enabling systems to scale in response to workload demands. z/Architecture supports a highly scalable virtual storage model, with address spaces up to 16 EB, supporting modern data-intensive applications.

I/O virtualization

IBM z17 supports highly flexible I/O virtualization, enabling shared or dedicated access to I/O resources across LPARs. Capabilities such as Multiple Image Facility, Parallel Access Volume (PAV), and HyperPAV enhance throughput and reduce latency for data-intensive workloads.

Dynamic I/O configuration enables systems to adapt without disruption, improving availability and operational agility. These capabilities are essential for supporting mission-critical workloads and high-performance data environments.

5.2.2 Cloud environments

Virtualization is critical to the viability of cloud service offerings because it provides the elasticity that is required to handle fluctuating demands for IT resources. On IBM z17, virtualization is deeply integrated across hardware, firmware, and operating systems, resulting in highly efficient resource use (up to 100% sustained usage) while maintaining strong isolation and security. These capabilities directly support the cost optimization, scalability, and resilience needs.

Cloud elasticity requirements are addressed through IBM z17 granularity offerings, including capacity levels, Tailored Fit Pricing for IBM Z (for variable and business-critical workloads), and CoD offerings. These features enable organizations to match infrastructure resources with workload demands, reinforcing IBM Z as a leading platform for enterprise and hybrid cloud environments.

Managing cloud environments requires tools that can efficiently orchestrate virtualized compute, storage, and network resources. OpenStack provides a flexible cloud management framework that automates infrastructure provisioning and supports a wide range of cloud service and deployment models.

IBM z17 can also be integrated with IBM Cloud® services, enabling organizations to transform infrastructure, applications, and data while simplifying operations. Cloud-native capabilities are delivered through IBM Cloud Paks, which provide pre-integrated, containerized solutions with a consistent operating model.

Through Red Hat integration, IBM Z extends its hybrid cloud capabilities with support for Red Hat OpenShift and Linux on IBM Z. This support enables containerized workloads, modern development practices, and access to a broad open-source ecosystem, which are key enablers for the application modernization and hybrid cloud strategies.

For more information about hybrid cloud capabilities, see Hybrid cloud with IBM Z.

Other offerings help automate and manage hybrid and multi-cloud environments:

- IBM Vault Self-Managed for enterprise secrets management
- IBM Terraform® for infrastructure as code (IaC), governance, and automated provisioning

For more information about these offerings, see this announcement.

IBM Infrastructure Suite for z/VM and Linux is the comprehensive software suite that can manage, monitor, and protect your virtualized workloads on IBM Z. It unifies infrastructure operations, including provisioning, backup, and performance monitoring, into a single solution that is tailored for enterprise hybrid cloud environments.

IBM Cloud Infrastructure Center is an OpenStack compatible platform that automates IaaS for traditional and containerized workloads for integrating IBM Z into broader enterprise hybrid clouds.

Building on these virtualization capabilities, IBM z17 provides flexible capacity options that enable systems to scale efficiently in response to workload demands.

5.3 IBM z17 capacity

Capacity features determine how systems are sized and scaled, enabling you to align resource allocation with workload demands and cost objectives over time.

IBM z17 provides increased capacity and performance compared to IBM z16. Several elements contribute to this effect, including many processors, individual processor performance, memory caches, simultaneous multithreading (SMT), and machine instructions, including single instruction multiple data (SIMD). Subcapacity settings continue to be offered.

The family of CoD offerings helps ensure a flexible addition of capacity when it is most needed, for example, during peak workload periods, scheduled maintenance, or in DR scenarios.

Tailored Fit Pricing for IBM Z options are designed to deliver unmatched simplicity and predictability of hardware capacity and software pricing, even in the constantly evolving era of hybrid cloud.

Note: Capacity and performance ratios are based on measurements and projections by using 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.

5.3.1 Capacity settings

IBM z17 expands the offer on subcapacity settings. Finer granularity in capacity levels enables the growth of installed capacity to follow more closely the enterprise growth for a smoother, pay-as-you-go investment profile. Many performance and monitoring tools are available on IBM Z environments that are coupled with the flexibility of the CoD options (see 5.3.2, “CoD offerings” on page 63). These features help you manage growth by making capacity available when it is needed.

IBM z17 ME1 capacity levels

IBM z17 ME1 offers four distinct capacity levels for the first 43 CPs[1]. Three of the levels are subcapacity levels (4xx, 5xx, and 6xx):

- One full capacity
- Three subcapacities

Full capacity CPs are identified as CP7. On IBM z17 ME1, up to 208 CPs can be configured as CP7. Up to 43 CPs can have subcapacity. The three subcapacity levels are identified by CP6, CP5, and CP4, and are displayed in hardware descriptions as Feature Codes on the CPs:

- Max43 (Feature Code 0571)
- Max90 (Feature Code 0572)
- Max136 (Feature Code 0573)
- Max183 (Feature Code 0574)
- Max208 (Feature Code 0575)

These processors deliver the scalability and granularity to meet the needs of medium-sized enterprises while also satisfying the requirements of large enterprises that have large-scale, mission-critical transaction and data processing requirements.

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

If more than 43 CPs are configured to the system, all must be full capacity because all CPs must be at the same capacity level. Granular capacity adds 129 subcapacity settings to the 208 capacity settings that are available with full capacity CPs (CP7), equaling 337 distinct capacity settings in IBM z17 ME1.

A processor is always set at full capacity when it is characterized as anything other than a CP, such as a zIIP, an IFL, or an ICF. A distinct pricing framework governs non-CP components, including acquisition and maintenance costs, along with diverse software licensing options that are tailored to these components.

To help size an IBM Z server to fit your requirements, IBM provides a no-cost tool that reflects the latest IBM Large Systems Performance Reference (LSPR) measurements: the IBM Z Processor Capacity Reference (zPCR). Download the tool from Getting Started with zPCR.

For more information about LSPR measurements, see LSPR for IBM Z.

IBM z17 ME2 and IBM z17 MER capacity levels

IBM z17 ME2 and IBM z17 MER offer 26 distinct capacity levels for each CP in the configuration for a total of 156 (26 x 6) capacity settings. These processors deliver the scalability and granularity to meet the needs of small and medium-sized enterprises.

IBM z17 ME2 and IBM z17 MER PUs are always set to full capacity when the PUs are characterized as anything other than a CP, such as a zIIP, an IFL, or an ICF.

[1] A central processor (CP) is the IBM Z standard processor for use with any supported operating system. It is required to run z/OS.

5.3.2 CoD offerings

IBM z17 maintains its CoD offerings, delivering flexibility, control, and administrative use while addressing diverse resource needs across multiple scenarios with refined granularity over allocated resources.

IBM z17 can perform concurrent upgrades, which provide an increase of processor capacity with no platform outage. Usually, with operating system support, a concurrent upgrade can also be nondisruptive to the operating system. These upgrades are based on the enablement of resources that are physically present in IBM z17.

Capacity upgrades encompass both permanent and temporary adjustments to installed capacity, manageable through the Customer Initiated Upgrade (CIU) feature (no IBM service personnel intervention is required). These modifications initiate online through IBM Resource Link, providing customers with self-sufficient control over their system's resource allocation.

Using the CIU facility requires a special contract between the customer and IBM. This contract specifies the terms and conditions for online CoD buying of upgrades, and other types of CoD upgrades are accepted. For more information, see IBM Resource Link.[2]

For more information about the CoD offerings, see *IBM z17 (9175) Technical Guide*, SG24-8579.

Permanent upgrades

Permanent upgrades of processors (CPs, IFLs, ICFs, zIIPs, and SAPs) and memory, or changes to a platform's Model-Capacity Identifier (up to the limits of the installed processor capacity on an existing IBM z17), can be performed by customers through the IBM Online Permanent Upgrade offering by using the CIU facility.

Temporary upgrades

Temporary upgrades of an IBM z17 can be done through On/Off CoD or Capacity BackUp (CBU) that is ordered from the CIU facility.

On/Off CoD function

On/Off CoD is a function that enables concurrent and temporary capacity growth of the CPC. On/Off CoD can be used for peak workload requirements for any period. On/Off CoD includes a daily hardware charge, and it can include an associated software charge. On/Off CoD offerings can be prepaid or post-paid.

When you use the On/Off CoD function, you can concurrently add processors (CPs, IFLs, ICFs, zIIPs, and SAPs), increase the CP capacity level, or both.

Prepaid On/Off CoD tokens: New prepaid On/Off CoD tokens do *not* carry forward to future systems.

Capacity BackUp function

CBU enables you to perform a concurrent and temporary activation of extra CPs, ICFs, IFLs, zIIPs, and SAPs, an increase of the CP capacity level, or both. This function can be used during an unforeseen loss of IBM Z capacity within the enterprise, or to perform a test of your DR procedures. The capacity of a CBU upgrade cannot be used for peak workload management.

[2] An IBM Resource Link ID is required.

CBU features are optional and require unused capacity to be available on CPC drawers of the backup system as unused PUs to increase the CP capacity level on a subcapacity system, or both. A CBU contract must be in place before the LICCC 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) and one disaster activation (up to 90 days). The record can be configured to be valid for up to 5 years. You can also order more tests for a CBU record in quantities of five, up to a maximum of 15 tests.

Suitable use of the CBU capability does not incur any other software charges from IBM.

5.3.3 IBM Z Flexible Capacity for Cyber Resiliency

IBM Z Flexible Capacity for Cyber Resiliency (Feature Code 9933 together with Feature Code 0376) can make production capacity available across sites between IBM z17 platforms for flexibility and elasticity in DR testing, planned maintenance, proactive outage avoidance, and actual DR scenarios.

IBM Z Flexible Capacity for Cyber Resiliency works with other temporary record types, for example, On/Off CoD and Tailored Fit Pricing for IBM Z Hardware Consumption Solution (TFP-Hardware).

With IBM Z Flexible Capacity for Cyber Resiliency, capacity can be allocated remotely without requiring onsite personnel (IBM or customer) after the initial setup. IBM Z Flexible Capacity for Cyber Resiliency provides the following benefits:

- Flexible duration of capacity (up to 1 year).
- Fully automated by using solutions such as IBM Geographically Dispersed Parallel Sysplex (IBM GDPS®)[3].
- Simplify compliance and improve confidence in DR scenarios, including tests.
- No need for CBU records by using this solution.

For more information, see *IBM Z Flexible Capacity for Cyber Resiliency*, REDP-5702.

z/OS capacity provisioning

Capacity provisioning helps you manage the CP and zIIP capacity of an IBM z17 server that is running one or more instances of the z/OS operating system. By using the IBM z/OS Capacity Provisioning Manager component, On/Off CoD temporary capacity can be activated and deactivated under control of a defined policy. Combined with functions in z/OS, the IBM z17 provisioning capability gives you a flexible, automated process to control the configuration and activation of On/Off CoD offerings.

5.3.4 Tailored Fit Pricing for IBM Z

TFP-Hardware provides an always-on, consumption-based capacity corridor that provides flexibility and control for unpredictable workload spikes throughout the day. It helps you scale IT demands and control behavior with a pay-for-use buffer and granular usage measurements.

[3] GDPS is a collection of software offerings that help with DR and business continuity. GDPS can be used for single-site or multi-site environments. For more information, see *IBM GDPS: An Introduction to Concepts and Capabilities*, SG24-6374.

Tailored Fit Pricing for IBM Z helps make embracing hybrid cloud simpler with IBM Z. The pricing option delivers simplicity, flexibility, and predictability of pricing across the stack, even with constantly increasing unpredictability in business demand.

IBM Z Flexible Capacity for Cyber Resiliency can be used with TFP-Hardware. The Flexible Capacity limit is based on the Base Capacity. The presence or activation of TFP-Hardware does not affect the amount of capacity that can be activated by a Flexible Capacity Transfer record.

The Flexible Capacity Transfer record is considered to be the first record that was activated, regardless of the order in which temporary records or TFP-Hardware were activated.

For more information about the CoD offerings and Tailored Fit Pricing for IBM Z, see *IBM z17 (9175) Technical Guide*, SG24-8579.

The available capacity must be considered together with system performance characteristics, which are influenced by processor design, memory hierarchy, and workload behavior.

5.4 IBM z17 performance

Understanding performance characteristics is essential for workload placement, tuning, and capacity planning, particularly for CPU- and memory-intensive applications.

The IBM Z microprocessor chip for IBM z17 has a high-frequency design that uses IBM leading microprocessor technology and offers more cache per core than other chips. Also, 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 IBM z17 is a result of the enhancements that are described in 5.1, “IBM z17 technology enhancements” on page 56.

5.4.1 LSPR workload suite: IBM z17 changes

To help you better understand workload variations, IBM provides a no-cost tool, zPCR, which is available at Getting Started with IBM Z Processor Capacity Reference.

IBM continues to measure the 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 millions of service units (MSU) ratings for reference.Capacity performance is closely associated with how a workload uses and interacts with a specific processor hardware design. Workload capacity performance is sensitive to the following major factors:

- Instruction path length
- Instruction complexity
- Memory hierarchy

The CPU Measurement Facility (CPUMF) data offers insight into the interaction of workload with the hardware design. CPUMF 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. The workload category is determined by the Level 1 Misses per 100 instructions (L1MP) and the *relative nest intensity* (RNI).

With the IBM Z platform, the LSPR introduced the following workload capacity categories that replace all prior primitives and mixes:

- LOW (RNI): A workload category that represents light use of the memory hierarchy.
- AVERAGE (RNI): A workload category that represents average use of the memory hierarchy. This category is expected to represent most production workloads.
- HIGH (RNI): A workload category that represents heavy use of the memory hierarchy.

These categories are based on the RNI, 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. CPU MF data can be collected by z/OS SMF record type 113 or z/VM Monitor by using CPU Measurement Facility Host Counters.

Beyond traditional low, average, and high classifications, the newest zPCR introduces the low-average and average-high mixed categories, offering enhanced precision 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 on IBM Z and z/VM environments.

The LSPR contains the internal throughput rate ratios for IBM z17 and the previous generations of processors. The report is based on measurements and projections by using standard IBM benchmarks in a controlled environment. The actual throughput that any user experiences varies depending on several factors, such as the amount of multiprogramming in the job stream, the I/O configuration, and the workload that is processed.

Experience demonstrates that IBM Z servers can run at up to 100% usage levels sustained. However, most users prefer to leave some available capacity 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 involve IBM Support when you plan for capacity.

For more information about IBM z17 performance, see *IBM z17 (9175) Technical Guide*, SG24-8579 and *IBM z17 (9176) Technical Guide*, SG24-8608.

5.4.2 IBM z17 throughput optimization

The memory and cache structure implementation in the CPC drawers of IBM z17 is enhanced compared to previous generations to provide sustained throughput and performance improvements. The memory is distributed throughout the CPC drawers and the CPC drawers have individual levels of cache that are private to the cores and shared by the cores. Nonetheless, all processors can access the highest level of cache and all the memory. Therefore, the system is managed as a memory-coherent symmetric multiprocessor.

Processors within the IBM z17 CPC drawer structure have different distance-to-memory attributes. CPC drawers are fully interconnected to minimize the distance. Other non-negligible effects result from data latency when grouping and dispatching work on a set of available logical processors. To minimize latency, the system attempts to dispatch and later redispatch work to a group of physical CPUs that share cache levels.

PR/SM manages the usage 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 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 closely together. They can cooperate to create an affinity between task and physical processor rather than between LPAR and physical processor, which is known as *HiperDispatch*.

5.4.3 HiperDispatch

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

In general, the PR/SM dispatcher assigns work to the minimum number of logical processors that are needed for the priority (weight) of the LPAR. On IBM z17, PR/SM attempts to group the logical processors into the same logical cluster or in the neighboring logical cluster in the same CPC drawer and, if possible, in the same chip. This configuration results in reduction of multi-processor effects, maximizing the usage of shared cache and reducing 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 improve 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 IBM z17. The entire IBM z17 stack (hardware, firmware, and software) tightly collaborates to obtain the full potential of the hardware. z/VM HiperDispatch provides support similar to the one in z/OS. You can dynamically turn on and off HiperDispatch without requiring an IPL.

IBM z17 includes several HiperDispatch algorithm enhancements, including the following examples:

- Improved memory affinity
- Improved LPAR placement based on IBM z16 experience
- Usage of a new chip configuration

Note: HiperDispatch requires SMT activation, which serves as an essential feature in IBM Z LSPR measurements that are all conducted within z/OS environments with HiperDispatch enabled. It is best practice to activate HiperDispatch for production workloads.

5.5 IBM z17 AI acceleration

Integrated AI capabilities enable real-time inferencing within transactional workflows, enabling AI models to be applied directly to operational data without requiring separate platforms or data movement.

IBM z17 enhances AI capabilities through a combination of on-chip and PCIe-attached accelerators. These capabilities are designed to support high-throughput, low-latency inferencing for both traditional machine learning (ML) models and emerging generative AI workloads.

5.5.1 AI acceleration architecture

IBM z17 provides AI acceleration through two primary components:

- Integrated AI Accelerator (on-chip):
  - Embedded in the IBM Telum II processor
  - Optimized for low-latency, in-transaction inferencing
  - Eliminates the need to move data off platform
- IBM Spyre Accelerator (PCIe-attached):
  - Designed for foundation models and generative AI workloads
  - Supports more complex model execution and larger datasets
  - Complements on-chip acceleration for hybrid AI use cases

Together, these components support a range of AI deployment patterns, from real-time decision-making to model-driven automation.

5.5.2 Integrated AI Accelerator

The IBM Telum II Integrated AI Accelerator provides improvements in inferencing performance:

- More than 24 trillion operations per second (TOPS)
- Approximately 4× increase in compute capability compared to previous generation
- Support for INT8 data types
- Shared access across cores for efficient usage

This design enables consistent, low-latency inferencing for high-volume transactional workloads.

5.5.3 AI workload integration

IBM z17 is designed to integrate AI processing directly into core business workflows:

- AI models can run within transactional applications
- Data remains within the platform, minimizing latency and improving security
- AI processing can occur alongside traditional application logic

This integration supports use cases such as the following ones:

- Fraud detection during financial transactions
- Risk scoring and compliance validation
- Real-time personalization and decision-making

5.5.4 Composite AI architecture

IBM z17 supports composite AI approaches that combine multiple model types within a single workflow:

- Traditional ML models (structured data)
- LLMs (unstructured data)
- Hybrid decision pipelines for improved accuracy

Here is an example:

- A structured data model evaluates transaction patterns.
- A language model analyzes associated text data.
- Results are combined to produce a final decision.

This approach improves prediction accuracy and reduces false positives while maintaining real-time response characteristics.

5.5.5 Generative AI and AI services

IBM z17 supports generative AI capabilities through integrated tools and accelerators:

- IBM watsonx® Code Assistant® for IBM Z:
  - Assists with code understanding, refactoring, and transformation
  - Supports COBOL-to-Java modernization scenarios
- IBM watsonx Assistant® for IBM Z:
  - Provides AI-driven operational assistance
  - Uses retrieval augmented generation (RAG) for domain-specific responses
  - Supports task automation through integration with tools such as Ansible and z/OSMF

These solutions improve developer productivity and reduce operational complexity.

5.5.6 AI development ecosystem

IBM z17 supports widely used AI frameworks through optimized and supported runtime environments:

- TensorFlow (IBM Z Accelerated for TensorFlow)
- PyTorch (IBM Z Accelerated for PyTorch)
- SnapML for model training and scoring
- Triton Inference Server for model deployment
- ONNX-based model execution with Deep Learning Compiler (DLC)

These frameworks have the following features:

- Optimized for the Integrated AI Accelerator
- Built and validated by using IBM secure engineering practices
- Available with enterprise-grade support through the AI Toolkit for IBM Z and LinuxONE

With this approach, you can use existing skills and tools while benefiting from platform-specific optimizations.

As system capabilities expand, considerations such as energy efficiency and environmental impact become more important.

5.6 IBM z17 sustainability

Energy efficiency and resource usage directly affect operational costs and environmental impact, which are important considerations in system design.

IBM z17 is designed for end-to-end sustainability, which can reduce TCO and energy costs while helping you adhere to global environmental regulatory requirements.

IBM z17 optimizes energy consumption by reducing power and cooling demands, reducing costs and carbon emissions. When compared to running the same workloads on previous generation systems, immediate savings in energy consumption are achieved through improved performance per kilowatt.

Also, IBM z17 provides sustainability instrumentation, enabling real-time power consumption reporting and monitoring at the LPAR level, with support from the operating systems, including z/OS[4], z/VM[5], and Linux on IBM Z. These features enable better tracking and management of power usage.

Alongside sustainability, IBM z17 continues to emphasize resiliency, helping ensure continuous operation and data protection in complex environments.

5.7 IBM Z resiliency and cyber resiliency

Resiliency capabilities determine how systems maintain availability and recover from failures or data corruption, which is critical for business continuity and risk management.

IBM Z resiliency addresses the need for continuous operations, data integrity, and service availability in modern enterprise environments. Organizations must handle infrastructure failures and sophisticated cyberthreats and data integrity risks. To meet these challenges, IBM Z combines traditional resiliency capabilities such as reliability, business continuity, and DR with advanced cyber resiliency mechanisms.

At its foundation, resiliency on IBM Z is built on a set of integrated capabilities that help ensure that systems remain available, recover quickly from disruptions, and protect critical business data. These capabilities include hardware and software reliability, clustering technologies, automated recovery, and secure data protection mechanisms.

5.7.1 Traditional resiliency foundations

Traditional resiliency foundations on IBM Z provide the baseline capabilities that are required to ensure continuous operations and service availability. These capabilities, centered on reliability, availability, and DR, are designed to minimize outages, maintain system stability, and enable rapid recovery from infrastructure failures. Together, they form the core building blocks on which more advanced resiliency and cyber resiliency capabilities are built.

[4] Power information is available by using SMF records. Post-processing provides approximations of z/OS service-class level power consumption.

[5] Power consumption metrics are provided within z/VM monitor, including CPC-specific and LPAR-specific level power consumption information.

Reliability, availability, and serviceability

IBM z17 continues the IBM Z tradition of industry-leading RAS capabilities, which are deeply integrated into both hardware and software. With mean time between failures measured in decades, the platform is designed to minimize outages and maintain consistent service levels.

The RAS strategy follows a building-block approach that includes the following components:

- Error prevention
- Error detection
- Recovery
- Problem determination
- Service structure
- Change management
- Measurement and analysis

These capabilities are continuously enhanced across generations to improve system stability, reduce downtime, and support evolving workload demands.

IBM Z also supports concurrent replace, repair, and upgrade operations across processors, memory, and I/O components, and nondisruptive firmware updates. Many capacity upgrades can be performed without system outages, helping ensure continuous operation.

High availability with IBM Parallel Sysplex

Parallel Sysplex extends RAS capabilities by enabling tightly coupled clustering of z/OS systems. This active-active architecture enables workloads to run simultaneously across multiple systems, providing near-continuous availability.

Parallel Sysplex supports shared access to data across LPARs and systems by using IBM Z I/O virtualization to maximize usage while maintaining strong isolation. In a failure, workloads continue to run on surviving systems, with the goal of achieving zero application downtime.

DR and automation with GDPS

GDPS enhances resiliency by providing automation for DR, storage replication, and system coordination. It manages failover and recovery across geographically dispersed environments, helping organizations maintain service continuity even during site-level failures.

GDPS integrates with Parallel Sysplex and storage subsystems to deliver highly automated recovery processes, enabling HA levels. These capabilities address the business continuity and uptime expectations.

For more information about IBM Z resiliency, see *Getting Started with IBM Z Resiliency*, SG24-8446.

5.7.2 IBM Z cyber resiliency

Building on these traditional resiliency capabilities, IBM Z cyber resiliency focuses on protecting against, detecting, and recovering from cyber-related disruptions, particularly ones that involve data corruption or malicious activity.

Cyber resiliency extends beyond infrastructure availability to help ensure that trusted, recoverable data is always available, even in the presence of sophisticated attacks.

IBM Z cyber resiliency includes these key elements:

- Immutable data and secure recovery

  The IBM Z Cyber Vault solution provides a structured approach to protecting and recovering critical data. It creates point-in-time immutable copies of production data that are stored in a secure, isolated environment.

  These copies are regularly validated to help ensure integrity and readiness for recovery. In a cyber incident, organizations can perform these tasks:

  - Forensic analysis
  - Targeted (surgical) recovery
  - Full system restoration

  The solution consists of three layers:

  - Storage: IBM DS8000 Safeguarded Copy for immutable backups
  - Automation: GDPS Logical Corruption Protection Manager
  - Environment: Isolated vault for validation and recovery

  For more information, see Server and cyber resiliency with IBM Z and *Getting Started with IBM Z Cyber Vault*, SG24-85111.

- Minimizing recovery time

  IBM Z System Recovery Boost (SRB) accelerates system restart, shutdown, and recovery operations by temporarily increasing processing capacity and parallelism. This capability reduces downtime and helps restore services more quickly after disruptions.

  For more information, see *Introducing IBM Z System Recovery Boost*, REDP-5563.

- Threat detection and response

  Early detection of cyberthreats is critical to limiting their impact. IBM Threat Detection for z/OS (IBM TDz) uses AI-based analytics to identify anomalous data access patterns and generate alerts for potential attacks.

  IBM TDz includes policy controls and filtering mechanisms to reduce false positives and provides actionable insights for rapid diagnosis and remediation. It also supports compliance requirements (like DORA) by strengthening security monitoring and reporting capabilities.

IBM Z resiliency combines traditional high availability and disaster recovery capabilities with advanced cyber resiliency features to address modern enterprise risks. By integrating strong RAS, clustering, automation, immutable data protection, and threat detection, IBM z17 enables organizations to maintain continuous operations and rapidly recover from both system failures and cyber incidents.

Complementing resiliency capabilities, IBM z17 provides comprehensive cybersecurity features to protect systems, data, and workloads.

5.8 IBM Z cybersecurity

Security capabilities influence how data, applications, and infrastructure are protected, and how compliance requirements are met across hybrid environments.

Although often confused, cybersecurity and cyber resiliency serve distinct roles. *Cybersecurity* focuses on preventing attacks, and *cyber resiliency* addresses the ability to withstand and recover from successful breaches. Together, they form a comprehensive approach to protecting enterprise systems and data.

IBM Z cybersecurity emphasizes proactive protection through a comprehensive set of integrated capabilities across hardware, firmware, and software layers. Rather than viewing IBM Z simply as the most secure platform, it is more accurate to describe it as the most securable platform, providing extensive security functions that organizations can tailor to meet their specific requirements.

Effective security on IBM Z requires understanding and deliberately applying these capabilities:

- Identity and access control

  Security begins with strong identity verification and access governance. Users must authenticate before accessing resources, and this process can be strengthened with MFA, such as IBM Z Multi-Factor Authentication, which adds extra verification steps.

  Once authenticated, access control systems enforce permissions that are based on defined roles. IBM Resource Access Control Facility (IBM RACF®) provides centralized identity management and supports RBAC,[6] helping ensure that users can access only the data and functions that are appropriate to their role.

- Encryption and data protection

  Encryption is a foundational element of IBM Z cybersecurity, protecting data at rest and in transit. IBM Z supports a wide range of cryptographic algorithms, including quantum-safe methods, helping ensure long-term data protection.

  IBM z17 enhances this capability with the following components:

  – Secure Boot mechanisms, helping ensure that only trusted firmware runs
  – Hardware root of trust, anchoring system integrity
  – Dual-signature verification, combining classical and quantum-safe cryptography

  These capabilities protect the system from unauthorized modification and help establish a secure operational baseline.

- Pervasive encryption

  Traditional security approaches relied on defining protected perimeters around data. However, modern environments, especially hybrid cloud, require a data-centric model.

  Pervasive encryption shifts protection from infrastructure boundaries to the data itself by encrypting the following data:

  – Data at rest
  – Data in flight
  – Data across systems and environments

  This approach helps reduce exposure risk without requiring application changes and supports regulatory requirements such as GDPR, PCI-DSS, and HIPAA.

  IBM z17 implements pervasive encryption through tight integration of hardware and software, including these key components:

  – CPACF for high-performance symmetric encryption
  – Crypto Express Hardware Security Modules (HSMs) for secure key storage and advanced cryptographic operations
  – Policy-based encryption for datasets, volumes, networks, storage, and Coupling Facility (CF) data

  For more information, see *Getting Started with z/OS Data Set Encryption*, SG24-8410.

[6] RBAC is a policy-neutral access control mechanism that is defined around roles and privileges. It can also be used to help achieve separation of duties.

- Quantum-safe security

  Emerging quantum computing capabilities pose potential risks to traditional encryption methods. IBM z17 addresses this situation with quantum-safe cryptography that is aligned with NIST standards.

  Supported algorithms include the following ones:

  - Federal Information Processing Standards (FIPS) 203 (Module-Lattice-Based Key-Encapsulation Mechanism (ML-KEM)) for key encapsulation
  - FIPS 204 (Module-Lattice-Based Digital Signature (ML-DSA)) for digital signatures

  IBM z17 also implements the following components:

  - Hybrid cryptographic schemes
  - Quantum-safe Secure Boot
  - Hardware-protected signature validation

  Supporting tools such as IBM Application Discovery and Delivery Intelligence (ADDI), IBM z/OS Integrated Cryptographic Service Facility (ICSF), and IBM Crypto Analytics Tool (CAT) help organizations assess and modernize their cryptographic environments.

  For more information, see *Transitioning to Quantum-Safe Cryptography on IBM Z*, SG24-8525.

- Encryption key management

  Secure and efficient key management is critical to maintaining encryption effectiveness. Unified Key Orchestrator for IBM z/OS (formerly IBM Enterprise Key Management Foundation-Web Edition) provides centralized lifecycle management for encryption keys across hybrid environments.

  UKO enables the following features:

  - Policy-based key generation
  - Role-based access and dual control
  - Auditability and compliance tracking
  - Key rotation and multi-tenancy

  This centralized approach simplifies key management while strengthening governance and compliance.

- Secure execution and confidential computing

  IBM Secure Execution for Linux provides a Trusted Execution Environment (TEE) that protects data in use by isolating workloads from unauthorized access, including system administrators and hypervisors.

  This capability helps ensure the following scenarios:

  - Confidentiality and integrity of workloads
  - Protection of sensitive data in memory
  - Secure workload deployment and attestation

  IBM Confidential Computing Platform extends these capabilities to containerized environments, enabling secure deployment of applications with enforced policies, encrypted contracts, and workload isolation in hybrid cloud scenarios.

  For more information, see *IBM Hyper Protect Platform: Applying Data Protection and Confidentiality in a Hybrid Cloud Environment*, SG24-8555.

- Secure workload isolation

  IBM Z provides advanced isolation mechanisms to protect workloads from both internal and external threats:

  - Secure Service Containers, which deliver fully encrypted and isolated environments
  - Restricted administrative access
  - Tamper-proof boot and runtime protection
  - API-only access models

  These features help ensure that sensitive applications and data remain protected throughout their lifecycle.

- Data-in-flight protection

  IBM Fibre Channel Endpoint Security enhances protection for data in motion between IBM Z systems and storage. It provides the following features:

  - Endpoint authentication
  - Encryption at the adapter level
  - Secure key exchange and management

  This capability helps ensure data integrity and confidentiality without requiring application changes.

  For more information and implementation details, see *IBM Fibre Channel Endpoint Security for IBM DS8900F and IBM Z*, SG24-8455.

- Compliance and governance

  IBM Z simplifies compliance through automation and integrated tools. The IBM Z Security and Compliance Center (zSCC) aggregates data across systems to streamline audits and enforce policies.

  Here are some of its features:

  - Predefined compliance profiles
  - Custom compliance goals
  - Integration with enterprise tools such as IBM Concert®
  - IBM Security® Patch Insights for z/OS

  These capabilities support regulatory requirements and reduce operational impact.

  For more information, see *Keeping Up With Security and Compliance on IBM Z*, SG24-8540.

- Emerging capabilities

  IBM continues to enhance cybersecurity through innovation, including planned AI-driven capabilities to automatically discover, classify, and tag sensitive data by using the IBM Telum II processor and ML technologies. These advancements will enable more targeted and effective data protection strategies.

IBM Z cybersecurity provides a comprehensive, layered security model that protects identities, data, workloads, and infrastructure. By combining strong authentication, pervasive encryption, quantum-safe cryptography, confidential computing, and automated compliance, IBM z17 enables organizations to proactively defend against modern cyberthreats while supporting secure hybrid cloud and AI-driven environments.

To support these capabilities in operational environments, IBM z17 integrates automation features that simplify management and improve efficiency.

5.9 Automation on IBM Z

Automation reduces operational complexity and enables consistent system management, enabling organizations to scale operations and improve responsiveness to changing conditions.

IBM automation solutions for IBM Z can provide real-time visibility, streamlined workloads, and security for application development and integration.

The key benefits of IBM Z automation help optimize IT processes, simplify management, reduce skill requirements, improve performance, and enable system self-management.

Automation on IBM Z is an established collection of tools and services that automate repetitive tasks across many diverse areas:

- Performance observability
- Resiliency
- Monitoring
- Data migration
- System management

For more information about the tools and services that make up IBM Z System Automation, see IT automation for IBM Z.

IBM Z integrates automated resiliency solutions to help ensure business continuity and improve data movement across hybrid cloud environments:

- IBM Z System Automation and NetView provides policy-driven failover automation, DR orchestration, and proactive incident response, reducing manual intervention in system failures.
- The IBM GDPS LCP manager provides DR, continuous availability, and data replication across multiple sites, helping ensure minimal downtime and data loss.

For more information about other automation products for managing hardware and software resources, see IBM Z System Automation.

In addition, automation through Red Hat Ansible Automation Platform and IBM Terraform improves both infrastructure provisioning and system management:

- Red Hat Ansible Automation Platform simplifies z/OS software management, system configurations, and hybrid cloud orchestration. With Ansible Certified Content for IBM Z, IT teams can automate LPAR provisioning, middleware deployment (like IMS and IBM CICS®), and security updates by using declarative, repeatable playbooks. This approach reduces skill requirements and accelerates DevOps adoption.
- IBM Terraform for IBM Z enables IaC, enabling teams to provision and update IBM Z resources with declarative configurations. This approach eliminates configuration drift, helps ensure consistency, and integrates IBM Z with cloud automation workflows.

Together, Ansible and Terraform unify automation across IBM Z and hybrid cloud environments, reducing complexity and accelerating IT modernization efforts.

For more information about infrastructure automation, see Red Hat Ansible Automation Platform and What is Terraform?.

AI-driven automation advancements with IBM Z also offer numerous benefits, including task automation, improved efficiency, reduced human error, and enhanced application development and integration. By pairing AI and automation, operations can be streamlined and improved service can be provided with IBM Z AIOps.

5.9.1 Automation in IBM Z AIOps

IBM Z AIOps uses AI and ML algorithms to analyze, detect, and predict system behaviors in real time. This approach helps reduce downtime, improve efficiency, and proactively resolve issues before they impact critical workloads.

On IBM Z, AIOps delivers the following features:

- Real-time performance observability: AI continuously monitors system performance, detecting anomalies and predicting failures to help ensure system stability.
- Automated incident response: By using ML-driven root cause analysis, IBM z17 reduces mean time to resolution by automatically identifying the source of issues.
- Optimized workload execution: AIOps dynamically adjusts workloads and system resources, helping ensure high performance while minimizing costs.

These capabilities enable IBM z17 to self-manage IT processes, reducing operational complexity while increasing system reliability. For more information, see AIOps for IBM Z.

5.9.2 Application development and integration

IBM z17 automates software development and integration workflows to optimize DevOps pipelines and enhance productivity with these offerings:

- IBM watsonx Assistant for IBM Z: AI-driven natural language automation helps operators troubleshoot and run tasks by using an AI-powered assistant that is trained on IBM Z operations.
- IBM watsonx Code Assistant® for IBM Z: Accelerates the application lifecycle and streamlines COBOL to Java modernization efforts.
- IBM Test Accelerator for IBM Z: Automates unit and integration testing, enabling developers to rapidly validate z/OS applications in virtual environments, improving code quality.
- z/OS Cloud Broker and Wazi Developer for Workspaces: Delivers self-service provisioning of z/OS environments and modern developer workspaces, enabling faster onboarding, streamlined access to mainframe resources, and tighter integration with CI/CD pipelines..

These tools help reduce software development time, improve application reliability, and enable AI-assisted DevOps automation on IBM z17.

For more information, see AI on IBM Z.

6

IBM z17 software support

Software support defines which operating systems and tools can use IBM z17 capabilities and determines compatibility, lifecycle planning, and modernization options.

This chapter outlines the supported operating systems and accompanying software for the IBM z17 and its respective features.

The following major operating systems are supported:

- IBM z/OS
- IBM z/VM
- IBM z/TPF
- 21st Century Software VSE^{n}
- Linux on IBM Z and the Kernel-based Virtual Machine (KVM) hypervisor

Note: Operating systems or software versions that reached end-of-service status are not supported.

Support and usage of hardware functions depend on the operating system version and release. The information in this chapter is subject to change. For the most current information, see Preventive Service Planning for IBM Z.

This chapter describes the following topics:

- 6.1, “Software support” on page 80
- 6.2, “Support by operating system” on page 83
- 6.3, “Software licensing” on page 85

© Copyright IBM Corp. 2025, 2026.

6.1 Software support

The software portfolio for IBM z17 includes various operating systems and middleware that support the most recent technologies. IBM Z software engineering prioritizes hybrid cloud capabilities, robust security features, resilience, AI integration, and support for application modernization initiatives. For more information, see Mainframe software for IBM Z.

6.1.1 Application development and languages

Software developers can use the multiple programming language environments that run on the IBM Z platform. Compilers and development tool support for the IBM Z platform are continuously extended to provide developers with agile and modern development methods.

Application development for Linux on IBM Z mirrors that area for Linux on distributed systems; therefore, this chapter concentrates on providing insights into the z/OS environment.

In addition to the traditional COBOL, PL/I, Fortran, and assembly languages, the IBM Z platform supports the following programming languages:

- C
- C++
- Java (including Java Platform, Enterprise Edition, and batch environments)
- Go
- Swift
- Python
- JavaScript
- Node.js

IBM actively embraces open-source projects to extend z/OS functions. For example, the Zowe open-source project brings together industry experts to drive innovation for the community of next-generation mainframe developers. This framework enables an ecosystem of software solutions that provide a simple, intuitive environment for various IT professionals performing administrative, development, test, and operation tasks on z/OS.

The IBM z/OS Container Extensions (zCX) feature enables you to run Docker containers natively under z/OS. For more information, see z/OS Container Extensions.

An extensive set of advanced integrated development environments (IDEs) and integration tools are available for continuous development, testing, and deployment of application code.

The IBM DevOps offerings, such as IBM Application Delivery Foundation for z/OS and IBM Concert, can be coupled with IBM Application Discovery and Delivery Intelligence (ADDI) to enable developers to understand the applications, gain cognitive insights into the process, and evolve those valuable older assets at speed with reduced risk to the enterprise.

Modern development practices are supported by IBM Concert and open-source-based Git Version Control Tools for IBM z/OS (ported by Rocket Software), which are modern source code managers that run on and support z/OS.

For more information about software for the IBM Z platform, see the Products catalog.

Note: Using up-to-date compiler versions helps ensure access to the most recent advancements in IBM Z technology, including novel cache structures, machine instructions, and instruction execution enhancements, which collectively bolster performance gains.

IBM z17 inherits the features and functions from its predecessor, such as on-chip compression and sort, and single instruction multiple data (SIMD) architectural notation, which provides efficient vector processing.

Java applications benefit from Guarded Storage Facility, which enables pause-less garbage collection.

Operating systems that run on IBM z17 can use IBM Z System Recovery Boost (SRB) to accelerate recovery after an outage.

The following security functions were introduced to complement the IBM Z security stack:

- Secure Execution: Provides better isolation and security for second-level guest systems that are running under the KVM hypervisor for IBM Z servers.
- Secure Boot: A feature for verifying an open-source operating systems' kernel to help ensure that it comes from the trusted provider.
- Quantum-safe cryptography algorithms.
- Fully homomorphic encryption.
- Processor Activity Instrumentation to count cryptographic operations.
- z/OS Validated Boot IPL[1].

6.1.2 Supported IBM compilers

Each new version of the following IBM z/OS compilers 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 delivers improved performance and advanced optimization. It supports the following items:

  - COBOL to Java Interoperability (For example, Java can call COBOL and vice versa.)
  - UTF-8
  - 64-bit JSON support
  - JSON Boolean support for Parse and Generate
  - Binary coded decimal to hexadecimal floating point conversions
  - Numeric editing operation
  - Zoned decimal operations
- IBM Automatic Binary Optimizer for z/OS

  The Automatic Binary Optimizer for z/OS improves the performance of compiled COBOL programs. It does not require source code, source code migration, or the tuning of performance options. It uses modern optimization technology to accelerate the performance of COBOL objects that are compiled by VS COBOL II 1.3 and later.
- Enterprise PL/I provides the following functions:
  - More hardware use: Perform Timing Facility Function instructions
  - Enhanced diagnostics for **SUBSCRIPTRANGE**

[1] Also supported by IBM z/VM 7.5

- Modernization enhancements:
  - Enhanced **COMPARE** for XML and JSON.
  - Support for Extended Binary Coded Decimal Interchange Code (EBCDIC) JSON.

- z/OS XL C/C++

  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 with IBM z/OS Problem Determination Tools.

  The latest version has the following features:

  - High-performance math libraries:
    - Mathematical Acceleration Subsystem (MASS).
    - Replace IBM Atlas Policy Suite with OpenBLAS.
  - Metal C for modernization of IBM High Level Assembler applications
  - Neural Network Processing Assist (NNPA) facility

- Java

  The latest version of Java supports the following features:

  - On-chip IBM zEnterprise Data Compression (zEDC) support
  - Crypto: Elliptic Curve Digital Signature Algorithm and Elliptic Curve Diffie-Hellman (ECDH) acceleration
  - Zoned decimal operations in the Data Access Accelerator library for enhanced interoperability
  - Java enablement of Deep Learning Containers (DLC) models that use IBM Z Integrated Accelerator for Artificial Intelligence (Artificial Intelligence Unit (AIU))

  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 use IBM Developer for z/OS to help boost productivity by editing, compiling, and debugging z/OS XL C and XL C++ applications from the workstation.

- IBM Open Enterprise SDK for Python

  Also available on z/OS is the IBM Open Enterprise SDK for Python (current version is 3.13), which is based on the popular Python interpreter from Python Software Foundation:

  - z/OS Language Environment-based 64-bit only
  - Runs on z/OS UNIX System Services environment with no prerequisites
  - Supports ASCII, EBCDIC, and Unicode (UTF-8)
  - Includes selected Python Package Index (PyPI) packages

For more information, see IBM Open Enterprise SDK for Python.

6.2 Support by operating system

This section lists the support of in-service operating systems. The current and minimum operating system levels that are required to support IBM z17 are listed in Table 6-1.

Table 6-1 Operating system requirements

Operating system	End of service	Notes
z/OS 3.2	September 30, 2030	See IBM Fix Category Values.
z/OS 3.1	September 30, 2028	
z/OS 2.5	September 30, 2026	
z/VM 7.5	March 31, 2031	
z/VM 7.4	March 31, 2029	
z/VM 7.3	April 30, 2027	
z/TPF 1.1	Not announced	
Linux on IBM Z	Support information is available for SUSE, Red Hat, and Canonical.	For more information about minimal and recommended distribution levels, see the Linux distributors' websites.
KVM hypervisor		
21st Century Software VSE^{n} 6.3.1 and later (plus Program Temporary Fixes (PTFs))	Not announced.	For more information, see 21st Century Software.

Note: PTFs and Fix Categories (FIXCATs): The usage of several features depends on a particular operating system. In all cases, PTFs might be necessary for the operating system level that is indicated. Review the IBM Z hardware FIXCATs before the operating system is installed. PTFs for z/OS, and z/VM 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, see 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.

For more information about IBM z17 and its features, see *IBM z17 (9175) Technical Guide*, SG24-8579.

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

6.2.1 z/OS

As a pivotal IBM Z operating system, z/OS supports IBM z17. However, IBM z17 features and functions vary based on the specific z/OS version that is deployed.

The z/OS release cycle was extended with IBM Software Support Services so that you may purchase extended defect support service for previous versions of the operating system.

The supported versions of z/OS include z/OS 3.1 + PTFs, z/OS 2.5 + PTFs, and z/OS 2.4[2] + PTFs.

For more information about the IBM z17 features and functions that are supported by z/OS, see *IBM z17 (9175) Technical Guide*, SG24-8579.

For more information, see IBM Support, IBM z/OS, and IBM Resource Link.

6.2.2 z/VM

IBM z16 support is provided with PTFs for z/VM 7.3 and 7.4.

For more information about z/VM, see IBM z/VM and for PTF availability, see z/VM Continuous Delivery News.

For more information about the IBM z17 features and functions that are supported by z/VM, see *IBM z17 (9175) Technical Guide*, SG24-8579.

6.2.3 z/TPF

IBM z17 support is provided by z/TPF 1.1 plus PTFs.

For more information, see z/TPF.

For more information about the IBM z17 features and functions that are supported by z/TPF, see *IBM z17 (9175) Technical Guide*, SG24-8579.

[2] z/OS 2.4 requires a service extension and offers toleration support only (no support of new IBM z17 functions).

6.2.4 Linux on IBM Z

The Red Hat, SUSE, and Ubuntu releases that are supported on IBM z17 are listed in Table 6-2.

Table 6-2 Linux on IBM Z distributions

Linux on IBM Z distribution	Version and release
SUSE Linux Enterprise Server	SUSE Linux Enterprise Server 16 with service
SUSE Linux Enterprise Server	SUSE Linux Enterprise Server 15 SP6 with service
SUSE Linux Enterprise Server	SUSE Linux Enterprise Server 12 SP5 with service (not as a Secure Execution KVM host)
Red Hat Enterprise Server	RHEL 9.4 with service
Red Hat Enterprise Server	RHEL 8.10 with service
Red Hat Enterprise Server	RHEL 7.9 with service (not as a KVM host)
Canonical	Ubuntu 24.04 LTS with service
Canonical	Ubuntu 22.04 LTS with service
Canonical	Ubuntu 20.04.1 LTS with service

For more information about the IBM z17 features and functions that are supported by the Linux on IBM Z distributions, see *IBM z17 (9175) Technical Guide*, SG24-8579.

For more information about Linux on IBM Z, see Linux on IBM Z.

6.2.5 Kernel-based Virtual Machine hypervisor

For IBM z17, the KVM hypervisor is delivered and supported by Linux distribution partners. For more information about KVM support for the IBM Z platform, see the documentation for your distribution and *Virtualization Cookbook for IBM Z Volume 5: KVM*, SG24-8463.

6.3 Software licensing

The IBM Z software portfolio includes operating system software (that is, z/OS, z/VM, and z/TPF) and middleware that runs on these operating systems. The portfolio also includes middleware for Linux on IBM Z environments. For IBM z17, 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 permission to use the product, the charge includes access to IBM product support during the support period. MLC pricing applies to z/OS and z/TPF operating systems. Charges are based on processor capacity, which is measured in millions of service units (MSU) per hour.

- Tailored Fit Pricing for IBM Z for software

  Tailored Fit Pricing for IBM Z is a flexible software pricing model that dramatically simplifies the pricing landscape through flexible deployment options that are tailored to your IBM Z environment.

Two new pricing solutions, Enterprise Consumption and Enterprise Capacity, offer alternatives to the Rolling 4 Hour Average (R4HA)-based pricing model for new and existing workloads.

- IBM International Program License Agreement (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 you to access IBM product support during the support period. With this option, you can also receive future releases and versions at no extra charge.

For more information about software licensing, see IBM Software Licensing. For more information about IBM Z pricing, see Pricing on IBM Z.

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

Subcapacity pricing is available to clients running on the z/VM 7 platform. Software pricing at less than full machine capacity can provide more flexibility and improved cost of computing as a client manages the volatility and growth of new workloads.

For more information about subcapacity pricing terms for z/VM and z/VM based programs, see Sub-capacity pricing terms for z/VM. Select z/VM based programs help improve flexibility and price and performance.

For more information about software licensing options, see *IBM z17 (9175) Technical Guide*, SG24-8579.

Abbreviations and acronyms

ADDI	IBM Application Discovery and Delivery Intelligence
AIU	Artificial Intelligence Unit
ASHRAE	American Society of Heating, Refrigerating, and Air-Conditioning Engineers
ASIC	application-specific integrated circuit
BMC	Baseboard Management Controller
BPA	Bulk Power Assembly
BTS	Backup Time Server
CAT	IBM Crypto Analytics Tool
CBU	Capacity BackUp
CCA	IBM Common Cryptographic Architecture
CE LR	Coupling Express Long Reach
CE2 LR	Coupling Express2 Long Reach
CE3 LR	Coupling Express3 Long Reach
CF	Coupling Facility
CFCC	Coupling Facility Control Code
CIU	Customer Initiated Upgrade
CP	central processor
CPACF	Central Processor Assist for Cryptographic Functions
CPC	central processor complex
CPUMF	CPU Measurement Facility
CSS	channel subsystem
CST	Coordinated Server Time
CTN	Coordinated Timing Network
DCM	dual-chip module
DLC	Deep Learning Containers
DPU	Data Processing Unit
EBCDIC	Extended Binary Coded Decimal Interchange Code
ECC	Elliptic Curve Cryptography
ECDH	Elliptic Curve Diffie-Hellman
EDA	Enhanced Drawer Availability
EP11	Enterprise PKCS #11
EPO	Emergency Power Off
ETS	External Time Source
FC	Fibre Channel
FCP	Fibre Channel Protocol
FIPS	Federal Information Processing Standards
FIXCAT	Fix Category
HMA	Hardware Management Appliance
HSA	Hardware System Area
HSM	Hardware Security Module
IaC	infrastructure as code
IBM	International Business Machines Corporation
IBM TDz	IBM Threat Detection for z/OS
ICF	Internal Coupling Facility
ICSF	IBM z/OS Integrated Cryptographic Service Facility
IDE	integrated development environment
IFL	Integrated Facility for Linux
IFLs	Integrated Facilities for Linux
IFP	Integrated Firmware Processor
IPLA	IBM International Program License Agreement
ISM	Internal Shared Memory
KEM	key encapsulation mechanism
KVM	Kernel-based Virtual Machine
LSPR	Large Systems Performance Reference
MASS	Mathematical Acceleration Subsystem
MCU	memory controller unit
ML	machine learning
ML-DSA	Module-Lattice-Based Digital Signature
ML-KEM	Module-Lattice-Based Key-Encapsulation Mechanism
MLC	monthly license charge
MSU	millions of service units
MTP	Multi-Fiber Termination Push-On
NNPA	Neural Network Processing Assist
NTP	Network Time Protocol
NTS	Network Time Security
OSA-Express	Open Systems Adapter-Express
OSC	oscillator card
PAV	Parallel Access Volume

© Copyright IBM Corp. 2025, 2026.

PCI HSM	Payment Card Industry HSM
PCIeCC	PCIe Cryptographic Coprocessor
PPC	Processor Power Card
PPS	pulse per second
PR/SM	IBM Processor Resource/Systems Manager
PRNG	Pseudo-Random Number Generator
PSU	power supply unit
PTF	Program Temporary Fix
PTP	Precision Time Protocol
PTS	Primary Time Server
PU	processor unit
R4HA	Rolling 4 Hour Average
RAG	retrieval augmented generation
RAIM	redundant array of independent memory
RNI	relative nest intensity
SAP	System Assist Processor
SE	Support Element
SHA	Secure Hash Algorithm
SIMD	single instruction multiple data
SMC	Shared Memory Communication
SMC-D	Shared Memory Communication-Direct Memory Access
SMC-R	Shared Memory Communication over Remote Direct Memory Access
SMP	symmetric multiprocessing
SMT	simultaneous multithreading
SRAM	static random access memory
SRB	IBM Z System Recovery Boost
STP	Server Time Protocol
TEE	Trusted Execution Environment
TFP-Hardware	Tailored Fit Pricing for IBM Z Hardware Consumption Solution
TKE	Trusted Key Entry
TOPS	trillion operations per second
TPS	transaction processing systems
TRNG	True Random Number Generator
VFM	IBM Virtual Flash Memory
zCX	IBM z/OS Container Extensions
zEDC	IBM zEnterprise Data Compression
zERT	IBM z/OS Encryption Readiness Technology
zIIP	IBM Z Integrated Information Processor
zPCR	IBM Z Processor Capacity Reference
zSCC	IBM Z Security and Compliance Center

Redbooks

IBM z17 Technical Introduction

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

SG24-8580-01

ISBN 0738462675

Printed in U.S.A.

Get connected