IBM LinuxONE Emperor 5

Unlocking innovation in enterprise IT

Highlights

End-to-end cybersecurity and privacy

Optimized IT for energy and cost savings

Built-in AI engineered for better outcomes



As enterprises face an escalating wave of cyberthreats, IBM® LinuxONE Emperor 5 arrives with state-of-the-art security capabilities, including **end-to-end data encryption, confidential computing**, and NIST-standardized **post-quantum cryptography.** The platform is designed to meet the highest standards of compliance, enabling businesses in regulated industries—such as finance, healthcare, and government—to secure sensitive data both at rest and in use. With IBM Secure Execution, LinuxONE Emperor 5 isolates workloads at the hardware level, providing protection against insider threats and unauthorized access. With up to 99.999999%[1] availability, industry leading performance and vertical scalability, IBM LinuxONE Emperor 5 is designed to reduce the physical server footprint, translating to significant energy and other resource cost savings. IBM LinuxONE Emperor 5 redefines how enterprises harness artificial intelligence by embedding AI acceleration directly into the platform. Businesses can now seamlessly deploy traditional and generative AI models right where their data resides, minimizing latency and maximizing insight, all with scale and security.

IBM LinuxONE Emperor 5 redefines the Linux® enterprise IT landscape, offering businesses the chance to achieve more with less through the integration of end-to-end security and advanced AI with data and applications. Built for a world where data protection and actionable insights are mission-critical, IBM LinuxONE Emperor 5 delivers a secure, AI-ready foundation for businesses navigating today's complex IT landscape. Ultimately, IBM LinuxONE Emperor 5 unlocks the full potential of your Linux infrastructure, driving real business value. It enhances transactional workloads, boosts data-serving capabilities, and promotes overall infrastructure efficiency, setting a new industry standard. IBM LinuxONE Emperor 5 stands as a beacon of innovation, empowering businesses globally by transforming their IT infrastructure into a robust driver of growth and value creation. With unmatched performance and security, IBM LinuxONE Emperor 5 sets businesses apart in today's competitive landscape. There's only one IBM LinuxONE.

↑ 35 billion

encrypted requests per day[2]

↑ 99.999999% availability

equivalent to 315 ms of downtime per year[3]

End-to-end cybersecurity and privacy

IBM LinuxONE Emperor 5 is a platform that drives business innovation through advanced security features. It utilizes integrated crypto accelerators to protect data at rest and in transit without impacting performance. The platform standardizes and simplifies data-in-use protection, integrates with IBM Vault Self-Managed and securely orchestrates the lifecycle of encryption keys across the enterprise. It also provides a single interface for both on-premises and multi-cloud protection, allowing for scaling and unifying encryption across the enterprise. Additionally, IBM LinuxONE facilitates the deployment of confidential containers to protect data and applications and addresses post-quantum cybersecurity risks through pioneering quantum-safe encryption. The platform leverages Linux crypto counters to highlight vulnerable cryptography and utilizes NIST-standardized post-quantum cryptography. It also provides confidential computing with integrated acceleration for AI, enabling the protection of AI models, data and applications.

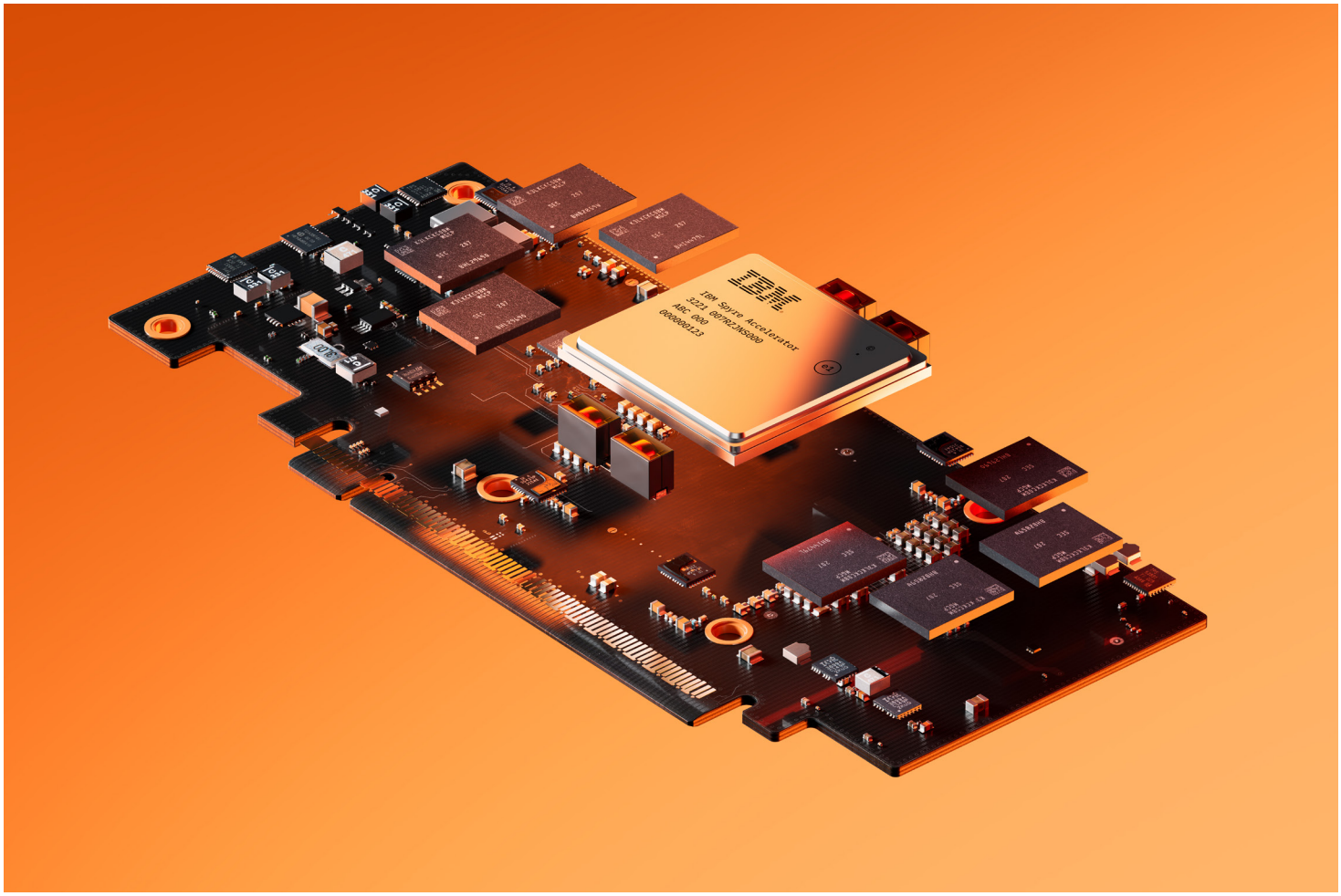
Optimized IT for energy and cost savings

IBM LinuxONE Emperor 5 is designed to provide high availability, flexibility and scalability while reducing costs and power consumption. It can achieve eight nines availability, making it suitable for businesses that require continuous uptime.[3] Organizations can save up to 94% in software costs over 5 years by moving cloud-native, containerized workloads from a compared x86 solution to IBM LinuxONE Emperor 5 running the same software products.[4] Additionally, the system can do the work of up to 2,944 cores of a compared x86 solution, making it a highly efficient option for data centers.[5]

Built-in AI engineered for better outcomes

IBM's AI accelerators provide a security-rich and efficient solution for building and deploying AI applications in a hybrid cloud environment. The benefits include trusted execution environments for protecting sensitive data, prebuilt training data and post-quantum cryptography for enhanced security. The on-chip AI accelerator on the IBM Telum® II processor enables near real-time scoring, gen AI and low-latency inferencing, while supporting multiple AI models to improve prediction accuracy. With up to 256 additional accelerator cores and 1 TB memory, the platform's scalable architecture and optimized energy consumption contribute to improved business efficiency and reduced operational costs. IBM LinuxONE Emperor 5 is designed to process up to 24 trillion operations per second (TOPS) shared across all cores on the chip[6] with IBM Integrated Accelerator for AI at full utilization.

The IBM Telum II data processing unit (DPU) reduces the power needed for input/output (I/O) management for a large IBM LinuxONE Emperor 5 system by over 90% compared to a similarly configured IBM LinuxONE Emperor 4.[7]



IBM LinuxONE Emperor 5 on-chip and off-chip innovative technology

IBM Telum II Processor

AI acceleration on chip coprocessor
The IBM Telum II processor integrates improved AI acceleration through an on-chip AI coprocessor to reduce latency and deliver outstanding performance for in-transaction inferencing. It now supports small language models (SLMs) where the number of variables is less than 8 billion. Organizations can embed AI directly into business processes and existing IBM LinuxONE Emperor 5 applications to help improve business outcomes and deliver customer value in each interaction at unprecedented scale and speed within stringent SLA response times.

I/O acceleration unit
A completely new DPU on the IBM Telum II processor chip is engineered to accelerate complex I/O protocols for networking and storage on IBM LinuxONE Emperor 5. The DPU simplifies system operations and can improve key component performance.

IBM Spyre Accelerator
The IBM Spyre™ Accelerator card, which will be available in 4Q 2025 via a PCIe card, will provide additional AI compute capability to complement the IBM Telum II processor. This component will extend and scale the AI capabilities of IBM LinuxONE Emperor 5 by providing compute to support gen AI use cases that need unstructured data such as text. Up to 48 Spyre™ Accelerators will be supported to scale gen AI for enterprise workloads that demand exceptional performance and rigorous security and resilience.

Configuration Table

IBM LinuxONE Emperor 5 Overview	**IBM LinuxONE Emperor 5 ML1** This configuration is designed for general-purpose use, offering a balance of performance, scalability, security and efficiency suitable for a wide range of applications.
Specifications	
Maximum number of engines	208
Maximum number of drawers	4
Maximum number of I/O drawers	12
Number of frames	4
Collocate with storage/switch	No
Frequency	5.5 GHz
Telum chip	Yes
Maximum memory	64 TB
Sizes	43, 90, 136, 183 and 208
Resources	
Specification sheets	LinuxONE Emperor 5 specifications
Energy efficiency	LinuxONE Emperor 5 carbon footprint

To learn more about IBM LinuxONE Emperor 5, contact your IBM Business Partner:

System Analysis Services, INC.

6176316946 | carl@sysgbs.com

http://systemsanalysisservices.com/

[1]**DISCLAIMER**: IBM internal data based on measurements and projections was used in calculating the expected value. Necessary components include IBM LinuxONE Emperor 5; IBM z/VM V7.3 systems or above collected in a Single System Image, each running RHOCP 4.14 or above; IBM Operations Manager; GDPS 4.6 or above for management of data recovery and virtual machine recovery across metro distance systems and storage, including Metro Multi-site workload and GDPS Global; and IBM DS8000 series storage with IBM HyperSwap. A MongoDB v4.4 workload was used. Necessary resilience technology must be enabled, including z/VM Single System Image clustering, GDPS xDR Proxy for z/VM, and Red Hat® OpenShift® Data Foundation (ODF) 4.14 or above for management of local storage devices. Application-induced outages are not included in the above measurements. Other configurations (hardware or software) may provide different availability characteristics.

[2]**DISCLAIMER**: IBM internal performance tests for the core consolidation study targeted a comparison of the following two servers. IBM Machine Type 9175 MAX 136 system consisting of three CPC drawers containing 136 configurable processor units and six I/O drawers to support both network and external storage. The x86 solution, which used a commercially available enterprise server with two fifth-generation Intel® Xeon® Platinum 8592+ processors, 64 cores per CPU. Both solutions had access to the same storage. The workloads consisted of a containerized online transaction processing (OLTP) WebSphere Liberty v25 application running on Red Hat OpenShift Container Platform (OCP) v4.17 and an EDB Postgres for Kubernetes v1.25 on the same OCP cluster simulating core online banking functions. Both solutions used Red Hat Enterprise Linux v9.5 and KVM. Results may vary.

The test results were extrapolated to a typical, complete customer IT solution that includes production and non-production IT environments isolated from each other. TCO included software, hardware, energy, network, data center space and labor costs. On the IBM z17 side, the complete solution requires one IBM z17 Type 9175 MAX 136, and on the x86 side, the complete IT solution requires 23 compared servers.

[3]**DISCLAIMER**: IBM internal data based on measurements and projections was used in calculating the expected value. Necessary components include IBM LinuxONE Emperor 5; IBM z/VM V7.3 systems or above collected in a Single System Image, each running RHOCP 4.14 or above; IBM Operations Manager; GDPS 4.6 or above for management of data recovery and virtual machine recovery across metro distance systems and storage, including Metro Multi-site workload and GDPS Global; and IBM DS8000 series storage with IBM HyperSwap. A MongoDB v4.4 workload was used. Necessary resilience technology must be enabled, including z/VM Single System Image clustering, GDPS xDR Proxy for z/VM, and Red Hat OpenShift Data Foundation (ODF) 4.14 or above for management of local storage devices. Application-induced outages are not included in the above measurements. Other configurations (hardware or software) may provide different availability characteristics.

[4]**DISCLAIMER**: IBM internal performance tests for the core consolidation study targeted a comparison of the following two servers. IBM Machine Type 9175 MAX 136 system consisting of three CPC drawers containing 136 configurable processor units and six I/O drawers to support both network and external storage. The x86 solution, which used a commercially available enterprise server with two fifth-generation Intel® Xeon® Platinum 8592+ processors, 64 cores per CPU. Both solutions had access to the same storage. The workloads consisted of a containerized online transaction processing (OLTP) WebSphere Liberty v25 application running on Red Hat OpenShift Container Platform (OCP) v4.17 and an EDB Postgres for Kubernetes v1.25 on the same OCP cluster simulating core online banking functions. Both solutions used Red Hat Enterprise Linux v9.5 and KVM. Results may vary.

The test results were extrapolated to a typical, complete customer IT solution that includes production and non-production IT environments isolated from each other. TCO included software, hardware, energy, network, data center space and labor costs. On the IBM z17 side, the complete solution requires one IBM z17 Type 9175 MAX 136, and on the x86 side, the complete IT solution requires 23 compared servers.

[5]**CLAIM**: IBM internal tests simulating a complete IT solution running containerized WebSphere Liberty and EDB Postgres workloads show that a single IBM LinuxONE Emperor 5 Max 136 can do the work of up to 2,944 cores of the compared x86 solution.

[5]**DISCLAIMER**: IBM internal performance tests for the core consolidation study targeted a comparison of the following two servers. IBM Machine Type 9175 MAX 136 system consisting of three CPC drawers containing 136 configurable processor units and six I/O drawers to support both network and external storage. The x86 solution, which used a commercially available enterprise server with two fifth-generation Intel® Xeon® Platinum 8592+ processors, 64 cores per CPU. Both solutions had access to the same storage. The workloads consisted of a containerized online transaction processing (OLTP) WebSphere Liberty v25 application running on Red Hat OpenShift Container Platform (OCP) v4.17 and an EDB Postgres for Kubernetes v1.25 on the same OCP cluster simulating core online banking functions. Both solutions used Red Hat Enterprise Linux v9.5 and KVM. Results may vary.

The test results were extrapolated to a typical, complete customer IT solution that includes production and non-production IT environments isolated from each other. TCO included software, hardware, energy, network, data center space and labor costs. On the IBM z17 side, the complete solution requires one IBM z17 Type 9175 MAX 136, and on the x86 side, the complete IT solution requires 23 compared servers.

[6]**DISCLAIMER**: Result is the maximum theoretical number of trillion operations per second (TOPS) in 8-bit precision that can be executed by a single IBM Integrated Accelerator for AI. Cores are running at 5.5GHz and have one IBM Integrated Accelerator for AI per chip. The IBM Integrated Accelerator for AI consists of 2 corelets, each with an array of 64 tensor cores capable of executing 4 integer-multiply-add operations (IMA) 8-way SIMD with no sparsity.

[7]**DISCLAIMER**: Comparison based on IBM lab measurements for the difference in power required for supporting I/O for FICON and OSA in an expected large IBM Machine Type 9175 configuration based on an actual historical large IBM Machine Type 3931 configuration. IBM Machine Type 9175 is Max 208 with a 23 TB memory, 56 active processors, 3 IBM Virtual Flash Memory, 14 ICA-SR 2.0, 7 PCIe+ I/O drawers with 69 FICON Express32 – 4P LX, 12 OSA-Express7S 1.2 GbE SX, 18 Network Express LR 10G and 4 Crypto Express 8S (2 HSMs). IBM Machine Type 3931 is configured to provide the same hardware capability. Results may vary.

© Copyright IBM Corporation 2025

IBM, the IBM logo, IBM Spyre, Spyre, IBM Telum, IBM Z, IBM z/OS, and IBM z17 are trademarks or registered trademarks of International Business Machines Corporation, in the United States and/or other countries. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on ibm.com/legal/copyright-trademark.

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.

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

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

Examples presented as illustrative only. Actual results will vary based on client configurations and conditions and, therefore, generally expected results cannot be provided.

Statement of Direction: IBM’s statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM’s sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

