IBM

IBM zSystems running Red Hat OpenShift Container Platform

Modernize for hybrid cloud – the cloud platform you want, with resiliency, security, and sustainability, without compromise

IBM zSystems

With IBM zSystems™ you can innovate with flexibility and agility and accelerate your modernization as you integrate IBM zSystems seamlessly into your hybrid cloud.

IBM zSystems technology and its cloud solutions empower developers with the agility to accelerate the development of cloud-native workloads and to modernize existing workloads, which can be integrated with digital services across the hybrid cloud – while keeping the data safe, encrypted and resilient.

The IBM z16™ product portfolio – IBM z16 multi frame, IBM z16 single frame, and IBM z16 rack mount – delivers breakthrough technologies with the highest security, performance and availability.

Red Hat® OpenShift® Container Platform

The Red Hat OpenShift Container Platform on IBM zSystems empowers your organization to integrate and modernize your applications with agility through integrated tooling and a secure and resilient foundation for cloud-native development and deployment.

Red Hat OpenShift is a trusted Kubernetes enterprise platform that supports, similar to IBM Cloud Paks®, modern, hybrid-cloud application development and a consistent foundation for enterprise applications anywhere – across physical, virtual, private, and public clouds.

Reduce your carbon footprint while improving your efficiency

IBM z16 is designed to make a powerful improvement in sustainability by decreasing electricity consumption, reducing the number of standing servers, and enabling high compute and resource utilization.

IBM z16 design is aligned with best practices for reducing electricity consumption, including that clients require a small number of physical systems, that IBM z16 is high energy-efficiency system, and that it enable high compute and resource utilization.

Consolidating Linux workloads on an IBM z16 single frame or rack mount instead of running them on compared x86 servers with similar conditions and location can reduce energy consumption by 75% and space by 67%.[1]

Running workloads on a centralized infrastructure such as IBM zSystems can contribute to fewer greenhouse gas emissions and a more environmentally sustainable IT environment.

Consolidating Linux® workloads on 5 IBM z16 systems instead of running them on compared x86 servers under similar conditions can reduce energy consumption by 75%, space by 50%, and the CO2e footprint by over 850 metric tons annually. This is equivalent to consuming about 362,000 fewer litres (95,600 gallons) of gasoline each year.[2]

Highlights

- Reduced energy consumption and physical footprint
- Data protection and privacy at scale through confidential computing
- Low latency and high throughput via colocation
- Consistent service delivery with a massively scalable system
- Reduced cost of computing while delivering superior quality of service

"The bank needed to increase their competitive business offerings by extending and modernizing integration with existing assets while optimizing SLAs and minimizing risk."

Data protection and privacy at scale

IBM z16 represents a breakthrough in data security. Quantum-safe cryptography is embedded in the system, along with classical cryptography, to protect against quantum-computing attacks now and in the future.

IBM z16 quantum-safe technology and key management services, were developed to help you protect data and keys against a potential future quantum attack like harvest now, decrypt later.[3]

IBM z16 represents a step forward, as clients have a safe and tested infrastructure that can deploy the more sophisticated and complex cryptography needed to protect sensitive data from cyber risks.

Integrity and confidentiality of data is protected with Crypto Express adapters (Hardware Security Modules) designed to meet the strong security requirements of Federal Information Processing Standards (FIPS) 140-2 Level 4.

Red Hat OpenShift containers on IBM zSystems can take advantage of the Crypto Express adapters via the 'Kubernetes device plug-in for IBM Crypto Express cards', available as a certified image in the Red Hat catalog.

Generate up to 100,000 certificates per second using protected keys exploiting Crypto Express 8S adapters running application pods on Red Hat OpenShift Container Platform on a single IBM z16 multi frame drawer.[4]

IBM z16 running Red Hat OpenShift is designed with 99.9999999 (7 9's) application availability, seamless on-demand scalability, and to execute disaster recovery actions to respond to unplanned events. Flexible capacity for cyber resiliency on IBM z16 enables you to transfer capacity easily and efficiently between different data centers for disaster recovery, regulatory compliance, maintenance, and other business needs.

Combined with IBM storage capabilities, this function delivers an extremely high availability solution for mission critical workloads.

IBM z16 multi frame and IBM z16 single frame, with IBM GDPS®, IBM DS8000® series storage with IBM HyperSwap®, and running a Red Hat OpenShift Container Platform environment, are designed to deliver 99.99999% availability.[5]

IBM z16 contains new capabilities to make compliance to PCI-DSS regulatory guidelines easier and more productive. Audit preparation times can be significantly reduced and require less staff to complete.

IBM z16 is integrated with the 'IBM Z Security and Compliance Center' to monitor and record system, network and application data for changes.

On IBM z16, use the 'IBM Z Security and Compliance Center' to run compliance scans on PostgreSQL without impacting the throughput of OLTP microservice applications running on Red Hat OpenShift.[6]

The unique combination of IBM zSystems plus Red Hat OpenShift creates a differentiated, security-rich solution.

Low latency and high throughput via colocation

Co-locating Red Hat OpenShift workloads side-by-side with IBM z/OS®, Linux, IBM z/VSE®, or z/TPF workloads on IBM zSystems, benefits not only from low latency, high throughput, and operational efficiency, but also leverages investments in existing applications.

When accessing your database while running an OLTP workload on Red Hat OpenShift Container Platform, achieve 4.2x more throughput by co-locating the workload on IBM z16 multi frame versus running the workload on compared x86 platform connecting remotely to the IBM z16 multi frame.[7]

Colocation on IBM zSystems can result in a streamlined IT infrastructure with few points of attack. This unique IBM zSystems capability is supported by high levels of scalability with granular sharing and certified workload isolation and protection.

On digital currency transactions run inferencing for fraud 85% faster by co-locating your application with Snap ML on IBM z16 multi frame versus running inferencing remotely using Scikit-learn on a compared x86 server.[8]

Deliver consistent service with a massively scalable system

IBM zSystems helps to avoid or recover from failures to minimize business disruptions, realized through component reliability, redundancy and features that assist in providing fault avoidance and tolerance, as well as permitting concurrent maintenance and repair.

IBM z16 is designed to deliver superior performance for mission-critical applications in transaction processing, data sharing and mixed workloads, where nothing can be compromised. The system is massively scalable with the ability to add capacity on demand and grow processing with minimal impact to energy usage, floor space and staffing.

IBM z16 is architected for balanced performance with multiple layers of cache, massive I/O capabilities, and integrated accelerators to drive high utilization and processor efficiency.

With IBM z16 multi frame, execute up to 20 billion HTTPS transactions per day with OLTP microservice applications running on Red Hat OpenShift.[9]

Organizations can scale containers on a single IBM z16 for nondisruptive vertical and horizontal growth to accommodate increases of workloads on demand. Resources can be shared and prioritized dynamically and efficiently between workloads, providing agility by delivering them whenever and wherever they are needed.

The IFL processors on the IBM z16 single frame also provide an optional multi-threading technology capability; with the multi-threading function enabled, the performance capacity of an IFL is expected to typically be up to 25% higher (range typically 10-40%) than without the multi-threading function enabled.[10]

On IBM z16 single frame, scale-out to 192 Red Hat OpenShift Container Platform Compute Nodes and deploy up to40.000 NGINX pods.

On IBM z16 multi frame, run the IBM Blockchain Platform deployed on Red Hat OpenShift Container Platform with on average 2x more throughput vs on compared x86 system. [12]

IBM z16 integrates new Artificial Intelligence (AI) acceleration via an on-chip AI hardware accelerator to reduce latency and deliver outstanding performance for in transaction inferencing.

Using one Integrated Accelerator for AI on an OLTP workload on IBM z16 multi frame matches the throughput of a compared remote x86 server running inferencing on 18 cores.[13]

IBM z16 allows for high workload density, usually resulting in a streamlined infrastructure with fewer components, lower management effort, and fewer software licenses compared to other platforms.

Reduced cost and increased ROI

Considering all the aspects mentioned above – reduced carbon footprint, privacy and protection, colocation benefits, consistent service with a scalable system – it seems obvious that they can also provide an economic advantage when running Red Hat OpenShift on IBM zSystems compared to other platforms.

IBM zSystems, integrated seamlessly into your hybrid cloud, can provide benefits that span across operations, cost optimization, business growth, and can accelerate your modernization for hybrid cloud.

Why IBM?

As you transform your business and differentiate yourself in a trust economy, IBM remains your partner. We have the total expertise in systems, software, delivery, and financing to help you create a secure and intelligent foundation for your on-premises cloud on IBM zSystems.

Our experts can help you design, configure, and implement Red Hat OpenShift, as well as IBM Cloud Paks®.

For more information

To learn more about IBM z16, IBM zSystems, and Red Hat OpenShift, please contact your IBM representative, your Red Hat representative, or IBM Business Partner®.

1. *Compared IBM Machine Type 3932 Max 68 model consisting of a CPC drawer and an I/O drawer to support network and external storage with 68 IFLs and 7 TB of memory in 1 frame versus compared 36 x86 servers (2 Skylake Xeon Gold Chips, 40 Cores) with a total of 1440 cores. IBM Machine Type 3932 Max 68 model power consumption was measured on systems and confirmed using the IBM Power estimator for the IBM Machine Type 3932 Max 68 model configuration. x86 power values were based on Feb. 2023 IDC QPI power values and reduced to 55% based on measurements of x86 servers by IBM and observed values in the field. The x86 server compared to uses approximately .6083 KWhr, 55% of IDC QPI system watts value. Savings assumes the Worldwide Data Center Power Utilization Effectiveness (PUE) factor of 1.55 to calculate the additional power needed for cooling. PUE is based on Uptime Institute 2022 Global Data Center Survey (https://uptimeinstitute.com/resources/research-and-reports/uptime-institute-global-data-center-survey-results-2022). x86 system space calculations require 3 racks. Results may vary based on client-specific usage and location.*
2. *Compared 5 IBM z16 Max 125 model consists of three CPC drawers containing 125 configurable cores (CPs, zIIPs, or IFLs) and two I/O drawers to support both network and external storage versus 192 x86 systems with a total of 10364 cores. IBM z16 power consumption was based on inputs to the IBM z16 IBM Power Estimation Tool for a memo configuration. x86 power consumption was based on March 2022 IDC QPI power values for 7 Cascade Lake and 5 Ice Lake server models, with 32 to 112 cores per server. All compared x86 servers were 2 or 4 socket servers. IBM Z and x86 are running 24x7x365 with production and non-production workloads. Savings assumes a Power Usage Effectiveness (PUE) ratio of 1.57 to calculate additional power for data center cooling. PUE is based on Uptime Institute 2021 Global Data Center Survey (https://uptimeinstitute.com/about-ui/press-releases/uptime-institute-11th-annual-global-data-center-survey). CO2e and other equivalencies that are based on the EPA GHG calculator (https://www.epa.gov/energy/greenhousegas-equivalencies-calculator) use U.S. National weighted averages. Results may vary based on client-specific usage and location.*
3. *IBM z16 with the Crypto Express 8S card provides hardware enabled quantum-safe APIs. The quantum-safe public key technology used in IBM z16 has been summitted to the PQC standardization process conducted by NIST. https://csrc.nist.gov/Projects/post-quantum-cryptography/round-3-submissions*
4. *Performance results is extrapolated from an IBM internal study designed to replicate secure certificate generation with Java on Red Hat OpenShift Container Platform (RHOCP) 4.10 on IBM z16 A01 using KVM. 2 microbenchmark pods (Signature RSA 2048 key is protected by AES master key in Crypto Express 8S adapters, certificate signatures are done with SHA-256 for x509 certificates) were run in parallel per compute node each driven locally with 20 parallel threads. IBM z16 A01 configuration: The RHOCP Management and Compute nodes ran on RHEL 8.5 KVM using mac VTap in a LPAR with 24 dedicated IFLs, 256 GB memory, IBM FlashSystem 9200 storage, CEX8S adapters in "2 HSM version" mode, one HSM per compute node. Packages used for benchmark: IBM Semeru Open 11 JDK 11.0.14.1.1 0.30.1-1 using Bouncy Castle packages for x509 certificate generation and SunPKCS11 JCE provider connected to Open cryptoki 3.16.0 CCA token for cryptographic operations. Results may vary.*
5. *IBM internal data based on measurements and projections was used in calculating the expected value. Necessary components include IBM z16; IBM z/VM V7.2 systems collected in a Single System Image, each running Red Hat OpenShift Container Platform 4.10 or above; IBM Operations Manager; GDPS 4.5for 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 v4.2workload was used. Necessary resiliency technology must be enabled, including z/VM Single System Image clustering, GDPS xDR Proxy for z/VM, and Red Hat OpenShift Data Foundation 4.10 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.*
6. *Performance result is based on IBM internal tests running on an IBM z16 LPAR with 24 dedicated IFLs, 560 GB memory and DASD storage an OLTP workload on Red Hat OpenShift Container Platform (RHOCP) 4.10 using RHEL 8.4 KVM. On 4 RHOCP Compute nodes 4 OLTP workload instances were running in parallel, each driven remotely from JMeter 5.2.1 with 128 parallel users. The KVM guests with RHOCP Compute nodes for the OLTP workload were configured with 12 vCPUs and 64 GB memory each. The KVM guests with RHOCP Management nodes and RHOCP Infrastructure nodes were configured with 4 vCPUs and 16 GB memory each. Compliance Operator deployed a configured from (https://www.ibm.com/docs/en/zscc/1.1.1). Results may vary.*
7. *This is an IBM internal study designed to replicate banking OLTP workload usage in the marketplace deployed on Red Hat OpenShift Container Platform (RHOCP) 4.9 on IBM z16 A01 using z/VM versus on compared x86 platform using KVM accessing the same PostgreSQL 12 database running in an IBM z16 A01 LPAR. IBM z16 A01 configuration: The PostgreSQL database ran in a LPAR with 12 dedicated IFLs, 128 GB memory, 1 TB FlashSystem 900 storage, RHEL 7.7 (SMT mode). The Compute nodes ran on z/VM 7.2 in a LPAR with 30 dedicated IFLs, 188 GB memory, DASD storage, and OSA connection to the PostgreSQL LPAR. LPAR with 2 IFL, 4 GB memory and RHEL 8.5 with RHOCP Proxy server. x86 configuration: The Compute nodes ran on KVM on RHEL 8.5 on 32 Cascade Lake Intel® Xeon® Gold CPU @ 2.30 GHz with Hyperthreading turned on, 192 GB memory, RAID5 local SSD storage, and 10Gbit Ethernet connection to the PostgreSQL LPAR. Results may vary.*
8. *Performance results based on IBM internal tests doing inferencing using a Scikit -learn Random Forest model with Snap ML v1.9.0 (tech preview) backend on IBM z16 A01 and with Scikit-learn v1.0.2 backend on compared x86 server. The model was trained on the following public dataset https://www.kaggle.com/datasets/ellipticco/elliptic-data-set. BentoML v0.13.1 (https://github.com/bentoml/BentoML) was used on both platforms as model serving framework. IBM z16 A01 configuration: Ubuntu 20.04 in an LPAR with 2 dedicated IFLs, 256 GB memory. x86 configuration: Ubuntu 20.04 on 9 IceLake Intel® Xeon® Gold CPU @ 2.80GHz with Hyperthreading turned on, 1 TB memory.*
9. *Performance result is extrapolated from IBM internal tests running in an IBM z16 A01 LPAR with 24 dedicated IFLs, 560 GB memory and DASD storage the Acme Air microservice benchmark (https://github.com/blueperf/acmeair-mainservice-java) on Red Hat OpenShift Container Platform (RHOCP) 4.9 using RHEL 8.4 KVM. On 4 RHOCP Compute nodes 4 Acme Air instances were running in parallel, each driven remotely from JMeter 5.2.1 with 384 parallel users. The KVM guests with RHOCP Compute nodes were configured with 12 vCPUs and 64 GB memory each. The KVM guests with RHOCP Management nodes and RHOCP Infrastructure nodes were configured with 4 vCPUs and 16 GB memory each. Results may vary.*
10. *Based on internal measurements. Results may vary by customer based on individual workload, configuration and software levels. Visit LSPR website for more details at: https://www-40.ibm.com/servers/resourcelink/lib03060.nsf/pages/lsprindex*
11. *Performance result is extrapolated from IBM internal tests running in an IBM Machine Type 3932 LPAR with 11 dedicated IFLs, 768 GB memory and FS9200 storage NGINX pods on Red Hat OpenShift Container Platform (RHOCP) 4.12 running on a RHEL 8.5 KVM host. 32 RHOCP Compute nodes with 220 NGINX pods were running in parallel. The KVM guests with RHOCP Compute nodes were configured with 2 vCPUs and 16 GB memory each. The KVM guests with RHOCP Management nodes were configured with 16 vCPUs and 128 GB memory each. Results may vary.*
12. *Performance results based on IBM internal tests running IBM (IBP) 2.5.2 on OpenShift Container Platform 4.9 running on an IBM z16 A01. Tests were run with sample cc chain code (https://github.com/hyperledger/fabric-test/tree/master/chaincodes/samplecc/go). IBP resources were defined as follows: 6 peers each with 0.25 -2 vCPU, 1 GB memory, 6 couch DB each with 0.375 -3 vCPU, 400 MB memory, 5 orderer each with 0.5 vCPU, 500 MB memory, 6 smart contract launcher each with 0.2 vCPU, 400 MB memory. Peer and couch DB vCPU resources scaled up proportionally with each data point. IBM z16 A01 configuration: 1 LPAR configured with 32 dedicated IFLs, 192 GB memory, RHEL 8.4 KVM. x86 configuration: 32 Intel® Xeon® Gold 5218 CPU @ 2.30GHz w/ Hyperthreading turned on, 192 GB memory, RHEL 8.4 KVM. Results may vary.*
13. *Performance results is extrapolated from IBM internal tests running an OTLP workload with credit card transaction using the Credit Card Fraud Detection (https://github.com/IBM/ai-on-z-fraud-detection) model on IBM z16 A01 vs running the OLTP workload https://github.com/IBM/megacard-standalone) on IBM z16 A01 and running inferencing on a remote x86 server running Tensorflow serving. IBM z16 A01 configuration: Ubuntu 20.04 in an LPAR with 6 dedicated IFLs, 256 GB memory, and IBM FlashSystem 900 storage. x86 configuration: Ubuntu 20.04 on 18 IceLake Intel® Xeon® Gold CPU @ 2.80GHz with Hyperthreading turned on, 1 TB memory, local SSDs. Results may vary.*

Learn more:

[IBM z16]

[IBM zSystems]

[Red Hat OpenShift]