The Advantages of Using Linux and Open Source Solutions

Delivering business agility to support the demands of the new application economy

IBM

ENTER »

1

Open Collaboration

Collaboration in a rich open source ecosystem

2

More Choice

More choice and speed with contemporary programming languages

3

Boosted Performance

Database performance improvements

4

Better Insight

Advancements in large-scale data processing

5

Easy Portability

Application portability and infrastructure flexibility

Better insights, faster response times, improved business agility

The mobile world and new app economy is driving change with unprecedented speed and scope—impacting the world's most powerful economies and developing nations, changing the game for business, industries and society at large, creating new requirements and expectations that have to be addressed.

In a world where consumer expectations continue to grow, the ability to keep pace with new opportunity and to differentiate through more innovation means your IT department is key to competing in the immediacy of a digital world. The criteria for success lies in your ability to:

- deliver at speed with lighting fast response times, all the time;
- provide access anytime, anywhere, and from any device, no matter how many users or transactions;
- ensure security for trusted engagements;
- enable personalization through analytics;
- achieve new levels of IT efficiency, agility and responsiveness using a cloud model.

The starburst effect on transaction growth rates from mobile can be dramatic, placing even greater pressure on IT. As a result, you need to take Linux to the next level where it can fully support the future of high-volume business-critical applications.

The intersection of traditional IT, the new application economy, open technology and community collaboration are central to serving these new requirements, redefining IT operational excellence and ultimately the customer experience.

Open source software applications and programs are bringing a broad set of capabilities and new solutions to IBM LinuxONE and z Systems—new programming languages and run-time environments, relational and NoSQL databases, more choices in big data analytics, and container technologies are becoming integrated into the enterprise IT tool box.

When these capabilities are combined with unique IBM LinuxONE and z Systems performance features—the ability to run 141 cores at 5.0 GHz and 10 TB of memory in one system, industry-leading I/O bandwidth and data processing throughput, EAL5+ security, and high-speed data compression and cryptography hardware—your business will obtain better insights, faster response times, and improved business agility.

As the open source ecosystem expands for both LinuxONE and z Systems, they become more accessible to open source developers striving to accelerate Linux application development within the enterprise.

Watch the coolest technology demo EVER. IBM Fellow Donna Dillenberger demonstrates the new IBM LinuxONE system for scalable financial trading. The demo shows multiple data loads (live data from the S&P 500 and Tweets) streaming via Maria DB, MongoDB, Spark Analytics, Chef, Docker and PostgreSQLContainers. In this LinuxONE demo, even with drastic upticks in CPU Utilization during the Greek financial crisis, response times are still lightning fast. Watch the demo here.

Explore the sections of the ebook to learn more.

Collaborate and innovate in a rich open source ecosystem

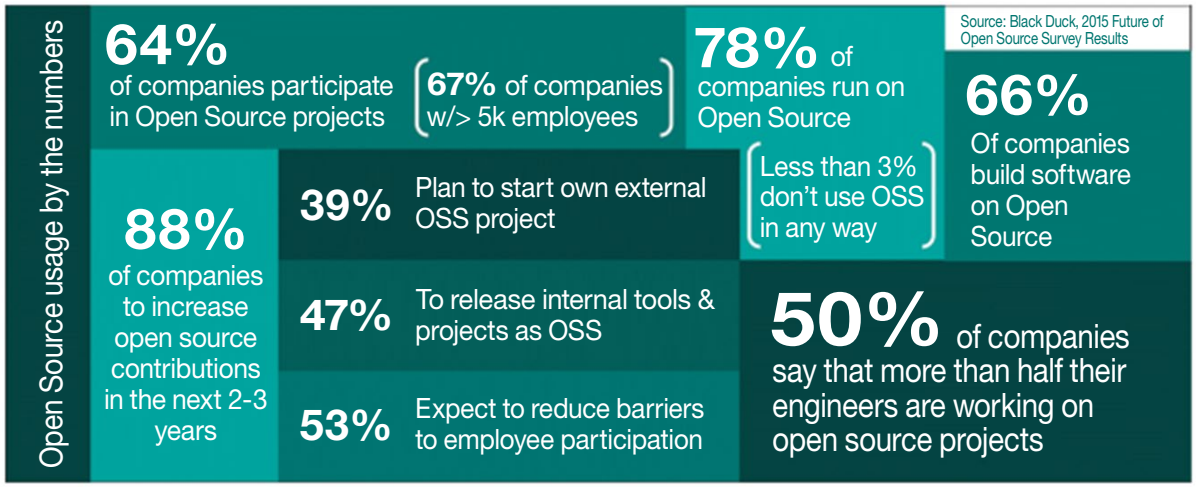
Speed plus capability delivers business agility

Agility is the ability to get to market quickly and effectively to solve the business problems you care about by leveraging best-of-breed **capabilities across eco-system**, security and management, while benefiting from industry leading **scale and performance**.

This ability is important to the success of organizations competing in an environment where new and disruptive technology innovations appear and are requested one after another. No sooner does an IT organization respond to and deliver applications to satisfy one business request, then another cooler, faster, more effective means of engaging with customers presents itself. The speed and pace at which a business needs to apply its IT skills is driving the need for transformation- challenging every level within the organization to improve time to market and drive deeper levels of engagement cross the entire value chain.

Open Source for business

Many of the world's most innovative organizations are utilizing Linux and open source technology solutions to deliver customer-facing, revenue-driving applications that serve millions of customers, clients and citizens. However, as Linux and open source deployments move from niche projects to being more pervasive in the data center to support business critical applications, it is imperative that these applications are fully supported by enterprise grade capabilities like non-disruptive scalability, unparalleled availability and continuous data protection. This is especially important for business-critical applications – essentially the solutions that support your entire organization and which require 24 x 7 availability.



Source: Black Duck, 2015 Future of Open Source Survey Results

Open source software brings a broad set of new capabilities to LinuxONE and z Systems in the form of more solutions that are optimized for use on an enterprise-grade platform that leverages inherent strengths—speed combined with ultra-security, application uptime, limitless capacity for growth, in addition to an inventive and collaborative ecosystem of developers and practitioners – all of which gives enterprises an important edge in improving their business agility.

Let's take a look at a few examples of how open technology can improve business agility.

Process more data, get better insights-faster. Discovering business intelligence requires the use of advanced analytics capabilities, such as Apache Hadoop and Apache Spark. Such capabilities are commonly deployed on distributed commodity hardware. However, when you are running the applications on LinuxONE or IBM z Systems (mainframes)—a hardware platform that allows up to 50 percent better response time in analytics than competing platforms[1]—given a fixed service-level agreement, you can process more data and receive better insights – which translates into better business agility.

Reduce risk, increase flexibility with diagonal scaling. Database partitioning or "sharding" is a common technique for scaling out a database that has become too large to fit within a single server. But sharding is complex in practice, it carries risks such as higher latency for aggregate queries and a lower level of data consistency, and the size of each shard is limited to the size of the servers. Using LinuxONE or z Systems allows you to scale out and scale up—increasing the amount of resources available to each shard server—in an approach termed "diagonal scaling.[2]" Diagonal scaling allows you to adapt to changing workloads with industry-leading performance, increased flexibility and reduced risks, and therefore offers better agility.

Meeting regulatory requirements. Another example of improved agility using Linux on z Systems is high-performance-secure-logging for auditing in the face of growing regulatory requirements. zEnterprise Data Compression (zEDC) facility allows IBM z Systems to offload main processors while speeding-up compression by up-to 10x. Linux on z Systems offers protected-key function which keeps the encryption key out of main memory and storage, keeping data-at-rest secure, while offering 50x better performance than secure-key function. When combining the speed and capability of zEDC and protected-key function, Linux on z offers unmatched agility for quickly, non-intrusively and securely logging snapshots of system state such as Docker instances or Spark Resilient Distributed Dataset (RDDs) for auditing purposes.

Free up valuable computing resources Apart from raw computing power, IBM z Systems can speed up compression and encryption with the zEnterprise Data Compression (zEDC) and Central Processor Assist for Cryptographic Functions (CPACF) features. Imagine compressing Spark RDDs (resilient distributed datasets) or Docker containers at high speed with little impact on CPU consumption; this frees up CPU cycles that can be used to perform more analytics, or faster, non-disruptive auditing of containerized applications. Imagine using OpenSSL and openCryptoki enhanced with CPACF; encrypted transactions will run faster, and data can be persisted securely and efficiently. This speed allows the computer to do more work, and gives you more agility.

Fifteen years of continuous innovation

IBM has a long history of involvement in open source software development, and continues to contribute to key open source technologies such as the Linux kernel, the Eclipse project, and many Apache projects, including the most recent, Apache Spark. IBM is a member of many open standard organizations and software governance consortia that help shape the future of open source software[3].

The IBM Linux Technology Center (LTC) is dedicated to enabling Linux adoption on IBM platforms and has been contributing continuously to core technologies such as the Linux kernel, glibc and GCC over the years. But architecture-specific code is only the tip of the ice-berg – and representing only a small portion of LTC contributions. For example, less than 2 percent of the kernel is z Systems-specific, with most of the 2 percent related to device drivers. Similarly, platform-specific code makes up about 0.5 percent and 0.3 percent of glibc and GCC, respectively. The majority of these platform details are transparent to application developers and end users.

Red Hat and SUSE ship enterprise Linux distributions for z Systems, and the Ubuntu, Debian and Fedora distributions also support the architecture.

The platform is mature and production-ready, and is able to run most traditional applications you would expect on a Linux server.

Today, more than 40 commonly used open source products run right out of the box, with more coming on board in future. By listening to clients, working with business partners and ISVs, and engaging the open source development community, IBM is working to bring more foundational open source technologies to LinuxONE and IBM z Systems, further enabling and encouraging the most sought after software developers to develop solutions that work on the platform – with a strong emphasis placed on programming languages and run-time technologies.

IBM is investing heavily to create a rich open source ecosystem to enable LinuxONE and IBM z System as the premier platform for new application deployment that exploits emerging technologies such as Node.js, MongoDB, PostgreSQL, MariaDB, Docker, Chef, Puppet and Spark. Additional focus is placed on open source workloads that can benefit from the reliability, availability, and serviceability (RAS) of the LinuxONE and IBM z Systems – bringing even more business values to users, including database management systems and cloud infrastructure, as well as analytics and big data solutions. Table 1 lists some of the key open source technologies that have already been enabled on the platform.[4] Over 40 packages have been enabled on the platform and continually growing with many others in the development pipeline based on input from our clients, Sale and business partners[4].

Table 1: Partial list of popular open source technologies available on IBM LinuxONE and IBM z Systems

Languages and Dev Environment	Database & Messaging	Cloud infrastructure
Node.js	MySQL	Docker
Ruby	PostgreSQL	Chef
Rails	MariaDB	Puppet
Python	MongoDB	Open Stack
LLVM	Cassandra	
OpenJDK	Redis	
GCCGO	CouchDB	
oCaml	Geode	
Erlang	RabbitMQ	
Apache HTTP Web Server		
PHP/Zend		
R language		
Clojure		
Scala		

The flexibility of Linux and open source combined with speed and processing power

The world's leading organizations are choosing Linux on IBM z Systems to run mission-critical applications that drive billions of business transactions every day leveraging the performance, reliability, security and virtualization capabilities of the enterprise grade platform for Linux to run databases (e.g. DB2 and Oracle) and middleware (e.g. WebSphere).

The explosive growth of mobile and cloud applications means that these systems are working harder than ever to drive data transactions on a global scale. In 2015, 27 percent of total workload on z Systems runs on Linux, and the percentage is growing. As innovations in open source technology gain traction in the market, demand for emerging technologies on IBM z Systems has increased, creating huge opportunities for open source developers and vendors.

Improved application performance, reliability and availability

IBM LinuxONE and z Systems are well-known for their reliability and availability- they ship with resilient and redundant hardware, which ensures continuous operations even in case of hardware faults. The machines boast high MTBFs (mean times between failures), and support hot-swapping of hardware, so they do not typically need to be taken offline for service. The use of IBM GDPS (Geographically Dispersed Parallel Sysplex) helps automate data replication and speeds up recovery from planned or unplanned outages.

IBM z Systems are also equipped with some of the fastest general-purpose processors in the world, such as the recently released IBM z13TM (z13) processor. Combined with its large and efficient memory cache hierarchy, the z13 is able to achieve high single-thread performance, which is ideal for scaling up applications to handle the massive volume of transactions that is typical in today's enterprise organizations.

Watch the video to learn more about IBM LinuxONE.

A better balance between risk and opportunity

Every organization is faced with the challenge of building both trust and value amidst uncertainty. Tasked with improving business performance while also reducing risk, business and IT leaders need to feel confident about risk exposures to data and infrastructure – all while meeting stringent regulatory requirements.

For example, as data and analytics become more critical to the development of business advantage, the need for comprehensive data security escalates. In response, many leaders are implementing stronger security and data privacy measures, in addition to governance policies, to protect their organization from both internal and external threats, proactively identifying and managing potential exposure to the risks associated with data breeches, compliance with industry regulations, and infrastructure vulnerabilities throughout the value chain.

The standard virtualization technologies in IBM z have been EAL5+ certified—the highest security level and isolation guarantees in the industry. These technologies offer low overhead and higher virtual machine (VM) density compared to other platforms, thanks to the advanced resource over-commit technology in the mainframe. As a result, applications that run on a mainframe can scale out to thousands of co-located virtual machines, as the mainframe essentially becomes a "data center in a box." What's more, they make the mainframe ready for cloud applications as they support rapid provisioning, multi-tenancy, and capacity scaling on demand.

The co-location of virtualized Linux guests brings significant benefits to applications that run on the mainframe. Using IBM HiperSockets, applications can transfer data from one VM to another in-memory instead of sending the data over the network; the transfer is more reliable (there is no physical connection to lose), more efficient (there is no network latency), and more secure (there is no wire to tap). Co-location is especially important where System of Engagement applications (e.g. mobile, cloud) frequently need to access System of Record data.

The IBM z/Architecture includes hardware support for cryptography, known as Central Processor Assist for Cryptographic Functions (CPACF). In the presence of a suitable cryptographic co-processor, CPACF supports the use of protected keys, which balances the speed of in-memory clear keys and the high security of hardware-backed, tamper-proof secure keys. Linux device driver support for the cryptographic co-processor has existed since 2012, and the OpenSSL and openCryptoki libraries are able to exploit these z Systems features to speed up clear key, protected key and secure key cryptography.[5]

Speed and performance improvements

The IBM zEDC Express adapter is another feature that distinguishes the mainframe as a data-processing powerhouse. It allows applications to off-load zLib-compatible compression work to a hardware co-processor, achieving a good compression ratio without consuming CPU cycles. For databases that employ on-the-fly compression, exploitation of the zEDC feature in Linux is expected[6] to improve performance by a factor of five. Users will be able to process more data in the same amount of time, and also save on storage costs.

Ease of use

You do not need to learn any mainframe commands to log on to a guest with SSH and be productive in a familiar Bash environment. Graphical user interfaces with remote connection capability, such as X11 and Xvnc, are also available if you prefer a graphical desktop environment. All the essential development tools are available for LinuxONE and z Systems. If some user-space applications are not readily available from distributions, most of them merely need a simple re-compilation to be able to run on the mainframe. Applications written for modern run-time systems, such as Java, Node.js, and PHP applications, simply run out of the box. Because of the similarity and compatibility with other enterprise Linux platforms, you can migrate whole applications stacks from other architectures to z Systems simply and easily, often with just a small amount of work, in hours or days rather than weeks or months."

More choices and speed with contemporary programming languages

Contemporary programming languages provide more choice and speed on the platform

Application developers on IBM LinuxONE and IBM z Systems can choose from a variety of modern, popular programming languages and run-time environments, including Java, Node.js (JavaScript), Python, Ruby-on-Rails, Scala, Erlang, and Go—and the list is growing. This variety allows new, modern applications to be built and run on Linux on z Systems easily.

Deploying Node.js

The most notable of these run-time technologies is Node.js. Node.js is the fastest growing ecosystem for application development- it has been growing at a rate three times faster than Java, and the number of Node.js modules contributed by developers worldwide has surpassed those of other languages.[14]

As a high-performance, highly scalable, event-driven, server-side JavaScript solution, Node.js has become popular for web application development—the "N" in the term "MEAN stack" stands for "Node.js."

To meet the huge demand for Node.js functionality on z Systems, IBM has ported the open source Node.js code to the platform. IBM released IBM SDK for Node.js 1.2 in early 2015. This SDK is fully compatible with Node.js version 0.12.[15] The SDK enables developers to write Node.js code on any platform, and test and deploy their applications on LinuxONE and z Systems.

With the wealth of corporate data that either resides or originates on the mainframe, there's a strong likelihood that cloud or mobile applications and APIs built on Node.js will require access to data and services on LinuxONE and IBM z. With Node.js, clients can expose the Systems of Engagement opportunities from Node.js, and tie them with the trusted Systems of Records on the mainframe.

Advantages of deploying Node.js on LinuxONE and IBM z

Substantial improvements in application throughput with reduced transaction response time

By deploying Node.js applications on IBM LinuxONE and IBM z, you gain all the strengths, security and reliability of the mainframe. For example, z Systems fast general-purpose processors are well suited to the Node.js single-threaded processing model. The large I/O capacity of z Systems also complements a strength of Node.js applications: the ability to scale and maintain thousands of concurrent connections.

Use the same set of tools and skills to diagnose both your Node.js and Java applications

- **Using IBM Health Center, you can monitor your Node.js applications and gain valuable insight into application hot spots, garbage collection activities, memory consumption and much more.**
- **The IBM Interactive Diagnostic Data Explorer (IDDE) is now aware of your Node.js runtime and JavaScript metadata, giving you the ability to scan the heap for JavaScript objects or walk JavaScript stack frames during postmortem analysis.**
- **Both Health Center and IDDE are available at no cost as part of IBM Support Assistant and on the Eclipse Marketplace. Support for IBM SDK for Node.js is also available upon request.**

IBM LinuxONE and z Systems can further provide an advantage to JavaScript applications that need to access enterprise data and transactions. Logical partitions on LinuxONE and z Systems can be linked together by high-speed networking HiperSockets, allowing memory-to-memory communication with drastically improved performance. JavaScript applications can leverage HiperSockets transparently to access data and transactions in other logical machines inside the box.

The IBM SDK for Node.js is 100 percent API compatible with the community version. The IBM SDK was built from the open source community code base, with extensions added for support on z Systems. Leveraging the strength of monitoring and debugging tools developed for Java, they have been extended to work with the IBM SDK for Node.js.

For more information, see Node.js and IBM SDK for Node.js

Enterprise users of the SDK will enjoy the performance and security afforded by the IBM LinuxONE and IBM z platform (on the AcmeAir benchmark[17], Linux on z Systems out-performs x86 by 52 percent), as well as the additional monitoring and debugging tools that ship with the SDK.

One of the most important benefits from deploying Node.js on IBM LinuxONE and z Systems is the ability to bring the processing of web requests to the platform where the data is hosted (co-location of application and data), which can improve application throughput by up to 200 percent, and reduce transaction response times by 60 percent. The enablement of Node.js on the platform also means that a large number of open source web application frameworks, such as Express and Sails.js, are now available to mainframe developers.

Node.js relies on the V8 JavaScript just-in-time compiler. For this reason, IBM first ported V8 to IBM z Systems. The availability of V8 on z Systems has allowed a number of other open source software that requires JavaScript functionality to run on the platform. IBM will be contributing the z Systems port to the V8 project and works closely with the community to further improve the performance of V8 on z Systems.

Boost database performance from SQL to NoSQL

Better per core performance and scalability

NoSQL? No problem. NoSQL databases are increasingly being adopted in big data and real-time analytics applications. Some of the most important NoSQL databases include MongoDB (the "M" in "MEAN stack"), MariaDB, Apache Geode, Apache Cassandra, Apache CouchDB, and Redis. These databases all run well on Linux on z Systems, offering you a healthy variety of choices for NoSQL applications on the mainframe.

PostgreSQL achieves up to 2.2x better pgbench throughput over alternative platforms

Open source relational database management systems such as PostgreSQL serve a great amount of data in many organizations. First released in 1996, PostgreSQL is a well-established enterprise-grade database system that powers many web sites and government agencies.[8] PostgreSQL 9.4 already runs well on z Systems, and it is able to capitalize on the strengths of the mainframe as a data-serving platform.

Open source relational database management systems such as PostgreSQL serve a great amount of data in many organizations. First released in 1996, PostgreSQL is a well-established enterprise-grade database system that powers many web sites and government agencies.[9] PostgreSQL 9.4 already runs well on z Systems, and it is able to capitalize on the strengths of the mainframe as a data-serving platform.

IBM LinuxONE and z Systems offers many features that others systems cannot match—including availability, stability, and flexibility—that combine to help PostgreSQL excel as a powerful enterprise database solution. An independent performance test[2] conducted by 2ndQuadrant found that PostgreSQL on the IBM z13 infrastructure provides greater throughput, regardless of the type of workloads, compared to an infrastructure based on a competitive architecture. According to a report by the PostgreSQL consultancy firm 2ndQuadrant,[10] PostgreSQL can offer better per-core performance of 1.6 to 2.2x via pgBench benchmark and scalability on the mainframe compared to distributed systems, as shown in Figure 1.

Figure 1 shows, PostgreSQL can offer better per-core performance and scalability on the IBM z compared to distributed systems.[11]

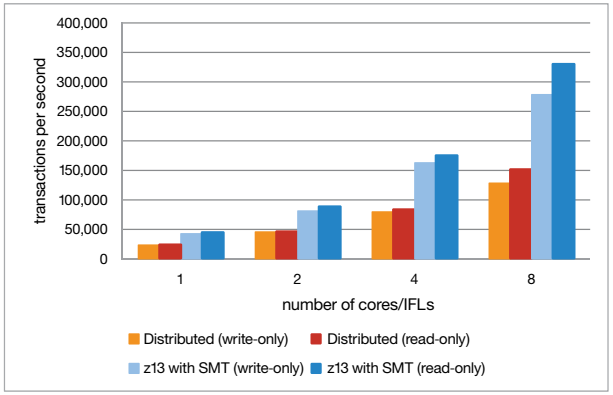


Figure 1. PGBench results (read-only and write-only, in-memory workloads) on IBM z13 vs. distributed systems

Global PostgreSQL support and consultancy firm, 2ndQuadrant, partners with IBM to support PostgreSQL on IBM z Systems. For more information, see 2ndQuadrant.

IBM also offers PostgreSQL as a solution on IBM Bluemix. For more information, see IBM Bluemix docs.

PostgreSQL 9.4 has been tested on enterprise Linux distributions (RHEL and SLES) that run on z Systems. For more information about the installation instructions, see Building PostgreSQL.

For more information, see PostgreSQL.

[1]: http://www.enterprisedb.com/postgres-plus-edb-blog/gary-carter/sql-server-users-what-you-dont-know-about-postgresql-can-cost-you

[2]: PostgreSQL performance on z/Linux done by 2ndQuadrant:

Porting, performance measurement and scaling advantages with MongoDB

MongoDB is a cross-platform document-oriented database. Released under a combination of the GNU Affero General Public License and the Apache License, MongoDB is free and open-source software. Since its initial release in 2009, MongoDB has been adopted as back-end software by a number of major websites and services. Today, MongoDB is the fastest-growing database ecosystem, with over 10 million downloads, thousands of customers, and more than 1,000 technology and service partners.

Classified as a NoSQL database, MongoDB eschews the traditional table-based relational database structure in favor of JSON-like documents without fixed schemas. The schema-less nature of MongoDB not only makes it better than relational databases at handling unstructured and sparse data, it also facilitates rapid prototyping and fast evolution of programs without paying the constant costs of updating database schemas and refactoring tables.

By adopting JSON/BSON as the document format, MongoDB allows developers to write data queries in JavaScript, a language with which they are already familiar, so that the need to mix SQL statements into the application code can be eliminated. It also eliminates the cost of ORM (object-relational mapping), because JSON documents can contain complex types and are by nature JavaScript objects. Because JSON is widely supported as a de-facto standard for information exchange, it is easy to consolidate data from multiple disparate sources into one MongoDB collection.

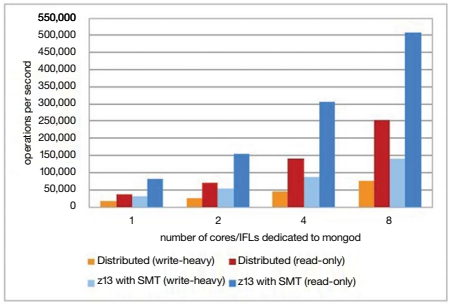
MongoDB also supports geospatial and time-series analytics, as well as scalability and high availability through sharding and replication, making it a highly capable engine for processing the vast amount of data generated online. All these factors help make programs easier to develop and maintain, boost programmer productivity, and improve business agility.

Big Data is becoming more prevalent as we start to leverage insights from structured and unstructured data together to deliver better business outcomes. MongoDB represents a powerful aggregation point where insights from traditional system-of-record data i.e. retail transaction history, can be joined with system-of-engagement data i.e. micro-beacon geodesic information, to provide a rich and comprehensive user-experience for the shopper. On top of this, MongoDB's database aggregation now enable businesses to bring these two worlds together for their own benefit, and gain access to insights to all of their data to facilitate critical business decisions.

Advantages of using MongoDB with LinuxONE and IBM z

You can easily build and run MongoDB on LinuxONE and IBM z Systems. All features are available including sharding, replication and compression. Applications that use any of the existing MongoDB drivers work without modification. With the wealth of corporate data that either resides on or originates from the mainframe, running MongoDB on LinuxONE or z Systems enables faster and more secure access to data sources for aggregation purposes. The performance and virtualization capabilities of IBM z also make it ideal for scaling out as well as scaling up NoSQL applications.

MongoDB achieves up to 2x better YCSB throughput on IBM z13 over alternative platforms



MongoDB YCSB (update- heavy) throughput on IBM z13 vs. distributed systems

Advantages of running MongoDB on LinuxONE and z Systems

By deploying MongoDB on z Systems, you gain all the strengths, security and reliability of the mainframe. IBM z Systems platforms are equipped with some of the fastest general-purpose processors in the world, and they are well-known for data processing throughput.

- **The large number of cores available in a z System and its high input/output bandwidth mean that MongoDB can scale up, as well as scale out.**
- **The enterprise-grade virtualization capabilities of z/VM support very high virtual machine density and resource over-commit, which allows MongoDB instances to achieve much better utilization than on other platforms.**
- **Co-located MongoDB instances (different shards, or replica set members within the same shard) can leverage HiperSockets to reduce network communication overhead.**
- **Proximity to enterprise data that resides on the z System, such as z/OS, allows MongoDB to perform ETL (extracting, transforming, and loading) on such data with low latency and high security.**

MongoDB achieves up to 2x better YCSB-A throughput on IBM z than on alternative platforms.

The superior single-thread performance and large memory capacity of IBM z allows MongoDB to scale up very well. On various YCSB (Yahoo Cloud Service Benchmark) workloads, MongoDB running on Linux on IBM z is able to achieve up to two times better throughput than the latest distributed platforms. Because the IBM z LPAR can support a maximum of 10 TB of memory, a single Linux system can accommodate a larger database that would have required sharding to multiple MongoDB servers on other platforms. This affords users the flexibility to reduce or avoid the risks and costs associated with sharding.

MongoDB drivers for various languages, such as C/C++, Node.js, and Ruby, either have been verified to work on Linux on z Systems or are being ported to the platform. IBM is working to contribute all the changes upstream to the latest version of the product.

In the case of MongoDB, some porting effort was required because it uses little-endian byte order pervasively, and IBM z System is a big-endian architecture.[12] Fortunately the byte order issue is well

understood and eliminating unnecessary assumptions of little-endian byte order in the code is often straightforward. Users do not need to be concerned with the native byte order of most programs, and it is easy to write portable code that can function correctly on either type of architecture.[13] In the end, only 0.14 percent of the code had to be modified to enable MongoDB to run on Linux on z Systems.

Three simple use cases using MongoDB on IBM LinuxONE and z Systems

- Aggregating data from multiple sources into a central repository to create a single 360-degree view of enterprise customers. Complex queries can be performed over the aggregated data efficiently.
- Exploiting the flexible data model and multiple options for scaling—including range-based, hash-based and location-aware sharding—to develop applications quickly, and to support the constantly changing requirements. Schemas can be radically changed while running in production, with zero impact on the user experience.
- Creating a read-forward cache of System of Record data for fast viewing or manipulation by a front-end system, such as a web application or a mobile application, which could be in use by a large number of concurrent users.

Compatible IBM products and services working with MongoDB

MongoDB is currently available on IBM Bluemix via MongoLab, a fully-managed cloud database service featuring high availability, automated backups, web-based tools, monitoring and support. Since version 10.5, IBM DB2 has provided JSON support.

One way that DB2 users can interact with JSON data is to deploy the MongoDB wire listener, which intercepts the MongoDB wire protocol and acts as a gateway in front of DB2 – which allows DB2 to accept and respond to requests from applications written against the MongoDB API. Any modern language with a driver that supports the MongoDB protocol can be used, including Node.js, PHP, Python and Ruby, as well as more traditional languages such as C, C++, Java and Perl.

For more information, see DB2 JSON capabilities, Part 1: Introduction to DB2 JSON

You can explore the MongoDB port for z Systems as open source code. Instructions for building and installing MongoDB on LinuxONE and z Systems are available in Building MongoDB.

Throughput improvements with MariaDB on LinuxONE and IBM z

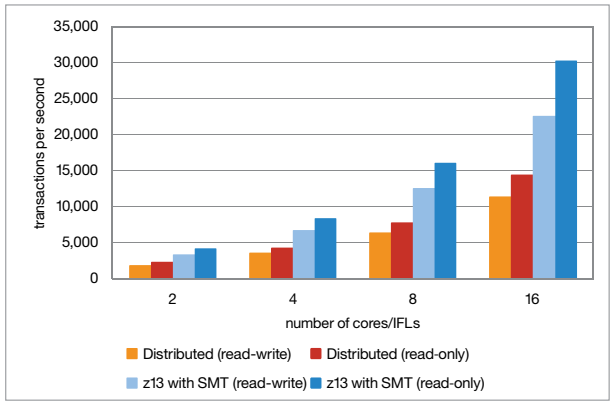
MariaDB achieves up to 2.1x better Sysbench OLTP throughput on IBM z over alternative platforms

MariaDB is a community-developed fork of the popular MySQL relational database management system, and it is intended to remain free under the GNU GPL. Being a fork of a leading open source software product, it is notable for being led by the original developers of MySQL, who forked it due to concerns over its acquisition by Oracle. Contributors are required to share their copyright with the MariaDB Foundation.

MariaDB delivers value through the following aspects:

- Backward compatibility with MySQL
- Perpetually open source
- More cutting edge technologies and more storage engines than MySQL

IBM LinuxONE and z Systems offer state-of-the-art hardware, world-class security and the capability to scale MariaDB with IBM's industry-leading mainframe virtualization technologies. Clients can easily consolidate clusters with many servers on one mainframe. IBM z Systems can host more servers per core than any other system with high-speed encryption, disaster recovery, and continuous availability solutions.



MariaDB: Substantial throughput improvements on IBM z System

Common use cases for MariaDB

Many IBM clients and partners use MariaDB as an alternative to MySQL and Microsoft SQL Server because MariaDB future-proofs their database investment, obtains better value, and developers and IT departments want to leverage the default database from their Linux distributions and simplify deployments as well.

The combination of Linux, Apache, MySQL/MariaDB, and PHP/Python/Perl—together known as LAMP—is perhaps the most commonly used web service solution stack. Common use cases include:

- Web and mobile applications
- Content management systems
- E-commerce and ticketing
- Search and advertising
- Business intelligence and analytics
- Games, entertainment, and social media

Collaboration with the MariaDB community

IBM actively contributes to the MariaDB community to ensure that the latest MariaDB works out of the box on IBM platforms. The effort includes fixing bugs, porting new features and optimizing for IBM platforms. For example, IBM has recently built and validated MariaDB with Galera Cluster capability on IBM z Systems. IBM partners with MariaDB, Inc. to deliver MariaDB Enterprise for IBM platforms and 24/7 support services to IBM customers with business and mission-critical application needs. This partnership has helped many IBM customers running Linux on Power, and IBM is working to extend the partnership to cover IBM z Systems as well.

Compatible IBM products and services working with MariaDB

MariaDB Inc. currently partners with IBM to deliver MariaDB Enterprise solution to Linux on Power Systems. MariaDB Enterprise extends MariaDB 10 with certified binaries and continuous value delivery to optimize MariaDB for the most challenging use cases. MariaDB is also certified on IBM z Systems running Red Hat Enterprise Linux and SUSE Linux Enterprise Server.

The MariaDB community version runs out of box on Linux on z Systems. You can install MariaDB 5.x through Yum on RHEL or Zypper on SLES. See the Installation guide for MariaDB 10.x. MariaDB Inc. and IBM are working together to bring MariaDB Enterprise to z Systems.

For more information, go to https://mariadb.org

For more information, see MariaDB.

Advances to large-scale data processing

"We believe strongly in the power of open source as the basis to build value for clients, and are fully committed to Spark as a foundational technology platform for accelerating innovation and driving analytics across every business in a fundamental way."

—Beth Smith, General Manager,
Analytics Platform, IBM Analytics

New options for analytics

Big data and real-time analytics are key to success to many enterprises. Apache Spark is an open source cluster computing framework that allows user programs to load data into a cluster's memory and query it repeatedly. As data and analytics are embedded into the fabric of business and society – from popular apps to the Internet of Things (IoT) – Apache Spark brings essential advances to large-scale data processing. First, it dramatically improves the performance of data dependent apps. Second, it radically simplifies the process of developing intelligent apps, which are fueled by data.

Apache Spark is agile, fast and easy to use. And because it is open source, it is improved continuously by a worldwide community. Spark is designed to perform both batch processing and new workloads, such as streaming, interactive queries, and machine learning.

New options for analytics allow data scientists to take advantage of IBM LinuxONE and z Systems advanced analytic capabilities without having to worry about the specific format of the data.
The enablement of Apache Spark for LinuxONE and z Systems allows data mining specialists to expand analytic software to leverage additional z Systems processing power. Data scientists can now use Apache Spark's common programming framework to take advantage of z Systems' advanced analytic capabilities without having to worry about the specific format of the data.

Spark running on IBM z produces 1.5x faster insights for real-time analytics than on alternative platforms.

IBM has tested the Spark-Perf benchmark suite on IBM z Systems, demonstrating that the platform can process 54 percent more data (see Figure 2) than distributed systems for model building, leading to real-time insights with higher accuracy.

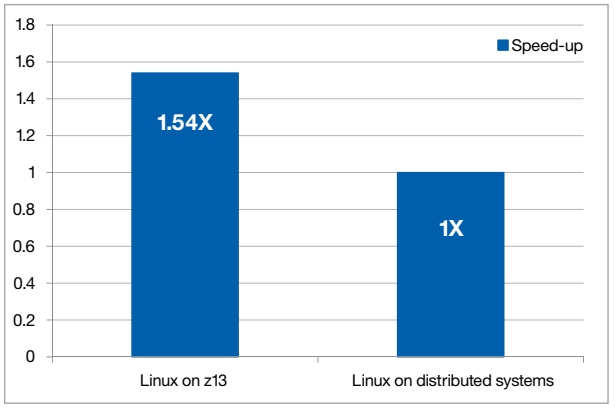


Figure 2. Spark-Perf "Spark" benchmark speed-up on z13 vs. distributed systems

Application portability and infrastructure flexibility

Docker delivers ease of use and application portability

Docker is an open, portable, light-weight run-time and packaging tool for Linux containers, and has gained a lot of momentum since its first release in 2013. Linux containers provide operating system-level virtualization for running multiple isolated applications on the same host. Using containers, it is easy to build and ship complex applications that have dependencies on a deep software stack or other infrastructure, without having to worry about interference from other applications that might have conflicting requirements. Compared to a standard virtual machine containers are much faster to boot, more efficient to run, and offer higher application density.

Docker provides developers with more ease of use, portability and the ability to quickly get up and running, enabling you to develop anywhere. Essentially, Docker has an engine/runtime that sits on top of the operating system and provides the virtual containers into which you deploy the software. Docker offers an elegant portability solution for application/software because it is lightweight and removes some of the complexities of managing a hypervisor, e.g. z/VM or VMware. However, if you want security, you need to package them as second level guests to get the isolation between the apps and data.

IBM has built Docker on Linux on z Systems.[7] It is straightforward to set up a base image and create Docker containers based on the image. Advanced application composition and container clustering are possible with Docker Compose and Docker Swarm, and IBM has provided build instructions for those tools as well.

Docker is an open source tool that provides a way of running isolated applications and software in a single Linux instance in what are called "containers".

- **Package apps together to more efficiently deploy them to get better density**
- **Develop the app package on any platform, and provided the binaries exist for the platform, deploy the app anywhere - on your laptop, in your data center, or in a public cloud**
- **Gaining traction as a way to contemporize Linux environments in datacenters**
- **Doesn't require any virtual machine resource overhead since the hypervisor is not required**
- **Enables full support of DevOps model through simple build process of containers, versioning, and deployment**
- **Simple scaling of stateless solution components (e.g. node.js instances).**

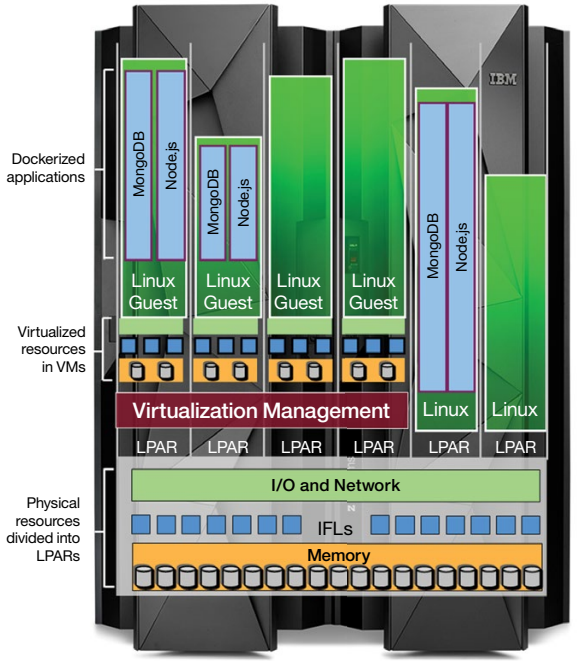


Figure 3. Different levels of isolation—LPARs, VMs and Docker containers

Docker binaries for both RedHat 7 and SUSE 12 are based on the Open Source version for Linux on IBM z Systems and IBM Power Systems. They are available for technology preview at IBM developerWorks. Managed images can be uploaded to "ibmcom" in Docker Hub. The uploaded files might renamed with s390 appended to the file name.

Advantages using Docker on LinuxONE and IBM z

The virtualization technology in LinuxONE and z Systems complements Docker by offering flexibility in the level of security isolation. For example, while developing or testing an application, you can obtain increased application density by deploying containers directly on one logical partition (LPAR), which eliminates the need for VM provisioning.

On the other hand, when deploying the application in production, you can put the containers inside second-level guests to get better isolation. For maximum security, you could run individual containerized applications inside separate LPARs, which give isolation on a bare-metal level unmatched by other platforms. See Figure 3 for an illustration of how containers can run on LPARs as well as VMs.

IBM z Systems application performance per Docker container is higher with a high number of containers per virtualized or physical resources. Its low hypervisor overhead allows Docker containers that are running on second level VM guests to have minimal impact on application performance while providing full security isolation in a multitenant enterprise environment. In addition to the main Docker tools, IBM is working to provide Dockerfiles that allow users to create containerized applications of their choice to run on IBM LinuxONE and z Systems.

Four simple use cases

1. *Faster more efficient deployment: write your applications once*
   You can write an app once, and send to other people and they know how to deploy. Now you don't have to develop an organization that knows how to install the app and knows what other packages are required. The containers provide the user with the automation for packaging with automated scripts inside the container, providing an easier, more efficient, and faster way to deploy and run apps.

2. *Break down the components of your applications: choose only what you need to build what you want*
   Let's say you have an app that requires lots of parts for a multi-tiered application- a workflow component, WebSphere component, database component, and math library component. You can put them in four different containers. Then, if you are not interested in the workflow component, the user only needs to deploy three of the containers. This allows the user to break down apps into different parts and use the pieces needed only. Software products often ship the entire product – in order to access the one component you need, you have to install the full complement of pieces. If your app requires all the components, you can put them in one container. If your app doesn't, you can keep them separate - as if the pieces are optional. The flexibility of being able to select only what you need – like building blocks, enables you to more efficiently build what you want.

3. *Simple application portability: Develop Java apps on one platform and deploy on multiple platforms*
   A developer can develop apps in Java on an Intel platform and deploy them on Linux on z Systems. Since Java doesn't require a recompile on different platforms, app portability is a given. Without containers, the developer didn't know if they had the right libraries, Java VM level, and middleware requirements. But now with containers, any software the app requires can be packaged into a container, provided it is Java. Although the container has to be rebuilt on the platform image you are using, but once you are in a container model this task becomes simple.

4. *Enable more applications in one system: Docker on LinuxONE and IBM z allows for larger density than VMs*
   Docker allows for larger density than VMs and enables more apps in one system. Before containers, a user could only run 10 WebSphere instances on an Intel system, but now with containers and enhanced memory efficiency, the user can run hundreds of the same instance, as you don't have to set up separate VMs with hypervisors.

IBM LinuxONE and z Systems offers you advantages over x86 with memory overcommit, and the overhead on an IBM z with CPU virtualization is relatively small. Thus, one will see a bigger advantage in the distrusted environment leveraging containers, but even more interesting is the combination of containers and security isolation due to this memory overcommit - lower VM overhead. Here's why:

In a Docker environment that leverages density, the user loses the security isolation between these apps because the hypervisor is not present. If security is not a priority, then running on bare metal gives good density at faster response time. However, if your organization is concerned about security, then you will want the isolation, and the overhead on the IBM LinuxONE and z Systems is minimal.

Organizations running mission critical applications can leverage density/bare metal in a development-test environment to rapidly put everything on one LPAR on LinuxONE or z Systems to eliminate the provisioning associated with hypervisors. In production, you can revert back to second level guests to get the isolation, and live with the hypervisor overhead as the production environment does not change rapidly. You can use this VM isolation on a tenant granularity to get the isolation on a tenant base, and on sufficient "mass" of applications to gain from the efficiency opportunities because VMs don't have as big overhead on the mainframe as compared with that on a distributed x86 platform.

With IBM LinuxONE and z Systems you can shape your environment with system virtualization and container elements according to your landscape and requirements, without limitations in performance. That means you can define IT structures according to your needs, not your system constraints.

Test drive your Linux and Open Source solutions

IBM is committed to working with open source development communities to enable them to develop, test, and continuously integrate code on LinuxONE and z Systems. IBM offers access to LinuxONE and z Systems hardware in a number of ways.

- IBM PartnerWorld members can sign up for the IBM Systems Application Advantage (also known as "Chiphopper") program, which helps partners port and support existing Linux applications on to IBM z Systems, for a limited time, at no cost. For more information, visit: **ibm.com**/partnerworld/wps/servlet/ContentHandler/isv_com_dvm_techval_chiphopper.
- Open source developers can get access through the Community Development System for Linux service. The access is for a limited time, but renewals are possible depending on the situation. Visit: **ibm.com**/systems/z/os/linux/support/community.html for details.
- IBM is working with partnering universities on Open Access Mainframe Community Cloud, which will provide z Systems resources to research projects and open source development communities.

For more information on LinuxONE, visit IBM on the web at **ibm.com**/LinuxONE.

For more information

To learn more about the using open source and Linux solutions in the enterprise, please contact your IBM representative or IBM Business Partner, or visit the following website: http://www-03.ibm.com/systems/z/os/linux/

[1] All performance data had been measured in a controlled laboratory environment. Actual results may vary.
[2] http://highscalability.com/strategy-diagonal-scaling-dont-forget-scale-out-and
[3] IBM is a member or sponsor of: The Linux Foundation, The OpenStack Foundation, The Node.js Foundation, The OpenJDK Governing Board, OASIS, and The Apache Software Foundation, among others.
[4] The central list of ported open source software can be found at our developerWorks community: https://www.ibm.com/developerworks/community/groups/community/lozopensource. Feel free to raise requests in the community forum for open source software that are not yet available on the platform.
[5] https://www-01.ibm.com/support/knowledgecenter/linuxonibm/com.ibm.linux.z.lxci/lxci_c_geninfo.html
[6] Linux kernel support for the zEDC Express adapter will be available in the near future.
[7] https://www.ibm.com/developerworks/linux/linux390/docker.html
[8] http://www.postgresql.org/about/users/
[9] http://www.postgresql.org/about/users/
[10] Mark Wong (2ndQuadrant Ltd.), OLTP Performance Benchmark of PostgreSQL 9.4 on IBM z Systems, http://2ndquadrant.com/en/support/support-ibm-z-systems/performance-analysis/
[11] http://2ndquadrant.com/en/support/support-ibm-z-systems/performance-analysis/
[12] To try out the z Systems port of MongoDB, see https://github.com/linux-on-ibm-z/docs/wiki/Building-MongoDB.
[13] Rob Pike, The Byte Order Fallacy, http://commandcenter.blogspot.ca/2012/04/byte-order-fallacy.html
[14] http://www.modulecounts.com/
[15] http://www.ibm.com/developerworks/web/nodesdk/
[16] https://github.com/acmeair/acmeair-nodejs

IBM Systems
Route 100
Somers, NY 10589

Produced in the United States of America
August 2015

ZSB03071-USEN-00