July 2021

Securing and Accelerating R3 Corda Enterprise with IBM LinuxONE

Authors:
IBM:
Robert Blessing-Hartley, Mori Ohara, Tatsushi Inagaki

© 2021 IBM Corporation

Neither this documentation nor any part of it may be copied or reproduced in any form or by any means or translated into another language, without the prior consent of the IBM Corporation.

© 2021 IBM Corporation

Table of Contents

© 2021 IBM Corporation

Executive Summary

R3 Corda is an enterprise blockchain platform that is designed to deliver privacy, interoperability, and scalability. Corda underpins top-of-stack applications, CorDapps, built by ecosystem participants.

IBM engaged with R3 to validate the performance characteristics of R3 Corda Enterprise 4.1 running on IBM Z®/LinuxONE and to define a roadmap for a combined platform of Corda Enterprise running on IBM Z and LinuxONE to deliver an enhanced performance and security proposition compared to an x86 deployment.

The IBM team defined a test harness around Corda Enterprise running Linux® on a dedicated logical partition on an IBM z15™ (Linux on z15 is comparable to an IBM LinuxONE III) and on a bare metal x86 system with an identical number of cores.

As a result of our efforts, we are able to publish repeatable results which demonstrate a significant increase in transaction throughput in a Corda 4.1 network on IBM LinuxONE. Running Corda 4.1 Enterprise on both z15 and compared x86 environments showed significant throughput performance improvement, up to 1.7x, on IBM z15 versus on the compared x86 system with up to a 50% decrease in the number of processors required to reach approximately 500 TPS through a single node. This is indicated in Figure 2 in the Results section of this paper. This rate is different from the overall transactional rate of a multi-node network since Corda Enterprise is designed to support many point-to-point transactions concurrently. Overall rate of a “real” network could be several orders of magnitude higher. Testing of a “real” network was not part of the intended scope of this paper. The LinuxONE III system delivered a consistently higher transactional rate (on average 1.56x and peak of 1.7x better), core for core, over the compared x86. At 32 cores LinuxONE III also scaled vertically up to 40% better than the compared x86 configuration.

IBM LinuxONE III/z15 delivers innovative security capabilities to the Corda platform which are also highlighted in this paper. LinuxONE servers are designed to deliver protection against both internal and external cyber threats – without requiring changes to applications.

For multi-tenant clouds, hardware virtualization with PR/SM™ provides “air-gap” level isolation between logical partitions, protecting against any peer access to critical data by other customers. It also enables new virtual machines to be provisioned faster and share resources.

LinuxONE includes hardware encryption. Faster encryption and decryption make it more practical for organizations to pervasively encrypt 100 percent of data.[1] Secure Service Containers provide a framework for deploying software appliances on LinuxONE. Once configured, access is permitted only through well-defined APIs and web interfaces. This is designed to prohibit access by rogue system administrators or other external threats to private data, provide validation of the appliance code at boot time, and automatically encrypt data.

For all IBM LinuxONE III hardware information, please refer to the datasheet at https://www.ibm.com/downloads/cas/ZV0EVLAK

© 2021 IBM Corporation

The Challenge

A consistent challenge to implementers of blockchain solutions is the comparatively low transaction rates that blockchain technology currently delivers. It is estimated, for example, that the Bitcoin network achieves a rate of 7 transactions per second (TPS).[2] With the growth in the variety of blockchain use cases, the need for improved performance through higher TPS has become increasingly important for these use cases to be viable.

Corda addresses this issue at a network level with its point-to-point, 'need to know' design which enables the throughput of a network to exceed that of any individual node. Some use-cases also require higher TPS through a single node and that is the focus of this report.

In addition, the data frequently stored on a blockchain network is "high value" and therefore a more likely target of malicious actors.

The security and performance of a blockchain network are at the core of its ultimate value. Delivering security and performance in a cost-effective manner is a significant challenge.

IBM LinuxONE Approach

The IBM LinuxONE platform is designed to deliver the capabilities necessary to support critical enterprise applications. As a leader in security capabilities, reliability, and scalability, IBM LinuxONE is designed to offer a high-performance infrastructure to run Corda Enterprise.

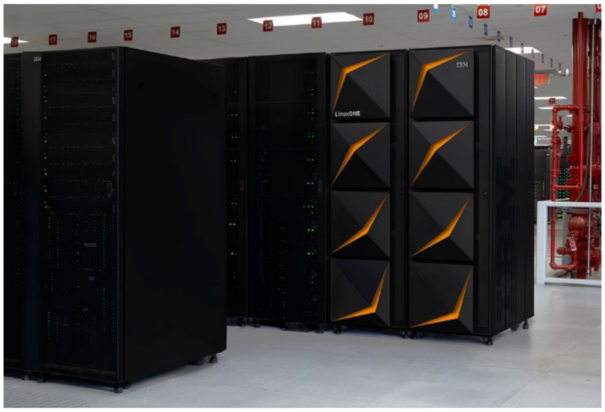

Figure 1. IBM LinuxONE

Amongst the many features that differentiate IBM LinuxONE from other platforms, several are particularly relevant to blockchain networks. [3]

Data Isolation: IBM LinuxONE logical partitions and Hyper Protect Virtual Servers provide multiple layers of isolation. Containers running in a Hyper Protect Virtual Server run in their own dedicated KVM instance providing isolation both at the Docker Namespace and at the KVM hypervisor layer. Each logical partition (LPAR) is designed to provide Common Criteria EAL 5+ certified isolation for virtual machines –– and as much as 16 TB of physical memory.

[2] https://www.ibm.com/downloads/cas/KJDPQKBE

[3] For all IBM LinuxONE III hardware information, please refer to the datasheet at https://www.ibm.com/downloads/cas/ZV0EVLAK

© 2021 IBM Corporation

Secure Service Container/IBM Cloud® Hyper Protect Virtual Servers: are designed to help protect blockchain (and other software) implementations against internal and external attack by encapsulating all software in a secure, signed, trusted appliance-style container, sealed and validated against tampering. This encapsulation, when properly configured, provides a blockchain or other software stack with protection against malware injection, misuse of privileged user credentials and deliberate or unintentional leakage of information.

IBM Hardware Security Modules: IBM's LinuxONE can further enhance security by encrypting all data in the blockchain container. Encryption keys for the Secure Service Container are stored in dedicated, tamper responsive Crypto Express cards. The IBM HSM is certified to the highest commercial level of NIST certification, FIPS 140-2 Level 4. When running with a Secure Service Container/Hyper Protect Virtual Server, this encryption is turned on, by default, and cannot be disabled.

Performance: IBM LinuxONE outperformed a compared x86 test system, on average 1.56x and at peak 1.7x. transactions per second. In addition, our tests showed that because of the unique processor architecture, that performance gain can be achieved with fewer cores, up to half as many at 32 x86 cores, compared to 16 IFLs.

Implementation

For the sake of simplicity and ease of repeatability we defined fixed sized systems against which we ran our test suite. Our test environments do not model a real-world implementation of Corda. It is our conjecture that all environments, when implemented across multiple dedicated systems (Database, Notary, etc.), would perform better than our test configuration. Likewise, our configuration was not designed for high-availability or disaster recovery.

Evaluation

Test Configuration

Please note: all results are based on this specific configuration. Because of the possibility of variations in your test environment, your results may vary.

Common Configuration

All Systems, both LinuxONE and the compared x86, ran Corda Enterprise 4.1.

A RAM Disk was used on all systems for PostgreSQL Database and persistent storage for ActiveMQ.

We used AdoptOpenJDK 8-OpenJ9 to execute the Corda nodes. We compiled native libraries bundled with Corda Enterprise since the binary executables for Z were not available in the delivery from R3.

The native libraries are:

- netty-4.1.39.Final
- netty-tcnative-2.0.25.Final
- activemq-artemis-2.6.2

© 2021 IBM Corporation

We compiled each of the libraries on both the x86 and z test platforms to make the software stack equivalent. We are working the respective communities to make these libraries available.

Netty and Netty TCNative were built to enable native OpenSSL and epoll libraries of the Netty project. We used OpenSSL for Netty TCNative, since the default, BoringSSL, does not currently support IBM Z/s390x.

Netty and Netty TCNative are built to enable native OpenSSL and epoll libraries of the Netty project. We used OpenSSL for Netty TCNative, since the default BoringSSL does not currently support IBM Z/s390x.

We observed that enabling the native ActiveMQ library had a significant impact to improve the throughput and to eliminate the performance fluctuations (on OpenJ9), compared to the default NIO-based ActiveMQ library.

We selectively disabled processors on each system for the scalability testing. The results reported here are with no assigned processor affinity. We kept all testing local to remove the potential for latency derived from network.

The use of OpenSSL is required to achieve some of the higher throughput numbers between the nodes.

Corda Enterprise throughput is dependent on the throughput of the underlying RDBMS. For our tests, we used a RAM disk to remove this dependency.

The latency between the node and the database should be kept to a reasonable minimum.

On the x86 Test bed

The Corda software configuration is running on a single bare metal system in IBM Cloud with the following characteristics:

- 384GB of RAM (12x32GB Hynix 32GB DDR4 2Rx4)
- Forty (40) 2.5GHz x86 (6248) processors of which 32 were used for the test bed to match the LinuxONE LPAR
- 1000 Mbps public and private network
- 1TB (SATA SSD)

RAM Disk used for PostgreSQL Database and persistent storage for ActiveMQ.

On the LinuxONE Emperor II (comparable to z14) Test bed

The Corda software configuration is running on a logical partition with the following characteristics:

- 256GB of DDR4 RAM
- Up to 32 Integrated Facilities for Linux (IFL) running at 5.2GHz
- 1000 Mbps public and private network
- 3390 DASD

On the LinuxONE III (comparable to z15) Test bed

The Corda software configuration is running on a logical partition with the following characteristics:

- 256GB of DDR4 RAM
- Up to 39 Integrated Facilities for Linux (IFL) running at 5.2GHz
- 1000 Mbps public and private network
- 3390 DASD

Test Case(s)

R3 provides a number of flows as part of the sizing and performance documentation located at https://docs.corda.net/docs/corda-enterprise/4.8/performance-testing/performance-results.html For our tests we focused on the Issue + Repeated Pay workflow. Issue + Repeated Pay is a simple, yet real-world transactional flow. This flow issues some states on Node A and then repeatedly transfers a fraction of those states to Node B via the Notary. It is possible to configure the number of output states generated and transferred to Node B in each transaction in order to demonstrate the effects of doing so on throughput.

It is important to note that this flow is much more complex in terms of the peer-to-peer communications than that description makes clear. Node B will never have seen the issuance transaction that contains the input state for the payment transaction. As a result, node B enters transaction dependency resolution to request the first transaction from node A, resulting in additional sub-flows and peer-to-peer communication.

We launched the flow using the RPC client. A limited number of flows are launched in parallel in order for the node to have enough load to reflect the performance expected and exploit the multi-threaded capabilities without overwhelming it with long queues of pending work (that will form a separate scenario as we develop the performance test suite on IBM LinuxONE further).

We measure the time taken from the time just before we request the execution of a flow from the RPC client to the time after we see the *Future* returned from startFlowRPC call complete on the client. At this point the transaction is recorded in all nodes that participate in the transactions and all sub-flows are complete.

In future versions of this paper, we may expand the range of scenarios to cover some in between, and some much more complex; involving more steps, greater variety of transaction sizes and/or more nodes with the hope that one of these could act as a proxy for your own workload. No workloads are the same and therefore any debate around sizing naturally leads to conversations around what type of flows, what size transactions involving what kinds of states and contracts. We can therefore only give you an idea of what might actually be required and/or possible.

© 2021 IBM Corporation

Results

We ran many iterations of our test suite across our 3 test systems to identify the most consistent results across each. Conveniently, the highest sustainable TPS for each platform was derived from the same basic configuration: No processor affinity assigned, OpenSSL, and PostgreSQL as the back-end database. The PostgreSQL instance was running in a RAM disk to minimize latency derived from I/O. Figure 2 below shows the performance characteristics of both systems we compared. The key take-aways are that the transactional throughput of z15 running Corda Enterprise 4.1 was essentially equivalent to a compared 32-core x86 system running with half the cores:

- in the *Issue* phase, as seen in the ramp up at the beginning of the test run
- and in the *Repeated Pay* phase of the test.

We observed that the *Repeated Pay* phase showed consistent performance on both tested platforms after the *Issue* phase. These results were consistent across numerous iterations of the tests, each of which was 30 minutes in duration. The compared x86 results are also broadly consistent with previously published performance tests performed by R3 and third parties.[4]

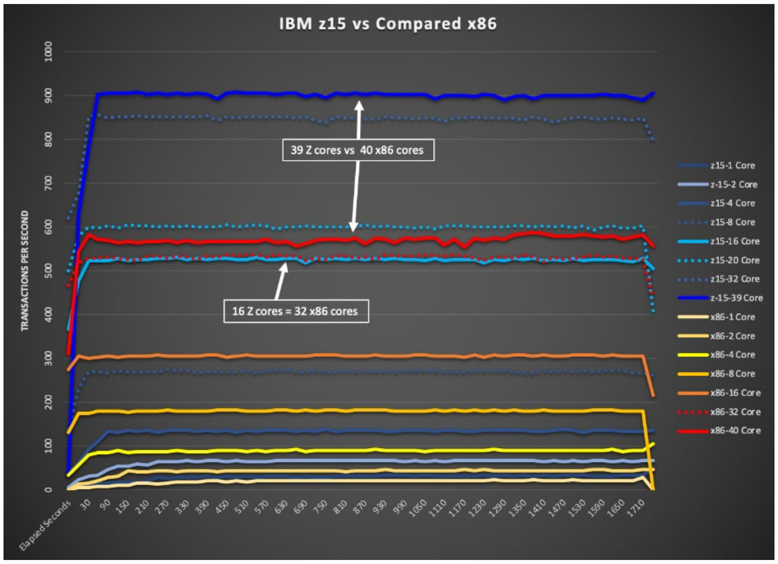

Figure 2. Comparative performance

Figure 3 demonstrates roughly linear scale with Linux on z15/LinuxONE III (2x TPS with 2x cores) up to 16 cores, after which the performance gain is still significant but less than 2x. This is indicative of running the entire stack in a single logical partition/system. As can been seen in Figure 4, the average CPU utilization of the 32-core system drops significantly, just over 20%, from the average CPU utilization of the 16-core system. As seen in Figure 3, Corda Enterprise 4.1 scales very well on IBM LinuxONE.

[4] Refer to https://docs.corda.net/docs/corda-enterprise/4.8/performance-testing/performance-results.html for details of the R3 testing.

© 2021 IBM Corporation

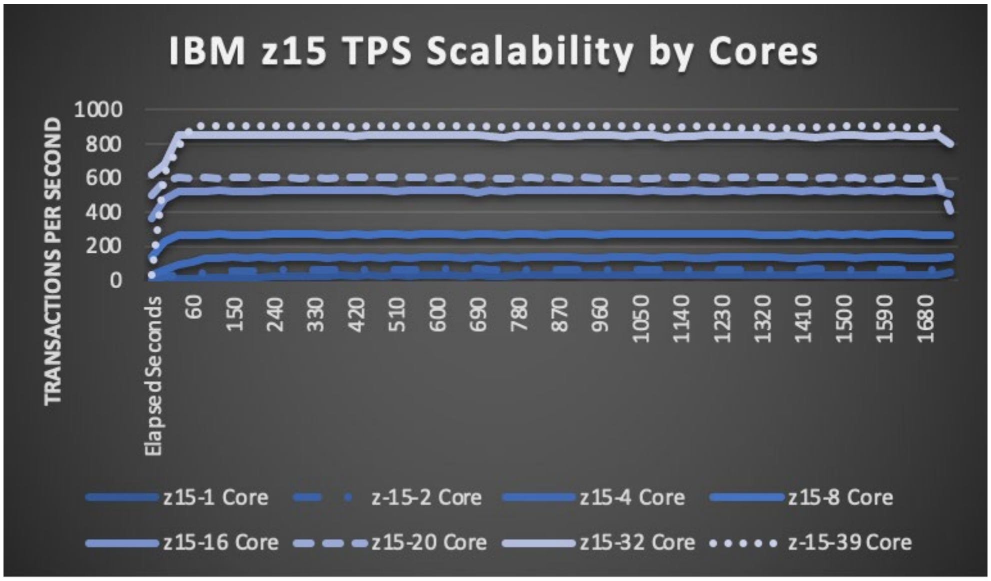

Figure 3. Scalability test on IBM z15

For the purposes of our test, we performed a simple, non-HA, installation of the Corda Enterprise platform. We expect that performance would benefit even more from a multi-LPAR or system configuration with a dedicated database node as long as the network latency between the nodes and the databases does not introduce additional delays.

Finally, Figure 4 shows the scalability characteristics of z15 and the compared x86 when the number of cores is varied. While both processors show non-linear speedup as the number of cores increased, z15 exhibits a better scalability than the compared x86. Running an identical Corda Enterprise workload, the z15 achieved 601 transactions per second using 20 cores compared to 571 transactions per second on a compared X86 configuration using 40 cores.

We observed z15 generally exhibits a lower CPU utilization (1-3% less) than the compared x86 at the same number of cores when running Corda Enterprise 4.1, z15 is expected to have a bigger potential to improve the performance further through software and deployment optimizations.

© 2021 IBM Corporation

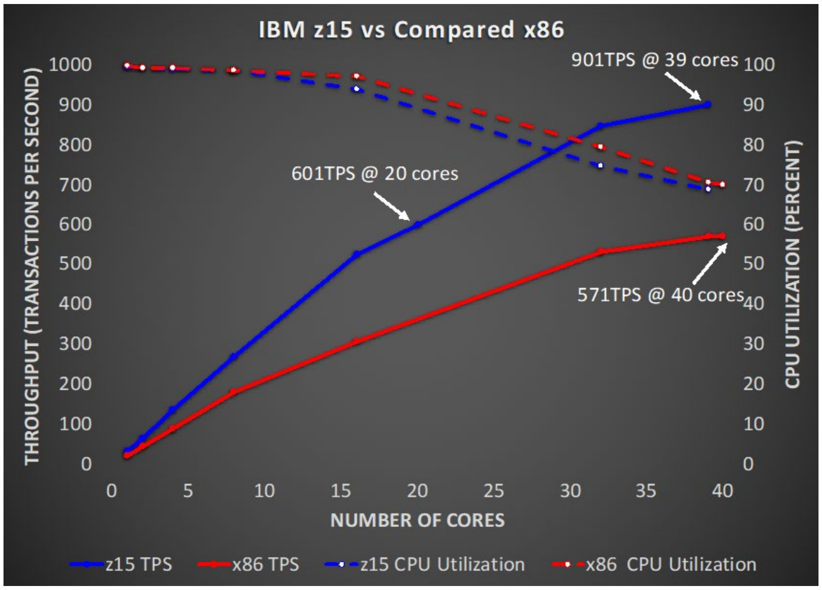

Figure 4. Scalability of z15 and x86

Observations Summary

- On the tested IBM LinuxONE III/z15 we achieved a peak average TPS of 901 running on the 39-core system (See figure 2)
- On IBM test systems, Corda Enterprise 4.1 on LinuxONE III/z15 achieves 1.7 greater throughput with a peak of 908 TPS utilizing 39 cores vs. the compared x86 system with a peak of 587 TPS utilizing 40 cores.
  (See figure 2)
- On IBM test systems, even a single core deployment of Corda Enterprise 4.1 offers 50% more transactional throughput on LinuxONE III (29 TPS average) versus the compared x86 system (19 TPS average). (See figure 2)
- In this deployment, we observed Corda Enterprise 4.1 can utilize servers in excess of 16 cores but peak efficiency for the stack is around 16 cores when running everything on a single system. (See figure 2)
- We observed that Corda Enterprise can scale more than 20x on IBM LinuxONE III/z15 by adding more cores for the nodes used in the benchmarks. We observed that a single core z15 delivered a peak of 32.8 TPS and a 39-core z15 delivered 908 at peak. (See figure 3)

Securing the Platform

Blockchain networks work on the basis of “Trust, but verify.” A node in a blockchain network is designed to not accept data from any other node without first validating the data and any historical data upon which it depends. This starts the foundation of security. The authority to transact in this fashion is delegated to the node by its owner/operator. A node's role is to work on behalf of its operator. This is the design of most blockchain platforms, including Corda. This model assumes that the environment in which a node is hosted can be trusted and that the primary threat vectors are external to the organization owning the nodes. It is, however, becoming more and more evident that malicious actors can come from a companies own employees or their service providers. Internal breaches of information or intentional reconfiguration of a company’s nodes could lead to significant disruption. Therefore, it is also important to ensure, to the greatest extent

© 2021 IBM Corporation

possible, that the infrastructure upon which a blockchain node depends (compute, storage, database) is uncompromised.

IBM Cloud Hyper Protect Virtual Servers (in the IBM Cloud or on-premises) using Secured Service Containers are designed to deliver application-level security protection to defend against internal and external attack vectors. The IBM Secure Service Container is unique to IBM LinuxONE III/z15 hardware.

IBM Hyper Protect Virtual Servers provide a secure virtualized infrastructure for private cloud deployments and protect the entire lifecycle of critical Linux workloads during their build, deployment and management on-premises.

Running Corda in Hyper Protect Virtual Server instances on the Secure Service Container partitions provides the following advantages in terms of security and integrity.

- System administrators do not need the access to the application data, memory, logs, secrets, applications or the operating system in the virtual server instances.
- Application developers do not need the secret to the production environment.
- Management of the virtual server instances does not require access to the application secrets.

Roadmap

The purpose of this collaboration between IBM and R3 is to define a long-term roadmap for the implementation of Corda Enterprise on IBM LinuxONE. At the time of writing, our initial approach is to assist R3 in the support of Corda Enterprise 4.x on IBM LinuxONE, z14, and z15 running in the IBM Hyper Protect Virtual Servers on-premises offering, as well as running on IBM Hyper Protect Virtual Servers in the IBM public cloud. Moving forward, we intend to investigate further avenues of integration.

Corda Enterprise support for IBM hardware security modules, both on-premises with the Crypto Express family and in the public cloud with IBM Cloud Hyper Protect Crypto Services may be included in the platform to deliver additional client security capabilities. Support for the IBM HSMs would provide support for the only FIPS140-2 Level 4 certified HSM available in several compared public cloud models in addition to providing that same level of security on-premises.

It is our further intent to develop and publish client-driven implementation patterns for scalable and secure implementations.

Conclusion

Choice of infrastructure can have a significant impact on both the performance and security of a Corda Enterprise network. IBM z15 and LinuxONE III provide a vertically scalable and high-performance platform. The capabilities of the IBM Secure Service Container utilized by the IBM Hyper Protect Virtual Servers offerings are designed to deliver a secured mode in which to run Corda Enterprise nodes. Whether running on-premises or in the public cloud, IBM Hyper Protect Virtual Servers provide a robust foundation of security and performance for Corda Enterprise deployments.

© 2021 IBM Corporation

Deployment Options

Hyper Protect Virtual Servers on-premises: Generally available November 2019
Hyper Protect Virtual Server in the IBM public cloud: Generally available November 2019 @ https://cloud.ibm.com/

About R3

R3 is an enterprise blockchain software firm working with a global ecosystem of more than 300 participants across multiple industries from both the private and public sectors to develop on Corda, its open source blockchain platform, and Corda Enterprise, a commercial version of Corda for enterprise usage.

R3's global team of over 200 professionals in 13 countries is supported by over 2,000 technology, financial, and legal experts drawn from its vibrant ecosystem.

The Corda platform is already being used in industries from financial services to healthcare, shipping, insurance and more. It records, manages and executes institutions' financial agreements in perfect synchrony with their peers, creating a world of frictionless commerce. Learn more at www.r3.com and www.corda.net.

Acknowledgements

The authors of this paper acknowledge, with gratitude, the support we received from the entire R3 team in authoring this paper. Specifically, Richard Parker, Moritz Platt and Richard Brown made significant contributions to the success of our tests and the material documented here.

© 2021 IBM Corporation

(C) Copyright IBM Corporation 2021
IBM Corporation
New Orchard Road
Armonk, NY 10504
U.S.A.
07/21

IBM, the IBM logo, IBM Z, IBM Cloud, PR/SM and z15 are trademarks or registered trademarks of the International Business Machines Corporation.

A current list of IBM trademarks is available on the Web at https://www.ibm.com/legal/us/en/copytrade.shtml, and select third party trademarks that might be referenced in this document is available at https://www.ibm.com/legal/us/en/copytrade.shtml#section_4.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

InfiniBand and InfiniBand Trade Association are registered trademarks of the InfiniBand Trade Association.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

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.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

OpenStack is a trademark of OpenStack LLC. The OpenStack trademark policy is available on the OpenStack website.

Red Hat®, JBoss®, OpenShift®, Fedora®, Hibernate®, Ansible®, CloudForms®, RHCA®, RHCE®, RHCSA®, Ceph®, and Gluster® are trademarks or registered trademarks of Red Hat, Inc. or its subsidiaries in the United States and other countries.

RStudio®, the RStudio logo and Shiny® are registered trademarks of RStudio, Inc.

TEALEAF is a registered trademark of Tealeaf, an IBM Company.

Temenos, Transact and T24 are registered trademark of Temenos AG.

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

Worklight is a trademark or registered trademark of Worklight, an IBM Company.

Zowe™, the Zowe™ logo and the Open Mainframe Project™ are trademarks of The Linux Foundation.

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

The information contained in this documentation is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this documentation, it is provided “as is” without warranty of any kind, express or implied. In addition, this information is based on IBM’s current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this documentation or any other documentation. Nothing contained in this documentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM (or its suppliers or licensors), or altering the terms and conditions of the applicable license agreement governing the use of IBM software.

References in these materials to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates. Product release dates and/or capabilities referenced in these materials may change at any time at IBM’s sole discretion based on market opportunities or other factors and are not intended to be a commitment to future product or feature availability in any way.

No other publication, distribution or use is permitted without the prior written consent of IBM. Customers who want a "deep drill down" on IBM IT Economics assessments should be directed to the IBM IT Economics Research & Consulting team under NDA. An NDA is required to explain IBM IT Economics methodologies, processes and competitive comparison. You can contact the IBM IT Economics team by sending an email to it.economics@us.ibm.com

16039016-USEN-02