IBM Z

IBM z14
OSA-Express7S 25 GbE
Performance Report

version 2019-04-19

© 2019 IBM Corporation

® Copyright IBM Corporation 2019
IBM Corporation
New Orchard Road
Armonk, NY 10504
U.S.A.
Produced in the United States of America,
4/2019

- The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States or other countries or both: IBM, IBM logo, IBM Z, IBM z14, zEnterprise and z/OS are trademarks or registered trademarks of the International Business Machines Corporation.
- 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.
- Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
- All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
- This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates. It is the user's responsibility to evaluate and verify the operation of any other products or programs with IBM products and programs.
- This document contains performance information. Performance is based on measurements and projections using standard IBM tools in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as machine driver level including OSA firmware level, LPAR configuration, operating system level, network hardware, network settings, the amount of multiprogramming in the user's job stream, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the numbers stated here.
- The information herein is provided "AS IS" with no warranties, express or implied, including without any warranties of merchantability, fitness for a particular purpose and any warranty or condition of non-infringement. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.
- Refer to www.ibm.com/legal/us for further legal information.

© 2019 IBM Corporation

Document Outline

- OSA-Express7S 25 Gigabit Ethernet adapter
- OSA Configuration Diagram
- Measurement Details
- Performance Charts
- Summary
- Key Terms
- Performance References (including where to find this document)
- Document History

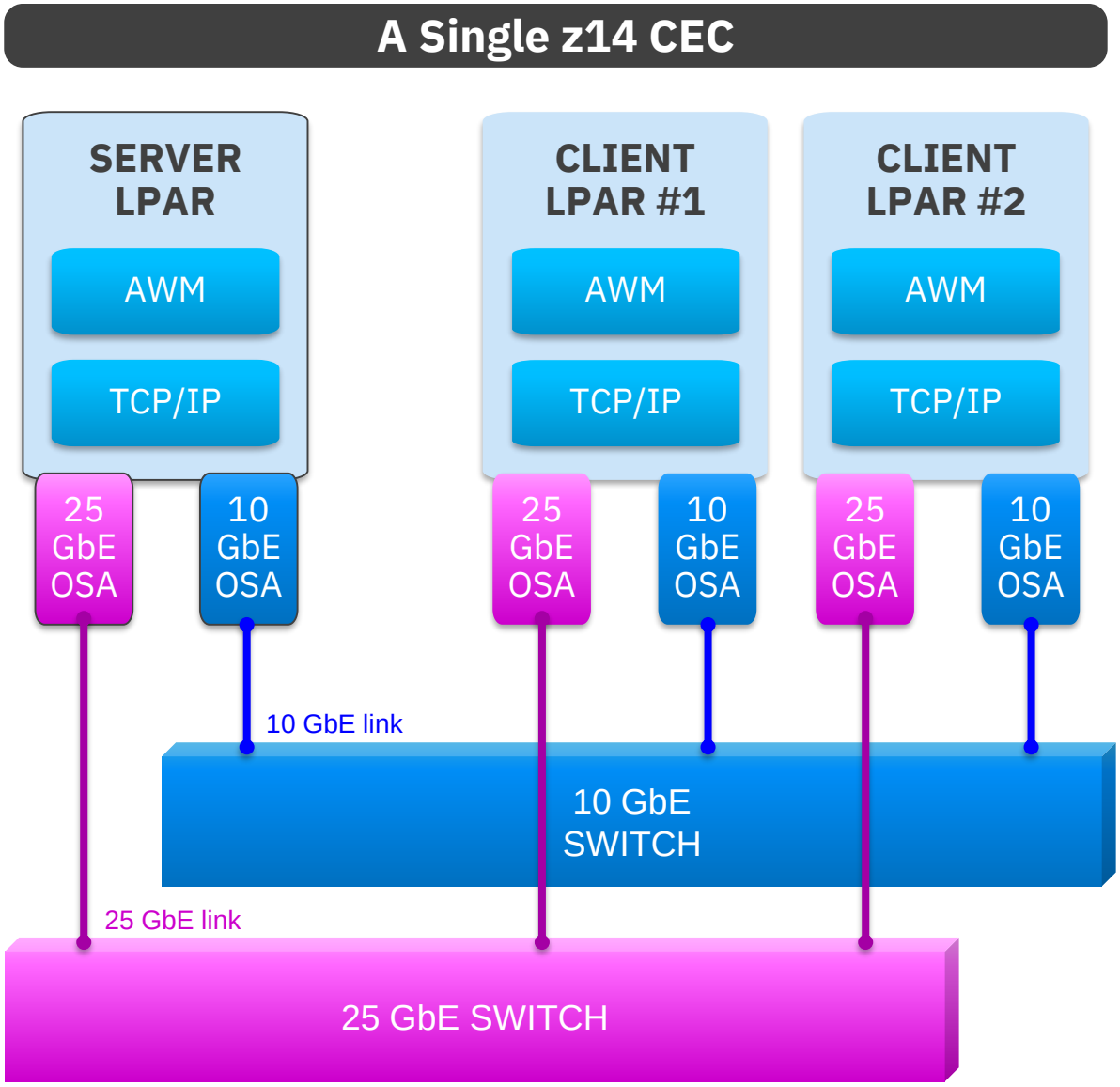
© 2019 IBM Corporation

OSA-Express7S 25 Gigabit Ethernet adapter

- Feature #0429
- Available as of April 9, 2019 on z14 (driver D36C)
- Benefits from prior-level OSA-Express6S 10 GbE performance
- Adds a new network interface controller (NIC) module unit and significant throughput
- There is one PCIe adapter and one port per feature (same as OSA-Express6S 10 GbE)
- This adapter allows for some potential consolidation of prior level adapters, depending on usage characteristics

© 2019 IBM Corporation

OSA Configuration Diagram



- The server OSAs and LPAR were the focus of the measurements
- OSA-Express7S 25 GbE and OSA-Express6S 10 GbE were measured independently
- AWM, an in-house socket application, was used to transfer network traffic via TCP/IP across the Ethernet between program memory on the client and and server LPARs
- The entire machine environment, including the network, was solely dedicated to the measurement

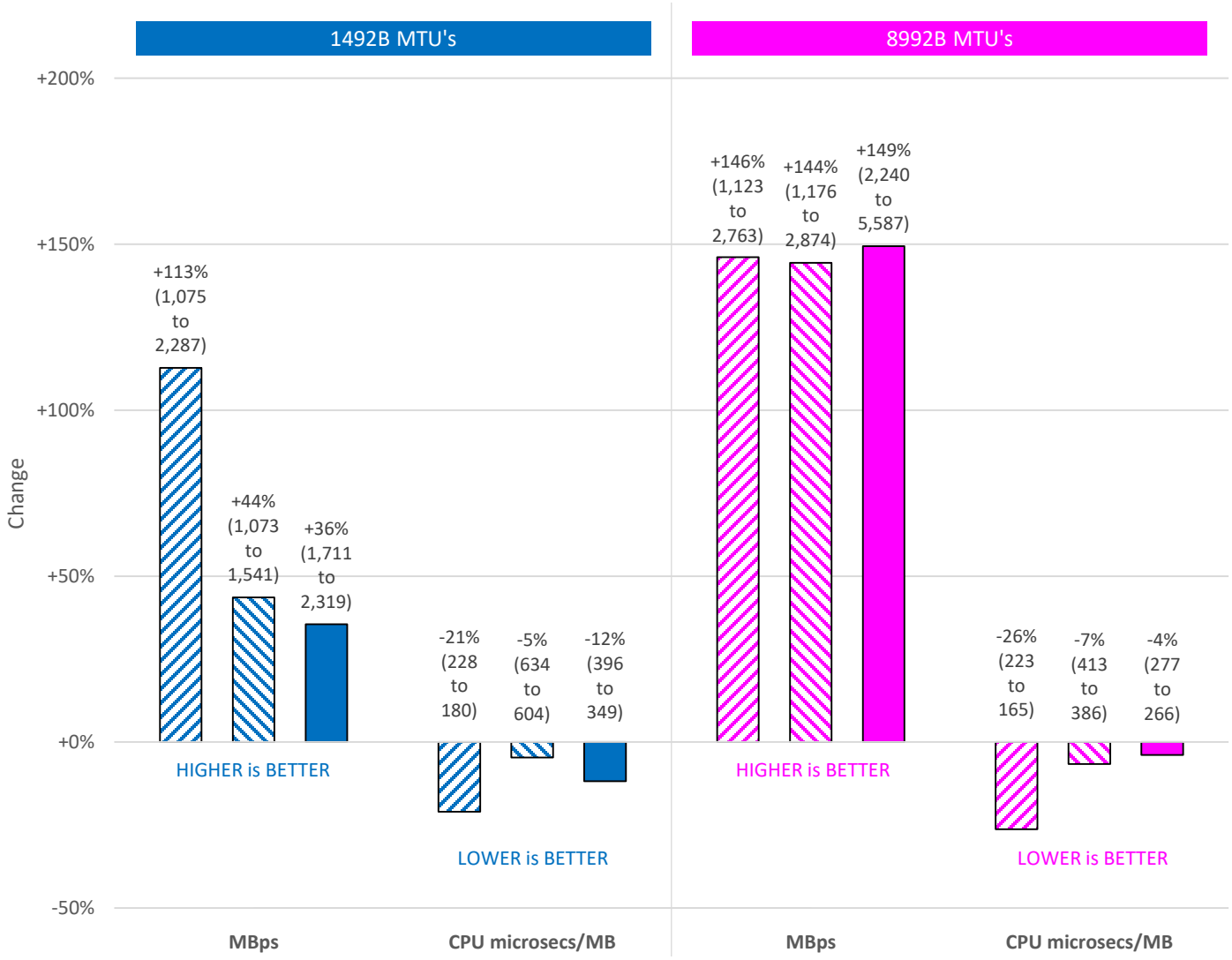
© 2019 IBM Corporation

Measurement Details

Machine	IBM z14 Type 3906 Model M05 (driver D36C)		
LPAR	Three LPARs each with 8 CPUs (dedicated)		
Operating system	z/OS V2R3 in each LPAR		
OSA	**OSA-Express7S 25 GbE SR (feature #0429)** • One per LPAR • QDIO Layer 3 mode (OSD) • Firmware at pre-GA level • 31 meter Ethernet fiber	VS	**OSA-Express6S 10 GbE LR (feature #0424)** • One per LPAR • QDIO Layer 3 mode (OSD) • Firmware at MCL P41446.008 • 22 meter Ethernet fiber
Ethernet switch	25 GbE switch		10 GbE switch
Protocol	TCP/IP version 4		
TCP settings	INBPERF = DYNAMIC with WORKLOADQ SEGMENTATIONOFFLOAD WORKLOADQ is referred to as Inbound Workload Queueing (IWQ) SEGMENTATIONOFFLOAD is referred to as or Large Send Offload (LSO) or TCP segmentation offload (TSO)		
Network buffers	65,535 byte socket send and receive buffers TCP send/receive buffers: 184,320 bytes on the server & 131,071 bytes on the clients Best practice would be to use 184,320 on all z/OS stacks		
Traffic bulk data streams	INBOUND to the server	OUTBOUND from the server	MIX of inbound and outbound
Network MTUs	1492B		8992B

© 2019 IBM Corporation

OSA-E6S 10 GbE to OSA-E7S 25 GbE Change in Streams Throughput & Host Costs
z/OS V2R3 on z14
1492B MTU's vs 8992B MTU's



LEGEND

OUTBOUND (5)

INBOUND (5)

MIX IN+OUT (20)

The number of active TCP connections is shown above in parentheses

COLOR SCHEME

- 1492B MTU's in blue
- 8992B MTU's in pink

THROUGHPUT

TCP payload MB per second

MEGABYTE (MB)

1 MB is 1,048,576 bytes

POINT SELECTION

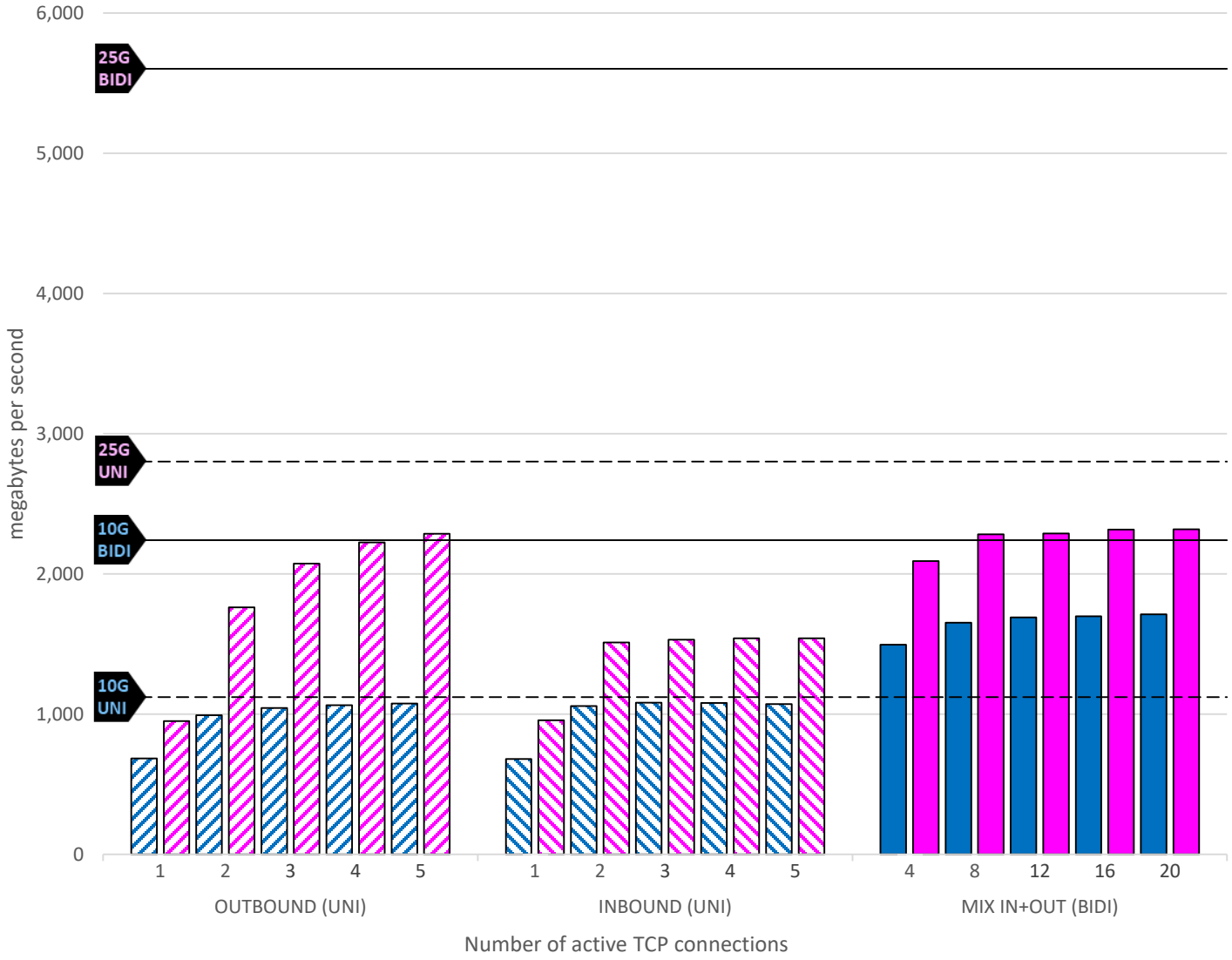
- The selected points do not necessarily represent the greatest achieved throughput
- For instance, the maximum throughput seen for 8992B MTU's is 5600 MBps

CONFIGURATION

See page Measurement Details

© 2019 IBM Corporation

OSA-E6S 10 GbE & OSA-E7S 25 GbE Streams TCP Payload Throughput
z/OS V2R3 on z14
1492B MTU's



COLOR SCHEME
- 10 GbE in blue
- 25 GbE in pink

MEGABYTE (MB)
1 MB is 1,048,576 bytes

EFFECTIVE LINK CAPACITY
A rough estimate of the bandwidth available for the TCP payload after considering TCP/IP and Ethernet overheads

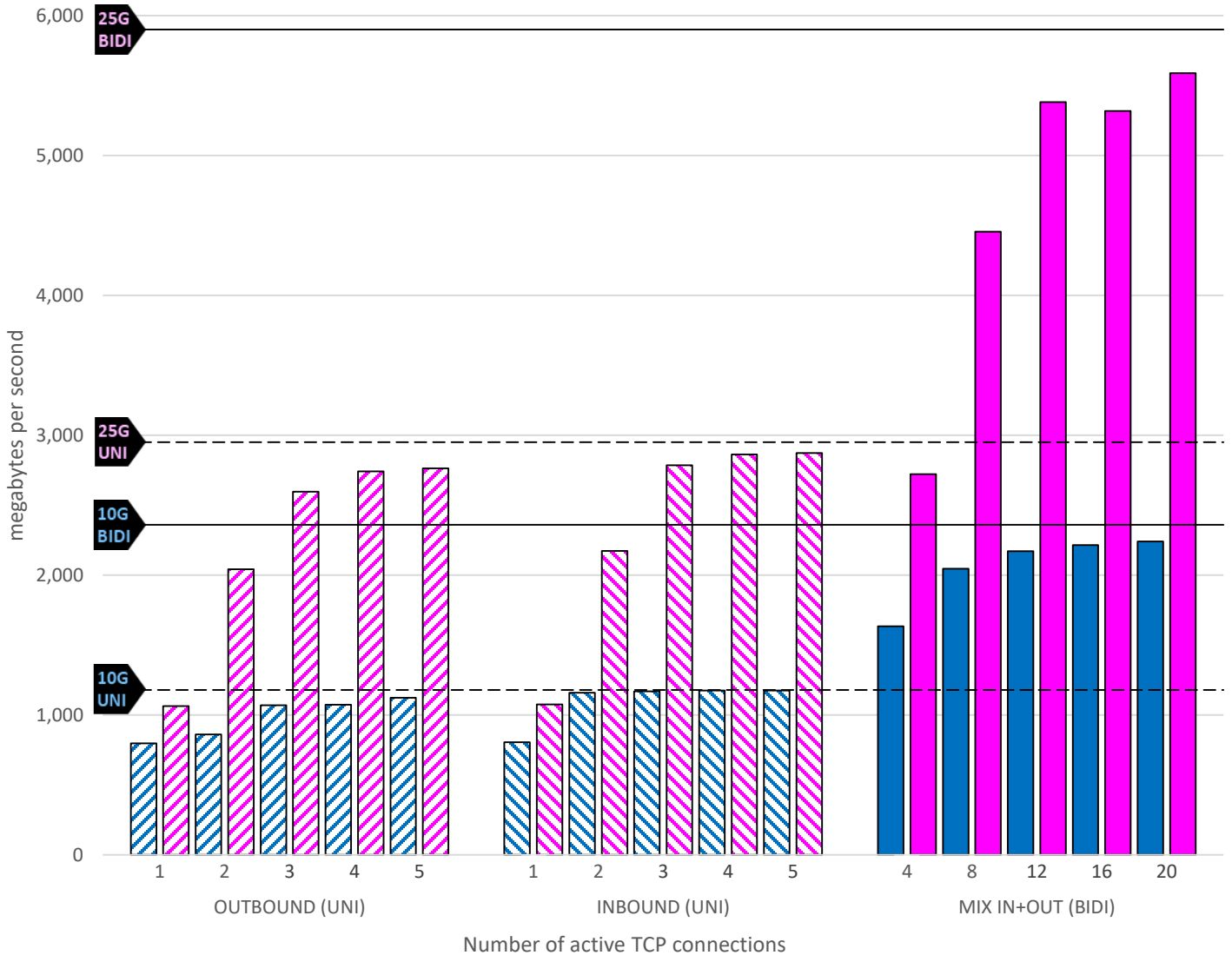
BIDI & UNI
- BIDI = bi-directional traffic
- UNI = uni-directional traffic

ANY SINGLE CONNECTION
The performance of any single connection depends on MTU size, transfer size, TCP Window size, the number of concurrent connections and other factors.

CONFIGURATION
See page Measurement Details

© 2019 IBM Corporation

OSA-E6S 10 GbE & OSA-E7S 25 GbE Streams TCP Payload Throughput
z/OS V2R3 on z14
8992B MTU's



COLOR SCHEME
- 10 GbE in blue
- 25 GbE in pink

MEGABYTE (MB)
1 MB is 1,048,576 bytes

EFFECTIVE LINK CAPACITY
A rough estimate of the bandwidth available for the TCP payload after considering TCP/IP and Ethernet overheads

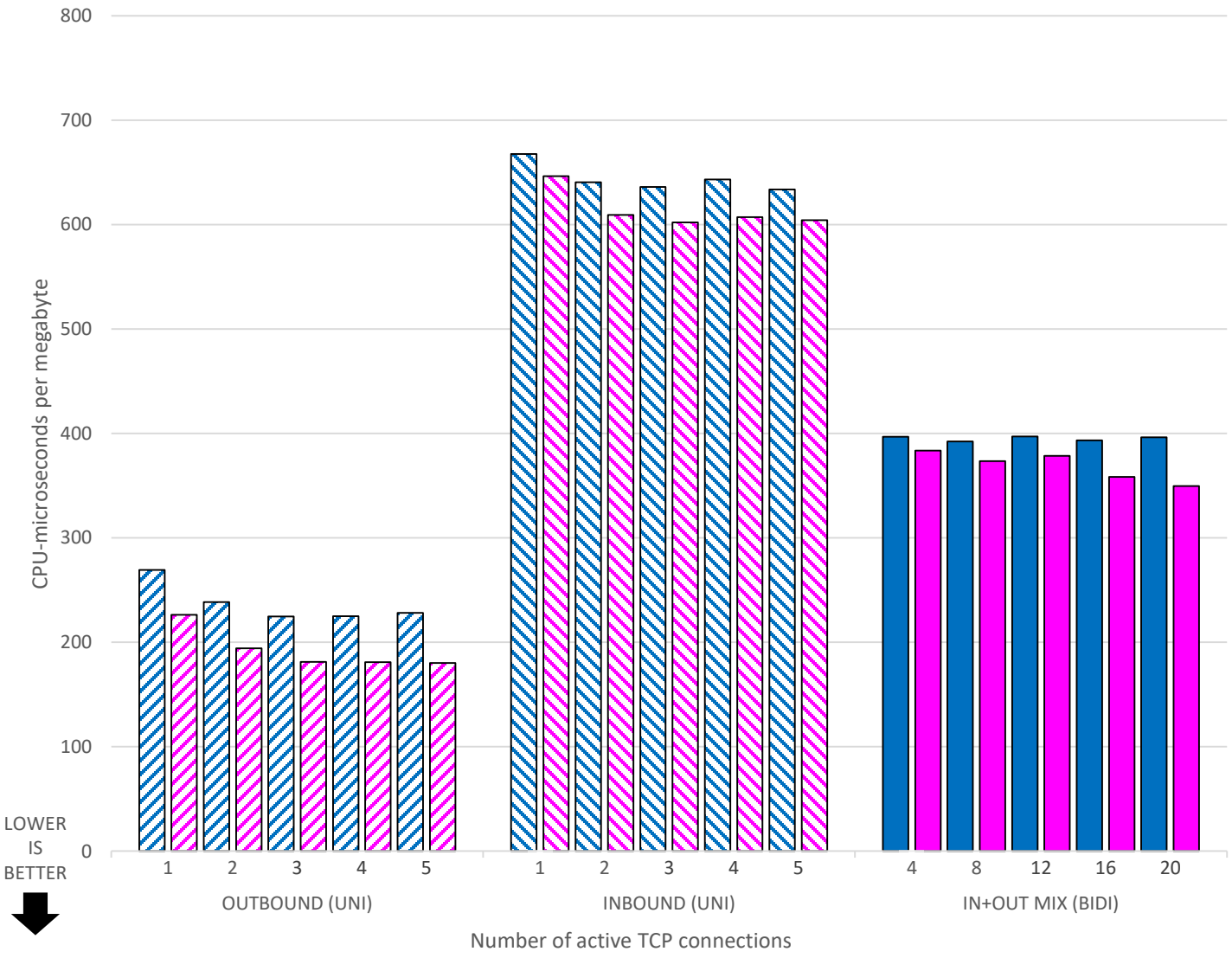
BIDI & UNI
- BIDI = bi-directional traffic
- UNI = uni-directional traffic

ANY SINGLE CONNECTION
The performance of any single connection depends on MTU size, transfer size, TCP Window size, the number of concurrent connections and other factors.

CONFIGURATION
See page Measurement Details

© 2019 IBM Corporation

OSA-E6S 10 GbE & OSA-E7S 25 GbE Server LPAR CPU Cost for Streams
z/OS V2R3 on z14
1492B MTU's



COLOR SCHEME
- 10 GbE in blue
- 25 GbE in pink

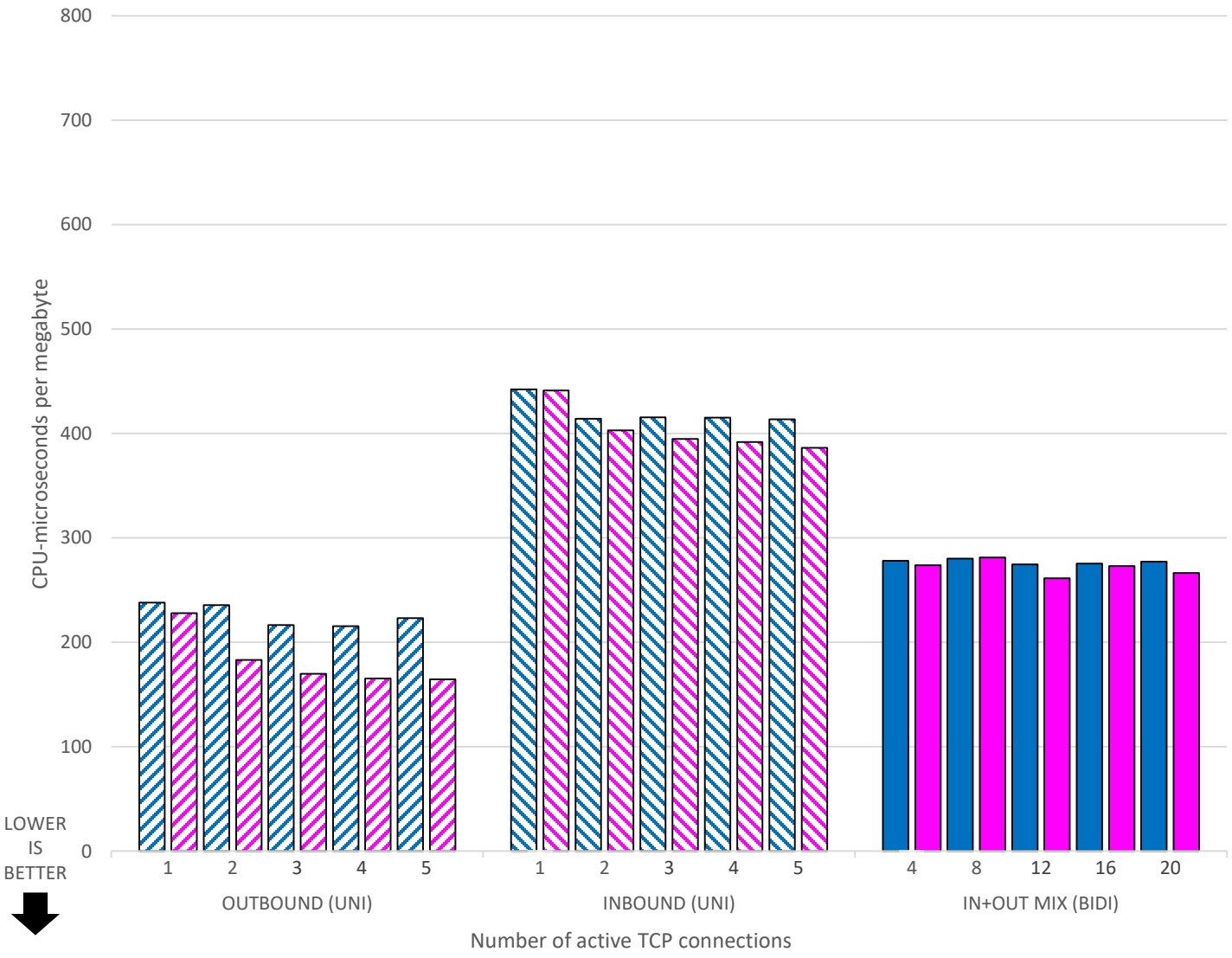
MEGABYTE (MB)
1 MB is 1,048,576 bytes

BIDI & UNI
- BIDI = bi-directional traffic
- UNI = uni-directional traffic

CONFIGURATION
See page Measurement Details

© 2019 IBM Corporation

OSA-E6S 10 GbE & OSA-E7S 25 GbE Server LPAR CPU Cost for Streams
z/OS V2R3 on z14
8992B MTU's



COLOR SCHEME
- 10 GbE in blue
- 25 GbE in pink

MEGABYTE (MB)
1 MB is 1,048,576 bytes

BIDI & UNI
- BIDI = bi-directional traffic
- UNI = uni-directional traffic

CONFIGURATION
See page Measurement Details

© 2019 IBM Corporation

Summary

Main Messages

OSA-Express7S 25 GbE, operating in QDIO mode (OSD) with 8992-byte MTUs, achieved **94%** of the effective link capacity when measuring the inbound, outbound, and mixed-direction TCP/IP streams workloads on z/OS. For outbound data streams this represents a **146%** improvement in throughput over OSA-Express6S 10 GbE, while improving host CPU costs per megabyte by **26%**.

In general, with 8992-byte MTUs OSA-E7S 25 GbE delivered nearly **150%** more streams throughput than OSA-E6S 10 GbE.

For 1492-byte MTUs, OSA-E7S 25 GbE increased outbound data streams throughput over OSA-E6S 10 GbE by **113%** over while improving CPU costs by **21%**. Throughput for inbound and mixed-direction streams increased by **44%** and **36%** respectively. These are important facts for those wanting to consolidate multiple 10 GbE OSAs onto a single OSA-E7S 25 GbE.

Additional Messages

If your networking infrastructure supports jumbo frames end-to-end consider increasing the size of the MTU to 8992 bytes, especially if the network is frequently moving streams of data or carrying interactive messages that span multiple standard size Ethernet frames.

Disabling LSO for outbound and mixed-direction streams type traffic will likely degrade performance for smaller MTU's.

© 2019 IBM Corporation

Key Terms (page 1 of 2)

AWM	• AWM is a TCP socket application that is focused on measuring the application socket API and the associated communications cost. AWM does not have any back-end business logic, Disk I/O or compute logic. This performance modeling approach is often referred to as a "micro benchmark" (in this case focused on the processing related to communications).
CPU	• The CPU is the central processing unit of the IBM z14 central processor complex (CPC). • In this report CPU microseconds is focused on capturing the measurement of the CPU utilization associated with cost of network communications.

Effective link capacity	• A rough estimate of the bandwidth available for TCP payload throughput taking into account the byte overheads associated with
	  • TCP/IP version 4 headers with 12 bytes of TCP options
	  • 25GBASE-R Ethernet with Reed-Solomon forward error correction
	• The calculation assumes
	  • All Ethernet packets carry a full payload consisting of one TCP/IP packet of the stated MTU size
	  • Ethernet packets are transmitted back-to-back onto the link
	  • Ethernet switch is not a bottleneck
	• The calculation does not include estimates of any other traffic (for example, Ethernet PAUSE frames, application control packets, and TCP/IP retransmissions)

Ethernet Frame	• A frame is essentially a container with a payload and various bits of information that help the frame reach the other end of the Ethernet link. The payload, in this report, consists of a TCP/IP packet. • A standard Ethernet frame can hold a 1500B MTU; a jumbo frame can hold a 9000B MTU. • In this report we use 1492B and 8992B in keeping with past measurements
Large Send Offload (LSO)	• The z/OS term for TCP Segmentation Offload
Megabyte (MB)	• In this report a megabyte (MB) refers to 1,048,576 bytes. • The standardized term for this unit is the mebibyte.

© 2019 IBM Corporation

Key Terms (page 2 of 2)

MTU	• MTU is Maximum Transmission Unit. This is the maximum number of bytes that can be put into the payload of an Ethernet frame. • For this report the MTU includes a 20-byte IP header and a 32-byte TCP header (including 12 bytes of options), thus the maximum TCP payload for 1492B and 8992B MTUs is 1440 bytes and 8992 bytes, respectively.	
NIC module unit	• The network interface controller (NIC) module unit on the OSA adapter is that component implementing the Ethernet network interface.	
Streams	• All transactions are initiated at the CLIENT. • An OUTBOUND stream transaction starts with the client sending a 20 byte request-message to the server followed by 20MB (approximately) from the server back to the client. OUTBOUND sends data traffic OUT of the server. • An INBOUND stream transaction starts with the client sending 20MB (approximately) to the server followed by 20 byte response-message in reply. INBOUND sends data traffic IN to the server. • MIXED is a blend of concurrently active INBOUND and OUTBOUND streams on the same Ethernet. • Each transaction is repeated, one after the previous, for three minutes on each active TCP/IP connection. • Performance is measured over a 60-second window during the middle of the run.	
Throughput	• Throughput refers to the successful rate of delivery of bytes or completion of transactions. • In this report throughput is the rate at which TCP/IP payload bytes are successfully transferred between the client and server LPARs. This is expressed in megabytes per second (MBps).	
Uni-directional (UNI) **Bi-directional (BIDI)**	• Uni-directional traffic is that in which there is data traffic flowing across the Ethernet link in only one direction at the time. • OUTBOUND and INBOUND streams are uni-directional.	• Bi-directional traffic is that in which there is data traffic flowing across the Ethernet link in both directions at the same time. • MIXED streams are bi-directional.

© 2019 IBM Corporation

z/OS Communications Server Performance References

- z/OS Communications Server performance index
  - http://www.ibm.com/support/docview.wss?rs=852&uid=swg27005524
  - An index to all published performance information for the z/OS Communications Server
  - The index is updated whenever document updates are made or documentation is added
- z/OS on the IBM Knowledge Center
  - https://www.ibm.com/support/knowledgecenter/en/SSLTBW
- IBM Knowledge Center
  - https://www.ibm.com/support/knowledgecenter/en
- Where to find this document
  - On the z/OS Communications Server performance index (above)

© 2019 IBM Corporation

Document History

VERSION DATE	DESCRIPTION
2019-04-19	• Initial release comparing OSA-Express7S 25 GbE and OSA-Express6S 10 GbE on z/OS

© 2019 IBM Corporation