

LOB COMPRESSION

DB2 V12

DANIEL KIM

12/14/16

INTRODUCTIONS

- Who am I?
- What are LOBs?
- What is zEDC?
- Why is it important?

LARGE OBJECT FILES

- Introduced in DB2 V6
 - From an application side, LOBs look like normal data column in a table
- Types of LOBs:
 - CLOB (Character LOB)
 - BLOB (Binary LOB)
 - DBLOB (Double Byte LOB)

ZEDC

- Industry standard data compression accelerator hardware card
 - Works as a PCIe card
- Can be installed into zOS and work with DB2 V12
- Get all the benefits of data compression, without the CPU degradation
- Why use it?

LOB COMPRESSION: RATIOS

- Different types of LOBs will yield different compression ratios
- More generally, CLOBs will compress better than BLOBs

Document Type	DSN1COMP	RBA
Customer Catalog	82% / 83%	74%
Word	66% ~ 73%	67% ~ 73%
XML	96%	50% ~ 94%
JPG	0%	0%

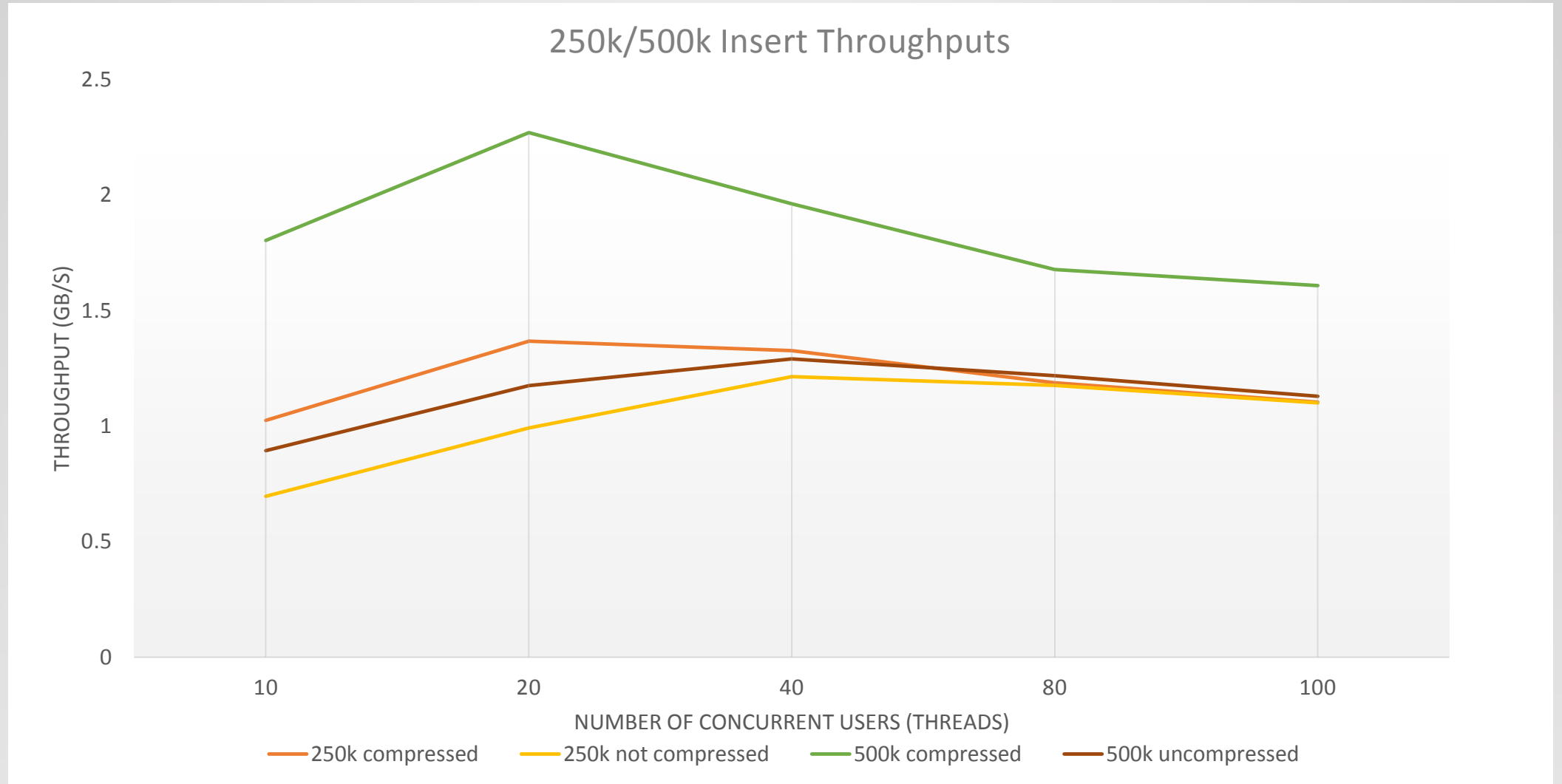
COMPRESSION BENEFITS: STORAGE SPACE

- A hypothetical situation...
 - 20K LOB
 - 4k Page size
 - 3k LOB with 4 k Page size...
- Prior to LOB compression, we would take up 5 pages to store this LOB
- Say we are able to achieve 60% compression:
 - 20K LOB size => 8k LOB size
 - 5 pages => 2 pages

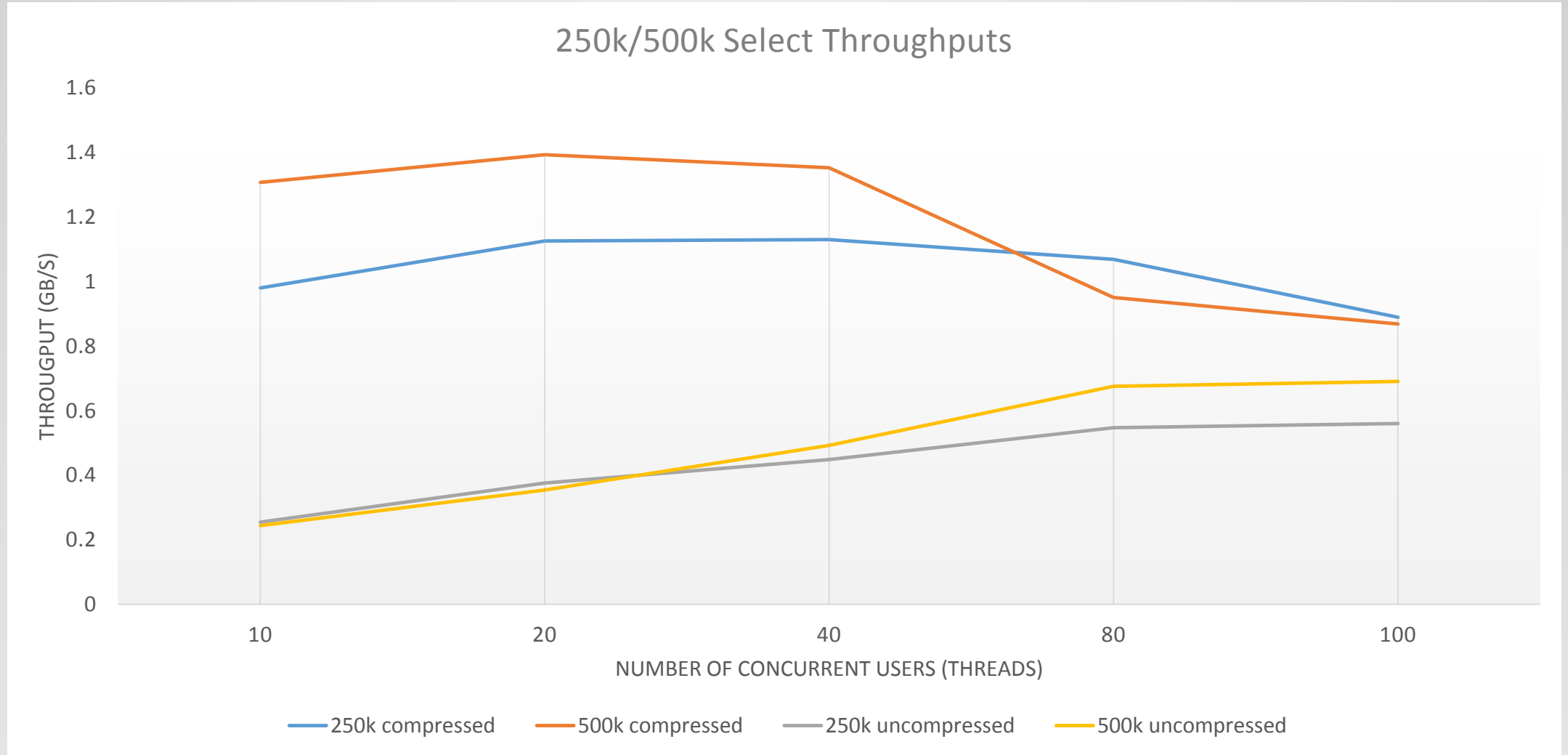
BENEFITS: THROUGHPUT

- When compression occurs, a LOB that has less pages will take less overhead to store
 - 5 pages written to DASD v.s. 2 pages
- I/O time saved v.s. CPU time taken for compression
 - The zEDC makes it possible to get best of both worlds
 - Results to follow shortly

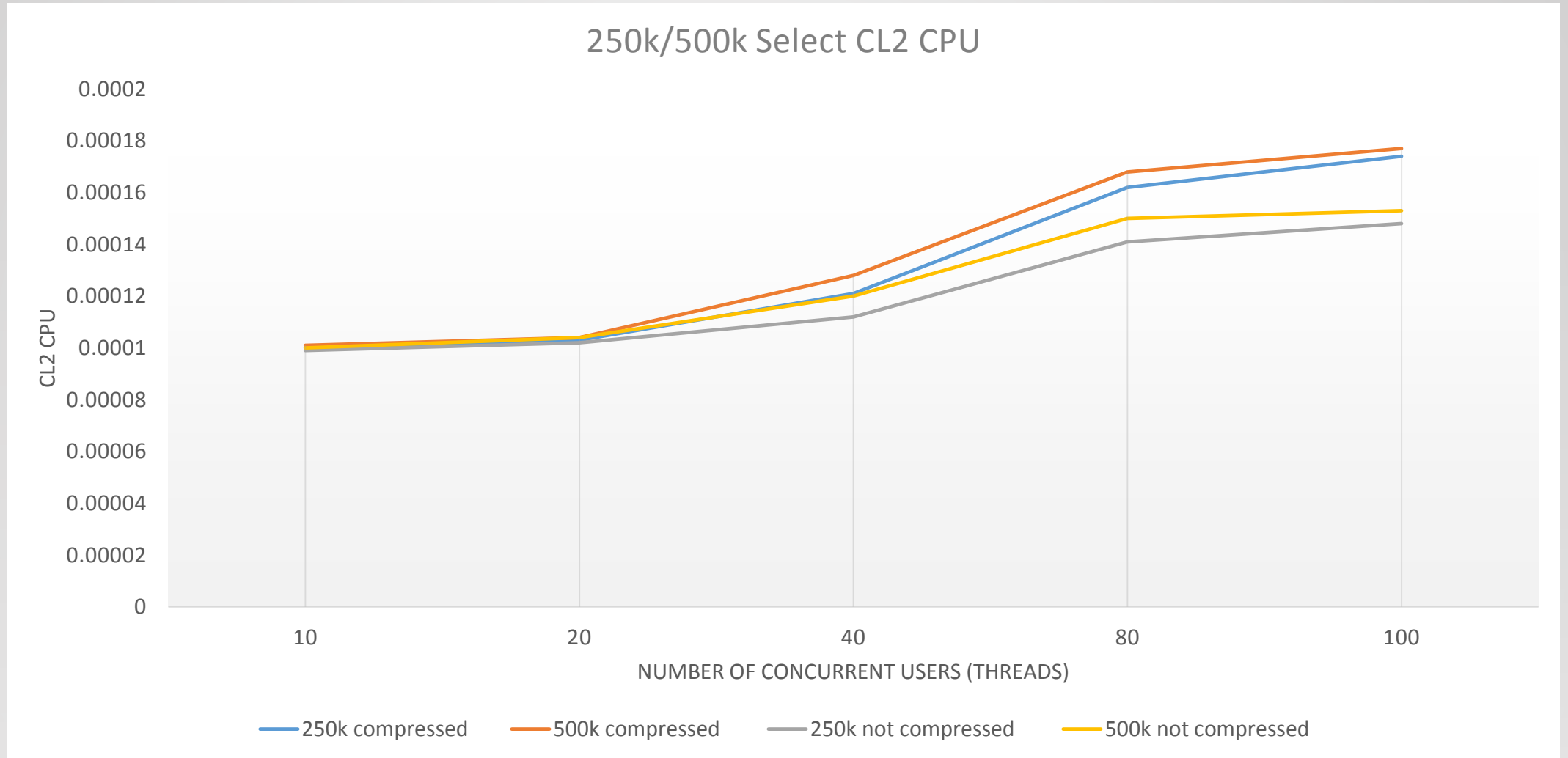
THROUGHPUT: INSERTS



THROUGHPUT: SELECT



CPU: SELECT



ZEDC BOTTLENECKS

- We see that increasing number of users does not result in direct linear growth
- There are two bottlenecks that are observable:
 - 1) 2 GB/S limitation as stated on the zEDC documentation
 - 2) zEDC queue times with more threads added

DSN1COMP

- Tool to view usage of storage on a LOB tablespace
- JCL:

```
//COMP1 EXEC PGM=DSN1COMP,PARM='LOB'
```

```
//SYSUT1 DD DISP=SHR,DSN=DSNDC1X.DSNDBD.DBLOB.LOBT001.I0001.A001
```

SAMPLE DSN1COMP TOOL OUTPUT

DSN1940I DSN1COMP COMPRESSION REPORT
DSN1COMP run on **compressed** LOB table space

LOB table space statistics

Number of LOBs 100 LOBs
Minimum LOB size 490 Bytes
Maximum LOB size 28 KB
Average LOB size 25 KB

Table space with compression (in KB)

Total uncompressed data 21 KB
Total compressed data 1,060 KB
Total LOB data 1,081 KB

Table space if uncompressed (in KB)

Total uncompressed data 2,546 KB
Total compressed data 0 KB
Total LOB data 2,546 KB

LOB compression ratio

Total LOB data compressed 1,081 KB
Total LOB data uncompressed 2,546 KB
Percentage of KB saved 58 %

Data pages needed for compressed LOB table space 180 Pages
Data pages needed for uncompressed LOB table space 540 Pages
Percentage of Data pages saved 67 %

Current page size 4 KB

NOTE

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice and at IBM's sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract.

The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.

NOTICES AND DISCLAIMERS

Copyright © 2016 by International Business Machines Corporation (IBM). No part of this document may be reproduced or transmitted in any form without written permission from IBM.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM.

Information in these presentations (including information relating to products that have not yet been announced by IBM) has been reviewed for accuracy as of the date of initial publication and could include unintentional technical or typographical errors. IBM shall have no responsibility to update this information. THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IN NO EVENT SHALL IBM BE LIABLE FOR ANY DAMAGE ARISING FROM THE USE OF THIS INFORMATION, INCLUDING BUT NOT LIMITED TO, LOSS OF DATA, BUSINESS INTERRUPTION, LOSS OF PROFIT OR LOSS OF OPPORTUNITY. IBM products and services are warranted according to the terms and conditions of the agreements under which they are provided.

IBM products are manufactured from new parts or new and used parts. In some cases, a product may not be new and may have been previously installed. Regardless, our warranty terms apply.”

Any statements regarding IBM's future direction, intent or product plans are subject to change or withdrawal without notice.

Performance data contained herein was generally obtained in a controlled, isolated environments. Customer examples are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual performance, cost, savings or other results in other operating environments may vary.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business.

Workshops, sessions and associated materials may have been prepared by independent session speakers, and do not necessarily reflect the views of IBM. All materials and discussions are provided for informational purposes only, and are neither intended to, nor shall constitute legal or other guidance or advice to any individual participant or their specific situation.

It is the customer's responsibility to insure its own compliance with legal requirements and to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.

NOTICES AND DISCLAIMERS: CONTINUED

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. IBM does not warrant the quality of any third-party products, or the ability of any such third-party products to interoperate with IBM's products. IBM EXPRESSLY DISCLAIMS ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents, copyrights, trademarks or other intellectual property right.

IBM, the IBM logo, ibm.com, Aspera®, Bluemix, Blueworks Live, CICS, Clearcase, Cognos®, DOORS®, Emptoris®, Enterprise Document Management System™, FASP®, FileNet®, Global Business Services®, Global Technology Services®, IBM ExperienceOne™, IBM SmartCloud®, IBM Social Business®, Information on Demand, ILOG, Maximo®, MQIntegrator®, MQSeries®, Netcool®, OMEGAMON, OpenPower, PureAnalytics™, PureApplication®, pureCluster™, PureCoverage®, PureData®, PureExperience®, PureFlex®, pureQuery®, pureScale®, PureSystems®, QRadar®, Rational®, Rhapsody®, Smarter Commerce®, SoDA, SPSS, Sterling Commerce®, StoredIQ, Tealeaf®, Tivoli®, Trusteer®, Unica®, urban{code}®, Watson, WebSphere®, Worklight®, X-Force® and System z® Z/OS, are trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at: www.ibm.com/legal/copytrade.shtml.

THANK YOU

- Questions?
 - Contact me at: chang.kim1@ibm.com
- A big thank you to everyone here and to Akiko Hoshikawa