System z Performance : More Than You Think

SHARE Pittsburgh Session 15709

EWCP

Steve Grabarits
IBM

August 6, 2014

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

AlphaBlox*
APPN*
CICS*
CICS/VSE*
Cool Blue
DB2*
DFSMS
DFSMShsm
DFSMSrmm
DirMaint
DRDA*
DS6000
DS8000
ECKD
ESCON*
FICON*
FlashCopy*
GDPS*
HiperSockets
HyperSwap
IBM*
IBM eServer
IBM logo*
IMS
Language Environment*
Lotus*
Large System Performance Reference™ (LSPR™)
Multiprise*
MVS
OMEGAMON*
Parallel Sysplex*
Performance Toolkit for VM
PowerPC*
PR/SM
Processor Resource/Systems Manager
RACF*
Redbooks*
Resource Link
RETAIN*
REXX
RMF
S/390*
Scalable Architecture for Financial Reporting
Sysplex Timer*
Systems Director Active Energy Manager
System/370
System p*
System Storage
System x*
System z
System z9*
System z10
Tivoli*
Tivoli Storage Manager
TotalStorage*
VSE/ESA
VTAM*
WebSphere*
zEnterprise
xSeries*
z9*
z10
z10 BC
z10 EC
z/Architecture*
z/OS*
z/VM*
z/VSE
zSeries*

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

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.
Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
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.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.
SAP, the SAP logo, and all other SAP product and service names mentioned herein are trademarks or registered trademarks of SAP AG in Germany and in several other countries all over the world.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

- **The performance observations are based on measurements and projections using standard benchmarks in a controlled environment.**
- **The actual throughput that a user will experience will vary depending upon 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 throughput improvements equivalent to the performance ratios stated here.**
- **References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates**

AGENDA

- **zEDC**
  - BSAM/QSAM
  - JAVA
  - System Logger
- **Large Memory Study**
  - Single System
  - Data Sharing
- **z/OS 2.1**

zEDC

- **IBM zEnterprise Data Compression (zEDC) Express is an IO adapter that does high performance industry standard compression**
- **Operating system requirements**
  - Requires z/OS 2.1 (w/PTFs) and the new zEDC Express for z/OS feature
  - z/OS V1.13 and V1.12 offer software decompression support only
- **Server requirements**
  - Available on zEC12 and zBC12
  - New zEDC Express feature for PCIe I/O drawer (FC#0420)
    - Each feature can be shared across up to 15 LPARs
    - Up to 8 features available on zEC12 or zBC12
  - Recommended high availability configuration per server is four features
    - This provides up to 4 GB/s of compression/decompression
    - Provides high availability during concurrent update (half devices unavailable during update)
    - Recommended minimum configuration per server is two features
  - Steps for installing zEDC Express in an existing zEC12/zBC12
    - Apply z/OS Service; Hot plug a zEDC Express adapter; IODF updates and Dynamic Activate

Our BSAM/QSAM Workload

- Used a mix of standard utility jobs to simulate writing and reading sequential data sets
  - IEBGENER – BSAM
  - IEBDG – QSAM
  - DFSORT – BSAM
    - Ability to write multiple outputs in same step with OUTFIL
    - Ability to read multiple inputs in same step with MERGE
  - IFASMFDP – QSAM
  - IDCAMS REPRO - QSAM
  - IEBCOMPR – QSAM
- Optimal buffering in JCL
- Test cases are I/O bound with nominal cpu usage
  - This was intentional to measure the cpu impact of compression
  - But it presents a best case scenario for elapsed time improvements

Real Data

- **LINKLIB, LPALIB – mostly binary**
- **PARMLIB, PROCLIB, MACLIB – text data**
- **SVCDUMP AND SADUMP**
- **SMF**
- **Site Directory, XML, JAVA**
- **Optim/z archive**

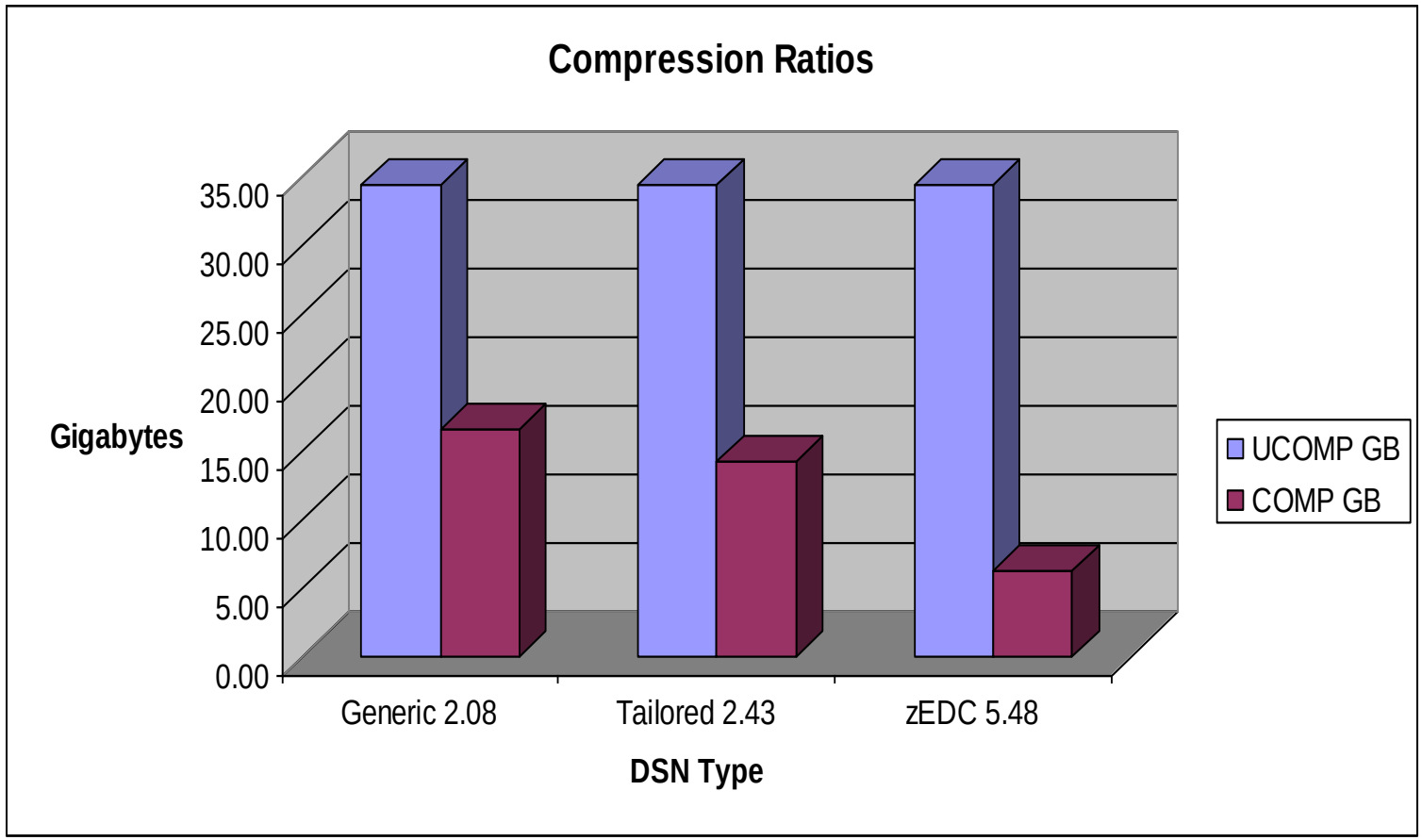
Test Scenarios

- **Single Thread**
  - Run only one job at a time to capture measurements without contention
- **Parallel**
  - Max of 9 jobs running in parallel to simulate contention for resources including the zEDC adaptors
  - Jobs had multiple steps repeating the same function to achieve longer duration of the execution
    - About 10-16 minutes for the R/W jobs
    - About 8-16 minutes for the Read Only jobs

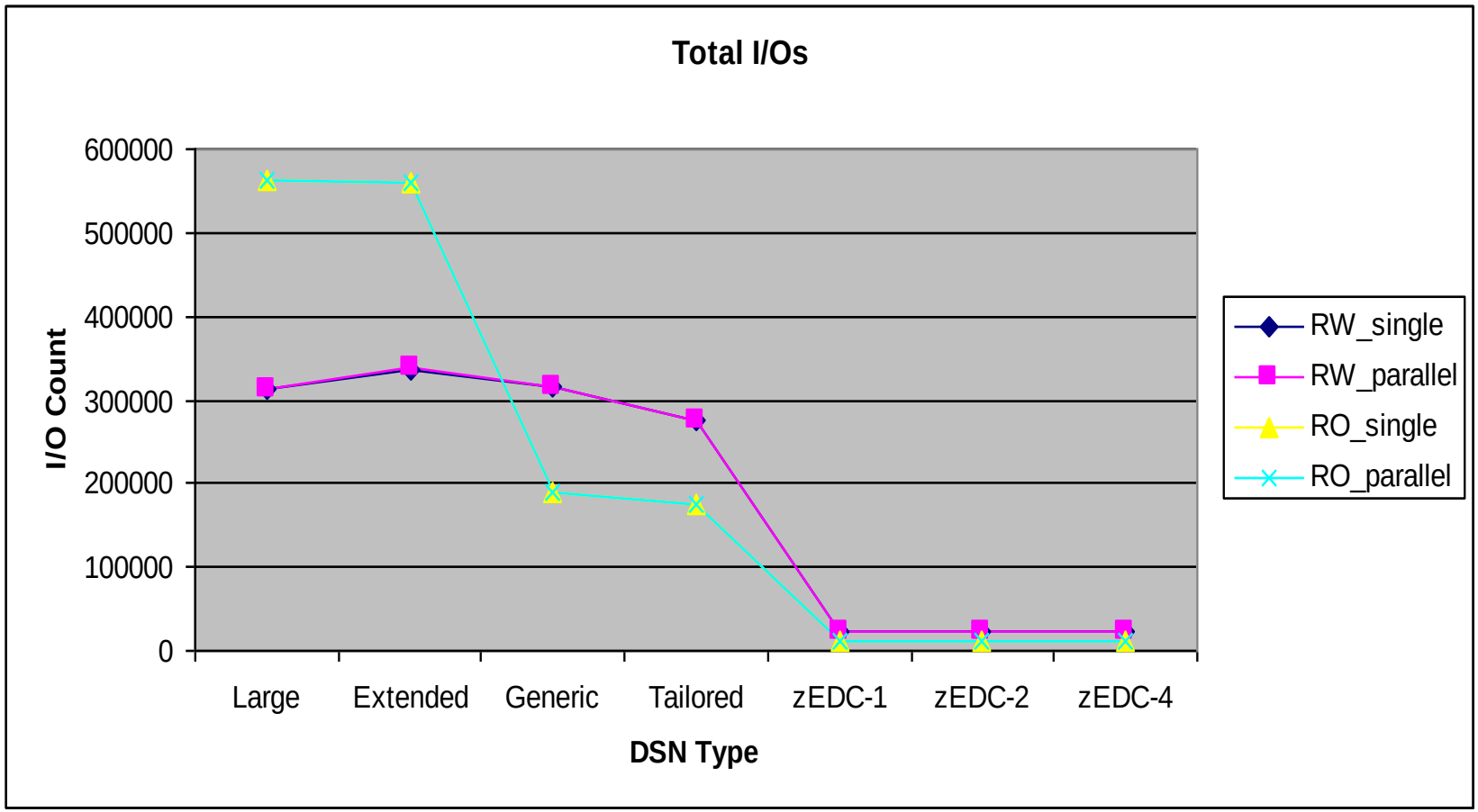
Comparison Points

- **Uncompressed**
  - Large Format
  - Extended Format
- **Existing DFSMS compression**
  - Generic
  - Tailored
- **zEDC**

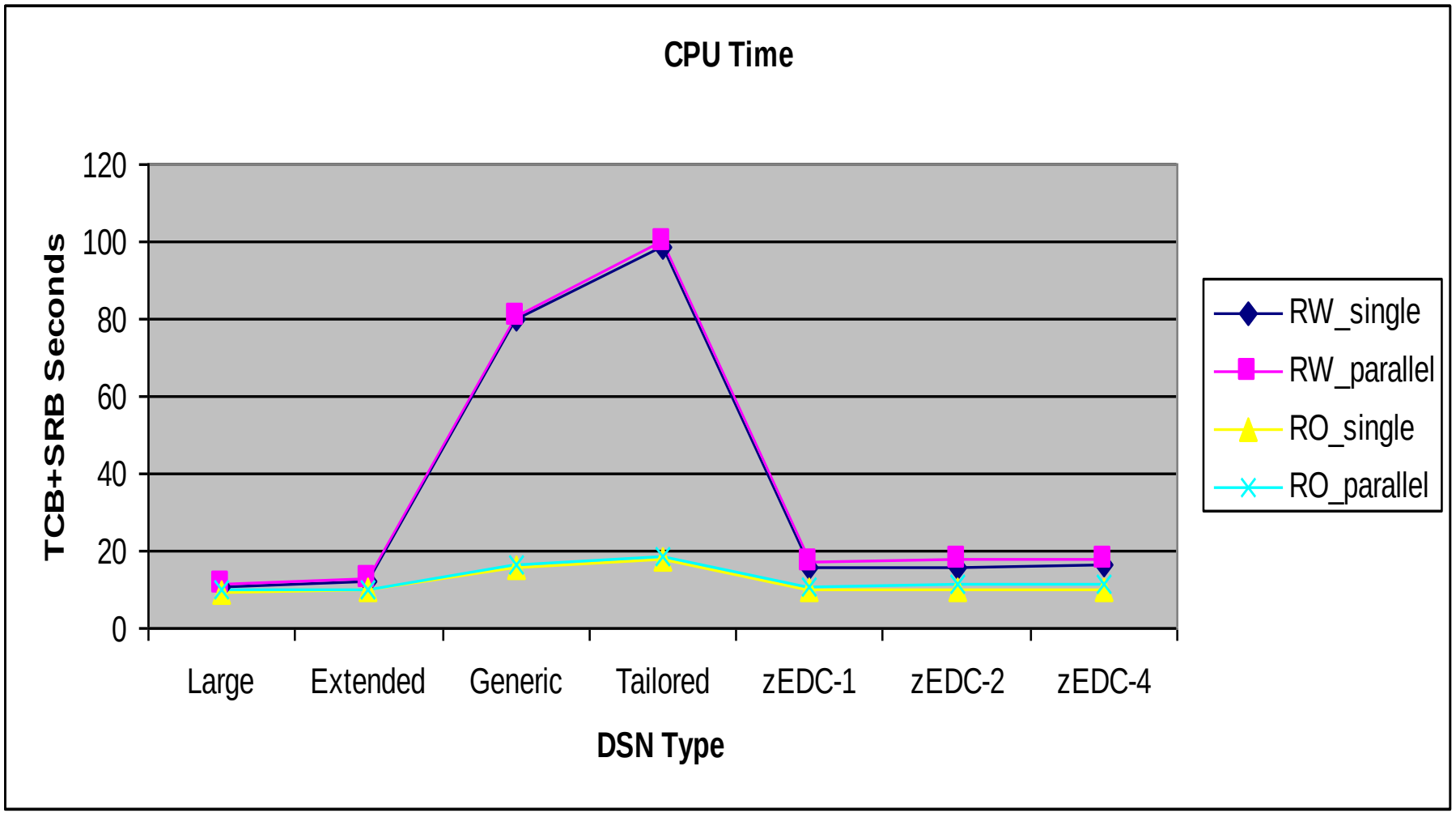
zEDC provides improved compression ratios



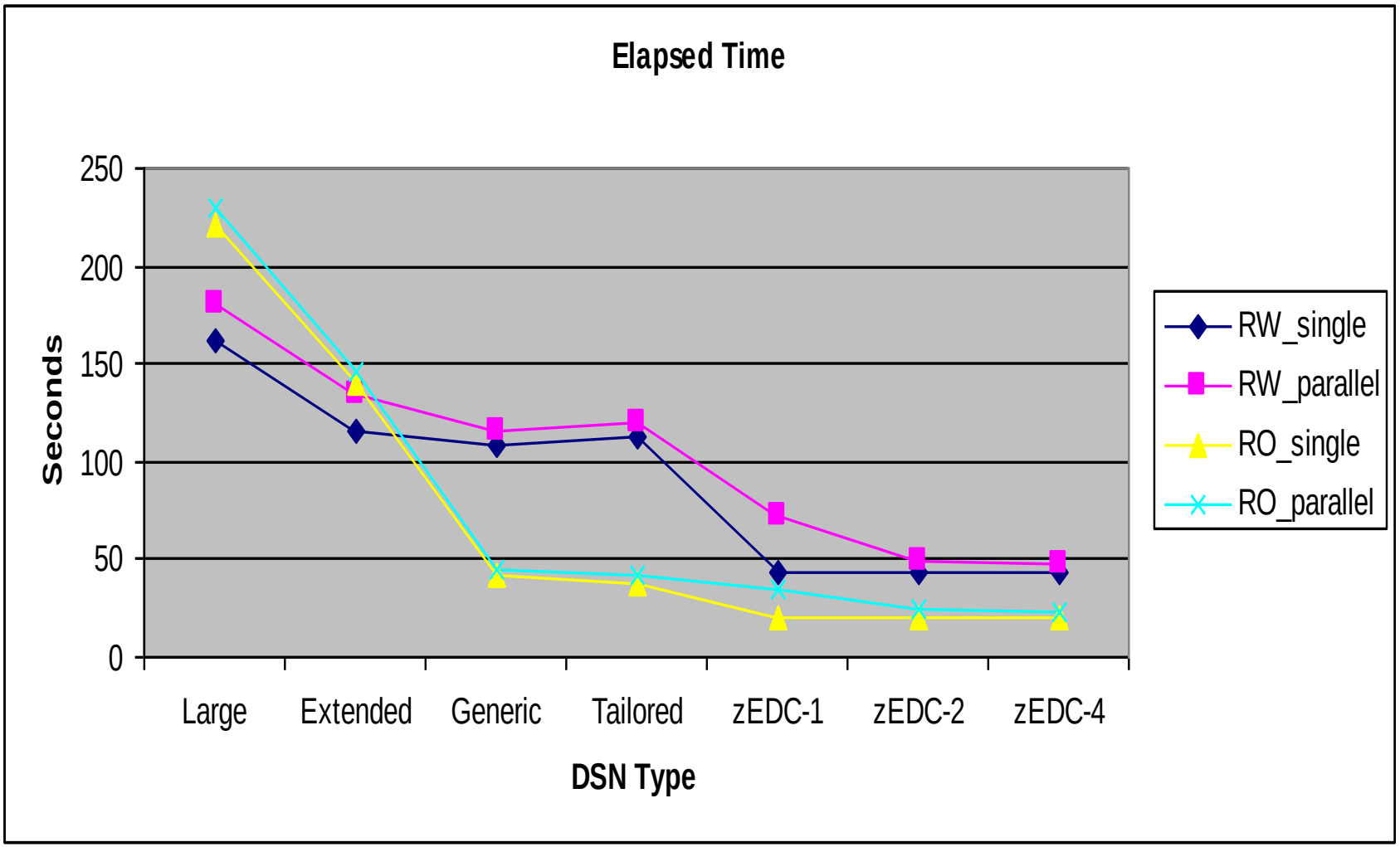
I/O Counts



CPU Time



Elapsed Time



BSAM/QSAM Data Set Compression with zEDC

- **BSAM/QSAM use of zEDC can provide a reduction in disk space of up to 4x over non-compressed BSAM/QSAM data, reducing I/O and elapsed time requirements which may provide a shorter batch window, with minimal CPU growth.**
- **BSAM/QSAM use of zEDC can provide a reduction in disk space up to double that of the existing BSAM/QSAM compression options, reducing I/O and elapsed time requirements which may provide a shorter batch window, while reducing the CPU cost for compression/decompression by up to 80%.**

Disclaimer : Based on projections and/or measurements completed in a controlled environment. Results may vary by customer based on individual workload, configuration and software levels.

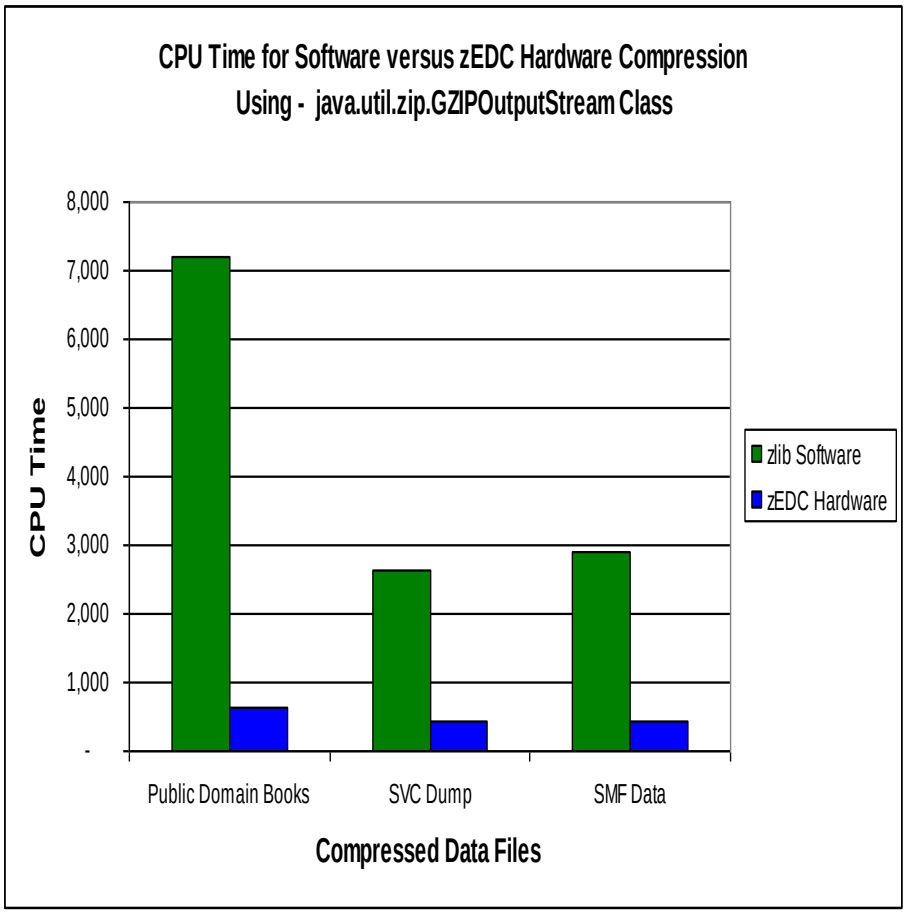
IBM System z Batch Network Analyzer and Compression

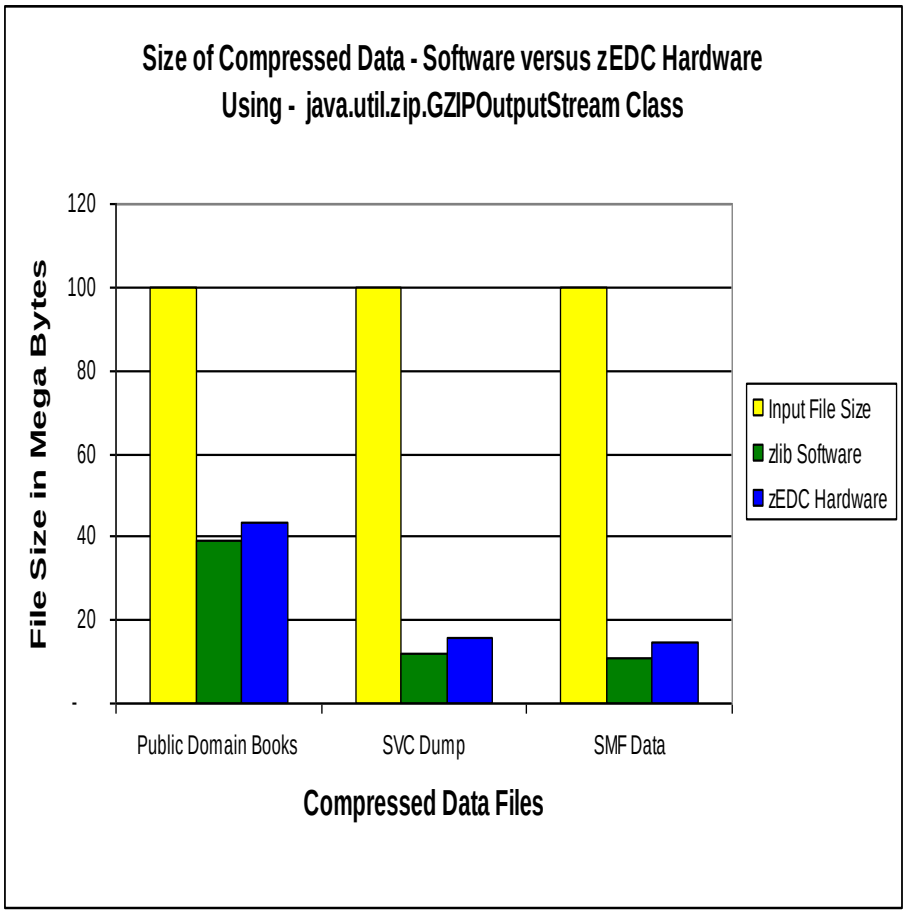
- **IBM System z Batch Network Analyzer 1.4.2**
  - A free, "as is" tool to analyze batch windows
  - Available to Customers, Business Partners and IBMers
  - PC based, and provides graphical and text reports
    - Including Gantt charts and support for Alternate Processors
- **Available on TechDocs**

  https://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS5132

- **zBNA identifies zEDC Compression Candidates**
  - Post-process customer provided SMF records, to identify jobs and data sets which are zEDC compression candidates across a specified time window, typically a batch window
  - Help estimate utilization of a zEDC feature and help estimate the number of features needed
  - Generate a list of data sets by job which already do hardware compression and may be candidates for zEDC
  - Generate a list of data sets by job which may be zEDC candidates but are not in extended format

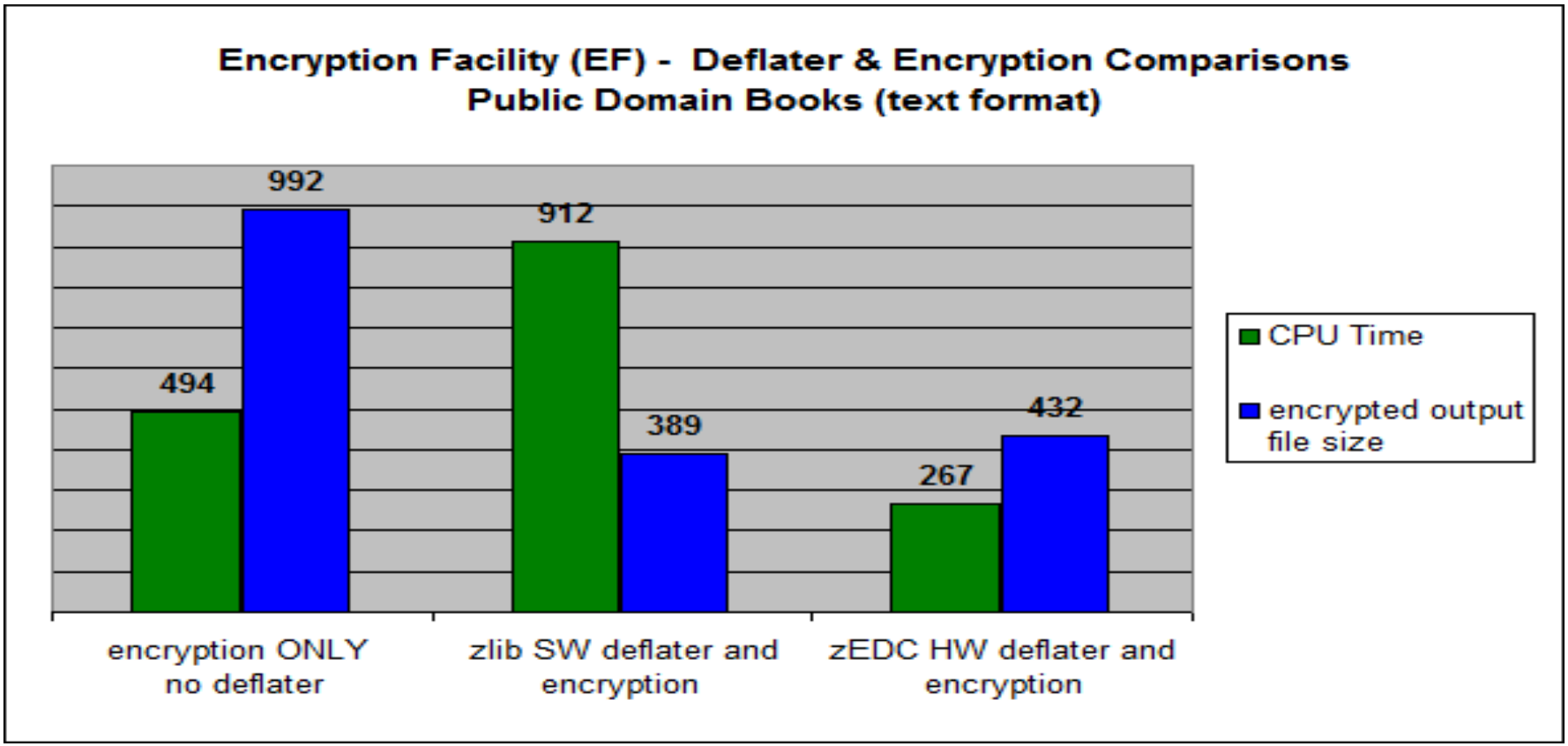
zEDC vs zlib Compression





- Java application to compress files using java.util.zip.GZIPOutputStream class
  - Up to 90% reduction in CPU time using zEDC hardware versus zlib software as shown
    - Up to 74% reduction in Elapsed time (not shown above)

Disclaimer : Based on projections and/or measurements completed in a controlled environment. Results may vary by customer based on individual workload, configuration and software levels.

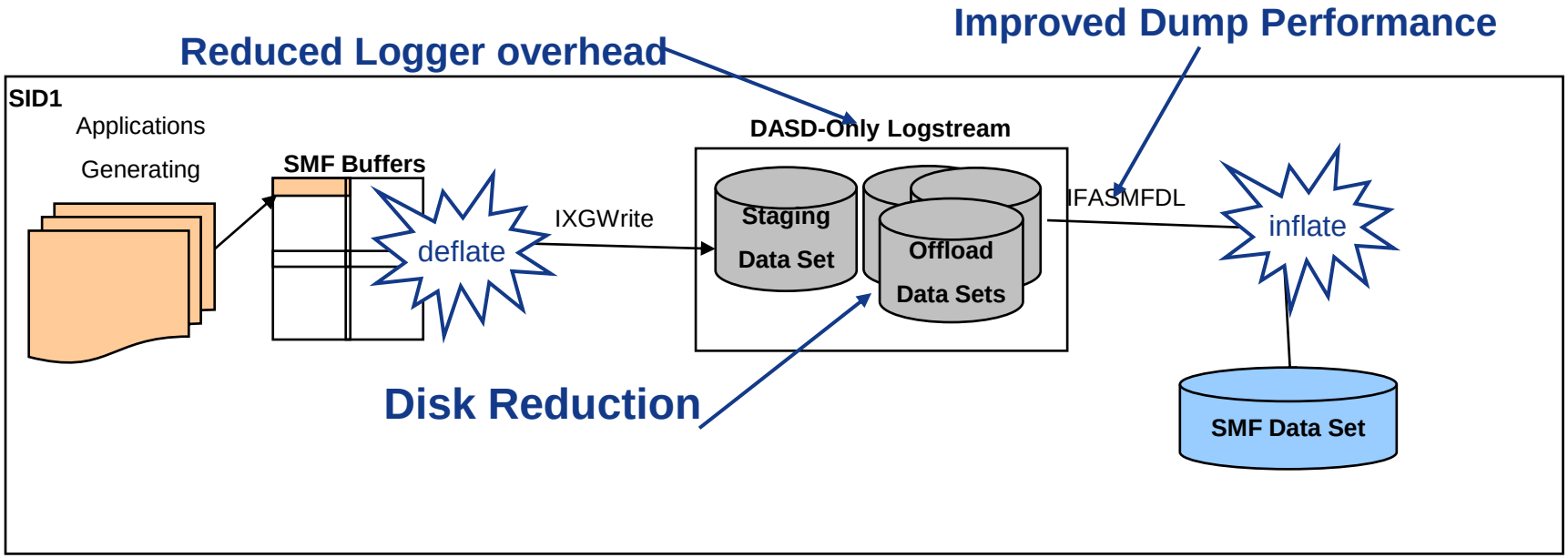


zEDC can provide IBM Encryption Facility users reductions of up to 60% in elapsed time and up to 70% in CPU time for environments where compression is already in use. For IBM Encryption Facility users not already using compression, compression with zEDC can provide IBM Encryption Facility users a reductions of up to 44% in elapsed time and up to 46% in CPU time

Disclaimer : Based on projections and/or measurements completed in a controlled environment. Results may vary by customer based on individual workload, configuration and software levels.

SMF using zEDC

This example shows a DASD-Only logstream used for SMF recording



Compression SMF logstreams reduce the amount of data in System Logger up to **4x** and reduce the elapsed time to extract IFASMFDL data up to **15%**

zEDC compression must be available on all systems that will access zEDC compressed SMF logstreams

Setup from SMFPRMxx either globally or per Logstream

Exploiting Large Memory for DB2 Buffer Pools using SAP Day Posting

- Brenda Beane, Seewah Chan, Paul Lekkas

Study Overview

- **IBM zEnterprise 196 and EC12 systems support…**
  - Up to 3 TB of real memory per server
  - Up to 1 TB per LPAR
- **DB2 10 and 11 for z/OS theoretically allow up to 1 TB of memory for all buffer pools per member**
- **Evaluated performance of using large amounts of memory for DB2 buffer pools in single system and data sharing**
- **Evaluated performance of reducing the number of buffer pools while keeping the total memory used for them constant**
- **Used SAP Banking Services (SBS) Day Posting workload**
  - Customer representative OLTP workload
  - Memory intensive, accesses a large number of tables, random I/O
- **Used DB2 11 for z/OS; z/OS 1.13 (SSI) and z/OS 2.1 (data sharing)**
- **Used 3-tier environment; SAP Database Server on IBM zEC12 (12 cps)**

Performance Metrics

- **Some of the performance metrics used in this presentation:**
  - **ETR:** External Throughput rate – Number of transactions ended per elapsed second (IMS trans, CICS trans, WAS trans, TSO trans, batch jobs…per second)
  - **ITR:** Internal Throughput Rate – Number of transactions per CPU busy second
  - **DB request time** – time for the application server to get a request back from the database server, a response time (network time + database processing time, but not any application server processing time)

Key Findings

- **More memory for DB2 local BPs and/or GBPs…**
  - Increases ITR… up to 25%; Increases ETR… up to 37%
  - Decreases response time… up to 70%;
  - Reduces sync read I/O… up to 97%
- **Reducing DB2 sync read I/O is key to performance improvements**
- **Performance benefit from adding memory to BPs is configuration and workload dependent… not every client will benefit from additional BP memory**
- **Significant performance improvements seen partly due to "pure DB2 workload" on z/OS; SAP Application Server is on Power7 blades**
- **Increasing local BPs may require additional GBP tuning**
- **Using fewer buffer pools simplifies management and tuning at little to no cost in performance in single system**

Techdoc available at

www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102461

Test Environment

System z Database Server

IBM zEC12 2827-HA1

- **12 CPs and up to 4 ICFs**
- **Up to 3 TB of real storage**
- **Up to 2 z/OS LPARs and 2 CFs**
- **DB2 11 for z/OS**
- **z/OS 1.13 - z/OS 2.1**
- **Up to 675 GB LFAREA**
- **1M large frames**

IBM System Storage

- Dual Frame DS8870 w/ 8 GB NVS, FICON Express8S for database
- DS8700 for DB2 logs

FICON Express8S

10 GbE Network

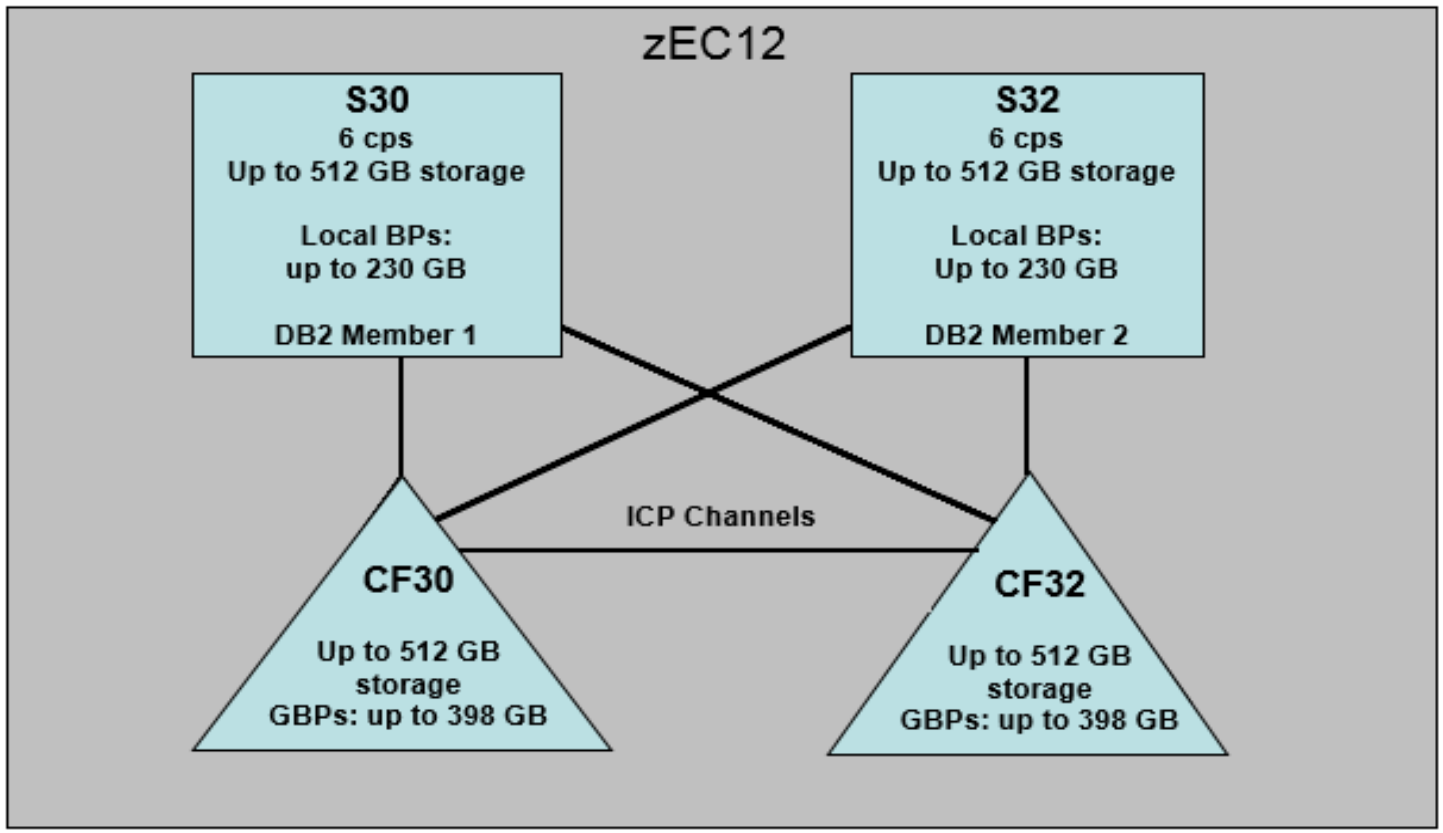
Application Servers

- Up to 13 IBM PS701 8406-71Y Blade Servers
- Each with 8 3.0 GHz processors and 128 GB memory
- AIX 7.1.0
- DB2 Connect 10.1 FP2

Presentation Server

- IBM 9133-55A with 4 2.1 GHz processors and 32 GB memory

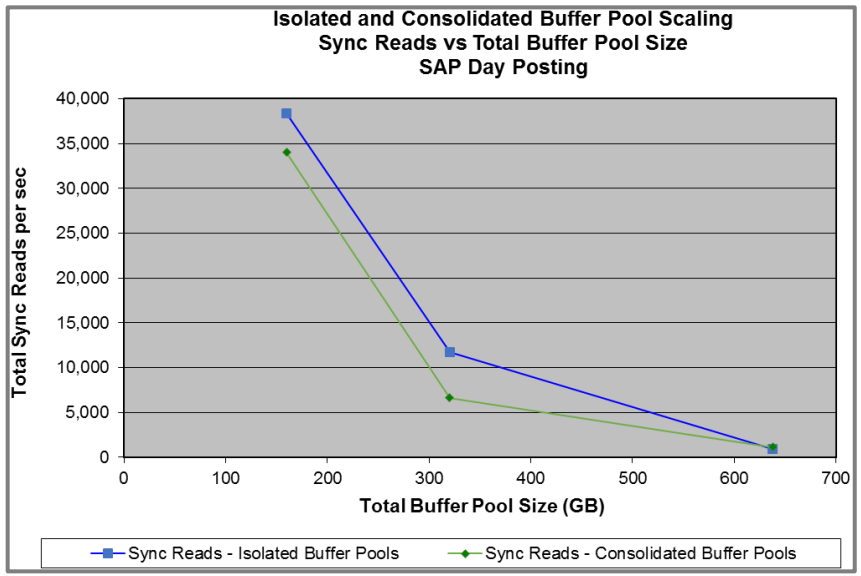
2-Way Data Sharing Configuration

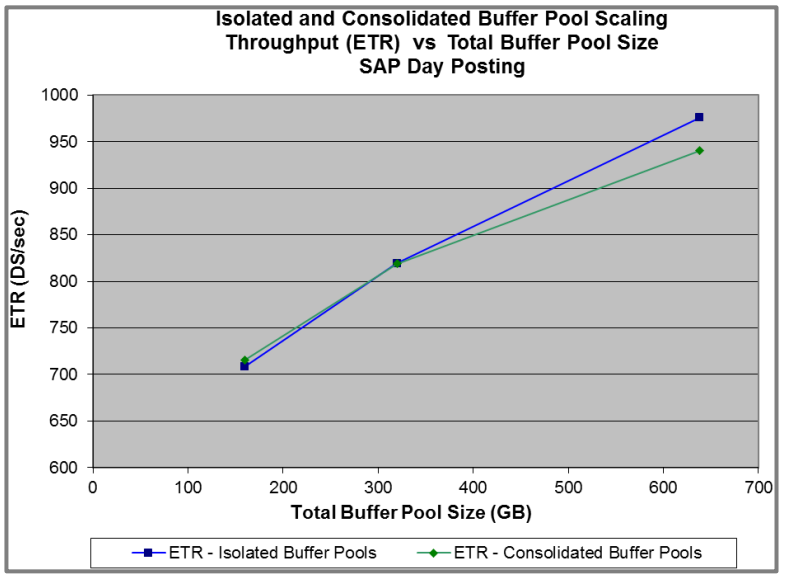


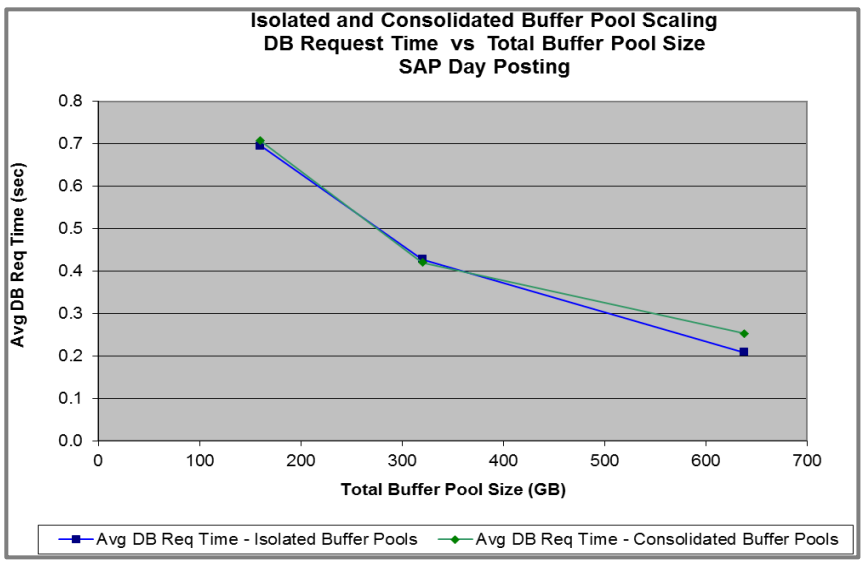
Single System Results – SAP Day Posting

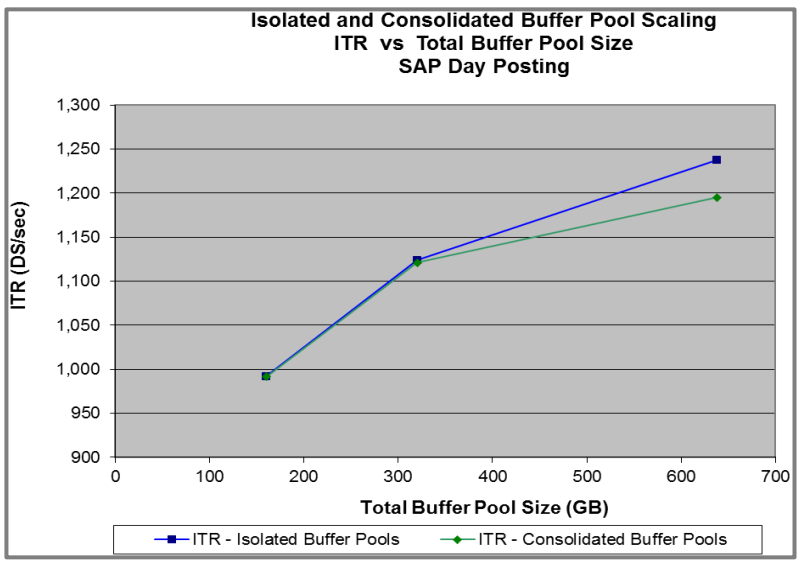
Run id	BP Size	ITR DS/sec	ITR Delta	%CPU	ETR DS/sec	ETR Delta	DB Req Time	DB Req Time Delta	Sync Reads	Sync Reads Delta
Isolated BPs (33)										
S30801B1	**161 GB**	991.84	n/a	71.46%	708.77	n/a	0.695	n/a	38.4K	n/a
S30802B1	**320 GB**	1123.63	**13%**	72.92%	819.35	16%	0.428	**-38%**	11.7K	**-70%**
S30906B2	**638 GB**	1237.43	**25%**	78.85%	975.71	38%	0.209	**-70%**	0.9K	**-97%**
Consolidated BPs (9)										
S31114B1	**161 GB**	991.31	n/a	72.19%	715.63	n/a	0.707	n/a	43.0K	n/a
S31115B1	**320 GB**	1121.18	**13%**	73.04%	818.91	14%	0.420	**-41%**	6.6K	**-85%**
S31118B1	**638 GB**	1195.20	**21%**	78.67%	940.27	31%	0.253	**-64%**	1.1K	**-97%**

Single System – Consolidated vs Isolated BPs









Data Sharing Measurements

- **2-way data sharing; IBM zEC12; two z/OS LPARs with 6 cps each**
- **Used a staged approach to understand the effects of adding memory to the local buffer pools, to the group buffer pools, and to both**
  - Local Buffer Pool Scaling
  - Group Buffer Pool Scaling
  - Scale both Local and Group Buffer Pools
- **Looked at how additional memory should be spread between LBPs and GBPs**
- **Checked if large GBPs (~ 200 GB) affected performance of GBP duplexing**
- **Checked if 4-way data sharing saw similar effects when using large amounts of memory.**

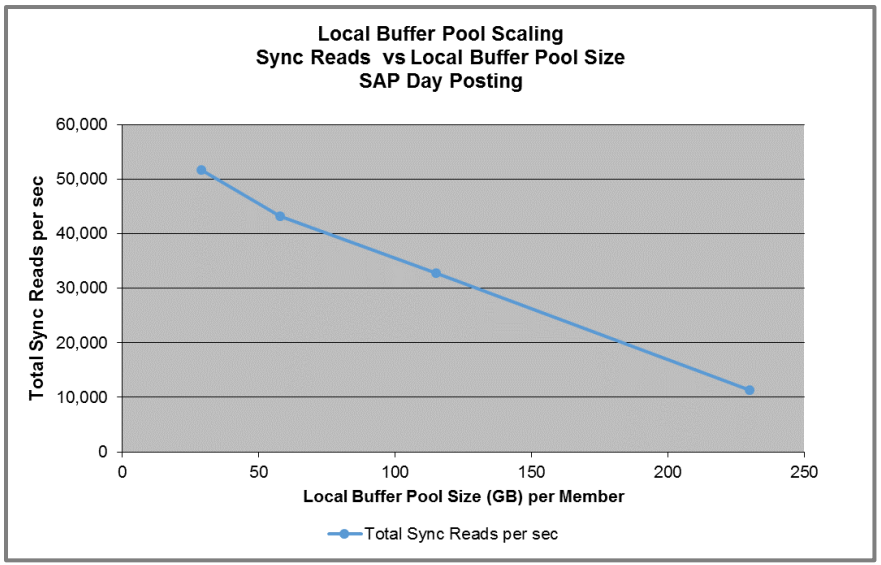
Data Sharing Results – Local Buffer Pool Scaling

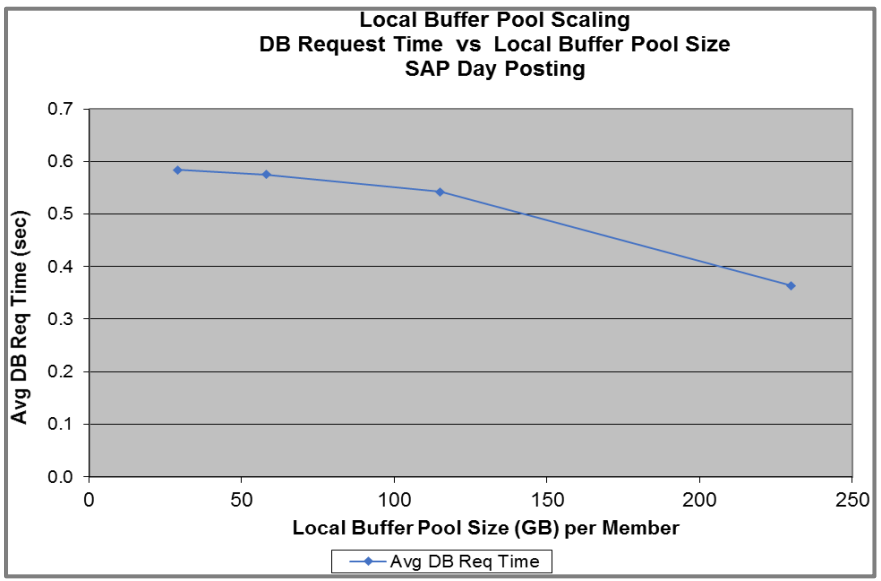
- Scaled LBPs; held size of GBPs constant
- GBPs tuned so no directory entry reclaims

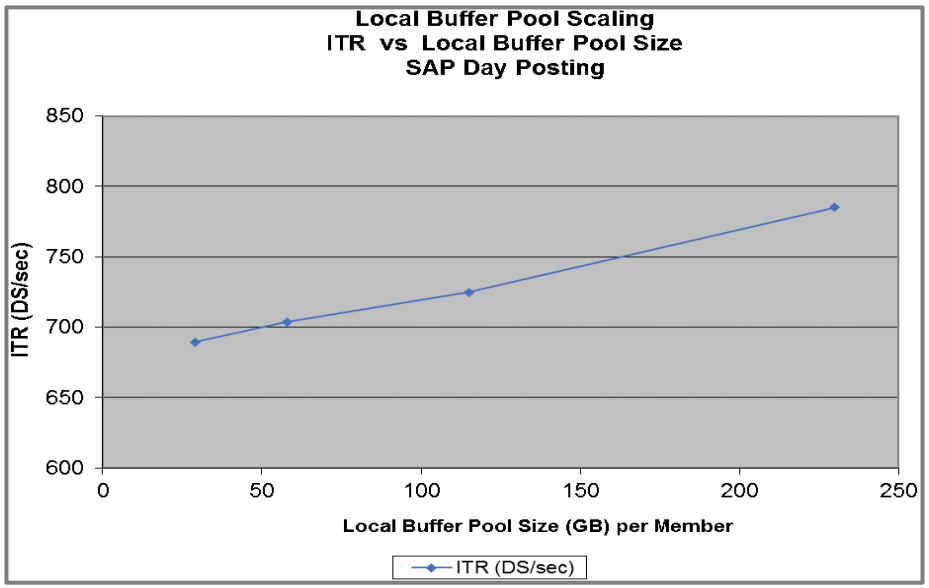
Run id	LBP Size GBP Size	ITR DS/sec	ITR Delta	%CPU	ETR DS/sec	ETR Delta	DB Req Time	DB Req Time Delta	Sync Reads	Sync Reads Delta
S40303B2	29 GB / 64 GB	689.45	n/a	84.37%	581.69	n/a	0.584	n/a	51.7K	n/a
S40303B1	58 GB / 64 GB	703.56	**2%**	84.47%	594.30	2%	0.575	**-2%**	43.2K	**-16%**
S40227B1	115 GB / 64 GB	724.81	**5%**	81.20%	588.51	1%	0.542	**-7%**	32.8K	**-37%**
S40209B1	230 GB / 64 GB	785.21	**14%**	81.80%	642.26	10%	0.364	**-38%**	11.4K	**-78%**

- Performance benefits of LBP scaling influenced by…
  - Amount of non-GBP dependent data in workload
    - SAP Day Posting has about 42% non-GBP dependent data
  - Use of DB2 member cluster feature in workload
  - Sizing GBPs to avoid directory entry reclaims

Data Sharing – Local Buffer Pool Scaling







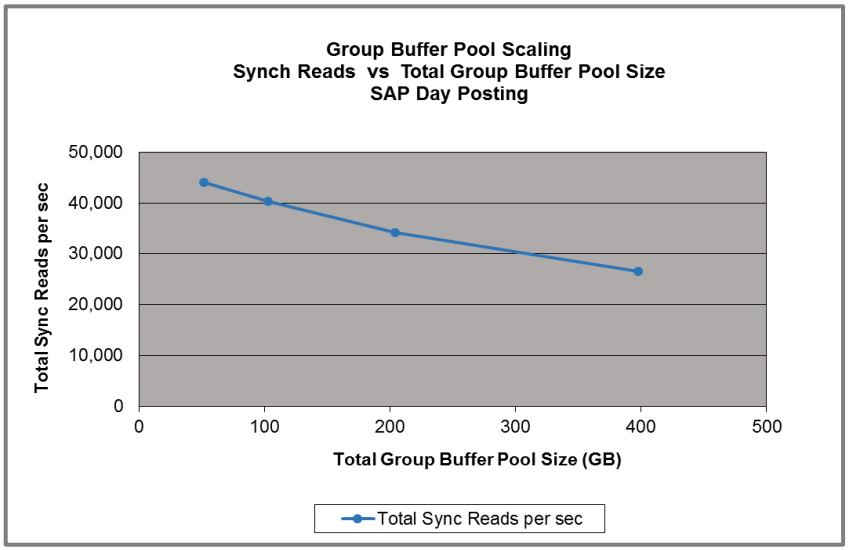
Data Sharing Results – Group Buffer Pool Scaling

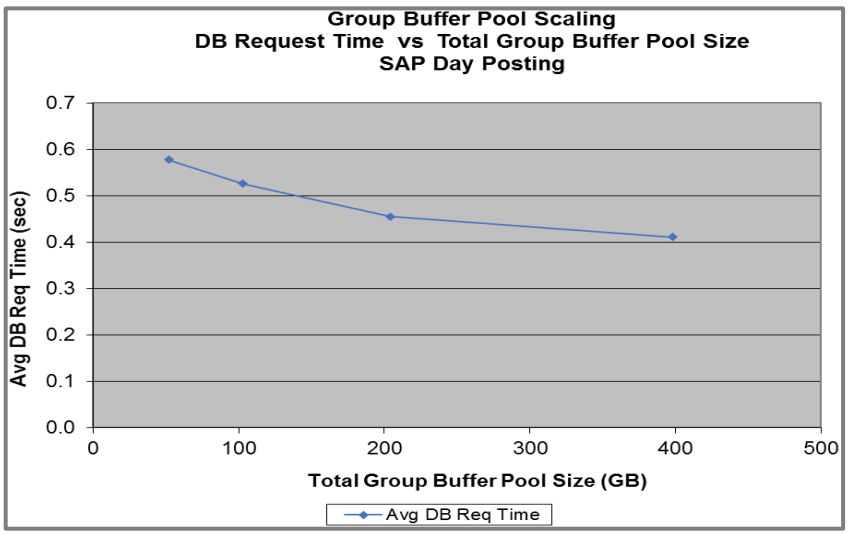
- Scaled GBPs; held size of LBPs constant
- GBPs tuned so no directory entry reclaims

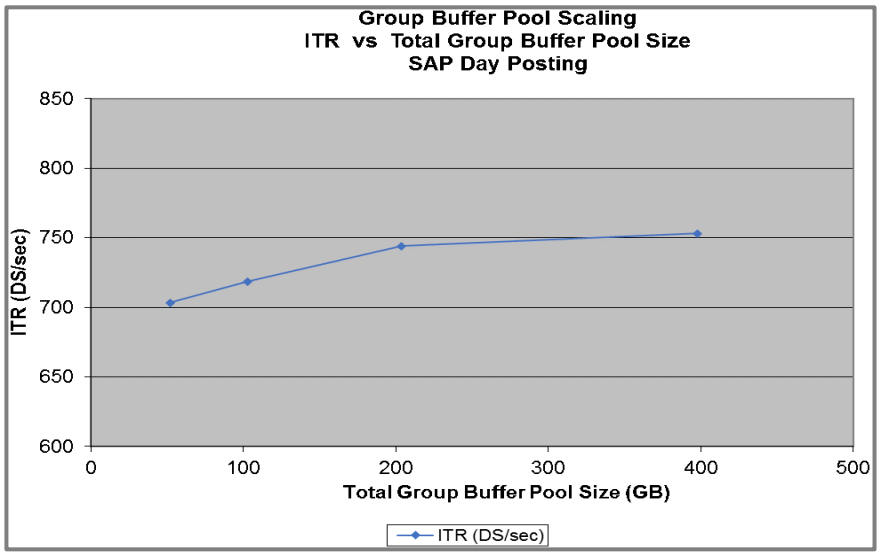
Run id	LBP Size GBP Size	ITR DS/sec	ITR Delta	%CPU	ETR DS/sec	ETR Delta	DB Req Time	DB Req Time Delta	Sync Reads	Sync Reads Delta
S40201B1	58 GB **52 GB**	703.28	n/a	83.33%	586.01	n/a	0.577	n/a	44.0K	n/a
S40218B1	58 GB **103 GB**	718.51	**2%**	80.94%	581.53	<1 %	0.526	**-9%**	40.3K	**8%**
S40211B1	58 GB **204 GB**	743.98	**6%**	79.34%	590.24	<1 %	0.455	**-21%**	34.2K	**-22%**
S40225B1	58 GB **398 GB**	753.05	**7%**	79.86%	601.46	3%	0.411	**-29%**	26.5K	**-40%**

- GBPs act as a second layer of cache
- Performance benefits of GBP scaling influenced by…
  - Amount of GBP dependent data in workload (GBPCACHE CHANGED)
    - SAP Day Posting has about 58% GBP dependent data
  - Use of DB2 member cluster feature in workload
    - Member cluster objects don't make effective use of GBP

Data Sharing – Group Buffer Pool Tuning





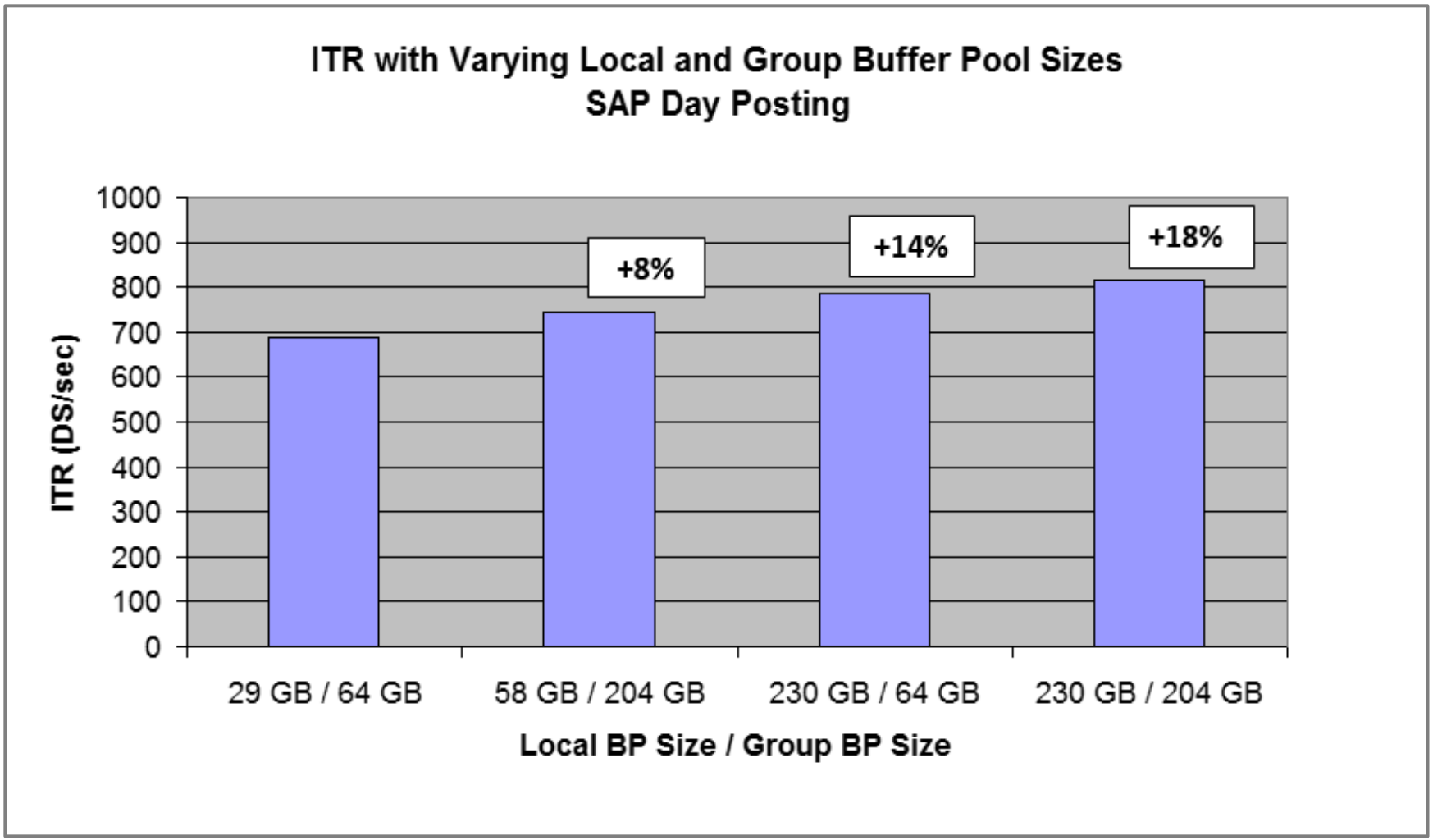


Data Sharing – Varying Local and Group BPs

Run id	LBP Size GBP Size	ITR DS/sec	ITR Delta	%CPU	ETR DS/sec	ETR Delta	DB Req Time	DB Req Time Delta	Sync Reads	Sync Reads Delta
S40303B2	29 GB 64 GB	689.45	n/a	84.37%	581.69	n/a	0.584	n/a	51.7K	n/a
S40211B1	58 GB 204 GB	743.98	**8%**	79.34%	590.24	1%	0.455	**-22%**	34.2K	**-34%**
S40209B1	230 GB 64 GB	785.21	**14%**	81.80%	642.26	10%	0.364	**-38%**	11.4K	**-78%**
S40409B1	**230 GB** **204 GB**	816.11	**18%**	78.30%	638.97	10%	0.382	**-35%**	7.9K	**-85%**

- In general, for most workloads, adding memory to both the local and group buffer pools will be the best approach.
- Always need to size GBPs to avoid directory reclaims when adding memory to local buffer pools.

Data Sharing – Varying Local and Group BPs



Key Findings (2)

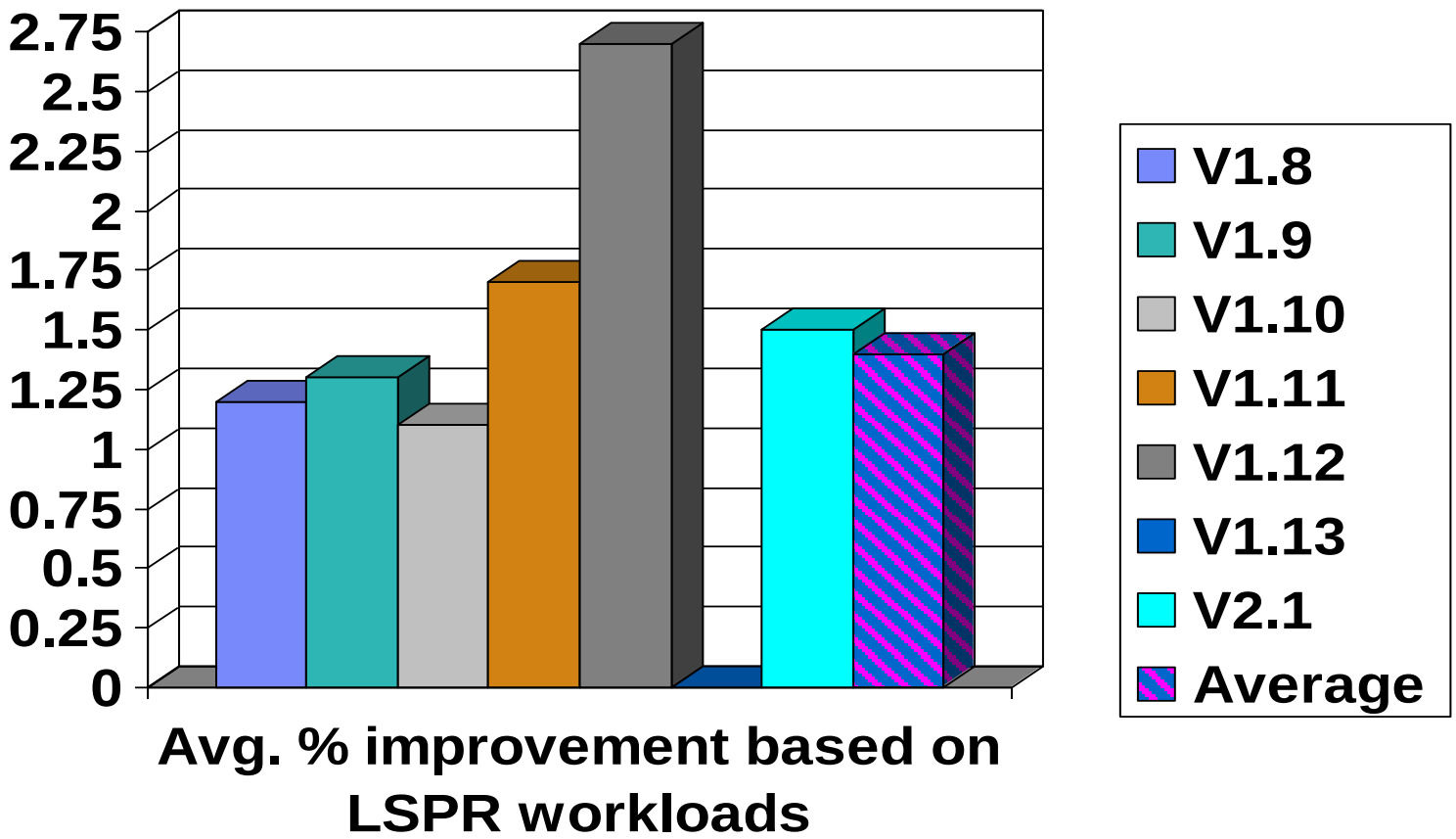
- SAP Day Posting workload got more benefit from larger LBPs than from larger GBPs… workload dependent!
- Some considerations for adding memory to LBPs/GBPs…
  - Current buffer pool tuning
    - Minimal sync I/O, good BP hit ratios; more memory won't help
  - DB2 configuration
    - number of DB2 members, GBP redundancy, etc
  - Workload characteristics
    - Data access patterns
      - random vs seq, use of member clustering, etc
    - Amount of GBP dependent data

Key Findings (3)

- Saw similar performance benefits with 4-way and 2-way data sharing as memory was added to LBPs and GBPs
- Saw similar performance benefits with and without GBP duplexing (with a GBP size up to 204 GB) as memory was added to LBPs and GBPs
- Still important to maintain a good DASD I/O subsystem. By adding memory to DB2 buffer pools, DB2 synchronous reads can be reduced, but there still will be DB2 asynchronous I/O, including prefetch and deferred writes, as well as the critical DB2 synchronous logging I/O.
- Highly recommended to “page fix” local buffer pools with I/O activity, provided there is real storage to back them.
- Highly recommended to use 1M fixed large frame support for the buffer pools that are page fixed

z/OS Release to Release Performance

CPU Usage (ITR)



References

1. **IBM Corp. 2014. DB2 11 for z/OS Buffer Pool Monitoring and Tuning Redpaper REDP-5092-00 http://www.redbooks.ibm.com/redpieces/abstracts/redp5092.html?Open**
2. **IBM Corp. 2013. DB2 for z/OS Best Practices – DB2 for z/OS Data Sharing Tuning Group Buffer Pools; John Campbell; https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/W7c1e3103820b_4d9e_add6_b5ade3280dcb/page/DB2%20for%20zOS%20Data%20Sharing%20Tuning%20Group%20Buffer%20Pools**
3. **IBM Corp. 2012. *IBM System zEnterprise, System Storage, and DB2 10 for z/OS: SAP Banking Services 7.0 150 Million Accounts Measurements***

   http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101978
4. **IBM Corp. 2013. *Large Systems Performance Reference, Document Number SC28-1187-17 https://www-304.ibm.com/servers/resourcelink/lib03060.nsf/pages/lsprindex/$file/SC28118717.pdf***
5. **IBM Corp. 2013. *DB2 11 for z/OS Technical Overview***

   http://www.redbooks.ibm.com/Redbooks.nsf/RedpieceAbstracts/sg248180.html?Open
6. ***SAP DBA Guide: DB2 11 for z/OS), Database Administration Guide: SAP on IBM DB2 for z/OS***

   https://websmp108.sap-ag.de/~sapidb/011000358700001053572013E (requires SAP user id)
7. **IBM Corp. 2009. *IBM System z and System Storage DS8800: Accelerating the SAP Deposits Management Workload With Solid State Drives***

   http://www-01.ibm.com/support/docview.wss?uid=tss1wp101442