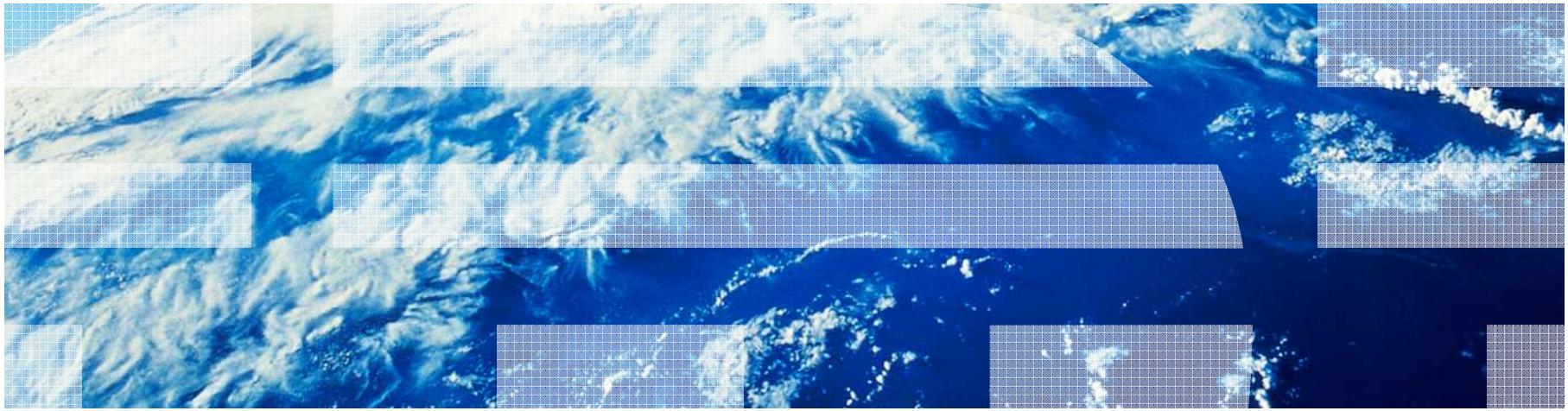
Marianne Hammer
IBM Corporation
Poughkeepsie, New York
hammerm@us.ibm.com

Session 10887: z/OS Performance Update
Share 118
Atlanta, GA
03/12/2012



Trademarks

IBM Corporation 2012

- IBM, the IBM logo and ibm.com are registered trademarks, and other company, product or service names may be trademarks or service marks of International Business Machines Corporation in the United States, other countries, or both. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml
- Adobe, the Adobe logo, PostScript, the PostScript logo, Cell Broadband Engine, Intel, the Intel logo, Intel Inside, the Intel Inside logo, Intel Centrino, the Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, IT Infrastructure Library, ITIL, Java and all Java-based trademarks, Linux, Microsoft, Windows, Windows NT, the Windows logo, and UNIX are trademarks or service marks of others as described under "Special attributions" at: http://www.ibm.com/legal/copytrade.shtml#section-special
- Other company, product and service names may be trademarks or service marks of others.
- 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.

Notes

- The performance observations are based on measurements and projections using standard IBM 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.

What's new in z/OS Performance - Agenda

- Introduction

- z/OS V1.12 Performance:
  - Release performance content – BCP focus
  - Results from LSPR workloads
  - Dispatcher timeslices option
  - Open Large Number of Data sets
  - SVC dump performance

- z/OS V1.13 Performance:
  - Release performance content and results
  - zFS Sysplex enhancements
  - RSM CTRACE reduced overhead
  - SDSF SORT CPU usage
  - IEBCOPY Performance

z/OS Performance Metrics and Tools

- Some of the performance metrics used in this presentation:
  - **ETR:** External Throughput rate – Number of transactions ended per second (IMS trans, CICS trans, WAS trans, TSO trans, batch jobs)
  - **ITR:** Internal Throughput Rate – Number of transactions per CPU busy second
  - **PL:** Pathlength – Instructions per transaction
- Performance data from RMF
- Internal version of HIS tool (customer Instrumentation)
- Other internal performance tools
- Release-to-release performance numbers in:
  - ATS Softcap tool
  - ATS zSoftcap tool – available spring 2011. New user interface.

Methodology for testing and comparing two z/OS releases

- Compare new release to previous release
  - Previous release (base for comparison): GA-level of z/OS
  - When using performance workloads: measure systems at high CPU utilization – 90%
- Both releases tested on:
  - Same processor HW configuration
  - Same I/O configuration
  - Same workload setup and tuning parameters
  - Same number of simulated users
  - Same database layout and size
  - Same SW stack level (DB2, IMS, CICS, WAS, Java)
- Everything is the same except for the SYSRES that is used for IPL, and any PARMLIB and catalog changes required for new release.

Workloads and Coverage for z/OS Release Testing

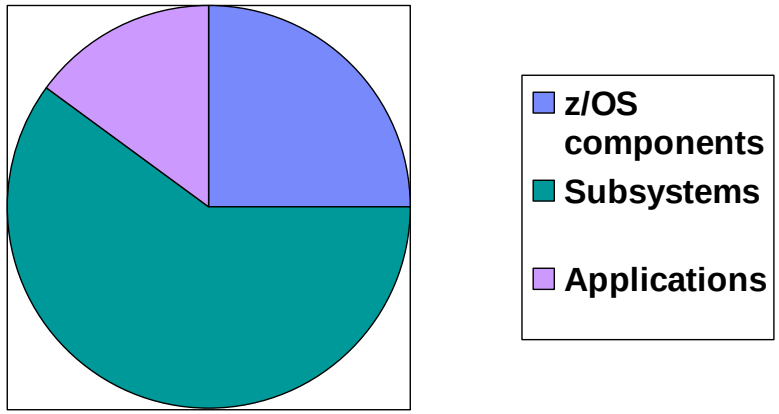
- WASDB (Daytrader): OLTP workload with WAS transactions and DB2 data base
- OLTP-T (IMS): OLTP workload with IMS transactions
- HIDI (WEB/DB2): OLTP workload with WAS front-end to CICS transactions using DB2 data base
- CB-L (CBW2): Commercial batch with long running jobs
- OLTP-W/PS HIDI workload in a two-system sysplex with DB2 datasharing:
- WASDB in a 3.tier setup with DB2 and WAS on separate LPARs using zIIPs and zAAPs.
- WASDB in a multi-LPAR configuration.
- MIDI: 75% OLTP-T and 25% CB-L (new workload for V1.13)
- LODI: 75% WASDB and 25% CB-L (new workload for V1.13)
- CB-S: Batch workload with short running jobs. CPU stress.
- Laddis: Using NFS V3.
- IOZONE: Using NFS V4
- USS Primitives: evaluates CPU usage for USS callable services.
- Unicode Primitives: evaluates CPU usage for the unicode services
- ZOSPERF: Primitives for some of the most common BCP functions (for example, getmain, wait, pause)

z/OS Release Performance Goals

- Performance improvement goals for z/OS:
  - 5% performance improvements for the z/OS component measured as the average result of the four LSPR workloads on a 32w single image.
    z/OS component represents on average 25% of total CPU usage for the LSPR workloads.
    This translates into an average of 1.25% overall system improvement compared to previous release.

  - VSCR goal: Reduce 31-bit common storage by 4 MB for the release (compared to previous release). The 4 MB goal is based on a 10% reduction of the perceived 40 MB z/OS contribution to 31-bit common (based on the LSPR workloads), and is measured against customer use of common virtual storage.

CPU usage LSPR workloads



Module CPU usage based on instrumentation data (CPU samples)

z/OS V1.12 Performance Content - BCP

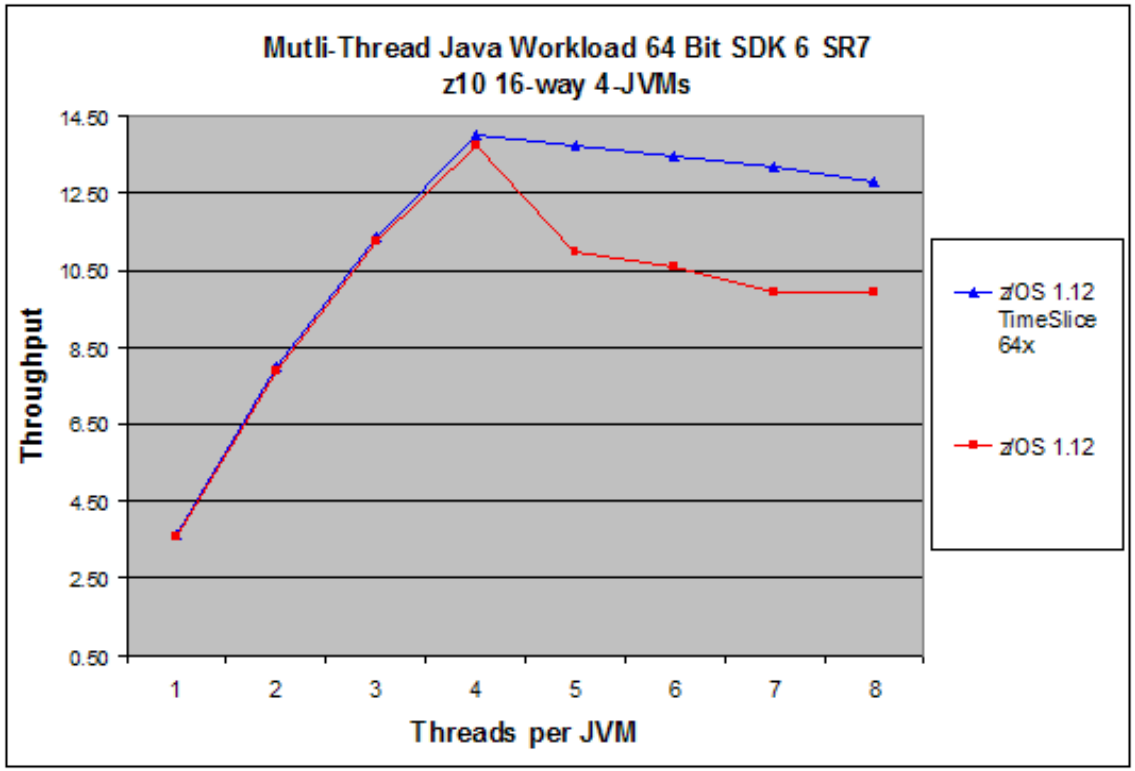
- RRS zero interest commit optimization for WAS transactions
- LE code optimization
  - Prefetch (STCMH) for heap pools
  - realloc() optimization: string manipulation intensive applications that use the realloc() function to grow the storage for the strings. Most benefits for applications that are heavy users of this function. Measured 20% improvement on z10 for a string manipulation test using realloc() (compared to V1.11).
- Overall system performance impact:
  - Changes to supervisor pause/release serialization – benefits to heavy users of pause/release (e.g. DB2)
  - Nucleus large page support
  - Misc hot cache items, and component re-compiles with higher optimization levels for top 200 z/OS csects (based on usage in the four LSPR workloads)
- Timeslices option for discretionary CPU-intenstive work
- Performance improvements when opening large number of data sets in a DB2 environment
- SVC dump performance improvements

z/OS V1.12 performance Results – LSPR Workloads

Workload	1w on z10	12w on z10 with HD=YES	32w on z10 with HD=YES	64w on z10 with HD=YES	Comments
WASDB (WAS and DB2 OLTP)	1.9%	1.5%	1.2%	2.5%	Mostly RRS and misc BCP
OLTP-W (WAS, CICS, DB2 OLTP)	-0.2%	0.1%	4.0%	3.4%	Pause-rel optimization and misc BCP
OLTP-T (IMS OLTP)	0.2%	1.7%	4.6%	-	Mostly VTAM OA33084 and misc BCP
CB-L (batch)	0.2%	-	1.2%	-	Misc BCP
Average	**0.5%**	**1.1%**	**2.7%**	**3.0%**	

z/OS V1.12 Dispatcher Timeslices Option

- The data shown are the result of a benchmark with more threads than CPs and very low pause times
  - only affects discretionary work that is CPU-intensive as determined by significant mean time to wait (MTTW)
- Tests done with default value and IEAOPTxx TIMESLICES=64 option
- 64 Bit multi-JVM had a 33 percent improvement at 7 Threads in each JVM
- TIMESLICES=255 at best had only a 2 percent improvement over TIMESLICES=64 in the Multi-JVM case



z/OS V1.12 Open DB2 Data Sets Performance

- DB2 PTF UK58205 – support up to 100000 open data sets in the DB2 DBM1 address space

- Allocation enhancement – MEMDSENQMGMT (Memory-based data set ENQ management)
  - ENQs managed in private storage instead of SWA blocks. Job will be non-restartable.
  - Enable via parmlib member or MVS command:
    - ALLOCxx: SYSTEM MEMDSENQMGMT(ENABLE)
    - SETALLOC SYSTEM,MEMDSENQMGMT=ENABLE
  - Exploit with DB2 APAR PM17542 (closed September 2010)

- GRS enhancement – special case for DB2 allocations:
  - Allow shared SYSZTIOT ENQs ahead of exclusive waiters to promote parallelism
    - Controlled by bit in JSCB (JSCBTIOD) which is set by program
    - Up to 50 shared request may jump ahead of an exclusive qname=SYSZTIOT
  - V1.12 APAR OA33633 (closed August 2010)
  - DB2 exploitation APAR PM18557 (closed August 2010)

z/OS V1.12 Open DB2 Data Sets – non-SMS Results

- Single member sysplex with GRS STAR on a z10 with 4 CPs (Pok tests)
- 20 concurrent batch jobs to open data sets
- DSNMAX=100000 in DSNZPARM

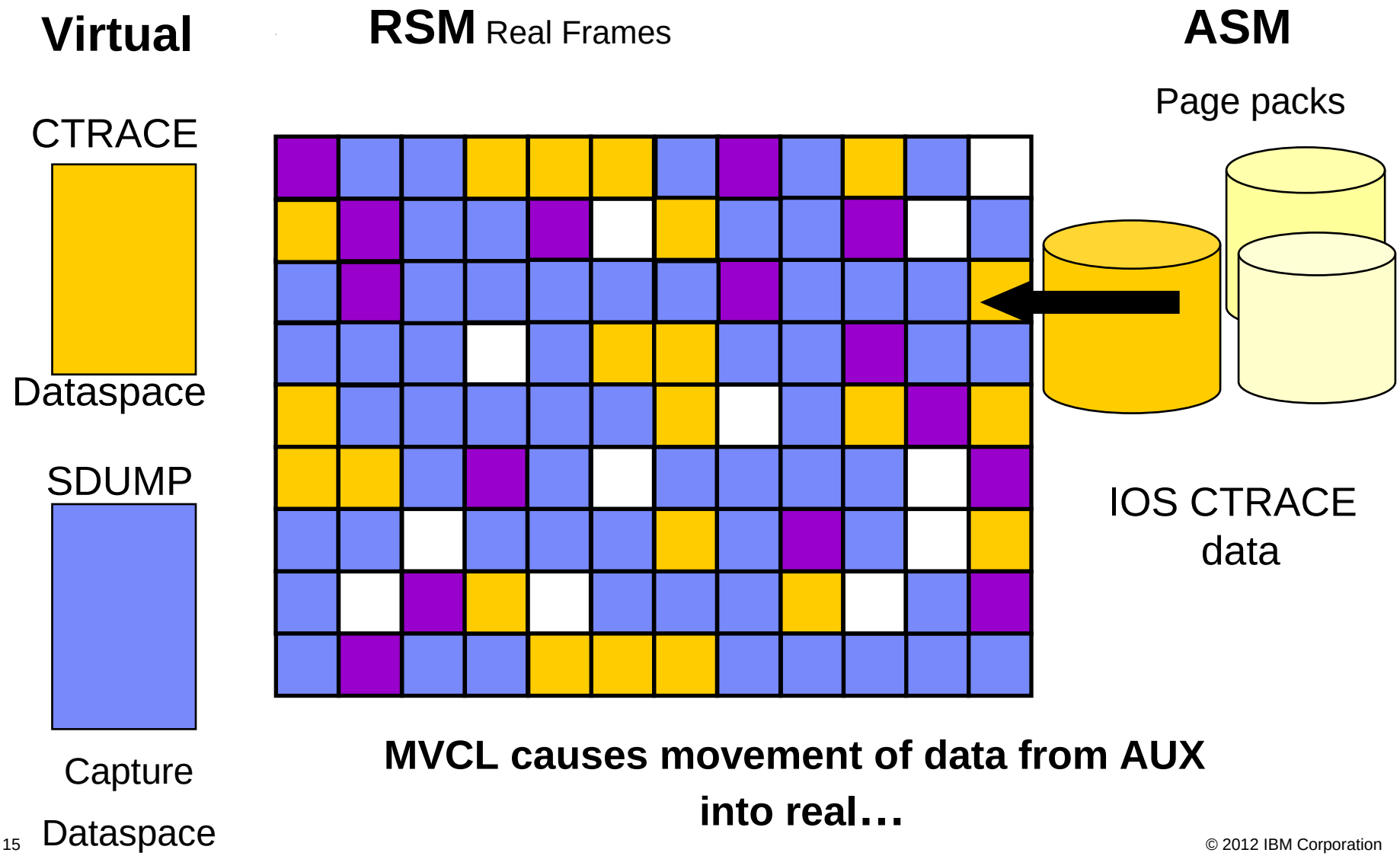
	V1.11	**V1.12 w/GRS exploitation**	**V1.12 w/MEMDSENQM GMT**	**Delta V1.11 to V.12 w/GRS and allocation**
Elapsed time	480.7	321.4	239.5	-50%
ETR (number of open DS per sec)	199.7	298.7	400.8	+100%
LPAR CPU%	59.5	66.5	75.7	+27%
ITR	335.6	449.1	529.3	+58%

z/OS V1.12 DB2 Open Data Sets – SMS-managed Results

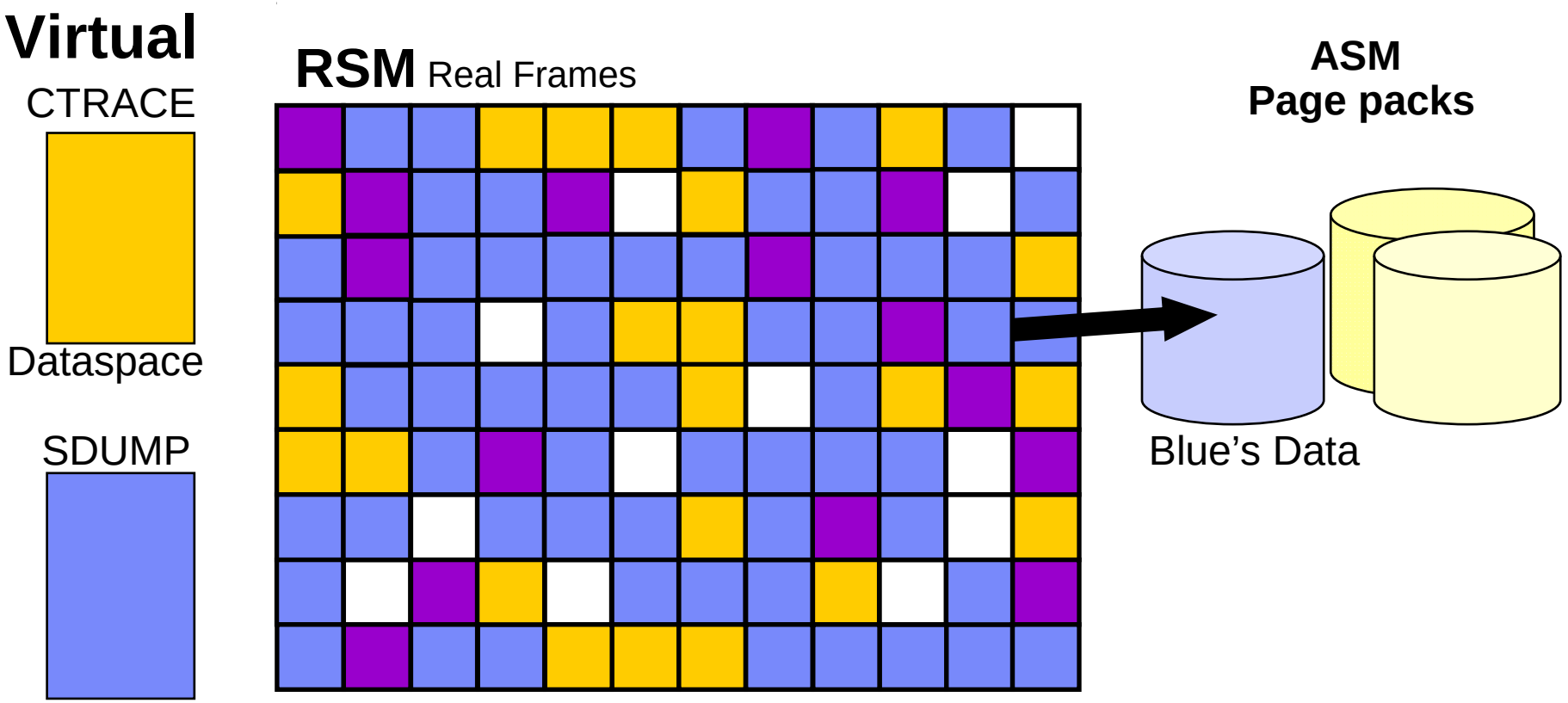
- Tests using SMS-managed data sets done in DB2 lab (San Jose)
- z10 with 8 CPs, sysplex, DB2 9 for z/OS, DS8300 disk
- Table shows difference between V1.11 and V1.12 w/MEMDSENQMGT and GRS exploitation:

	Delta V1.11 to V.12 with GRS and MEMDSENQMGT
Elapsed time in sec	-76%
Avg. open time per data set	-69%
DBM1 TCB time	-85%
ETR	+4.3x

SVC Dump Exit Data Capture Problem



SVC Dump Exit Data Capture Problem

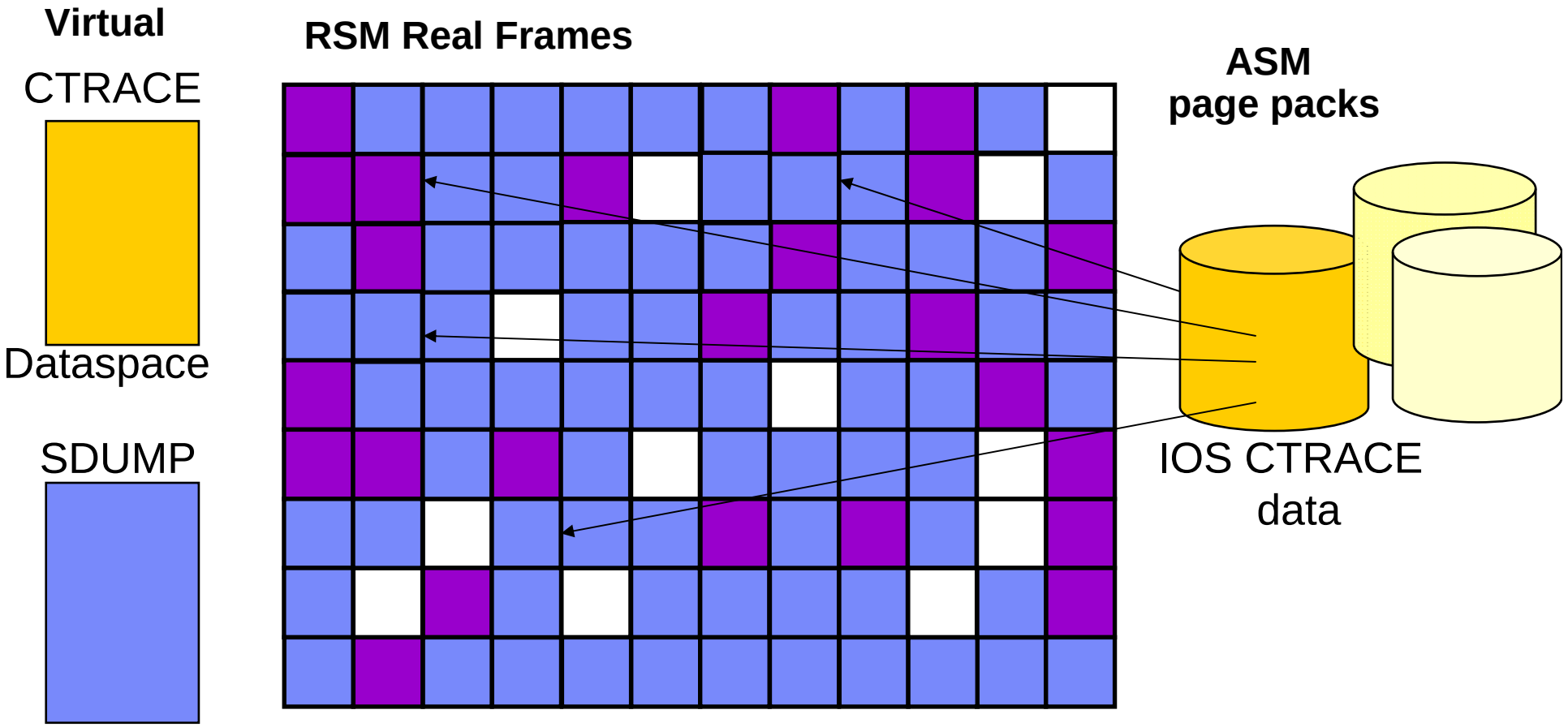


…which in turn forces page-out of other more important data

z/OS V1.12 SVC Dump Solution

- Reduce memory pressure when capturing exit data
- Do a smart copy such that if the source data is out on AUX we do an I/O directly into the SDUMP buffer space to capture the data while leaving the source data out on AUX
- In the previous example this would mean that the IOS component trace data will not be brought into real and an I/O will be done to copy the data from AUX directly into the SDUMP capture dataspace
- Avoid changing the reference pattern of the source data due to capture
- Copy the data via a special RSM service such that if the source data did not appear referenced before the capture it still remains unreferenced after the capture
- Data that has been captured will not cause other perhaps more important data to be paged out
- Data in the SDUMP capture space will be made to look old so that this data will be paged out before any important workload data is paged out

z/OS V1.12 SVC Dump Exit Data Capture Solution



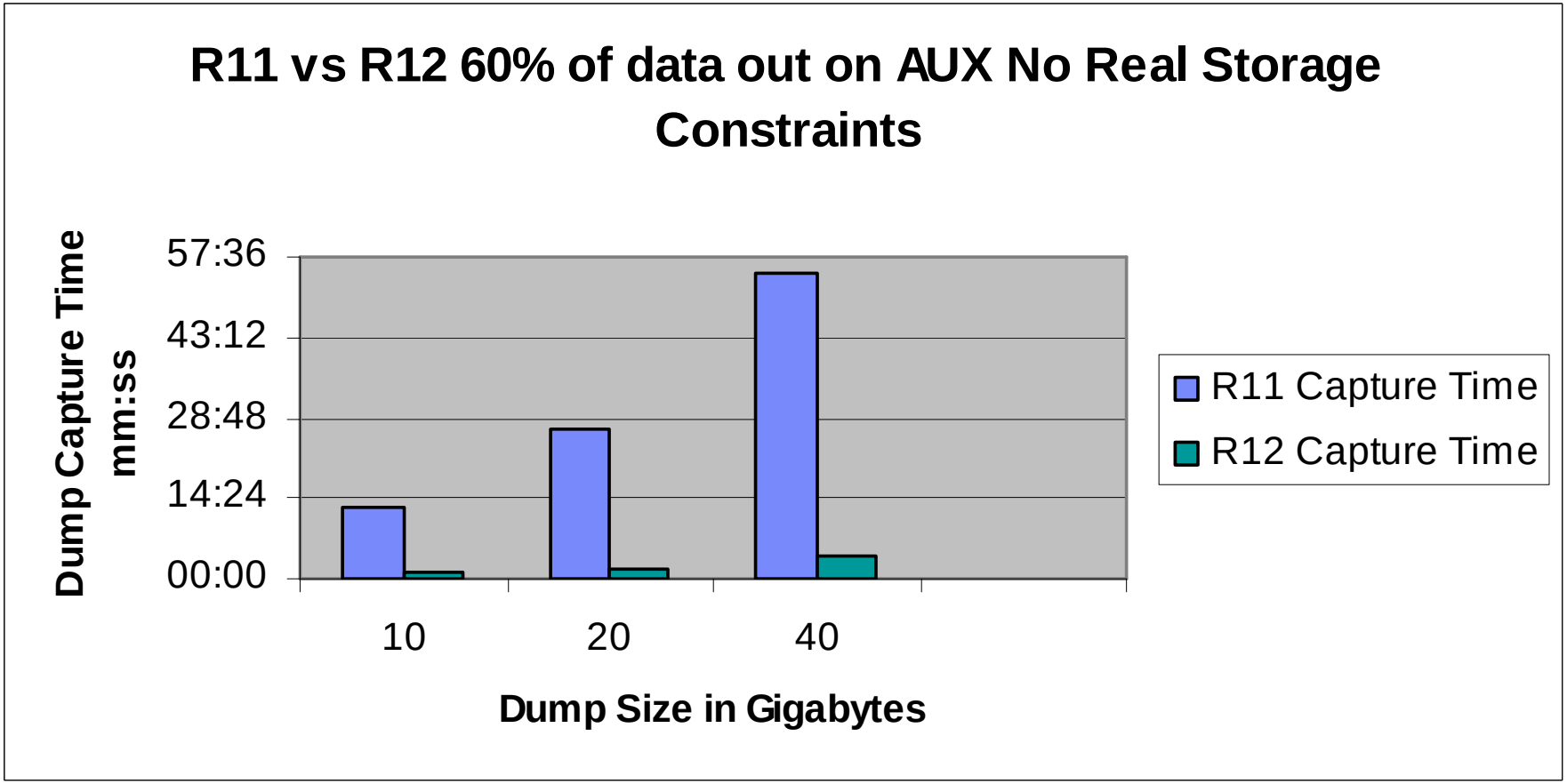
Smart Copy direct to SDUMP dataspace

z/OS V1.12 SVC Dump Test Configuration and Scenarios

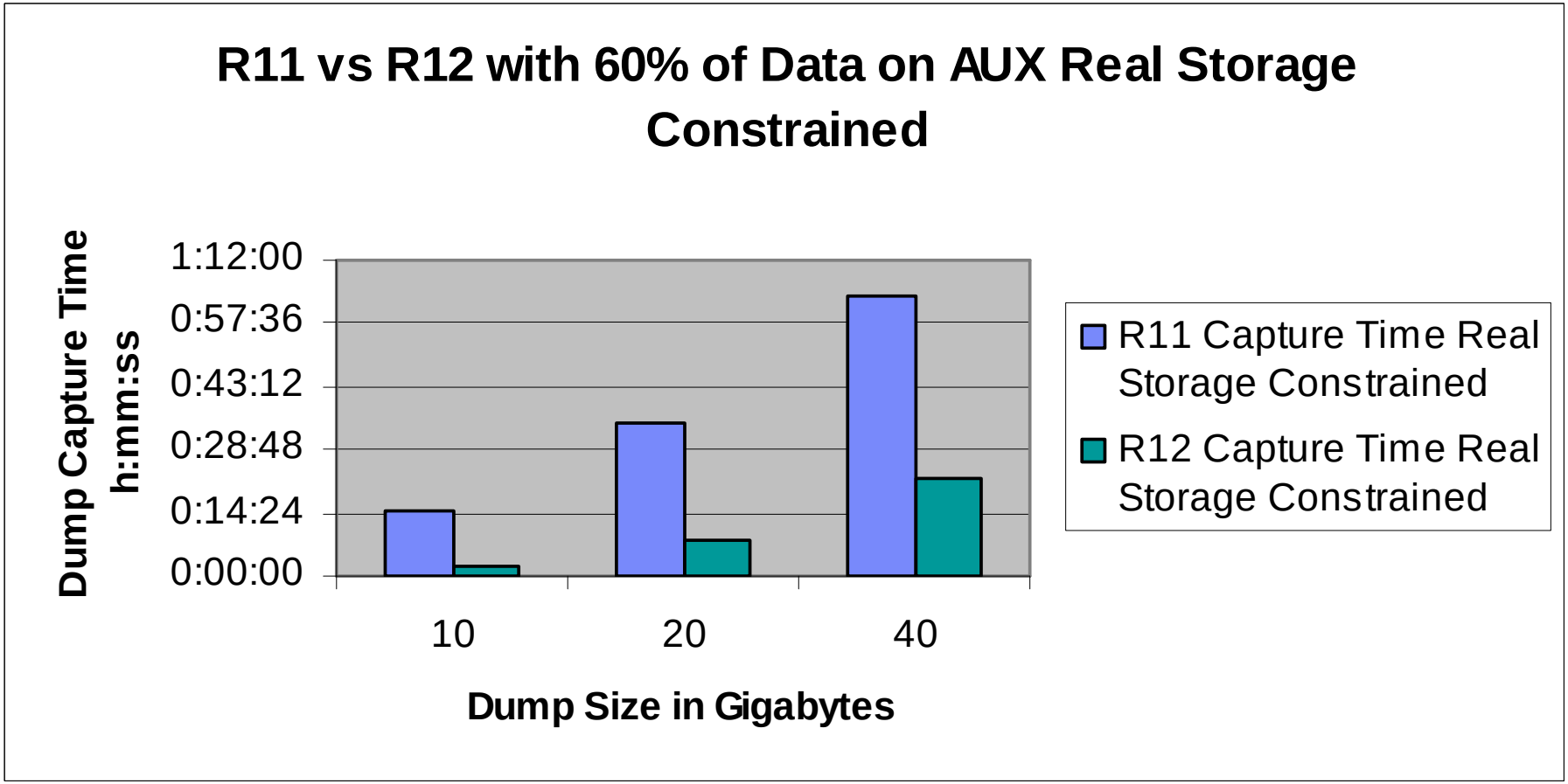
- Three variations of amount of storage on page data sets: 0%, 5%, and 60%
- Tested with real storage constrained and non-constrained environments.
- Tests on a z10 with 4 CPs

Dump Size	Real storage	Page data sets	Dump data sets
10 GB	25 GB	15 x 4 GB page data sets on 2105-800	2 x 54 GB dump data set on 2107-922 w/116.7 GB cache
20GB	45 GB	15 x 4 GB page data sets on 2105-800	2 x 54 GB dump data set on 2107-922 w/116.7 GB cache
40GB	85 GB	15 x 6.74 GB page data sets on 2105-800	2 x 54 GB dump data set on 2107-922 w/116.7 GB cache

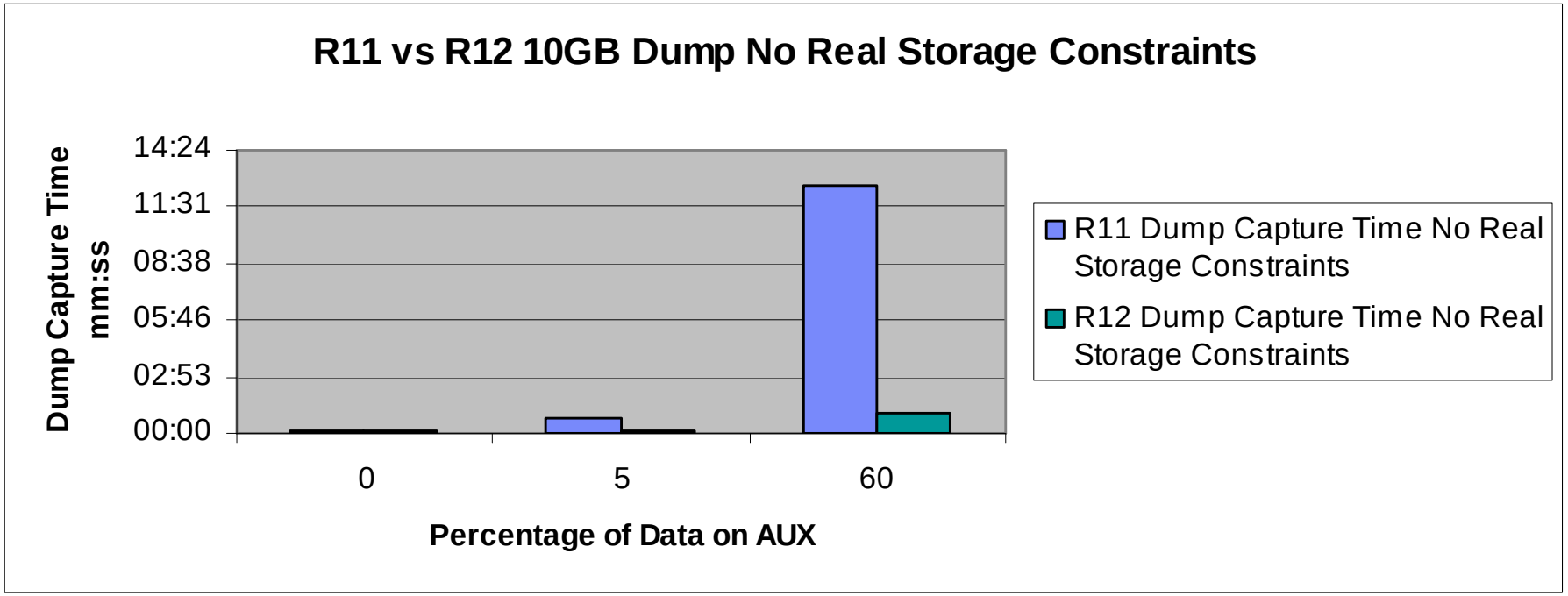
z/OS V1.12 with 60% of Data on Aux and No Real Storage Constraints



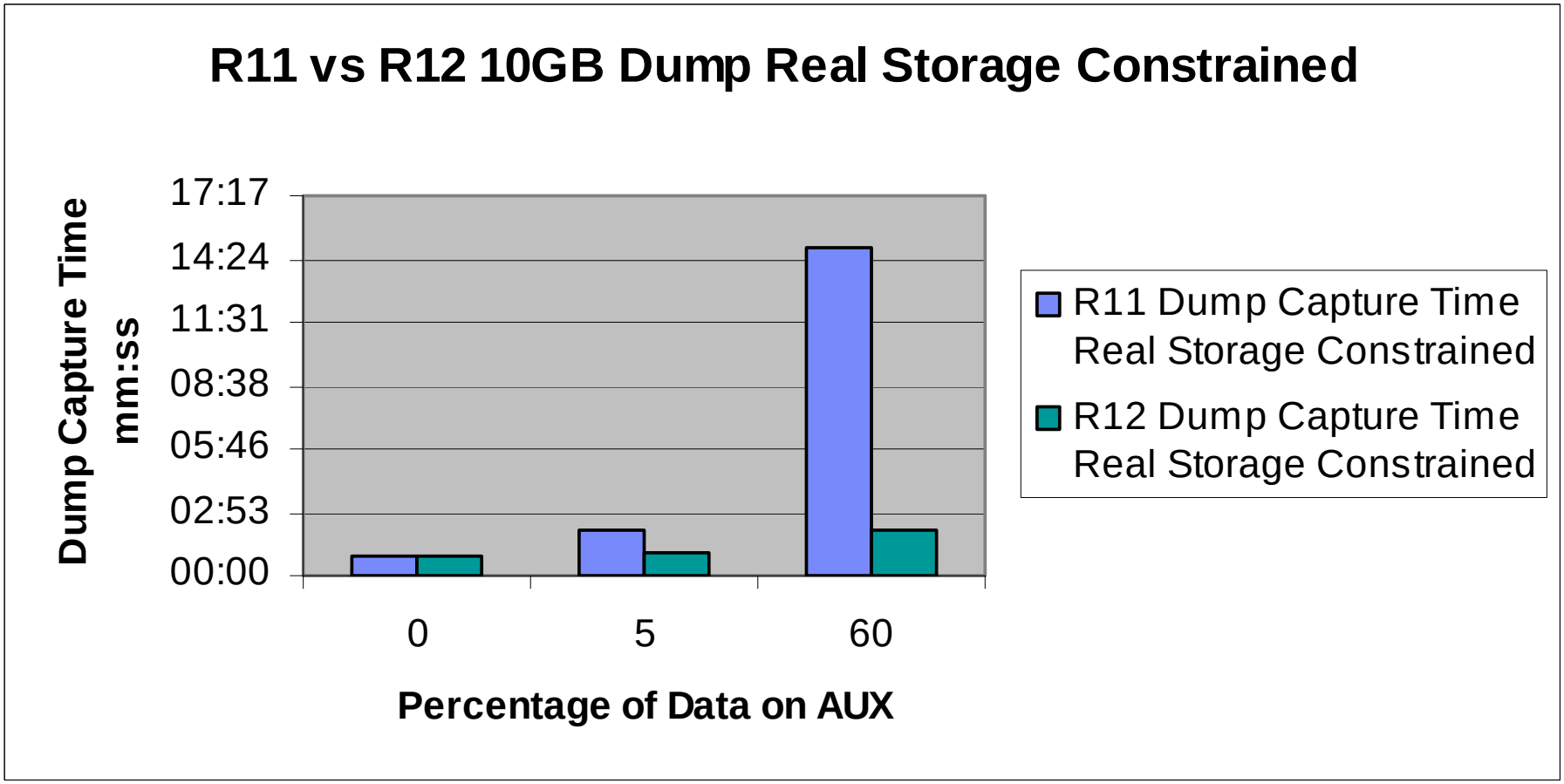
z/OS V1.12 with 60% of Data on Aux and Real Storage Constraints



z/OS V1.12 10 GB Dump and No Real Storage Constraints



z/OS V1.12 10 GB Dump and Real Storage Cosntraints



z/OS V1.12 SVC Dump Improvements with 60% of Data on Aux

Dump Size	No real storage constraint	Real storage constrained
10 GB	91%	86%
20 GB	92%	76%
40 GB	93%	65%

- Over 90% performance improvement measured in systems without real storage constraints, and only slightly increasing benefits as size of dump increases.

- In real-storage-constrained systems, the performance benefits are significant, although less pronounced (65%-86%). As the size of the dump increases, the observed performance benefits decrease from 86% to 65%.

z/OS V1.12 SVC Dump Improvements with 5% of Data on Aux

Dump Size	No real storage constraint	Real storage constrained
10 GB	78%	50%
20 GB	79%	52%
40 GB	82%	50%

- About 80% performance improvement measured in systems without real storage constraints, and only slightly increasing benefits as size of dump increases.

- In real-storage-constrained systems, the performance benefits are significant, although less pronounced (around 50%).

- Performance improvements are roughly 60% better in non-real storage constrained environments as compared to real-storage constrained environments.

z/OS V1.12 SVC Dump Improvements Summary

- Dramatic performance improvements observed in capture time for address spaces with high percentage on AUX, especially in environments with no-real-storage constraints.
  - 40 GB dump, 60% on AUX, 55min-R11 vs. 4 min-R12
- Significant performance improvements also observed in capture time for address spaces with as little as 5% on AUX, in constrained and non-constrained environments
  - 40 GB dump, 5% on AUX, 50-80% improvement
- Future performance runs will focus on the improvements in capture time for common storage (whole system non-dispatchable).

z/OS V1.13 Performance Results

V1.13 Performance Content and Results

- Based on z/OS release content the expectation was performance equivalence (0%) for V1.13 compared to V1.12 for LSPR workloads.
- Tests performed on z196 configurations: 1w, 8w, 32w, and 80w
- Results: Average system ITR delta with LSPR workloads on V1.13 for all configurations on z196 is 0%.
- Unicode services: Reduced CPU cost per conversion (character, case, normalization, collation):
  - 11% to 97% for 8-byte data size
  - 1% to 18% for 4K data size
- PKI (Public Key Infrastructure) Services:
  - CRL (Certificate Revocation List) processing greatly improved (99% reduction in processing time).
  - New DB2 support can handle 5X more transactions per second than VSAM.
- zFS
  - Equivalent to V1.12 for monoplex.
  - Most sysplex environments will see between 1.5 - 2.5X improvement.
  - Application location is less important
- RSM CTRACE: reduced overhead
- SDSF SORT: reduced CPU cost when sorting large amounts of output
- IEBCOPY:
  - APF requirement removed
  - Performance improvements

z/OS V1.13 RMF Serialization Delay Report

- RMF XML report using RMF XML Toolkit
- Start RMF III to collect data in SMF record type 72 subtype 5
- Postprocessor:
  - //XRPTS DD name for output data set
  - REPORTS(SDELAY)
- Install RMF XML Reporter from RMF Web site:
  - http://www.ibm.com/systems/z/os/zos/features/rmf/

RMF Postprocessor Interval Report [System P32] : Serialization Delay Report

z/OS V1R13 Start: 05/25/2011-10.28.20 Interval: 10:49:000 minutes

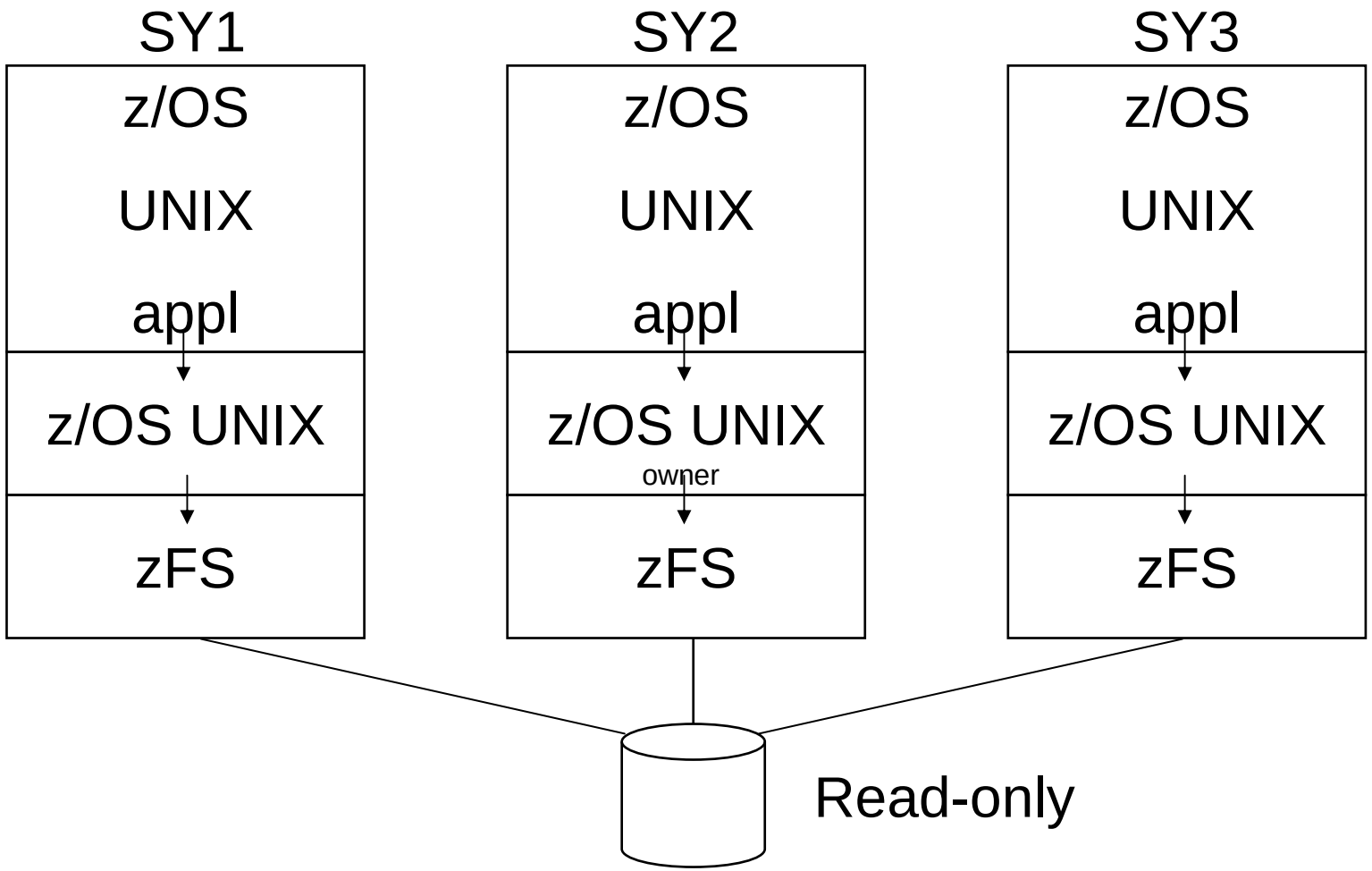
V1R13 RMF End: 05/25/2011-10.39.08 Cycle: 1000 milliseconds

Serialization Delay Summary

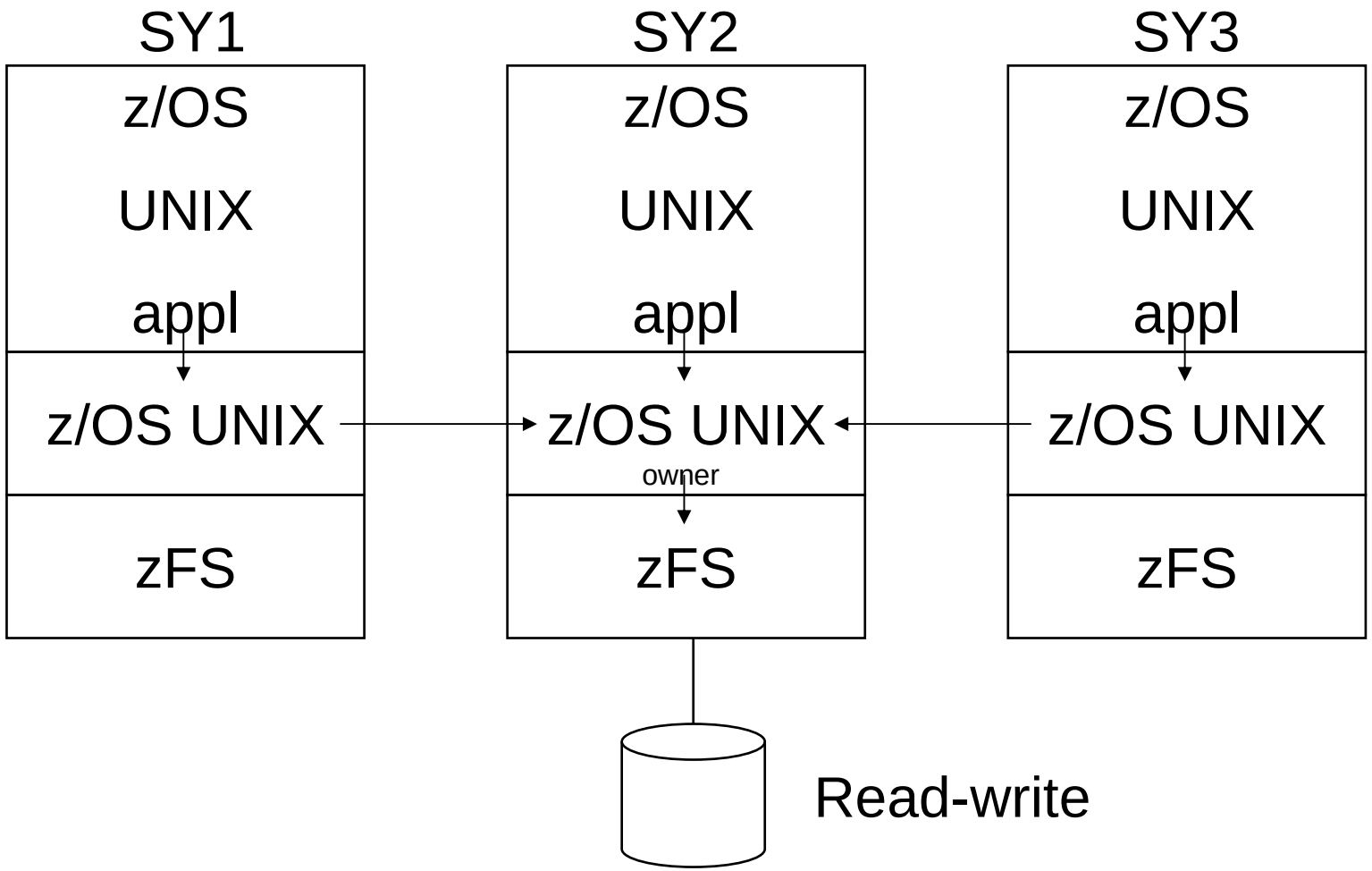
System Locks

Lock Type	Total Contention Time	Avg Contention Time	Total Contention Count	Contention Count with QLen>1
CMS	0	0.00	15	0
CMSEQDQ	0	0.00	5	0
CMSLatch	0	0.00	1	0
CMSSMF	3553	0.01	306436	35253
Local	125093	0.00	18338201	1755528
CML Owner	1337257	0.24	5376109	3233965

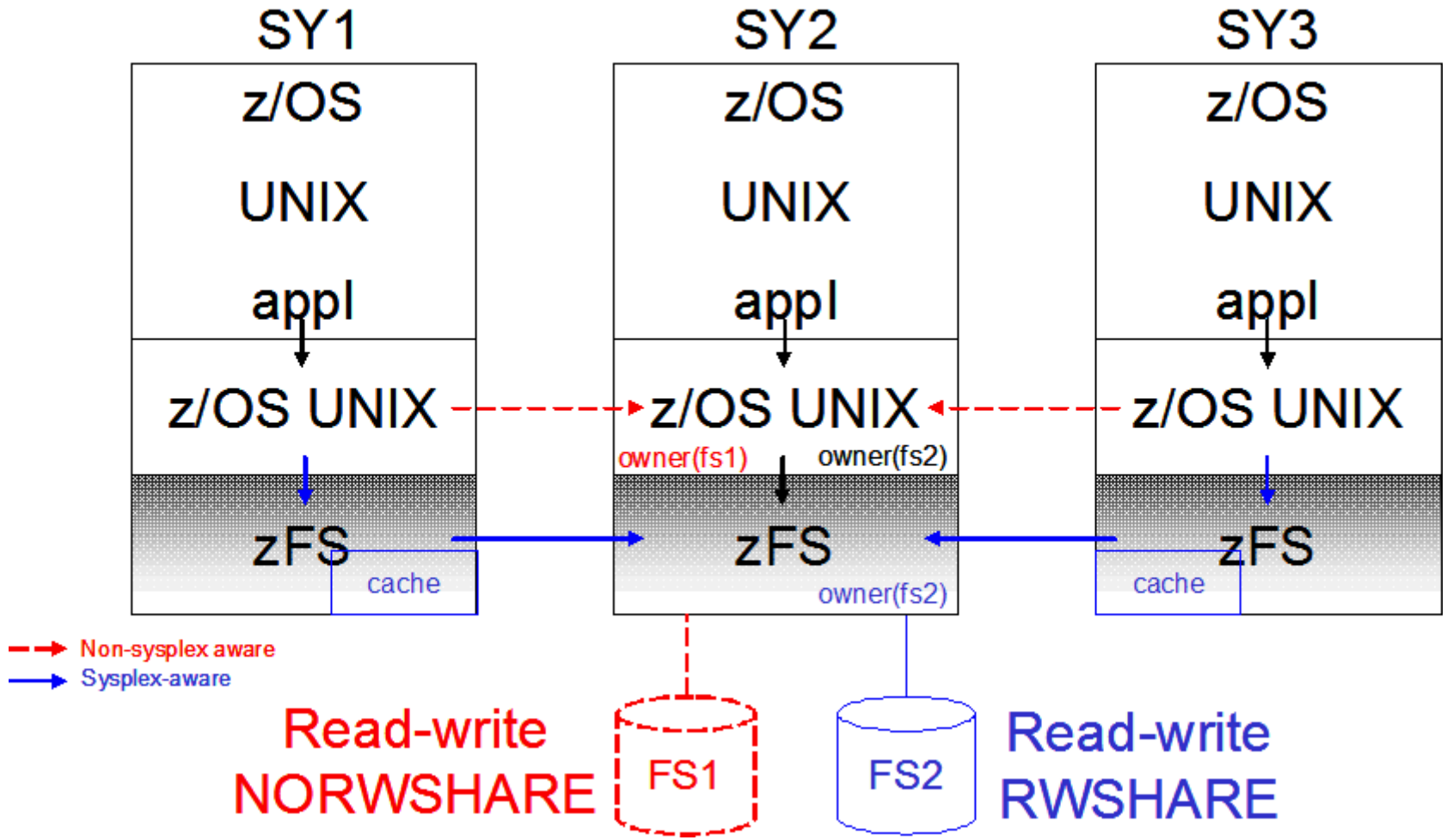
zFS File System Sharing Read-only



Sharing a zFS Filesystem Read-Write prior to z/OS V1.11

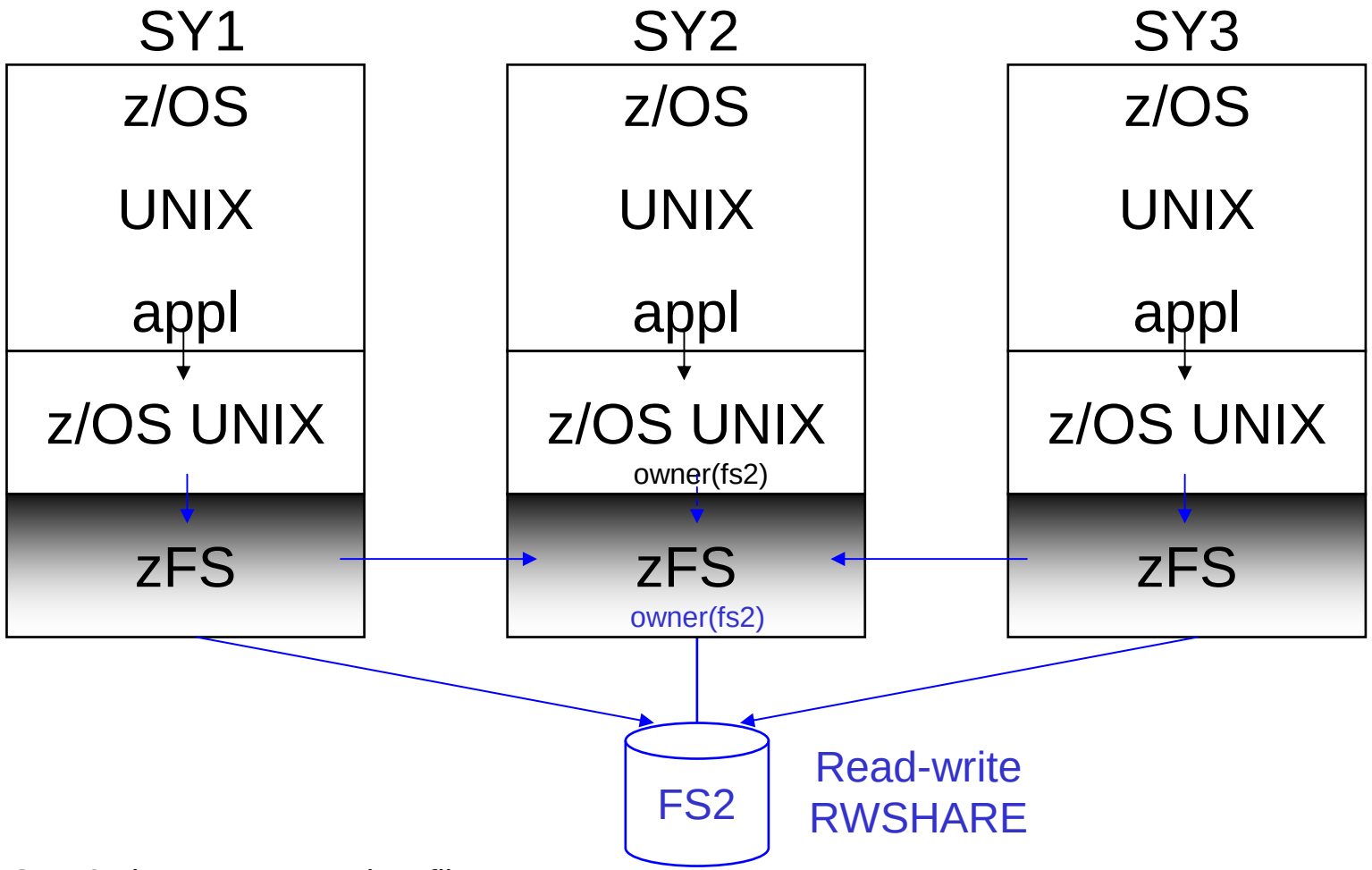


Sharing a zFS Filesystem Read-Write in V1.11 with apar OA29619



IOEPRMxx: SYSPLEX=FILESYS
BPXPRMxx: SYSPLEX(YES)

V1.13 zFS Direct I/O



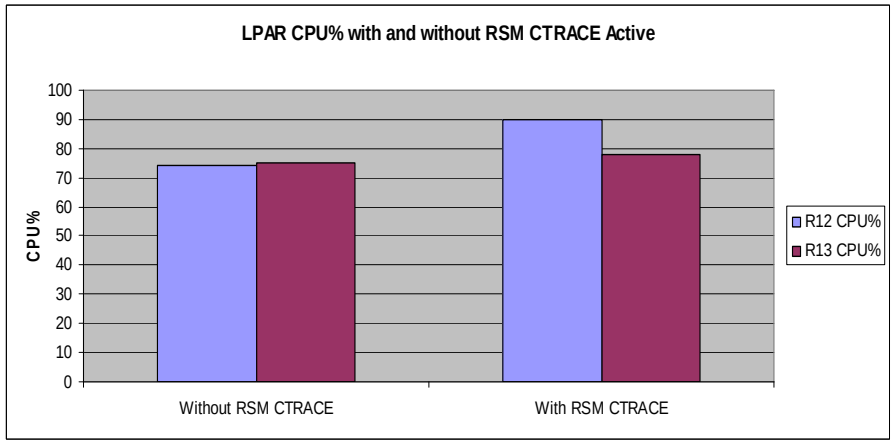
zFS R13 always runs sysplex=filesys

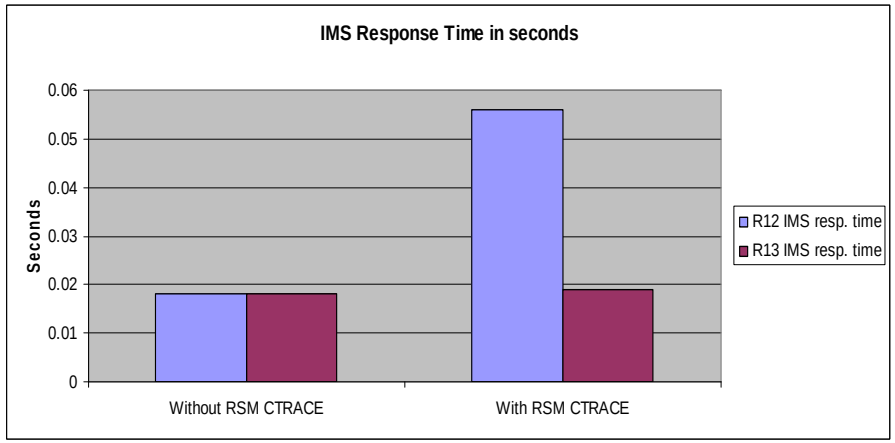
z/OS V1.13 Migrating to Shared zFS

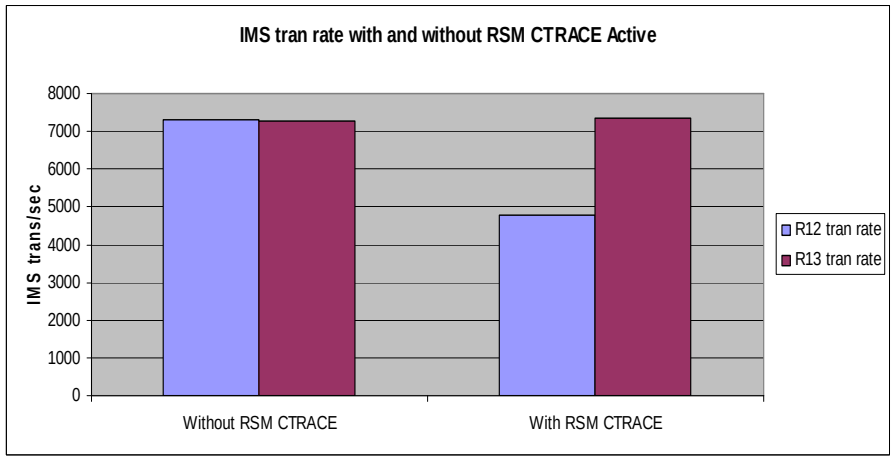
- zFS R13 always runs sysplex=filesys (and it needs all other systems in the shared file system environment to be running sysplex=filesys)
- Migration to zFS R13 is a two step process:
  - Install toleration APAR OA32925 (PTF UA55765) on all zFS R11 and R12 systems and make it active with a rolling IPL.
  - Change your zFS IOEFSPRM file to sysplex=filesys on all systems and make it active with a rolling IPL.
    (Default is that all zFS read-write file systems will be non-sysplex aware - NORWSHARE)
- There is a zFS migration health check to verify the sysplex=filesys option in prior releases (ZOSMIGV1R13_ZFS_FILESYS)

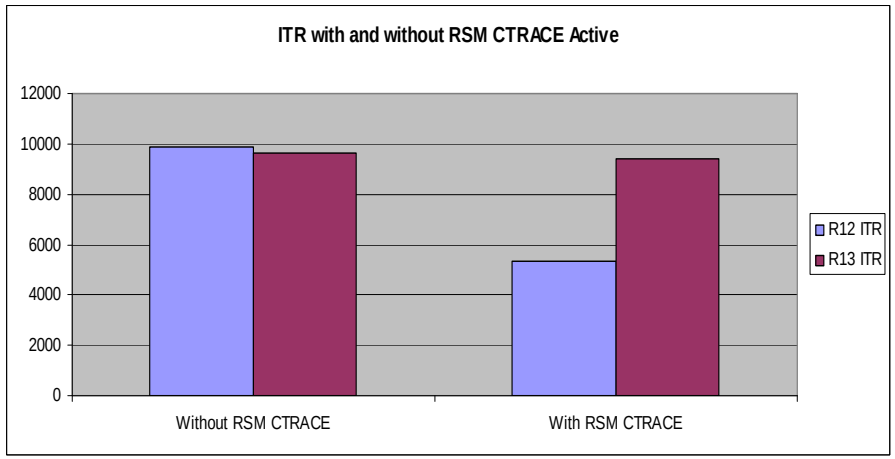
z/OS V1.13 RSM CTRACE – Reduced Overhead

- RSM CTRACE improvements in V1.13 In a system with many active CPs:
  - CPU overhead is reduced
  - IMS response times and throughput not impacted when running the CTRACE



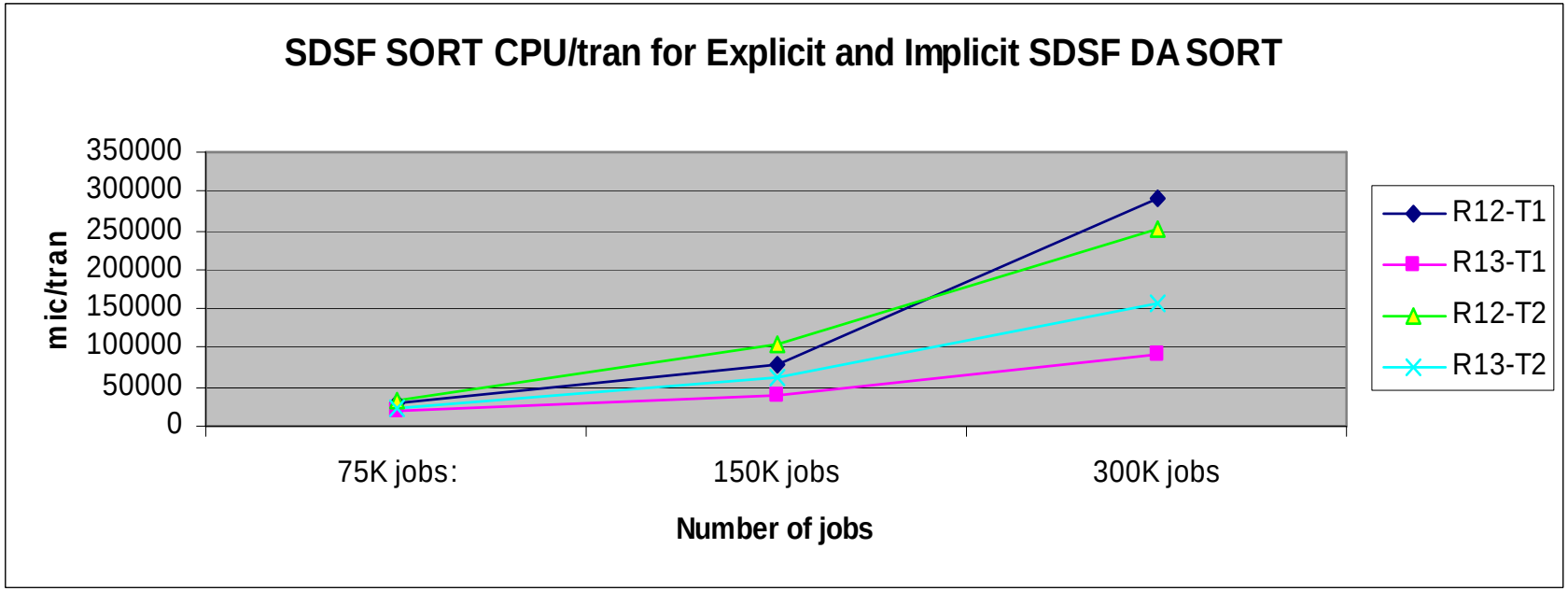






z/OS V1.13 SDSF SORT CPU Usage

- V1.13 provides reduction in CPU usage for ISPF/SDSF sort routine when sorting large number of jobs
- Test#1: SDSF O command with implicit sort
- Test#2: SDSF O command with explicit sort
- Most reduction in SDSF CPU usage when displaying and sorting the largest number of jobs.
- Improved ETR and ITR as well as 25% reduction in real storage usage.



z/OS V1.13 IEBCOPY Performance – PDS to PDS

- Copy PDS to PDS fixed block record format:
  - Copy 1500 members from PDS source to PDS target (source PDS 90% full)
  - Elapsed time reduced by 67% and CPU% reduced by 75%

- Copy PDS to PDS variable block record format:
  - Copy 1500 members from PDS source to PDS target (source PDS 90% full)
  - Elapsed time reduced by 60-68% and CPU usage reduced by 75-80% depending on record size

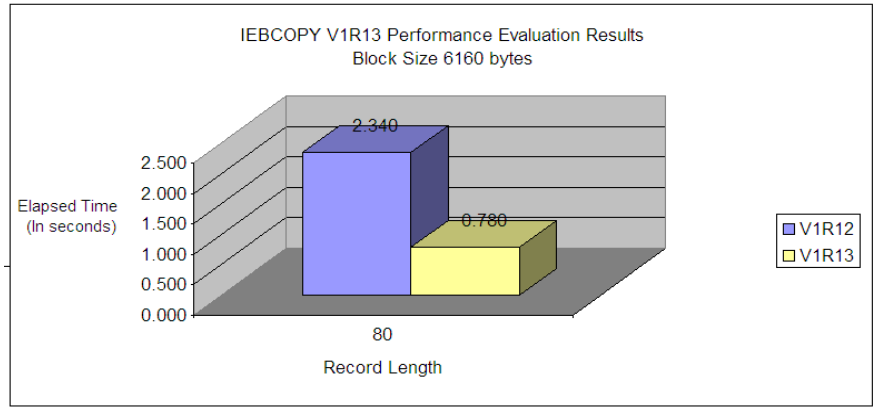
Copy PDS to PDS Testing results
Block size 6160 Format (FB)

Elapsed Time measurements

LRECL	V1R12	V1R13	Delta (%)
80	2.340	0.780	-66.67
132	*	*	---
4096	*	*	---

CPU Time measurements

LRECL	V1R12	V1R13	Delta (%)
80	0.004	0.001	-75.00
132	*	*	---
4096	*	*	---



Note: * .- Record length is inconsistent with block size for this record format. Test not executed for this variation

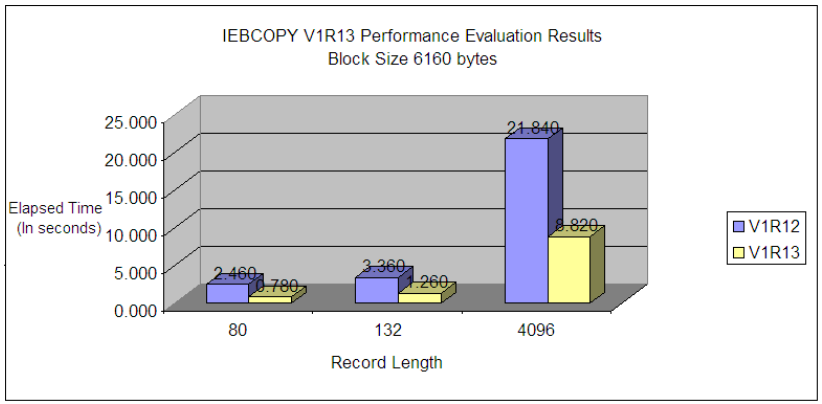
Copy PDS to PDS Testing results
Block size 6160 Format (VB)

Elapsed Time measurements

LRECL	V1R12	V1R13	Delta (%)
80	2.460	0.780	-68.29
132	3.360	1.260	-62.50
4096	21.840	8.820	-59.62

CPU Time measurements

LRECL	V1R12	V1R13	Delta (%)
80	0.004	0.001	-75.00
132	0.005	0.001	-80.00
4096	0.005	0.001	-80.00



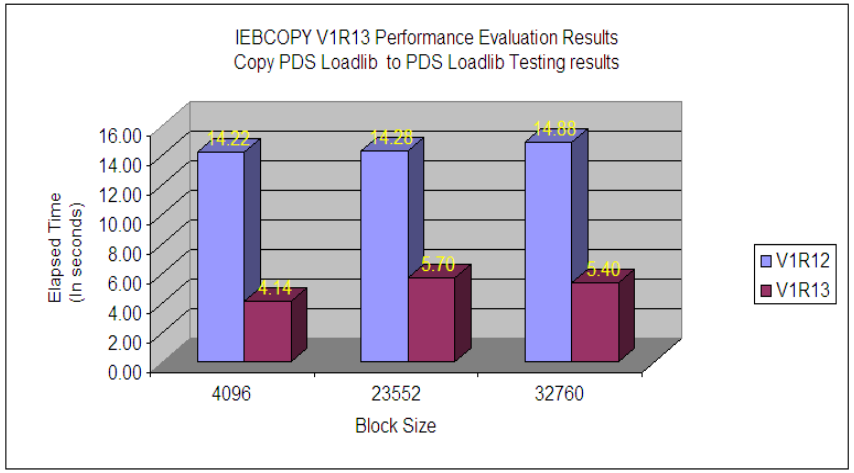
z/OS V1.13 IEBCOPY Performance – PDS Loadlib

- Copy PDS loadlib to PDS loadlib:
  - Copy 1000 members from PDS source loadlib to PDS target loadlib (source PDS 90% full)
  - Elapsed time reduced by 64 -71% depending on block size
- Copy PDS loadlib to sequential:
  - Copy 1000 members from PDS source loadlib to sequential target (source PDS 90% full)
  - Elapsed time reduced by 31-72% depending on block size

Copy PDS Loadlib to PDS Loadllib Testing results
LRECL=0 Format (U)

Elapsed Time measurements

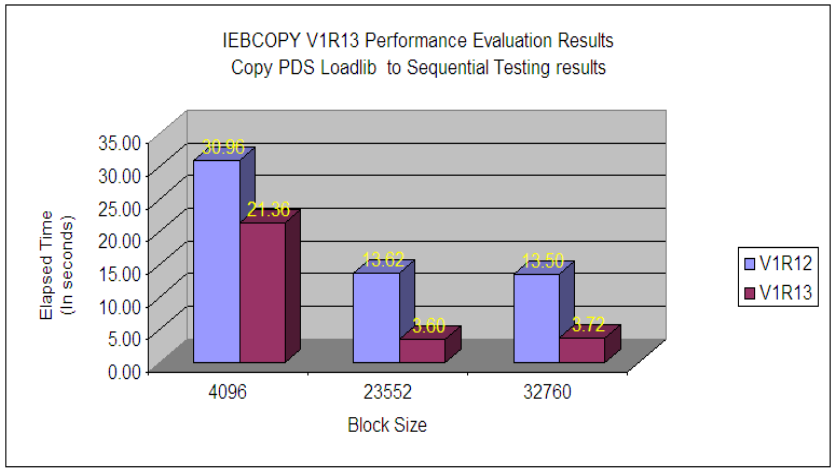
BLKSIZE	V1R12	V1R13	Delta (%)
4096	14.22	4.14	-70.89
23552	14.28	5.70	-60.08
32760	14.88	5.40	-63.71



Copy PDS Loadlib to SEQ Testing results
LRECL=0 Format (U)

Elapsed Time measurements

BLKSIZE	V1R12	V1R13	Delta (%)
4096	30.96	21.36	-31.01
23552	13.62	3.60	-73.57
32760	13.50	3.72	-72.44



z/OS V1.13 IEBCOPY Performance - Compress

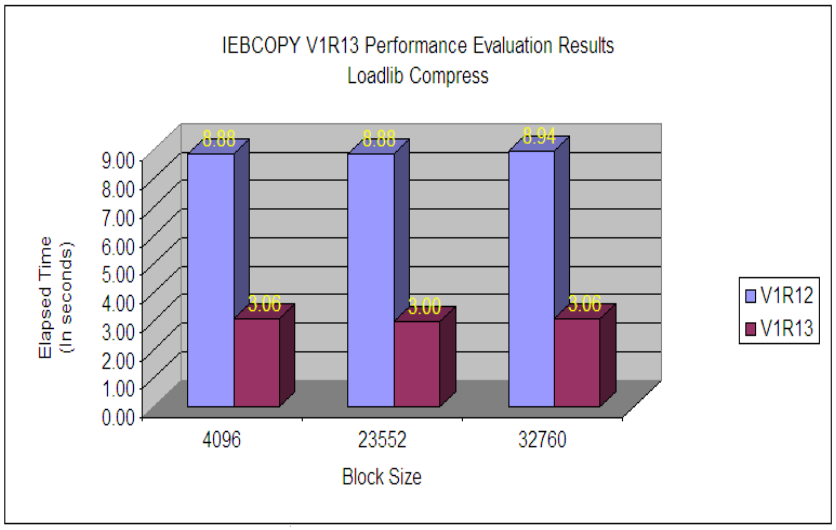
- Compress PDS data set record format undefined:
  - Delete 500 members and compress (target PDS had 500 members and was 90% full)
  - Elapsed time reduced by 66%

- Compress PDS data set record format fixed:
  - Delete 500 members and compress (prior to delete the target PDS had 1500 members and was 90% full)
  - Elapsed time reduced by 34-63% depending on block size

Loadlib Compress Testing results
LRECL=0 Format (U)

Elapsed Time measurements

BLKSIZE	V1R12	V1R13	Delta (%)
4096	8.88	3.06	-65.54
23552	8.88	3.00	-66.22
32760	8.94	3.06	-65.77



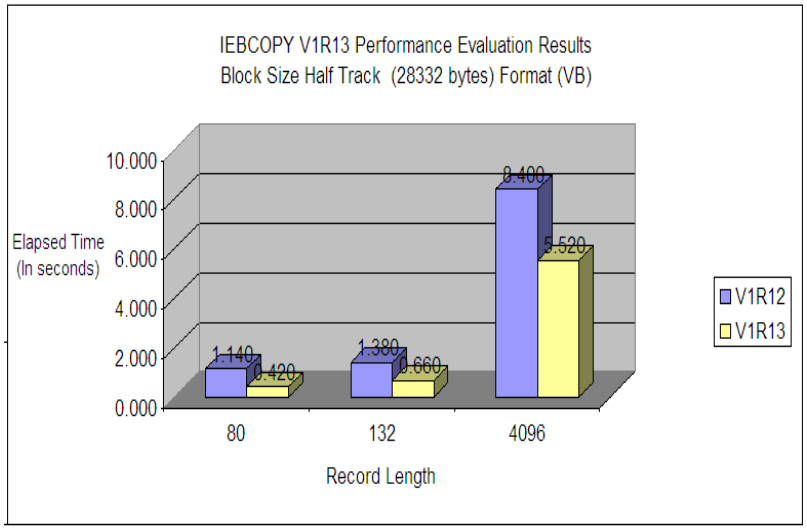
Compress PDS Testing results
Block size 28332 Format (VB)

Elapsed Time measurements

LRECL	V1R12	V1R13	Delta (%)
80	1.140	0.420	-63.16
132	1.380	0.660	-52.17
4096	8.400	5.520	-34.29

CPU Time measurements

LRECL	V1R12	V1R13	Delta (%)
80	0.000	0.000	0.000
132	0.001	0.000	-100.00
4096	0.004	0.004	0.000



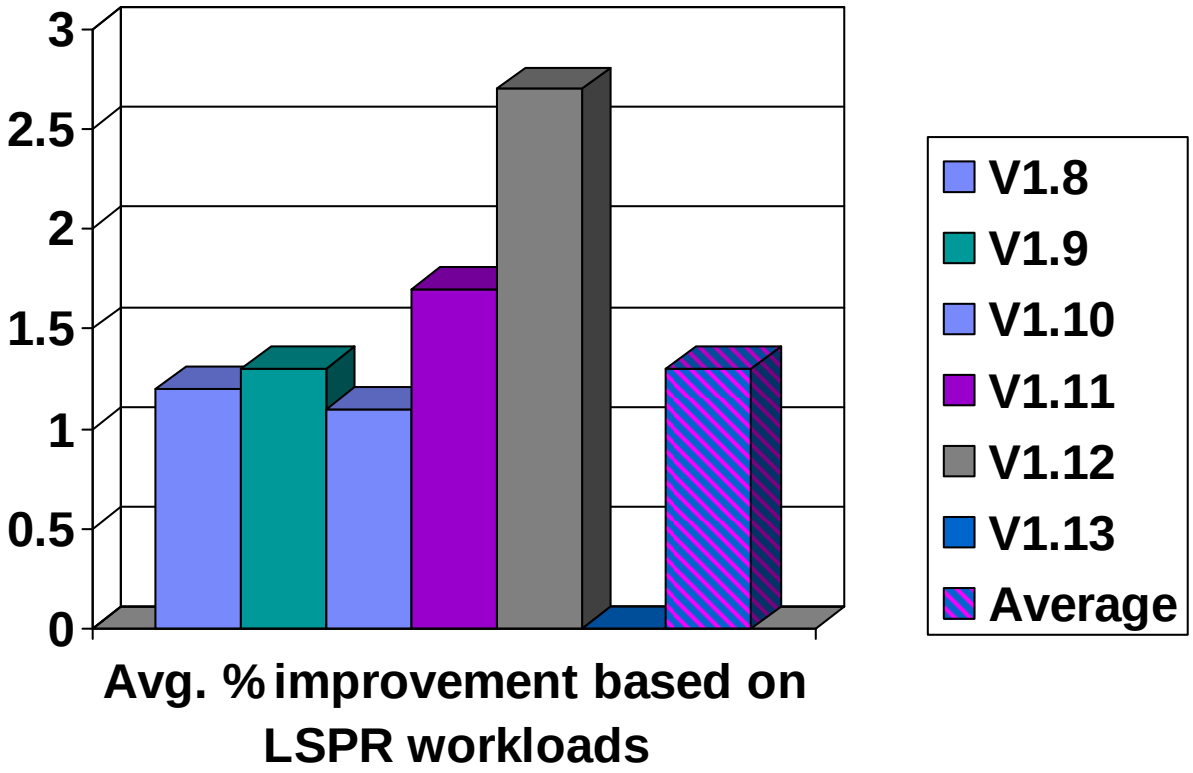
z/OS V1.12 and V1.13 Results Summary

- V1.12:
  - CPU performance goal: Overachieved: avg. 2.7% vs goal of 1.25%
  - VSCR results:
    - Across workloads : 8-12 MB savings (target 4 MB)
    - RMF Cpools to 64-bit : ~10 MB
    - MMSB to 64-bit : ~ 1 MB

- V1.13:
  - CPU performance goal: equivalent to V1.12
  - VSCR results:
    - SSRB move to 64-bit: benefit depending on workloads
    - Communication Server ctrace move to 64-bit:
      - VIT: saved 4MB of storage below the bar
      - TN3270 ctrace saved about 256MB below the bar

z/OS Release-to-Release Performance Improvements

- Focus on performance improvements:
  - Future z/OS releases
  - z/OS SW stack



z/OS Performance References

- System z ATS zSoftCap tool:
  - Customer version: http://www.ibm.com/support/techdocs
  - IBM internal version: http://w3.ibm.com/support/americas/wsc/cpsproducts.html
  - IBM business partners: http://partners.boulder.ibm.com/src/atsmastr.nsf/WebIndex/PRS1762
- LSPR web site (LSPR workloads):
  - https://www.ibm.com/servers/resourcelink/lib03060.nsf/pages/lsprindex

SHARE sessions:

- 10625: Significant Enhancements in z/OS V1R13 zFS
  - Tuesday March 13: 4:30-5:30 in Ballroom D (Richard Theis)
- 10891: Workload Management Update for z/OS 1.13 and 1.12
  - Monday March 12: 4:30-5:30 in Juniper (Stefan Wirag)