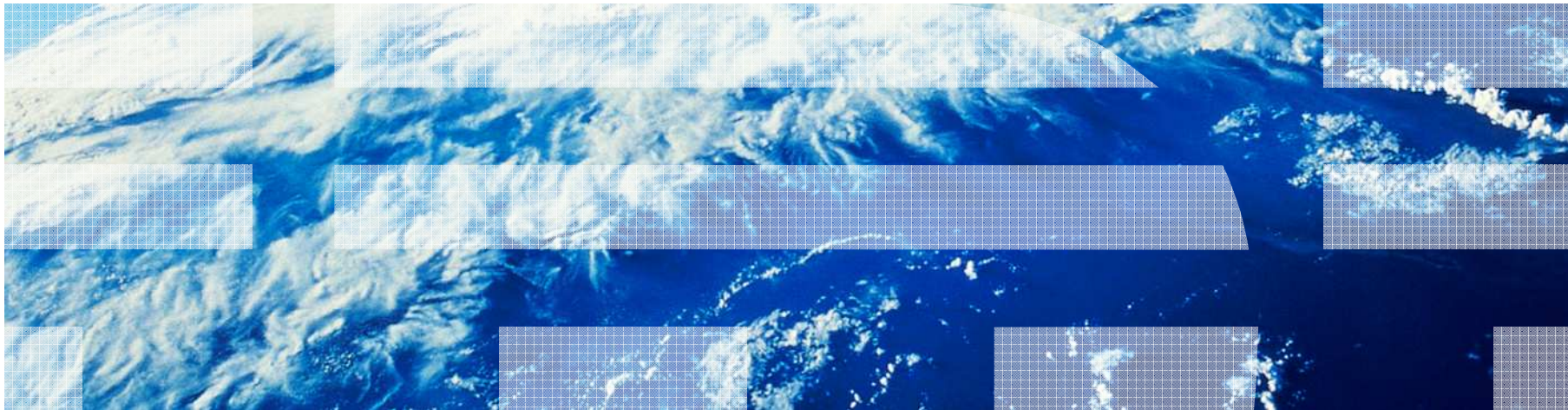


z/OS V1.13 - A Performance Update

Share Session: 12915

San Francisco, CA

02/05/2013



Trademarks

IBM Corporation 2013

- 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](http://www.ibm.com/legal/copytrade.shtml)” 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.

z/OS V1.13 – A Performance Update - Agenda

- 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 V1.13 and zEC12 Transactional Execution Facility

- z/OS V1.13 RSM Enablement Offering (FMID JBB778H) :
 - zFlash Express exploitation
 - 1 MB Pageable large pages

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.

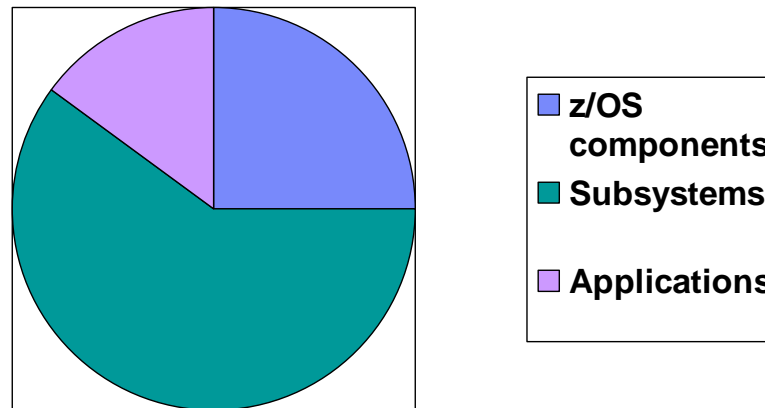
Workloads and Coverage for z/OS Release Testing

- **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 z/OS 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)

V1.13 Performance Content and Results

- Based on z/OS release content the expectation was performance equivalence (0%) for the LSPR workloads on V1.13 compared to V1.12.
- Tests performed on z196 configurations: 1w, 8w, 32w, and 80w
- LSPR regression test results: average system ITR delta for V1.13 compared to V1.12 for all configurations on z196 is 0%.
- V1.13 Performance improvements:
 - zFS
 - Equivalent to V1.12 for monoplex.
 - Most sysplex environments will see between 1.5 - 2.5X improvement.
 - Application location is less important
 - 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.
 - RSM CTRACE: reduced overhead
 - SDSF SORT: reduced CPU cost when sorting large amounts of jobs in output queue
 - 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

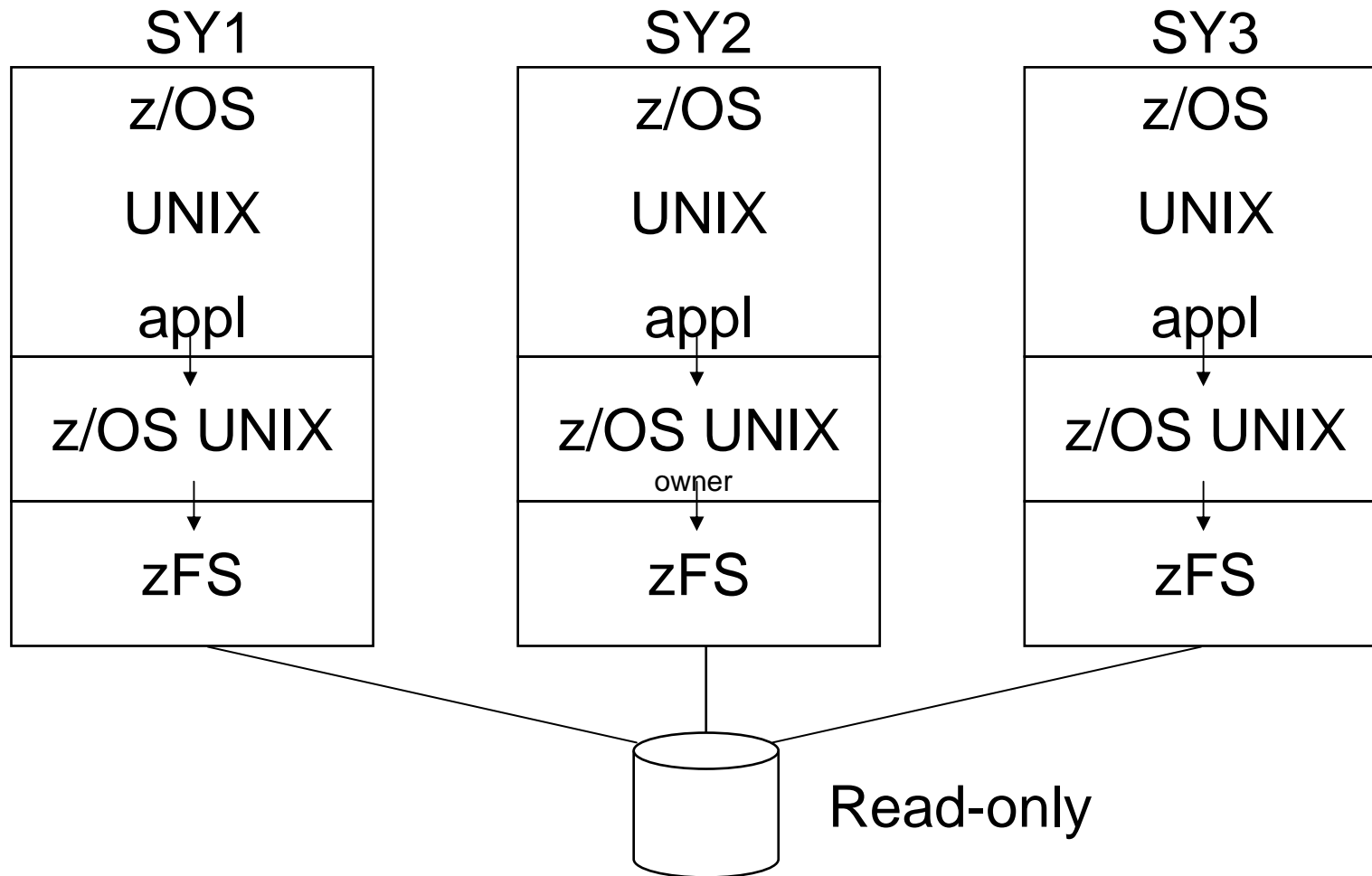
V1R13 RMF End: 05/25/2011-10.39.08 Cycle: 1000 milliseconds

Serialization Delay Summary

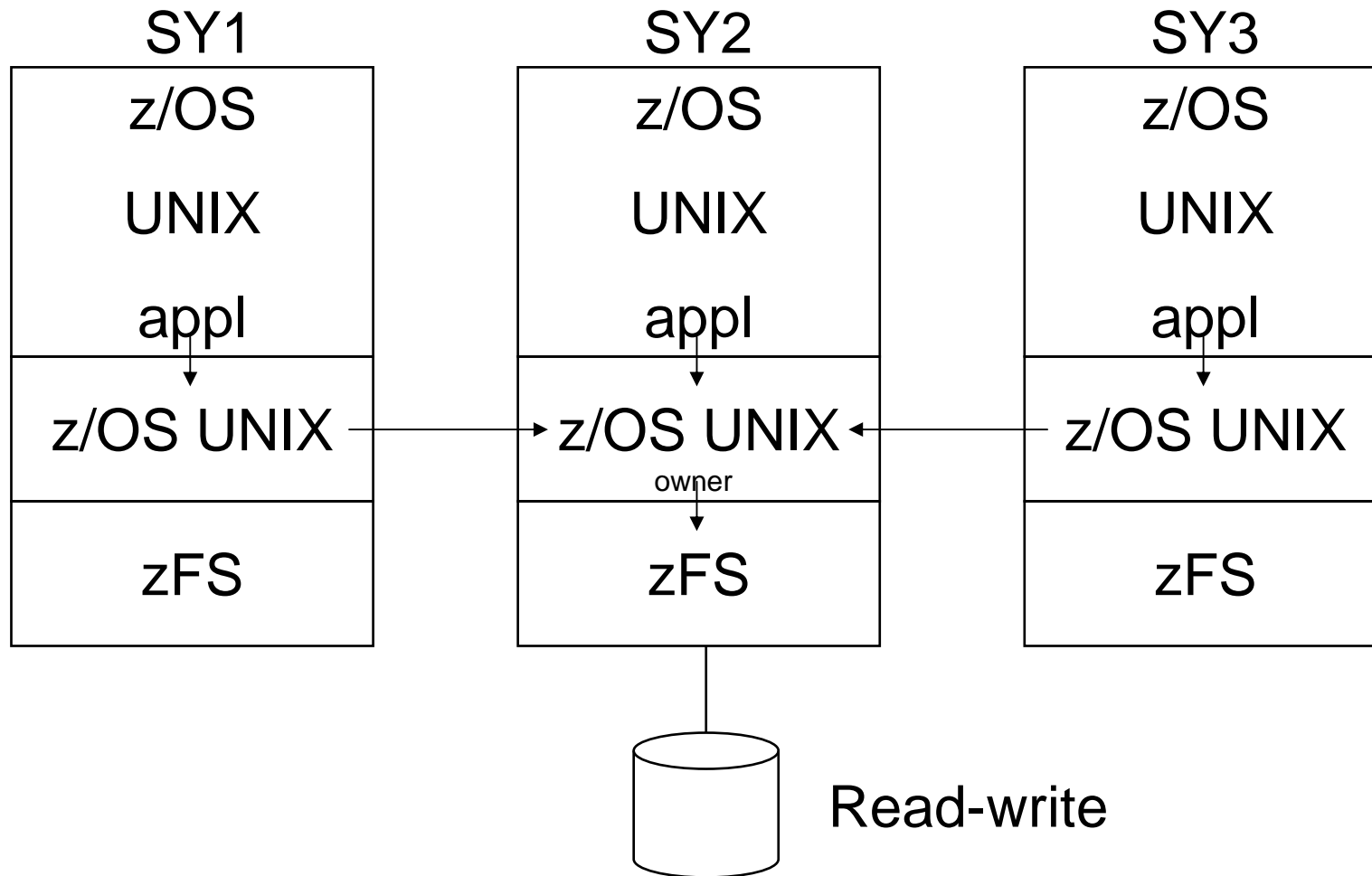
System Locks

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

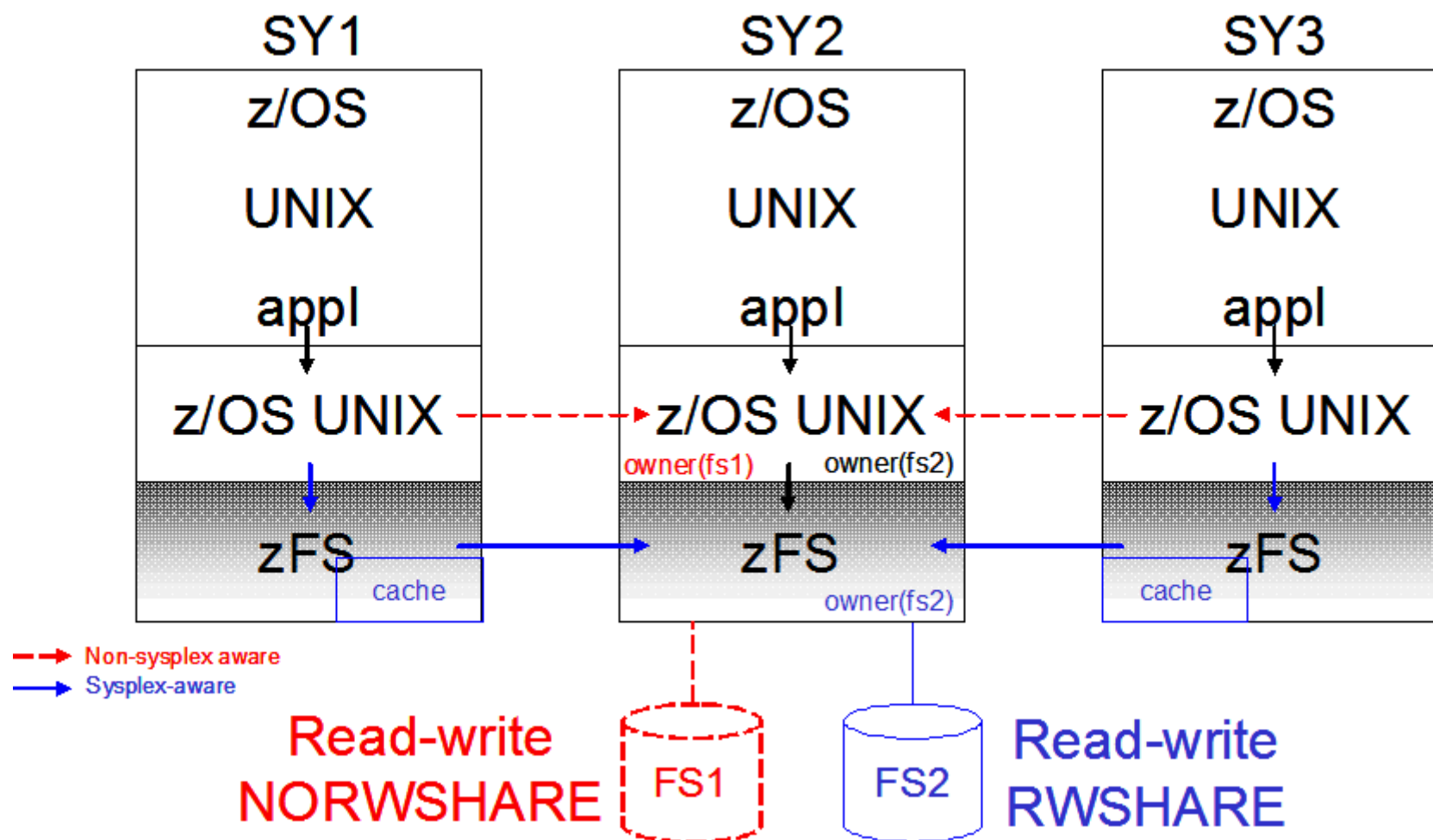
zFS File System Sharing Read-only



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

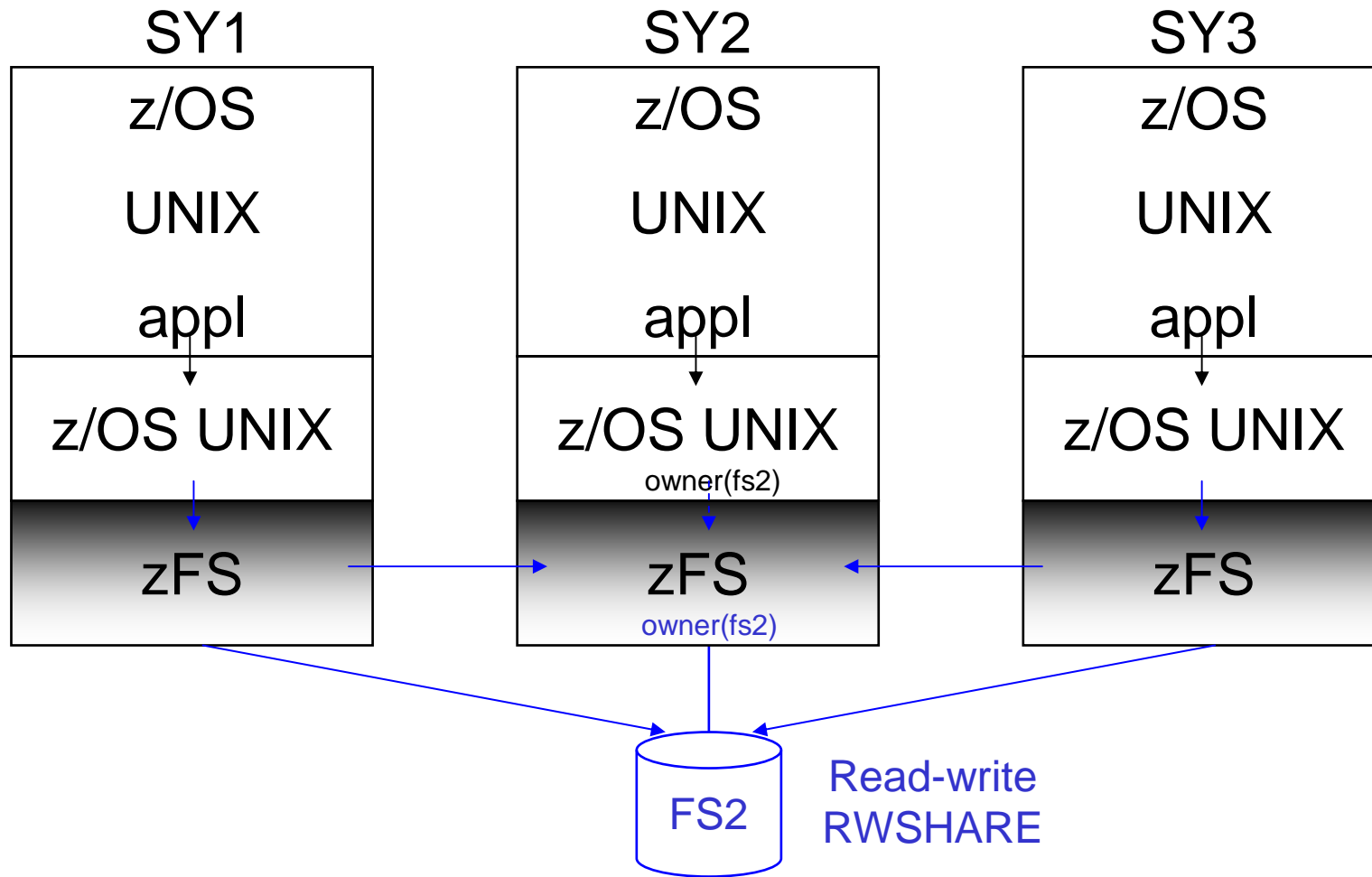


Sharing a zFS Filesystem Read-Write in V1.11 with apar OA29619



IOEPRMxx: SYSPLEX=FILESYS
BPXPRMxx: SYSPLEX(YES)

V1.13 zFS Direct I/O



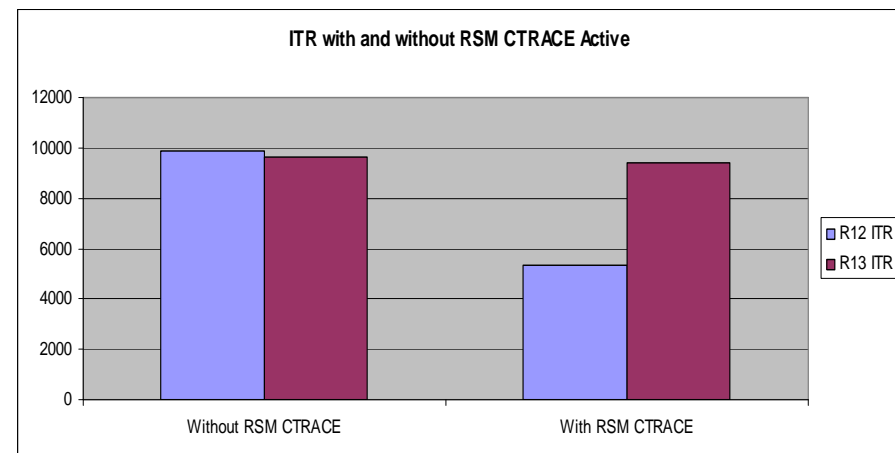
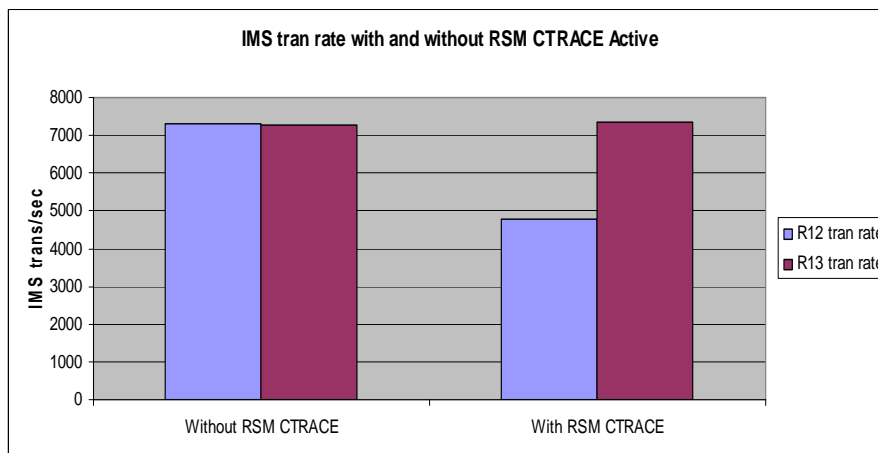
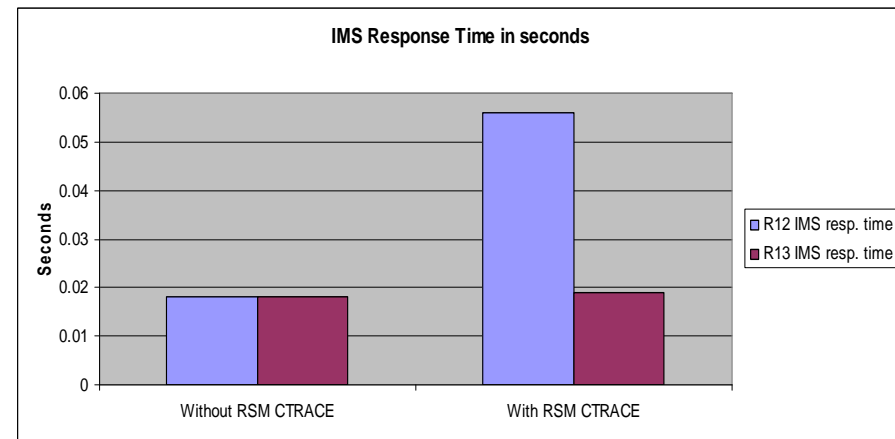
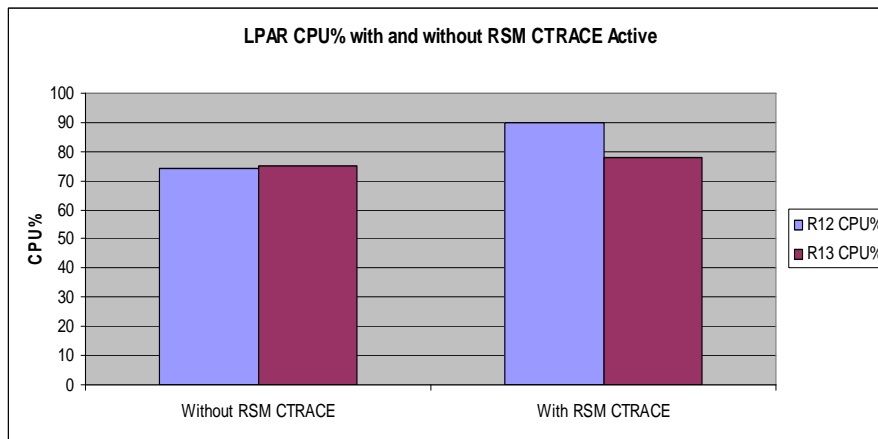
zFS R13 always runs sysplex=filesys

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`)

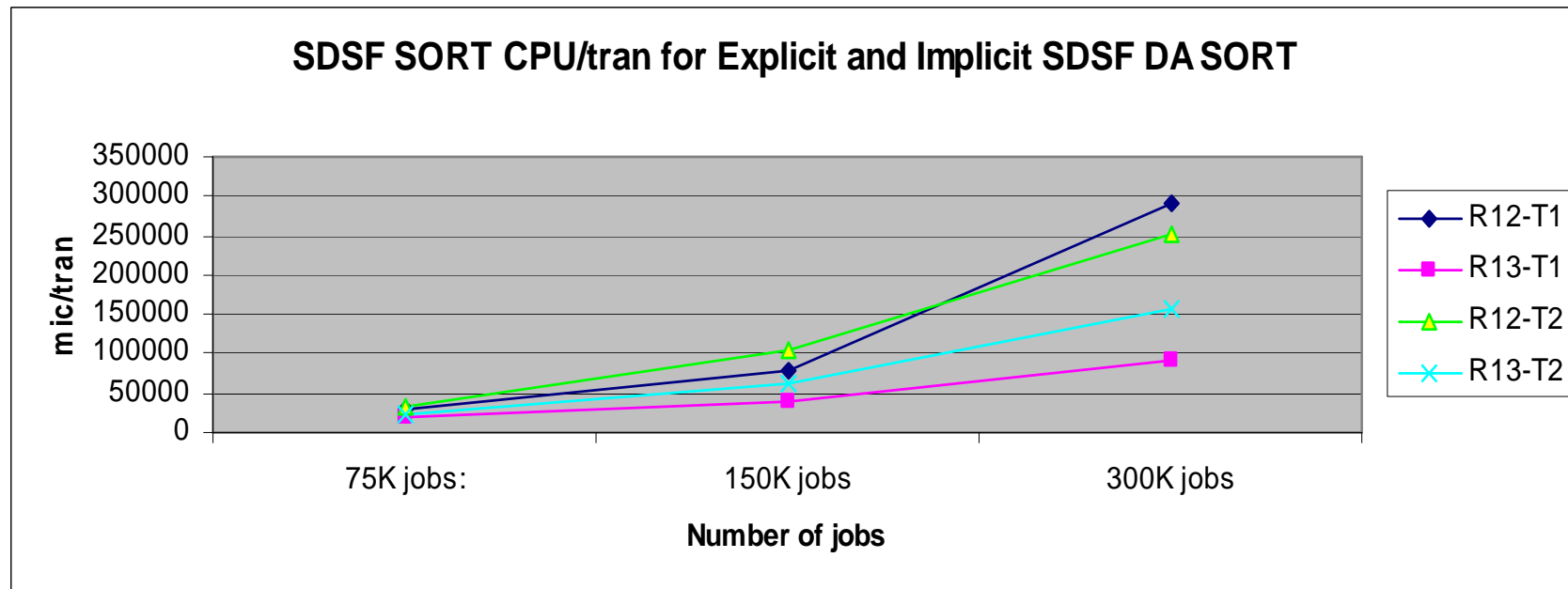
z/OS V1.13 RSM CTRACE – Reduced Overhead

- 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 on the output queue.
- 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 on the output queue.
- 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

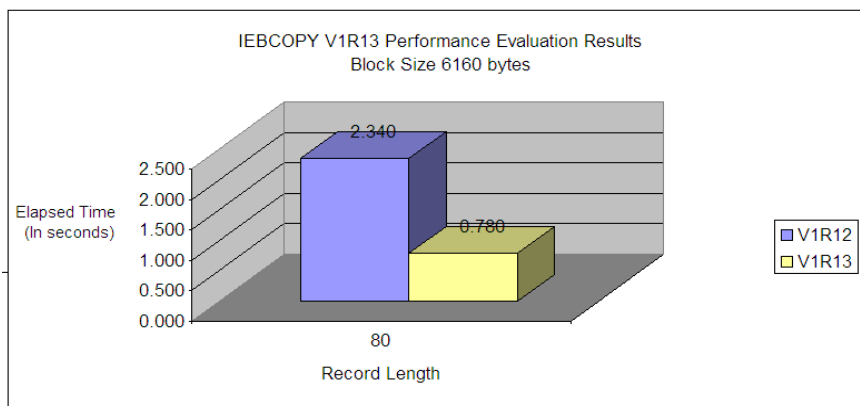
Copy PDS to PDS Testing results
Block size 6160 Format (FB)

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

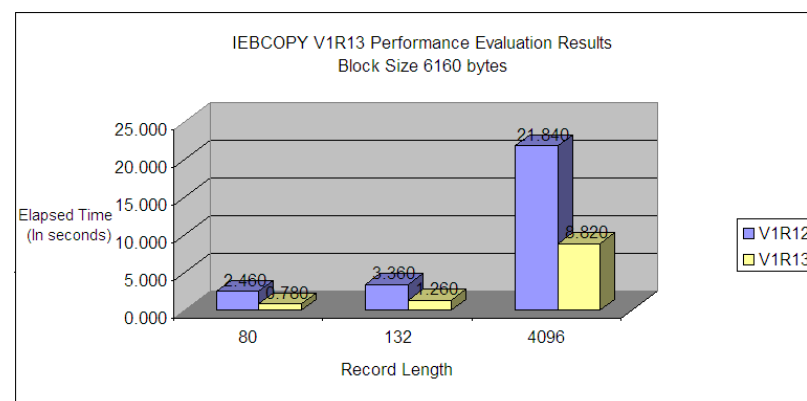
Copy PDS to PDS Testing results
Block size 6160 Format (VB)

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

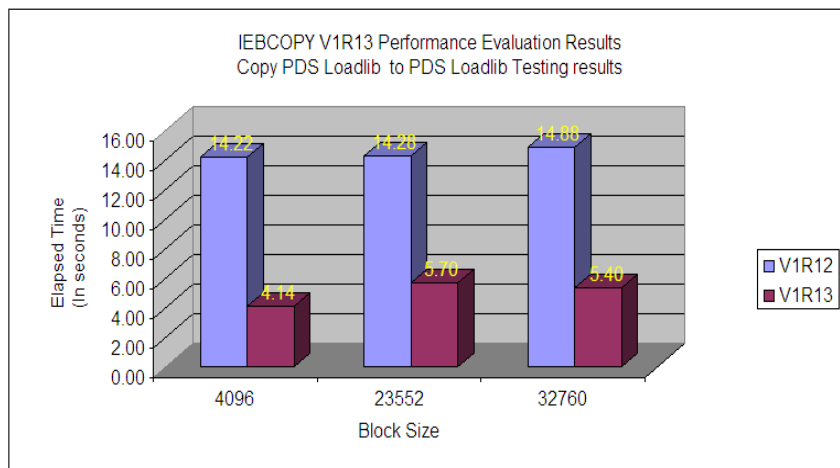
- Copy PDS loadlib to PDS loadlib:
 - Copy 1000 members from PDS source loadlib to PDS target loadlib (source PDS 90% full)
 - Elapsed time reduced by 64 -71% depending on block size

- 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 Loadlib Testing results
LRECL=0 Format (U)

Elapsed Time measurements

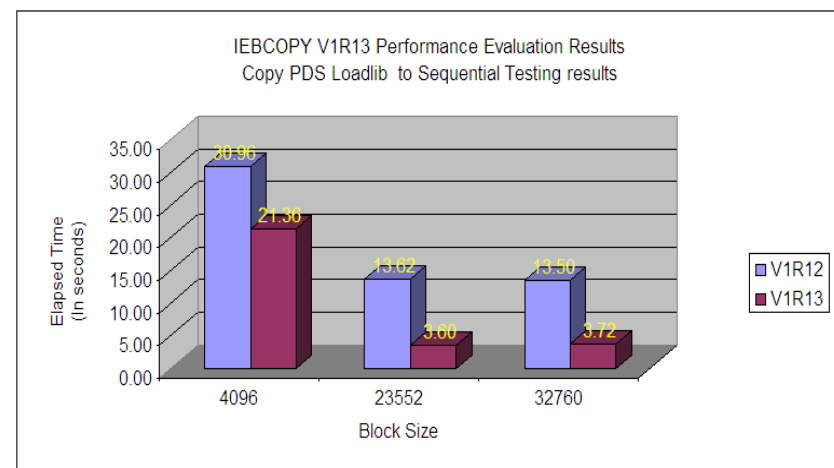
BLKSIZE	V1R12	V1R13	Delta (%)
4096	14.22	4.14	-70.89
23552	14.28	5.70	-60.08
32760	14.88	5.40	-63.71



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



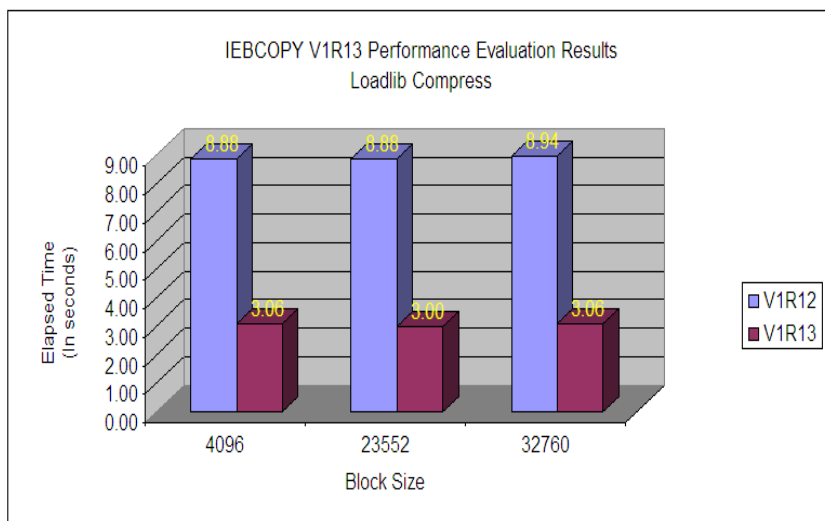
z/OS V1.13 IEBCOPY Performance - Compress

- 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 record 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



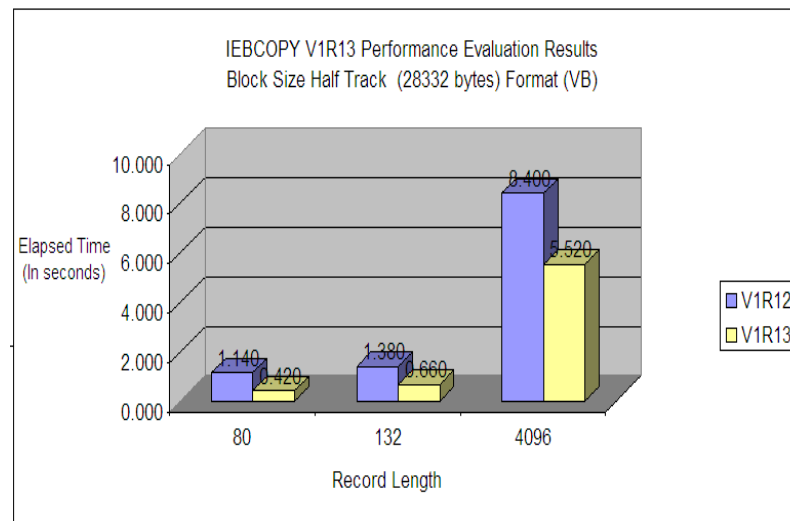
Compress PDS Testing results
Block size 28332 Format (VB)

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

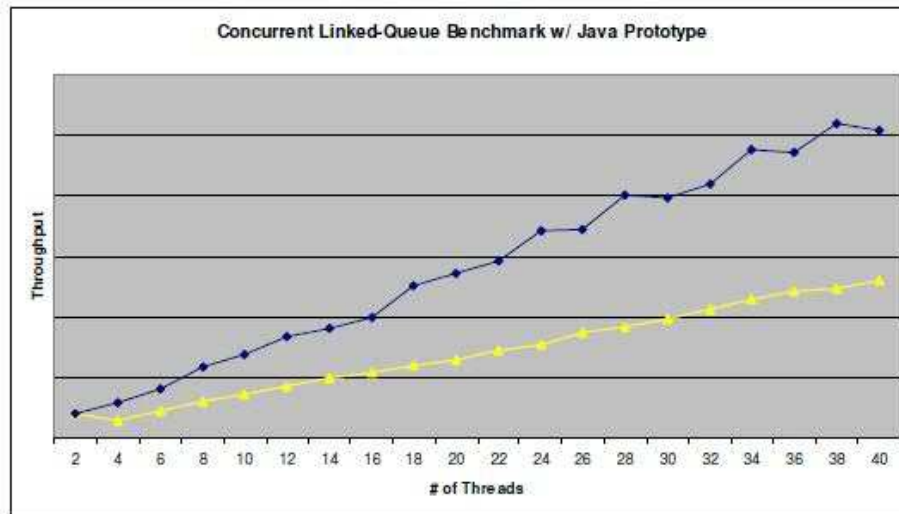
CPU Time measurements

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

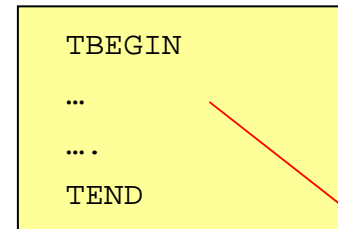


zEC12 Transactional Execution Facility

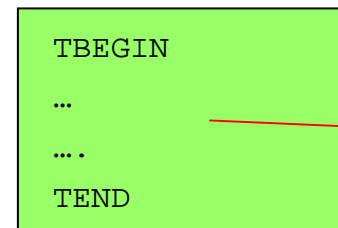
- zEC12 first system to offer Transactional Execution Facility
- New serialization option in a multi-threaded environment
 - Reduce use of locks and latches
- New set of instructions allows defining groups of instructions to be executed atomically.
- z/OS support:
 - V1.13 with apar OA38829 for test
 - Java 7 SR3
 - V2.1 full support with diagnostic aids and tools
- Performance:
 - Java: up to 2x improved scalability of `java.util.concurrent.ConcurrentLinkedQueue`
 - Example of multi-threaded work queue:



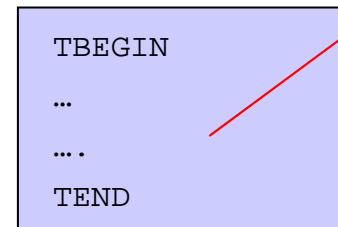
Thread A



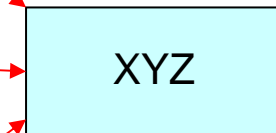
Thread B



Thread C

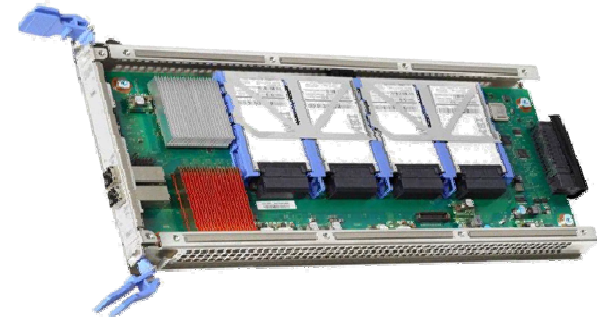


Data

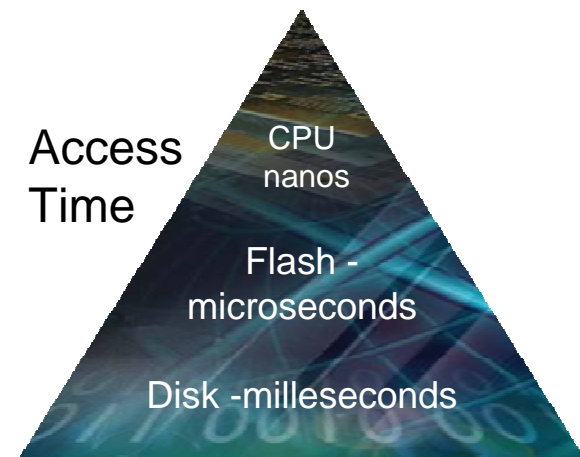


zFlash Express on zEC12

- Physically comprised of internal storage on Flash SSDs
- A new tier of memory, Storage Class Memory (SCM)
- Uses standard PCIe I/O drawer
- Supported on z/OS V1.13 plus RSM Enablement Offering
- Flash Express cards delivered as a RAID 10 mirrored card pair
 - Each card pair provides **1.4 TB** usable storage
 - Maximum 4 card pairs (4 X1.4=5.6 TB)
- Immediately usable
 - No capacity planning needed
 - No intelligent data placement needed
 - Full virtualization of card across partitions
- Robust design
 - Designed for long life
 - Designed for concurrent replacement or upgrade
- Security Characteristics
 - Data encrypted on the flash express adapter with 128 bit AES encryption
 - Keys stored on smart cards plugged into the System z SE
 - Removal of smart cards renders data unusable



One Flash Express Card



zFlash Express Support

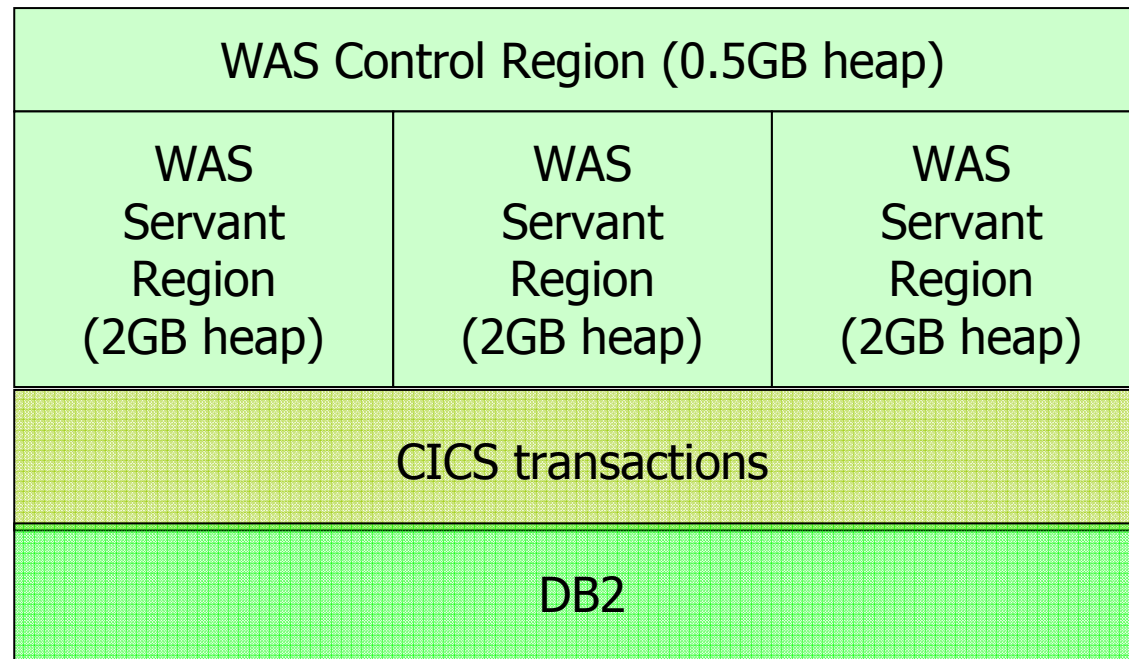
- Optional priced feature on zEC12 machine
- z/OS support:
 - V1.13 RSM Enablement Offering (FMID JBB778H) available 12/14/2012
 - V2.1: Planned GA 09/30/2013
 - IEASYSxx: PAGESCM=ALL | NONE (Default=ALL)
- Exploitation:
 - RSM using zFlash for paging:
 - 4KB pages and 1MB pageable large pages
 - Required for use of 1MB pageable large pages
 - 1MB pageable large page exploitation:
 - LE in 31-bit mode for stack and heap
 - Java 6.0.1 (626) SR4 and Java SDK 7 SR3
 - Java heap and codecache 31-bit and 64-bit
 - IMS V12 with apar PM66866
 - CQS Interface Buffers (V1.13 and V2.1)
 - CQSPUT buffers (V2.1 only)
 - DB2 11
 - Buffer pools and buffer pool control blocks

zFlash Exploitation – Test Environment

- zEC12 LPAR with:
 - 8 CPs
 - 36 GB of real storage
 - 128 GB of SCM storage
- DASD: DS8800 Model 2107-951 with 60GB cache for page data sets
 - the DS8800 sub was dedicated to the LPAR with no other I/O activity than the paging I/O
 - cache hit rates 95% - 100%
 - 16 local page data sets spread across 8 LCUs
- z/OS V1.13 with RSM Enablement Offering FMID JBB778H
- Workload:
 - Websphere front-end to a CICS/DB2 OLTP workload
- Test scenarios:
 - Workload transition
 - SVC dump
 - Stand-Alone Dump

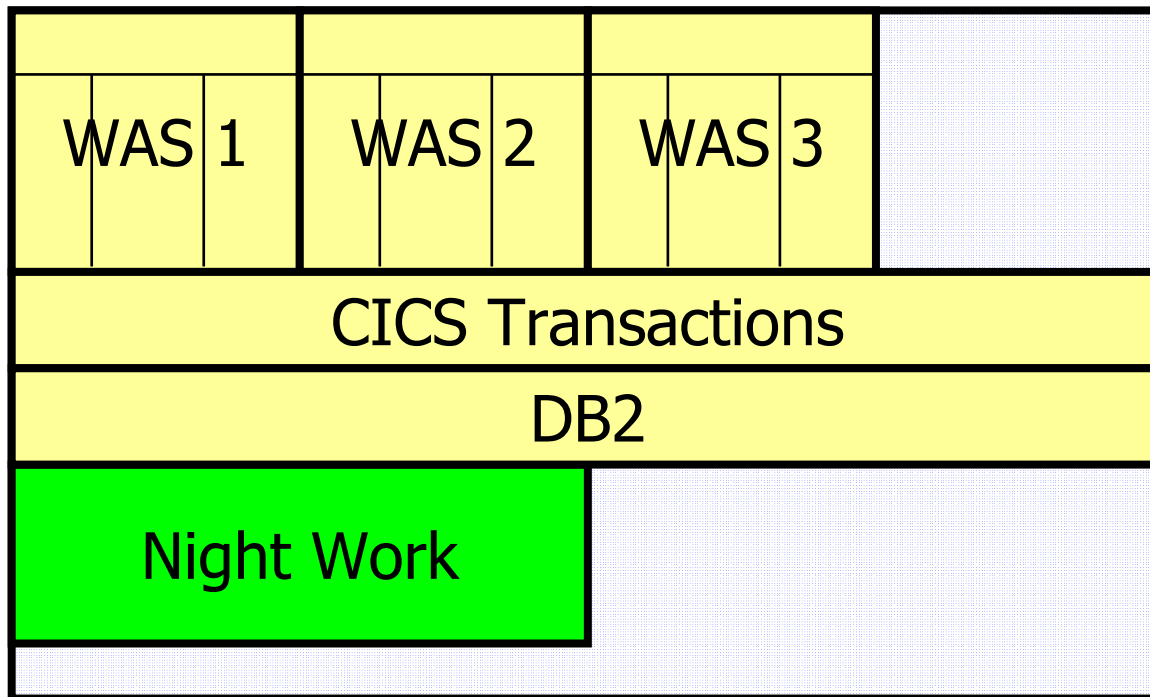
zFlash Workload -Websphere accessing CICS and DB2

- Each WAS instance has a WAS Control Region and 3 WAS Servant Regions.
- Each WAS Control Region has a 0.5GB heap plus a JIT Code cache.
- Each WAS Servant Region has a 2GB heap plus a JIT Code Cache.
- 500 clients per WAS Control Region.



zFlash – Workload Transition

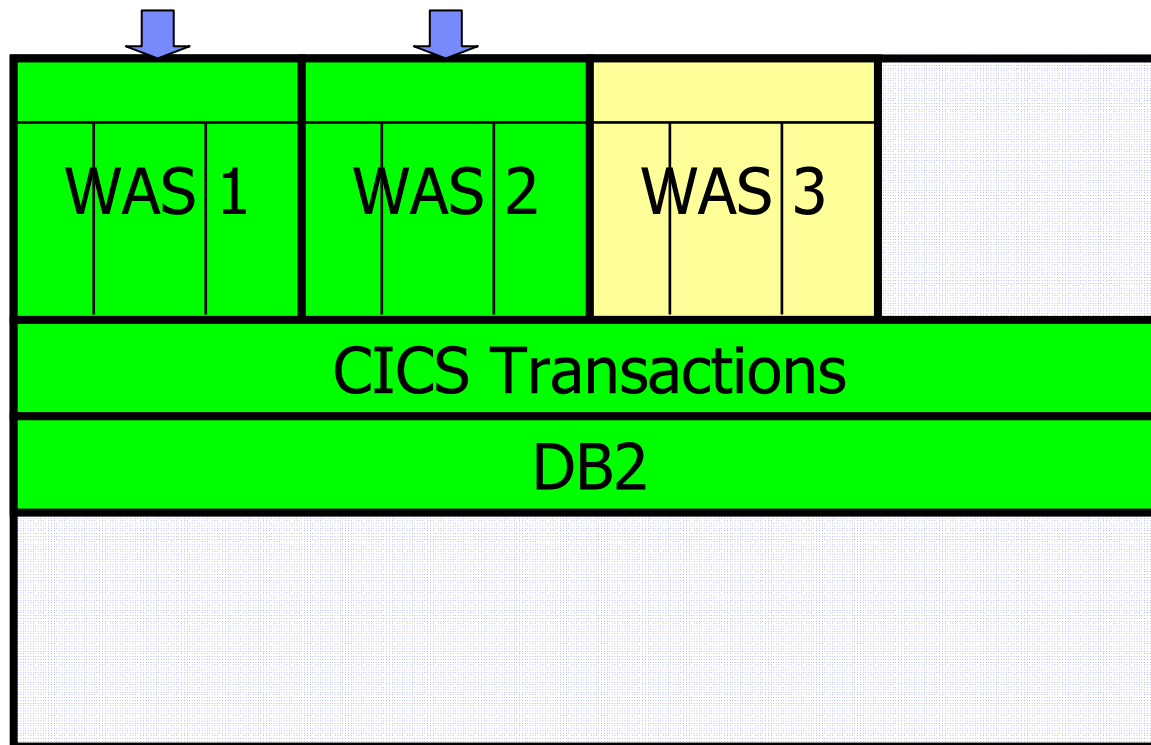
- WAS workload to CICS and DB2 represents OLTP work which is stopped
- Simulated overnight work consumes real storage pushing other pages out



~21GB is pushed out to Auxiliary storage

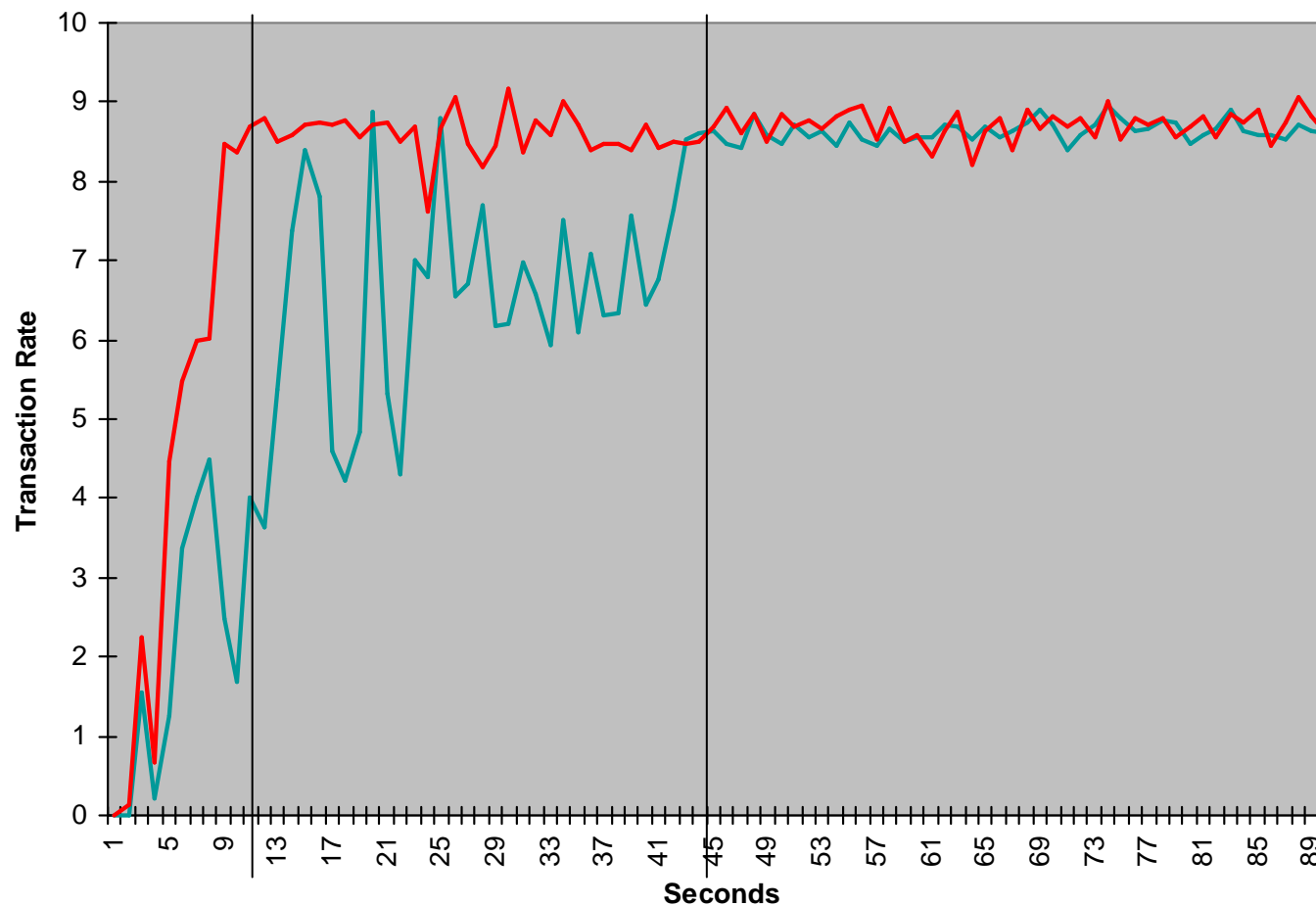
zFlash Workload Transition – Night batch transition to OLTP

- The “Night Work” is stopped and OLTP work is started (WAS 1 and WAS 2)
- Measure the time needed to bring the OLTP work to full speed.



~14GB is
paged back
in from
Auxiliary
storage

zFlash Workload Transition Results



Paging to DASD, it took about 44 seconds for the workload to reach full speed

Tran/sec DASD
Tran/sec Flash

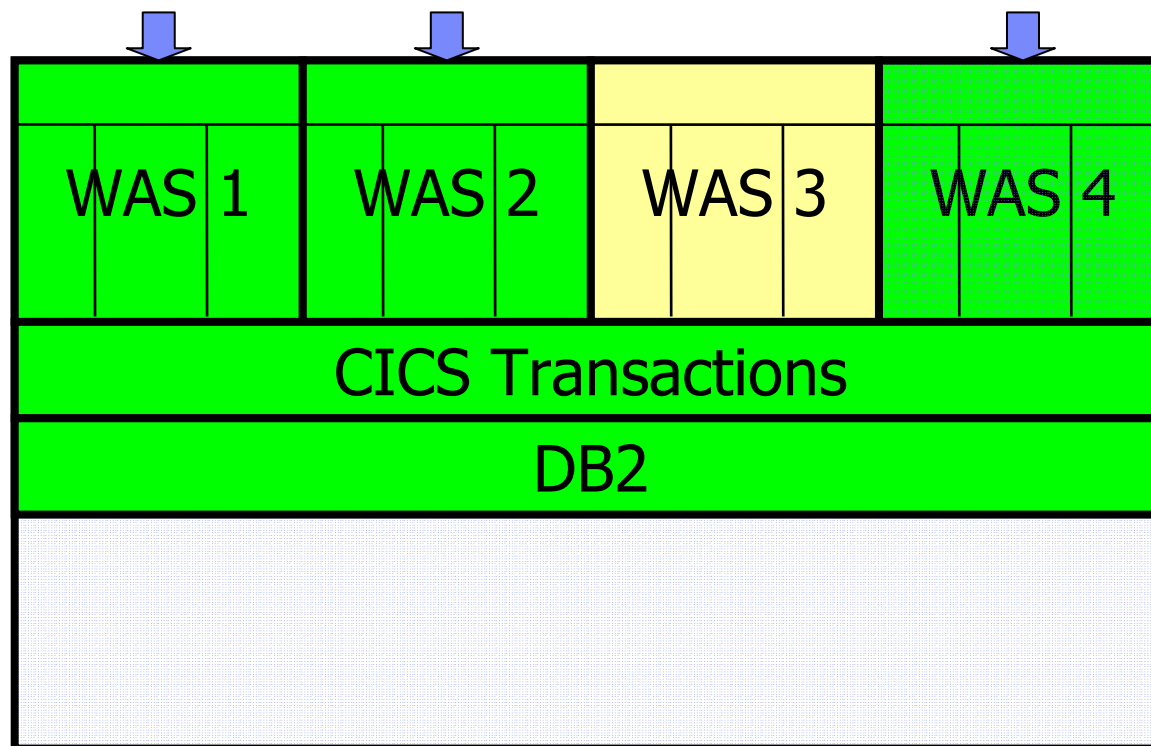
Paging to Flash, it took only 10 seconds for the workload to reach full speed

zFlash Workload Transition - Results

	DASD	Flash	Improvement
Total Transactions within <u>first 45 seconds</u>	251	343	37% increase
Average response time within <u>first 45 seconds</u>	0.62	0.06	90% reduction

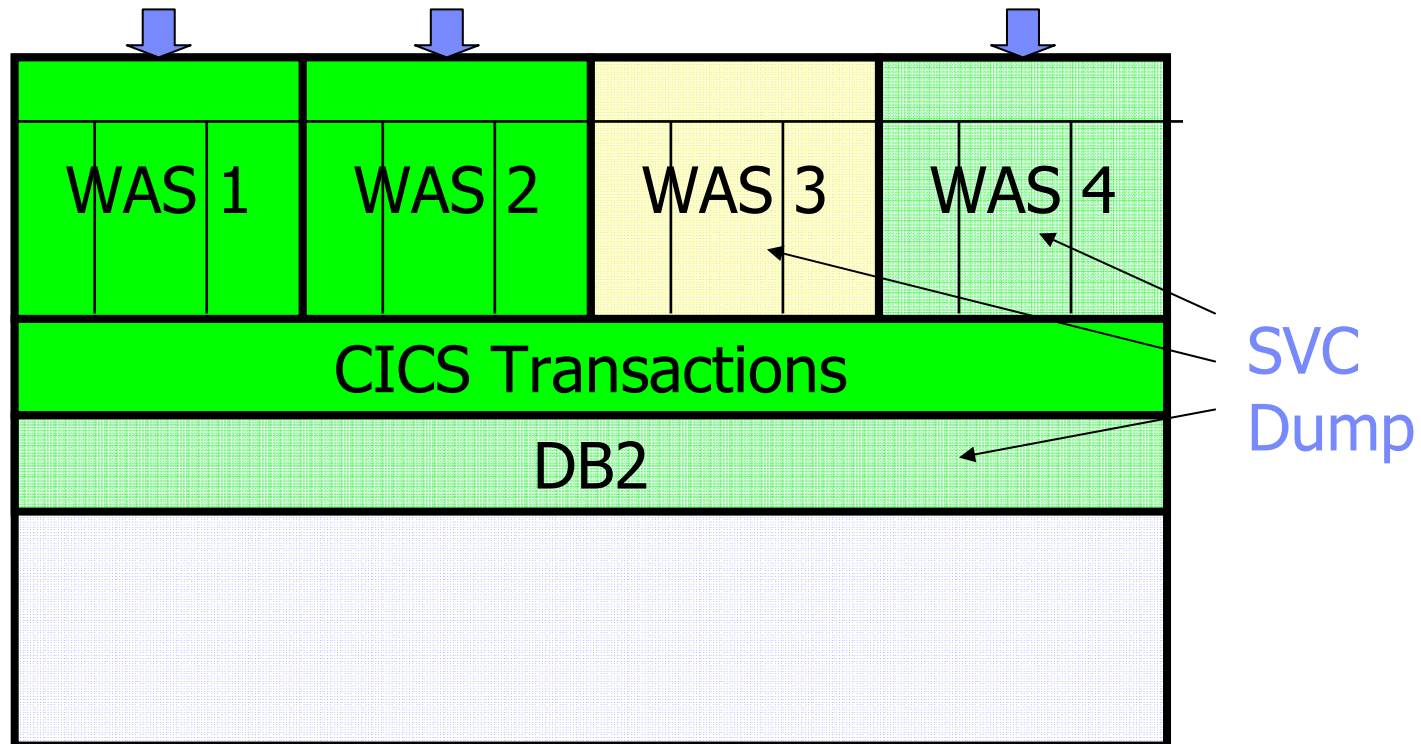
zFlash SVC Dump

- SVC dump with data pages out to Aux
- Three of four WAS instances were active.
- One WAS instance was stopped and most pages were paged out.



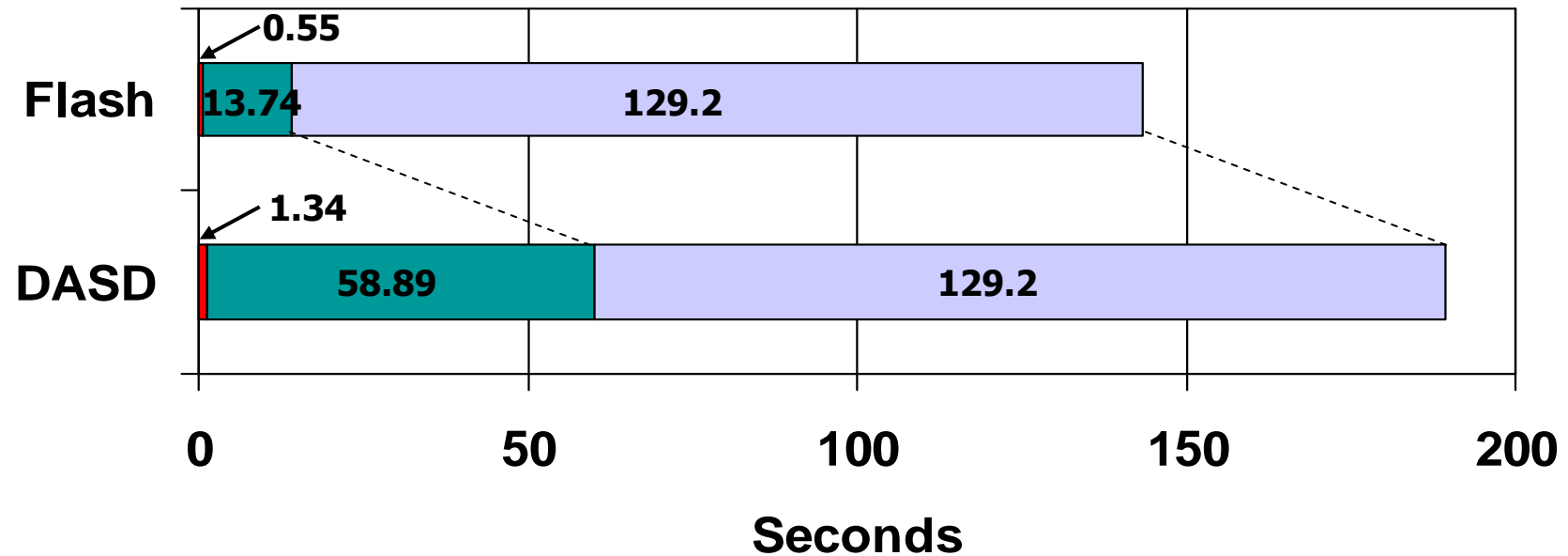
zFlash SVC Dump

- Capture an SVC dump of WAS instance 3 and 4, and DB2.
- Measure the capture time for the SVC dump.
- Size of SVC dump: 17 GB



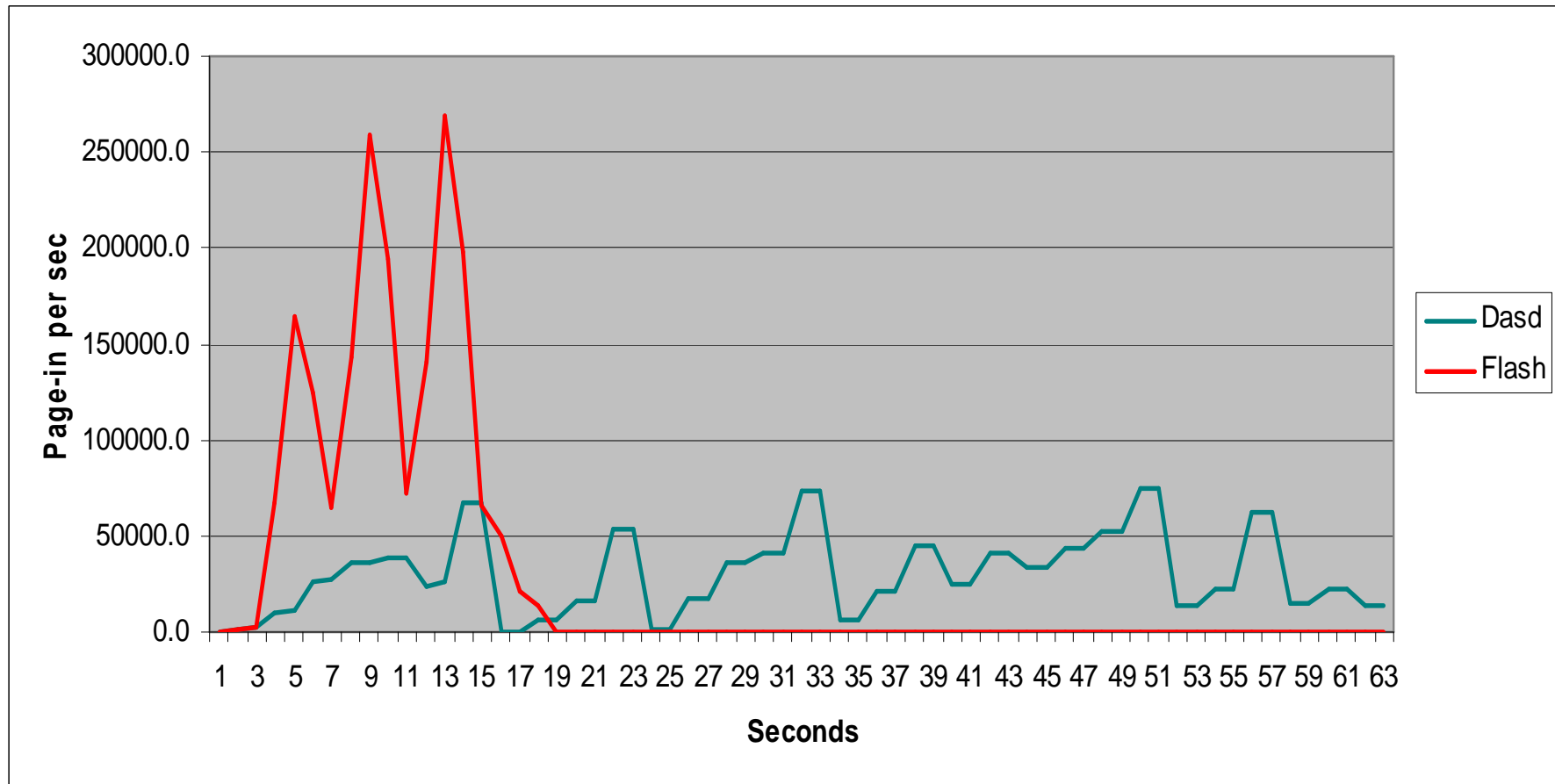
zFlash SVC Dump - Results

- System non-dispatchable time reduced by 59%
- Max address space non-dispatchable time reduced by 77%



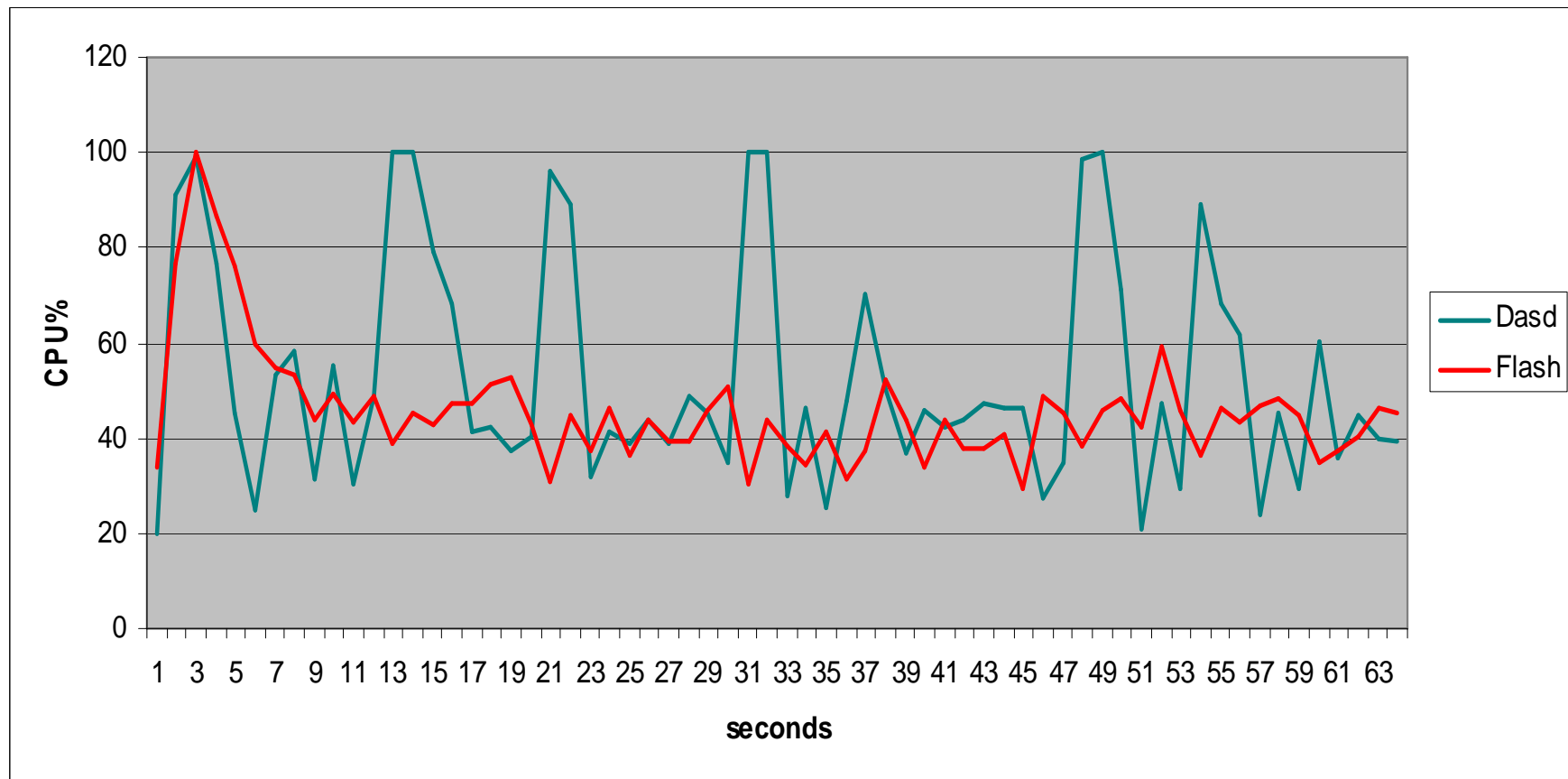
zFlash SVC Dump – Page-in Rate

- Peak page-in rate with Dasd: 75,000 pages per sec
- Peak page-in rate with SCM: 260,000 pages per sec



zFlash SVC Dump – CPU Usage

- CPU peaks correspond to peaks in page-in rates
- Several peaks when using DASD while one peak when using SCM



zFlash SVC Dump - RMF Page Data Set Report Example

- RMF Page Data Set report: average over 6 minutes

```

                                P A G E   D A T A   S E T   A C T I V I T Y

z/OS V1R13                      SYSTEM ID P41                      DATE 10/09/2012          INTERVAL 05.59.585
                                RPT VERSION V1R13 RMF              TIME 14.30.28          CYCLE 0.050 SECONDS

NUMBER OF SAMPLES =    7,190                      PAGE DATA SET AND SCM USAGE
                                -----
PAGE                               %   PAGE                               V
SPACE  VOLUME  DEV  DEVICE          SLOTS  ---- SLOTS USED  ---  BAD          IN  TRANS  NUMBER  PAGES  I
TYPE   SERIAL  NUM  TYPE           ALLOC   MIN   MAX   AVG   SLOTS     USE  TIME   IO REQ  XFER'D O  DATA SET NAME
PLPA   41PAG0  5473 33903          98999  14655 14655 14655     0   0.00  0.000    0      0   SYS1.P41.PLPA
COMMON 41PAG0  5473 33903           89999    61    61    61     0   0.00  0.000    2     32  SYS1.P41.COMMON
LOCAL  41PAG0  5473 33903          410399    0     0     0     0   0.00  0.000    0      0 Y  SYS1.P41.LOCAL
SCM    N/A     N/A  N/A           33554K 6030K 6108K 6061K     0   4.24  0.000  721516 17.19M  N/A

```

zFlash Stand-Alone Dump

- APAR OA40842 - Improvements in Stand-Alone Dump time when dumping data that are paged out
- Overall 30 sec reduction in dump time due to faster page-in of data on aux

Tests	Total dump time	Paging I/O wait time	Batch read rate MB/sec	Total GB dumped	GB of data from aux
Dasd Page data sets	00:04:05.14	00:00:41.30	438.06	50.9	17.7
Flash for paging	00:03:35.94	00:00:10.38	1612.30	50.6	16.3

z/OS V1.13 1 MB Pageable Large Page Exploitation

- Benefits of large pages:
 - Better performance by decreasing the number of TLB misses that an application incurs
 - Less time spent converting virtual addresses into physical addresses
 - Less real storage used to maintain DAT structures
- Fixed large pages vs pageable large pages:
 - Fixed large pages are backed at allocation. Pageable large pages are backed when referenced.
 - Use of fixed large pages for unauthorized users is controlled by a RACF profile (IARRSM.LRPAGES). No RACF authorization to use pageable large pages.
 - Fixed large pages stay as 1 MB pages while pageable large pages may be demoted to 4K pages in certain situations.
- Performance:
 - Java: performance with pageable 1MB large pages is equivalent to 1MB fixed large pages for java heap: up to 5% ITR impact
 - IMS using pageable large pages: up to 1% system ITR improvement. Expect more with z/OS V2.1.
 - DB2 using pageable large pages: up to 3% system ITR improvement.

z/OS V1.13 Results Summary

- V1.13:
 - Overall CPU performance goal: equivalent to V1.12 (avg. of four LSPR workloads)
 - Performance enhancements for zFS, IEBCOPY, Unicode services, and PKI services

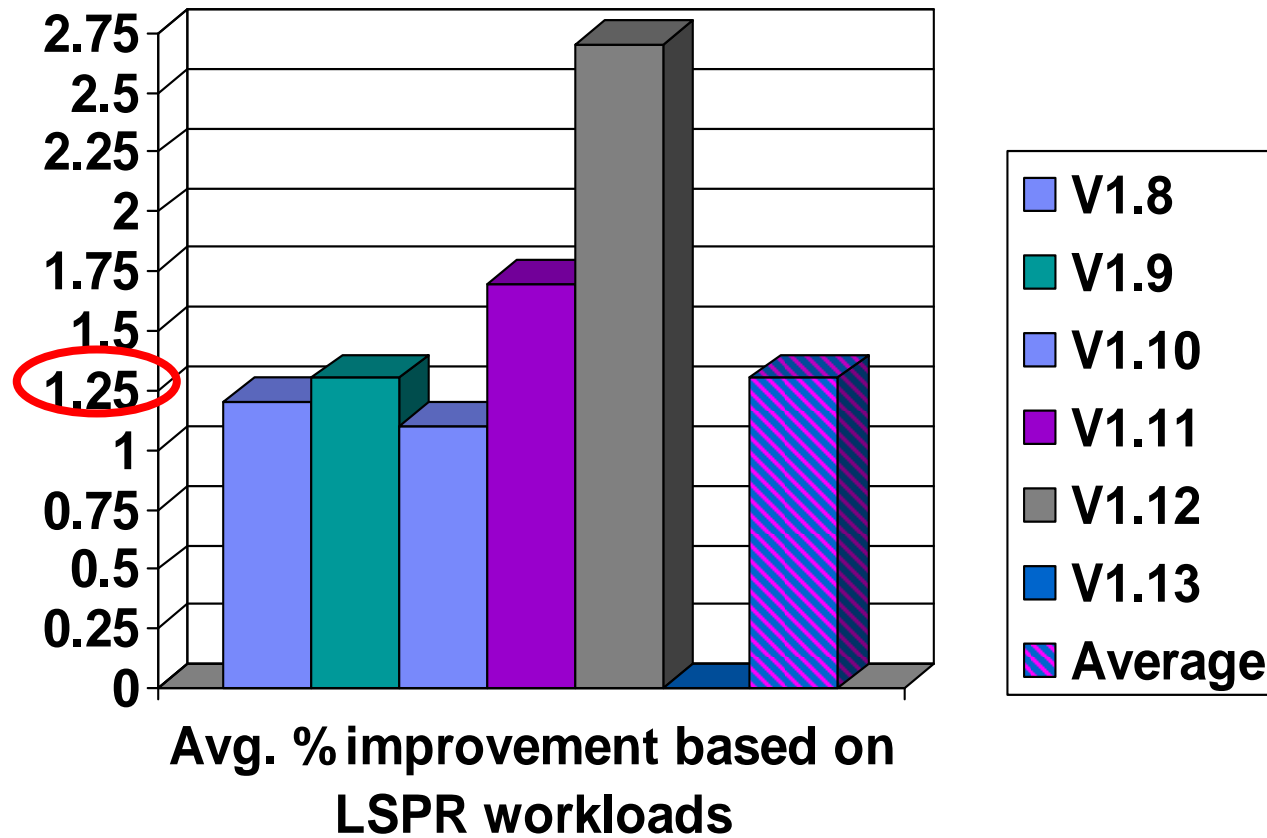
 - VSCR results:
 - SSRB move to 64-bit: benefit depending on workloads
 - Communication Server ctrace move to 64-bit:
 - VIT: saved 4MB of storage below the bar
 - TN3270 ctrace saved about 256MB below the bar

- V1.13 and zEC12 Transactional Execution Facility
 - Potential to reduce impact of lock contention

- V1.13 w/web deliverable FMID JBB778H and zFlash Express on zEC12:
 - zFlash Express: performance improvements for page-in
 - 1 MB pageable large pages: up to 5% performance improvement depending on exploitation

z/OS Release-to-Release Performance Improvements

- Focus on performance improvements:
 - Future z/OS releases
 - z/OS SW stack
- Results V1.8 through V1.13:



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/atmastr.nsf/WebIndex/PRS1762>
- LSPR web site (LSPR workloads):
 - <https://www.ibm.com/servers/resourceink/lib03060.nsf/pages/lsprindex>

Related SHARE sessions San Francisco:

- 13086: zFlash Introduction, Uses and Benefits
 - Tuesday February 5: 1:30-2:30 in Imperial A (Elpida Tzortzatos)
- 13057: zFlash Setup, Management and Configuration
 - Tuesday February 5: 3:00-4:00 in Imperial A (Elpida Tzortzatos)
- 12917: Experiences with IBM zEC12 Flash Express and IBM zAware Features
 - Wednesday February 6: 8:00-9:00 in Yosemite (Mary Astley)
- 12353: Running IBM Java on zEC12
 - Thursday February 7: 3:00 - 4:00 in Grand Ballroom B (Marcel Mitran)

System z Social Media Channels

- Top Facebook pages related to System z:

- [IBM System z](#)
- [IBM Academic Initiative System z](#)
- [IBM Master the Mainframe Contest](#)
- [IBM Destination z](#)
- [Millennial Mainframer](#)
- [IBM Smarter Computing](#)

- Top LinkedIn groups related to System z:

- [System z Advocates](#)
- [SAP on System z](#)
- [IBM Mainframe- Unofficial Group](#)
- [IBM System z Events](#)
- [Mainframe Experts Network](#)
- [System z Linux](#)
- [Enterprise Systems](#)
- [Mainframe Security Gurus](#)

- Twitter profiles related to System z:

- [IBM System z](#)
- [IBM System z Events](#)
- [IBM DB2 on System z](#)
- [Millennial Mainframer](#)
- [Destination z](#)
- [IBM Smarter Computing](#)

- YouTube accounts related to System z:

- [IBM System z](#)
- [Destination z](#)
- [IBM Smarter Computing](#)

- Top System z blogs to check out:

- [Mainframe Insights](#)
- [Smarter Computing](#)
- [Millennial Mainframer](#)
- [Mainframe & Hybrid Computing](#)
- [The Mainframe Blog](#)
- [Mainframe Watch Belgium](#)
- [Mainframe Update](#)
- [Enterprise Systems Media Blog](#)
- [Dancing Dinosaur](#)
- [DB2 for z/OS](#)
- [IBM Destination z](#)
- [DB2utor](#)

