

z/OS 2.2 Performance Enhancements

Session 17910

Carol Creter

IBM



#SHAREorg



SHARE is an independent volunteer-run information technology association that provides **education, professional networking and industry influence.**

Copyright (c) 2015 by SHARE Inc.  Except where otherwise noted, this work is licensed under <http://creativecommons.org/licenses/by-nc-sa/3.0/>



Trademarks

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

CICS*	DFSMSshm	FICON*	IBM logo*	PrintWay*	WebSphere*	zArchitecture*
CICS Explorer	DFSMSrmm	FlashCopy*	IMS	REXX	z10 BC	zEnterprise*
DB2*	DFSORT	GDPS*	InfinBand*	RMF	z10 EC	z Systems
DFSMS	DS6000*	HiperSockets	Infoprint*	System z9*	z13	z/OS*
DFSMSdftp	DS8000*	HyperSwap*	Language Environment*	System z10	z/Architecture*	HyperWrite™
DFSMSdss	Easy Tier\	IBM*	Parallel Sysplex*	Tivoli*	zEnterprise*	z/OS*

* 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.

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.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

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.

Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and

Linux is a registered trademark of Linus Torvalds 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.

OpenStack is a trademark of OpenStack LLC. The OpenStack trademark policy is available on the [OpenStack website](#).

TEALEAF is a registered trademark of Tealeaf, an IBM Company.

Windows Server and the Windows logo are trademarks of the Microsoft group of countries.

Worklight is a trademark or registered trademark of Worklight, an IBM Company.

UNIX is a registered trademark of The Open Group in the United States and other countries.

* Other product and service names might be trademarks of IBM or other companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any 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.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the 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.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This information provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g. zIIPs, zAAPs, and IFLs) ("SEs"). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT"). No other workload processing is authorized for execution on an SE. IBM offers SE at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.



Agenda

- zEDC enablement for DFSMSdss and HSM
- Batch Storage Usage
- Tamper Resistant SMF
- SRB Dispatching Priority
- RSM Support for Large Storage
- Standalone Dump Enhancements
- New WLM Support for DB2 Buffer Tuning
- JES Support Larger Workloads
- OPT Members to support 4TB
- VSAM Locking
- WLM Health Based Routing
- Miscellaneous Performance

Complete your session evaluations online at www.SHARE.org/Orlando-Eval

z/OS V2R2 Preview*



Usability and Skills

z/OSMF as a base element of z/OS; TCP/IP configuration; z/OSMF plug-in setup workflow; Updates to WLM, RMF, Incident Log, Software Management, WebISPF applications; New z/OSMF External Applications API; DJC and Deadline Scheduling for JES2; System Symbol enhancements...

Application Development

Web Enablement Toolkit, EU ordering rules for Unicode, ISPF improvements, DFSORT Date Functions, Enhanced RESTful data set and file APIs, Parallel Batch Scheduling, Improved JES3 symbol and JCL support, ...

Scalability & Performance

More threads for z/OS UNIX® System Services, AMODE64 File System Services for zFS & NFS, CA-Level Locking for RLS, zFS performance, Even More Jobs for JES2, ...



Enhancing Security

Signed SMF records, RFC 4556 X.509 support in Kerberos, RRSF Dynamic Node Reassignment, Multiple certificate approvers, PKI RFC 6277 Support, System SSL RFC 2560 OCSP Support, z/OS UNIX security improvements, BCPii audit records, ...

Availability

Dynamic JES2 Checkpoint Tuning & Expansion, Private Area Virtual Storage Tracking in PFA, Dynamic TDS (LDAP) Compatibility Upgrades, Multi-target PPRC, Incremental FlashCopy, XCF message processing, LOGREC deallocation, O/C/EOV Dynamic Exits, ...

Systems Management

Smarter Subsystem Interface processing, DFSMSHsm Storage Tiers Extensions, Health-Based Workload Routing, RMF Reporting Enhancements, Generic Tracker Improvements, ...

Networking

64-bit TCP/IP Stack, RoCE Improvements, DVIPA Limit, CICS Sockets, Enterprise Extender Scalability, NIST SP800-131a, TLS Session Reuse, Resolver Improvements, ...

Complete your session evaluations online at www.SHARE.org/Orlando-Eval



* Statements regarding IBM future direction and intent are subject to change or withdrawal, and represent goals and objectives only.

zEDC enablement for DFSMSdss and HSM

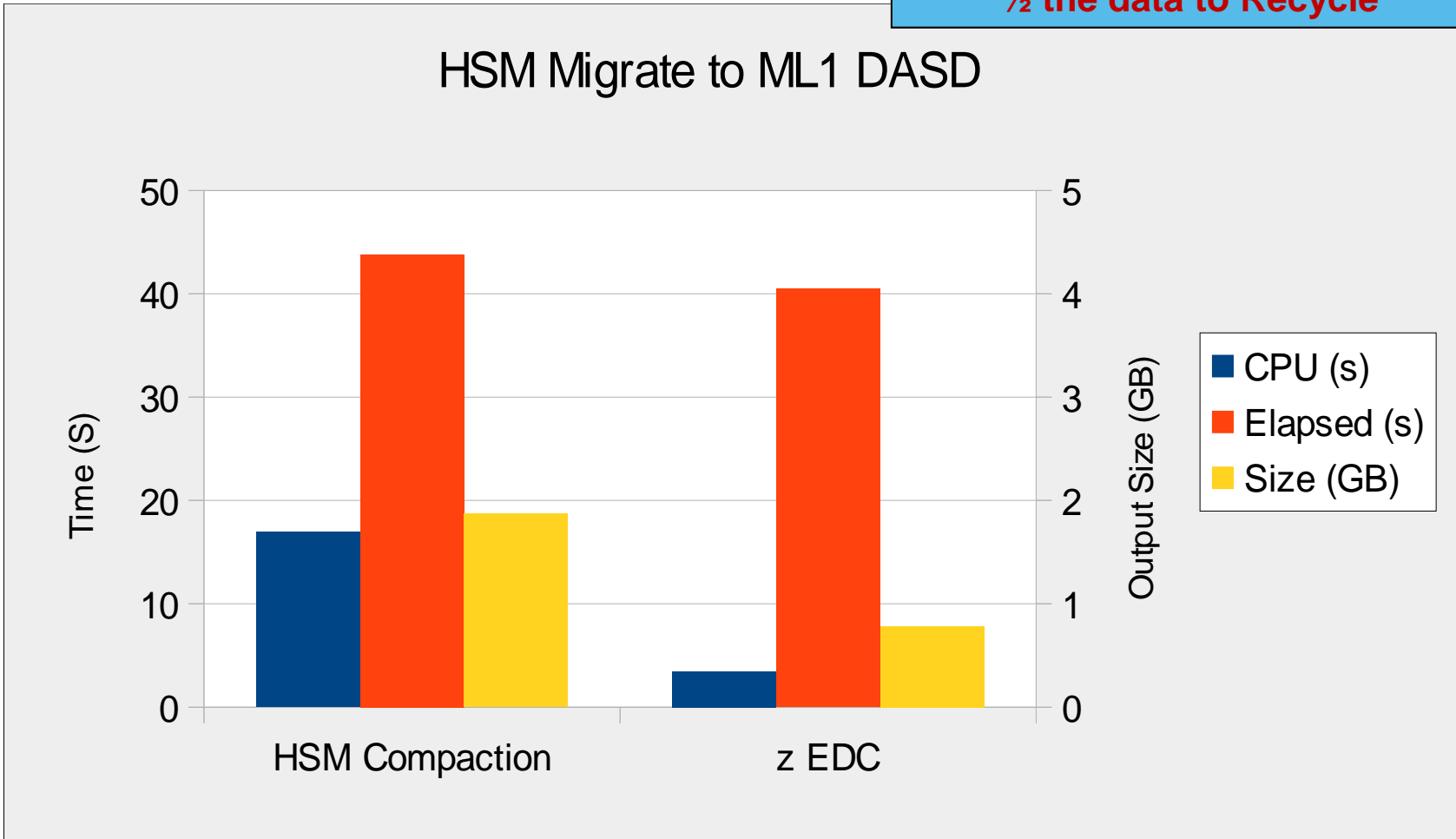


- Enable z/OS V2R1 DFSMSdss and hsm to utilize zEDC
 - Reduce MIPS
 - Improve compression ratio
 - Improve throughput
- DFSMSHsm can use zEDC, when DSS is the datamover, for the following functions:
 - Migrate / Recall
 - Backup / Recover
 - Dump / Restore
- PDS data sets are moved entirely by HSM, so are not yet supported
- When specified, DFSMSHsm will invoke DSS with the zCOMPRESS(PREFERRED) option

zEDC Measurement

Up to 80% CPU Reduction
Up to 53% Space Reduction

1/2 the data for SSM
1/2 the data to Recycle

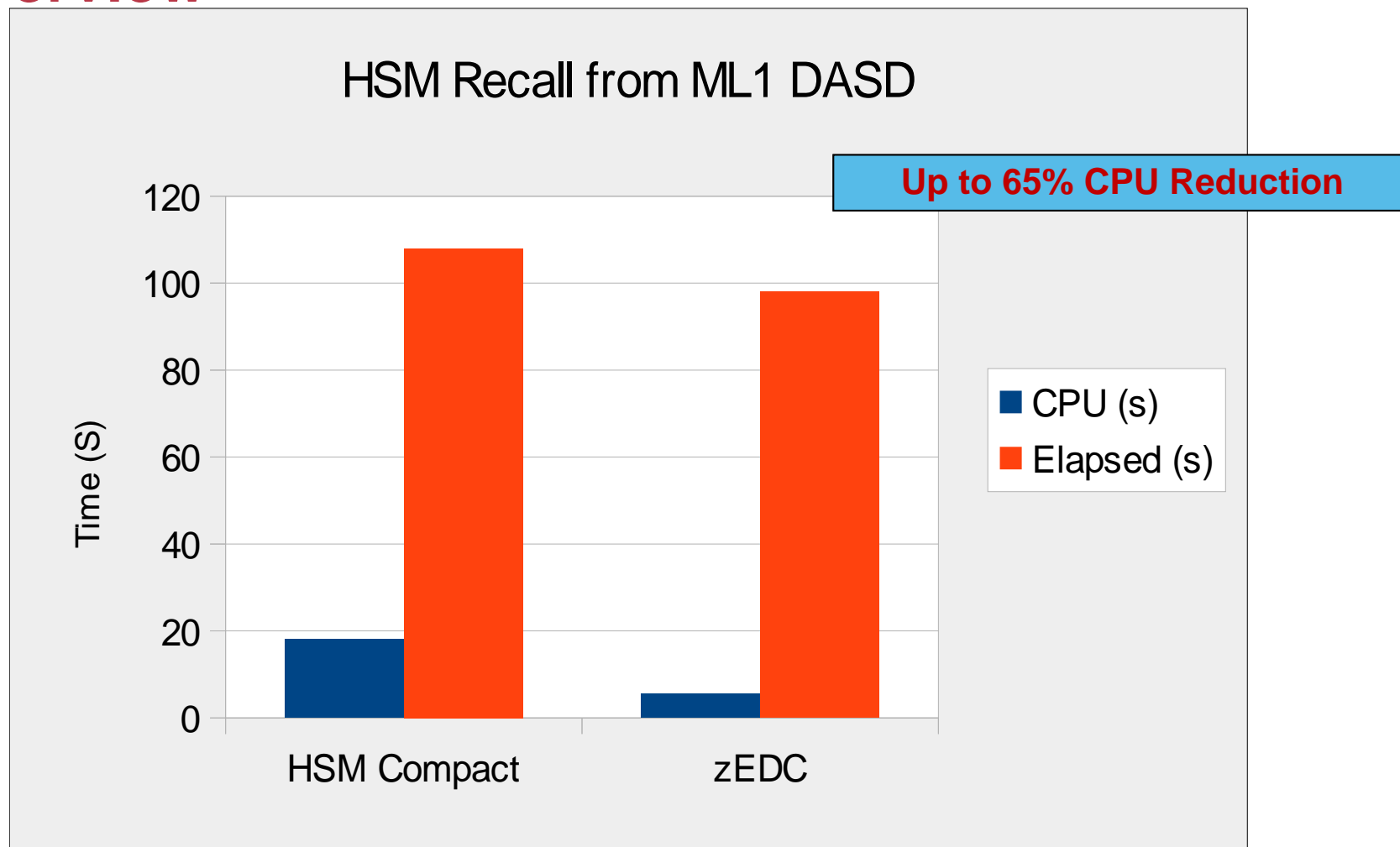


* Disclaimer: Based on projections and/or measurements completed in a controlled environment. Results may vary by customer based on individual workload, configuration and software levels.

Complete your session evaluations online at www.SHARE.org/Orlando-Eval



Overview



* 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.
Complete your session evaluations online at www.SHARE.org/Orlando-Eval

Batch Storage Usage

- Allows users more control in amount of below-the-line and above-the-line storage that must be available to process jobs
- Provide a new **REGIONX=** JCL keyword that supports *two* storage specifications and allows the end-user to explicitly specify values for their below-the-line storage and above-the-line storage needs..
- Syntax: **REGIONX=(below,above)**
 - *below* is in **K** or **M** units, less than 16384K or 16M
 - *above* is in **K**, **M**, or **G** units, less than 2G
- Support for the new **REGIONX** keyword is provided:
 - On the **JOB** statement
 - Including when a job is started by a START command
 - On the **EXEC** statement
 - Including procedure overrides
 - Including when a procedure is started by a START command

Batch Storage Usage (cont..)

- Region Size

CODED	BELOW	ABOVE
REGIONX=	DEFAULT	DEFAULT
REGIONX=(,)	DEFAULT	DEFAULT
REGIONX=1M	1M	DEFAULT
REGIONX=(1M,)	1M	DEFAULT
REGIONX=(,1G)	DEFAULT	1G

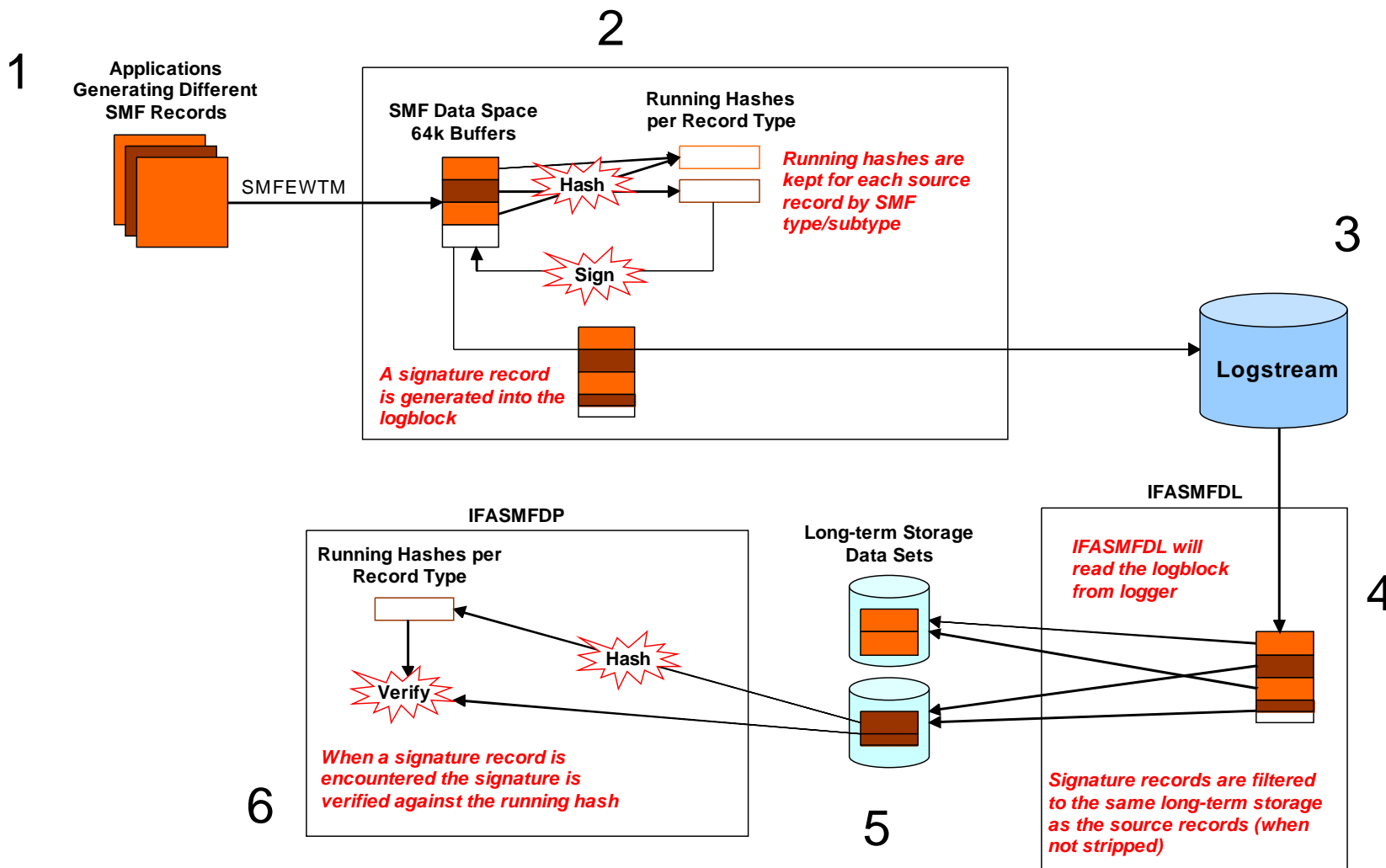
- Where Region Size Obtained

JOB	EXEC	RESULT
REGIONX	REGION	ABEND
REGION	REGIONX	ABEND
REGIONX	REGIONX	EXEC
REGION	REGION	JOB

Tamper Resistant SMF

- Increase the value of SMF data by making it verifiable using industry standard encryption
- Applications recording to SMF can transparently leverage this support
- The SMF data is signed on the way to System Logger
 - As each record is written to the logstream it is hashed
 - Running hash maintained per unique SMF type/subtype
 - Periodically, the hash will be digitally signed and that signature data will be recorded to the logstream as a signature record
 - On the global interval a signature is created for all data hashed during the interval and recorded to the logstream
 - These operations are performed with the private key

Tamper Resistant SMF (cont.)



Tamper Resistant SMF (cont.)

- Implementation
 - The first step is create a public/private key pair via ICSF
 - Update the SMF configuration to sign record SMFPRMxx
 - New option *RECSIGN* can be specified globally or per *LSNAME*
 - Default is *NORECSIGN*
 - IFASMFDL can carry signature data with the SMF records
 - By default IFASMFDL will drop signature records
 - The *NOSIGSTRIP* option can be used to have signature records written to OUTDD data sets
 - IFASMFDP can carry signature records and perform validation
 - New IFASMFDP parameters *SIGSTRIP* and *SIGVALIDATE*
 - *SIGSTRIP* behaves the same as with IFASMFDL
 - *SIGVALIDATE* indicates that signature validation is to be performed
 - Default: *NOSIGVALIDATE* (don't perform validation)

Tamper Resistant SMF (cont.)

Dissecting an IFASMFDP SIGVALIDATE report

RECORD VALIDATION REPORT FOR SY1								
RECORD TYPE	RECORD SUBTYPE	VALIDATION FAILURE	VALIDATION START DATE-TIME	VALIDATION END DATE-TIME	RECORDS VALIDATED	GROUPS VALIDATED	INTERVALS VALIDATED	
128	*	N	10/23/2014-11:00:00	10/23/2014-13:00:00	60	10	2	
145	1	N	10/23/2014-11:00:00	10/23/2014-13:00:00	3	3	2	
160	*	N	10/23/2014-11:00:00	10/23/2014-13:00:00	10	2	2	

VALIDATION SUCCEEDED



Indicates successful validation of this record type and subtype



When all data validates the report ends with this message. On a failure this would provide additional information

SRB Dispatching Priority

- The target SRB of a schedule could be more important to the system than the scheduler but currently is just thrown into the mix of available work units.
 - Allow the SRB scheduler to indicate that the target SRB is to 'take over' the current dispatch
 - Improved responsiveness (when exploited)
- IEAMSCHD provides new optional keyword:
 - TRANSFER=NO | YES
 - **NO** (the default) specifies the SRB will be dispatched normally
 - **YES** specifies an attempt to transfer control to the SRB if the requestor is running enabled and in task mode.
 - If the transfer is successful the requestor will be redispached normally.
 - TRANSFER=YES is ignored for the following cases a SYNCH=YES request running in SRB mode the caller holds any locks

RSM Support for Large Storage

- The changes to RSM enable the use of more memory on each system which can in turn improve performance
- Substantial latency reduction for OLTP workloads
 - Significant response time reductions and increased transaction rates
- In-memory databases dramatic gains in response time by avoiding IO wait
- Batch window reduction
 - More concurrent workloads and shorter elapsed times for jobs
- Reduced time to insight for analytic workloads
 - Processes data more efficiently helping organizations keep pace with influx of data
 - Reduces time it takes to get from raw data to business insight

RSM (cont.)

- Real Storage Management (RSM) has made extensive internal changes to frame management and serialization.
 - These changes enable large memory usage (up to 4TB on z13), better performance and increased parallelism.
- RSM is modified to support large amounts of real storage:
 - Initialization of real storage is more efficient
 - Management of 1MB frames is improved
 - Improvements to configuring storage online and offline
 - Steal processing can find eligible frames more efficiently
- RSM is modified to increase concurrency both at the system and application levels
 - Multiple page faults can be dealt with concurrently within an address space or with common storage
 - Page fixing and unfixing can occur concurrently within an address space or with common storage
 - Getmain/Freemain of storage can occur concurrently with page fixing/freeing and page faults
 - Less contention on available frame queues

RSM Migration & Coexistence Considerations

- Change in default for LFAREA INCLUDE1MAFC
 - With APAR OA41968, a new keyword, INCLUDE1MAFC, was added to the LFAREA parameter in the IEASYSxx parmlib member
 - When specified 1MB large frames (as defined in the LFAREA system parameter) are included in the available frame count (RCEAFC) and also included in results of SYSEVENT STGTEST
 - Therefore RSM performs less paging when there is an abundance of available fixed 1MB pages in the system
 - The default at V2R1 is not to include 1MB frames in the available frame count
 - At V2R2 the default changes to include 1MB frames. This is equivalent to specifying LFAREA=(INCLUDE1MAFC,...)
 - To retain the previous behavior specify LFAREA=(INCLUDE1MAFC(NO),...)
- Related support rolled down to V2R1 via OA44207 and OA44436.
 - This support is included in V2R2 base

Standalone Dump Enhancements

- z/OS systems can be up to 4TB resulting in larger dumps
- Need to leave more central storage available for dumping
- New option REAL= parameter for stand-alone dump
 - Default of REAL=ALL dumps all real storage
 - REAL=USED dumps only used real storage skipping any unused (AKA available) real frames
 - Can be set with REAL=USED at generation of ADSADMP or at runtime with SET REAL(USED)

New WLM Support for DB2 buffer tuning

- Helps optimizing response times and reduce CPU consumption
- New criteria to choose which bufferpools to adjust
 - If multiple bufferpools qualify to be shrunk, chose one used by the least important periods first
 - Consider the fact that DB2 does not necessarily use up the complete size of a bufferpool as recommended by WLM
- Implementation
 - ALTER BUFFERPOOL AUTOSIZE(YES[NO])
 - Causes DB2 registers bufferpool to WLM when YES
 - WLM will recommend to grow the size of the bufferpool when the Performance Index of a Service Class Period is impacted and bufferpool delays are a significant contributor
 - WLM will recommend to shrink the size of the bufferpool due to donation to a suffering Service Class Period

JES Support for Larger Workloads

- JES2 manages the setting of the values for checkpoint
 - Adjusted based on workload, MAS size, etc.
 - Tuning MASDEF HOLD and DORMANCY automatically
 - Simplified operations and better performance
 - To enable automatic checkpoint cycle management:
`$T MASDEF,CYCLEMGT=AUTO`
- Deal with performance problems related to larger number of jobs by updating limits on job, JOE, and BERT numbers
 - JOBNUM limit now 1,000,000
 - JOENUM limit now 2,500,000
 - BERTNUM limit now 2,500,000
 - Increased BLOB size
 - Cache of available track groups increased from 256 to 1024
- New command to display duplicate job families
 - `$D DUPJOB`

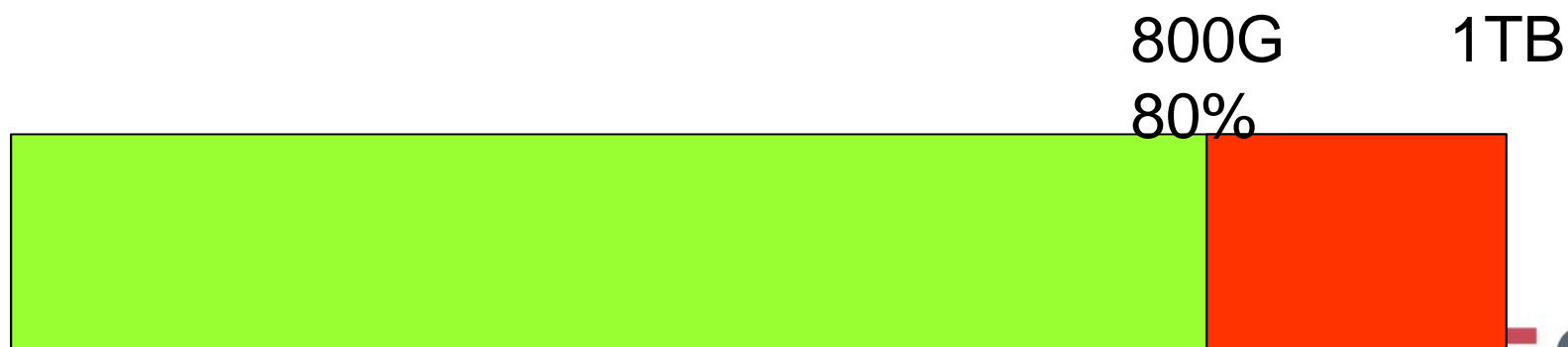
OPT Parameters to support 4TB

- z/OS now supports up to 4TB
- OPT parameters need to be updated to support large targets for large LPARs
 - MCCFXTPR
 - RCCFXTT
 - RCCFXET
- Changes the OPT parameter calculations
 - Simpler to set storage targets with the OPT parameter
 - Run more workload in parallel

Usage & Invocation – OPT Parameter MCCFXTPR

- The *MCCFXTPR* keyword in the IEAOPTxx parmlib member specifies the percentage of online storage that may be page fixed before a pageable storage shortage is detected and message IRA400E is issued.

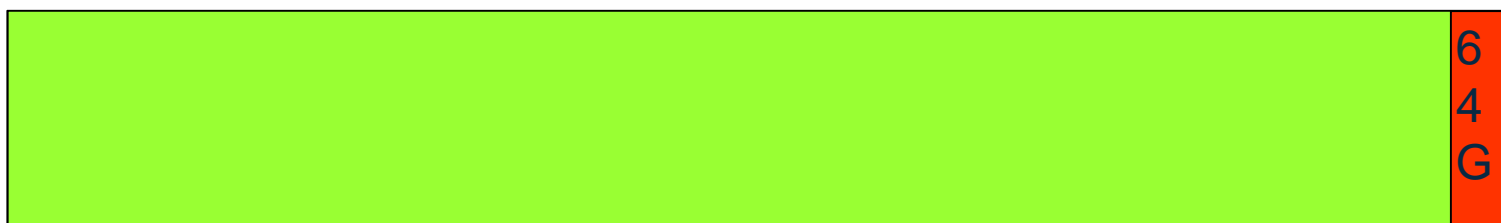
The current *MCCFXTPR* default of 80% requires that 20% (100 minus *MCCFXTPR*) of storage remain pageable regardless of the amount of online storage. On systems with large amounts of central storage, the *MCCFXTPR* default of 80% can result in a pageable storage shortage being detected, when there is still plenty of pageable storage.



Usage & Invocation – OPT Parameter MCCFXTPR ...

- In zOS 2.2 the required pageable storage for large systems (>320G) is always 64G. This allows to fix much more storage without running into a pageable storage shortage

1TB

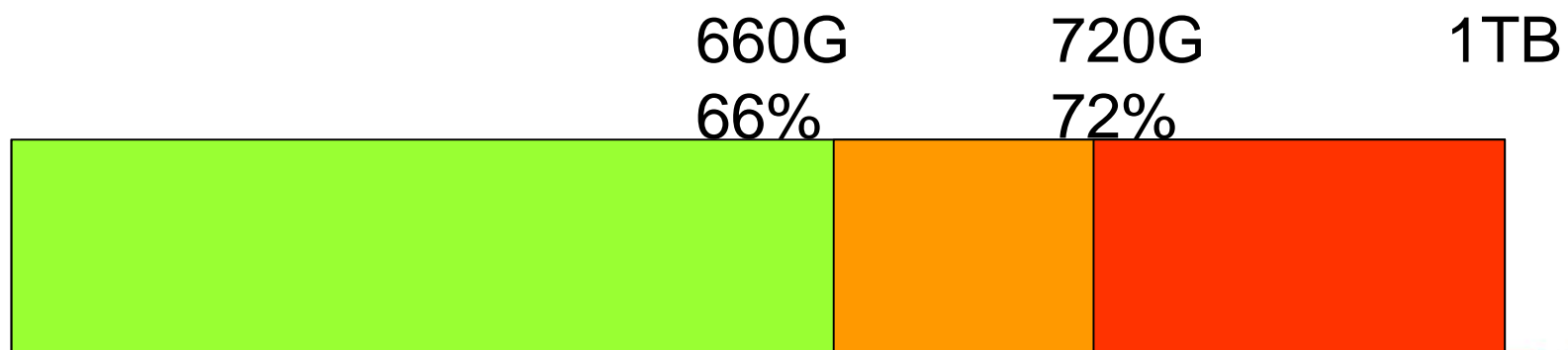


- Rollback to zOS 2.1 is available: OA44668
- Recommendation let default

Usage & Invocation – OPT Parameter RCCFXTT



- The *RCCFXTT* keyword in the IEAOPTxx parmlib member specifies the low and high threshold of fixed real storage. SRM uses these thresholds to determine if the system MPL needs to be increased/decreased. The default is 66% and 72%.
- On small systems such percentages are not a problem.
- On a 1T machine these percentages mean, when 660G of storage is fixed the system is not longer under utilized and WLM will stop increasing the MPL.



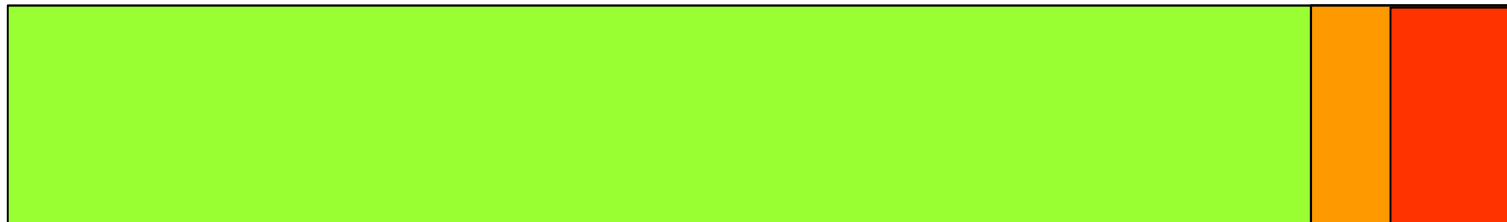
Complete your session evaluations online at www.SHARE.org/Orlando-Eval



Usage & Invocation – OPT Parameter RCCFXTT ...

- In zOS 2.2 SRM is able to calculate the optimal RCCFXTT lower and upper threshold
- The **RCCFXTT=AUTO** definition activates the threshold calculation by SRM
 - The default is still: $RCCFXTT=(66, 72)$
 - Recommendation $RCCFXTT=AUTO$

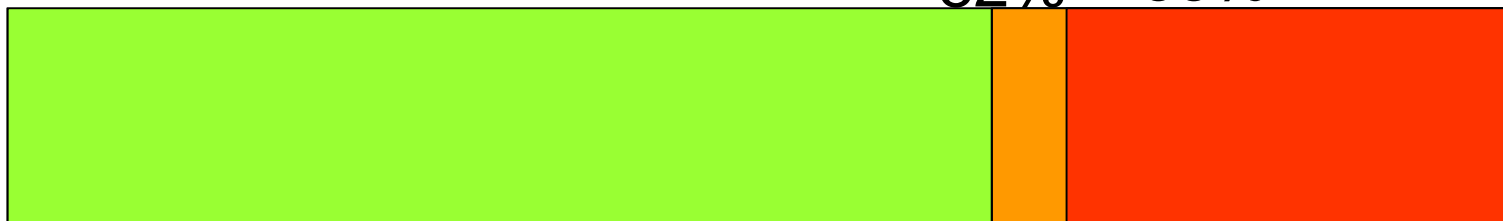
1TB



Usage & Invocation – OPT Parameter RCCEXTT

- The *RCCEXTT* keyword in the IEAOPTxx parmlib member specifies the low and high threshold of fixed real storage below 16M. SRM uses these thresholds to determine if the system MPL needs to be increased/decreased. The default is 82% and 88%.
- This OPT keyword is also enhanced, mainly to stay consistent with the *RCCFXTT* keyword. The ***RCCEXTT=AUTO*** definition activates the threshold calculation by SRM
 - The default is still: *RCCEXTT=(82, 88)*
 - Recommendation *RCCEXTT=AUTO*

	13,1M	14.4M	16M
	82%	88%	



VSAM locking

- CA(Control Area)-Level Locking for RLS
 - Today an entire data set's index is locked for a number of operations
 - ✓ Notably CI splits, CI reclaims, spanned-record processing
 - z/OS V2R2 is designed to lock the index at the CA level
 - For all KSDS and RRDS (including AIXes and Catalogs)
 - CA split and reclaim still need the data set level lock
 - Expected to improve performance and make much larger data sets practical with high update activity

WLM Health Based Routing

- Health-Based Workload Routing Enhancements
 - WLM “anti-storm drain” infrastructure planned to support:
 - Health weight adjustment based on “finger-pointing”
 - WLM differentiates between health values of the address space reported by itself or reported by another space
 - The algorithm for determining the health indicator for an address space is changed. The value is no longer the last value being reported but the minimum of the values reported by the different callers.
 - An additional function of IWM4HLTH refers to RAS considerations regarding a server's health state. The RESET function restarts setting of a composite health value by specifying an initial value and discarding the values reported by other callers before.
 - Callers of service can specify a reason for cause of change
 - Callers can identify themselves by a subsystem type and subsystem name. WLM uses these parameters to recognize different callers of the service.
 - XCF plans to exploit when a Parallel Sysplex member is not keeping up with its messages
 - RTD to report servers with health values less than 100 (the maximum)

Miscellaneous Performance

- **LLA Improvements**

- Will be designed to make it more likely that certain program objects, such as those compiled using COBOL Version 5 (5655-W32), can be cached by LLA in VLF
- Intended to help to improve performance for such program objects in LLA-managed libraries
- Also, planned to be available for z/OS V1R13 and z/OS V2R1 with the PTF for APAR OA45127

- **SDSF Improvements**

- Some CPU-Intensive SDSF work can be offloaded to a zIIP
- Sorting for “large” panels

Miscellaneous Performance

- **DFSORT support for zHPF**
 - z/OS V2R2 DFSORT is supporting zHPF
 - For SORTIN, SORTOUT and OUTFIL
 - Will provide significant performance benefits where zHPF is available

z/OS Support Summary



Out of service
 Service Extension available
 Service Withdrawal Dates

z/OS	z9 [®] EC z9 BC	z10 EC [™] z10 BC [™]	z196 z114	zBX	zEC12 zBC12	z13	DS8000 [®]	TS1140 TS7700	End of Service	Coexists with z/OS...	Planned Ship Date ²
R11	X	X	X	X	X		X	X	9/12	R13	
R12	X	X	X	X	X	X ³	X	X	9/14 ¹	V2R1	
R13	X	X	X	X	X	X	X	X	9/16 ²	V2R2 ²	
<u>V2R1</u>	X	X	X	X	X	X	X	X	2H18 ²	V2R3 ²	
V2R2 ²		X	X	X	X	X	X	X	2H20 ²	V2R4 ²	2H15 ²

1. Fee-based service extension available
2. All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
3. Fee-based service extension required for support, or for some features

Complete your session evaluations online at www.SHARE.org/Orlando-Eval



IBM z13 Functions and Features



Five hardware models
<p>Up to 141 processors configurable as CPs, zIIPs, IFLs, ICFs or optional SAPs (no zAAPs)</p> <ul style="list-style-type: none"> •100-way on z/OS® V1.12 or V1.13 •Up to 141-way on z/OS V2.1 (non-SMT mode) •Up to 128-way on z/OS V2.1 (SMT mode) <ul style="list-style-type: none"> - max active threads in SMT mode is 213
<p>Up to 10 TB of Redundant Array of Independent Memory (RAIM)</p> <ul style="list-style-type: none"> ▪1 TB per z/OS LPAR on z/OS V1.12 or V1.13 ▪Up to 4 TB per z/OS LPAR plan for z/OS V2.2 <ul style="list-style-type: none"> - Also planned for z/OS V2.1 (SoD)
Changed (node) cache structure
96 GB Fixed HSA
<p>Up to 85 LPARs (Up to 60 LPARs with z/OS V1.12 on any LPAR)</p>
Up to six logical channel subsystems (CSSs)
4 Subchannel Sets per CSS
Single Instruction Multiple Data (SIMD) instruction set
Two-way simultaneous multithreading (SMT) support for up to 128 cores (IFLs and zIIPs)
New and enhanced instructions
<p>XL C/C++ ARCH(11) and TUNE(11) exploitation: New z13 hardware instruction support, SIMD (Vector support) and Vector data, Decimal Floating Point packed conversion facility support, Performance improvements</p>



(z/OS support in blue)

g/Orlando-Eval

IBM zAware: z/OS and Linux® on IBM zSystems™
CPU Measurement Facility
Flash Express (Storage Class Memory-SCM)
CF exploitation of Flash Express
IBM z Systems Data Compression (zEDC) capability using zEDC Express
OSA Express5S
Shared RoCE Express Support
Greater than 256 PFID support
PCIe extended address translation
Enhanced the PCIe function definition
PCIe function measurement block changes
FICON Express16S
FICON® Dynamic Routing
High Performance FICON for z Systems (including zHPF extended distance II)
Fabric Priority for I/O write requests
CryptoExpress5S: Next Generation Coprocessor support, Support architecture for up to 85 Domains, Format Preserving Encryption (FPE)
Integrated Coupling Adapter (ICA) Links
Increased number of coupling CHPIDs, from 128 to 256 per CEC
zBX Model 004 support