
Oracle on Linux on System z with z13 Update -- Why this is one of the Best and Fastest Growing Linux on z Solutions Available

June 26, 2015

David Simpson
Oracle Technical Specialist
z Systems Advanced Technical Solutions
simpson.dave@us.ibm.com

Gaylan Braselton
IBM North America Solution Leader
Oracle on z Systems
gbrasel@us.ibm.com



Trademarks

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

AIX*	Geographically Dispersed Parallel Sysplex	OS/390*	VSE/ESA
APPN*	HiperSockets	Parallel Sysplex*	VTAM*
CICS*	HyperSwap	PR/SM	WebSphere*
DB2*	IBM*	Processor Resource/Systems Manager	z/Architecture
DB2 Connect	IBM eServer	RACF*	z/OS*
DirMaint	IBM e(logo)server*	Resource Link	z/VM*
DRDA*	IBM logo*	RMF	z/VSE
Distributed Relational Database Architecture	IMS	S/390*	zSeries*
e-business logo*	InfoPrint*	Sysplex Timer*	
ECKD	Language Environment*	System z	
Enterprise Storage Server*	MQSeries*	System z9	
ESCON*	Multiprise*	TotalStorage*	
FICON*	NetView*	Virtualization Engine	
GDPS*	On demand business logo	VSE/ESA	

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Intel is a trademark of Intel Corporation in the United States, other countries, or both.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

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

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

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


Topics -- Oracle On Linux on System z with z13 Update

- Oracle on Linux on z Systems Solution, Certification and Support
- zVM and IBM z13 and Linux with Oracle
- Oracle and I/O and z13 Performance
- IBM Flash System with Oracle

Oracle Technology Solutions for IBM z Systems

Oracle DB Server on z/OS

Oracle DB Server on Linux

Database	Oracle Database 10gR2	Oracle DB 10g Release 2 (10.2.0.5) Terminal Release	Oracle DB 10g Release 2 (10.2.0.5)
	Oracle Database 11gR2	No	Oracle DB 11g Release 2 (11.2.0.2, 11.2.0.3, 11.2.0.4)
	Oracle Database 12c	No	Oracle DB 12c Release 1 (12.1.0.1, 12.1.0.2) 

Oracle Fusion Middleware on z/OS

Oracle Fusion Middleware on Linux

Fusion Middleware	Oracle FMW 11gR1	Not Applicable Oracle DB/FMW	WebLogic Server 11g Release 1 (11.1.1.6+) SOA 11g Release 1 (11.1.1.6+) WebCenter 11g Release 1 (11.1.1.6+) Tuxedo 11g Release 1 (11.1.1.2+) WebLogic Portal 11g Release 1 (10.3.2)
	Oracle FMW 12cR1	Not Applicable Oracle DB/FMW	WebLogic Server 12c Release 1 (12.1.1+, 12.1.2+) Note: 12.1.2 includes Coherence
Enterprise Manager	Oracle Enterprise Cloud Control Agent	Not Applicable Oracle DB/FMW	12c Release 2 (12.1.0.3 ⁽¹⁾ , 12.1.0.4) Oracle DB 12c (12.1.0.1) Oracle DB 11gR2 (11.2.0.3 ⁽¹⁾ , 11.2.0.4)
	Oracle GoldenGate	11g Release 2 (11.2.1.0.5+) DB2 v10, 9.1 for z/OS (11.2.1.0.1+) DB2 v10, 9.1, 8.1 for z/OS 11g Release 1 (11.1.1.1.1) DB2 v10, 9.1, 8.1 for z/OS	12c Release 1 (12.1.0.1, 12.1.0.2) Oracle DB 11gR2 (11.2.0.2, 11.2.0.3)
Golden Gate	Oracle GoldenGate		11g Release 2 (11.2.1.0.2+) Oracle DB 10g R2 (10.2.0.4+)

Oracle on Linux on z – Strengths and Growth

- Solution Current with Oracle 12c
- Entire stack officially certified by Oracle (support.oracle.com)
- Oracle Patch Set Updates and Critical Patches for Loz same date as other platforms
- Over 1,000 Customers running Oracle on Loz, all sizes, industries
- Oracle is 'Best Fit' Solution for Loz
- SUSE and RedHat Support
- Dedicated IBM Team, continued growth, new z customer, over 12 POCs underway now
- IBM Oracle Alliance and International Competency Center
- Dedicated System z team in Oracle Development
- Dedicated System z team in Oracle Support
- Oracle presentations on Loz at SHARE, COLLABORATE, SIG, and Webcasts
- Over 10 Redbooks on Oracle z, latest May 14, Oracle participation
<http://www.redbooks.ibm.com/abstracts/sg248159.html?Open>
- System z 'My Oracle Support Community' on support.oracle.com
- Active z Oracle User Group, Volunteer led, Oracle and IBM Supported
www.oraclezsig.org (longest running still active official Oracle user group, 27 years)
- 28th Annual z Oracle SIG User Group Annual Conference, NY, April 2015, Oracle and IBM Supported
www.oraclezsig.org (includes presentations from past conferences)
- Customer Testimonials
<https://www.youtube.com/watch?v=yIBy6aNTqhQ> (Radixx)
<https://www.youtube.com/watch?v=o1SJyv7rNIA&feature=youtu.be> (Evertec)

Oracle License Requirements With Linux on z Systems and z13

For **Oracle Enterprise Edition (EE)** Database and **Core Based Pricing**:

- Any z eligible, Oracle licenses are not processor specific, and are transferrable

 - 1 IFL = 1 Core for any and all z Systems IFLs

- The Oracle Core Factor for any/all System z is 1.0

 - www.oracle.com/us/corporate/contracts/processor-core-factor-table-070634.pdf

- There are no special contracts or Oracle licenses for z Systems,
same licenses, same backup, DR etc terms apply to z as all other systems

- z LPAR's are considered 'Hard Partitions' by Oracle
for SubCapacity Pricing, where applicable

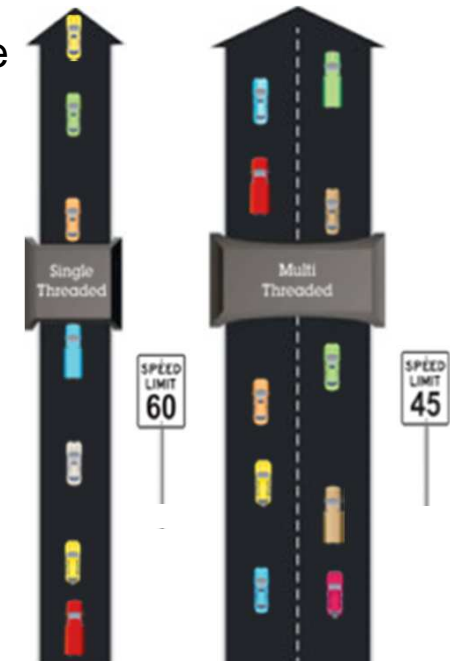
- Oracle license requirements not changed by multiple threads/core etc (SMT, z13)

IBM z13: SMT – Simultaneous Multi-Threading

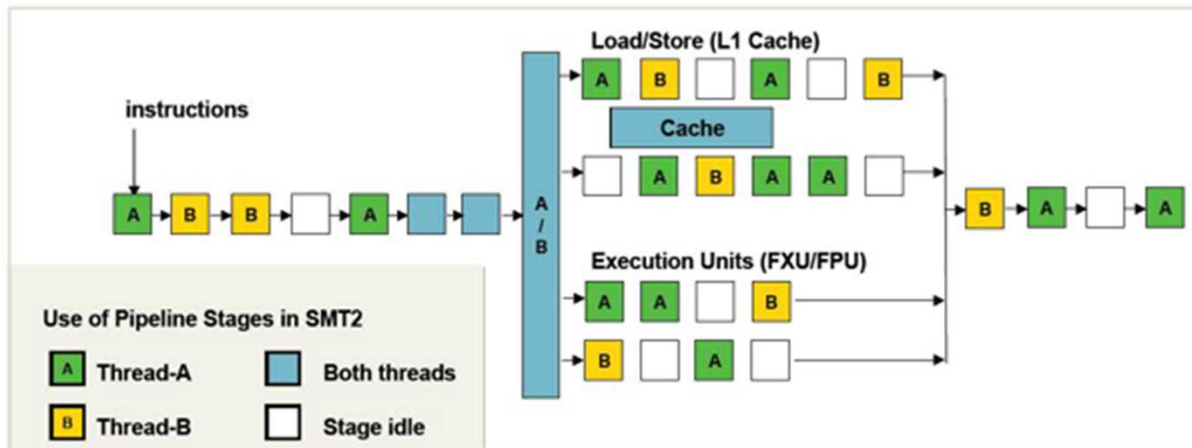
- Double the number of hardware threads per core
 - Independent threads can be more effective pipeline utilization

- Threads share resources – may impact single thread performance
 - Pipeline (eg. physical registers, fxu, fpu, lsu etc)
 - Cache

- Throughput improvement is workload dependent



Two zIIP lanes handle more traffic overall



- **Oracle (12c) 12.1.0.2** was released for System z on November 14, 2014 (same date as other platforms). Patches continue to be released for System z on the same date as **ALL** other platforms and are right up to date.
- **IBM z13** supports up to 141 Oracle configurable cores with up to 10 TB of memory.
 - Clock speed of new cores - 5 GHz
 - Simultaneous multithreading (SMT) support with z/VM 6.3
 - z/VM support for Single Instruction Multiple Data (SIMD): In a future deliverable

  **Patch 20299023: DATABASE PATCH SET UPDATE 12.1.0.2.3 (APR2015)**

Last Updated Apr 14, 2015 1:18 AM (Tuesday)

Product Oracle Database - Enterprise Edition
(More...)

Release Oracle 12.1.0.2.0

Platform IBM: Linux on System z

Size 50.9 MB

Download Access Software

Classification Security

Patch Tag All Database

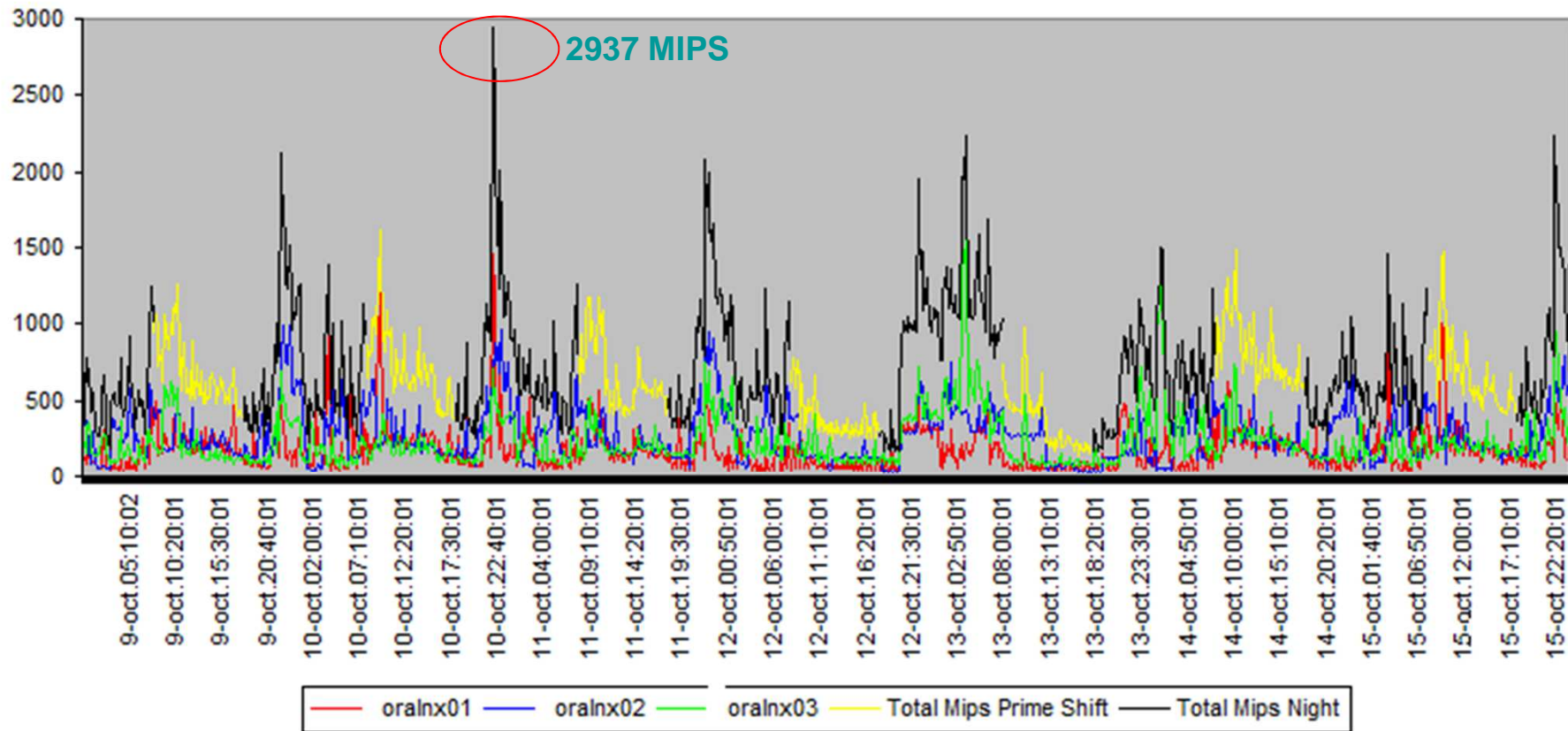
Oracle Database 12.1 Support Update for Linux on System z

Linux on System z specifics

- It's Fast
 - Built using PDF (Profile Directed Feedback).
 - Approximately 5% Faster even with all the new features.
- New Features – less resources
- EM agent 12.1. enabled
 - OEM Cloud Control 12cR3 or 12cR4
- IBM Redbook
 - **Experiences with Oracle Database 12c Release 1 on Linux on System z** **SG248159** <http://www.redbooks.ibm.com/abstracts/sg248159.html?Open>

Sizing Consolidated CPU consumption – equivalent MIPS

October 2012 - equivalent MIPS (wo z/VM)



zEC12 vs z13 Testing Parameters



- Test results in this presentation are my own for Educational purposes only.
- Test results should not be construed as typical for a particular customer workload.
- z/VM development recommend getting good MONWRITE data BEFORE moving to z13 and initially disable SMT if possible.
- Use the z/VM CPUMF / SMTMET tool to extract SMT metrics
<http://www.vm.ibm.com/perf/reports/zvm/html/1q5smt.html>
- REALLY Important to be on the recommended z/VM service and Linux kernel levels: Suse 11 SP3+ (3.0.101-0.40.1) / Red Hat 6.6+ (2.6.32-504.16.2.el6) per <http://www-03.ibm.com/systems/z/os/linux/resources/testedplatforms.html>

Testing on zEC12 with 2 Dedicated IFLs



Instance Efficiency Percentages (Target 100%)

Buffer Nowait %:	100.00	Redo NoWait %:	100.00
Buffer Hit %:	100.00	In-memory Sort %:	100.00
Library Hit %:	99.99	Soft Parse %:	87.07
Execute to Parse %:	99.99	Latch Hit %:	100.00
Parse CPU to Parse Elapsed %:	100.00	% Non-Parse CPU:	99.99
Flash Cache Hit %:	0.00		

Top 10 Foreground Events by Total Wait Time

Event	Waits	Total Wait Time (sec)	Wait Avg(ms)	% DB time	Wait Class
DB CPU		239.6		99.6	
db file sequential read	328	.1	0.33	.0	User I/O
control file sequential read	298	.1	0.36	.0	System I/O

- Silly Little Oracle Benchmark (SLOB)
- Logical I/O (Random memory access to Oracle SGA)
- Want to have 99% + DB CPU and 100% Buffer Hit Ratio for a clean test from Oracle Automatic Workload Repository (AWR) Report.

z13 versus zEC12 hardware comparison

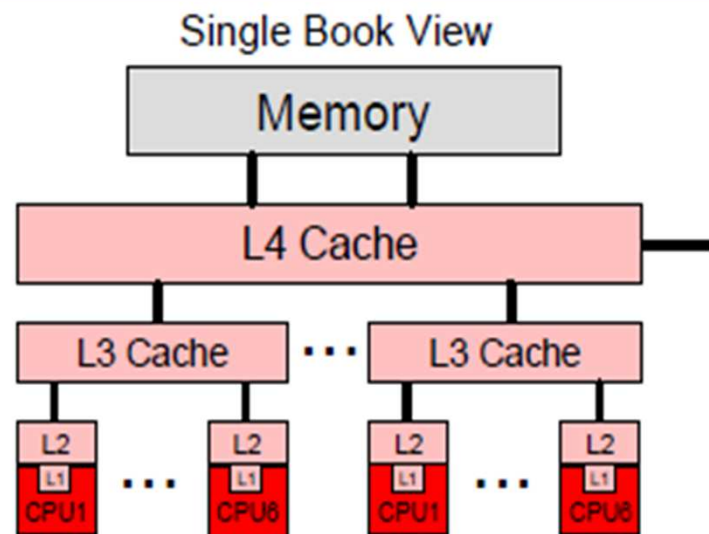
■ zEC12

▶ CPU

- 5.5 GHz
- Enhanced Out-Of-Order

▶ Caches

- L1 private 64k i, 96k d
- L2 private 1 MB i + 1 MB d
- L3 shared 48 MB / chip
- L4 shared 384 MB / book



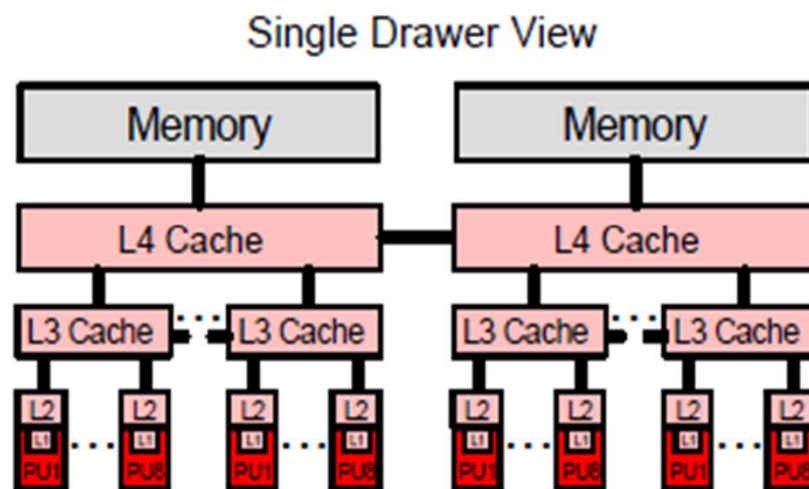
■ z13

▶ CPU

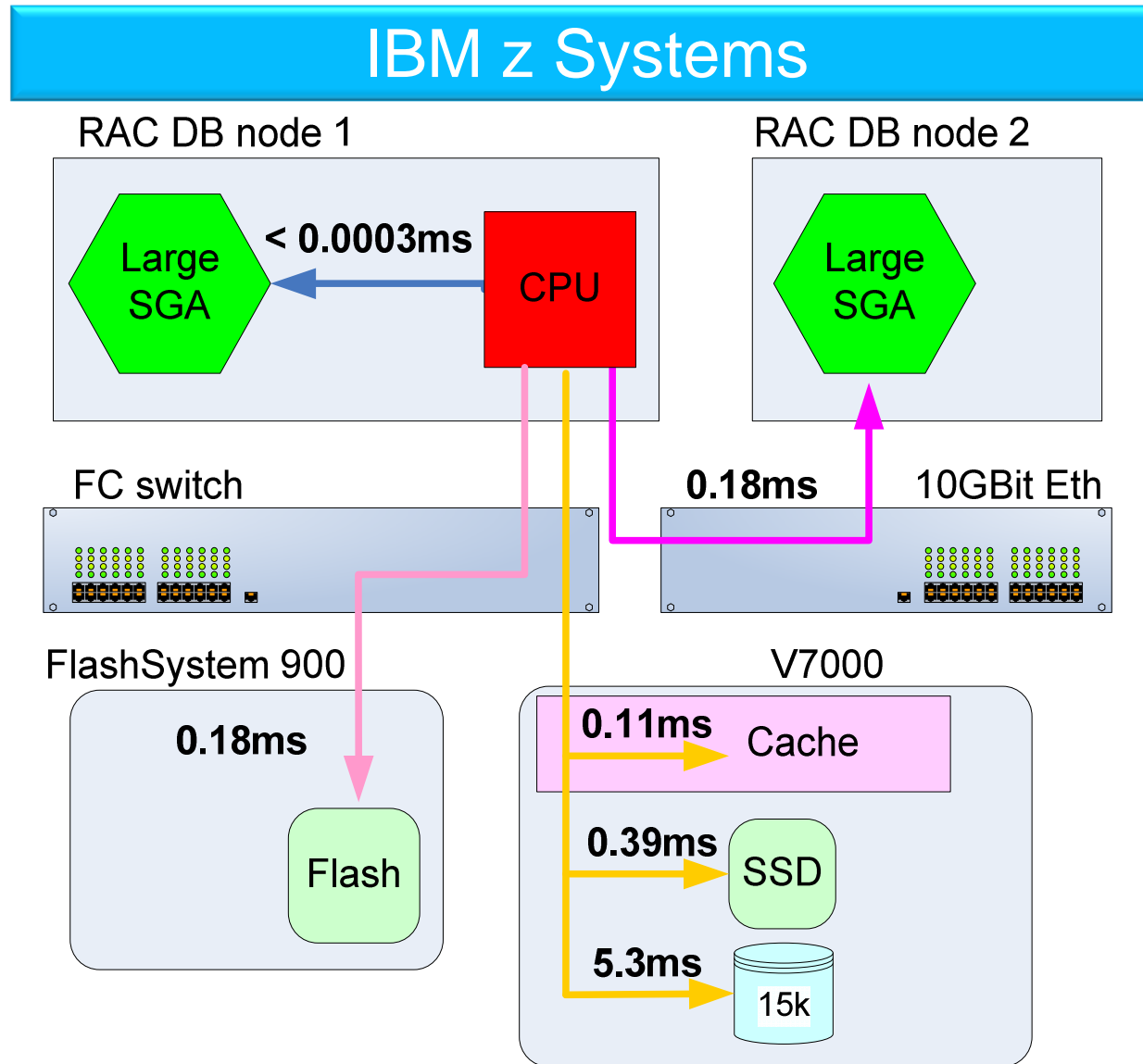
- 5.0 GHz
- Major pipeline enhancements

▶ Caches

- L1 private 96k i, 128k d
- L2 private 2 MB i + 2 MB d
- L3 shared 64 MB / chip
- L4 shared 480 MB / node
 - plus 224 MB NIC

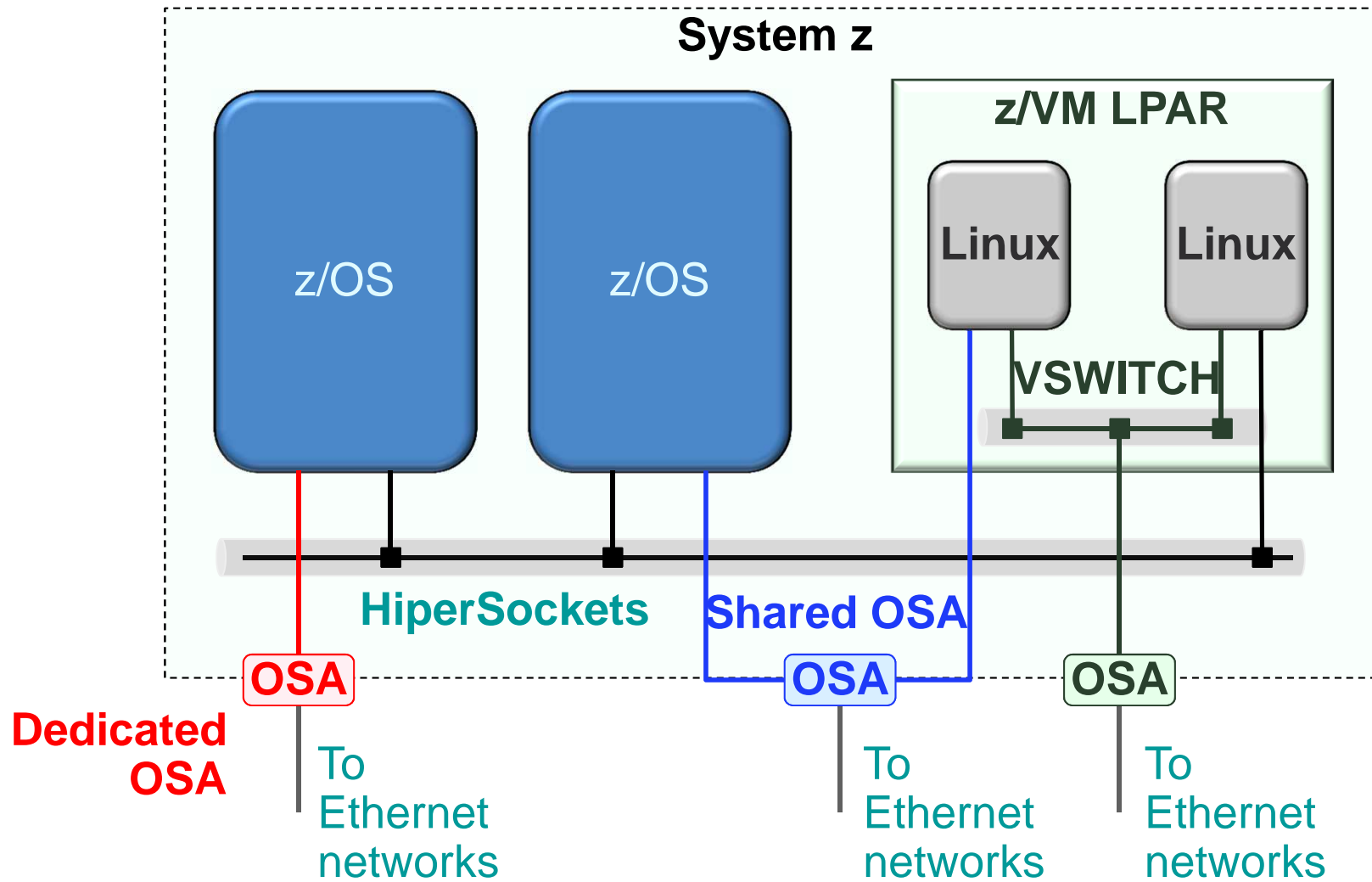


Cache Data Close to the DB

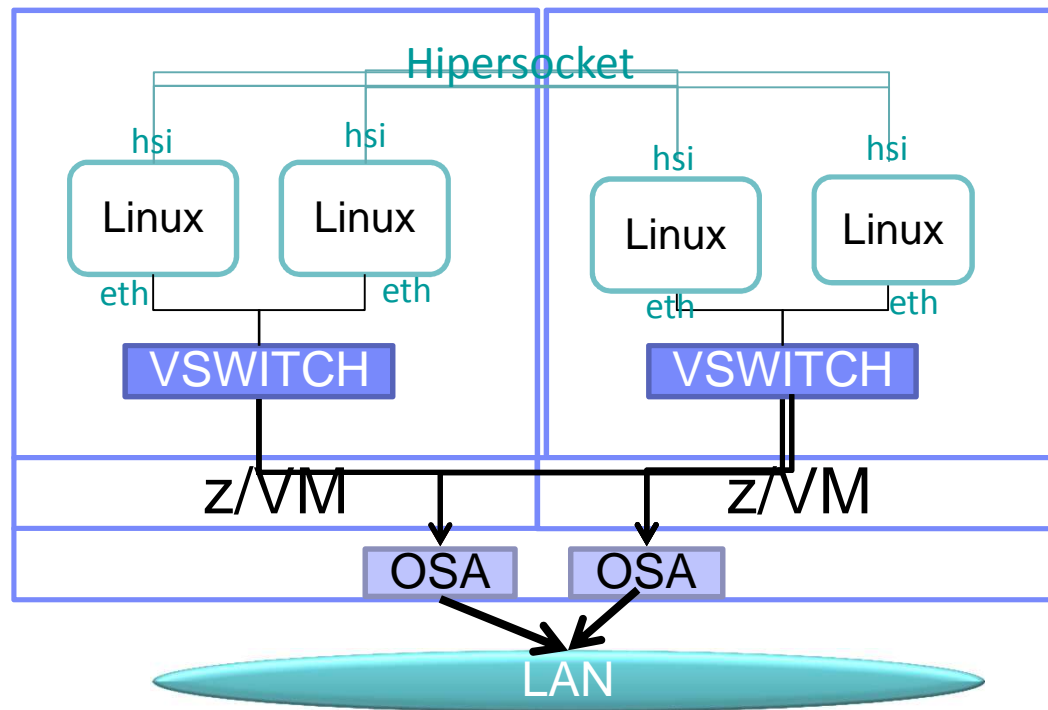


- The service times shown on this slide reflect the expected random read service time for an 8KB block of data, some are estimated.

System z Networking




- Choose correct MTU size (should match with App or DB Server)
- Adjust Network queue length kernel parameter
- For high transactional applications try to reduce number of hops (latency) between app and db server



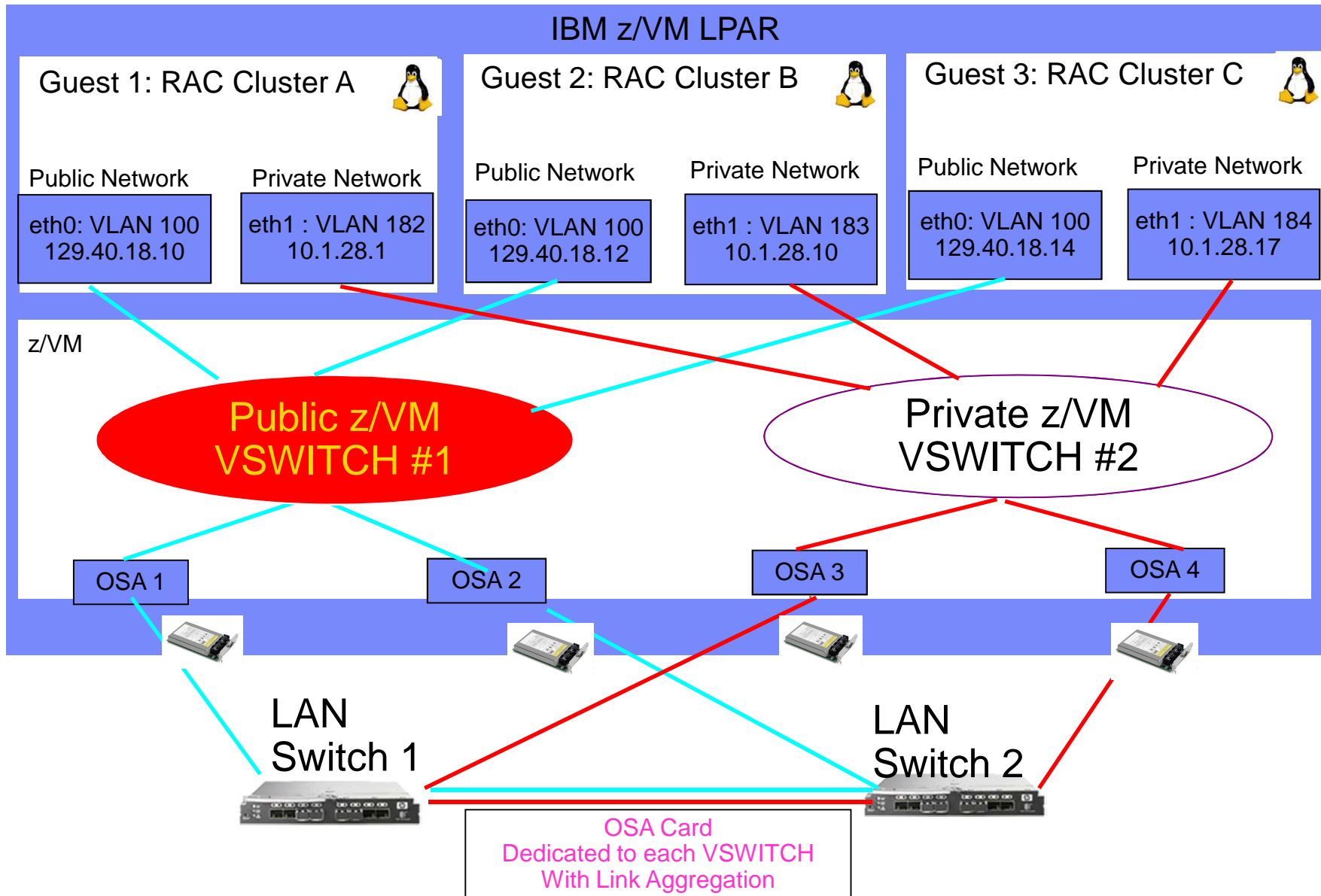
Choose the Correct Network MTU size

netstat -s of Interconnect	MTU Size of 1492 (default)	MTU Size of 8992 (with 8K DB block size, especially RAC)
Before reassemblies	43,530,572	1,563,179
After reassemblies	54,281,987	1,565,071
Delta assemblies	10,751,415	1,892

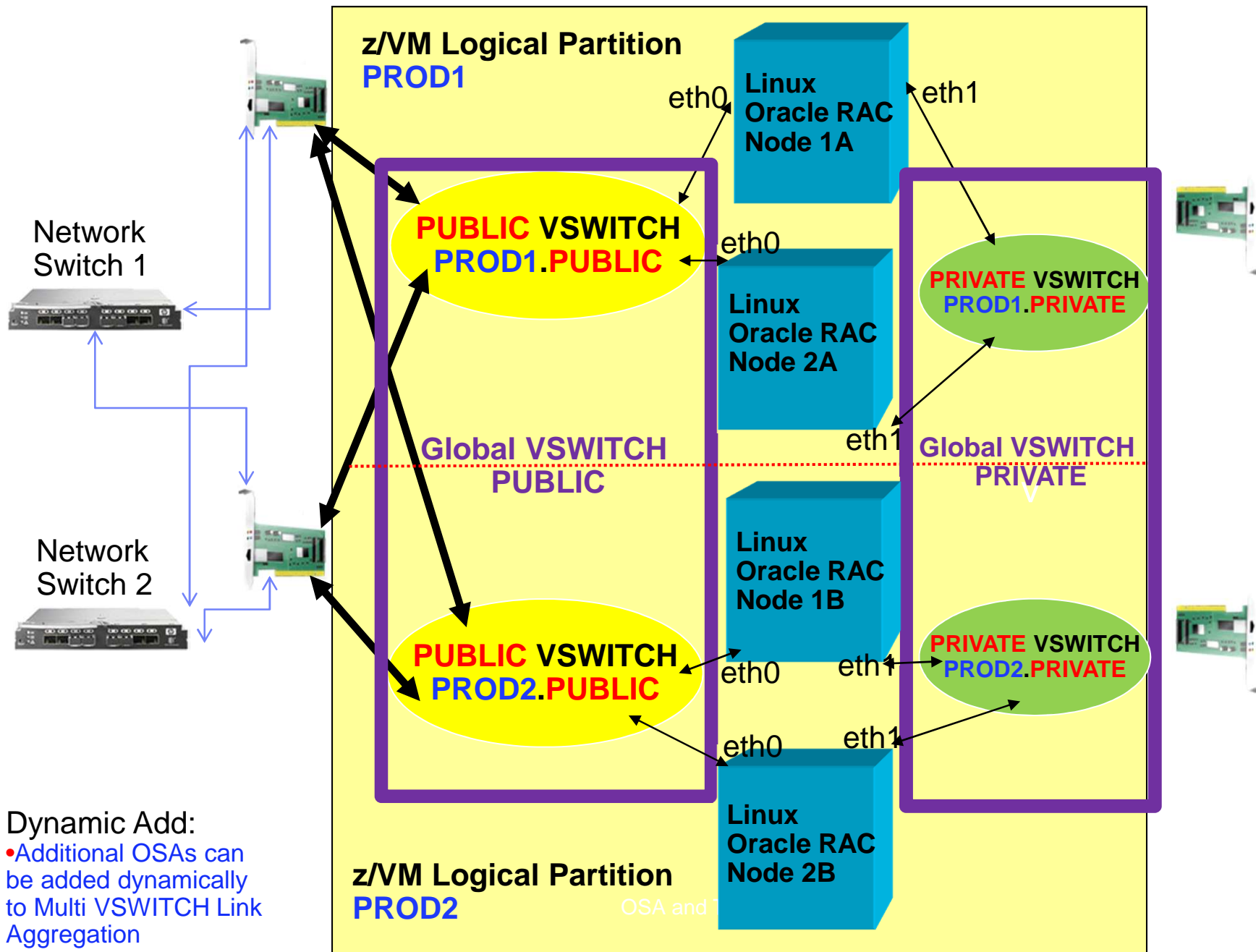
Oracle High Availability Networking Options:

- **Link Aggregation** – (Active / Active) Allow up to 8 OSA-Express adapters to be aggregated per virtual switch. Each OSA-Express feature must be exclusive to the virtual switch (e.g. OSA's can now be shared ).
- **Linux Bonding** – create 2 Linux interfaces – e.g. **eth1** & **eth2** and create a bonded interface **bond0** made up of eth1 and eth2.
- **Oracle HAIP** – Oracle 11gR2+ can now have up to 4 Private interconnect interfaces to load balance interconnect traffic.

Oracle Interconnect with z/VM Link Aggregation (zEC12)



Oracle RAC with z/VM Multi VSWITCH LAG



- z/VM 6.3 with APARS VM65583 and PI21053.
- OSA-Express4S & OSA-Express5s support for Multi-Vswitch Link Aggregation requires IBM z13
- A port group (LAG) can be connected to up to 16 LPARS (single CEC). A port group cannot span multiple CECs.
- *Please See Rick Tarcza's presentation <http://www.vm.ibm.com/virtualnetwork/63lnkag.pdf> for more information*

System z & IBM Flash System: Highest Reliability, Maximum Performance



Now you can leverage the “Economies of Scale” of Flash

- Easily added to your existing SAN
- Accelerate Application Performance
- Gain Greater System Utilization
- Lower Software & Hardware Cost
- Save Power / Cooling / Floor Space
- Drive Value Out of Big Data



IBM FlashSystem is certified ([reference SSIC](#)) to attach to Linux on System z, with or without an SVC, to meet your business objectives

Would you like to demo this architecture?

You can now demo hardware either in person or virtually.

Demo Location: Benchmark Center in Poughkeepsie, NY

Performance of Linux on System z with FlashSystem

I/O bound relational databases can benefit from IBM FlashSystem over spinning disks.

- **21x** reduction in response times*
- **9x** improvement in IO wait times*
- **2x** improvement in CPU utilization*

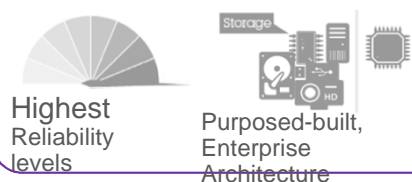
* IBM internal test results

Why IBM FlashSystem for Linux on System z?

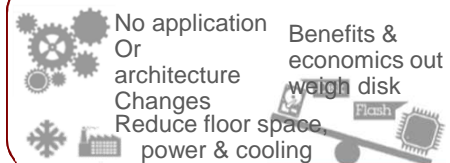
Extreme Performance



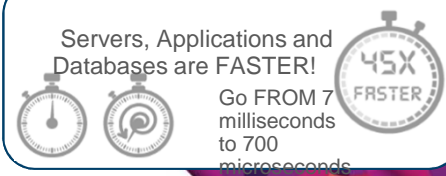
Enterprise Reliability



Macro Efficiency



IBM MicroLatency™



Summary

- I/O bound AWR's can benefit from "Flash" Storage in the following way:
- **21x** reduction in response times
- **957%** improvement in IO wait time

Top 5 Timed Foreground Events

Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
db file sequential read	16,543,046	126,308	8	47.26	User I/O
log file sync	453,034	81,895	181	30.64	Commit
DB CPU		42,771		16.00	
local write wait	42,921	2,858	67	1.07	User I/O
db file parallel read	94,662	1,847	20	0.69	User I/O

Other Recommendations

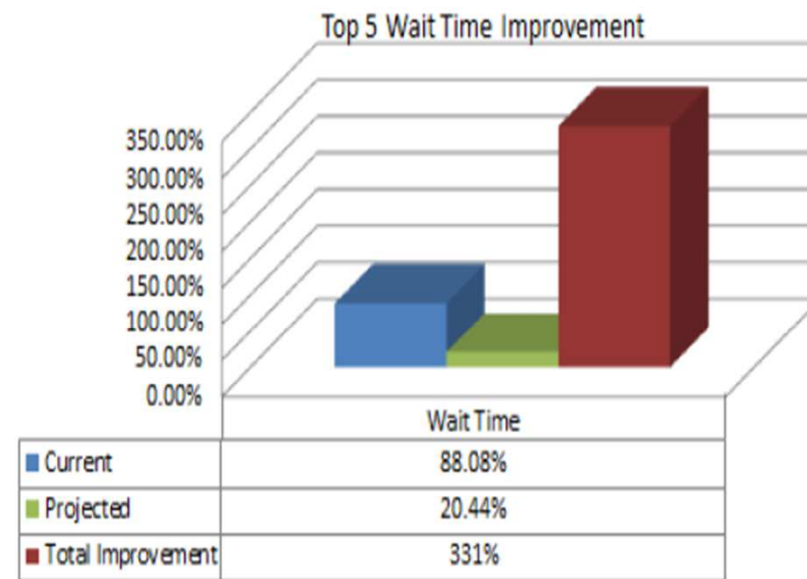
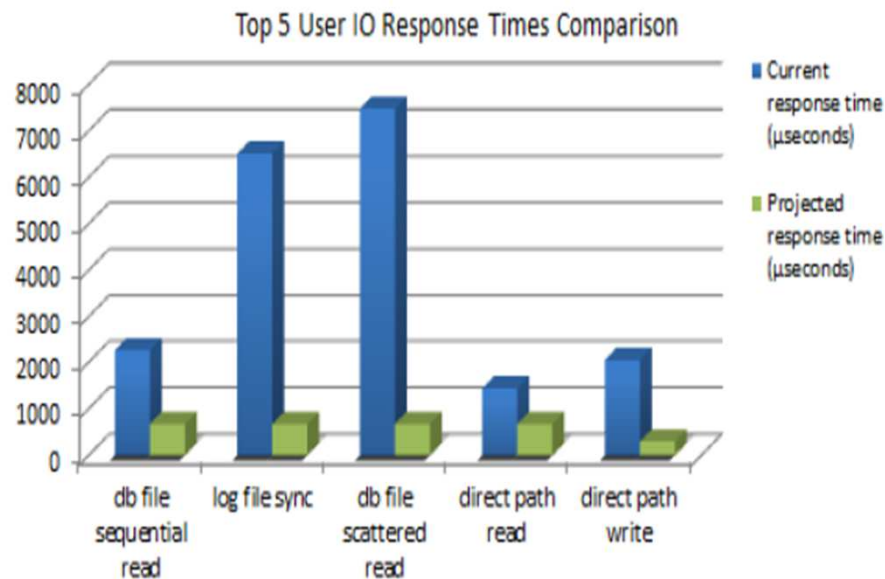
- Rerun AWR reports after implementing to validate outcomes
- Use application partitioning on the app-tier to ensure data files and log files are located on the faster storage.

Top 5 Timed Foreground Events

Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
DB CPU		4,497		78.46	
log file sync	2,047,256	982	0	17.14	Commit
library cache: mutex X	212,427	105	0	1.82	Concurrency
db file scattered read	58,832	44	1	0.77	User I/O
db file sequential read	43,311	18	0	0.31	User I/O

Aggregating factors for FlashSystem implementation

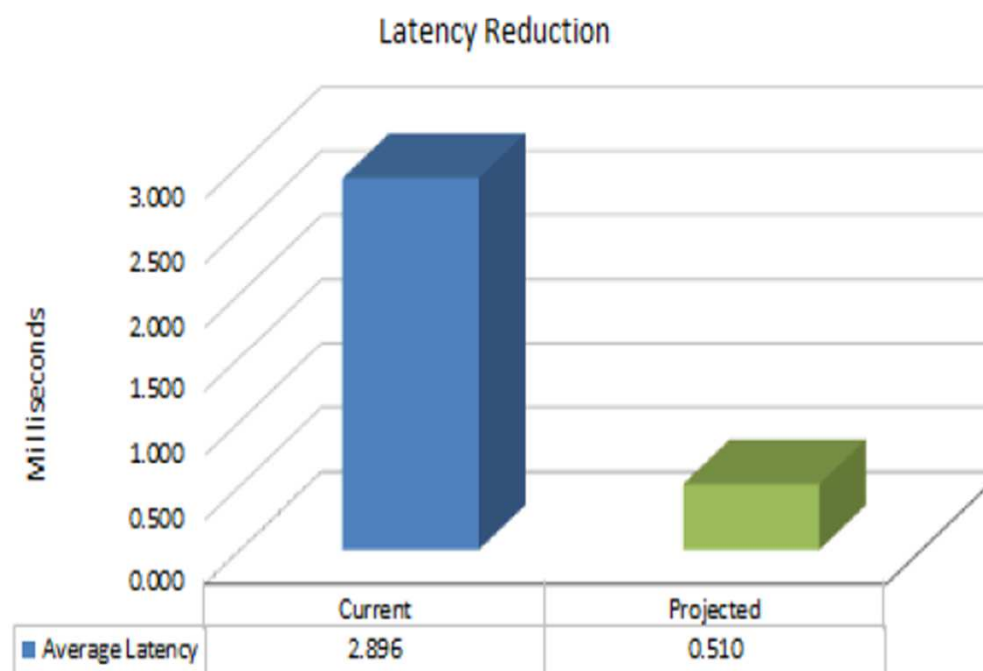
- Reduce User IOWait time



User IOWait events indicate a reduction in IOWait times are possible with a FlashSystem implementation. Db file sequential read is causing the majority of the disk contention across all three AWRs. The IOWait time would decrease from 88.08% of overall wait time to 22.44%, an improvement of **331%**.

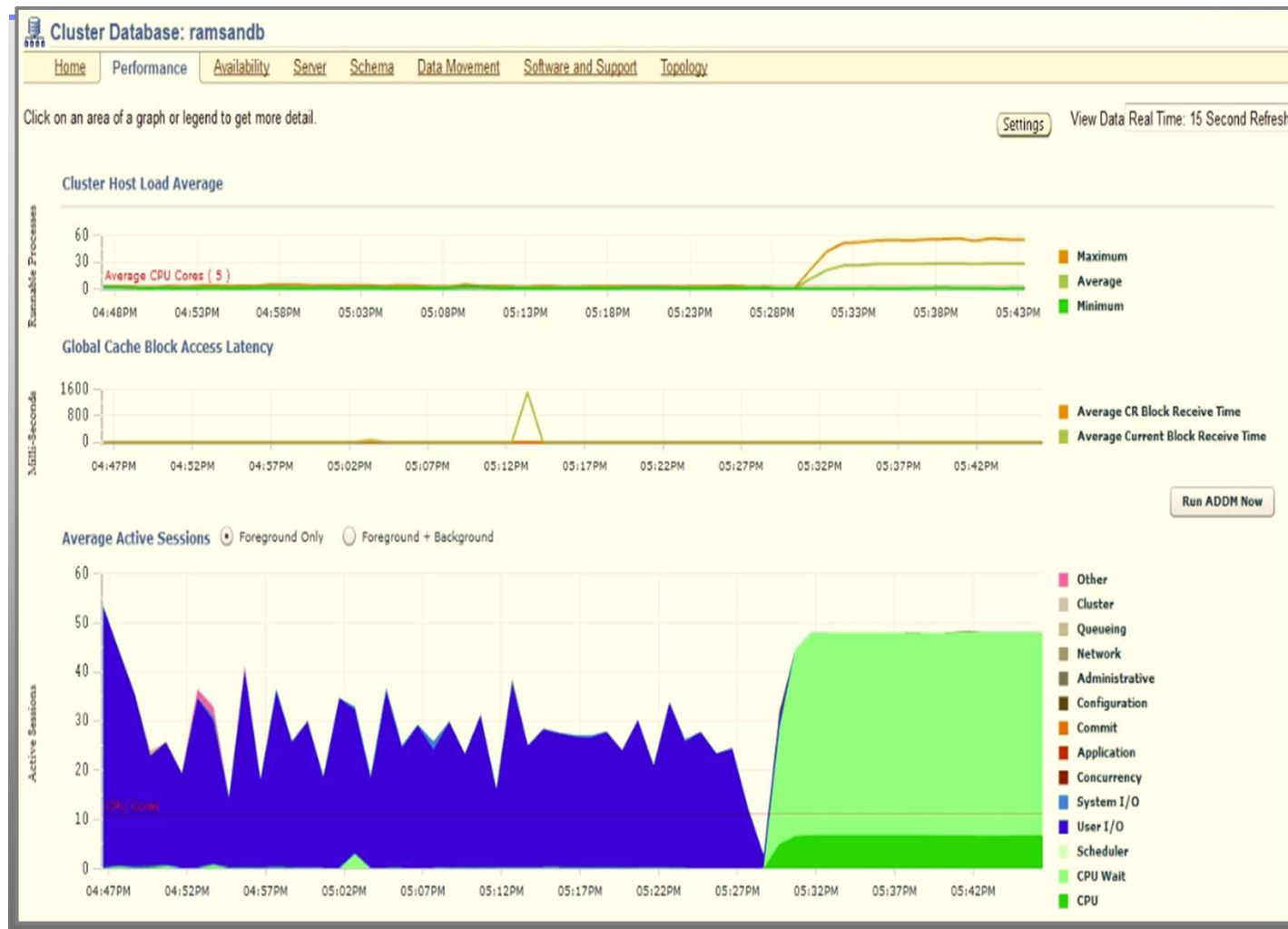
Aggregating factors for FlashSystem implementation

- Reduce Response Time / Latency



The microsecond response times of the FlashSystem would significantly reduce latency while driving higher utilization at the server and application level. Average Latency would decrease from 2,896 microseconds to 510 microseconds.

Acceleration of Database with IBM FlashSystem



After switching to FlashSystem

(05:27 PM) Disk IO wait disappears and waiting is now on host CPU. This graph shows the effect of the low latency of FlashSystem and how it increases the host CPU utilization.

IBM Flash Systems vs. Competitor Flash Array

In same space and power, IBM FlashSystem delivers **2.4x higher IOPS** and up to **14x the usable capacity** at around **1/5th the cost** per usable Terabyte

	Other Vendor	IBM FlashSystem 900 (5.7 TB cards)	IBM FlashSystem 900 (2.8 TB cards)	IBM FlashSystem 900 (1.2 TB cards)
Max 8k Read IOPS	400,000 ¹	955,000 est	955,000 est	845,000 est
Raw Storage	12.8 TB ¹	105.5 TB	52.7 TB	21.4 TB
Usable Storage	3.7 TiB ²	51.8 TiB ⁴	26.3 TiB ⁴	10.9 TiB ⁴
3-year TCO (List)	\$330,880 ³	\$947,700	\$498,900	\$254,100
\$/Usable TB (List)	\$89k	\$18k	\$19k	\$23k

*** In 2014, IBM was the #1 seller of Flash Based storage systems**

¹ Data Sheet for 3 servers - 1,036,000 IOPS

² ASM High redundancy assumed

³ Includes disk license and 3 years maintenance

⁴ With Variable Stripe RAID

Oracle Certified Virtualized Platforms



- **Oracle VM** & **IBM z/VM** Hypervisors are CERTIFIED to run Oracle workloads. (IBM PowerVM, z PR/SM support LPAR virtualization as well)
- VMWARE supported but NOT certified by Oracle.
- **Oracle VM** cannot do memory overcommit – maximum recommended overcommit of virtual to real processors is 2:1
- **IBM z/VM** handles over commitment of Memory and Virtual processors very well. (You still need to conserve resources where possible!)



- Performance
 - Oracle runs well on System z for both memory access (Logical I/O)
 - Integration with Flash Systems allows Oracle to run well with Physical I/Os
- Consolidation
 - z/VM can virtualize / overcommit resources well.
 - System z can run Oracle at very high cpu utilization rates with little degradation.
 - System z can dynamically add system resources (memory, network, cpu)
- Highly Available
 - System z runs Oracle workloads highly available (hardware) and in some cases can avoid configuring Oracle RAC for availability.
 - Linux HA solutions can be leveraged to increase application availability.
- Security
 - Oracle on System z can be ran highly secure with FIPs (US Govt.) 140-2 compliance at z/VM and Oracle levels.
 - SSL Crypto card support for Oracle SQL*net network traffic.

Information Sources

- <http://www.oracle.com/ibm>
 - Oracle IBM Partner Relationship
- <http://otn.oracle.com>
 - Oracle Select “Downloads”
- <http://www.vm.ibm.com/perf/tips>
 - General z/VM Performance & Tuning Tips, Capacity planning
 - <http://www.vm.ibm.com/virtualnetwork/63lnkag.pdf>
- <https://support.oracle.com>
 - Oracle Support Webpage (My Oracle Support)
- <http://www-124.ibm.com/developerworks/oss/linux390/index.shtml>
 - Lot’ s of information on Linux for zSeries, IBM DeveloperWorks
- <http://www-128.ibm.com/developerworks/linux/linux390/perf/index.html>
 - Hints and Tips for tuning Linux on System z
- <http://www.zseriesoraclesig.org>
 - Special Interest Group of Oracle users on the mainframe (z/OS and Linux)
- <http://www.mail-archive.com/linux-390%40vm.marist.edu/>
 - Marist List Server
- <http://www.ibm.com/redbooks>
 - SG24-6482-00 Experiences with Oracle Database 10g on Linux for zSeries
 - SG24-7191-00 Experiences with Oracle 10gR2 Solutions on Linux for System z
 - SG24-7573-00 Using Oracle Solutions on Linux on System z
 - SG24-7634-00 Experiences with Oracle Solutions on Linux for IBM System z
 - SG24-8159-00 Experiences with Oracle Database 12c Release 1 On Linux on System z
 - REDP-4788-00 Installing Oracle 11gR2 RAC on Linux on System z

Questions?



Thank
YOU