#SHAREorg

SHARE

Leading-edge Technology on System z

David Rhoderick
IBM Corporation

Tuesday 11th March 2014
Session Number 15033

Test link: www.SHARE.org

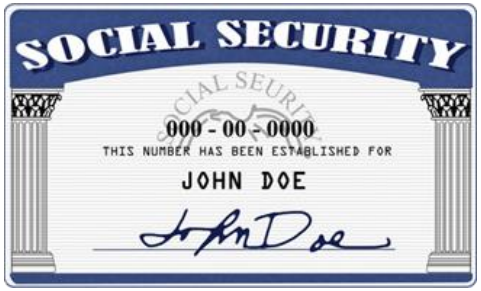
SHARE in Anaheim

Fifty years ago, IBM introduced the first mainframe computer…

System 360 – April 7, 1964

It helped put men on the moon…

It touches all of us from the day we were born…



It was revolutionary…

It was innovative…

It changed the world!

It has changed the way we live and work…

Source: IBM

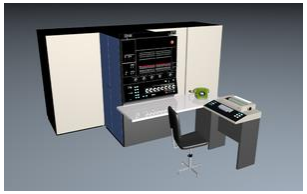
Initial innovation optimized hardware and software for centralized processing of 3GL apps (OLTP and batch)

Customer investment horizon 5-10 years

- 24-bit addressing (32-bit architecture)
- 1 or 2 cores
- 16MB storage
- 24K core memory

- VM operating system

- 24-bit or 31-bit virtual addressing
- Fully integrated monolithic memory
- 256 channel architecture
- Virtual storage



- MVS, IMS, CICS, and DB2

S/360 S/370

1964 ***1970***

Then came the PC, networking and client-server…

System 360 – April 7, 1964

It is still revolutionary…

It is still innovative…

It is still changing the world!

…Did mainframe innovation *stop* when PCs came along?

IBM model 5150 – c.1981

Then came the PC, networking and client-server…

System 360 – April 7, 1964

NO!! – IBM continues to invest $BILLIONS in mainframe technology

It is still revolutionary…

It is still innovative…

It is still changing the world!

IBM model 5150 – c.1981

Customer demand and technical leadership have lead to *continuous* re-invention of the mainframe

Hardware carry-forward + continuous application compatibility

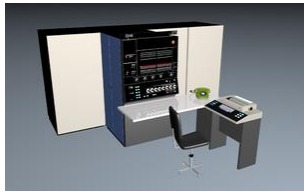
- 24-bit addressing (32-bit architecture)
- 1 or 2 cores
- 16MB storage
- 24K core memory

- VM operating system

S/360

1964

- 24-bit or 31-bit virtual addressing
- Fully integrated monolithic memory
- 256 channel architecture
- Virtual storage



- MVS, IMS, CICS, and DB2

S/370

1970

- CMOS processors
- More than 1,000 MIPS
- Parallel sysplex
- Enterprise Systems Architecture (ESA)

S/390

1990

- Specialty engines
- Hardware-assisted compression and encryption
- Decimal floating point
- 64-bit superscalar architecture

- WebSphere

zSeries

2000

- zEC12: up to 120 cores, 5.5GHz speed, 78,000+ MIPS
- zBC12 for mid-range
- Hybrid computing
- zBladeCenter extension and zManager

- Rational Development & Test

zEnterprise

2010

The IBM zEnterprise server – ready for the business challenges of today and the future

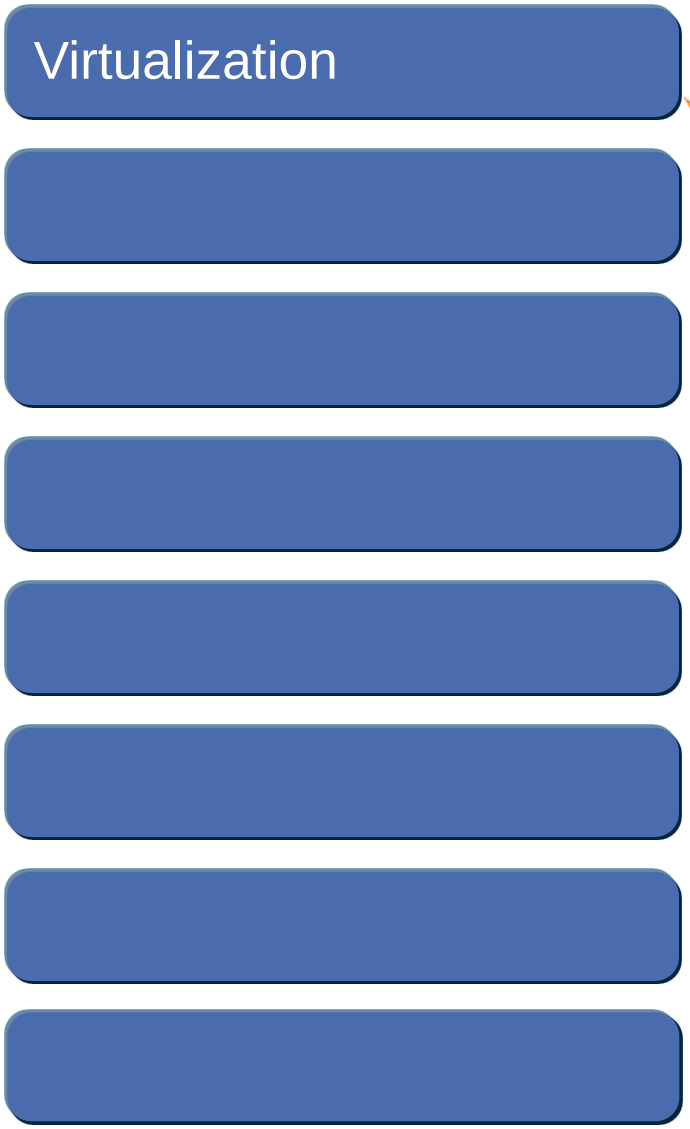
IBM zEnterprise EC12

IBM zEnterprise BC12

- The *fastest*, most *available* and *secure* platform commercially available
- Supports today's newest workloads
  - Data and analytics
  - Cloud
  - Mobile
- A multi-architecture platform for hybrid workloads
- Lowest total cost of ownership for most enterprise workloads

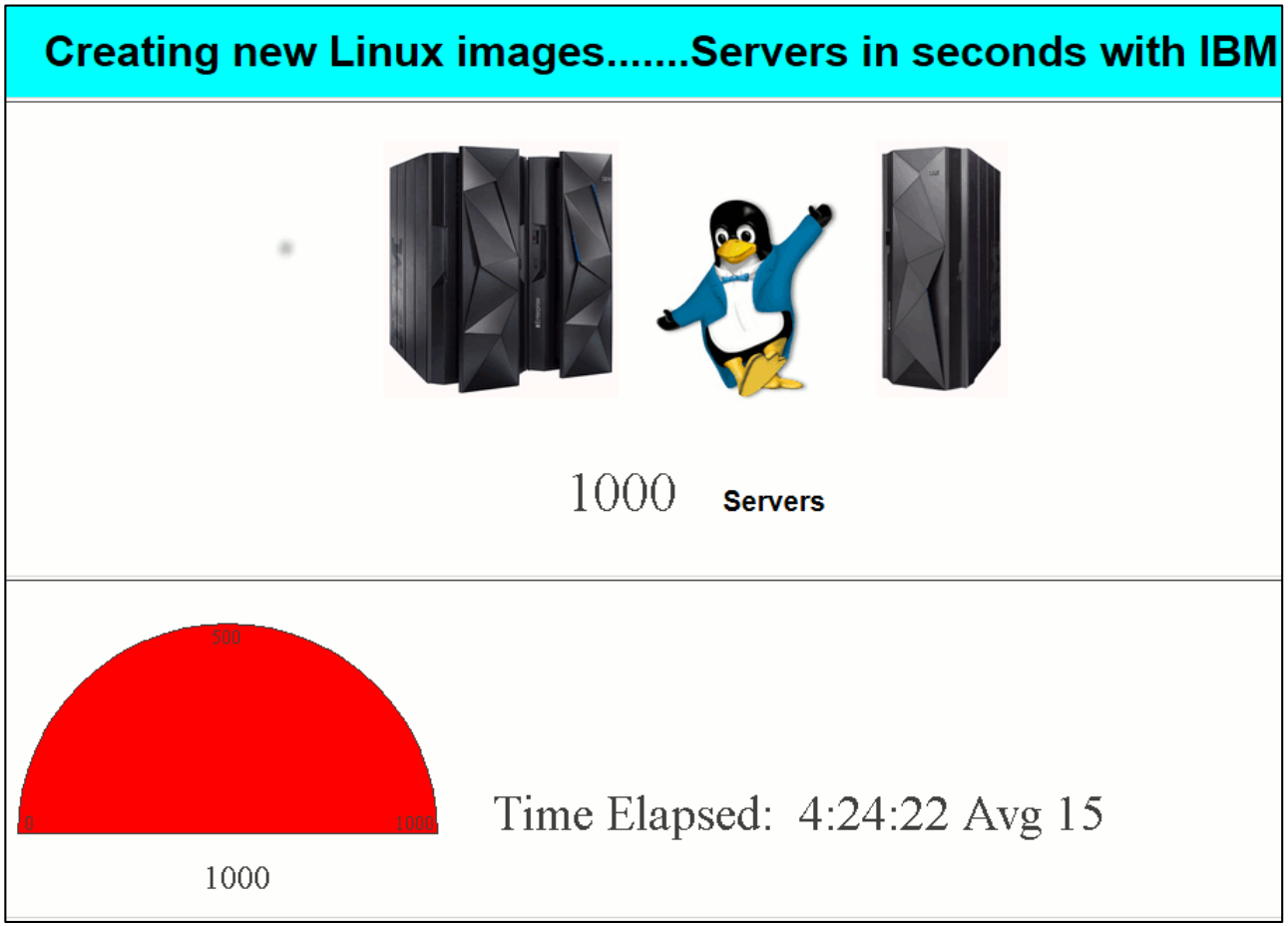
All due to the key mainframe innovations from the last 50 years…

The IBM mainframe was the world's first virtualized server

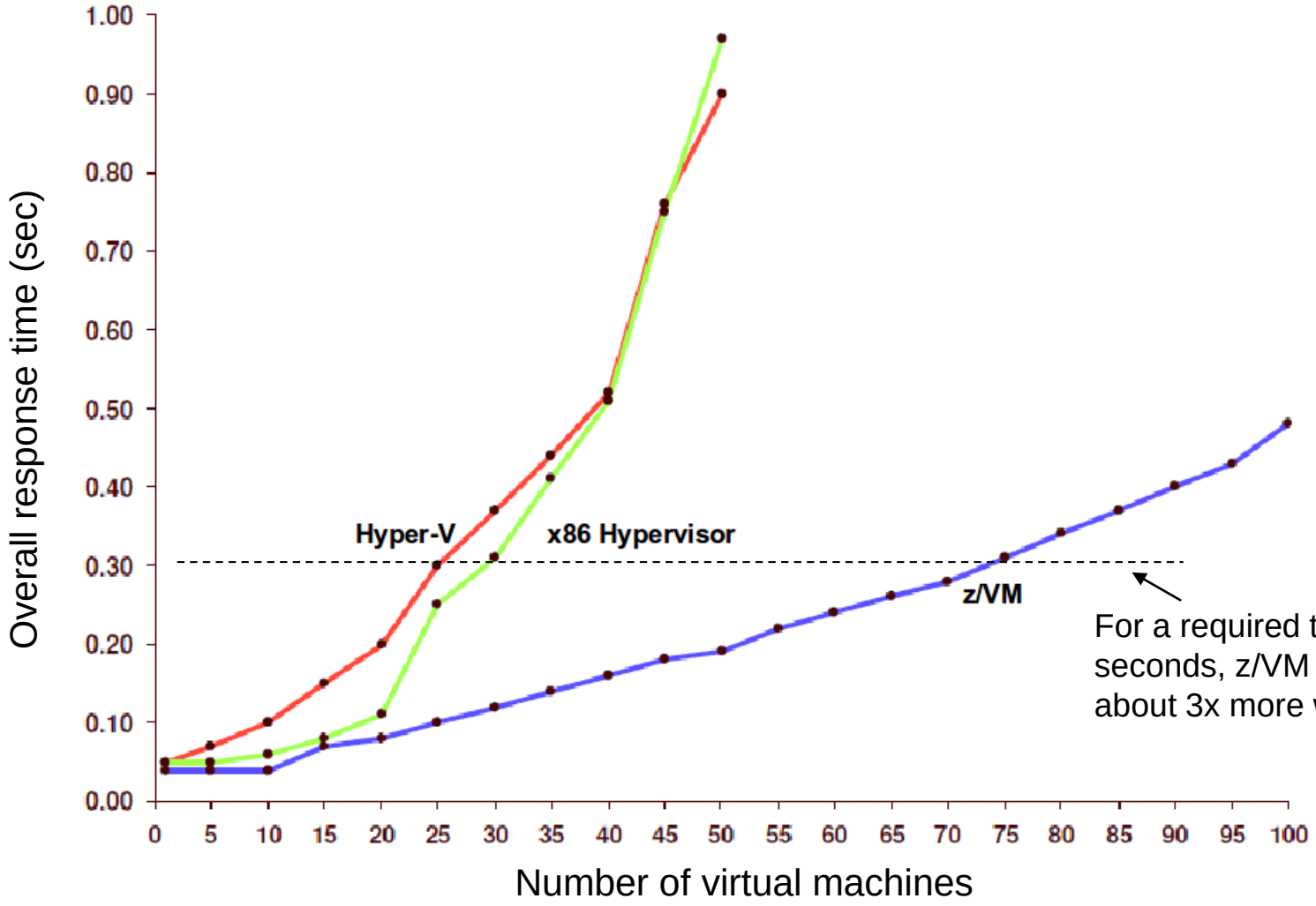


- Shared-everything design
- Virtualization built into the microcode
- Thousands of virtual guests
- Near 100% utilization
- Ideal choice for cloud deployments

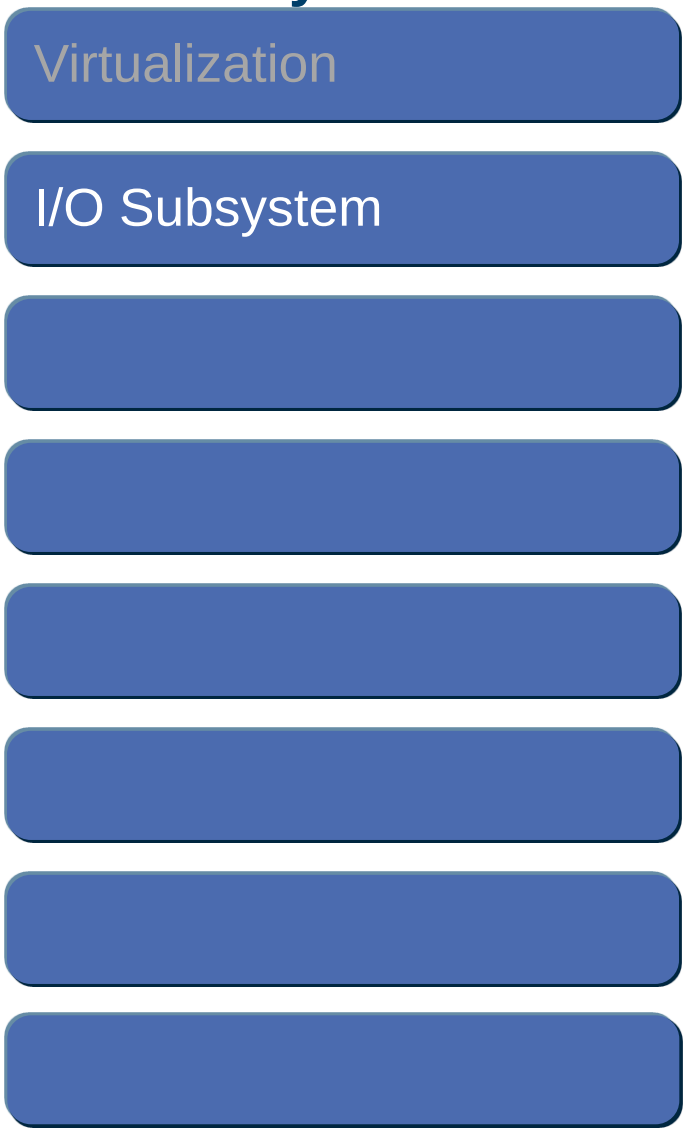
DEMO: How many virtual machines can zEnterprise create?



Compared to leading distributed hypervisors, z/VM demonstrates better scalability

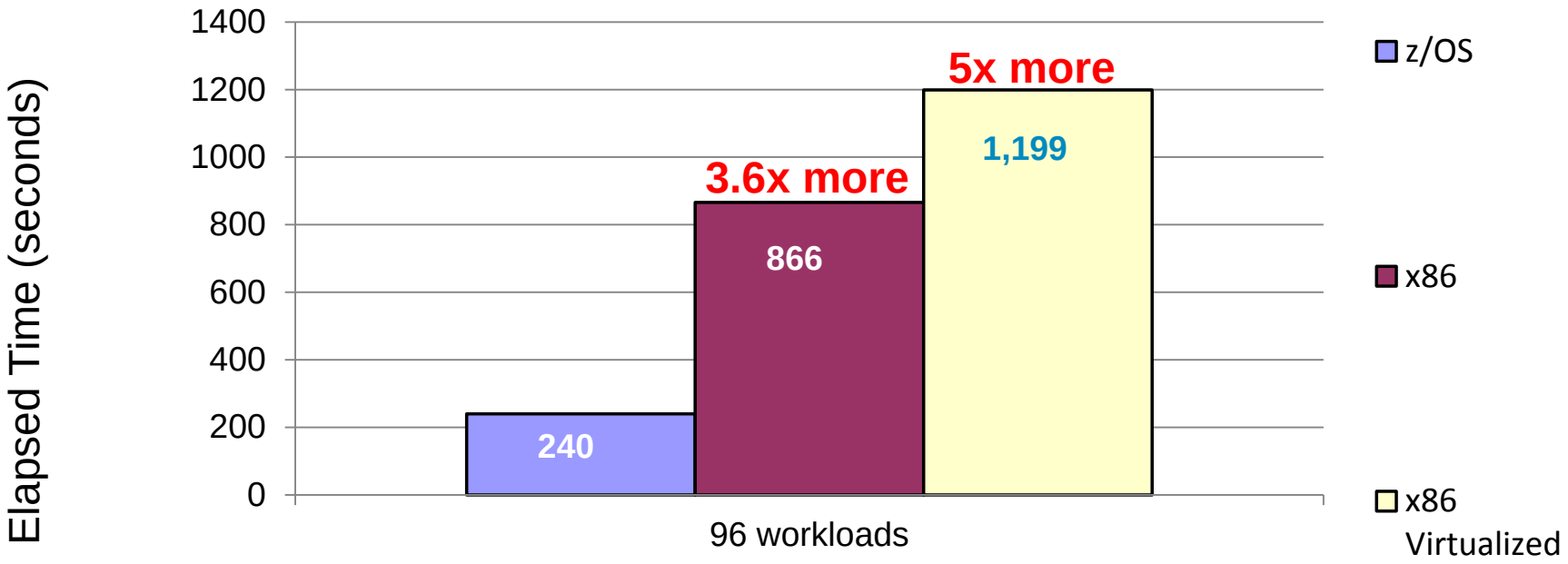


A unique zEnterprise feature not found on other servers is the I/O subsystem



- Reduces CPU usage by offloading I/O overhead
- Reduces number and cost of software licenses
- Improves I/O performance for batch and high performance OLTP
- Allows introduction of new facilities into existing I/O subsystem

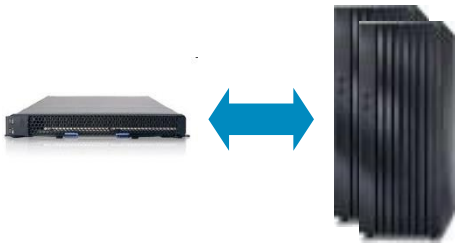
In comparison tests of I/O load capacity, Intel times were significantly slower



Performance comparison test of an I/O intensive workload with identical enterprise class storage. zEC12 had **8** core. Westmere EX server had **40** core @2.4GHz. Each system connected via 4 x 8Gb links to DS8800. zEC12 running against 8 SSD DASD CKD volumes. Intel server running against 8 SSD LUNs FB volumes. Note: Storage limitations came into effect at workload counts greater than 96.

Source: IBM CPO

Batch workloads take advantage of zEnterprise capability to support high I/O capacity

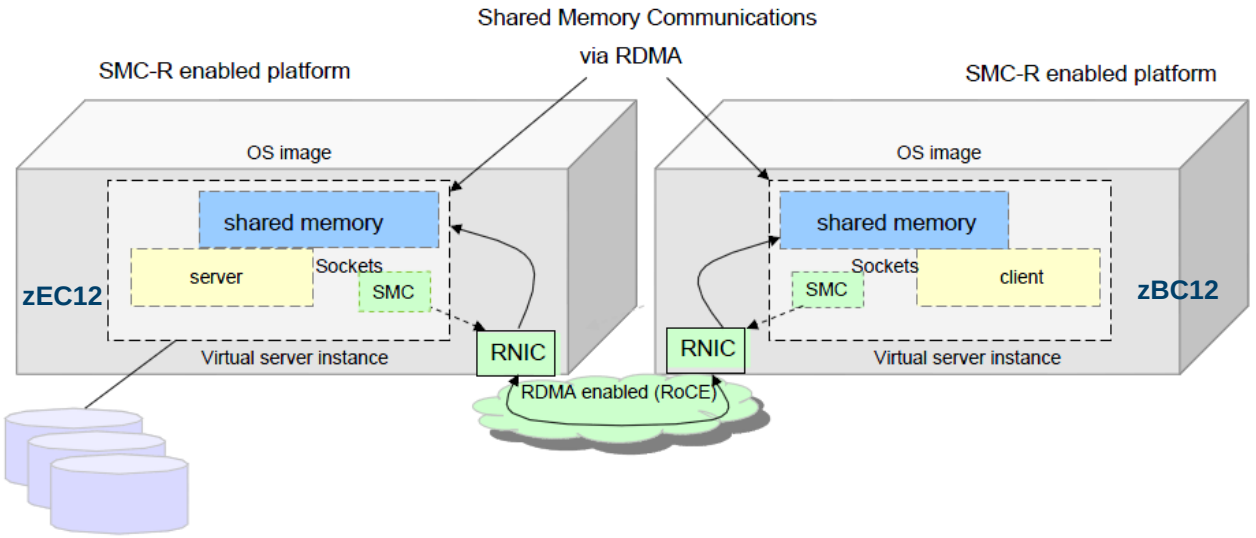
Intel Xeon E7-4870 + DS8300	z/OS + DS8800
40 processors 128 GB RAM Linux Sort	8 processors 128 GB RAM DFSORT
	
Sorting Average CPU 89%	Sorting Average CPU 72%
SORT *Job: Sort a 3 GB transaction file – Repetitions: 300*	
Total Elapsed 2,657 secs Bytes Per Sec **240 MB**	Total Elapsed 515 secs Bytes Per Sec **3,000 MB**
MERGE *Job: Merge 30 sorted files into a 90 GB master file – Repetitions: 10*	
Total Elapsed 4,051 secs Bytes Per Sec **157 MB**	Total Elapsed 446 secs Bytes Per Sec **3,460 MB**

13x more I/O bandwidth than Intel

Intel Batch window is 7x longer than z/OS

Source: IBM Internal Study. Results may vary based on customer workload profiles/characteristics.

IBM continues to innovate with new PCIe features – Shared Memory Communications (SMC-R) introduced in 2013



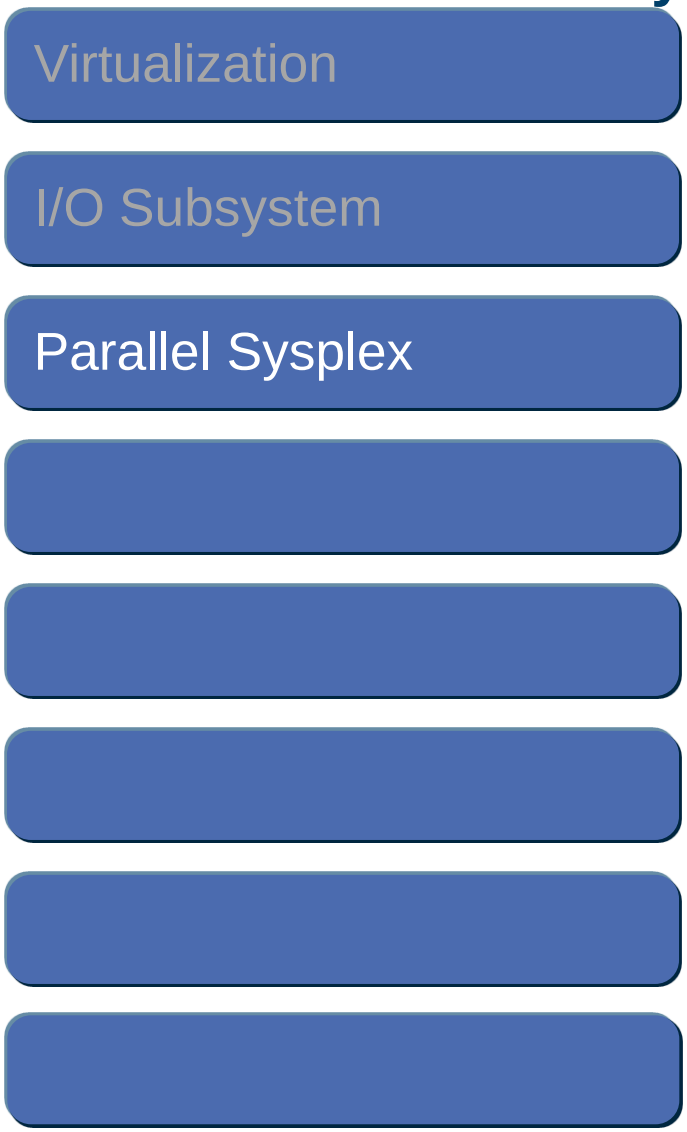
Network latency reduced up to

*80%**

- 10GbE RDMA over Converged Ethernet (RoCE) Express card
- Helps reduce latency and CPU resource consumption
- Runs over TCP/IP across z/OS systems
- Can be used seamlessly by *any* z/OS TCP sockets-based without any changes

* Based on internal IBM benchmarks of modeled z/OS TCP sockets-based workloads with request/response traffic patterns using SMC-R vs. TCP/IP. The actual throughput that any user will experience will vary.

Parallel sysplex gives zEnterprise continuous availability with near linear scalability



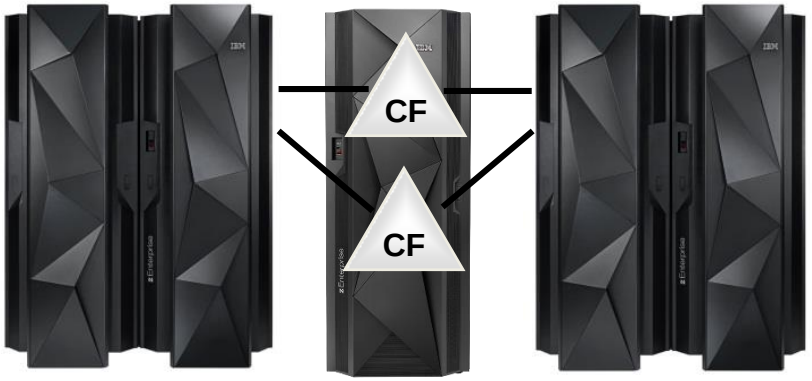
- Optimized to support IBM middleware
- Provides a single image across the cluster
- Centralized design optimizes data sharing
- Enables near-infinite elasticity

zEnterprise parallel sysplex clusters provide unmatched processing power and availability

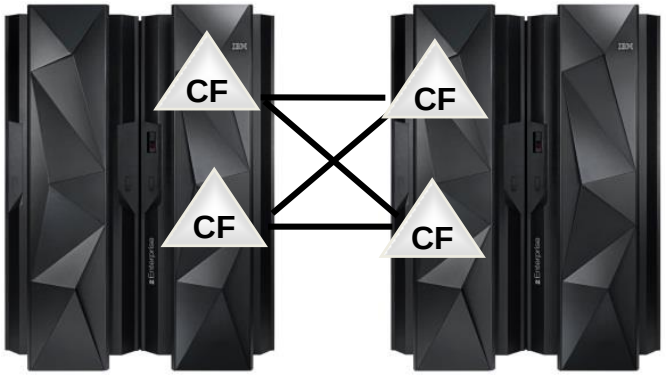
Single System Sysplex

- Clustering driven by specialty engines (Coupling Facility)
- Presents a single system image of a z/OS workload
- Potentially **2.5M MIPS** per 32-way cluster*

*Equivalent to about 240 of the largest Oracle servers



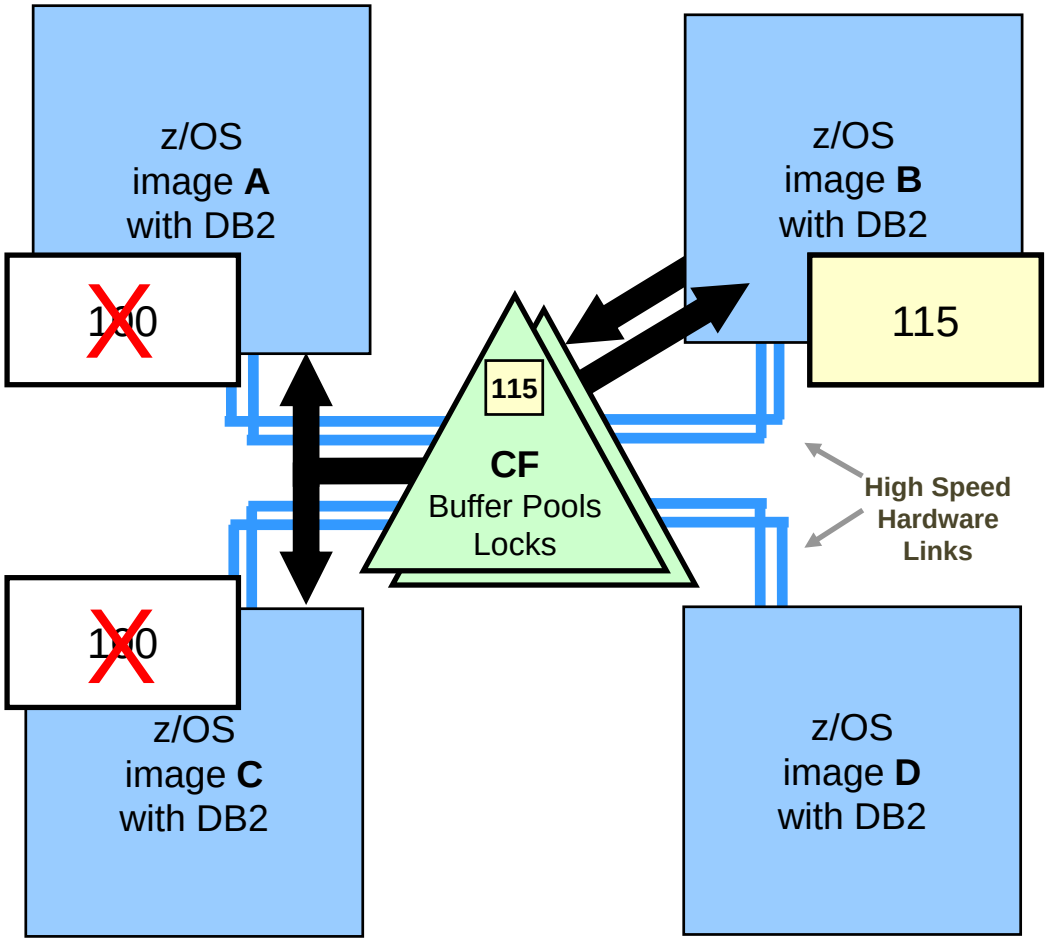
External Coupling Facility
(Can be different class server)



Cross Connected Servers with internal Coupling Facilities

- Enables rolling updates
- Supports continuous access to business services and data – from anywhere, at anytime
- Designed for **99.999%** availability

zEnterprise's centralized Coupling Facility permits efficient lock and cache management in DB2



A and C have data in local buffer pool without locks

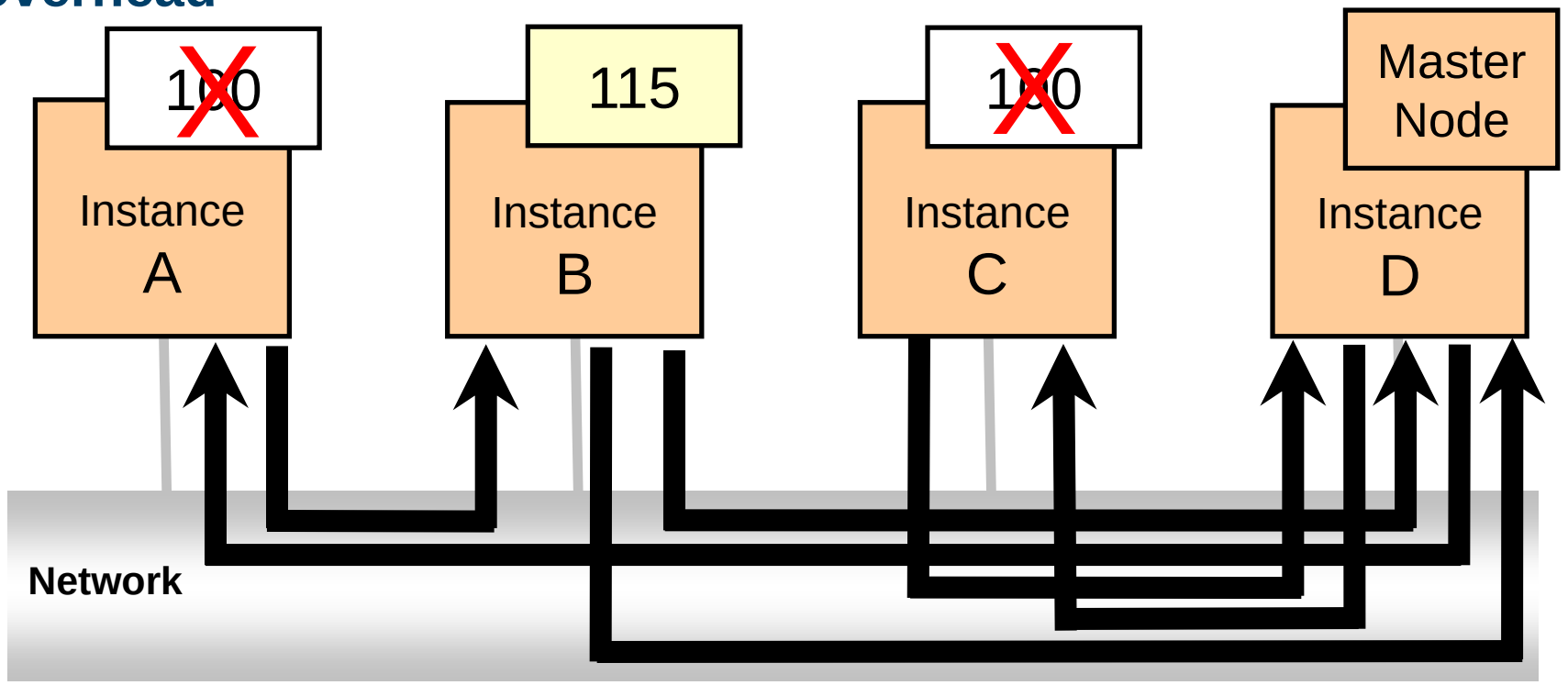
1. B registers page to CF and obtains write lock
2. B updates data
3. B commits update

B caches update in group buffer pool

CF invalidates all cached copies without interrupting processors

Cache and locks are maintained with no inter-node disturbance!

Oracle RAC's distributed lock management design causes overhead



Lock assume

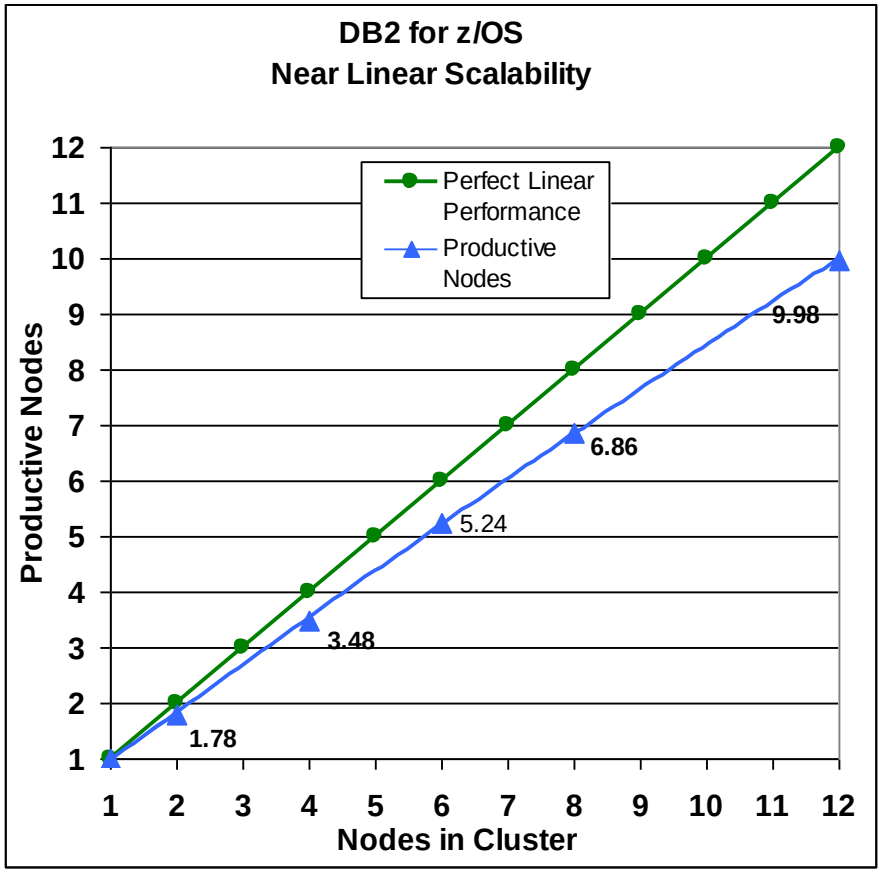
7. B updates local copy

Inter-node connections: **6**

In a cluster with 4 nodes, an update operation may need 6 network connections and two in-memory calls (not shown).

Example based on Oracle's US Patent 7,107,319 B2.

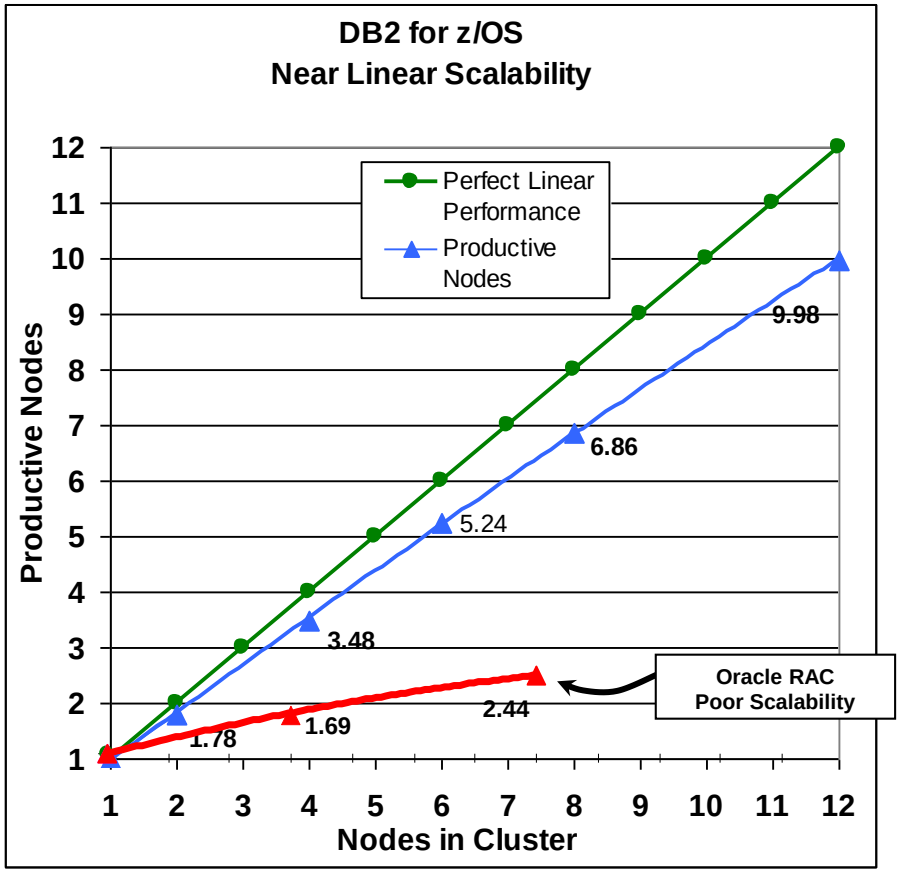
DB2 for z/OS in a parallel sysplex scales efficiently and transparently



- DB2 leverages unique Parallel Sysplex clustering design to achieve near linear scaling
  - No data partitioning required
  - No transaction routing required
  - No cluster awareness required in applications
- Elastic processing capacity
  - Applications are not tied to database partitioning schemes
  - Automatically balances workload across cluster

DB2 for z/OS OLTP result (ITG '03)

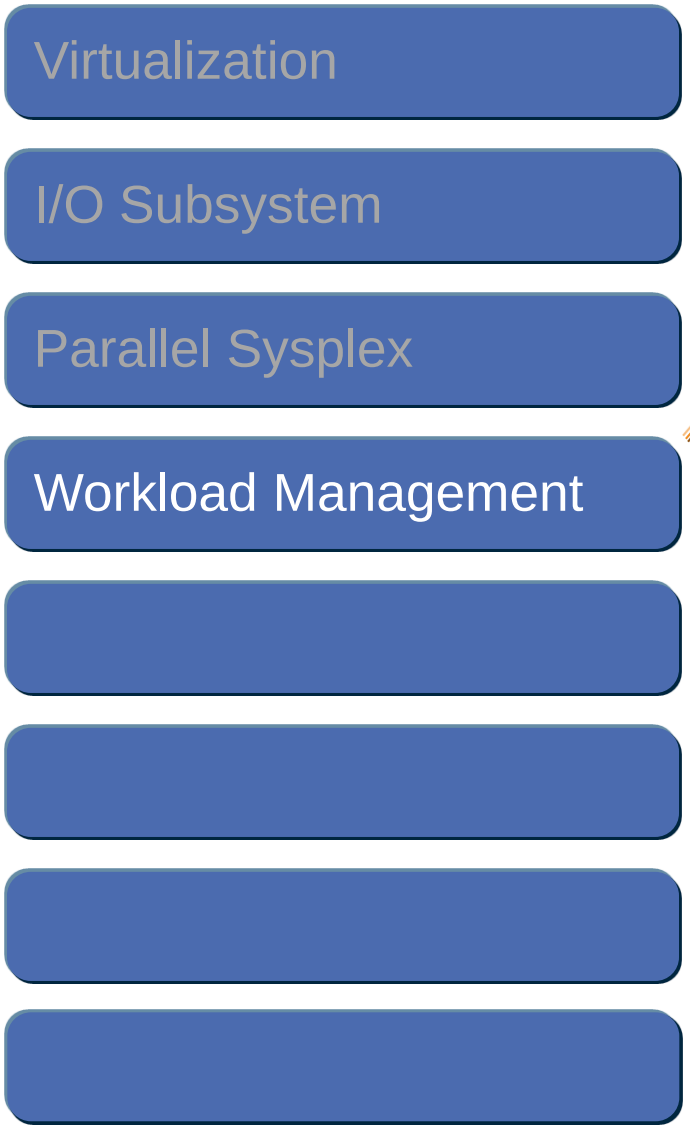
The only option for Intel-based servers is Oracle RAC



- Oracle RAC's lock and cache system is inefficient by design
  - Scaling RAC requires complex tuning and partitioning
  - Application partition awareness makes it difficult to add or remove nodes
- Published studies demonstrate difficult or poor scalability
  - Dell (shown in chart): Poor scalability despite using InfiniBand for RAC interconnect
  - CERN: Four month team effort to tune RAC, change database, change application
  - Insight Technology: Even a simple application on two node RAC requires complex tuning and partitioning to scale

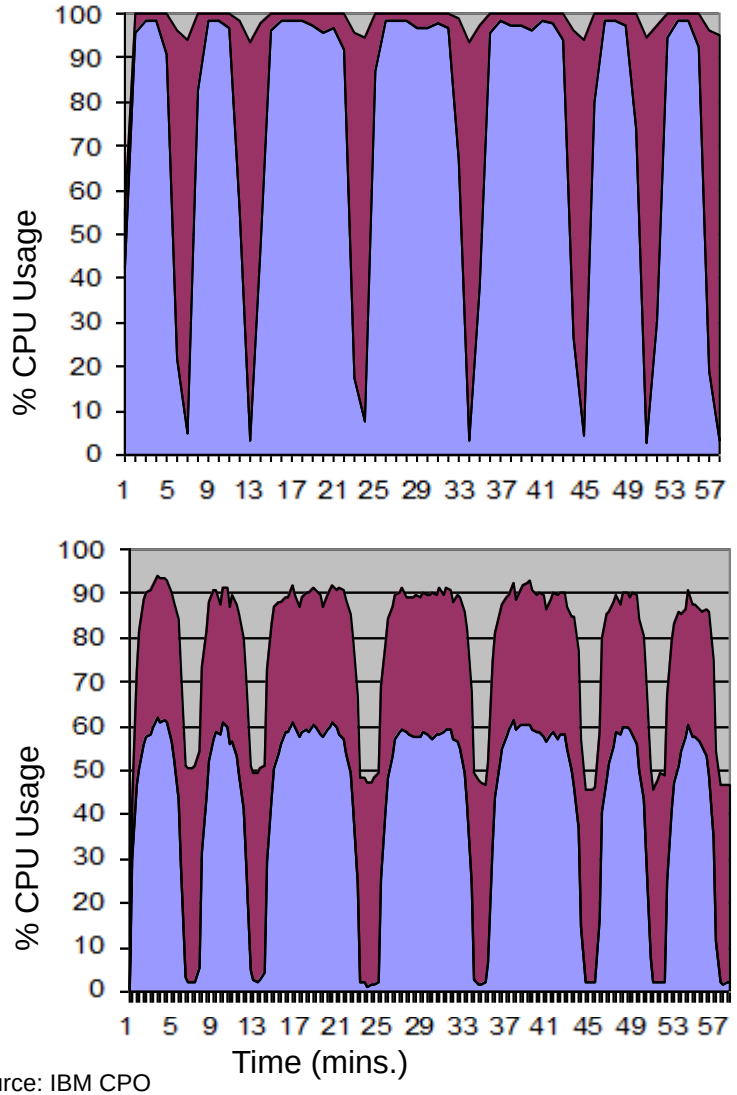
Oracle RAC characteristics as shown in Dell RAC InfiniBand Study http://www.dell.com/downloads/global/power/ps2q07-20070279-Mahmood.pdf
CERN (European Organization for Nuclear Research) http://www.oracleracsig.org/pls/apex/RAC_SIG.download_my_file?p_file=1001900
Insight Technology http://www.insight-tec.com/en/mailmagazine/vol136.html

The zEnterprise demonstrates “perfect” workload management



- Applies across all resources, not just CPU
- Ensures priority workloads meet service level agreements
- Cross platform
- Covers heterogeneous platforms

System z virtualization enables mixing of high and low priority workloads without penalty



System z

- No degradation of high priority workloads
- Low priority workloads use up all but 2% of available resources (high utilization)
- **Result**: Consolidate workloads of different priorities on the same platform and still meet service level agreement

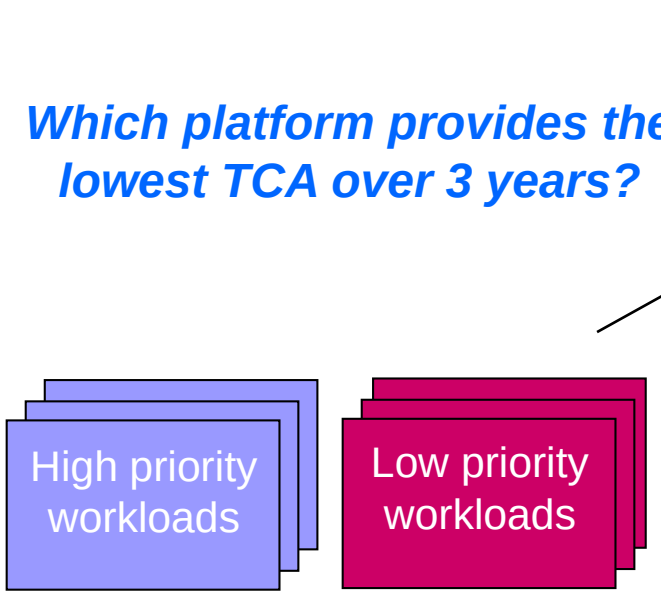
Common Intel hypervisor

- High priority workloads degrade when low priority workloads were added
- Low priority workloads used too much resources, and overall CPU utilization was not nearly as high
- **Result**: Inefficient, unreliable workload management means separate servers are required to insure service level agreement is met

Source: IBM CPO

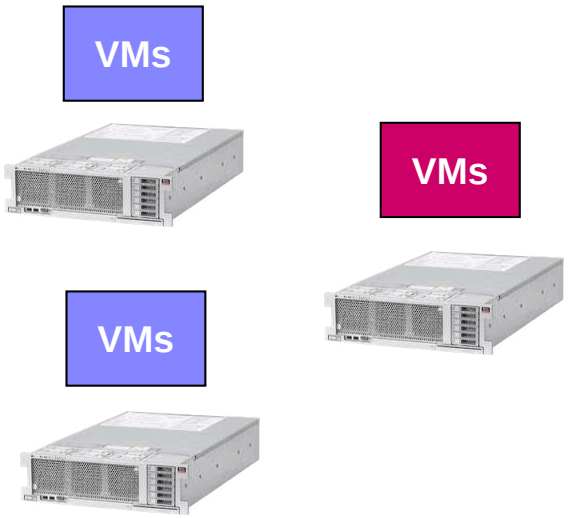
zEnterprise easily manages mixed priority workloads and lowers costs

Which platform provides the lowest TCA over 3 years?



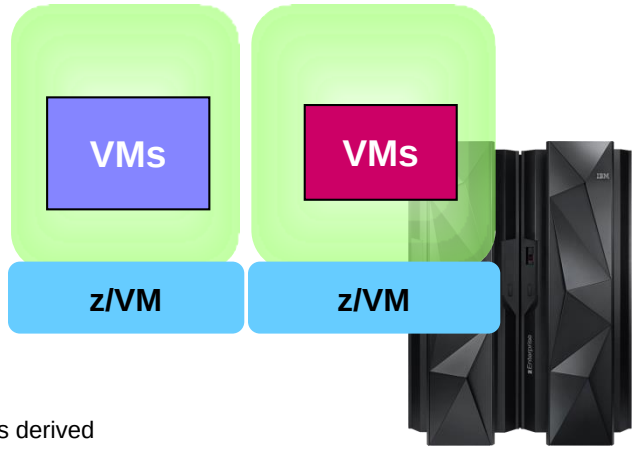
- **IBM WebSphere 8.5 ND**
- **IBM DB2 10 AESE**
- **Monitoring software**

High priority online banking workloads driving a total of **9.1M** transactions per hour and low priority discretionary workloads driving **2.8M** transactions per hour



Virtualized on 3 Intel 40 core servers

$15.9M (3 yr. TCA)



z/VM on zEC12
32 IFLs

$6.5M (3 yr. TCA)

59% *lower cost!*

Consolidation ratios derived from IBM internal studies.. zEC12 numbers derived from measurements on z196. Results may vary based on customer workload profiles/characteristics. Prices will vary by country.

Only zEnterprise offers numerous options for optimizing workloads to reduce costs

- Virtualization
- I/O Subsystem
- Parallel Sysplex
- Workload Management
- Workload Optimization
-
-
-

- Specific workloads can be moved to custom hardware
- Hybrid workloads supported by multi-architecture platform
- Reduces costs and improves price/performance ratio

Workload optimizations are achieved via special I/O cards

zEnterprise Data Compression (zEDC) introduced in 2013

- Compatible with current coprocessor-based compression
- Specifically designed for large amounts of bulk data
- Cost effective – reduces CPU overhead, and storage overhead
- Optimizes cross-platform exchanges
  - Compatible with zlib compression – an industry standard widely used across all platforms

Up to **4x** data compression

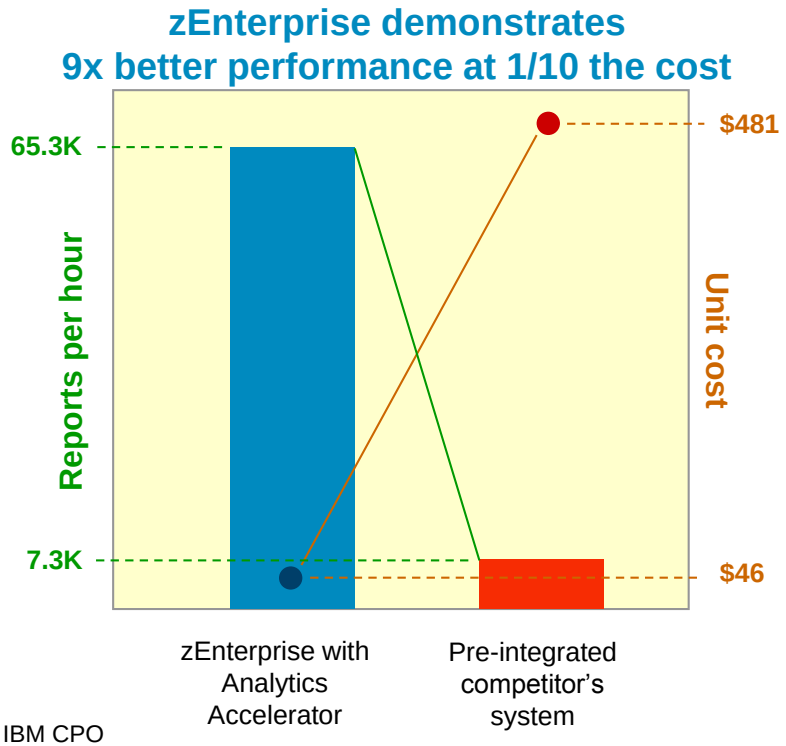
Up to **118x** reduction in CPU

Up to **24x** throughput improvement with zlib

Source: IBM

IBM DB2 Analytics Accelerator speeds up deep analytics queries

- A workload-optimized, blade-based appliance that runs queries in seconds versus hours
- Integrated with DB2 for z/OS, and transparent to applications
- Drives down the costs of data warehousing and business analytics



Source: IBM CPO

zEnterprise extends to support hybrid computing

zEnterprise BladeCenter Extension (zBX) and Unified Resource Manager (URM)

- Industry's first multi-architecture platform
  - zBX includes Power, x86, and accelerator blades
- URM extends System z governance extended to zBX blades
  - Provides resource and workload management across mainframe and blades
- Supports application integration with Microsoft Windows, Linux and AIX
- Greater opportunities for consolidation and simplification
- Consistent business controls across applications and platforms

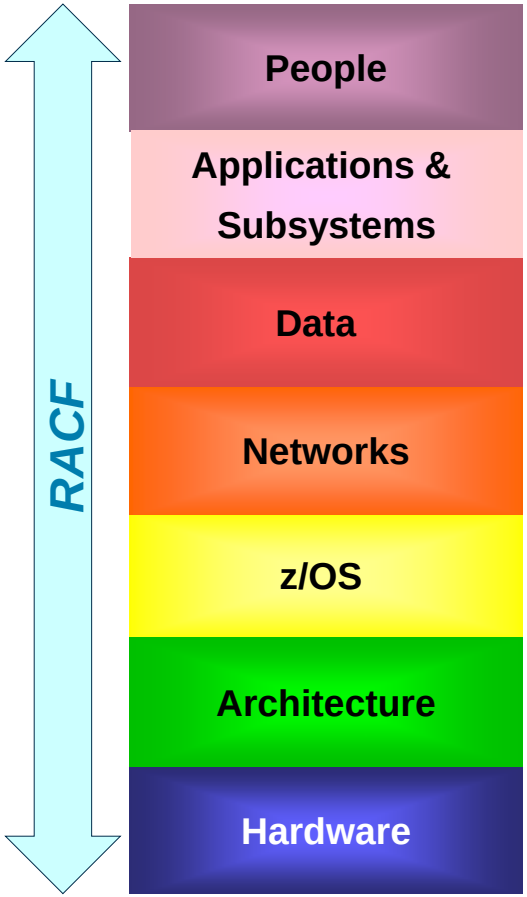
zEnterprise BladeCenter Extension

zEnterprise – the *most* secure commercially available platform

- Virtualization
- I/O Subsystem
- Parallel Sysplex
- Workload Management
- Workload Optimization
- Security
-
-

- Highest commercially available EAL ratings
- Multiple encryption options
- Provides full function Public Key certificate authority
- APIs extend encryption services across the enterprise
- State of the art security monitoring

Remote Access Control Facility (RACF) provides security throughout the entire zEnterprise stack



- Tools, reporting, auditing
- Access control to all classes of resources
- Integrated into the operating system
- Provides Enterprise Identity Management
- Supports cryptographic services
- Supports digital certificates

Virtualized System z security is superior to other platforms and augmentation costs less

Security Natively Covered by Platform

Security Level Description	IBM System z	x86	Competative UNIX
Normal corporate	100.00%	18.16%	30.26%
Credit card processing involved	99.00%	11.04%	18.28%
Banking	94.00%	5.26%	10.22%
Healthcare	100.00%	3.24%	8.51%
Research	92.50%	2.86%	4.16%
Defense	85.54%	0.26%	1.86%

Major security deficiencies on distributed platforms

Distributed platforms require ***considerable additional expense***

On System z most security requirements are standard

Incremental Cost to Achieve Required Security

Security Level Description	IBM System z	x86	Competative UNIX
Normal corporate	0.00%	32.54%	12.37%
Credit card processing involved	2.32%	46.27%	29.53%
Banking	2.07%	51.31%	26.58%
Healthcare	0.00%	67.26%	35.89%
Research	4.28%	91.26%	64.28%
Defense	11.36%	125.41%	102.26%

Little additional augmentation required on System z

Source: "Tracked, Hacked and Attacked?"

zEnterprise's reliability, availability and serviceability are legendary

- Virtualization
- I/O Subsystem
- Parallel Sysplex
- Workload Management
- Workload Optimization
- Security
- Reliability

- Comprehensive, multi-layered strategy for reliability and serviceability
- Supports large number of concurrent operations during maintenance
- "Five 9s" availability
- Lowest costs

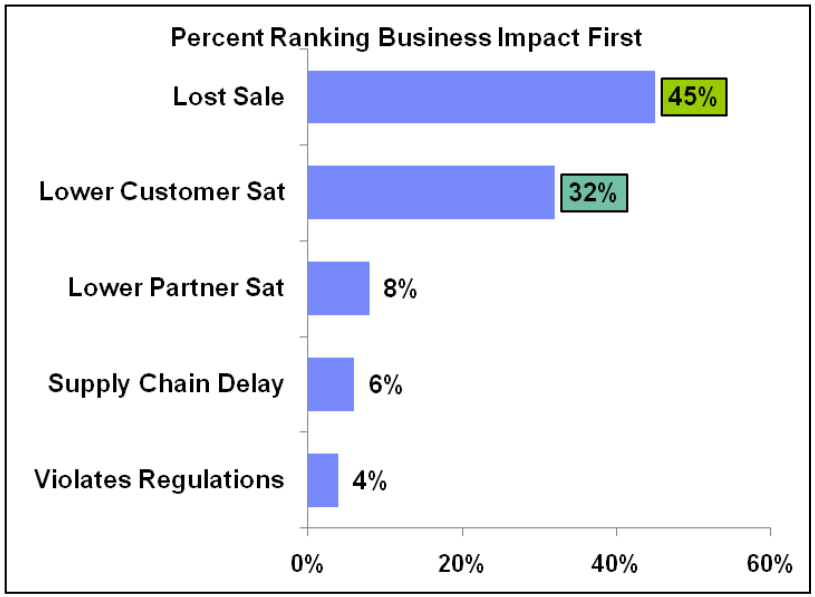
zEnterprise supports concurrent operations during maintenance

Capability	zEC12	x86
ECC on Memory Control Circuitry	**Transparent While Running**	Can recognize/repair soft errors while running; limited ability with hard errors
Oscillator Failure	**Transparent While Running**	**Must bring server down to replace**
Core Sparing	**Transparent While Running**	**Must bring server down to replace**
Microcode Driver Updates	**While Running**	Some OS-level drivers can update while running, not firmware drivers; reboot often required
Book Additions, Replacement	**While Running**	**Must bring server down**
Memory Replacement	**While Running**	**Must bring server down**
Memory Bus Adaptor Replacement	**While Running**	**Must bring server down**
I/O Upgrades	**While Running**	**Must bring server down to replace (limited ability to replace I/O in some servers)**
Concurrent Driver Maintenance	**While Running**	**Limited – some drivers replaceable while running**
Redundant Service Element	**2 per System**	"Support processors" can act as poor man's SE, but no redundancy

Single book systems may not support concurrent memory upgrades

Downtime seriously effects sales, revenue, customer satisfaction

Business Impact of 10 Minutes Of Downtime



Source: IBM Customer Survey

Revenue Impact of Downtime Per Hour

Figure 1 Cost of downtime by industry segment

Average = $2.7M

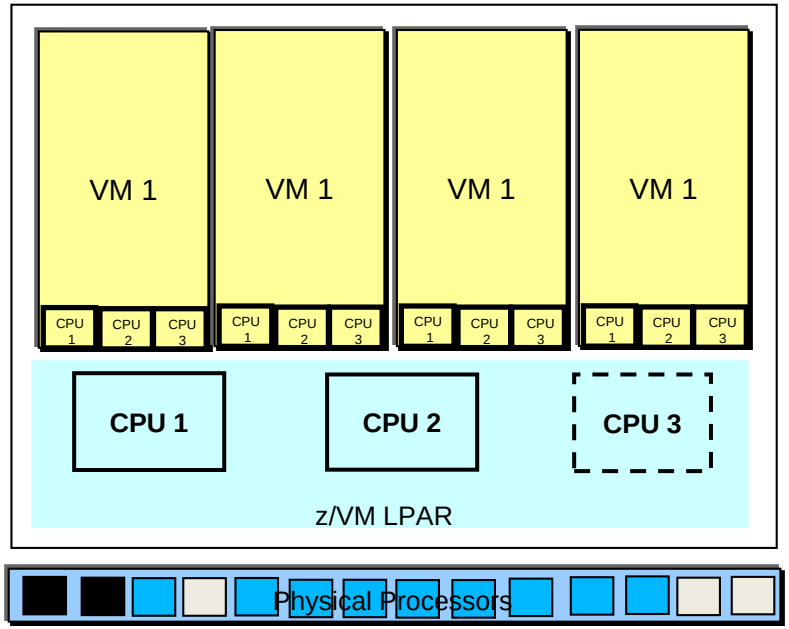
Industry/Sector	Revenue/Hour
Energy	$1,468,798
Telecommunications	$4,611,604
Financial	$8,213,470
Information Technology	$3,316,058
Insurance	$2,582,382
Pharmaceuticals	$2,058,710
Banking	$1,145,129
Consumer Products	$989,795
Chemicals	$1,071,404
Transportation	$1,463,128

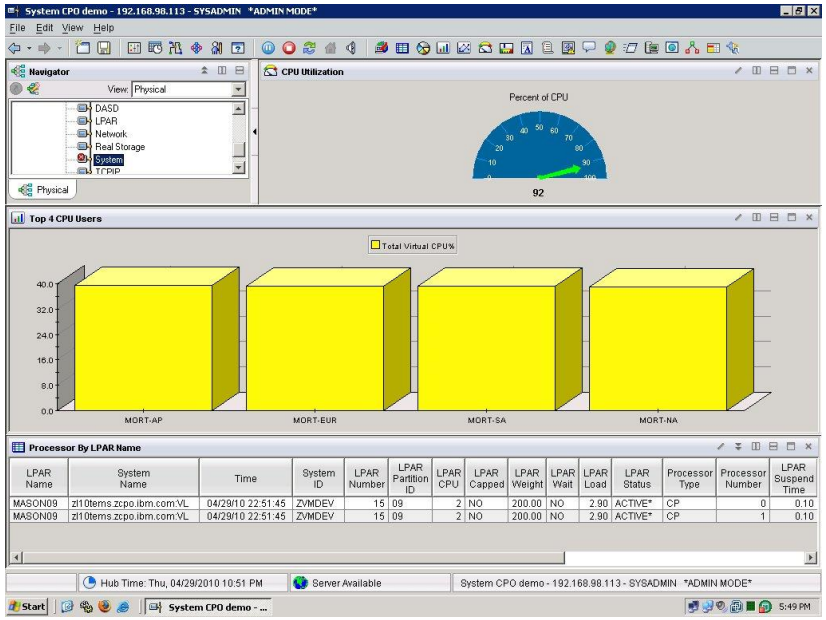
Source: Robert Frances Group 2006

Profit Impact of Downtime

A Telco	%	Profit 2009	Profit/Hr	Profit/Min
Wireless	68%	$3,000,000,000	$342,466	$5,708
Cable	29%	$1,300,000,000	$148,402	$2,473
Media	3%	$120,000,000	$13,699	$228
Total	**100%**	**$4,420,000,000**	**$504,566**	**$8,409**

DEMO: Dynamically add processing capacity to z/VM LPAR to handle increased workload… *without disruption*





Tivoli Enterprise Portal

- Guest VMs run without disruption
- Dynamically add logical processors to z/VM LPAR
- Dynamically add processors shared among LPARs

Today's mainframe – 50 years of *continuous* innovation and decades more to come

- Virtualization
- I/O Subsystem
- Parallel Sysplex
- Workload Management
- Workload Optimization
- Security
- Reliability
- Multi-tenancy OS

- Efficient use of all resources
- Faster response to load
- Easy to manage
- Ability to specify Service Level
- Reduced TCO

Other OSs are rarely used for multi-tenancy and lack the pressure to innovation

- PCs, Windows, Unix originated as single-use systems
  - Multi-tenancy is usually achieved through multiple virtual machines
  - VMware, KVM, Hyper V … are add-ons

- Virtual machine multi-tenancy imposes an extra layer of control – the hypervisor can't communicate directly with the application

- Result: z/OS more efficient than virtual machine clusters
  - Virtual memory management (e.g. 'ballooning')
  - Workload management, including I/O prioritization
  - Ubiquitous shared services and resource pools
  - In-memory communication between cooperating programs (vs VLAN/IP stack)
  - Address spaces
  - Central security
  - ….

zEnterprise efficiency – lower cost for multi-tenant database workloads

Which platform can achieve the lowest cost per workload?

I/O Intensive Database Workload

Brokerage High Volume Trading workload driving a minimum* of **243** transactions per second on 200GB database (using SSD storage)

1 workload on 16-core quarter unit

Pre-integrated DB Competitor V2 Multi-Tenant Private Cloud

$2.27M
per workload

5 multi-tenant workloads on zBC12

4 GPs + 3 zIIPs

DB2 10 for z/OS on zBC12

$1.70M
per workload

25% *lower cost*

* Maximum TPS was measured at 270 based on 70 ms injection interval for customer threads. SLA requires no more than 10% degradation in throughput, yielding a Minimum TPS of 243.

Today's mainframe – 50 years of *continuous* innovation and decades more to come

- Virtualization
- I/O Subsystem
- Parallel Sysplex
- Workload Management
- Workload Optimization
- Security
- Reliability
- Multi-tenancy OS

IBM zEnterprise EC12

The basis for many new opportunities for mainframe workloads


thank
you!