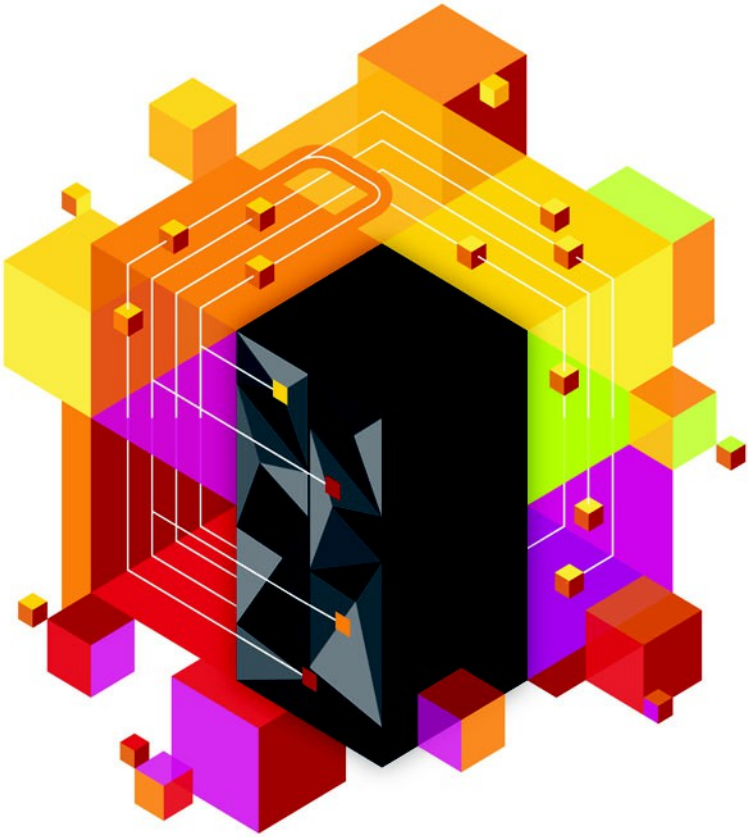
Make The Right Platform Decision to Grow Your Business

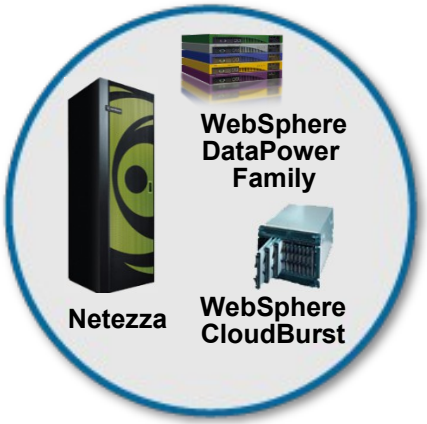


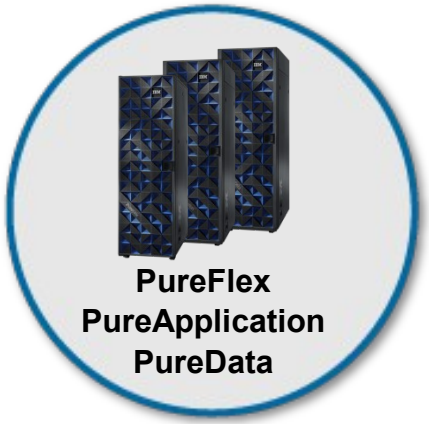
John J Thomas

IBM Competitive Project Office

IBM Offers A Broad Range Of Systems To Meet Your Workload Needs

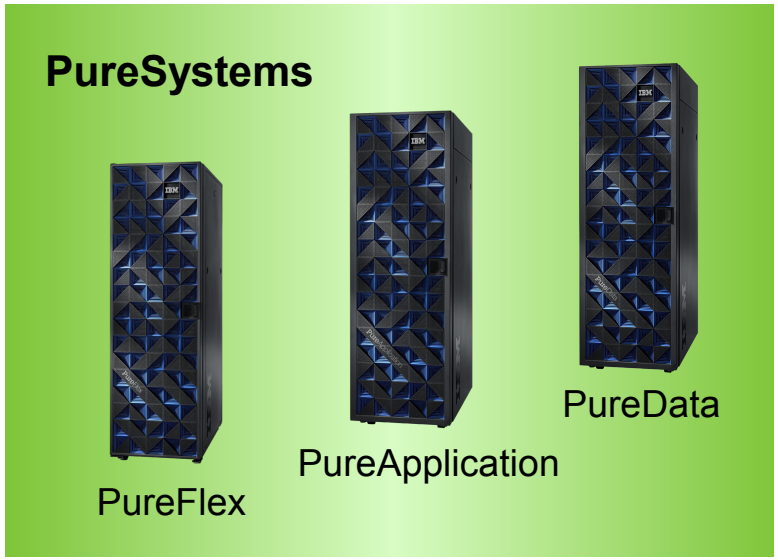
More capability and choice than the competition





Appliances	IBM PureSystems	IBM Power Systems	IBM zEnterprise
Integrated function	Blade economics	Optimized performance	Global scale transaction processing
Simple Setup	Flexible choice	Direct attached SSD	Critical data
	Integrated expertise		

PureSystems



- Optimized for blade workloads
- Simplify and speedup delivery of blade infrastructure
- Expert integrated systems

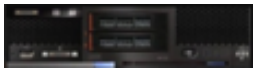
3

IBM Flex System Goes Beyond Blades

Building Blocks: IBM Flex System components

Chassis
14 half-wide
bays for nodes

Compute Nodes
Power 2S/4S
x86 2S/4S



Storage Node
V7000 (optional)

Management Appliance (Optional)

Networking
10/40GbE, FCoE,
IB 8/16Gb FC

Expansion
PCIe, Storage

Build to Order
(Choice of Compute Node, Storage and Networking)

Flex System

- Flexible choice
- Integrated design
- Pre-assembled hardware
- On-site set up services

IBM PureFlex System Simplifies Set-Up And Management

Built with Choice of Compute Nodes, Storage and Networking

Flex System

Custom Built (wide choice of components)

Pre-configured PureFlex System

- Factory integrated and pre-configured
- Built-in Patterns of Expertise (Infrastructure Patterns)
- Faster deployment and lower cost
- Includes cloud management

Express, Standard and Enterprise Configurations

IBM PureApplication System - *Optimized For Speed, Simplification, And Less Customer Labor*

IBM PureApplication System

Deep integration and optimization

Integrated Cloud Management

Self service provisioning, resource groups, automatic workload scaling, metering

+

Integrated Virtualization Management

Automated and policy driven pattern deployment

+

Web Application Middleware Platform

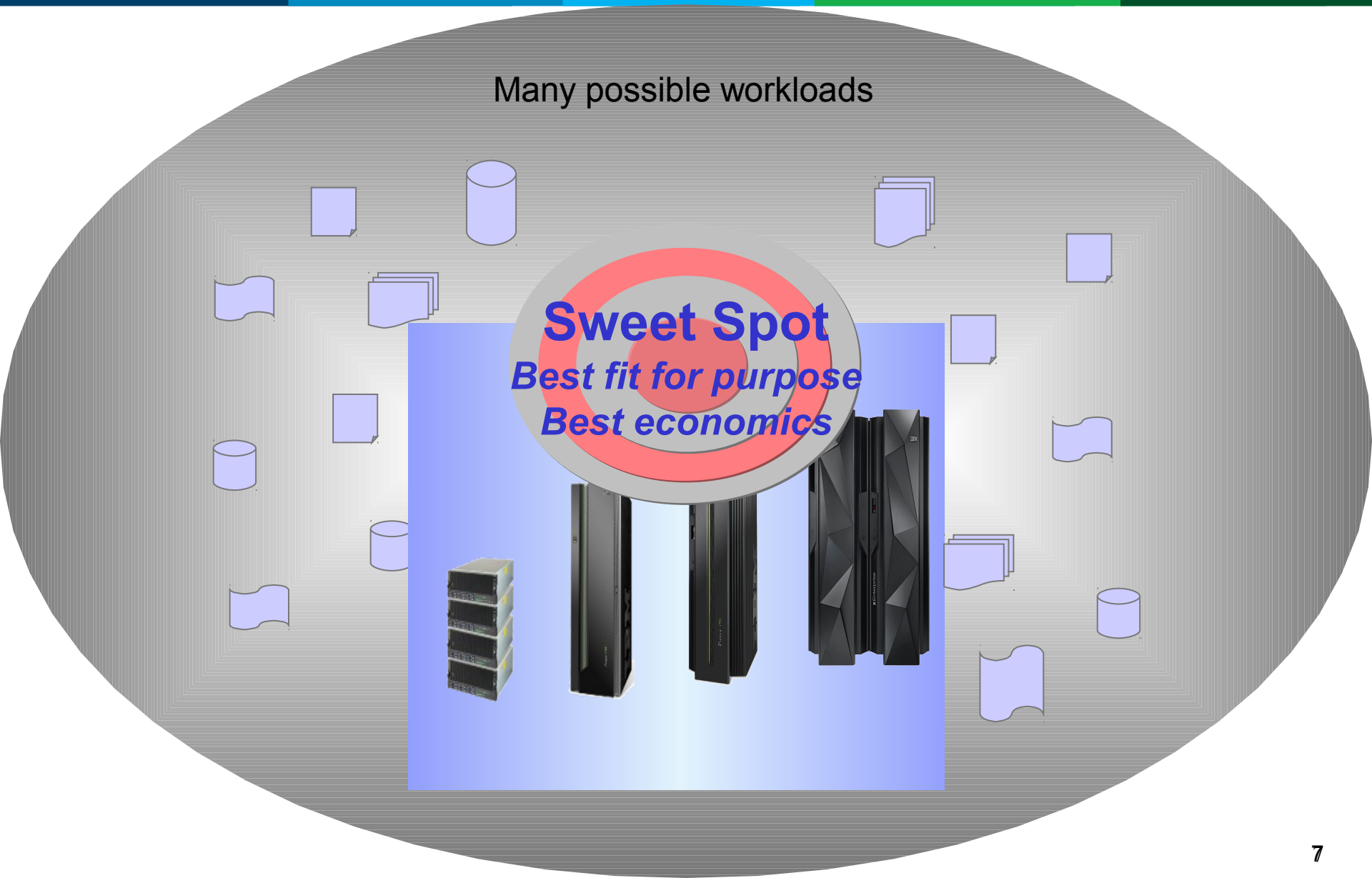
IBM WAS and DB2 licenses included

+

IBM PureFlex System Hardware

Compute + Network + Storage + Hardware Management

What Workloads Are Best Fit On High End Systems?



Sweet Spot Workloads

System z

- Global scale critical data workloads
- Transaction processing
- Batch processing
- Co-located analytics
- Consolidated on one platform

Power Systems

- Large critical data workloads
- CPU intensive and cache intensive processing applications
- Consolidated on one platform

What Makes System z Optimum For These Workloads?

- Concentrated processing power in a single complex
- Dedicated I/O sub-system with large scale I/O bandwidth
- DS8000 storage systems capacity and performance
- DB2 Analytics Accelerator facilitates co-located analytics
- “Perfect” workload management
- Better labor productivity
- Industry-leading RAS and security

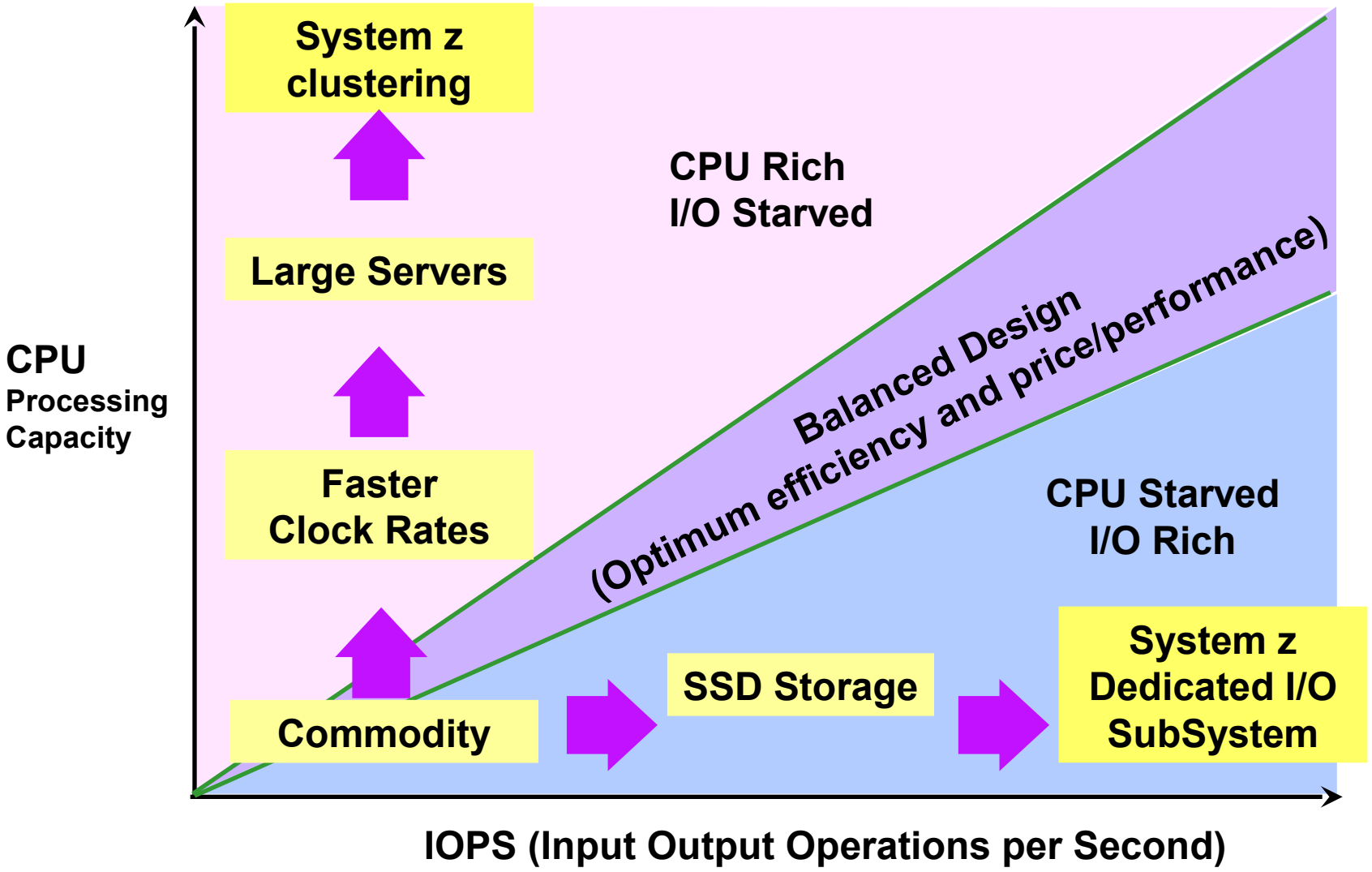
Result: Unbeatable Performance With Best Economics

What Makes Power Optimum For These Workloads?

- Large capacity servers with up to 1024 threads
- Cache structures optimized for larger working sets
- Bus attached SSD to match high processing capacity and performance
- DS8000 storage systems capacity and performance
- “Near Perfect” workload management
- Best-in-class RAS and security

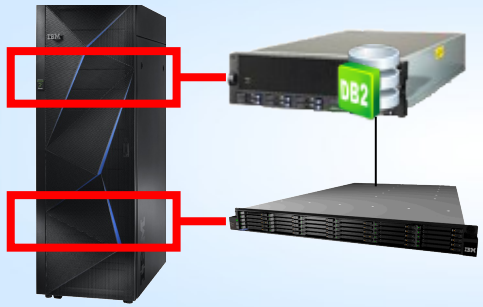
Result: Unbeatable Performance With Best Economics

IBM High End Systems Deliver Balanced Capabilities For Maximum Efficiency



Bus Attached SSD Helps Power Beat Pre-Integrated Database Competitor (Database Workload Classic)

IBM DB2 Advanced Enterprise Server Edition v10



2S x 8c = 16 cores
AIX 7.1, 64-bit

EXP30 SSD Drawer
25 SSDs

IBM Power 780+

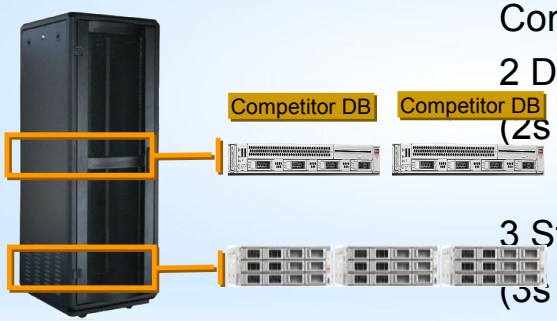
8,111 Transactions/sec

$285 per Trans/sec

✓ **3.4x** **Faster**

✓ **70%** **Lower cost per transaction**

Pre-integrated Database Competitor



Competitor Linux- ¼ Rack
2 Data Nodes
(2s x 8c = 16 cores)

3 Storage Nodes
(3s x 8c = 24 cores)

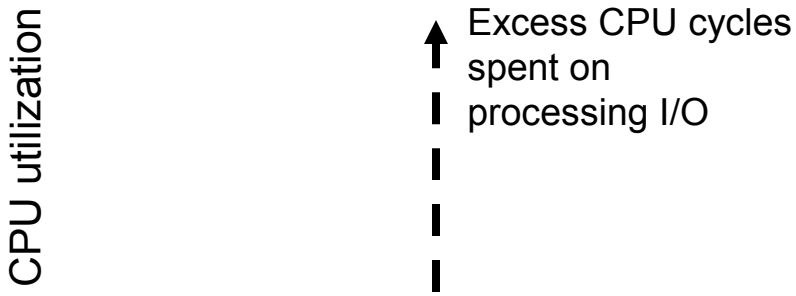
2,363 Transactions/sec

$959 per Trans/sec

This is an IBM internal study designed to replicate a typical IBM customer workload usage in the marketplace. The results were obtained under laboratory conditions, and not in an actual customer environment. IBM's internal workload studies are not benchmark applications, nor are they based on any benchmark standard. As such, customer applications, differences in the stack deployed, and other systems variations or testing conditions may produce different results and may vary based on actual configuration, applications, specific queries and other variables in a production environment.

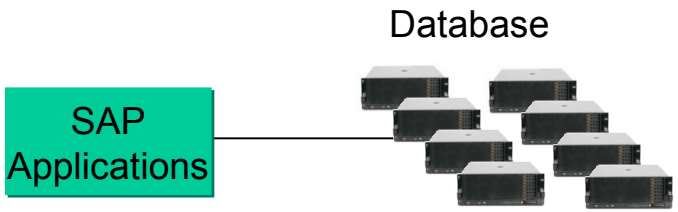
Intel Performance Degrades As I/O Demand Increases

- No dedicated I/O subsystem
- Test case scenario: Run multiple virtual machines on x86 server
  - Each virtual machine has an average I/O rate
  - x86 processor utilization is consumed as I/O rate increases



zEC12 Sysplex With Dedicated I/O Sub System Achieves World Record SAP Banking

Competitor DB on Intel



8x 3850 x5 with 16 cores
(dual active clusters)

128 DB cores

DB2 on z/OS

zEC12 2-way data sharing Sysplex

44 DB cores

14 CF cores

A world record at half the cost!

Database Unit Cost (5yr TCA)
$0.30/Postings per hour

Postings per Hour	42.0M
# of Accounts	90M
Hardware	$0.63M
Software	$11.98M
Total Cost	$12.61M

Database Unit Cost (5yr TCA)
$0.15/Postings per hour

Postings per Hour	59.1M
# of Accounts	150M
DB2 Solution Edition (HW+SW)	$7.49M
Capacity Backup (CBU)	$1.24M
Total Cost	$8.73M

Note: Cost of platform infrastructure for benchmark transaction production. Cost of packaged application software not included. List prices used.

DS8000 SSD Helps zEC12 Beat Pre-Integrated Competitor

Which platform can achieve the lowest cost per workload?

High Volume Trading Workload

I/O-intensive DB workloads, each driving a minimum* of **243** transactions per second on 200GB database

1 workload on **16**-core quarter unit

Pre-integrated Competitor Multi-Tenant Private Cloud

$2.27M per workload

5 multi-tenant workloads on zEC12 **4**-cores

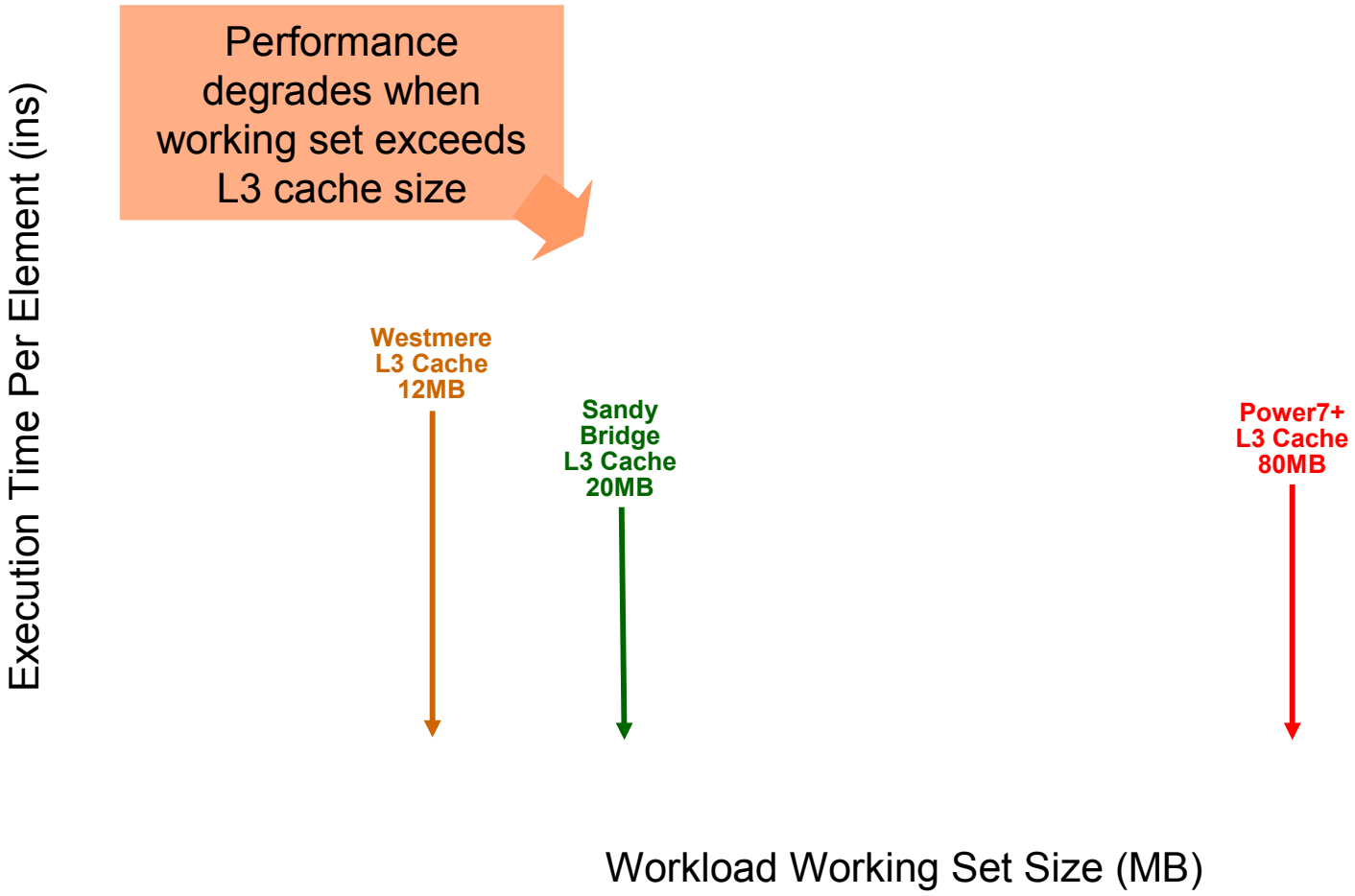
DB2 10 for z/OS on zEC12 with DS8000 SSD

$1.73M per workload

***20x** core density*
***25%** lower cost*

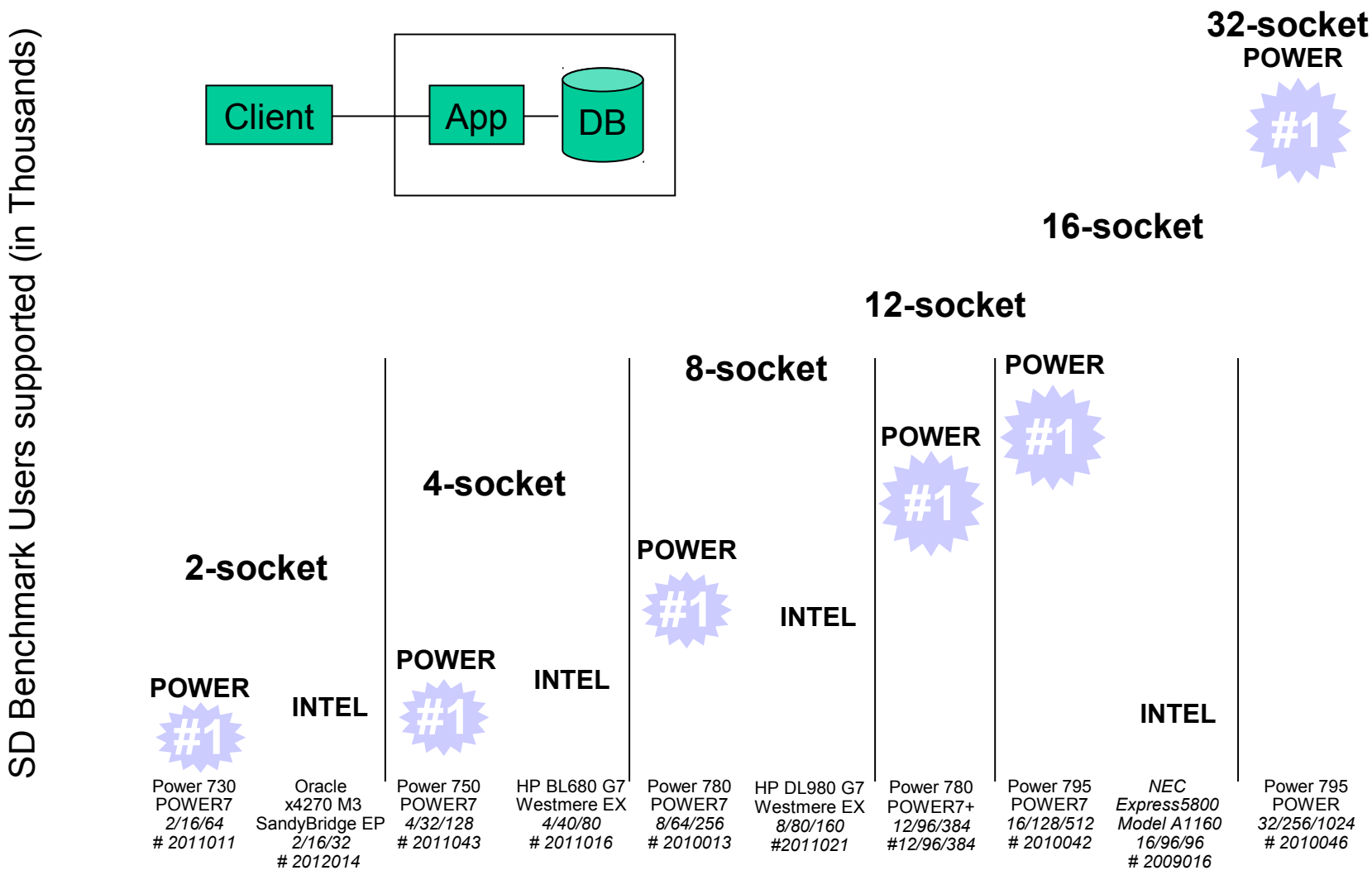
* Maximum TPS was measured at 270 based on 70 ms injection interval for customer threads. SLA requires no more than 10% degradation in throughput, yielding a minimum TPS of 243

Larger Power L3 On-Chip Cache Supports Workloads With Larger Working Sets



Test case: Addition operation to each element in array with defined size

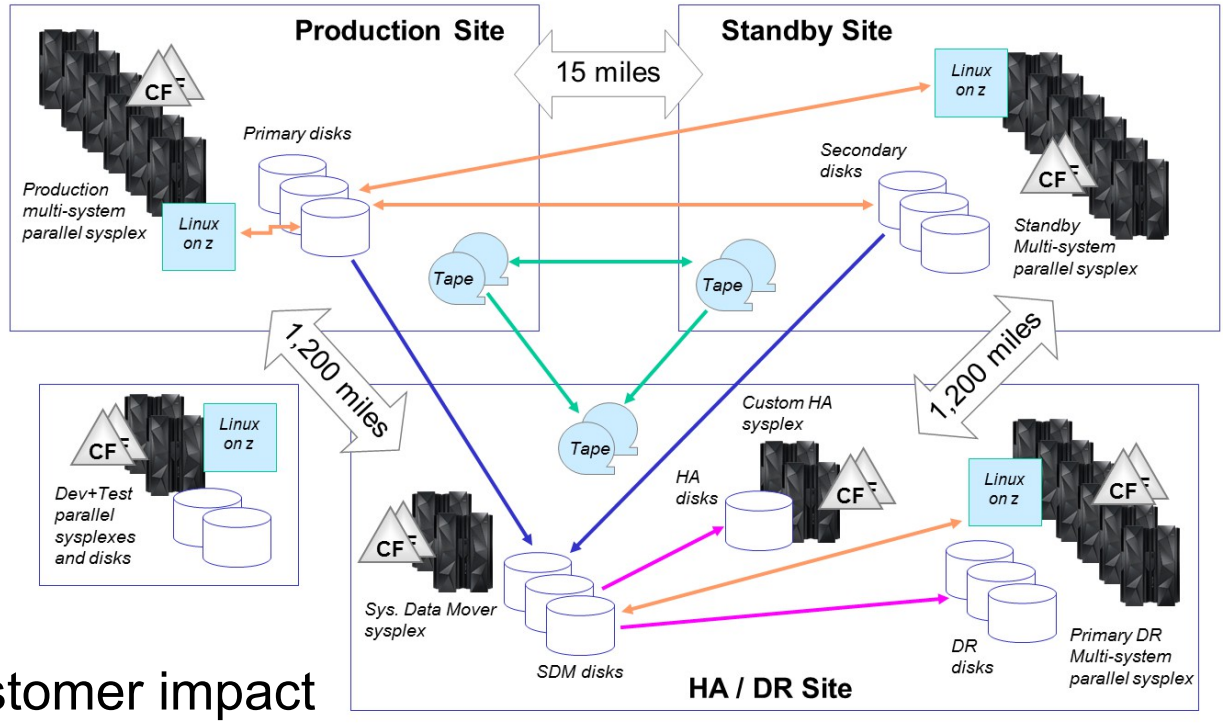
More Threads And Larger Cache Help Power Beat Int
In SAP Benchmark (S&D 2-Tier)



Configuration and results are on the two-tier SAP SD standard application benchmark running SAP enhancement package 4 for the SAP ERP 6.0 application (Unicode):
Power result is with DB2 9.7 database and HP server is with MaxDB 7.8 database. The numbers below the server shows no of processors / no of cores / no of threads and SAP certificate number.
Results valid as of 02/15/2012. Source: http://www.sap.com/benchmark

Global Scale Transaction Processing With System z

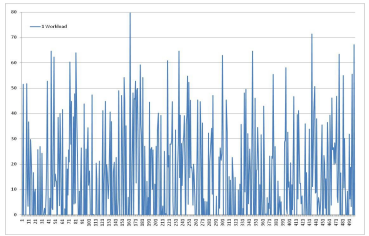
- 1B CICS trans/day
- 4,000 IMS trans/sec
- 14M ACH transactions in 2.5 hours
  - 30ms response
- Production site
  - 6 mainframes
  - 6 way sysplex
  - 216 CPUs, 200K MIPS



- Zero outages, zero customer impact
- Linux is Active-Active in the two data centers, with zero downtime
  - 15% Linux, growing at 30%
- *"Crazy about security overall, and the z system has a fortress around it"*

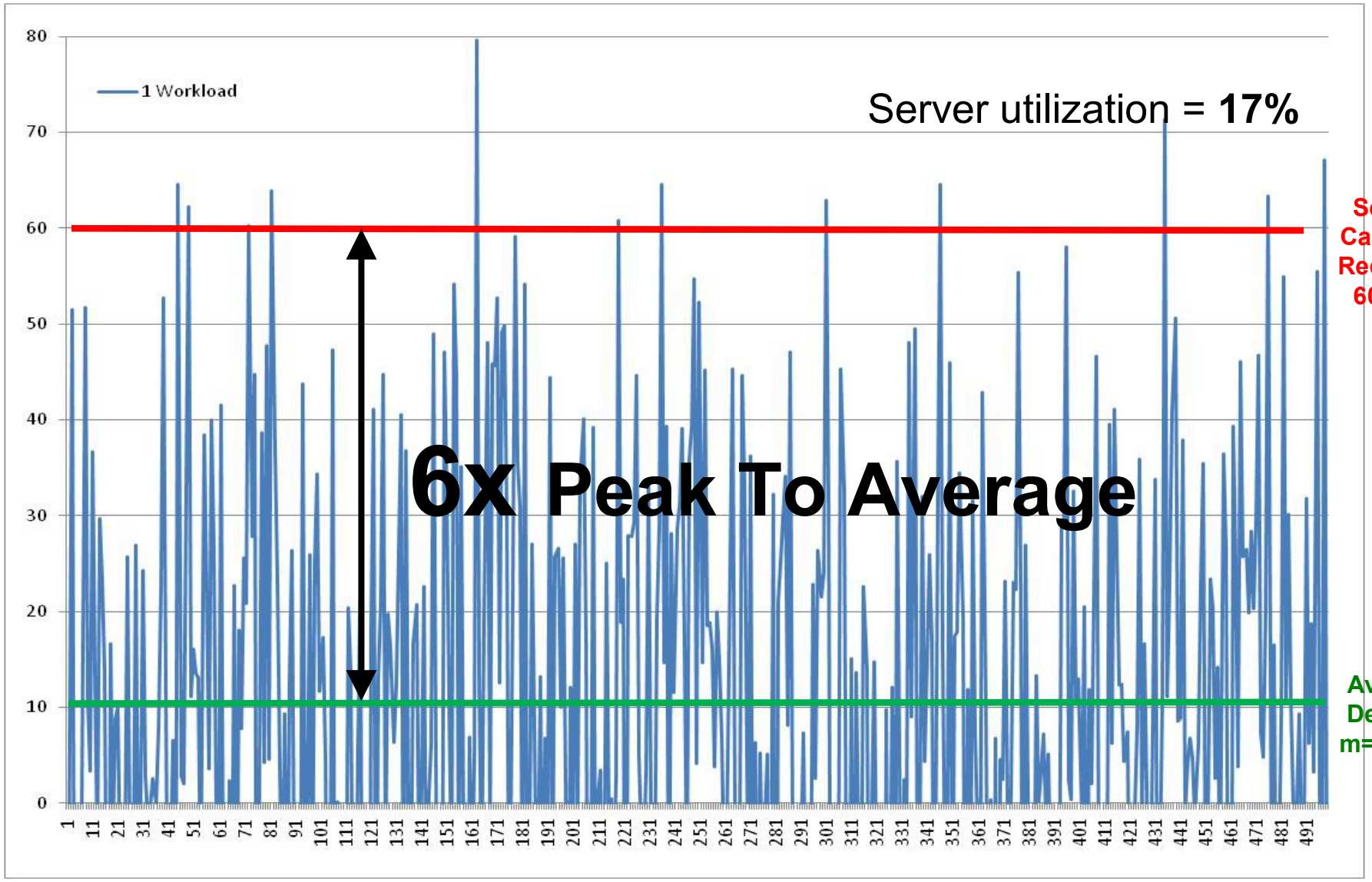
Larger Servers With More Resources Make More Effective Consolidation Platforms

- Most workloads experience variance in demand



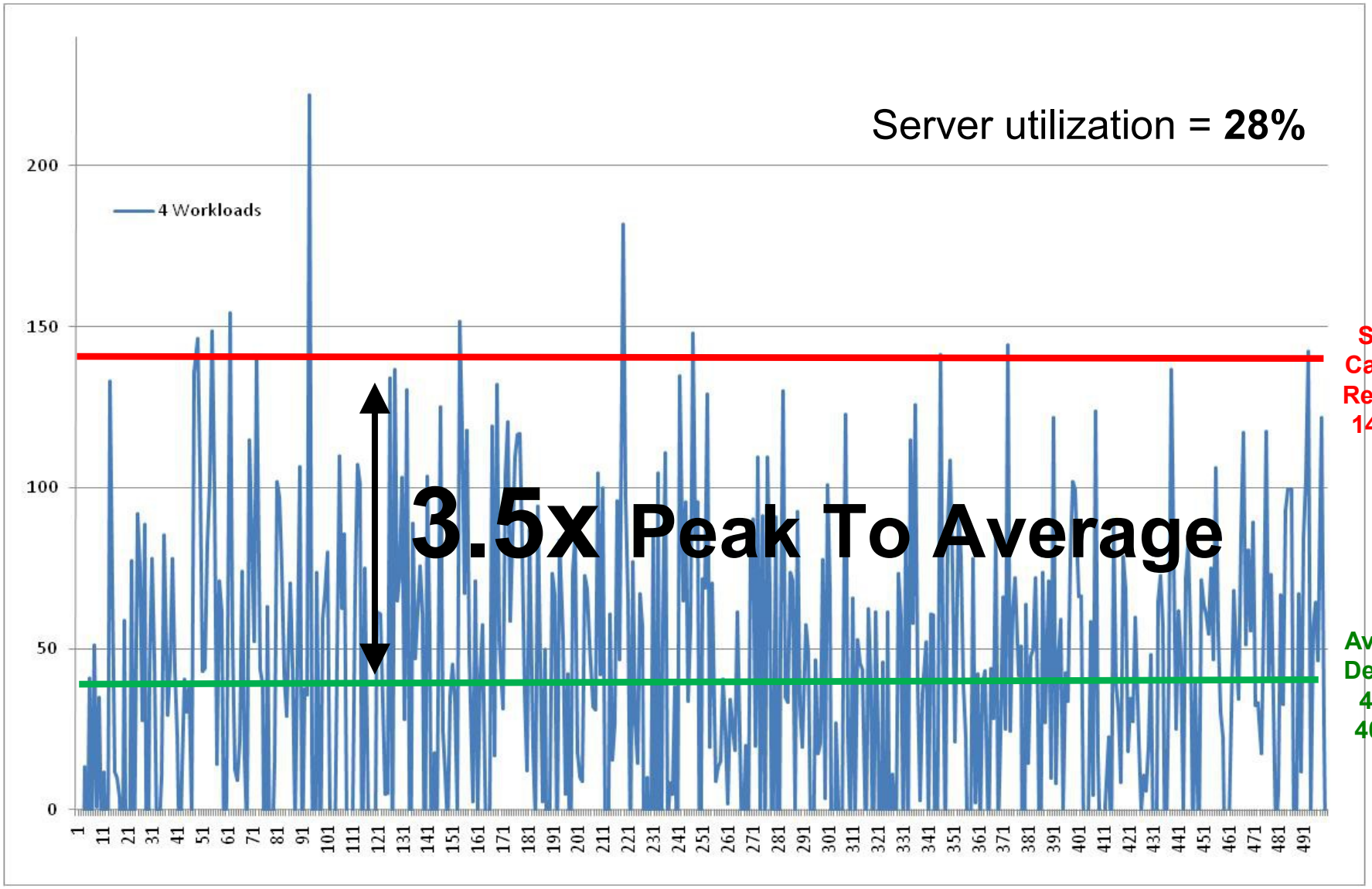
- When you consolidate workloads with variance on a virtualized server, the variance of the sum is less (statistical multiplexing)
- The more workloads you can consolidate, the smaller is the variance of the sum
- Consequently, bigger servers with capacity to run more workloads can be driven to higher average utilization levels without violating service level agreements, thereby reducing the cost per workload

A Single Workload Requires a Machine Capacity Of 6x the Average Demand



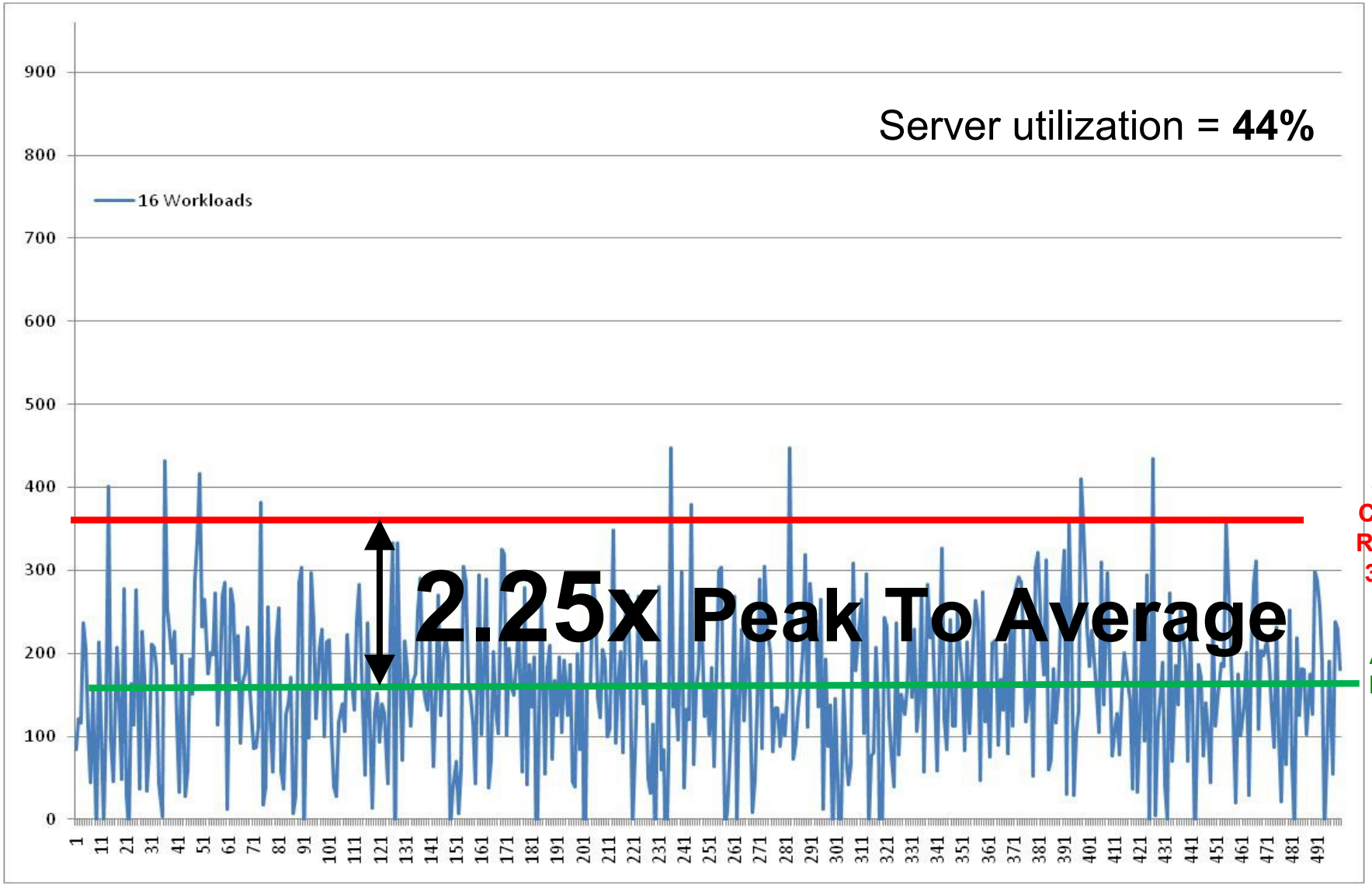
Assumes coefficient of variation = 2.5, required to meet 97.7% SLA

Consolidation Of 4 Workloads Requires Server Capacity Of 3.5x Average Demand



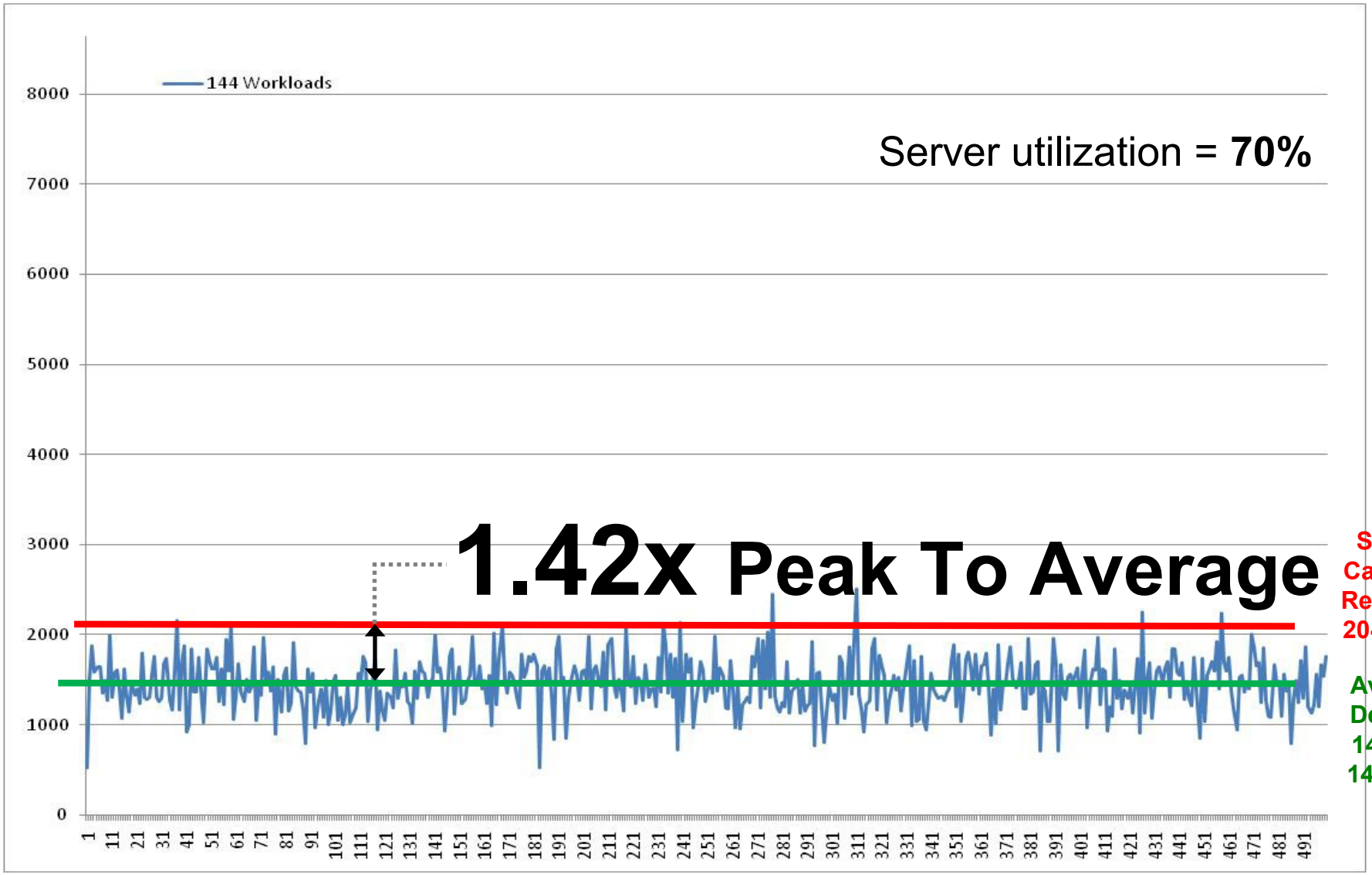
Assumes coefficient of variation = 2.5, required to meet 97.7% SLA

Consolidation Of 16 Workloads Requires Server Capacity Of 2.25x Average Demand



Assumes coefficient of variation = 2.5, required to meet 97.7% SLA

Consolidation Of 144 Workloads Requires Server Capacity Of 1.42x Average Demand



Assumes coefficient of variation = 2.5, required to meet 97.7% SLA

Let's Look At Actual Customer Data

- Large US insurance company
- 13 Production POWER7 frames
  - Some large servers, some small servers
- Detailed CPU utilization data
  - 30 minute intervals, one whole week
  - For each LPAR on the frame
  - For each frame in the data center
- Measure peak, average, variance

Detailed Data Example: One Frame

Frame	LPAR	Min	Max	Std. Dev.	Average	Variance	Max Cores
MSP159	PA3APDC	10.44	59.57	6.46	22.37	0.83	1.19
MSP159	PC2APDC	14.40	45.29	5.19	19.11	0.69	0.91
MSP159	PC18PDC	10.36	41.48	5.19	14.45	0.94	1.24
MSP159	PB5BPDC	9.49	32.92	3.23	11.83	0.89	0.99
MSP159	PB4EPDC	9.26	37.16	3.54	11.57	1.11	1.11
MSP159	PAF5PDC	6.00	95.27	11.78	11.25	3.73	4.76
MSP159	PFE2PDC	4.43	46.23	6.63	9.33	1.98	0.92
MSP159	PB3EPDC	7.83	14.31	0.60	8.53	0.34	0.29
MSP159	MSP159VIO2	4.33	14.95	1.86	8.51	0.38	0.45
MSP159	PCB1PDC	0.79	88.48	17.73	7.88	5.12	5.31

Customer Data Confirms Theory

Servers with more LPARs have less variance in their utilization!

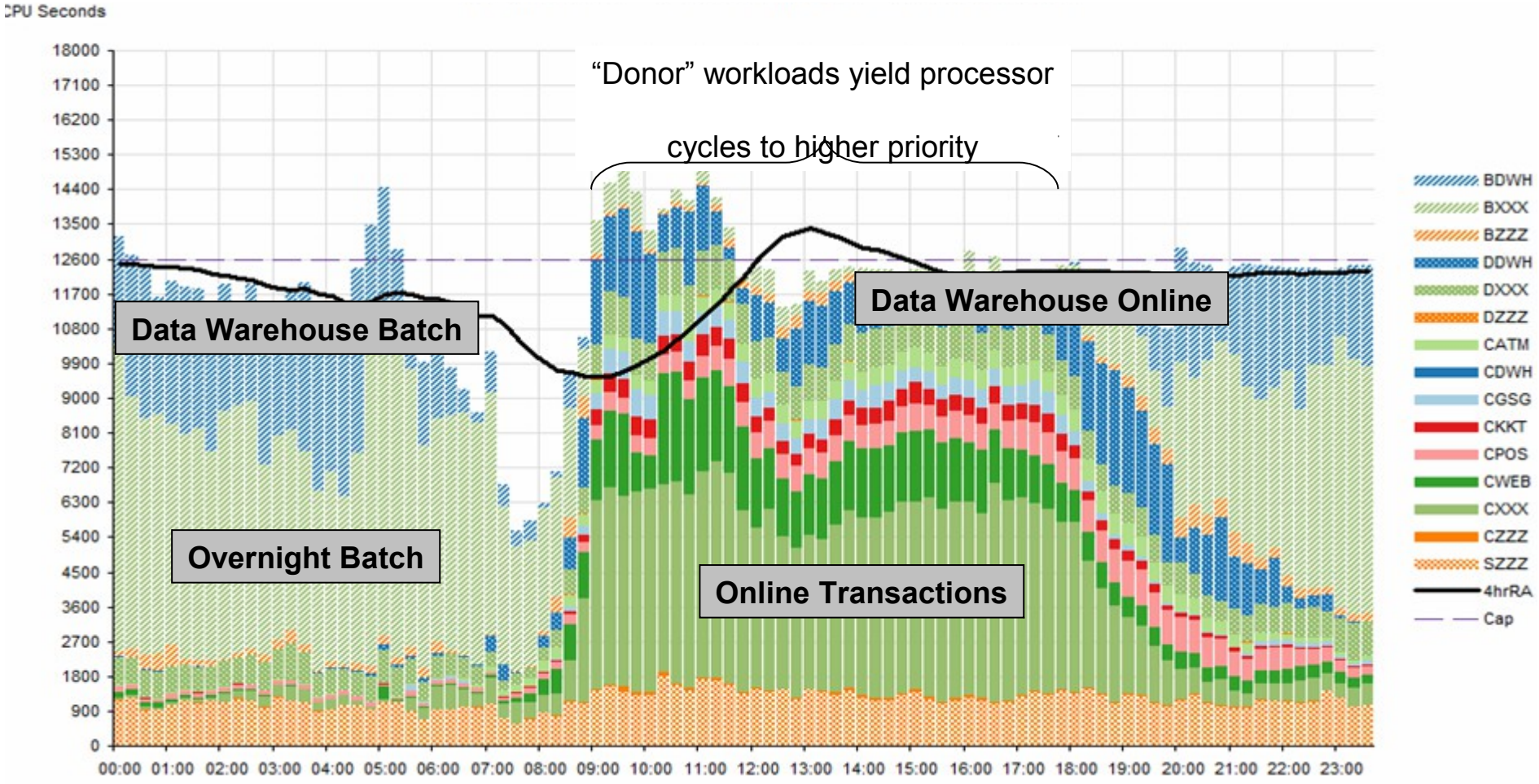
Observations

- There is a benefit to large scale servers
  - The headroom required to accommodate variability goes up only by sqrt(n) when n workloads are pooled
  - The larger the shared processor pool is, the more statistical benefit you get
  - Large scale virtualization platforms are able to consolidate large numbers of virtual machines because of this

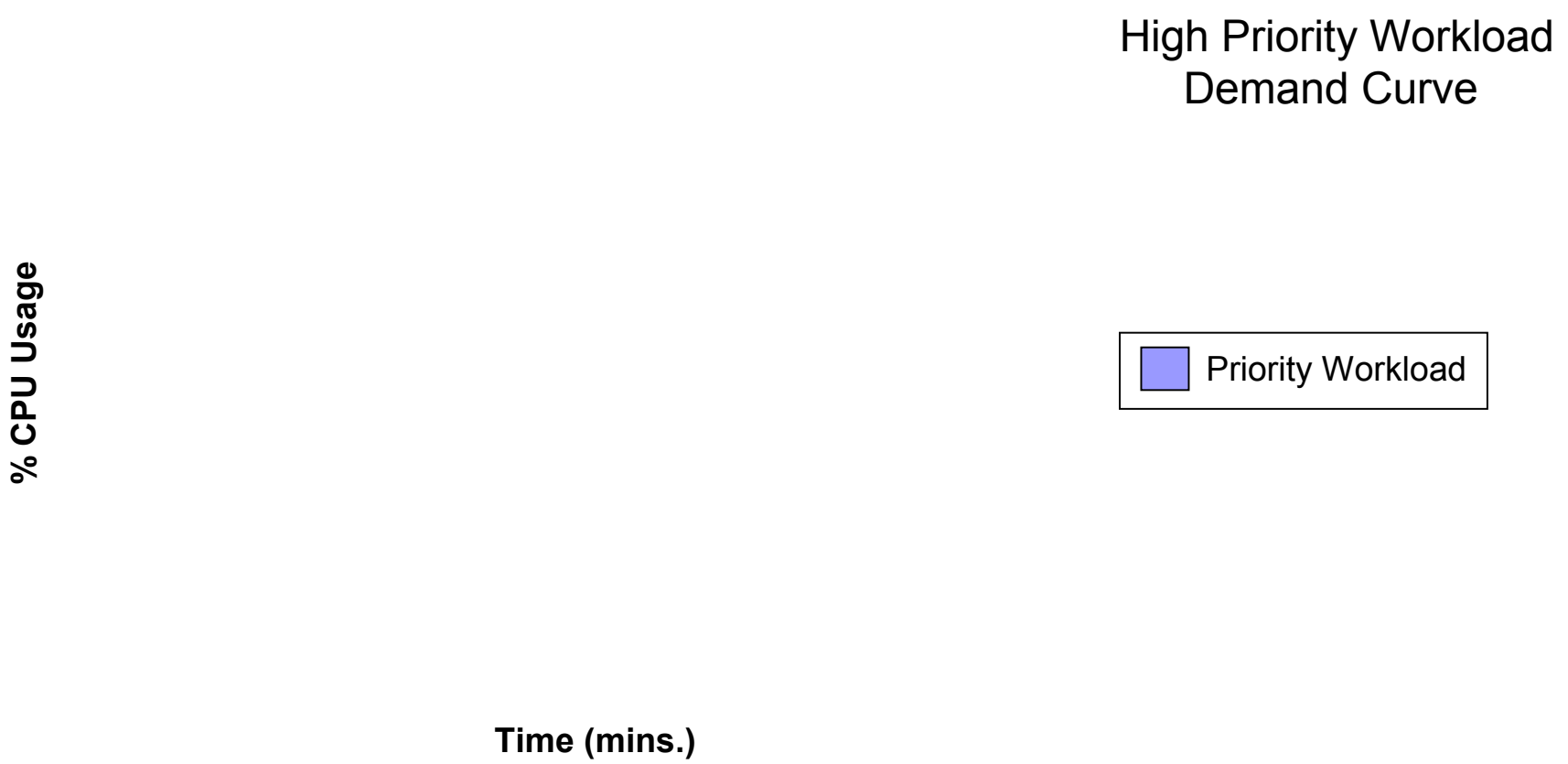
- Servers with capacity to run more workloads can be driven to higher average utilization levels without violating service level agreements

Perfect Workload Management Drives Even More Efficiency

Run multiple consolidated workloads on the same platform



High Priority Web Workload With Varying Demand Running Standalone On System z



Capacity Used
High Priority - 72.2% CPU Minutes
Unused (wasted) - 27.8% CPU Minutes

Priority Workload Metrics
Total Throughput: 9.125M
Avg Response Time: 140ms

High Priority Workload On System z Does Not Degrade When Low Priority Workload Is Added

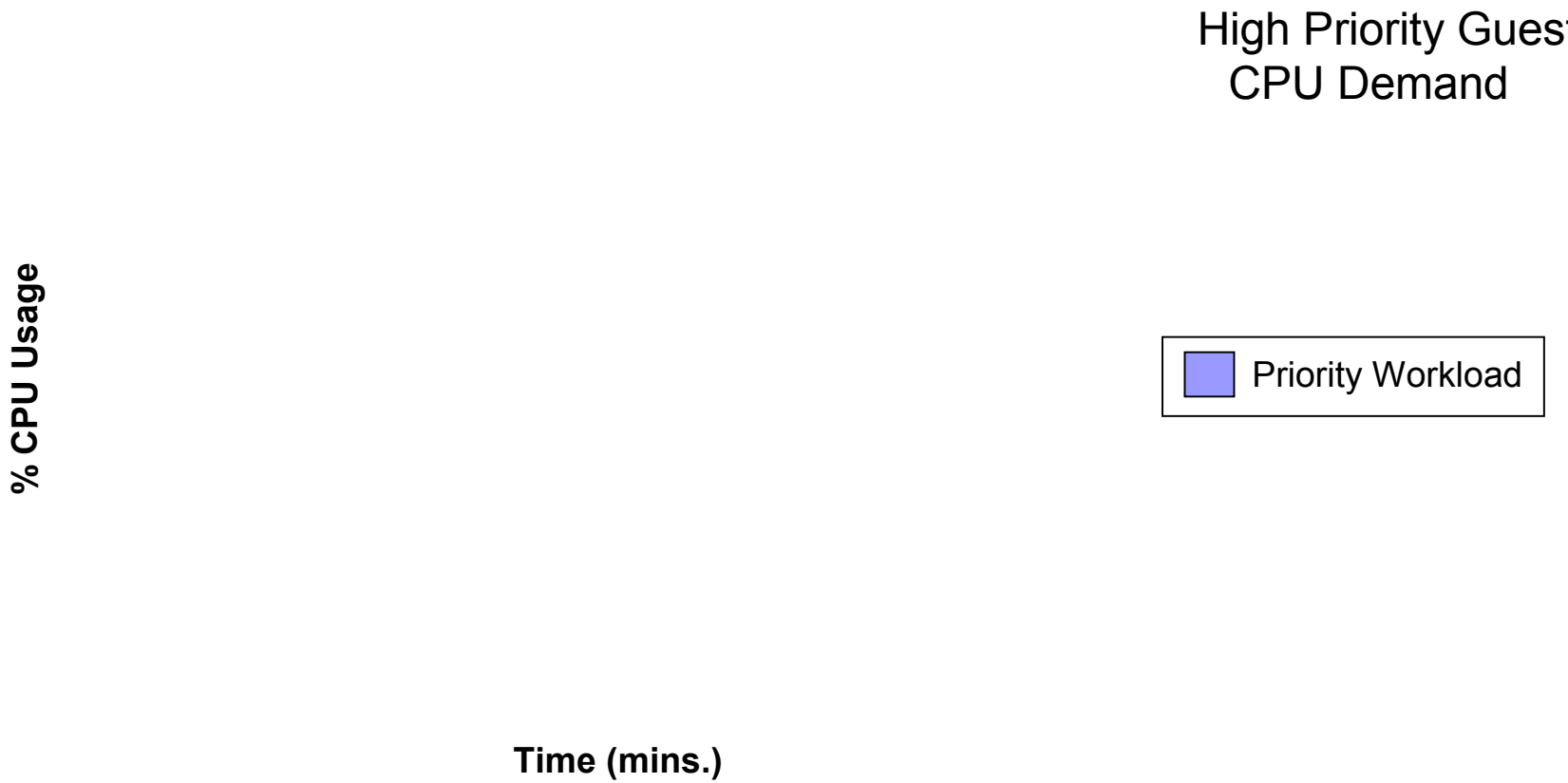
Run High Priority
And Low Priority
Workloads Together

NO
throughput leakage
NO
response time increase

Capacity Used
High Priority - 74.2% CPU Minutes
Low Priority - 23.9% CPU Minutes
Wasted – 1.9% CPU Minutes

Priority Workload Metrics
Total Throughput: 9.125M
Avg Response Time: 140ms

High Priority Web Workload With Varying Demand Running Standalone On Leading Intel Hypervisor

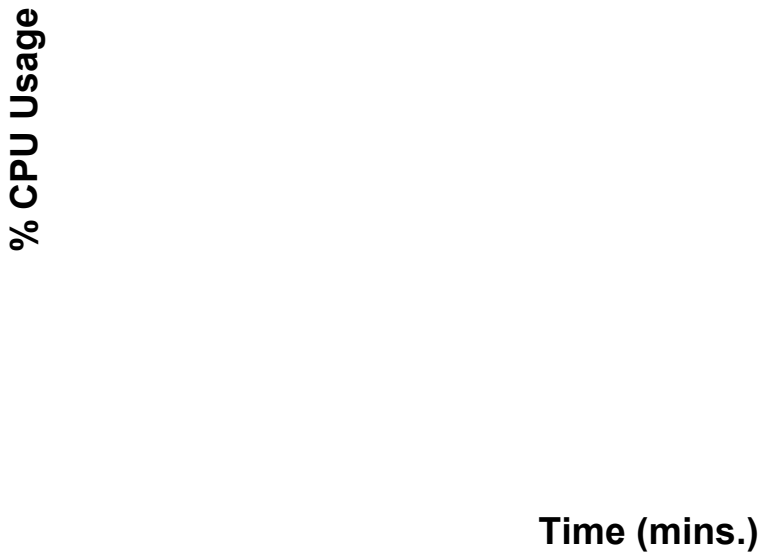


Capacity Used
High Priority - 57.5% CPU Minutes
Unused (wasted) – 42.5% CPU Minutes

Priority Workload Metrics
Total Throughput: 6.47M
Avg Response Time: 153ms

High Priority Workload On Leading Intel Hypervisor Degrades Severely When Low Priority Workload Is Added

Run High Priority
And Low Priority
Workloads Together



30.7%
throughput leakage
45.1%
response time increase
21.9%
wasted CPU minutes

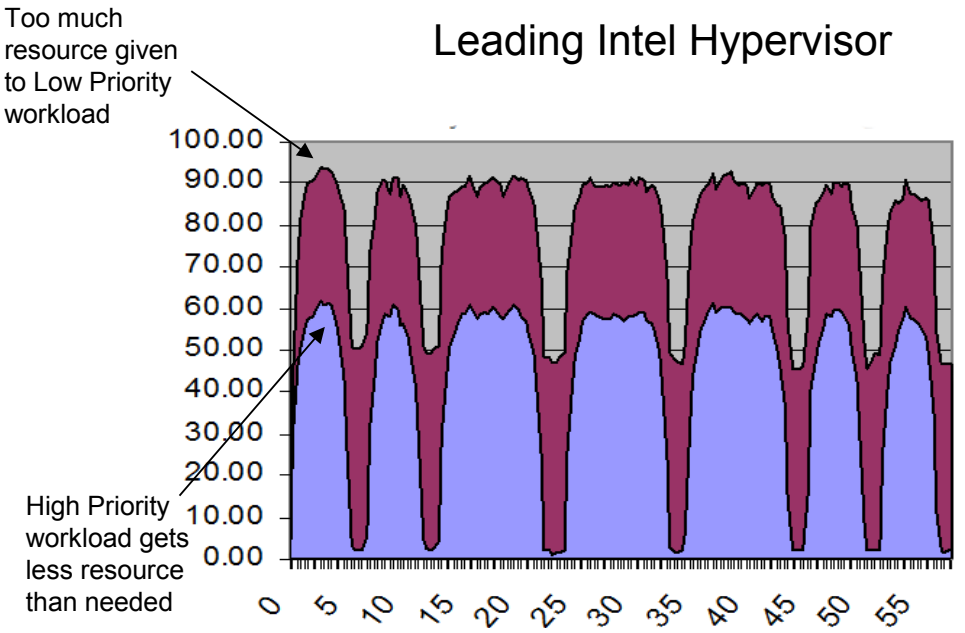
Capacity Used
High Priority - 42.3% CPU Minutes
Low Priority – 35.8% CPU Minutes
Wasted – 21.9% CPU Minutes

Priority Workload Metrics
Total Throughput: 4.48M
Avg Response Time: 220ms

System z Virtualization Enables Mixing Of High And Low Priority Workloads Without Penalty

System z

Leading Intel Hypervisor

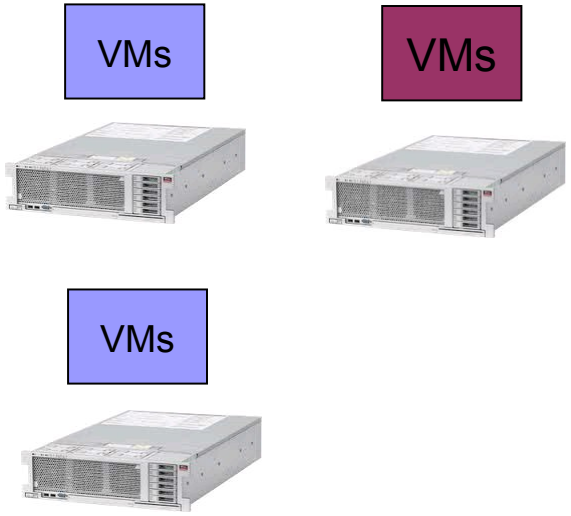


- Perfect workload management
- Consolidate workloads of different priorities on the same platform
- Full use of available processing resource (high utilization)

- Imperfect workload management
- Forces workloads to be segregated on different servers
- More servers are required (low utilization)

Consolidate High and Low Priority Workloads Together While Maintaining Response Time SLA

Servers required to achieve the same level of performance for high and low priority workloads



Virtualized on 3 Intel 40 core servers

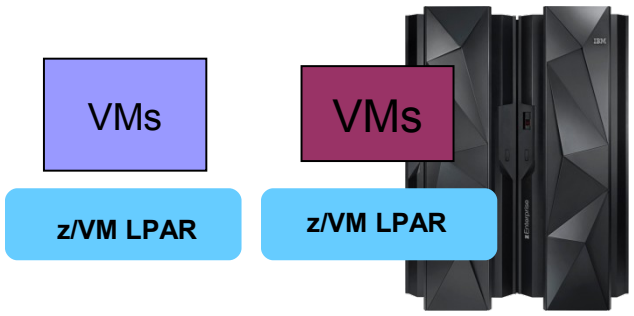
$13.7M (3 yr. TCA)

High priority workloads

Low priority workloads

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

High priority online banking workloads driving a total of **11.9M** transactions per hour and low priority discretionary workloads



z/VM on zEC12
32 IFLs

$5.77M (3 yr. TCA)

58% *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.

Large Consolidation Systems With Centralized Management Deliver Better Labor Productivity

HP Servers + ISV

Production Servers
HP 9000 Superdome rp4440
HP Integrity rx6600

Dev/Test Servers
HP 9000 Superdome rp5470
HP Integrity rx6600

Claims per year **327,652**

$0.12 per claim

$0.79 per claim

Mainframe support staff has 6.6x better productivity

IBM System z CICS/DB2

Total MIPS 11,302

MIPS Used for commercial claims processing production/dev/test **2,418**

Claims per year **4,056,000**

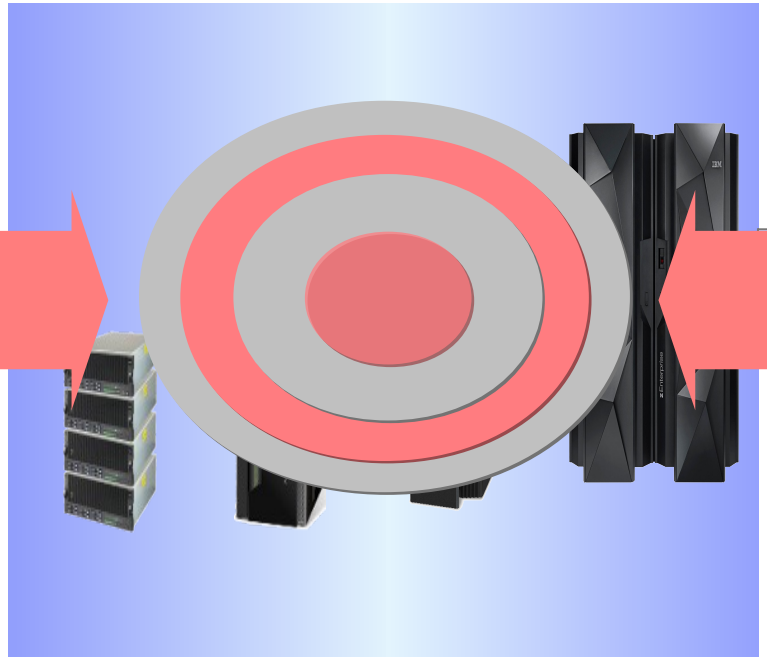
Resilient Enough To Survive An Earthquake

A scene from our Tokyo datacenter after an 8.9-magnitude earthquake on March 2011

There were <u>no</u> service interruptions, and there was no need to switch over to a disaster recovery site.

Run These Workloads In the High End Systems Sweet Spot

- Global scale critical data workloads
- Transaction processing
- Batch processing
- Co-located analytics
- Consolidated on one platform



- Large critical data workloads
- CPU intensive and cache intensive processing applications
- Consolidated on one platform

Best price/performance

Prove it with an Eagle Fit for Purpose study!