The New zEnterprise – A Cost-Busting Platform

TCO Lessons Learned, Establishing Equivalence

(aka Improving ROI with CICS and the Mainframe)

Total Cost of Ownership In Context:
The Mainframe Is IBMs Premier Platform

- IBM arguably makes the best x86 and UNIX hardware
- However, the mainframe remains the pinnacle of IBM technology
- Technologies pioneered on the mainframe eventually get emulated in limited ways on x86 and UNIX systems

- As we already know, today the mainframe has:
  - The best hardware – world’s fastest microprocessor (5.5 GHz), most I/O channels, hardware assisted clustering for scale out, ...
  - The best software – able to smoothly integrate the value of the existing (z/OS, CICS, COBOL, PL/I, etc.) with the newest technologies of today (Linux, J2EE, Java, XML, REST, etc.)
  - Best for efficiency – heavily automated, world-leading problem determination and performance analysis, minimal administrative footprint (centralized deployment = very few LPARs, eg. <5)
  - Best for environmentals – lowest electrical, networking and floor space usage per unit of work processed compared to others

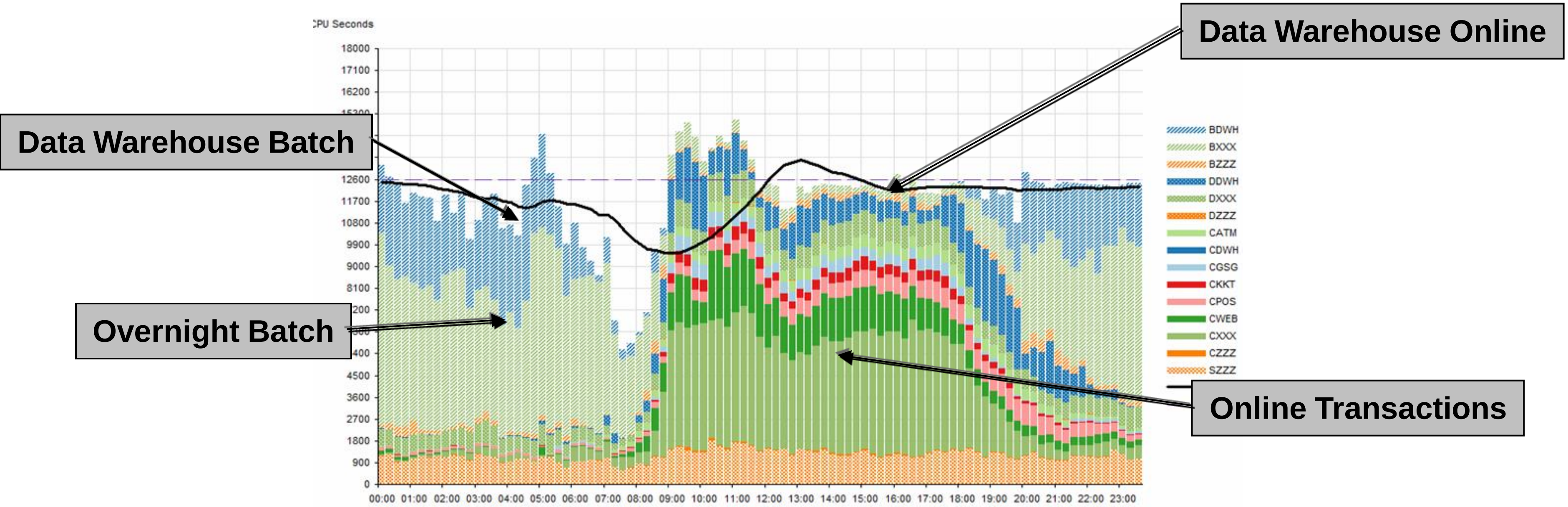
What's So Special About A Mainframe Anyway?

- A few things come together to make mainframes special
  - Centralized computing is inherently more efficient than distributed or modular computing
  - High availability at the application level, not just the hardware, and easy DR
  - Ability to reuse existing assets in new ways and via new technologies gives huge investment protection
  - Unmatched workload management enables "stuffing the pipe" without impact to important work – run continuously at 100% safely
  - Balanced compute and I/O hardware for ultimate scale up
  - Hardware assisted clustering (used by all middleware) for ultimate scale out

- The results can be measured – for instance, throughput per core on mainframes is MUCH higher than on other machines
  - Simply switching to a modular deployment model while remaining on mainframe hardware requires a 6x expansion in compute capacity
  - Mainframes routinely provide the highest availability to their end users
  - Reusing (sometimes with some modernization) existing assets is fast, cheap and safe compared to the other options
  - Flexibility in scale and ability to dynamically react to workload spikes without manual intervention gives peace of mind against changing conditions

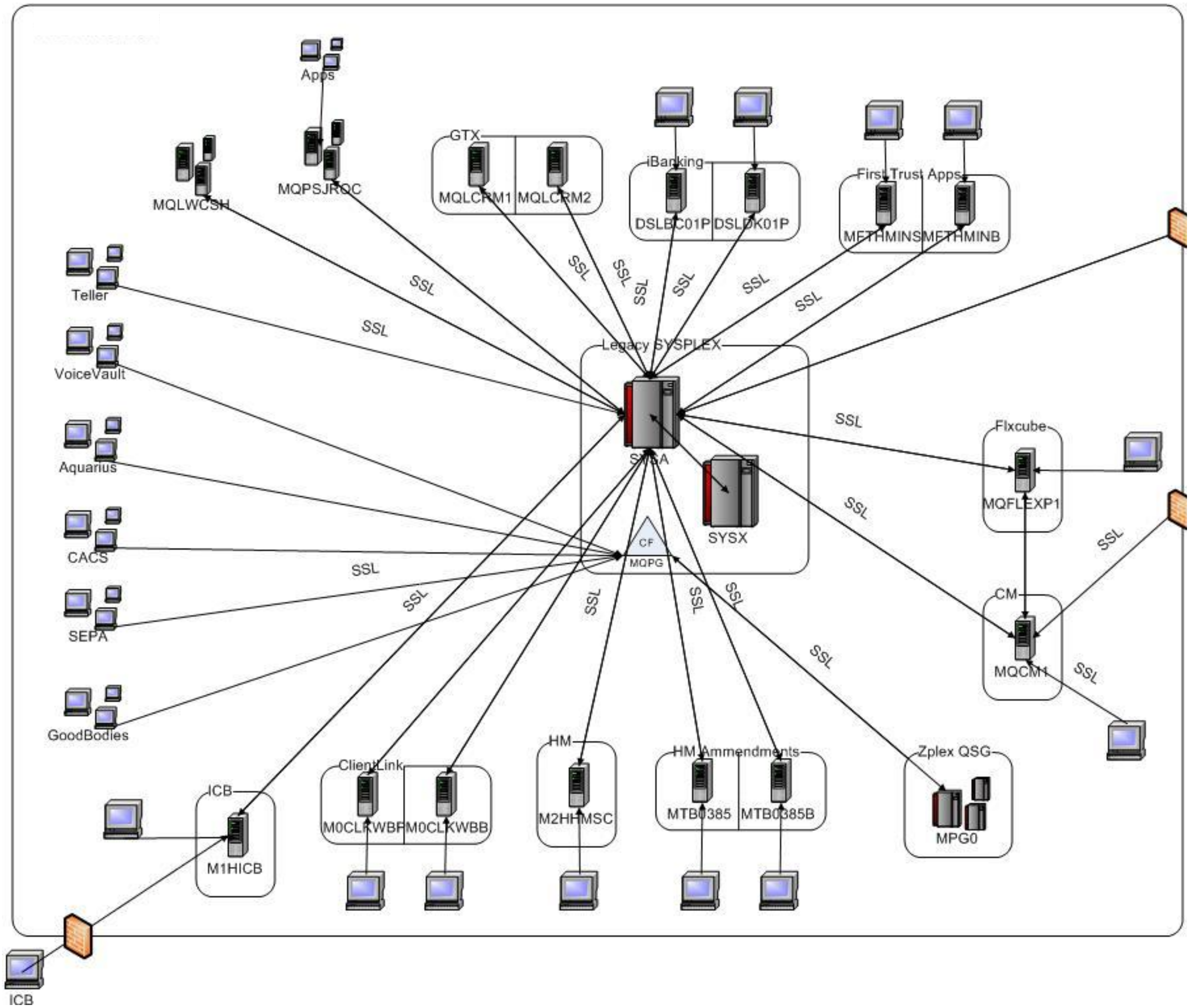
Customer Data: System z Utilization Unheard Of On Other Hardware Platforms

Example: Core banking workloads running simultaneously with high utilization



- Compare to a medium-sized EMEA bank with a range of Sun/Oracle and x86 servers, average over whole estate <10%
  - 1125 total servers, 228 no data (MS Windows)
  - 624 have average utilization 10% or less (70% of estate)
  - 511 have 5% or less (60%), only 5% avg over 40% utilization

Typical Customer Deployment Pattern Shows The Importance Of The Mainframe



- Many applications depend on the core mainframes for their data and services
- In the diagram it is striking that there are almost no connections between individual client nodes, communications travel to/from the core
- The mainframes provide a highly available central cluster of core services and data to the whole business

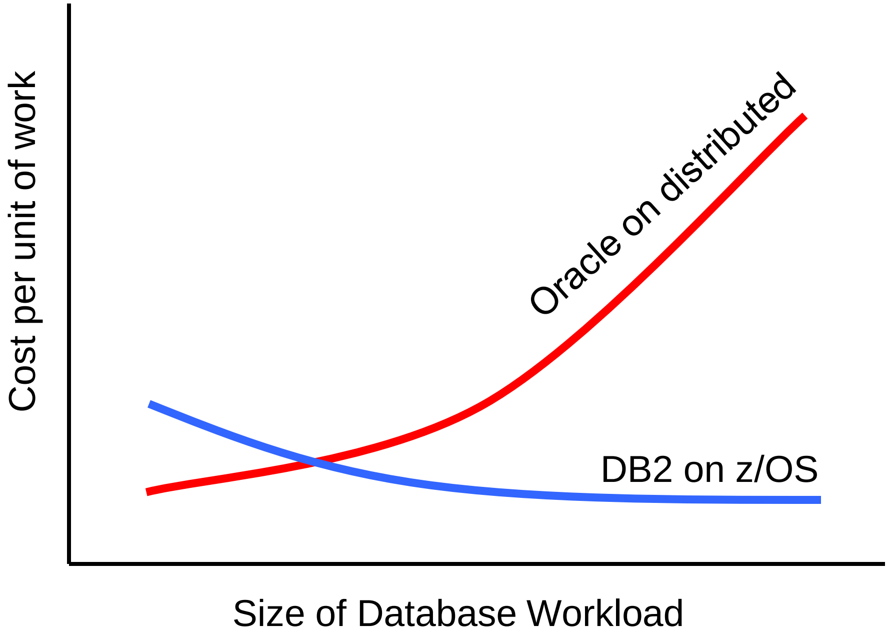
Example of The Cost Differences Of System z vs. Distributed

Why do distributed database costs per unit of work **rise** as workloads grow?

- Core proliferation
  - Software priced per-core
- Oracle database software costs
  - Largest component of cost
- Hardware cost escalates faster as number of cores (SMP) grows

Why do System z database costs per unit of work **fall** as workloads grow?

- Linear scalability
- Mainframe software pricing reflects economies of scale



Core Proliferation For A Small Workload Shows Relative Efficiency (Actual Customer)

6x 8-way Production / Dev
2x 64-way Production / Dev
Application/MQ/DB2/Dev partitions

2x z900 3-way Production / Dev / QA / Test

6 processors
(1,660 MIPS)

176 processors
(800,072 Performance units)

$25.4M (5 yr. TCO)

$17.9M (5 yr. TCO)

***29x** more cores!*

Still a mainframe customer (primary workloads moved to 2 new mainframes), this equivalence shows only the “filler” applications that moved off in an application portfolio management initiative

Core Proliferation Is Not Related To Size

- De-consolidation of applications to dedicated servers
  - Dedicated servers / images for functional roles – application, database, security, batch, systems management
  - Separate servers for production, development, quality assurance test
  - Low utilization due to provisioning for the peak on each server and pre-provisioning for growth
  - NOT SOLVED by virtualization due to capacity reservations in production!

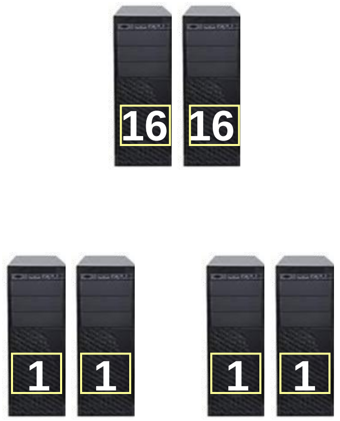
- Disaster Recovery
  - 100% coverage doubles the number of cores required

- Processing comparisons
  - Mainframe has dedicated processors for I/O, distributed does not
  - Processing expansion (CICS/COBOL path lengths are highly optimized)
  - Converting IMS hierarchical database to relational results in a 3x expansion
  - No networking on mainframe reduces computation (and latency)
  - Storage expansion and increased data management

Core Proliferation For A Smaller Project (Actual Customer Experience)

2x 16-way Production / Dev / Test / Education
App, DB, Security, Print and Monitoring
4x 1-way Admin / Provisioning / Batch Scheduling

z890 2-way Production / Dev / Test / Education
App, DB, Security, Print, Admin & Monitoring



0.88 processors
(332 MIPS)

36 Unix processors
(222,292 Performance Units)

$17.9M (4 yr. TCO)

41x more cores

$4.9M (4 yr. TCO)

Almost 5 Year Migration

670 Performance Units per MIPS

Plus:
2x HP SAN Servers (existing)
Many (existing) Windows servers

No Disaster Recovery

Move To x86 – Worst Core Proliferation Ever! (Actual Customer Experience)

3x HP DL580 (2ch/20co)
Production / Dev / Test
(2011 x86 technology)

z800 Production /
Dev / Test
(2002 mainframe technology)

2.1 processors
(499 MIPS)

60 Linux processors
(383,022 Perf Units, 8 MIPS/core)

29x more cores (with the 9 year technology gap)

Equivalent to 180x more cores with current technology

1.5 Year Migration

768 Performance Units per MIPS

Works In Reverse Too! Typical UNIX-Style Data Workloads Benefit From System z

Which platform provides the lowest TCA over 3 years?

Oracle DB workload

3 Database Workloads
Oracle Enterprise Edition
Oracle Real Application Cluster

3 Oracle RAC clusters
4 server nodes per cluster

12 total Oracle T3-4 servers (768 cores)

$28.7M (3 yr. TCA)

3 Oracle RAC clusters
4 nodes per cluster
Each node is a Linux on z guest

zEC12 with 27 IFLs

$5.7M (3 yr. TCA)

***80%** less cost!*

TCA includes hardware, software, maintenance, support and subscription. Workload Equivalence derived from a proof-of-concept study conducted at a large Cooperative Bank and projecting to T3-4 servers using published TPC-C Results normalizing them to Relative Performance Units as available from Ideas International

Chargeback (or Cost Accounting) Can Mislead If Allocations Are Not Truly Aligned With Platforms

- The mainframe is often the computing platform that has been around the longest in the data center
- It is also the platform that is best at accounting and billing for usage
  - Great for supporting cloud deployments!
- Therefore many datacenter fixed costs that are not platform specific are recovered or accounted for under the mainframe
  - The Eagle TCO team has now come across 3 customers who recover the company jet costs under the heading of mainframe computing...
  - Many “small” costs on a per-server basis are ignored and recovered as overhead to the whole data center... via the mainframe!
- While most important applications that run off the mainframe continue to use the mainframe, everyone gets charged their fair share
- However, this chargeback / accounting scheme gives the impression that the mainframe is expensive compared to other options
- Unfortunately this creates incentives to move off, and it is only when customers eventually update their chargeback that they discover they were led astray

Other Savings The Mainframe Provides

- Special pricing metrics for new workloads and integration
- Sub-capacity pricing for utility pricing of workloads
- Consolidation with Linux on System z reduces the software and management costs of modular UNIX-style workloads
- Specialty processors and devices to reduce cost further
- Fast, simple hardware refresh with investment protection
- Easy and low cost disaster recovery
- Minimal additional capacity for non-production environments (development, test)
- Gold standard in tooling for automation and problem determination
- Enjoy the simplicity, security and performance advantages of collocating important workloads and data together

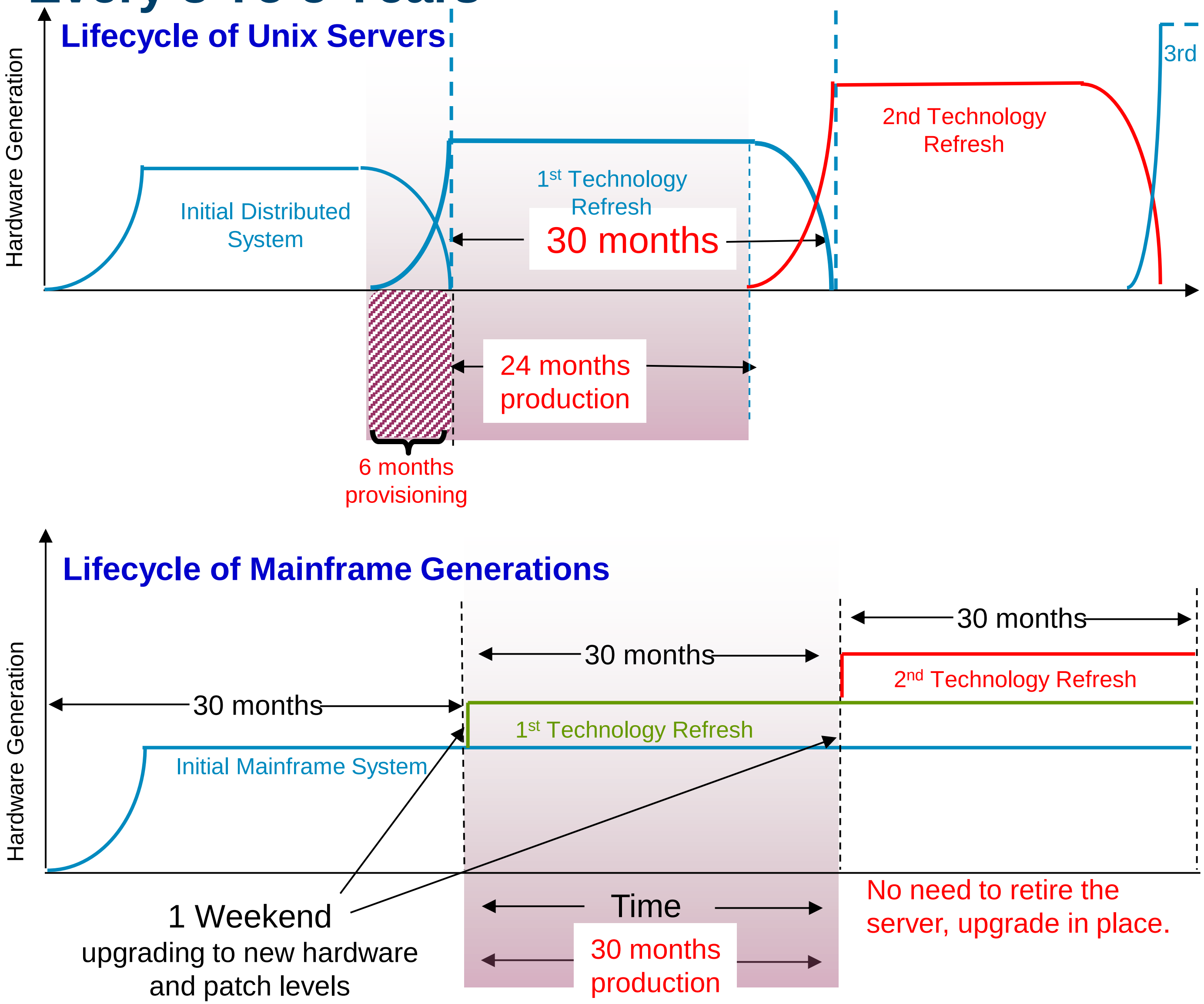
System z Linux Consolidation Saves Money

- Large financial services company
  - Mixed Oracle and WAS environments on Intel
  - Scaling out rapidly – up to 172 images on 836 cores

	Refresh existing x86	*Consolidate on Linux on z*	*Observation*
Software	9.76M	3.70M	Software costs down 62%
Hardware	2.32M	5.22M	Hardware costs up 125%
Labor/Facilities	3.83M	0.69M	Charge to department down 82%
Migration	0.18M	0.41M	
Cost Avoidance	-	-0.61M	
Total	**16.1M**	**9.41M**	

- 5 Year Savings: 6.6M
- 26:1 core consolidation from virtualized x86 to Linux on z

Distributed Servers Need To Be Replaced Every 3 To 5 Years



Refresh is normally even worse than just re-purchasing existing capacity as this real customer demonstrates:

Non-mainframe systems must co-exist for months at a time while being refreshed, requiring space, power, licenses etc. In this case only 24 months of productive work is realized for each 30 month lease period and the leases overlap up to 6 months

The mainframe by contrast is upgraded over a weekend and is fully productive at all times

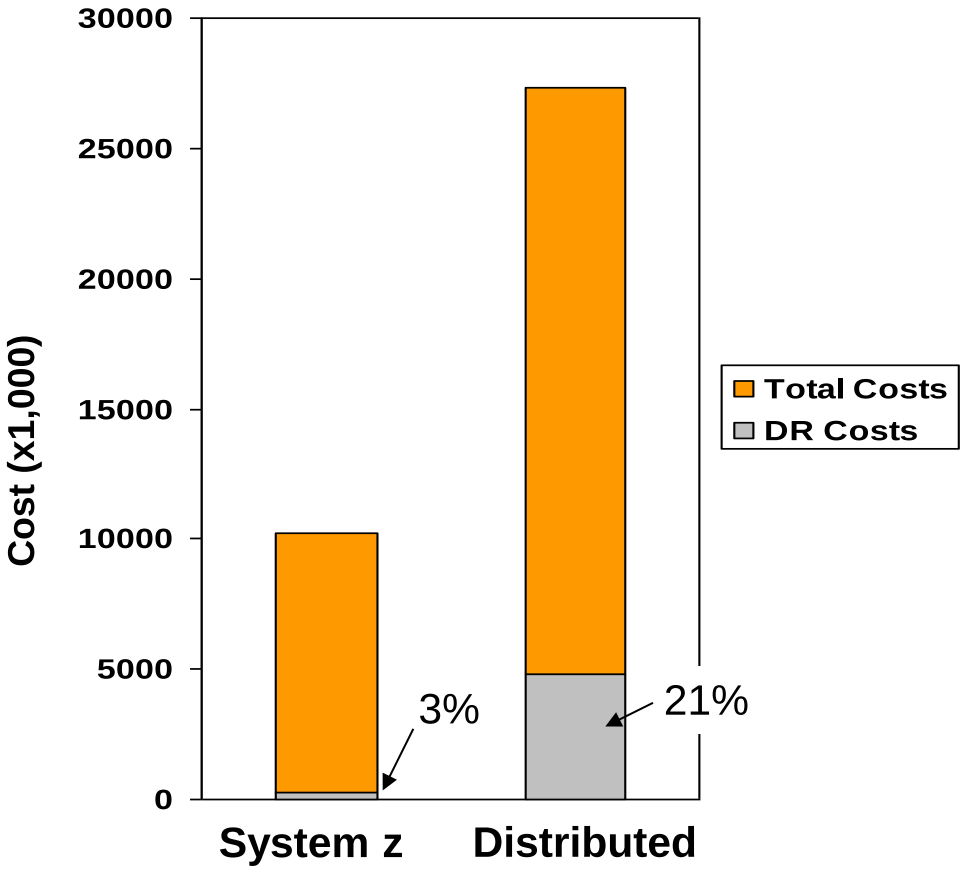
Disaster Recovery On System z Costs Much Less Than On Distributed Servers

A large European insurance company with mixed distributed and System z environment at :

Disaster Recovery Cost as a percentage of Total Direct Costs:

System z – **3%**

Distributed – **21%**

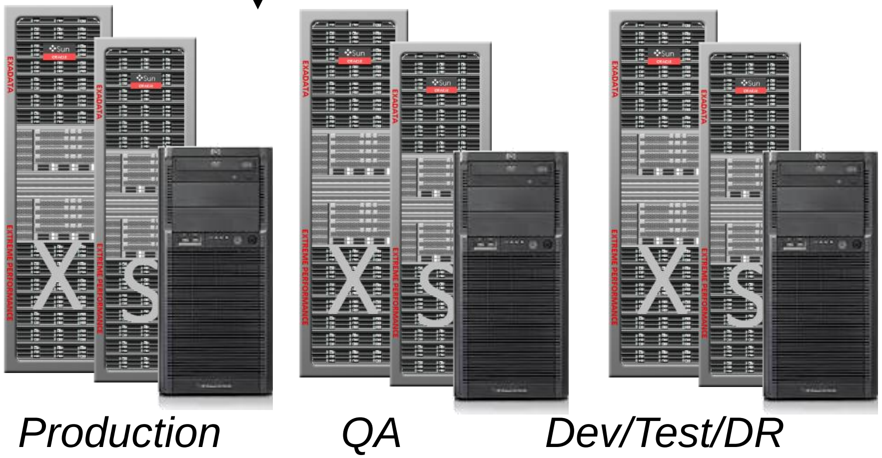


Two mission-critical workloads on distributed servers had DR cost > 40% of total costs

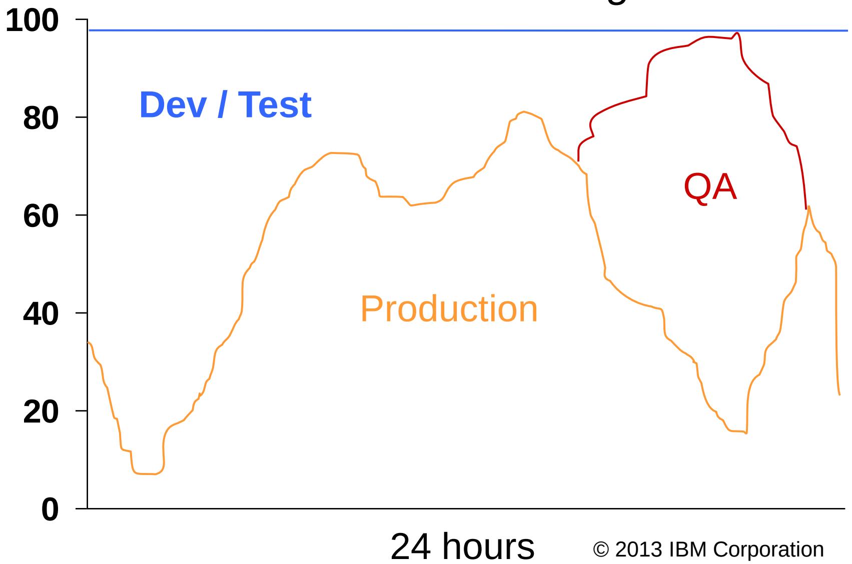
Non-Production Environments Require Fewer Resources On The Mainframe

- High Availability mechanisms for Production
  - Dedicated failover (Prod x 2.5)
  - N+1 clustered (Prod x 2 worst case)
  - Mainframe (usually Prod x 1, sometimes less!)

- Development and Test Capacity
  - Mainframe – Prod +20%
  - Distributed – a range, often Prod +200%

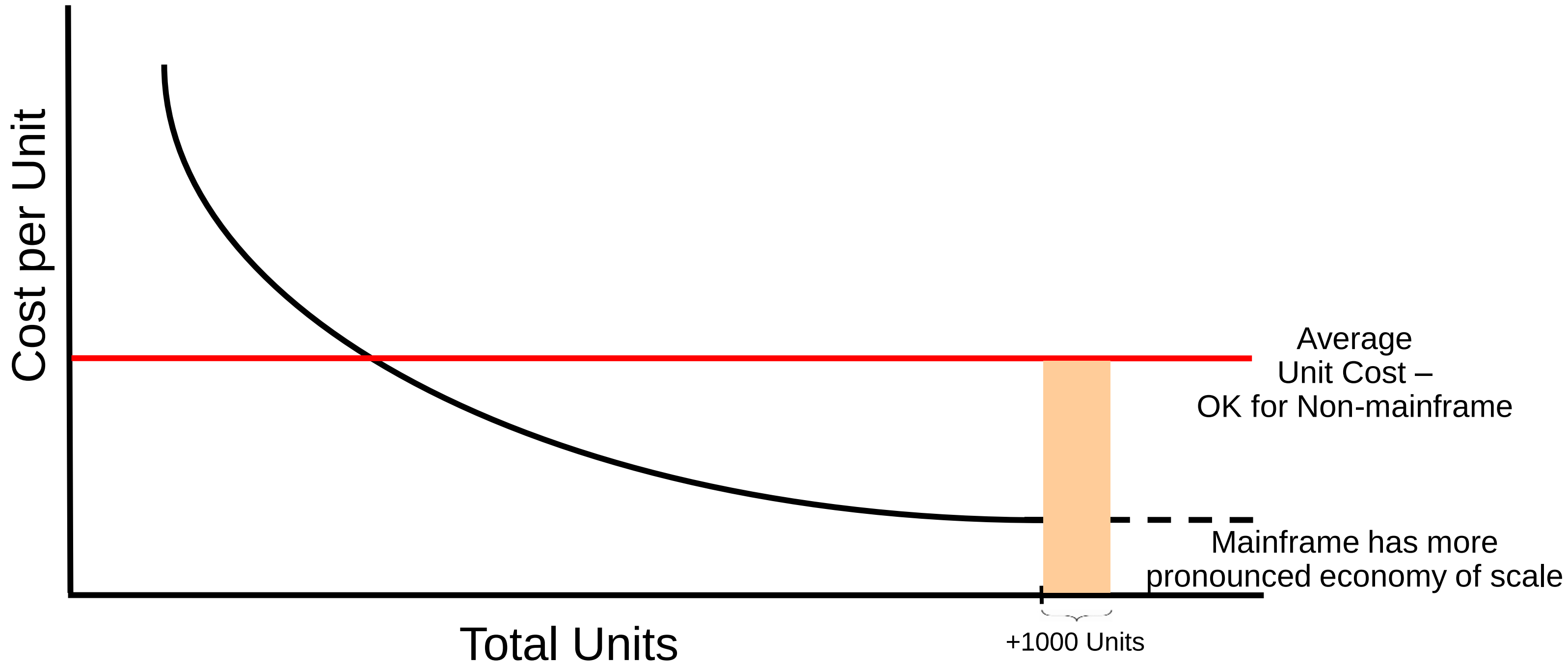


Mainframe Usage Profile

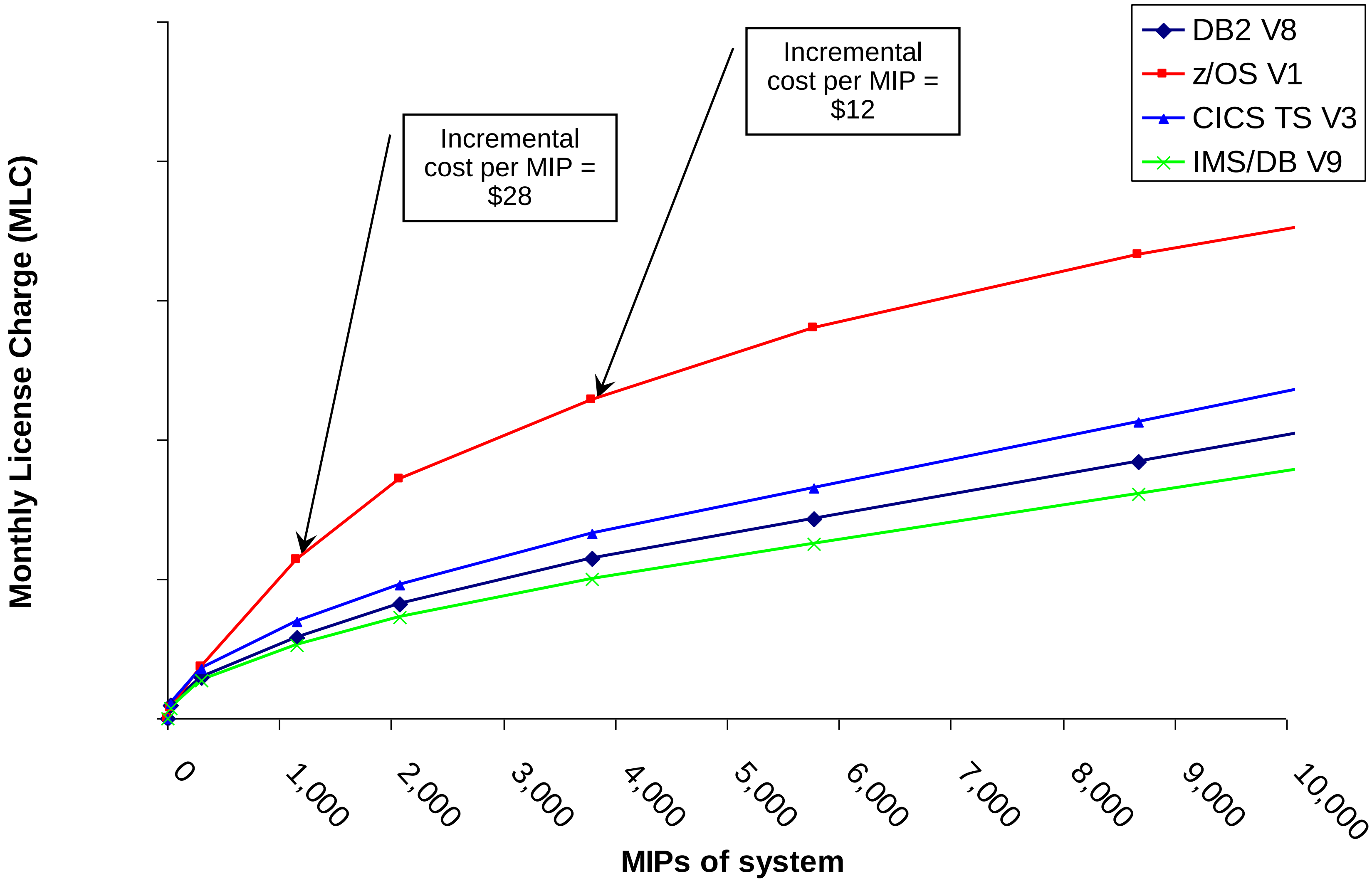


Reality Of Incremental Costs For Different Platforms Is Quite Different

- Mainframes are priced to deliver a substantial economy of scale as they grow
- Doubling of capacity results in as little as a 30% cost growth for software on z/OS
- Average Cost is significantly more than incremental cost



Mainframe Pricing Curves Favor Growth (Software Example Shown)



Future Mainframe Directions

- Consolidate and improve analytics
  - Keep data where it originates rather than move it around
  - Allows change from delayed data warehouse model to real-time analytics with feedback loop into online systems
  - Offload to specialty devices where necessary for cost, but crucially, do this dynamically under computer control, not manually – so retain a single access paradigm for all work

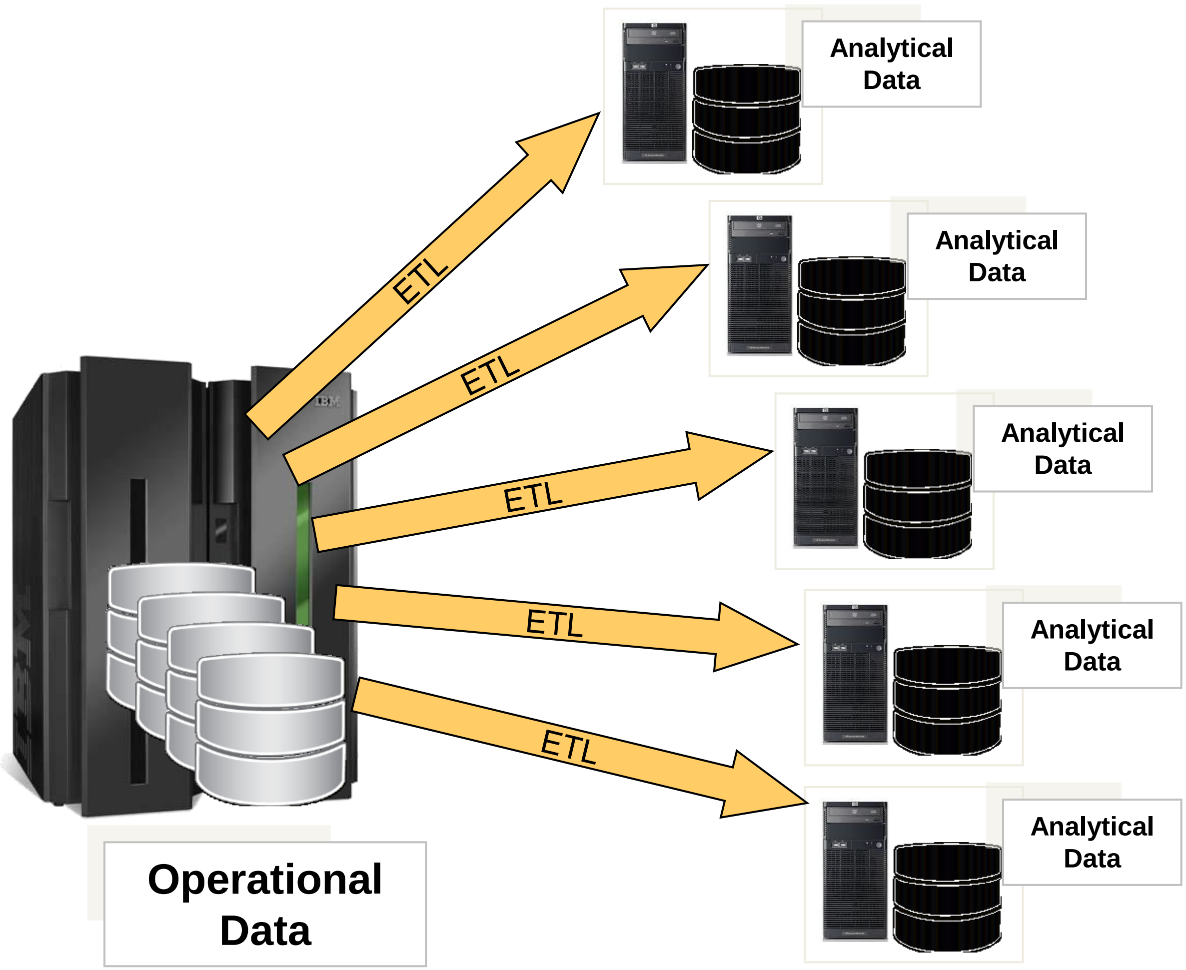
- Continue to enable new technologies to access existing services and data
  - eg. CICS enables HTTP, then XML and now RESTful access to existing transactions and data

- Enable adaptation to changing business requirements
  - eg. CICS embeds WAS Liberty for batch allowing batch tasks to run at any time during the day in cooperation with online work

- Continue to provide the most efficient centralized deployment model
  - Ever larger, faster and better machines with middleware to exploit

- Enhance existing cloud and management capabilities
  - eg. Smart Cloud Provisioning with workload patterns support

Many Customers Have Built Complex Data Movement Solutions With Unforeseen Costs



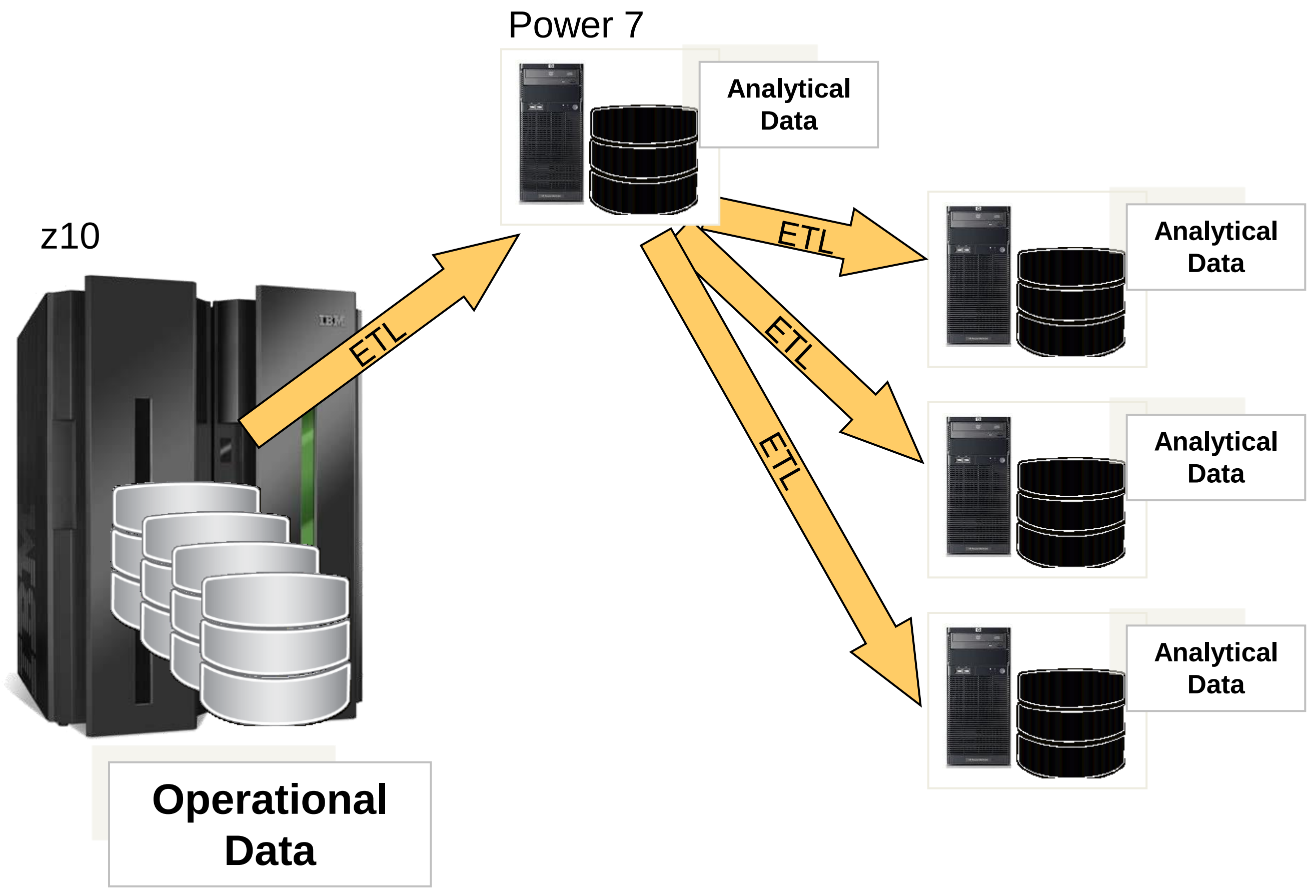
A large European bank:

- 120 database images created from bulk data transfers
- 1,000 applications on 750 cores with 14,000 software titles
- ETL consuming 28% of total distributed cores and **16% of total MIPS**

A large Asian bank:

- One mainframe devoted exclusively to bulk data transfers
- ETL consuming 8% of total distributed core and **18% of total MIPS**

Clearly Moving Data Is Not Free
Here is a Typical Situation…



4 yr. amortized cost summary

System costs = $8,046,198

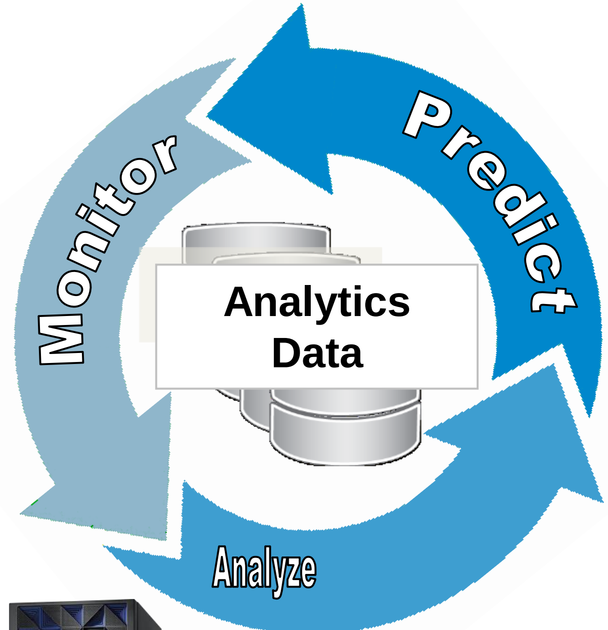
Labor costs = $223,137

Total = $8,269,335

IBM zEnterprise System is Optimized for *All* Critical Data

Run the business

Grow the business



zEnterprise EC12

IBM DB2 Analytics Accelerator

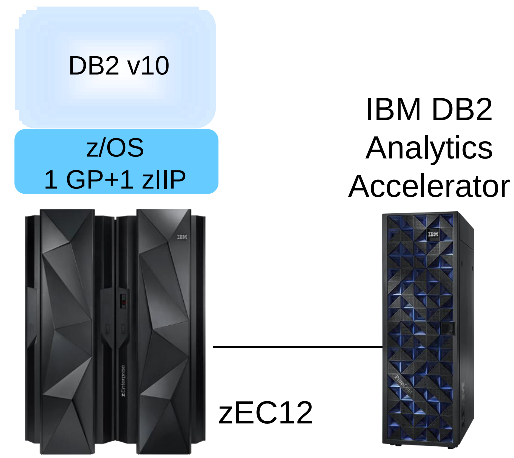
Analytics on zEnterprise with the Accelerator Beats the Competition

Standalone Pre-integrated Competitor – Quarter Unit

Unit Cost (3yr TCA) **$905/RpH**

Workload Time	3,043 mins
Reports per Hour (RpH)	3,178
Competitor ¼ Rack (HW+SW+Storage)	$2,876,561

IBM zEnterprise



Unit Cost (3yr TCA) **$71/RpH**

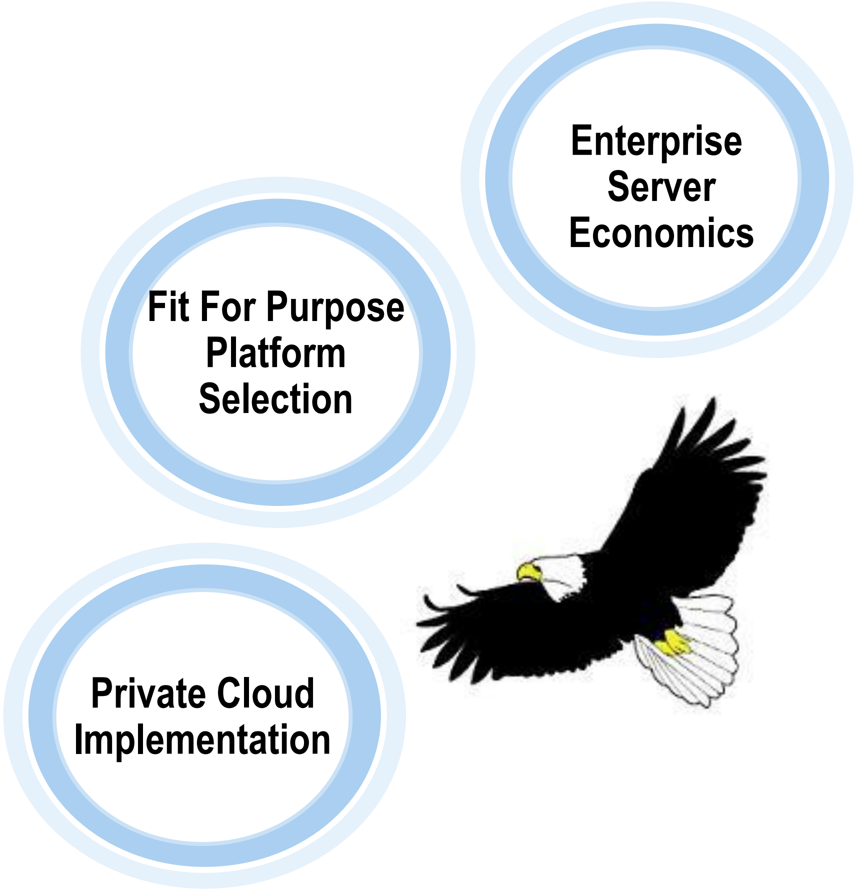
Workload Time	294 mins
Reports per Hour (RpH)	32,891
zEC12 (1 GP + 1 zIIP, HW+SW+50TB Storage)+Accelerator	$2,337,400

12x better price performance!

Source: IBM Customer Study.

The IBM Eagle team helps customers understand mainframe costs and value

- **Worldwide** team of senior technical IT staff
- **Free of Charge** Total Cost of Ownership (TCO) studies
  - Help customers evaluate the lowest cost option among alternative approaches
  - Includes a one day on-site visit and is **specifically tailored to a customer's enterprise**
- Studies cover POWER, PureSystems and Storage accounts in addition to System z
  - For both IBM customer and Business Partner customer accounts
- Over 300 customer studies since formation in 2007
- Contact: eagletco@us.ibm.com



What happens in a TCO study?

Workload identified for analysis

frontends

databases

other components

Application XYZ (Prod, Dev, QA)

Deployment Choices

Do nothing

Optimize current environment

Deploy on other platforms

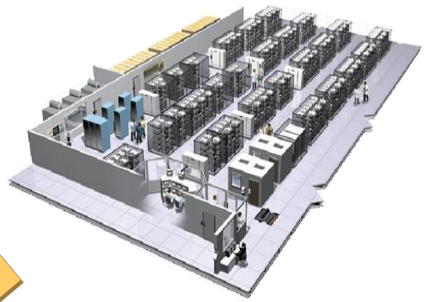
Key steps in analysis

1. Establish equivalent configurations
- Needed to deliver workload

2. Compare Total Cost of Ownership
- TCO looks at different dimensions of cost

How can we determine equivalent configurations?

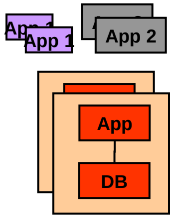
Real world aspects determine accurate equivalence

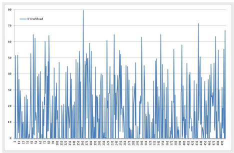


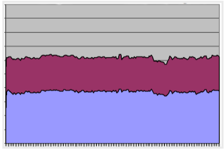
Top Down approach

What we see in customer environments

Bottoms up approach

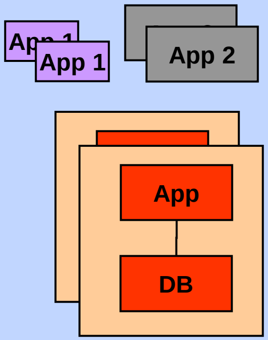


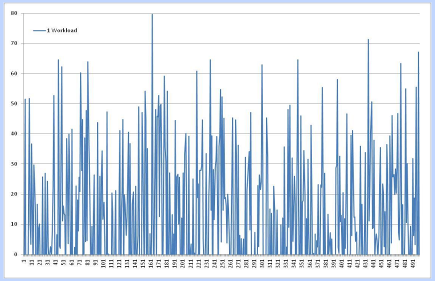


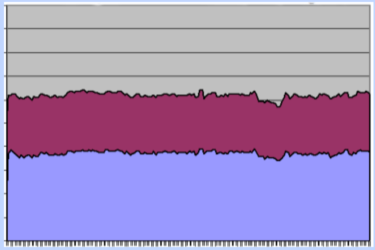


What we know about platforms and measure in atomic benchmarks

Platform differences and atomic benchmarks set a baseline for establishing equivalence







Platform factors

GHz, cache, I/O, co-location

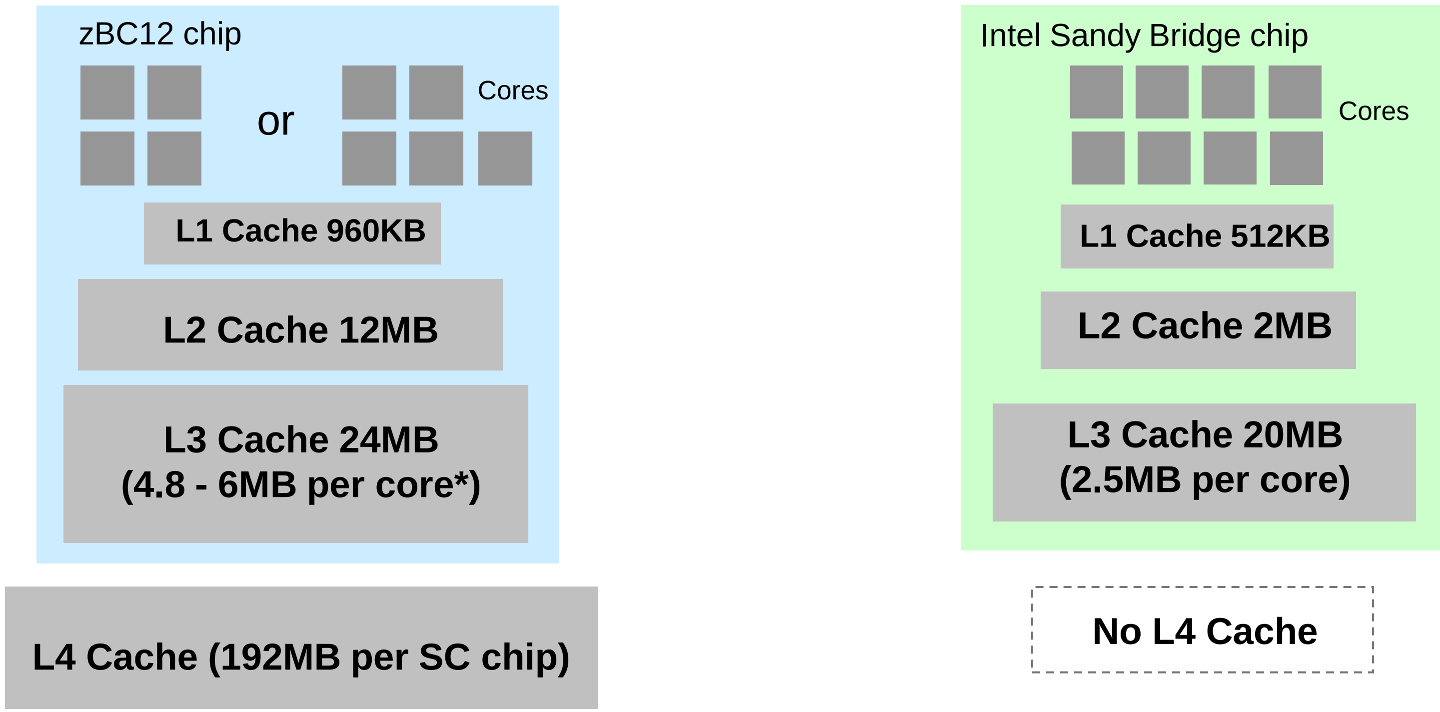
Variability in demand

Different size servers

Workload Management

Mix workloads with different priorities

Like zEC12, new zBC12 has larger cache structures to support more concurrent workloads



Advantages of large cache:

- Fewer cache misses help maintain thread processing speed
- Improves database performance by holding larger working sets
- Improves consolidated workload performance by supporting more working sets

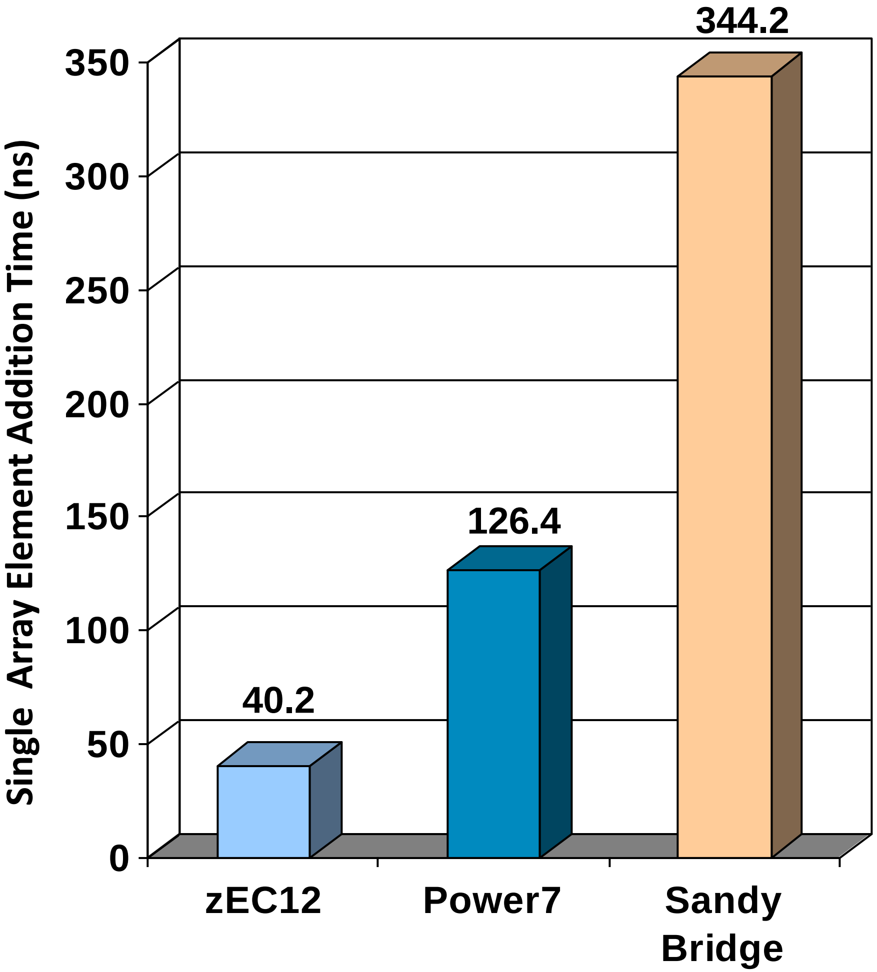
* Six core PU chips using 4 and 5 active core per PU chip. 4.8 MB L3 cache if 5 active core per chip. 6MB L3 cache if 4 active core per chip.

Intel servers slow down under cache intensive workloads

- Multiple concurrent processes introduces
  cache contention
  - Example: 5 processes each with 70MB working set size

- Intel workloads significantly slowed due to cache contention

- System z with z/OS showed results 8x faster than Intel system



Larger cache is beneficial for SAP workloads – as well as CICS, VSAM and Batch workloads

Cost advantage for smaller scale SAP database:

SQL Server on Intel

SAP Applications

4 x HP DL980 2.13GHz 4ch/32co

128 DB cores

SAP Applications

DB2 on z/OS

zEC12 with 3 GP + 2 zIIPs

5 cores

29% *lower unit cost*

Database Unit Cost
$86/User

# of Users	23,000
Hardware	$0.34M
Software	$1.64M
Total (3 yr. TCA)	$1.98M

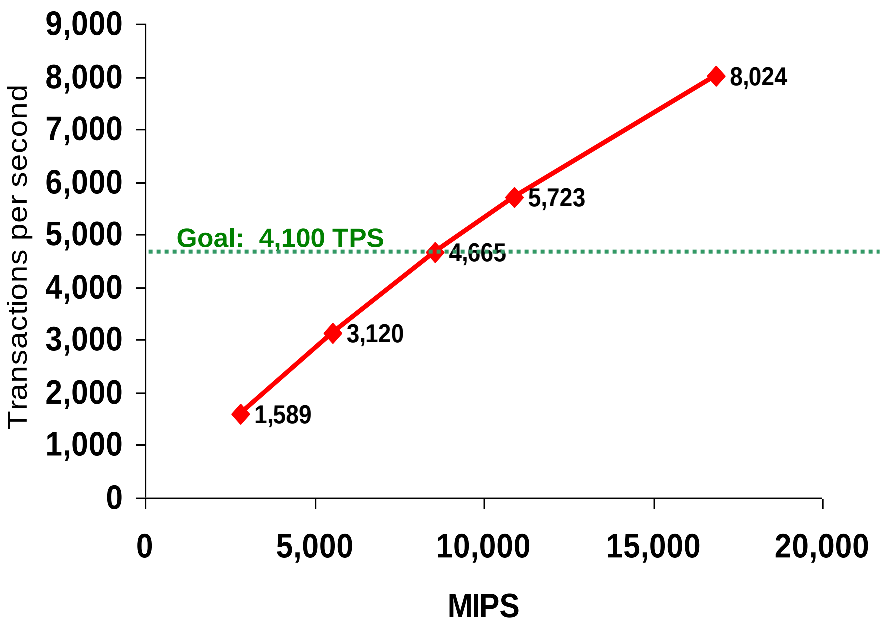
Database Unit Cost
$61/User

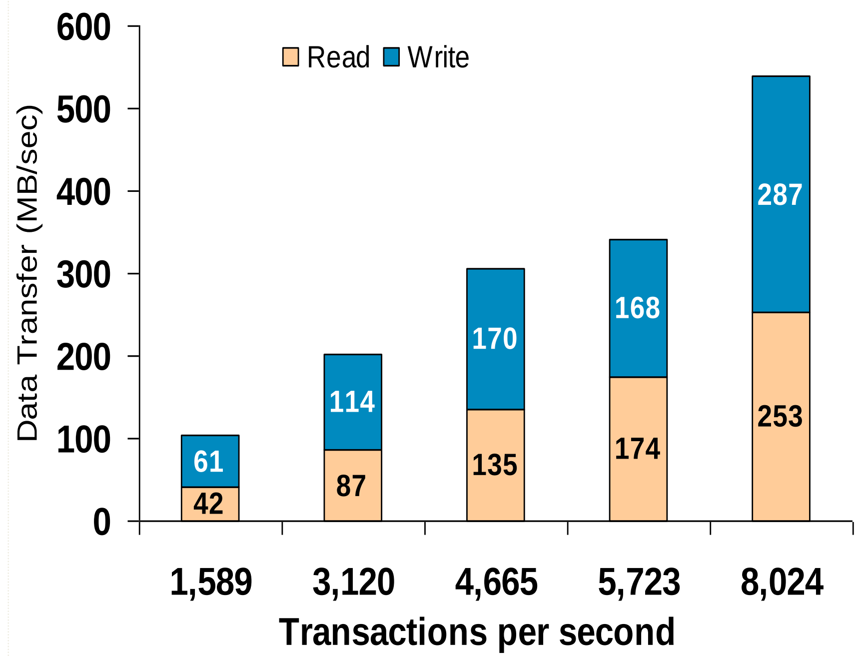
# of Users	23,000
DB2 Solution Edition(HW+SW)	$1.40M
Total (3 yr. TCA)	$1.40M

Note: Workload Equivalence established from a large US Retailer SAP DB offload incorporating estimated CPU Savings from DB2 for z/OS upgrade (107 Performance Units per MIPS). Upgrading from DB2 V8 to V10 reduces average CPU usage by 28%. DB2 V10 for z/OS on zEC12 and SQL Server 2008 on Intel

Dedicated I/O subsystem means System z is ideal for high bandwidth workloads

Capacity benchmark for Bank of China:



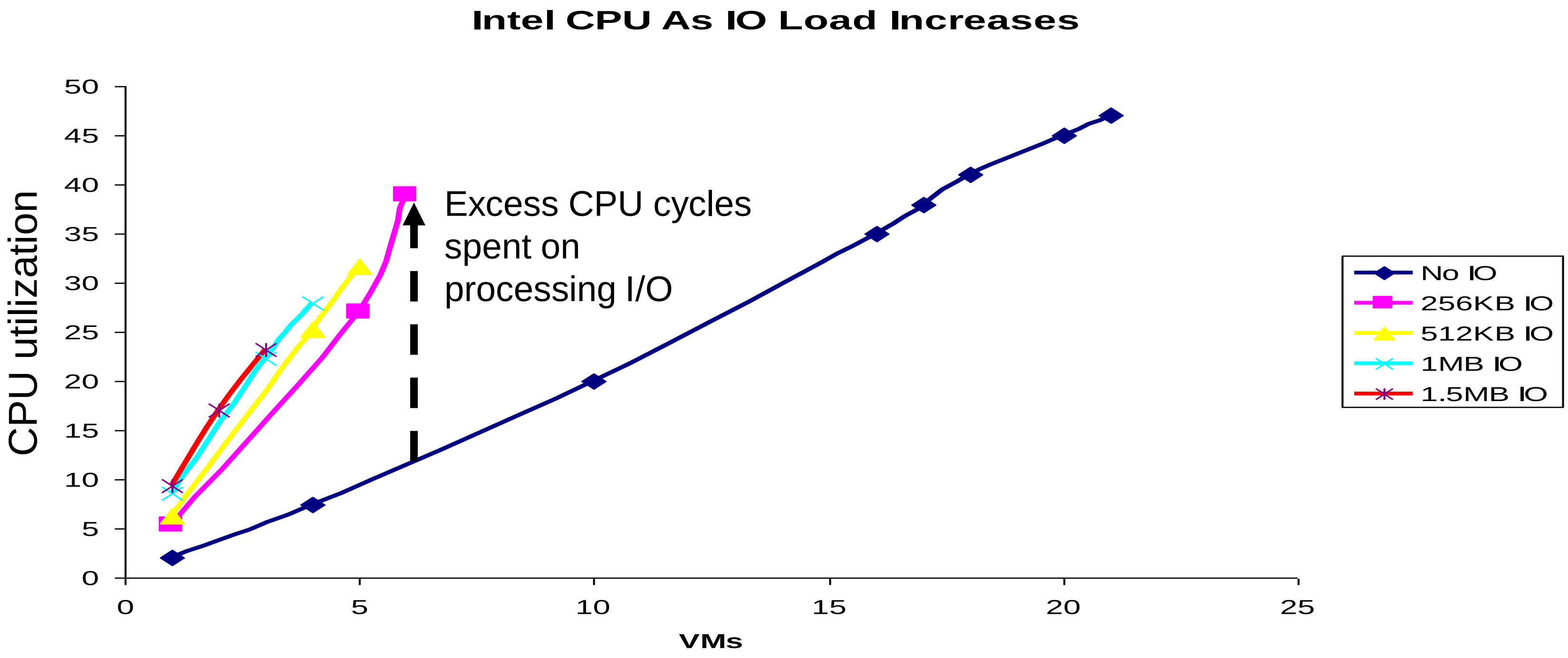


System z easily surpassed benchmark goal, and demonstrates near linear scalability

Reads and writes are well-balanced and scale linearly, demonstrating no constraints on I/O constraint

Tests show Intel's performance degrades as I/O demand increases

- 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
- With no dedicated I/O subsystem, Intel's performance degrades



Multi-tenant database testing also demonstrates System z's superior ability to handle I/O load

Which platform can achieve the lowest cost per workload?

I/O Intensive Database Workload

Brokerage high volume trading workload, each driving a minimum* of **243** transactions per second on 200GB database

1 workload on 16-core quarter unit

Pre-integrated DB Competitor V2 Multi-Tenant Private Cloud

$2.27M/workload

5 multi-tenant workloads

on zEC12

2 GPs + 2 zIIPs

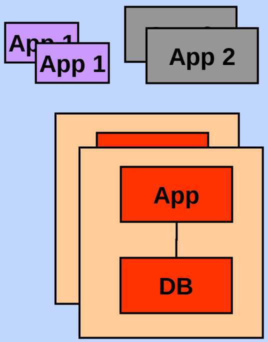
DB2 10 for z/OS on zEC12

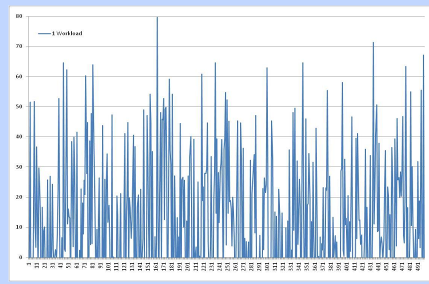
$1.73M/workload

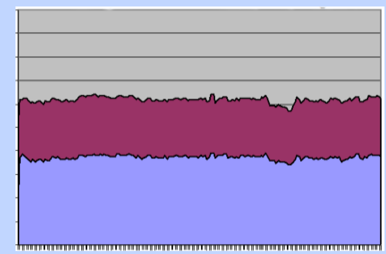
25%
lower cost!

* 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

Platform differences and atomic benchmarks set a baseline for establishing equivalence







Platform factors

GHz, cache, I/O, co-location

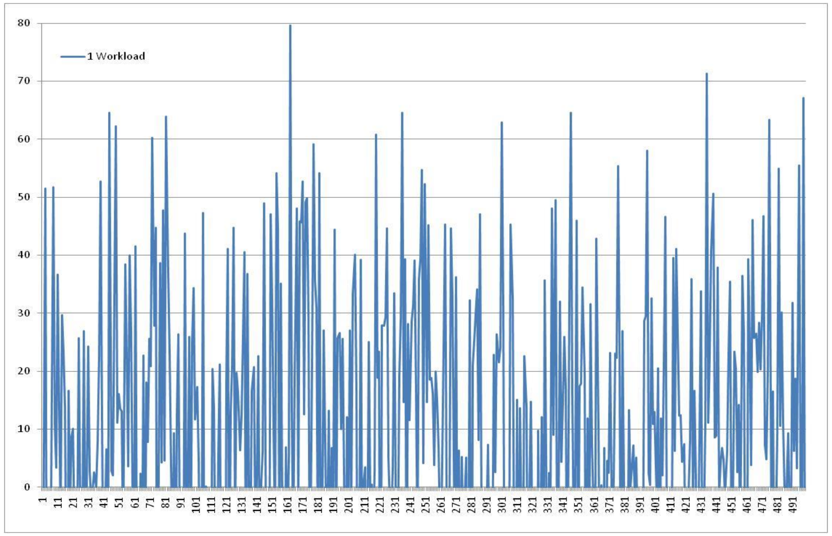
Variability in demand

Different size servers

Workload Management

Mix workloads with different priorities

Larger servers with more resources make more effective consolidation platforms



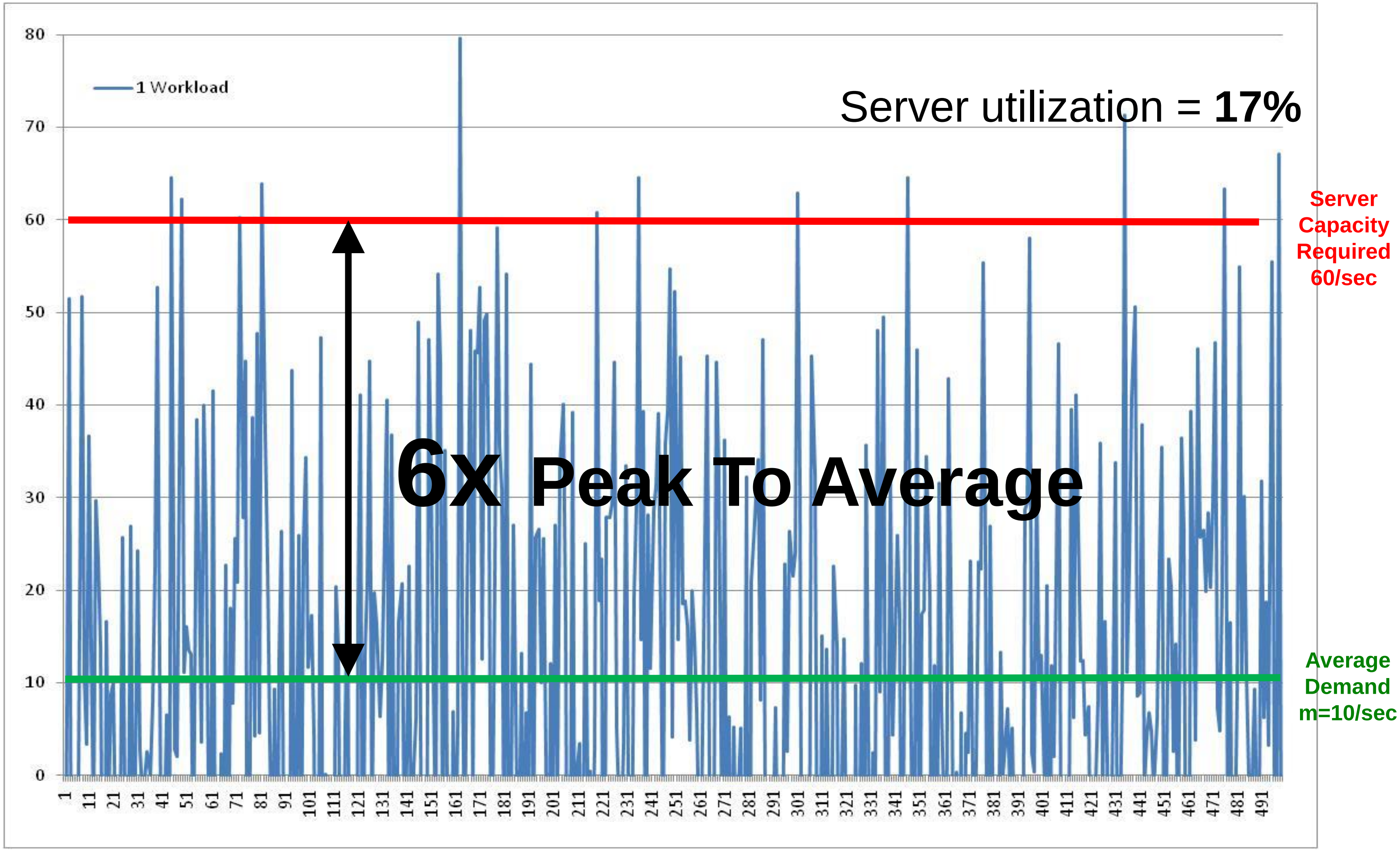
- Most workloads experience variance in demand

- 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

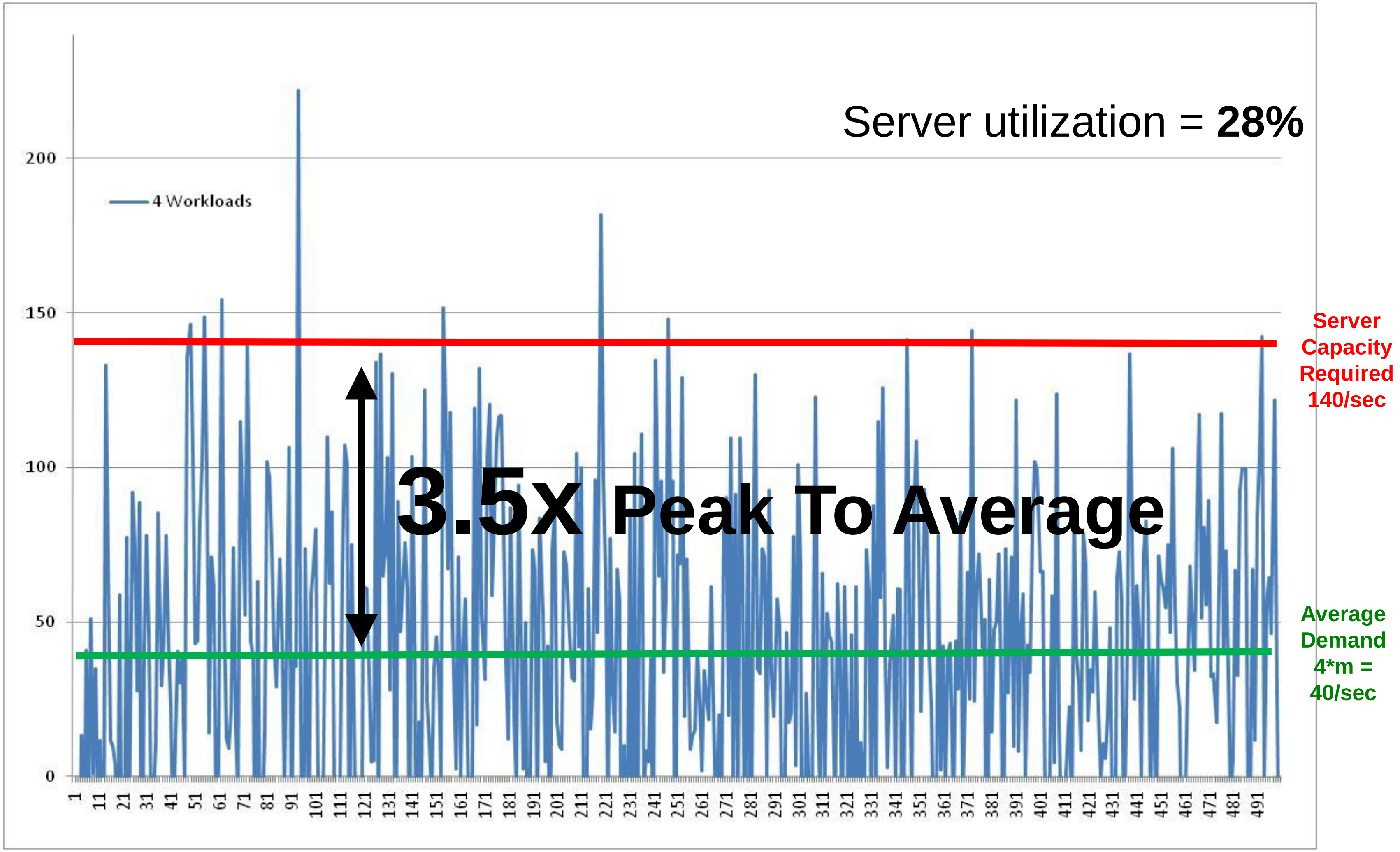
- Consequently, bigger servers with capacity to run more workloads can be driven to higher average utilization levels without violating service level agreements, thereby reducing the cost per workload

A single workload requires a machine capacity of 6x the average demand



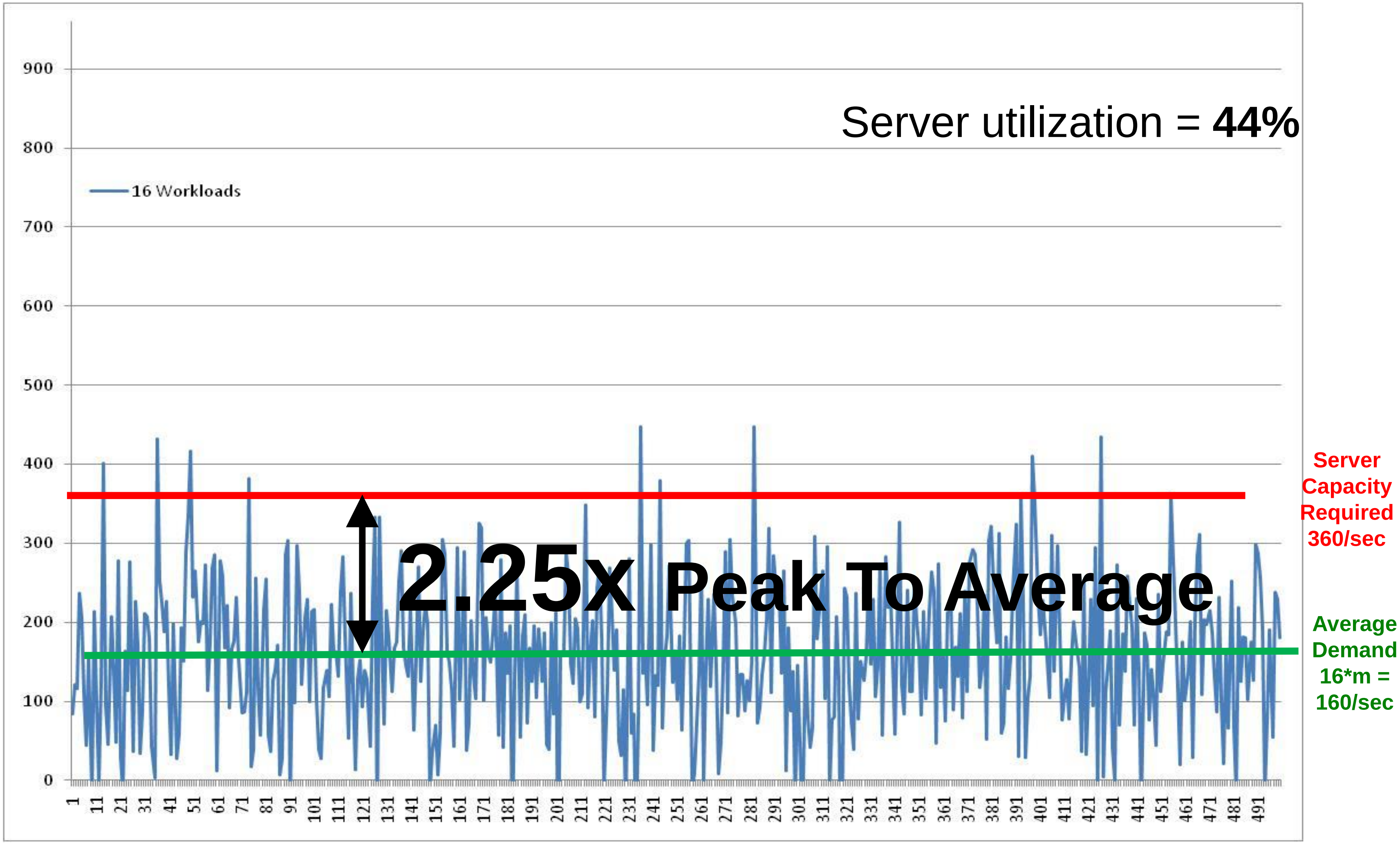
Assumes coefficient of variation = 2.5, required to meet 97.7% SLA

Consolidation of 4 workloads requires server capacity of 3.5x average demand



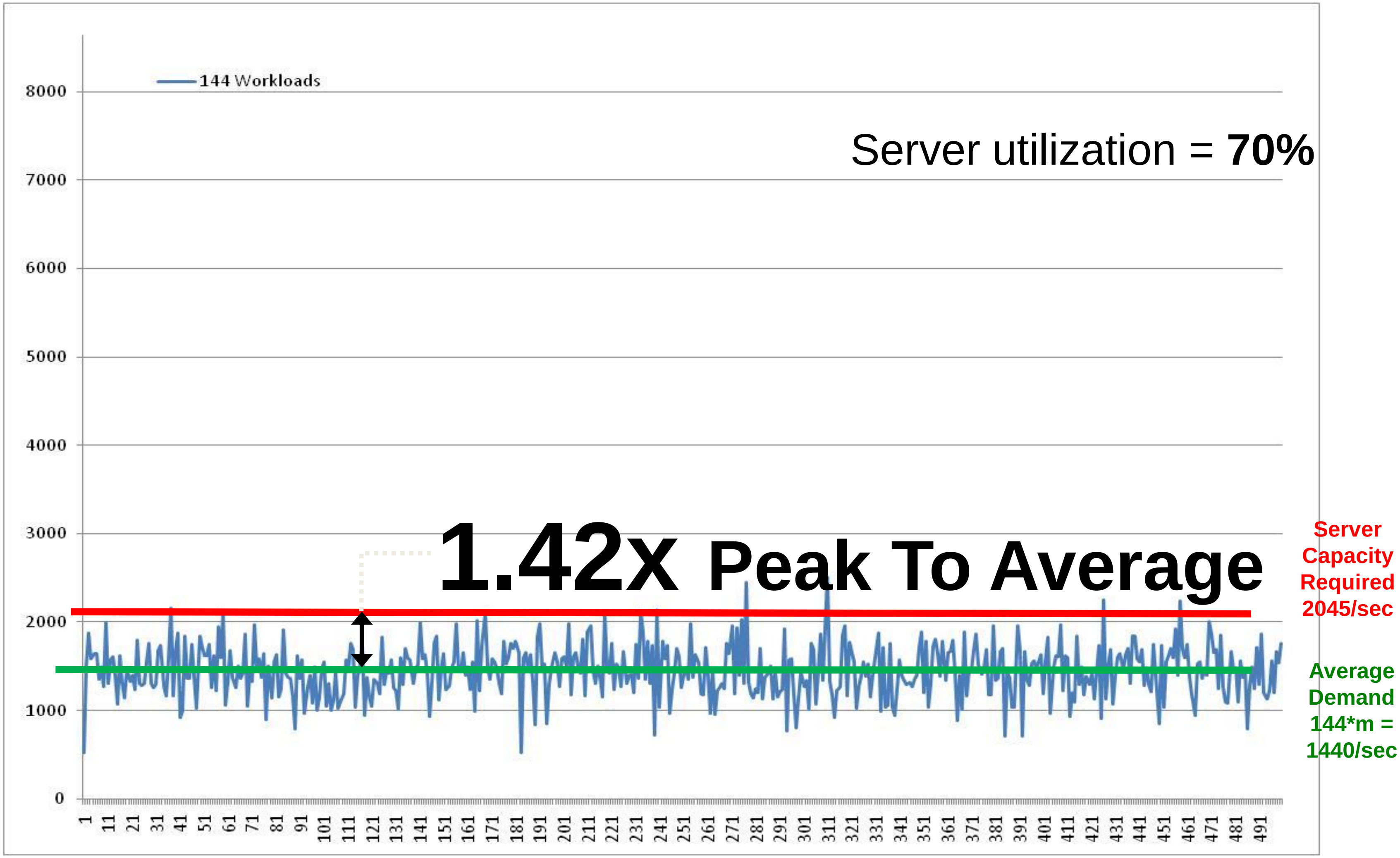
Assumes coefficient of variation = 2.5, required to meet 97.7% SLA

Consolidation of 16 workloads requires server capacity of 2.25x average demand



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

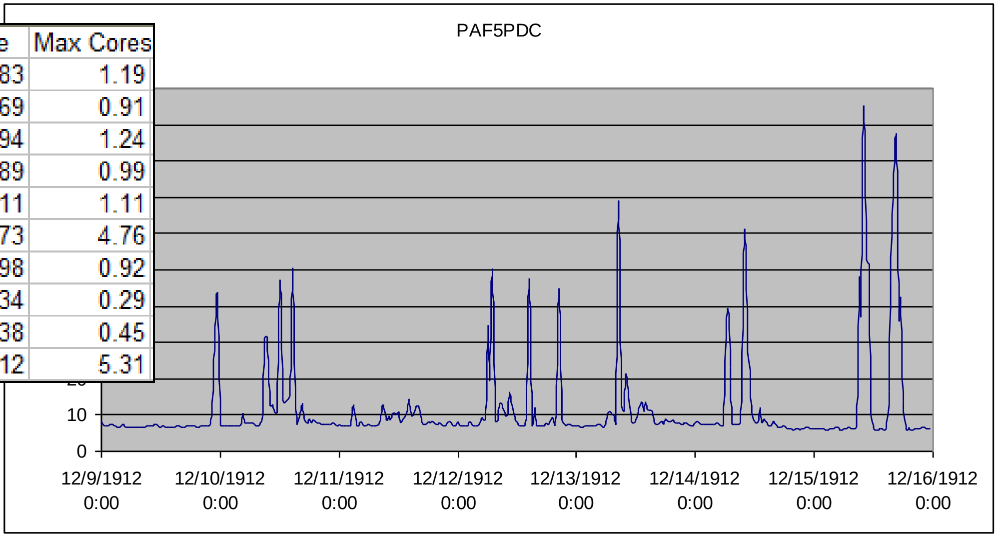
Consolidation of 144 workloads requires server capacity of 1.42x average demand



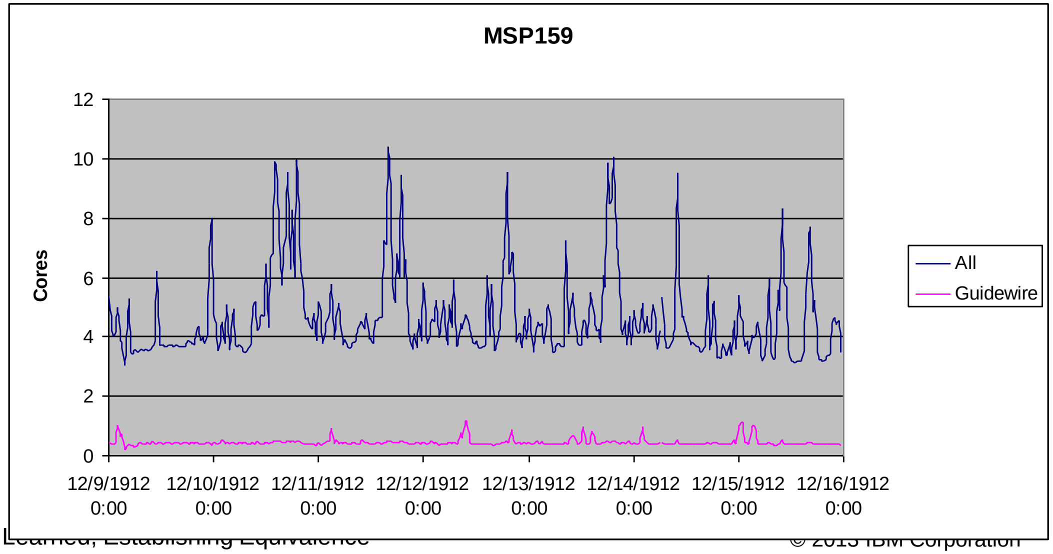
Assumes coefficient of variation = 2.5, required to meet 97.7% SLA

Actual data from a POWER customer demonstrates how statistical multiplexing applies to all large scale virtualization platforms

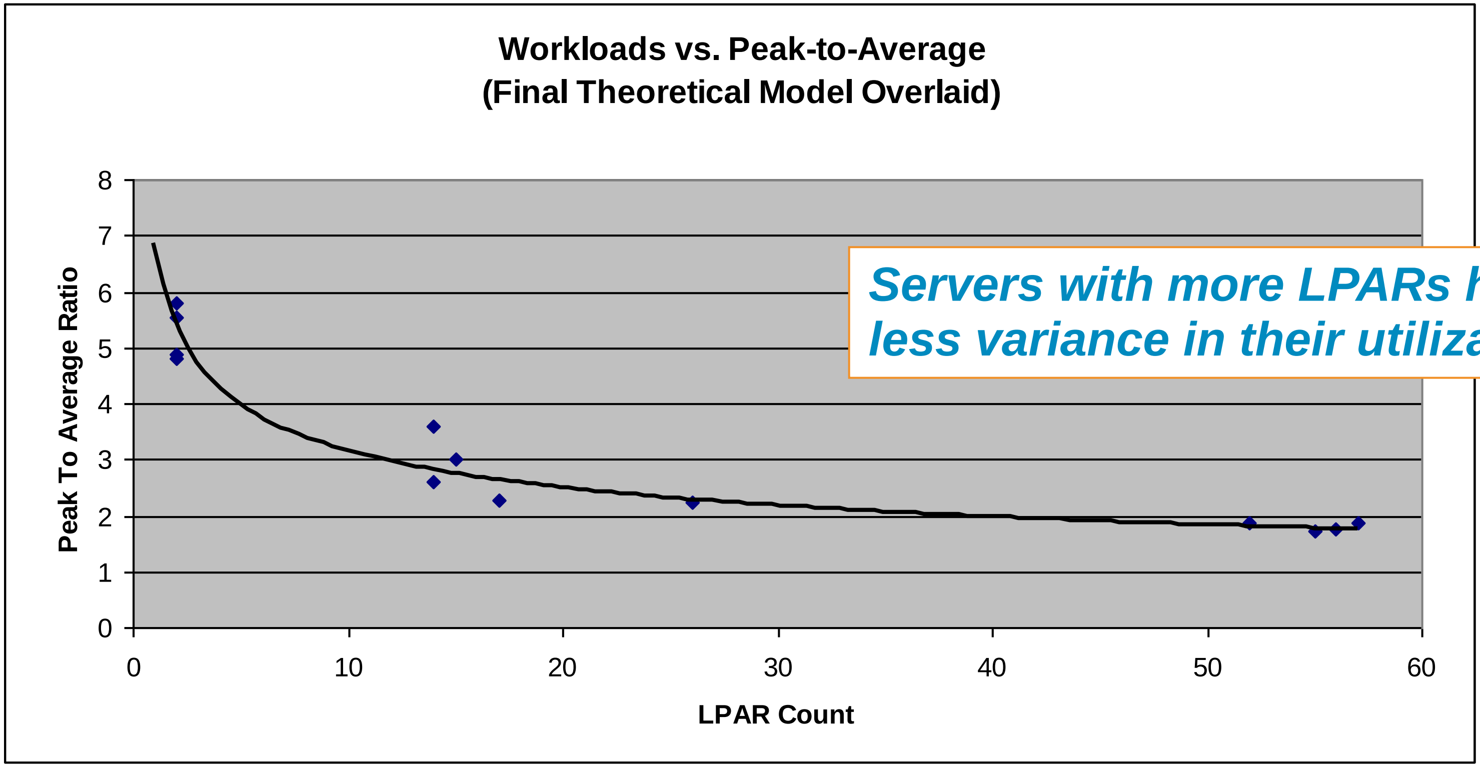
Frame	LPAR	Min	Max	Std. Dev.	Average	Variance	Max Cores
MSP159	PA3APDC	10.44	59.57	6.46	22.37	0.83	1.19
MSP159	PC2APDC	14.40	45.29	5.19	19.11	0.69	0.91
MSP159	PC18PDC	10.36	41.48	5.19	14.45	0.94	1.24
MSP159	PB5BPDC	9.49	32.92	3.23	11.83	0.89	0.99
MSP159	PB4EPDC	9.26	37.16	3.54	11.57	1.11	1.11
MSP159	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



- 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



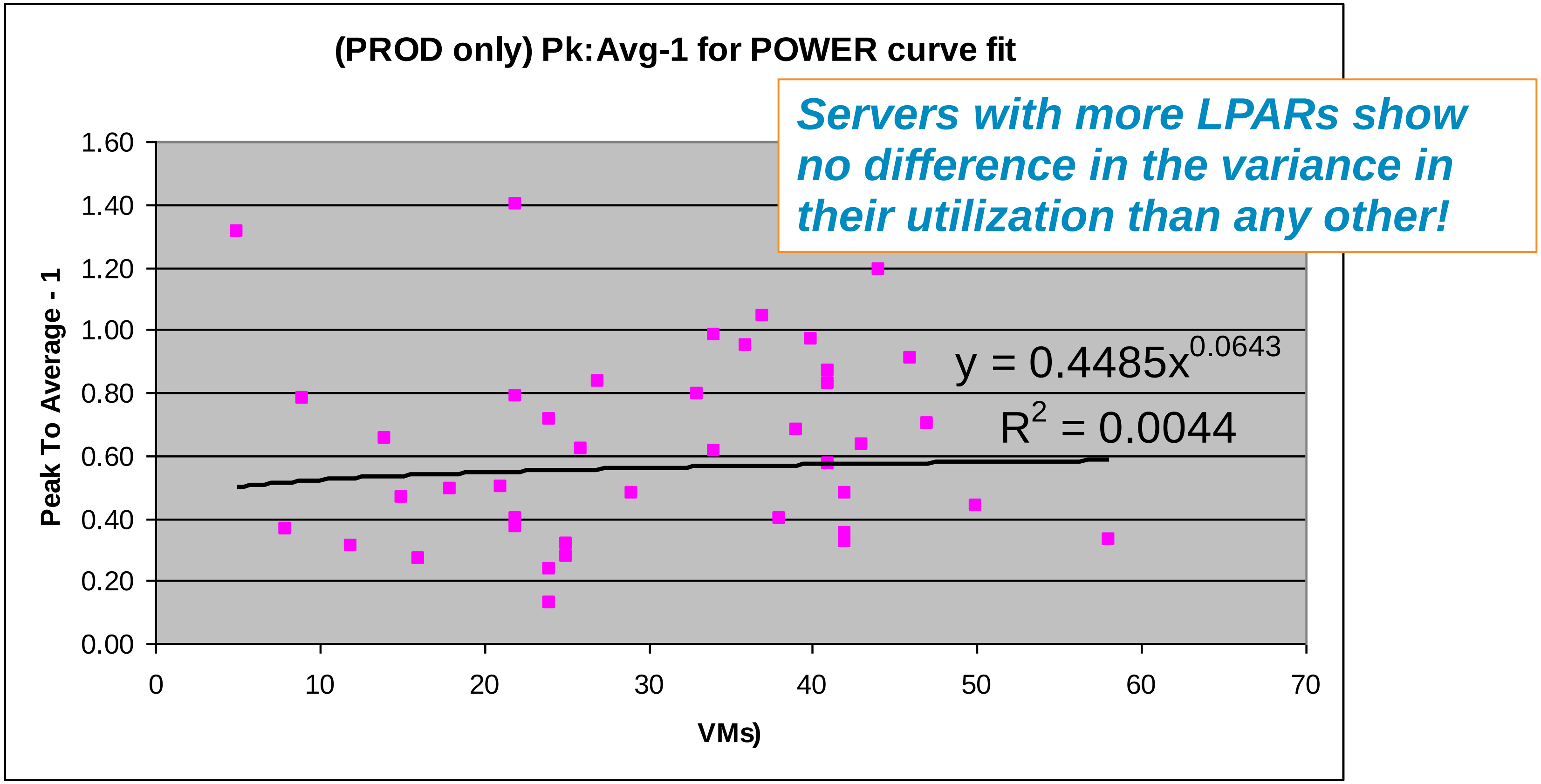
Customer data confirms statistical multiplexing theory



Servers with more LPARs have less variance in their utilization!

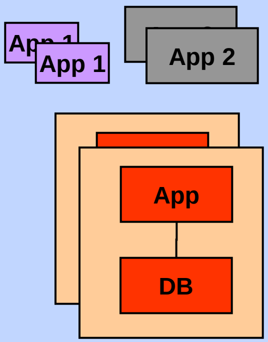
- The larger the shared processor pool, the greater the statistical benefit
- Large scale virtualization platforms are able to consolidate large numbers of virtual machines because of this
- Servers with capacity to run more workloads can be driven to higher average utilization levels without violating service level agreements

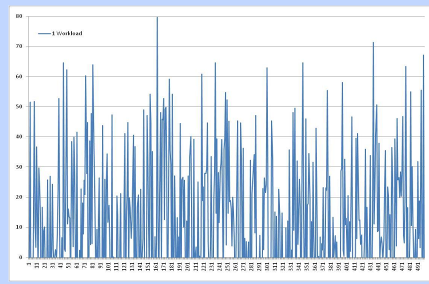
Just for grins, what does a typical x86 VMware “consolidation” look like?

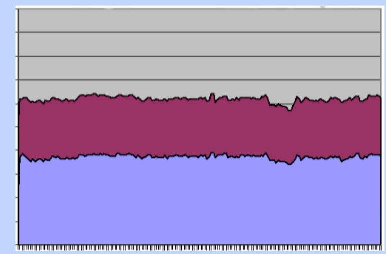


- This looks more like a random number generator than a consolidation curve!
- No apparent correlation between workloads and consolidation...

Platform differences and atomic benchmarks set a baseline for establishing equivalence







Platform factors

GHz, cache, I/O, co-location

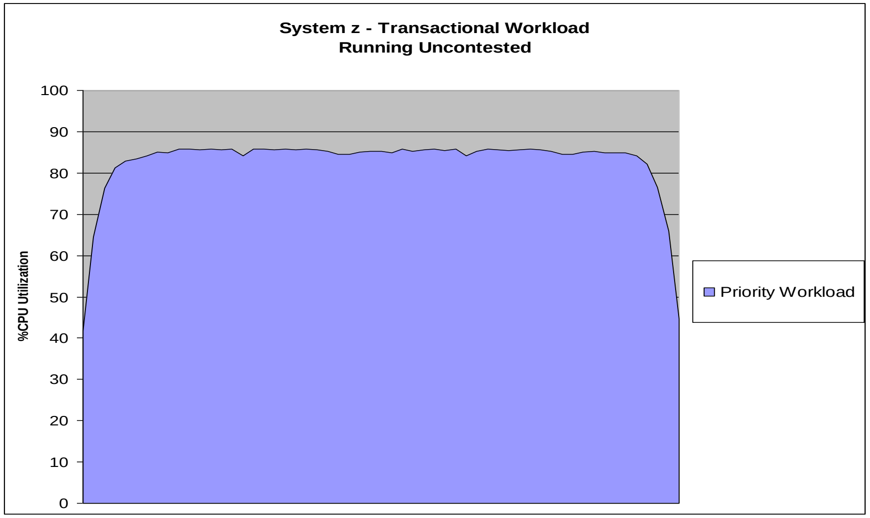
Variability in demand

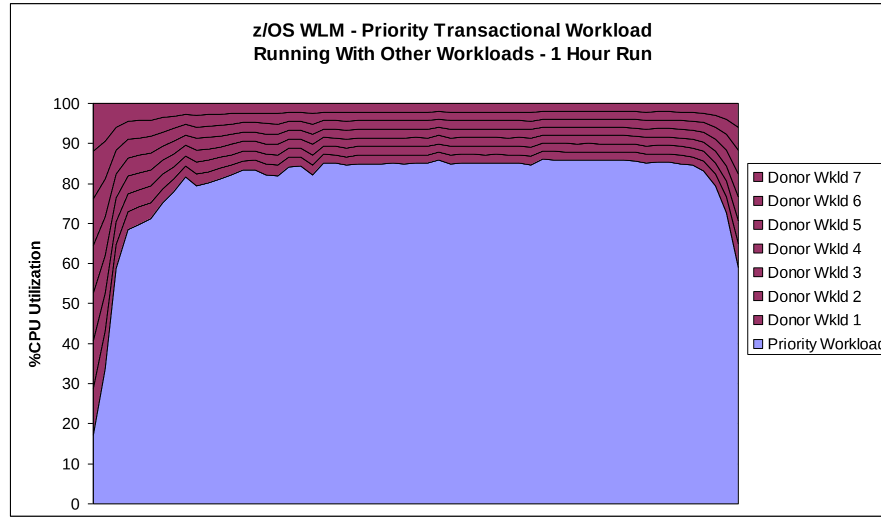
Different size servers

Workload Management

Mix workloads with different priorities

Priority transactional workload does not degrade when low priority workloads added





Capacity Used
High Priority Steady State - 85.2% CPU Minutes
Unused (wasted) - 14.8% CPU Minutes

Priority Workload Metrics
Total Throughput: 417.8K
Maximum TPS 129.7

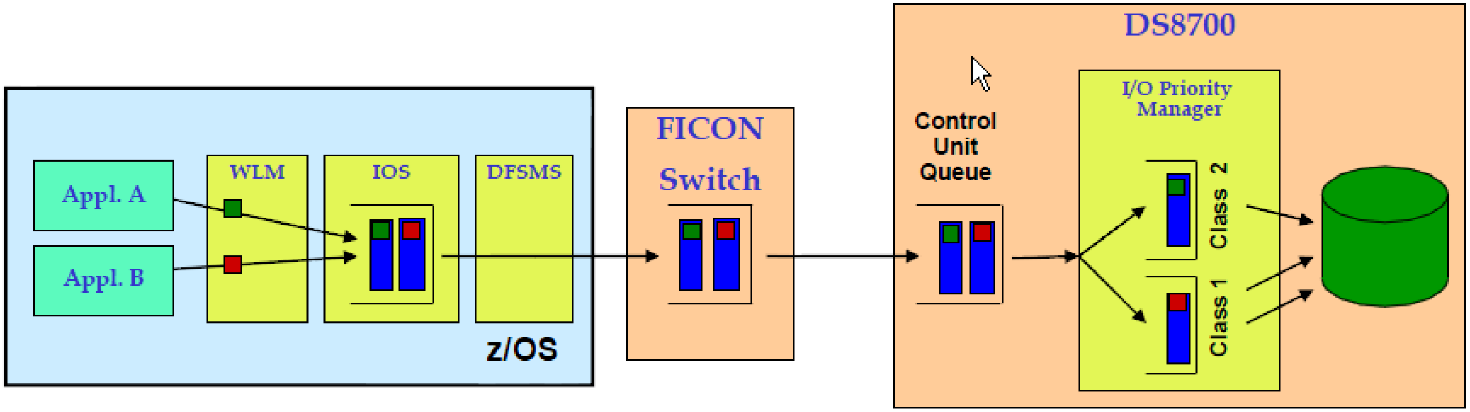
Capacity Used
High Priority Steady State - 85.3% CPU Minutes
Unused (wasted) - 0% CPU Minutes

Priority Workload Metrics
Total Throughput: 414.7K
Maximum TPS 128.1

NO *steady state CPU usage leakage*
1% *total transaction leakage*

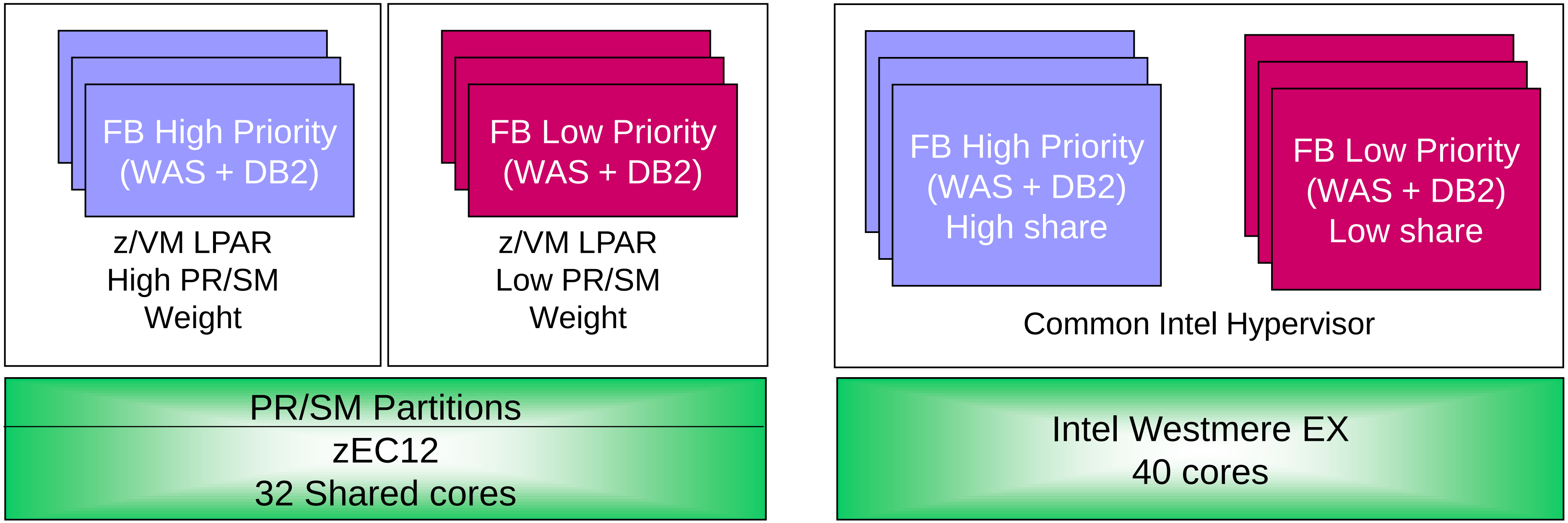
z/OS Workload Manager (WLM) extends priority all the way down to storage

- FICON protocol supports advanced storage connectivity features not found in x86
- Priority Queuing:
  - Priority of the low-priority programs will be increased to prevent high-priority channel programs from dominating lower priority ones

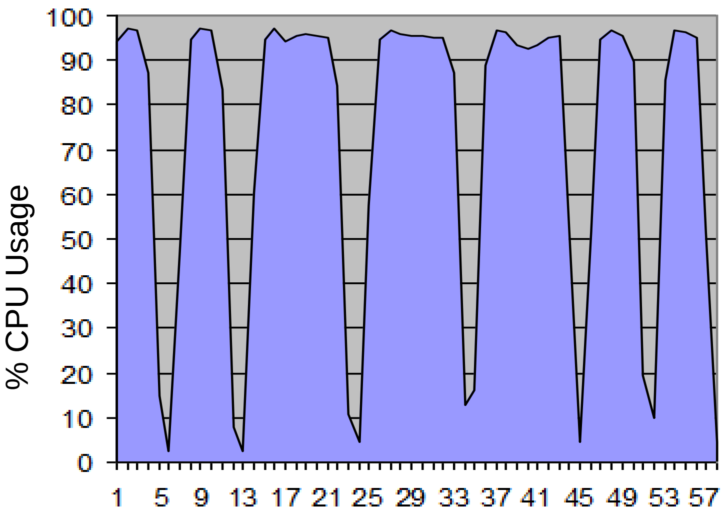


Tests demonstrate comparison of System z PR/SM virtualization to a common hypervisor

- High Priority web workload has defined demand over time
- SLA requires that response time does not degrade
- Low Priority web workload has unlimited demand
- It “soaks up” unused CPU minutes

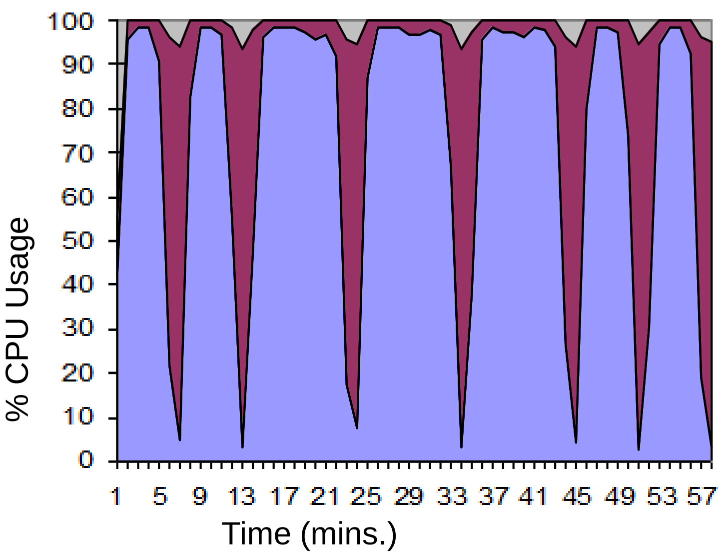


System z demonstrates perfect workload management…



Demand curve for 10 high priority workloads running in 1 z/VM LPAR (PR/SM weight = 99)

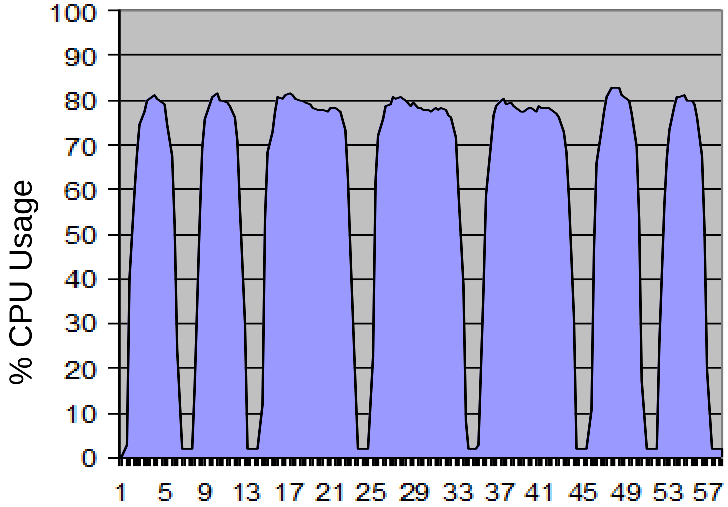
- **Workloads consume 72% of available CPU resources (28% unused)**
- **Total throughput: 9.13M**
- **Average response time: 140ms**



Demand curve when 14 low priority (PR/SM weight = 1) workloads are added in a second z/VM LPAR

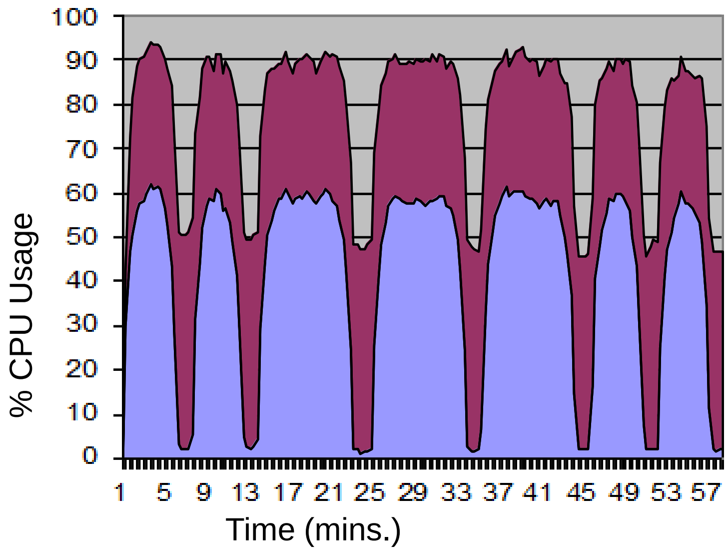
- **All but 2% of available CPU resources is used (high=74%, low=24%)**
- **High priority workload throughput is maintained (9.13M)**
- **No response time degradation (140ms)**

…Unlike this common Intel hypervisor which demonstrates imperfect workload management



Demand curve for 10 high priority workloads running on a common Intel hypervisor (high share)

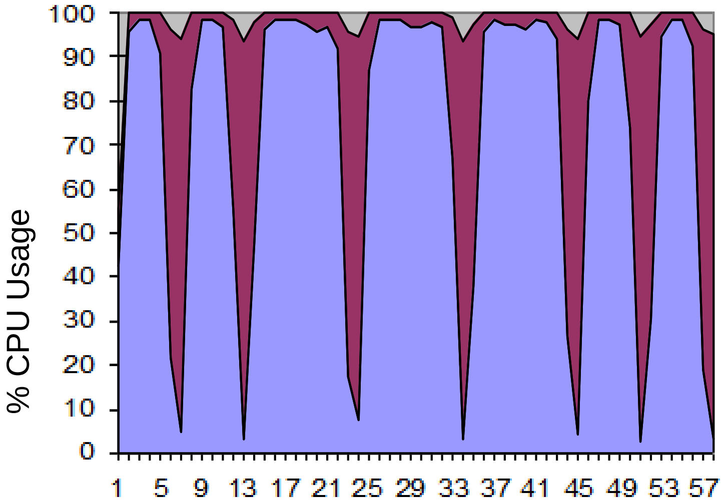
- **Workloads consume 58% of available CPU resources (42% unused)**
- **Total throughput: 6.47M**
- **Average response time: 153ms**



Demand curve when 14 low priority (low share) workloads are added

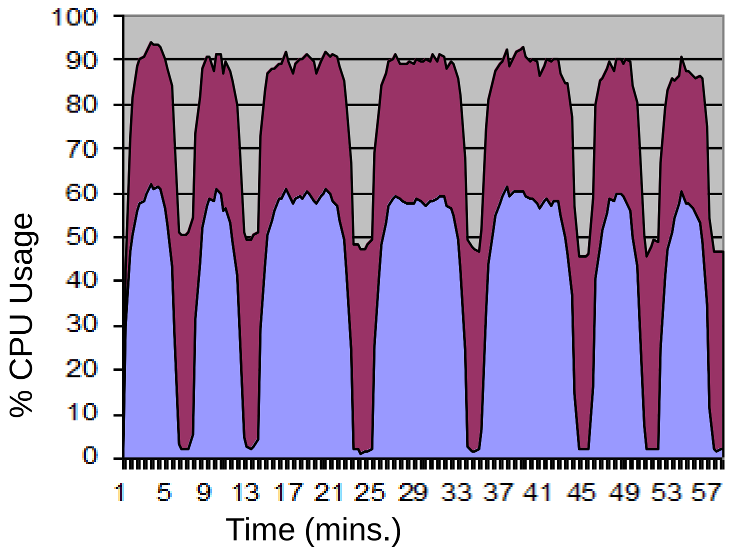
- **22% of available CPU resources is unused (high=42%, low=36%)**
- **High priority workload throughput drops 31% (4.48M)**
- **Response time degrades 45% (220ms)**

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



System z

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



Common Intel hypervisor

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

Imperfect workload management leads to core proliferation and higher costs

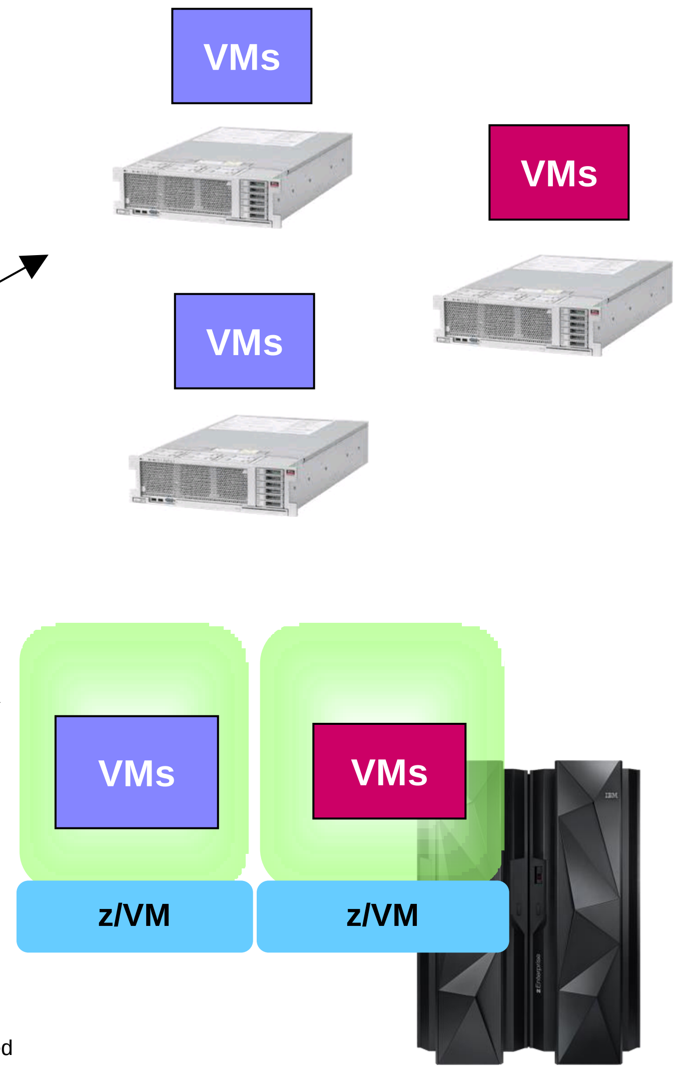
Which platform provides the lowest TCA over 3 years?

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 **9.1M** transactions per hour and low priority discretionary workloads driving **2.8M** transactions per hour



Virtualized on 3 Intel 40 core servers

$13.7M (3 yr. TCA)

z/VM on zEC12
32 IFLs

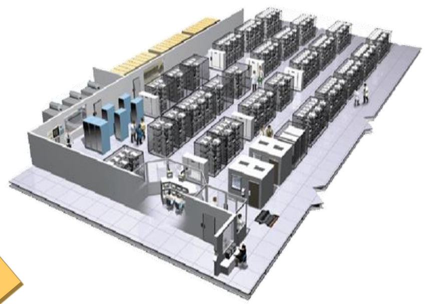
$5.77M (3 yr. TCA)

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.

How can we determine equivalent configurations?

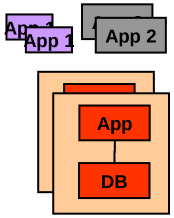
Real world aspects determine accurate equivalence

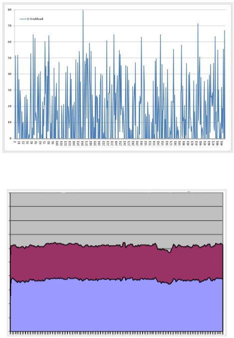


Top Down approach

What we see in customer environments

Bottoms up approach





What we know about platforms and measure in atomic benchmarks

Why are rehosting costs underestimated?

From HP's "Mainframe Alternative Sizing" guide, published in 2012…

MIPS Level	z196 Models	Actual MIPS	z10 EC Models	z10 Actual MIPS	z10 BC Models	z10 BC Actual MIPS	z114 Models	z114 Actual MIPS	HP Cores Estimate	Total HP equivalent MIPS
1,000	2817-701	1,202	2097-701	889	2098-Z02	1250	2818-Z01	782	2	866
2,000	2817-702	2,272	2097-702	1,667	2098-Z03	1784	2818-Z03	2026	5	1,860
3,000	2817-703	3,311	2097-704	3,114	2098-Z05	2760	2818-Z05	3139	8	3,021

Can a 2-chip, quad-core x86-based Blade server really replace 3,000+ MIPS?

- Simple core comparisons are inherently inaccurate…
- Real world use cases suggest this number is off by a factor of **10-20 times**

Eagle TCO study shows this mid-sized workload was *not* cheaper on the distributed platform

6x 8-way (x86) Production / Dev
2x 64-way (Unix) Production / Dev
Application/MQ/DB2/Dev partitions

2x z900 3-way Production / Dev / QA / Test

1,660 MIPS
(6 processors)

?

176 processors

$25.4M (5 yr. TCO)

29x
more cores!

$17.9M (5 yr. TCO)

482 Performance Units per MIPS

Eagle TCO Study shows a pure Intel offload was not cost-effective…

3x HP DL580 (2ch/20co)
Production / Dev / Test
(2011 x86 technology)

z800 Production /
Dev / Test
(2002 mainframe technology)

499 MIPS

(2.1 processors)

60 processors

Despite a 9-year technology gap,
the Intel platform still required
29x more processors

768 Performance Units per MIPS

Is there a cross over point? 1,000 MIPS? 500 MIPS?

A sampling of Eagle TCO data suggests there is no minimum MIPS value that automatically makes an offload financially beneficial…

Customer	z (MIPS)	distributed (PUs)	5-Year TCO z	5-Year TCO distributed	5-Year TCO z/dist %
Average	**1,166**	**218,472**	**9,050,451**	**16,325,492**	
SA Government Agency	475	241,291	19,773,442	25,261,624	78.27%
German Financial	1,200	263,177	3,939,889	4,701,033	83.81%
NA Financial Services	2,526	308,144	3,456,611	5,939,476	58.20%
US utility company	456	163744	6,157,295	13,380,866	46.02%
European Insurance	904	171,062	13,019,980	15,877,484	82.00%
US Manufacturer	900	453,168	11,277,266	16,019,269	70.40%
Asian Bank	1,416	136,013	2,342,300	7,237,681	32.36%
US Retailer	1,700	215,124	3,543,154	8,951,851	39.58%
US County Government	88	43,884	4,717,394	8,108,668	58.18%
US Retailer	1,500	184,732	9,254,186	20,861,515	44.36%
AP bank	1,336	168,113	17,300,000	27,200,000	63.60%
AP bank	300	24,162	5,200,000	11,500,000	45.22%
US Manufacturer	1,917	261,040	4,758,313	7,350,216	64.74%
US Food Services	1,600	424,952	21,966,475	56,167,206	39.11%

The determining factor is really the *nature* of the workload…

Eagle TCO study shows this small workload was *not* cheaper on the distributed platform

2x 16-way (Unix) Production / Dev / Test / Education
App, DB, Security, Print and Monitoring
4x 1-way (Unix) Admin / Provisioning / Batch Scheduling

z890 2-way Production / Dev / Test / Education
App, DB, Security, Print, Admin & Monitoring

36 processors

332 MIPS
(0.88 processors)

$17.9M (4 yr. TCO)

41x
more cores!

$4.9M (4 yr. TCO)

670 Performance Units per MIPS

Eagle TCO study shows even this VERY small workload was not cheaper on the distributed platform

4x p550 (1ch/2co)
Application and DB

z890 Production / Test

88 MIPS
(0.24 processors)

8 processors

$8.1M (5 yr. TCO)

33x
more cores!

$4.7M (5 yr. TCO)

499 Performance Units per MIPS

What happens in a TCO study?

Workload identified for analysis

frontends

databases

other components

Application XYZ (Prod, Dev, QA)

Deployment Choices

- **Do nothing**
- **Optimize current environment**
- **Deploy on other platforms**

Key steps in analysis

1. Establish equivalent configurations
 - Needed to deliver workload

2. Compare Total Cost of Ownership
 - TCO looks at different dimensions of cost