Analyzing IT Value and Cost Considerations – Maximizing The Value of Your Mainframe

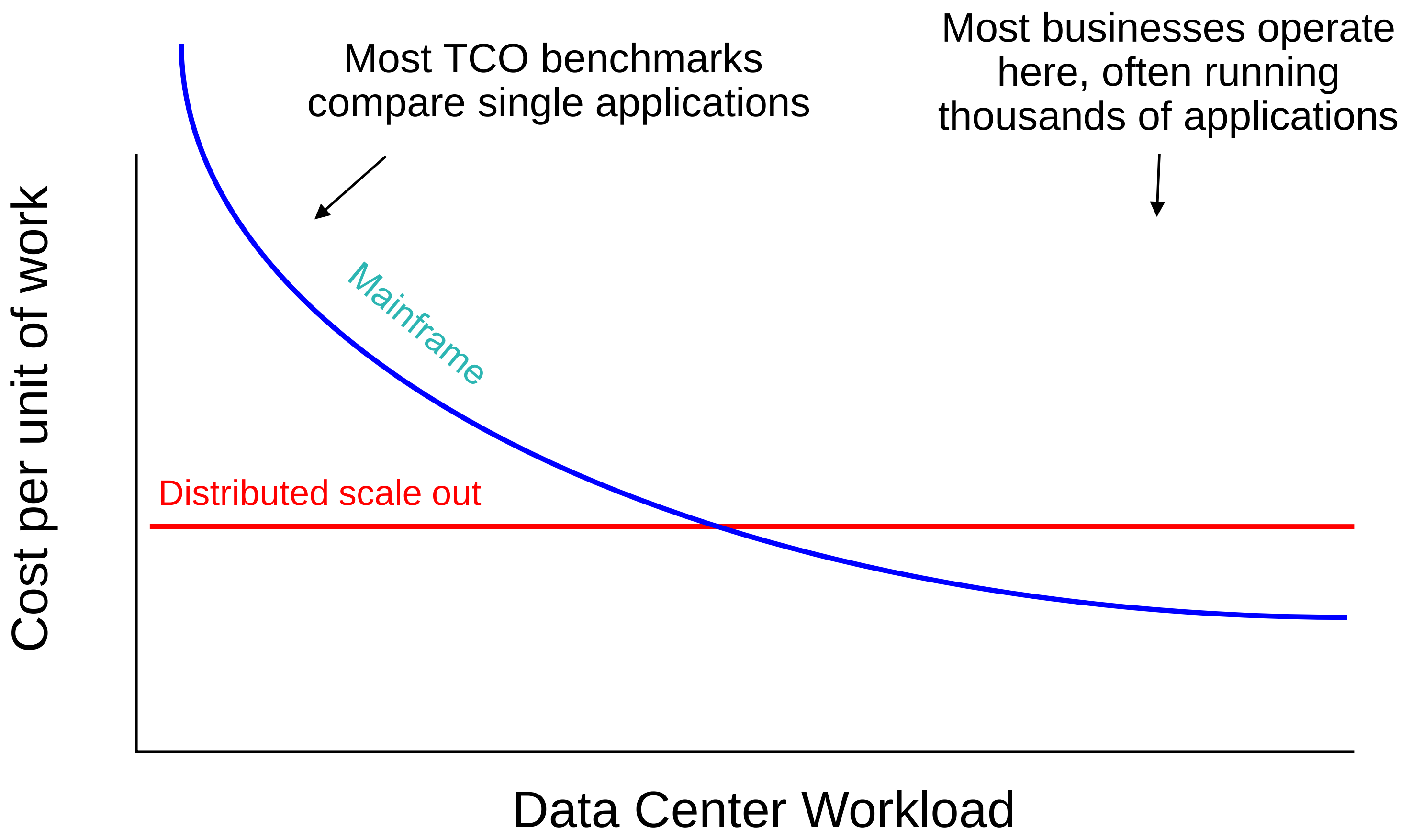
Ray Jones, Vice President,

Worldwide System z Software

IBM Software Group

March 2013

Mainframe Cost/Unit of Work Decreases as Workload Increases



Strategies to achieve breakthrough reductions in IT cost

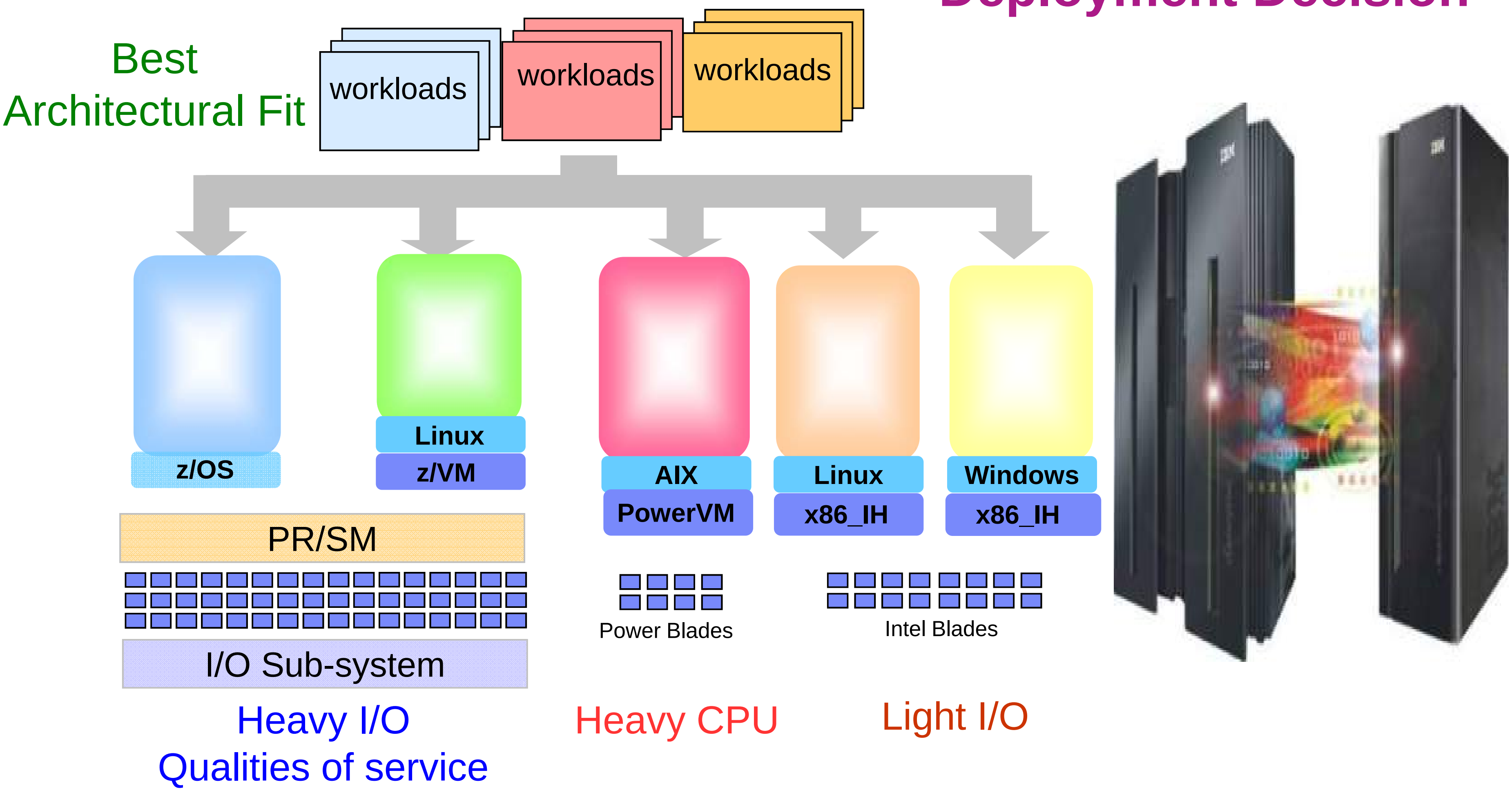
Ascertain true elements of cost:

Hardware/Software/Maintenance
Networking
Energy
Labor
Storage

New metric for the age of Smarter Computing

COST PER WORKLOAD

Workload Characteristics Influence The Best Fit Deployment Decision



Deploy or consolidate workloads on the environment best suited for each workload to yield lowest cost

Deploying Stand Alone Workloads With Heavy CPU Requirements

Benchmark to determine which platform provides the lowest TCA over 3 years

Heavy CPU workloads

- **IBM WebSphere ND**
- **Monitoring software**
- **On 8 core Nehalem servers**

Online banking workloads, each driving **460** transactions per second with light I/O

2 workloads per Intel blade

Scale to 16 cores

Virtualized on Intel
16 core HX5 Blade
$200,055 per workload

Best Fit

1 workload per POWER7 blade

PowerVM on PS701
8 core POWER7 Blade
$216,658 per workload

10 workloads per 32-way z/VM

z/VM on z196 CPC
32 IFLs

$328,477 per workload

Consolidation ratios derived from IBM internal studies. HX5 2.13GHz 2ch/16co performance projected from x3550 2.66GHz 2ch/12co measurements. zBX with x blades is a statement of direction only. Results may vary based on customer workload profiles/characteristics. Prices will vary by country.

Deploying Stand Alone Workloads With Light CPU Requirements

Benchmark to determine which platform provides the lowest TCA over 3 years

Light workloads

- **IBM WebSphere ND**
- **Monitoring software**
- **On 4 core “older” Intel**

Online banking workloads, each driving **22** transactions per second with moderate I/O

47 workloads per Intel blade

Virtualized on Intel
16 core HX5 Blade
$8,165 per workload

28 workload per POWER7 blade

Fast low cost threads

PowerVM on PS701
8 core POWER7 Blade
$7,738 per workload

Best Fit

155 workloads per 32-way z/VM

z/VM on z196 CPC
32 IFLs

$21,192 per workload

Consolidation ratios derived from IBM internal studies. HX5 2.13GHz 2ch/16co performance projected from x3550 2.66GHz 2ch/12co measurements. zBX with x blades is a statement of direction only. Results may vary based on customer workload profiles/characteristics. Prices will vary by country.

Deploying Stand Alone Workloads With Heavy I/O Requirements

Benchmark to determine which platform provides the lowest TCA over 3 years

Heavy I/O workloads

- **IBM WebSphere ND**
- **Monitoring software**
- **On 4 core "Older" Intel**

Online banking workloads, each driving **22** transactions per second, with **1 MB I/O per transaction**

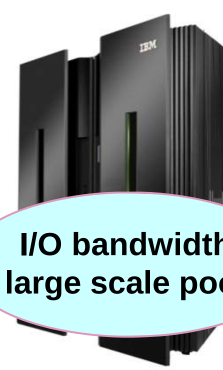
1 workload
per Intel blade

Virtualized on Intel
16 core HX5 Blade
$400,109 per workload

1 workload
per POWER7 blade

PowerVM on PS701
8 core POWER7 Blade
$216,658 per workload

40 workloads
per 32-way z/VM



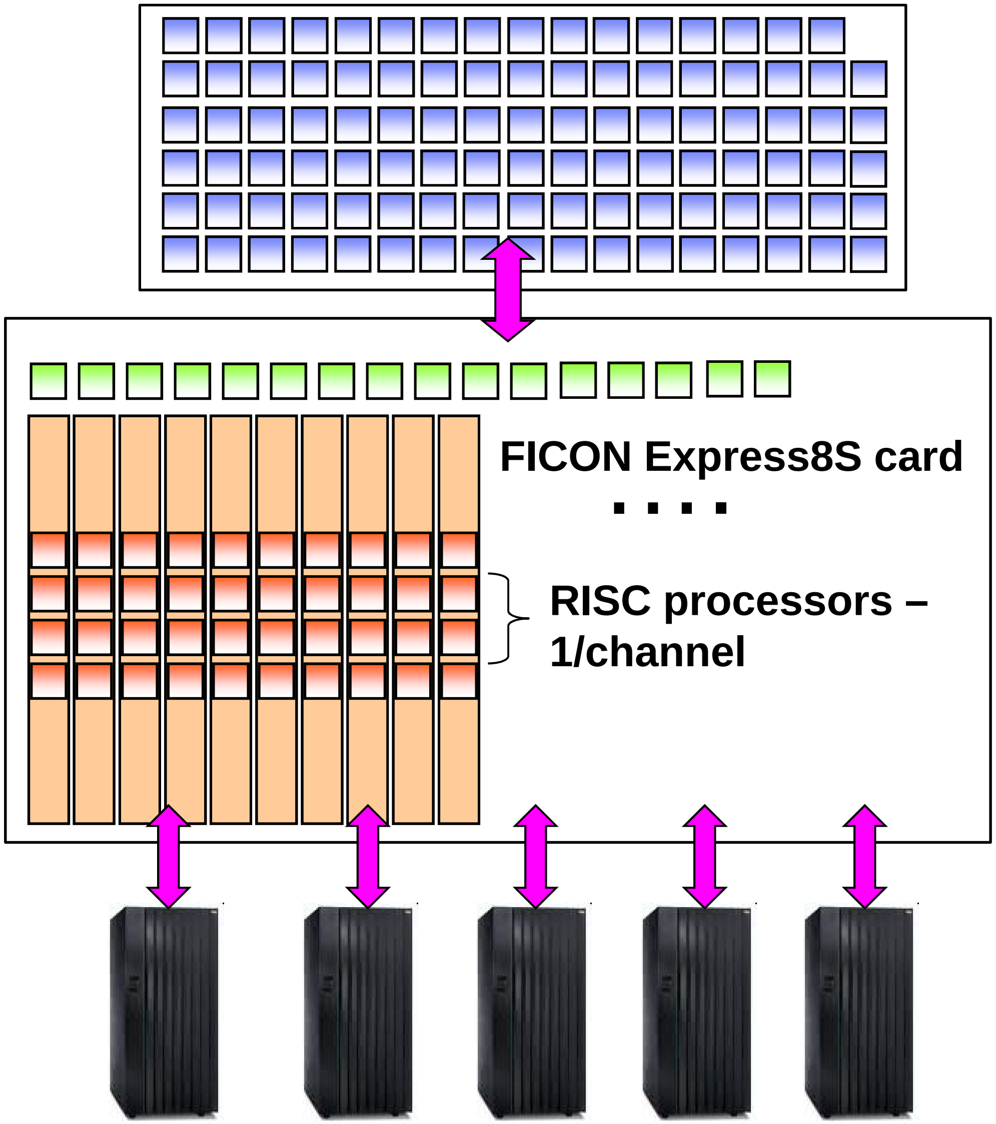
I/O bandwidth large scale pool

z/VM on z196 CPC
32 IFLs
$82,119 per workload

Best Fit

Consolidation ratios derived from IBM internal studies. HX5 2.13GHz 2ch/16co performance projected from x3550 2.66GHz 2ch/12co measurements. zBX with x blades is a statement of direction only. Results may vary based on customer workload profiles/characteristics. Prices will vary by country.

zEnterprise Has A Dedicated I/O Subsystem For High I/O Bandwidth



EC12

- **Up to 101 general purpose processors or Specialty Engines**
  – Execute business logic
- **Up to 16 System Assist Processors to manage I/O requests**
  – Can sustain up to **2.4M IOPS***
- **Up to 160 physical FICON cards for I/O transfers**
  – Up to **320 RISC processors**
- **Up to 1,024 channels**
- **IBM DS8800 Storage System**
  – Up to **440K IOPS capability**
- Delivers efficiency at scale

* Recommend 70% max SAP Utilization – 1.7M IOPS
Numbers represent High Performance FICON traffic

zEnterprise Efficiency At Scale – Lower Cost Per Consolidated Workload

Which platform can achieve the lowest cost per workload?

200GB TPC-E
250 tps

Brokerage TPC-E workload, each driving **250** transactions per second on 200GB database

1 workload
on 16-core
quarter unit

Pre-integrated
Competitor
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

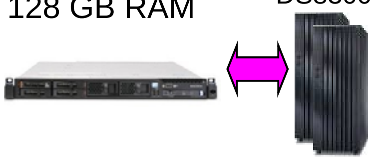
25% lower cost

* Projected from z196 using 1.25 performance factor

Benchmarks Show System z And z/OS Are Optimized For Batch Processing

Intel x3550

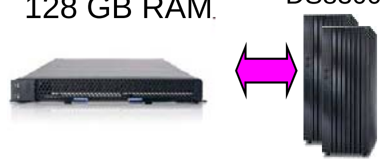
12 processors
128 GB RAM



Sorting Average CPU 89%

Power PS701

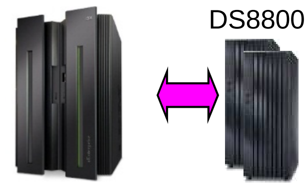
8 processors
128 GB RAM



Sorting Average CPU 92%

Linux on z

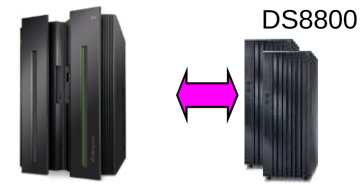
8 processors 128 GB RAM



Sorting Average CPU 90%

z/OS

8 processors 128 GB RAM



Sorting Average CPU 72%

SORT *Job: Sort a 3 GB transaction file – Repetitions: 300*

	Intel x3550	Power PS701	Linux on z	z/OS
Total Time (secs)	7,680	6,900	2,590	644
Concurrency	12	20	18	45
Rate (MB/sec)	240	280	746.2	3,000

MERGE *Job: Merge 30 sorted files into a 90 GB master file – Repetitions: 10*

	Intel x3550	Power PS701	Linux on z	z/OS
Total Time (secs)	11,709	7,920	2,799	558
Concurrency	10	10	10	10
Rate (MB/sec)	157	244	690.5	3,460

Results:

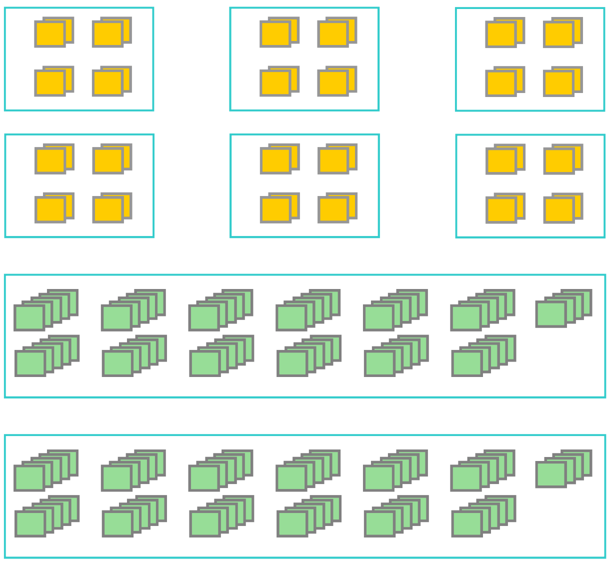
1. **Running same software, x86 batch window is 3.6x greater than System z**
2. **On System z, Linux batch window is 4.5x greater than z/OS**
3. **Off-loading batch from z/OS to x86 leads to as much as 16x increase in batch window**

Source: Internal IBM study

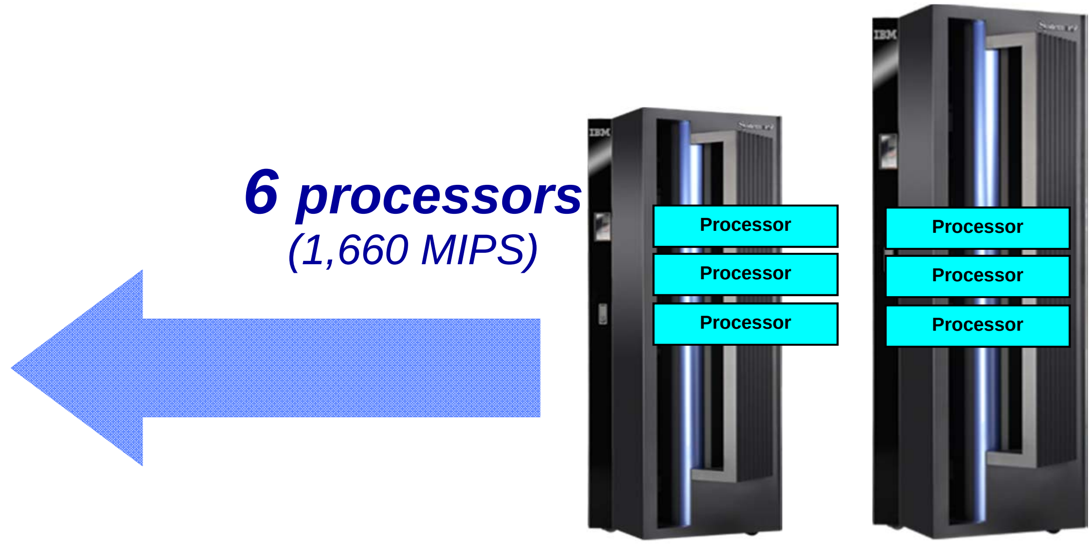
Core Proliferation for a Mid-sized Offload Project

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 distributed processors
(800,072 Performance units)

$25.4M TCO (5yr)

$17.9M TCO (5yr)

**482 Performance Units
per MIPS**

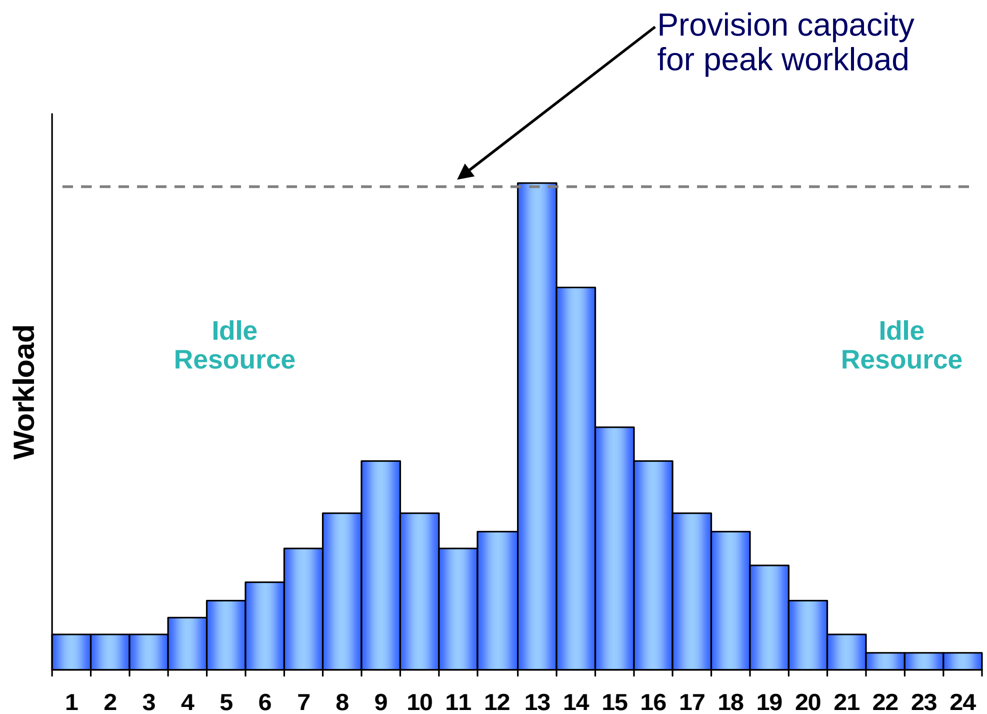
Utilization of Distributed Servers & Storage

Typical utilization of:

Windows Servers	5-10%
UNIX Servers	10-20%
System z Servers	85-100%

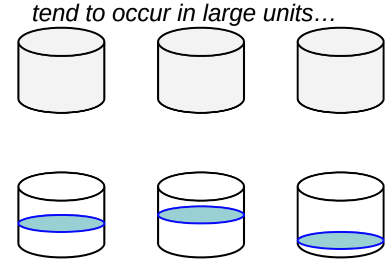
Server dedicated to one application

The cost of storage is typically three times more in distributed environments



- **Storage Allocation**
  - Application-specific resulting in over-allocations
  - Fine grained storage allocation mechanisms characteristic of mainframe storage are uncommon in distributed environments.
- **Storage Utilization**
  - Single digit utilization for distributed environments is not uncommon
  - Storage utilization of 80% + is typical for mainframe
- **Storage Management**
  - Data disaster recovery, synchronization, and transfer requirements add complexity and cost

Application specific storage allocations tend to occur in large units…



resulting typically in single digit utilization

What Is A Typical Value Of Sigma?

IBM Survey Of Workload Variability In 3200 Servers

Type Of Workload	Average Utilization	Peak Utilization	Sigma
Infrastructure	**6%**	**35%**	**2.5 * Mean**
Web Server	**4%**	**24%**	**2.5 * Mean**
Application	**4%**	**34%**	**3.75 * Mean**
Database	**5%**	**37%**	**3.25 * Mean**
Terminal	**6%**	**45%**	**3.25 * Mean**
E-Mail	**4%**	**34%**	**3.75 * Mean**

IBM System x™ Servers and VMware Virtual Machine Sizing Guide

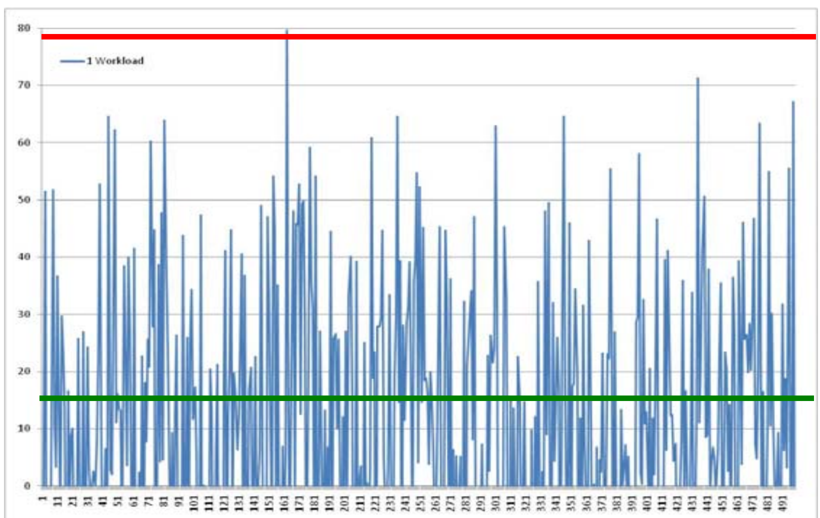
Legacy workloads on XEON 2.5-2.8GHz Servers

Normal probability distribution

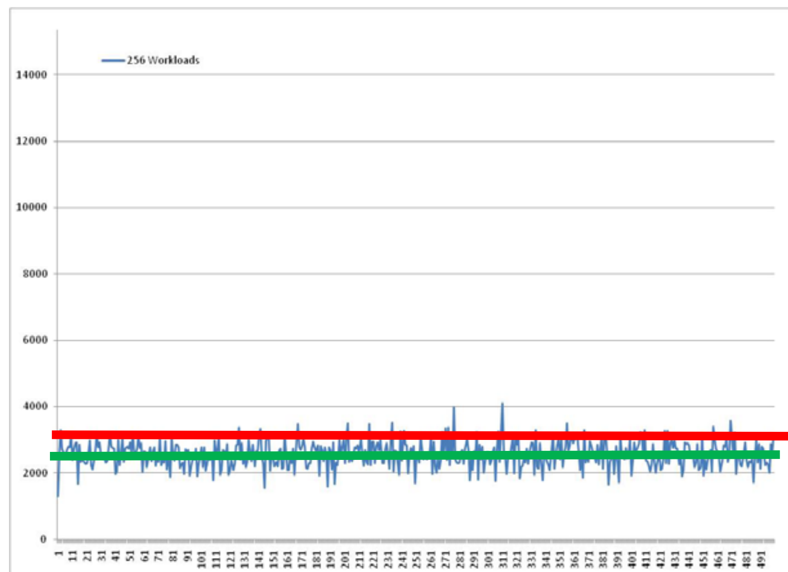
New Workload Scenarios – Beware Benchmarks

- **Stress test benchmarks have no variability!**
  – They drive the system under test to100% utilization with no variation
  – Comparing mean throughputs at 100% utilization doesn't give a realistic view of the resources required for deployment

Running a new workload with variability Sigma=2.5*Mean requires processing capacity equal to **6 times the Mean** workload demand

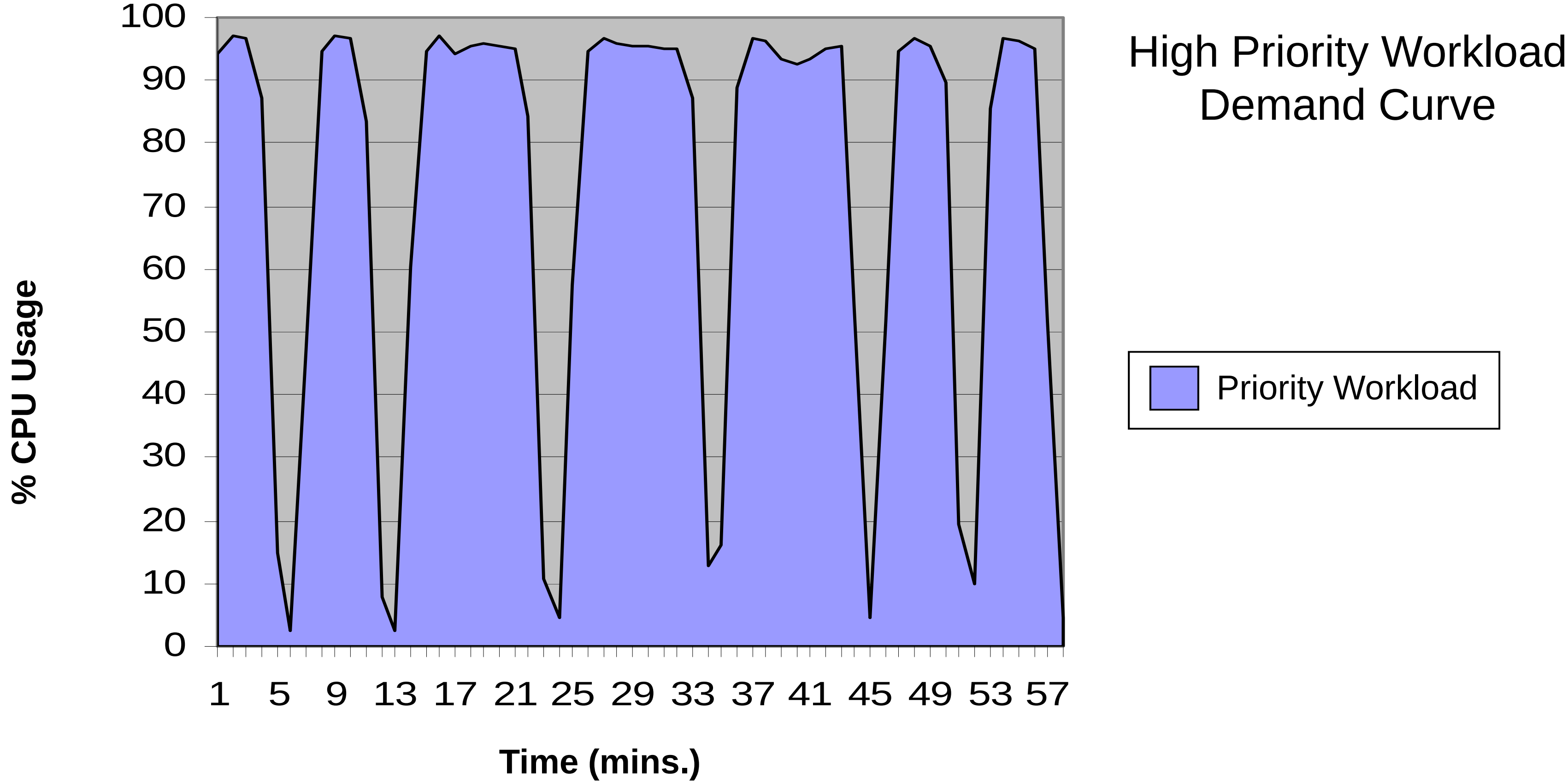


Adding a new workload to a pool of 256 existing workloads will require incremental processing capacity equal* to the **Mean** workload demand



* If we add one more workload to a pool of 256 consolidated workloads the computing resource required for the pool goes up by 1.00047 * Mean

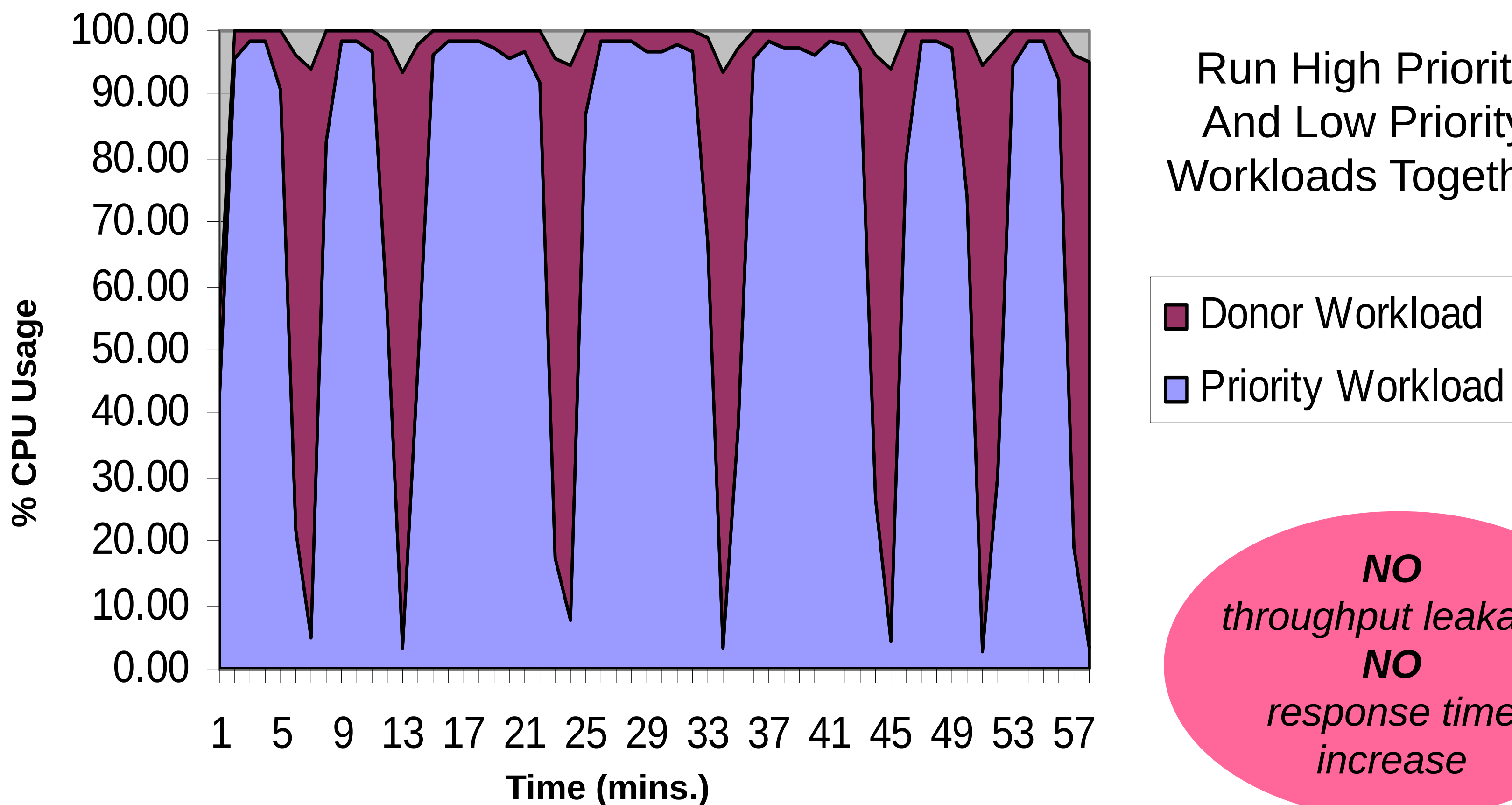
Priority Workload With Varying Demand Running Standalone On System z PR/SM



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

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



Run High Priority And Low Priority Workloads Together

NO* *throughput leakage
NO* *response time increase

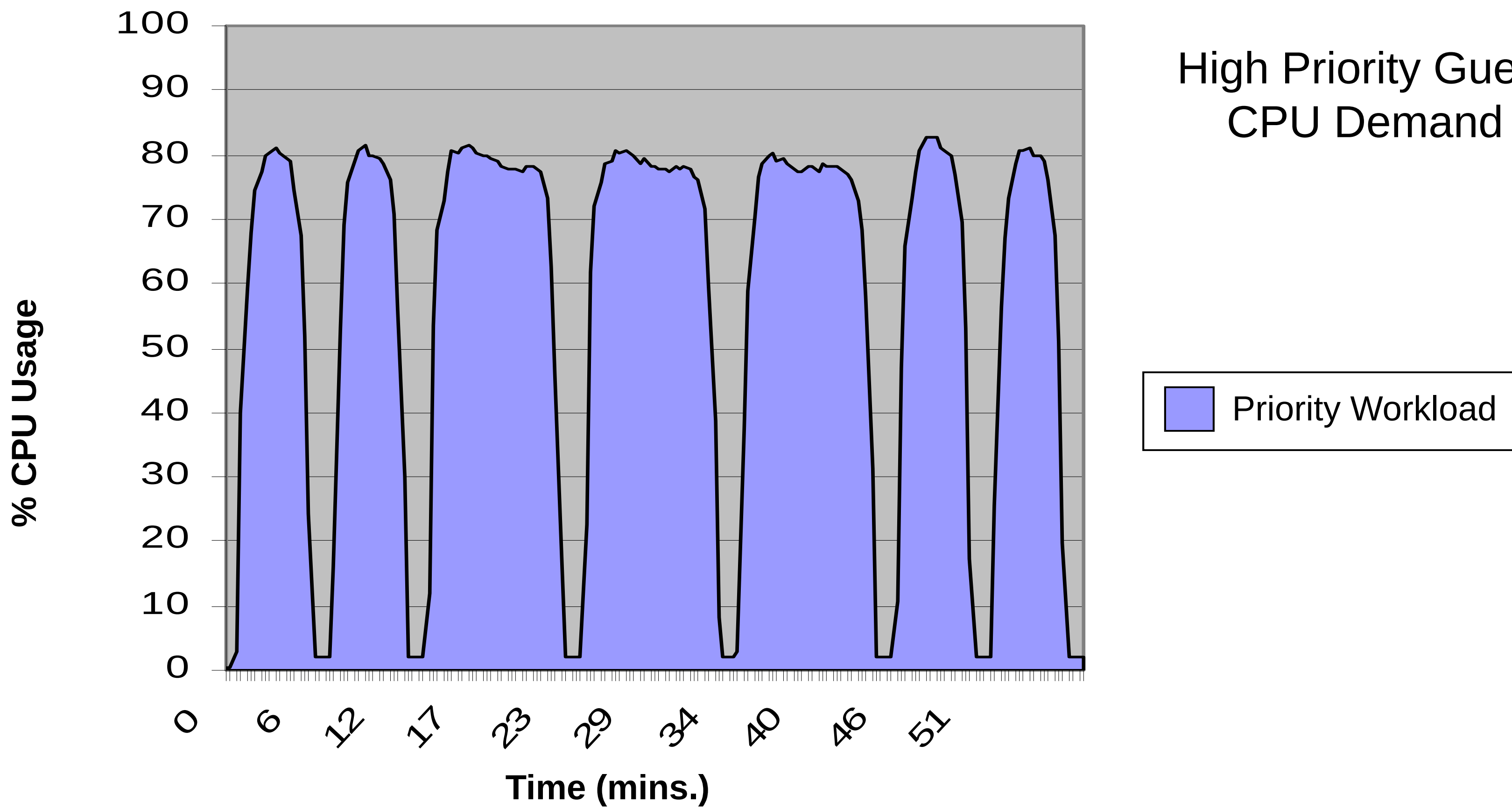
Capacity Used
High Priority - 74.2% CPU Minutes
Low Priority - 23.9% CPU Minutes
Wasted – 1.9% CPU Minutes

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

IBM

Priority Workload With Varying Demand Running Standalone On x86 Hypervisor

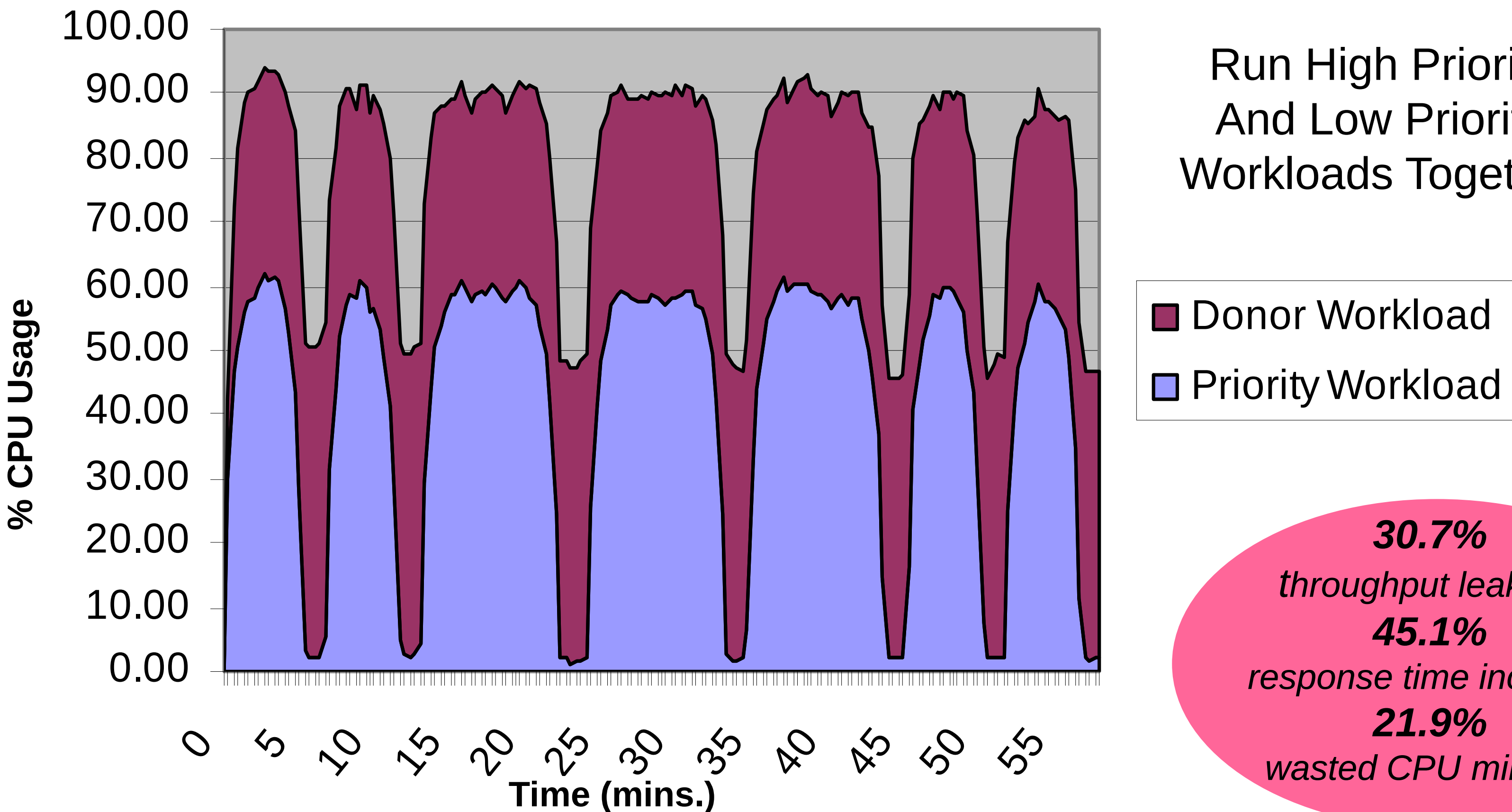
High Priority Guest CPU Demand



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

Priority Workload On x86 Hypervisor Degrades Severely When Low Priority Workload Is Added



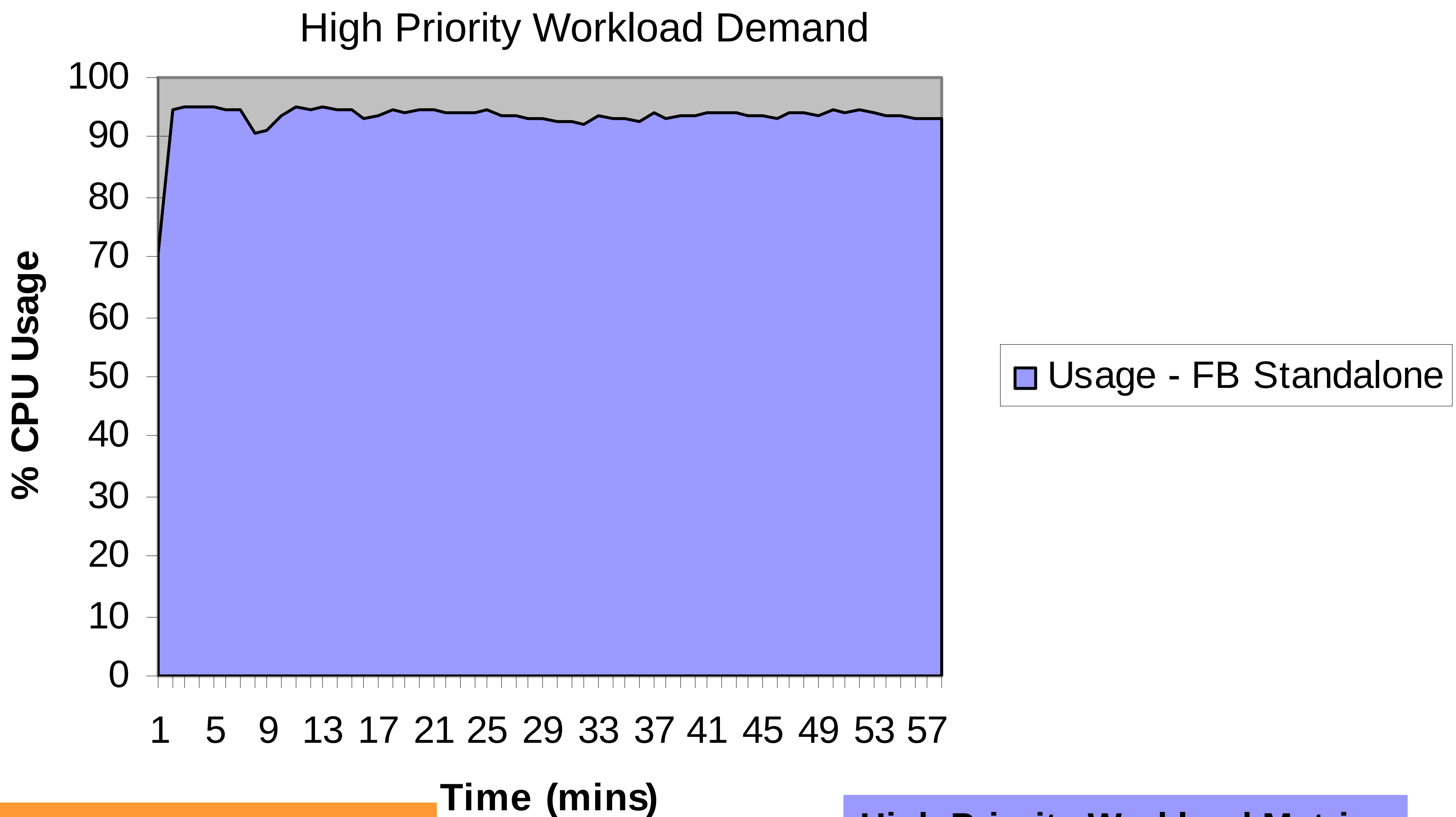
Run High Priority And Low Priority Workloads Together

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

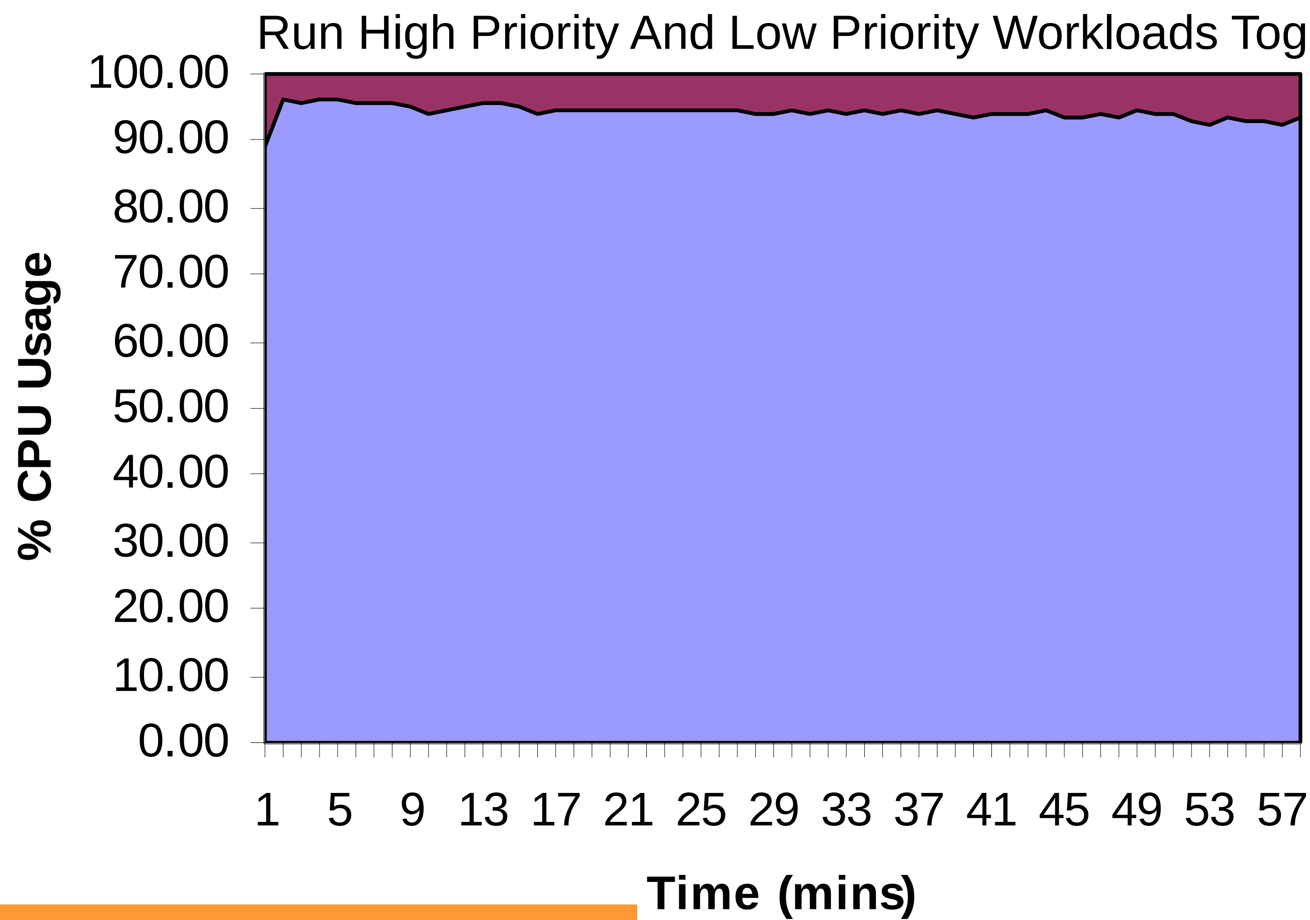
High Priority Web Workload with Constant Demand Running Standalone on System z



Capacity Used
High Priority – 93.4% CPU Minutes
Wasted – 6.6% CPU Minutes

High Priority Workload Metrics
Total Throughput: 11.95M
Avg Response Time: 149ms

High Priority Workload on System z Does Not Degrade When Low Priority Workload is Added



Low Priority
High Priority

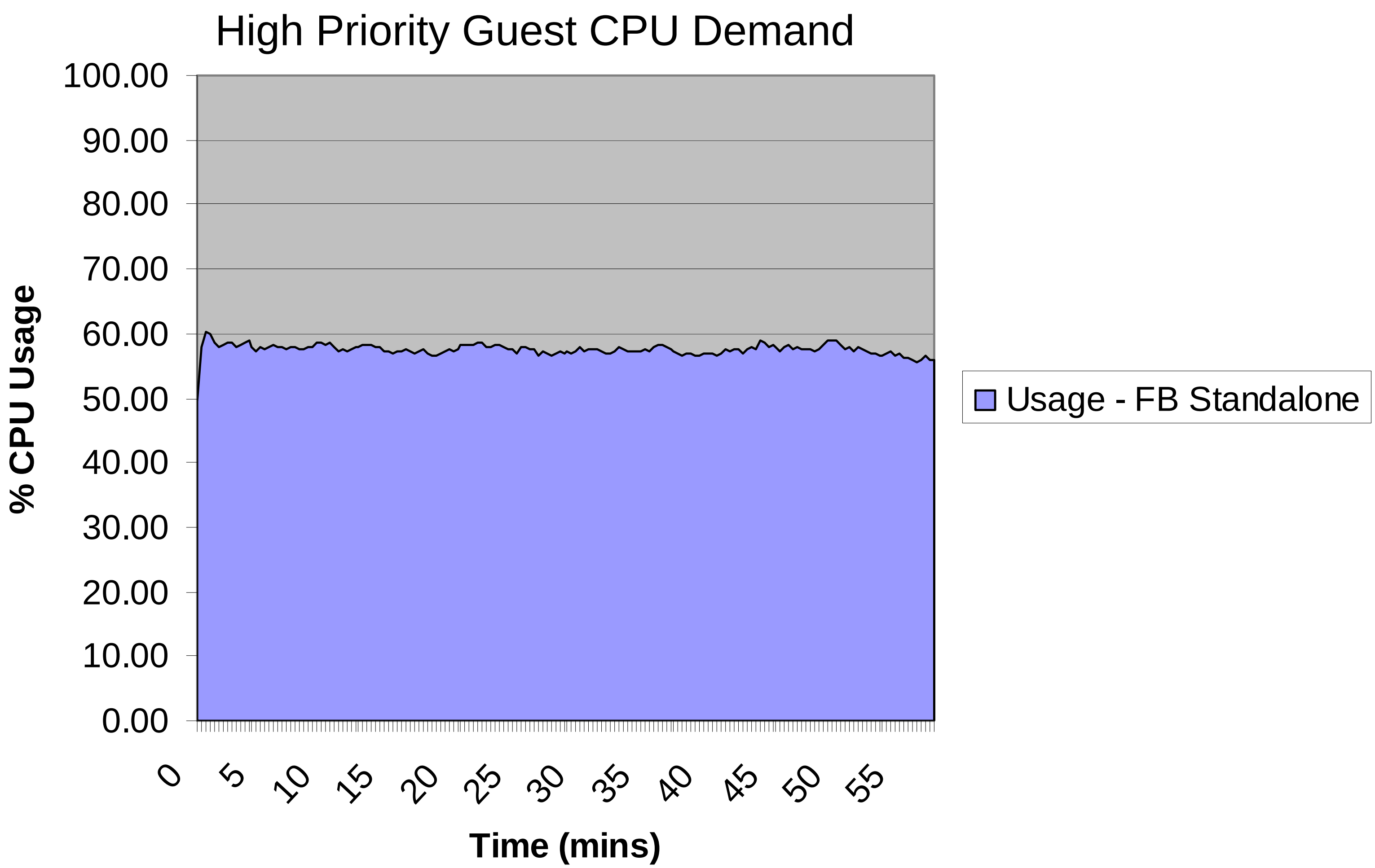
0.5% throughput decrease

0.7% response time increase

Capacity Used
High Priority – 94.3% CPU Minutes
Low Priority – 5.7% CPU Minutes
Wasted - 0% CPU Minutes

High Priority Workload Metrics
Total Throughput: 11.89M
Avg Response Time: 150ms

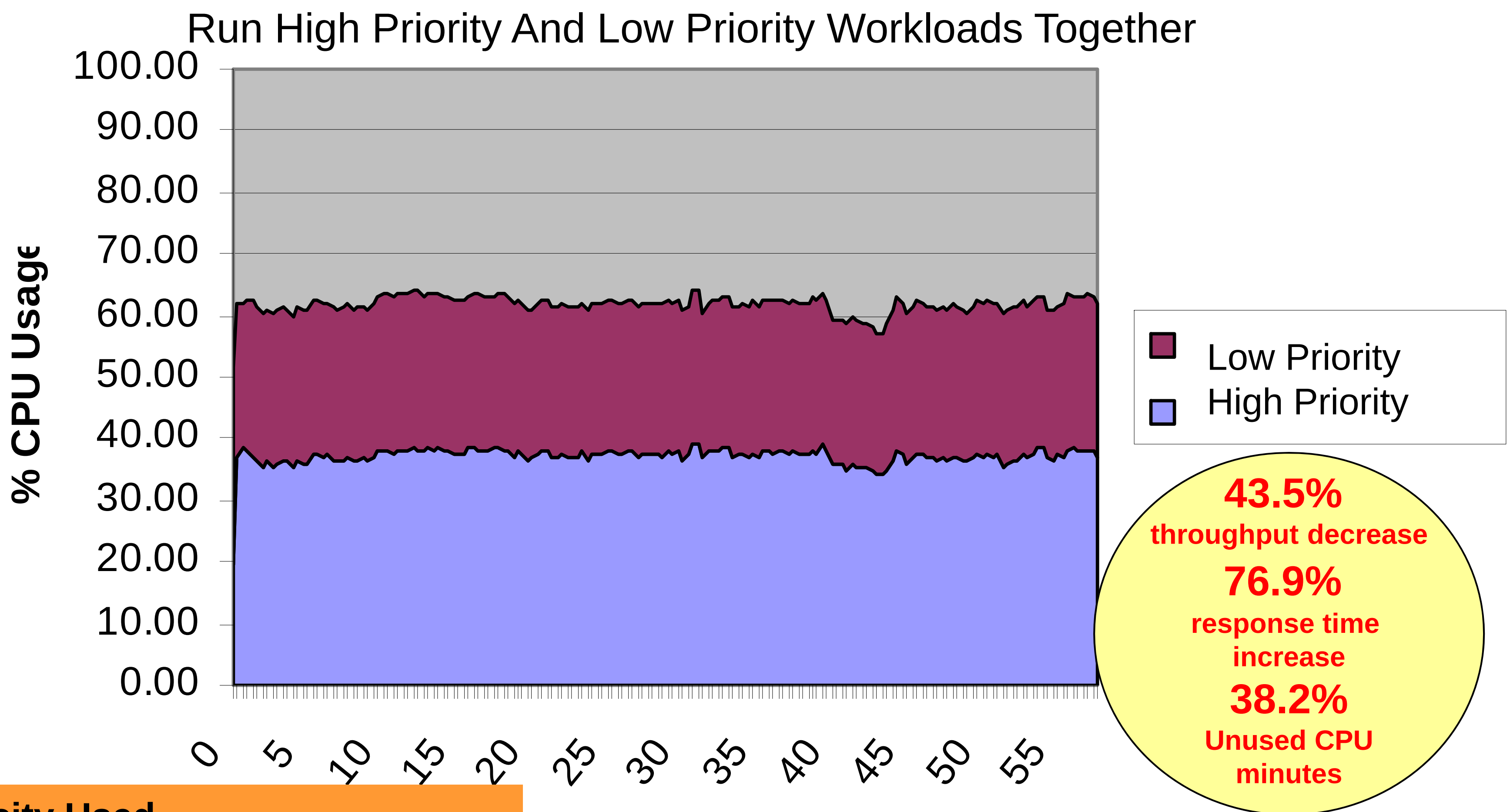
High Priority Web Workload with Constant Demand Running Standalone on x86/Common Hypervisor



Capacity Used
High Priority – 58% CPU Minutes
Wasted – 42% CPU Minutes

High Priority Workload Metrics
Total Throughput: 9.68M
Avg Response Time: 277ms

High Priority Workload on x86/Common Hypervisor Degrades Severely When Low Priority Workload is Added



Capacity Used
High Priority – 37.1% CPU Minutes
Low Priority – 24.7% CPU Minutes
Wasted – 38.2% CPU Minutes

High Priority Workload Metrics
Total Throughput: 5.47M
Avg Response Time: 490ms

Deliver High And Low Priority Workloads Together While Maintaining SLA

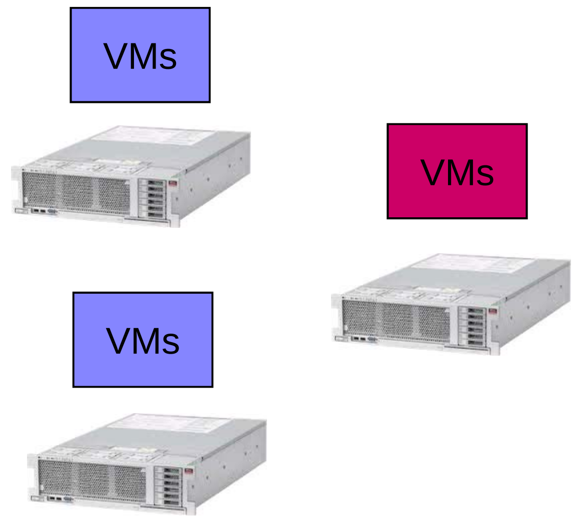
Comparison to determine which platform provides the lowest TCA over 3 years

High priority workloads

Low priority workloads

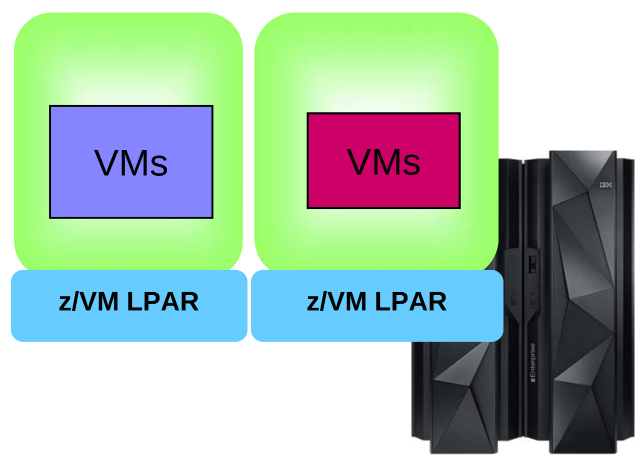
- **IBM WebSphere 8.5 ND**
- **IBM DB2 10 AESE**
- **Monitoring software**

High priority online banking workloads driving a total of **11.89M** transactions per hour and low priority discretionary workloads



Virtualized on 3 Intel 40 core servers

$13.66M (3 yr. TCA)



z/VM on zEC12

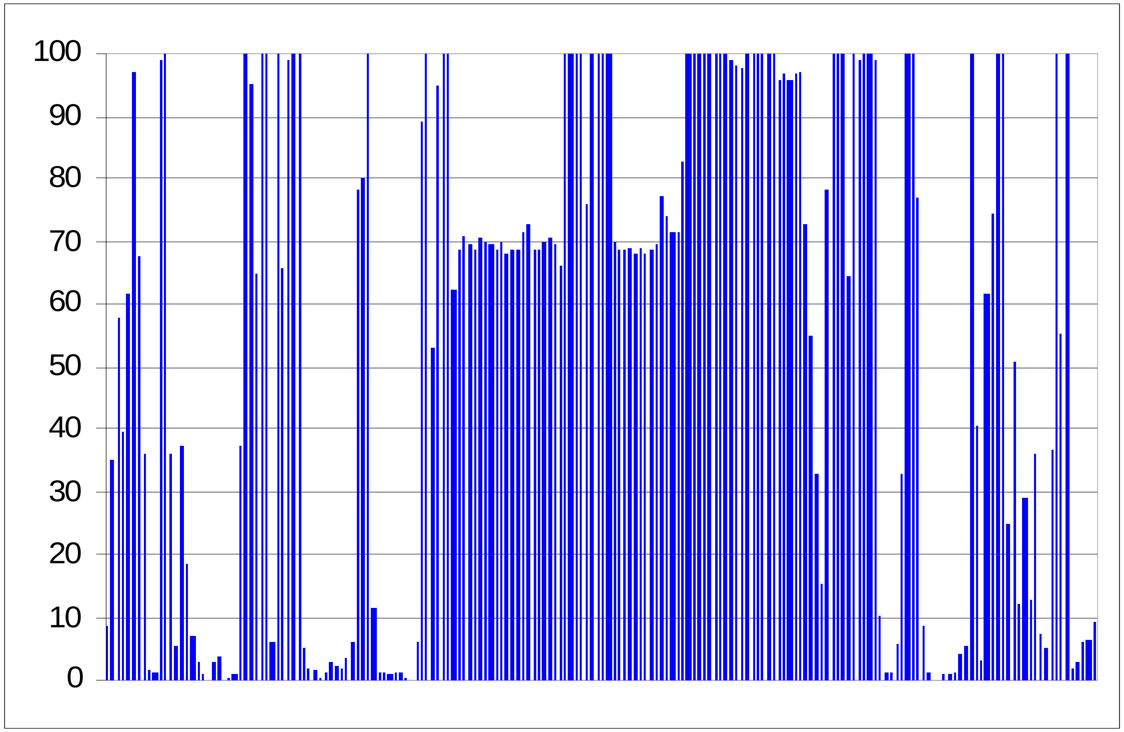
32 IFLs

$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.

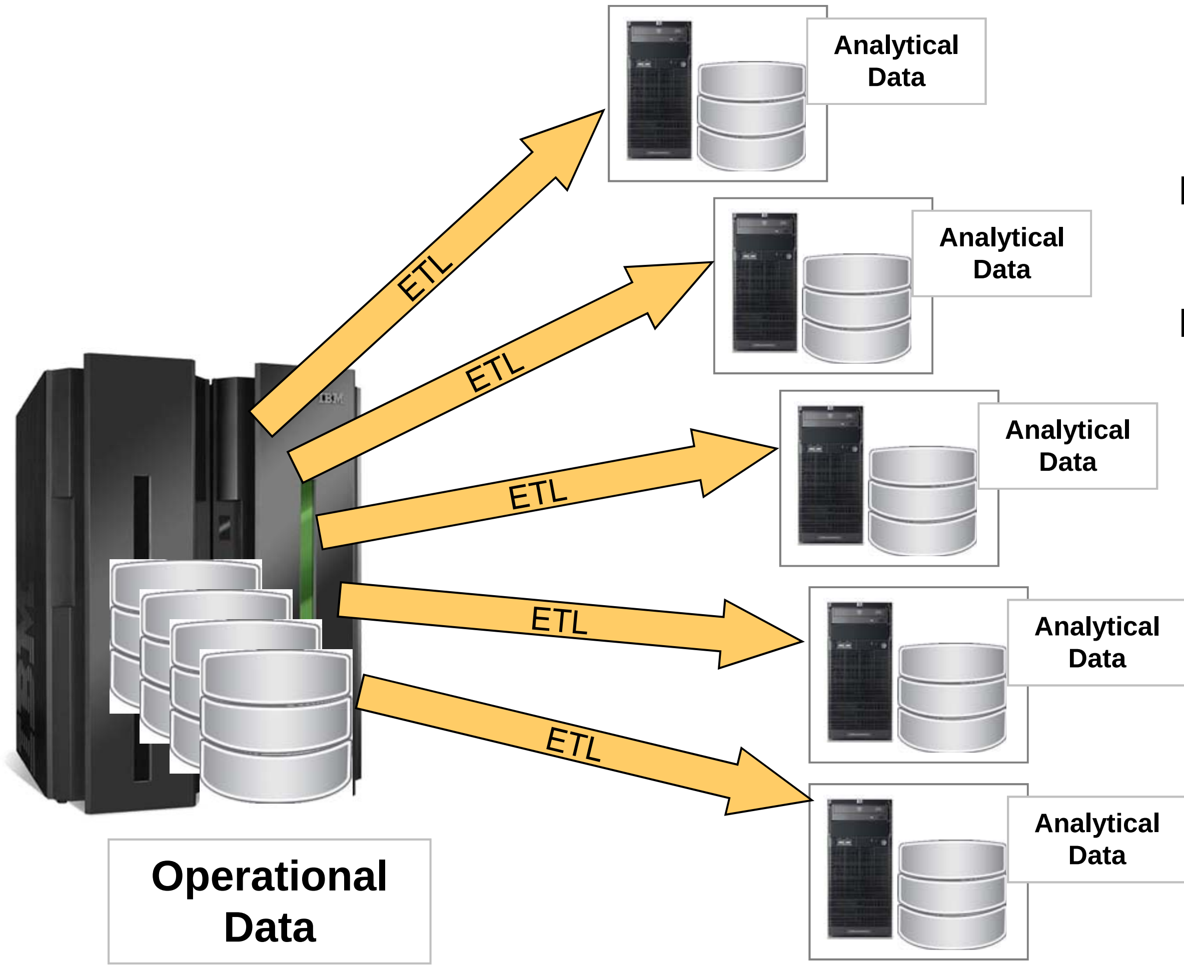
Sample LPAR - ETL Server - Bulk Data Movement



Classic ETL or Data Warehouse Pattern. Very High Utilization for multiple Hours. But also many Missing Data Points

* graph shows (only) 195 hourly data points.
00:00 mon 05 march to 24:00 sun 18 march.

What is "Mainframe Blockade"?

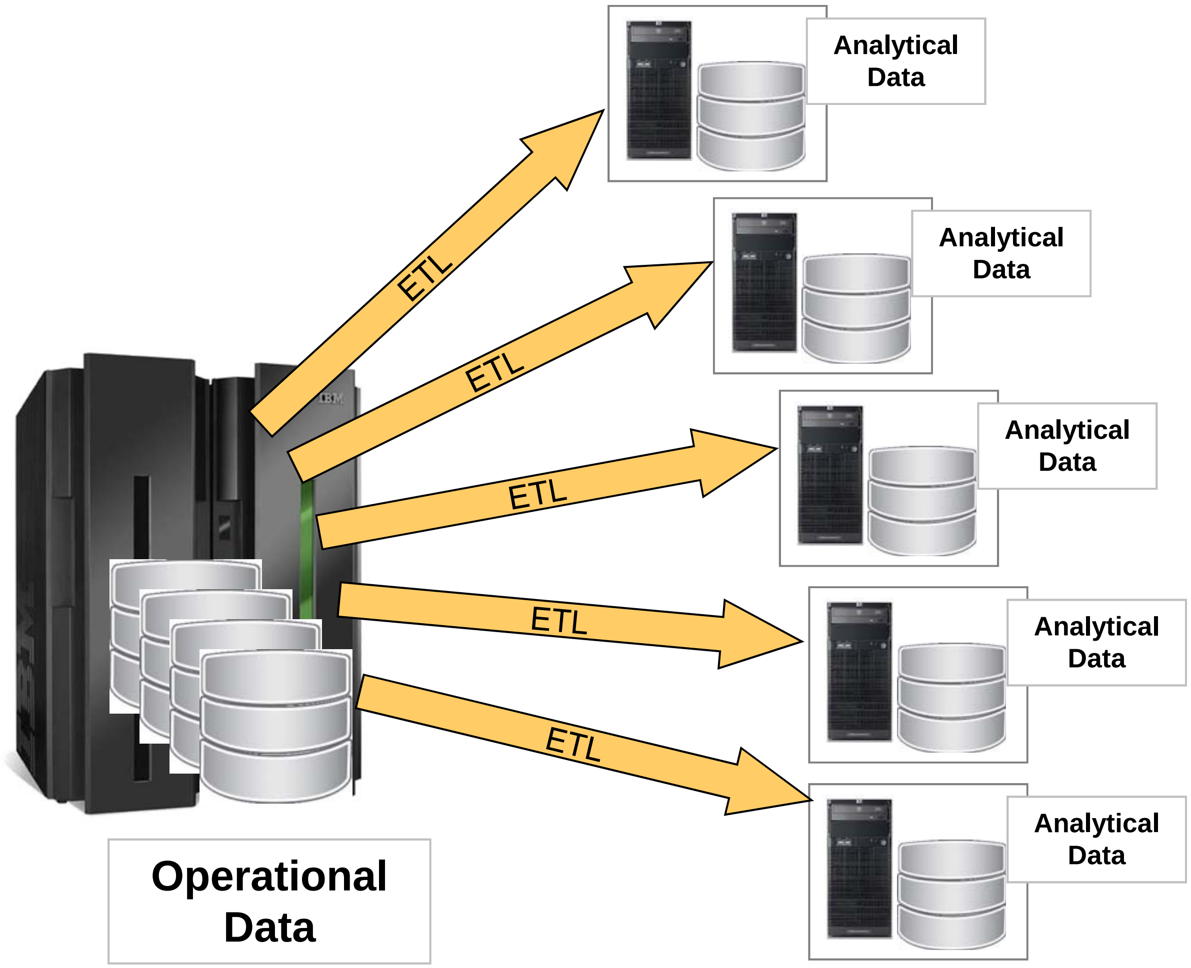


- Isolation of the mainframe
- Large proliferation of local solutions (applications and databases)

Businesses fall for common misperceptions:

- Distributed servers are cheap
- Offloading will reduce costs by reducing MIPS
- Cost of data transfers in insignificant

"Mainframe Blockade" Can Result in Significant Capacity Burn



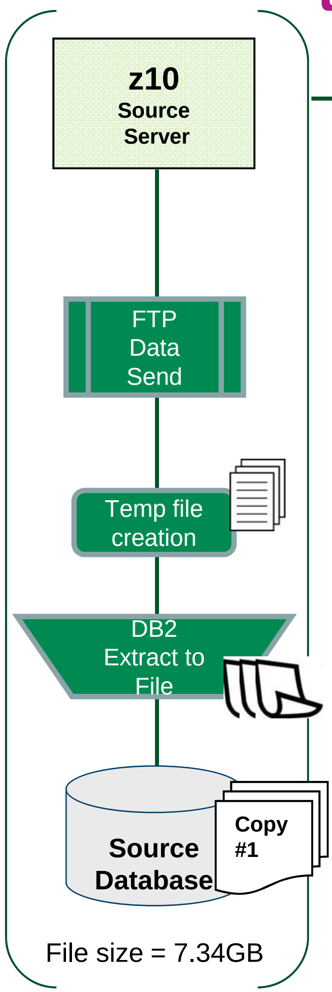
A large European bank:

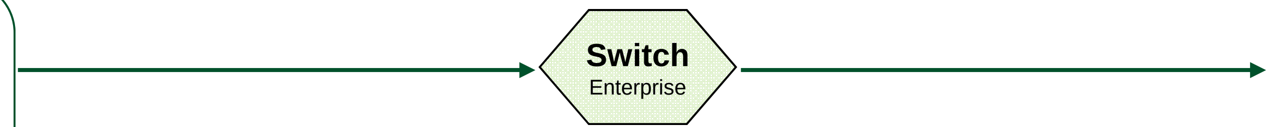
- 120 database images created from bulk data transfers
- 1,000 applications on 750 cores with 14,000 software titles
- ETL consuming 28% of total distributed cores and **16% of total MIPS**

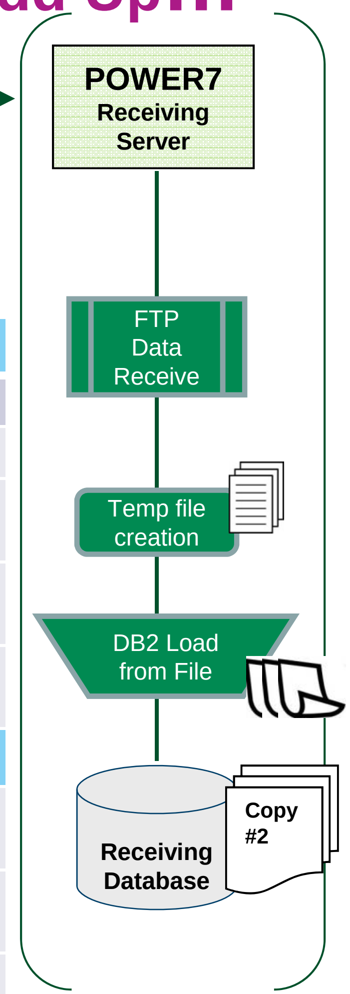
A large Asian bank:

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

Data Transfer is Mistakenly Perceived to be Insignificant, But Tests Show Times Add Up…







Switch	Send	Recv
Utilization	72.6%	72.6%

z10

FTP	Send	Recv
Transmit Time	153.69 sec	156.03 sec
Utilization	18%	16%
FTP Setup Labor*	56 sec	n/a
FTP Cmd Labor*	88 sec	n/a
FTP Confirm Labor*	12 sec	n/a
ETL	**Extract**	**Load**
ETL Execution Time	126 sec	n/a
ETL Utilization	52%	n/a
ETL Setup Labor**	300 sec	n/a
ETL Confirm Labor**	12 sec	n/a

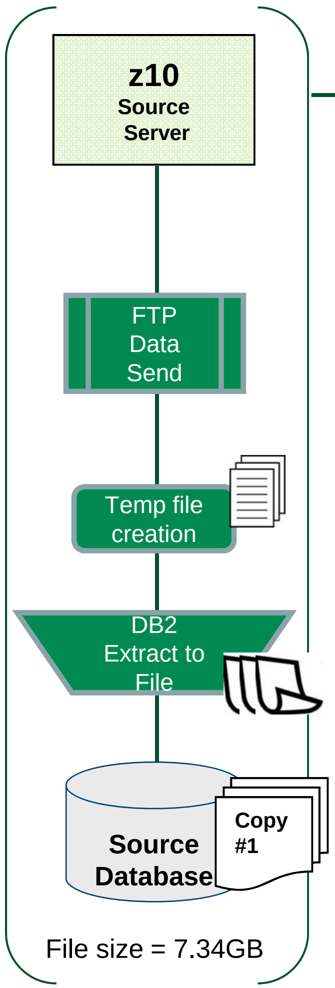
POWER7

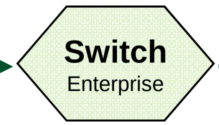
FTP	Send	Recv
Transmit Time	156.03 sec	153.69 sec
Utilization	20%	20%
FTP Setup Labor*	56 sec	n/a
FTP Cmd Labor*	88 sec	n/a
FTP Confirm Labor*	12 sec	n/a
ETL	**Extract**	**Load**
ETL Execution Time***	n/a	345 sec
ETL Utilization***	n/a	74%
ETL Setup Labor**	n/a	300 sec
ETL Confirm Labor**	n/a	12 sec

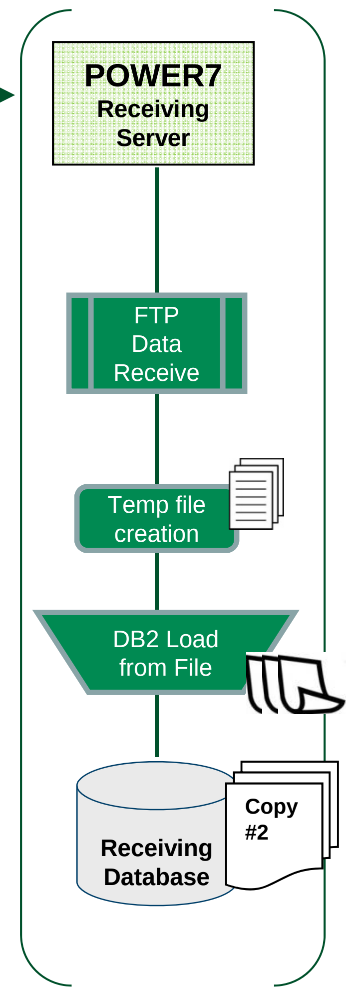
* Estimates based on measurements from previous FTP test
** Estimate based on work performed

*** Estimate based on Characteristics of ETL Workload on z and AIX study

And the Cost of Data Transfer Adds Up Also!



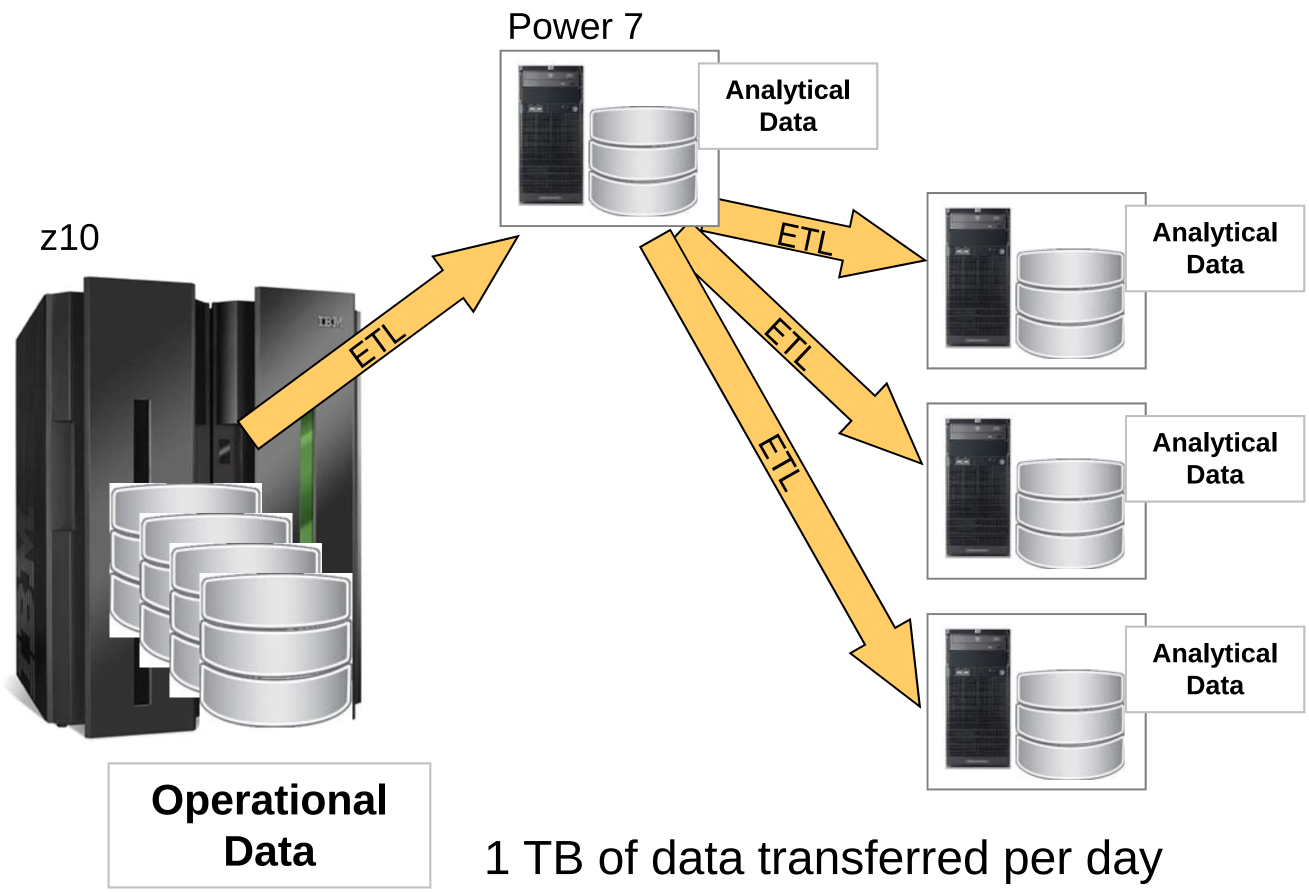




4 yr. amortized cost summary

System z Extract and Send	**$1.96/GB**
Distributed Receive and Load	**$0.437/GB**
Network	**$0.2924/GB**
System z Storage	**$49.33/GB**
Distributed Storage	**$59.68/GB**
System z Admin	**$15.21/job**
Distributed Admin	**$14.00/job**
System z Storage Admin	**$5.88/GB/yr**
Distributed Storage Admin	**$12.99/GB/yr**

Here is a Typical Situation…



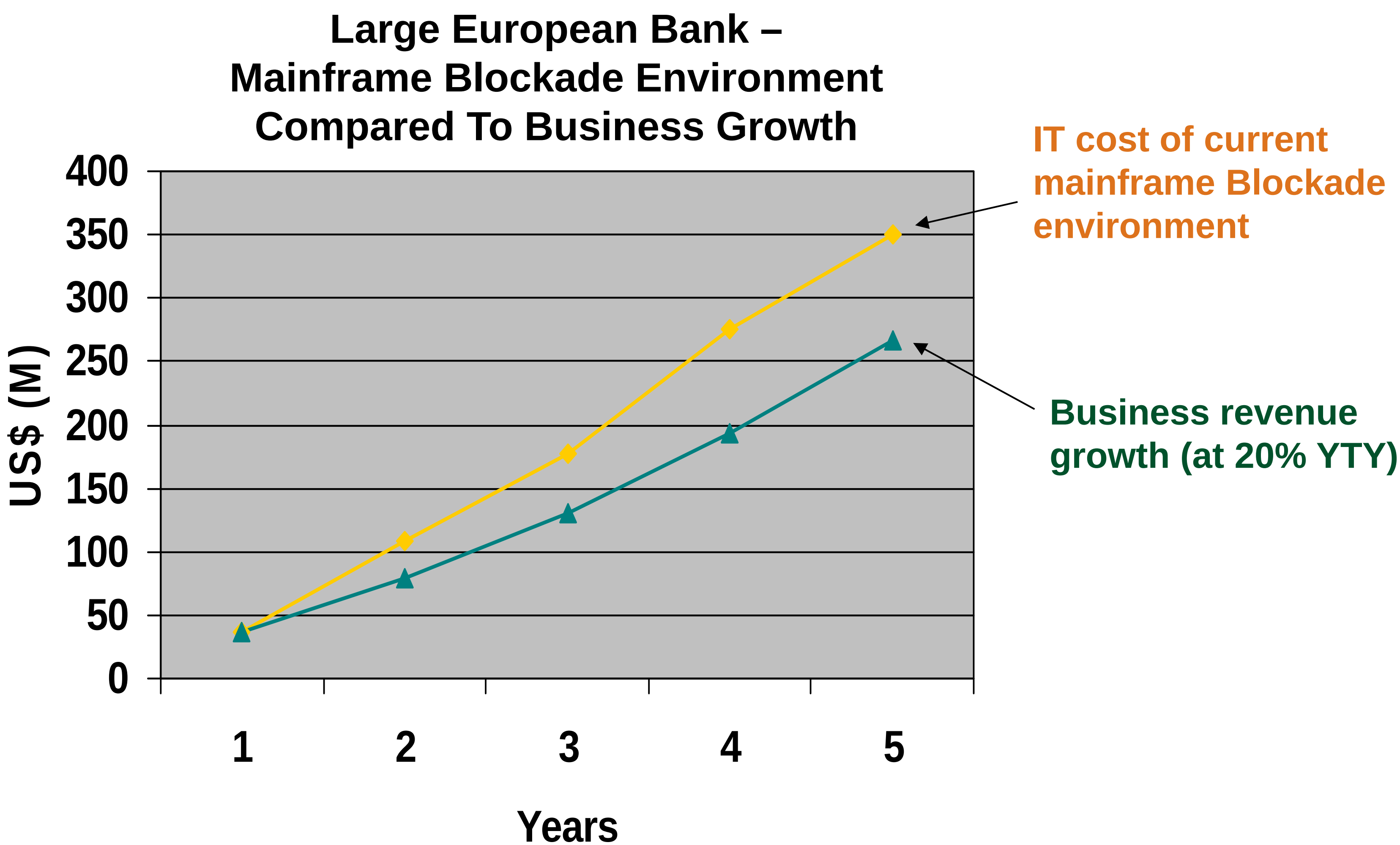
1 TB of data transferred per day – one initial copy, plus three derivative copies

4 yr. amortized cost summary

System z Extract and Send	$2,861,600
Distributed Receive and Load	$4,466,140
Network	$430,408
System z Storage	$49,330
Distributed Storage	$238,720
System z Admin	$22,207
Distributed Admin	$143,090
System z Storage Admin	$5,880
Distributed Storage Admin	$51,960

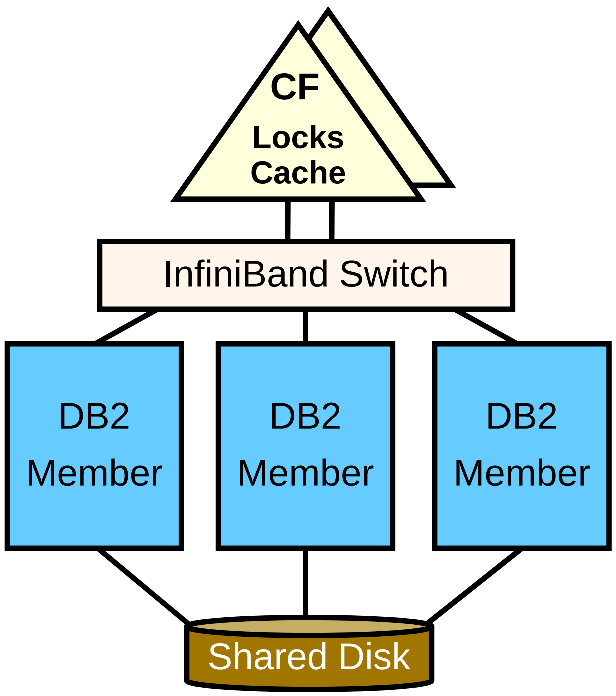
Source: CPO internal study. Assume dist. send and load is same cost as receive and load.. Also, assume 2 switches and 2 T3 WAN connections.

Businesses are Finding the Cost of a "Mainframe Blockade" Strategy is Not Sustainable!



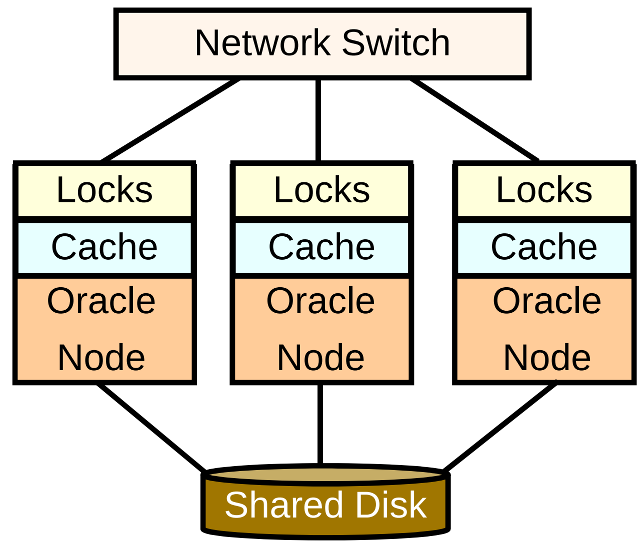
Clusters Grow Database Processing Power Beyond Single Server Solutions

DB2 for z/OS
Centralized Coupling Facility Design



Efficient lock and buffer management achieve near linear scalability

Oracle RAC
Distributed Design



Inefficient distributed locking and buffer management limits scaling

zEnterprise Is Optimized For Operational Analytics

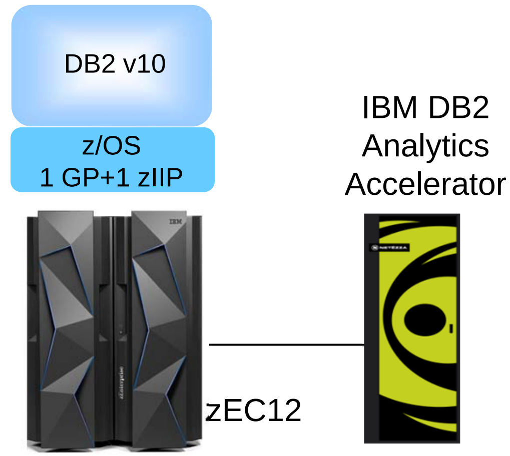
Standalone Pre-integrated Competitor

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



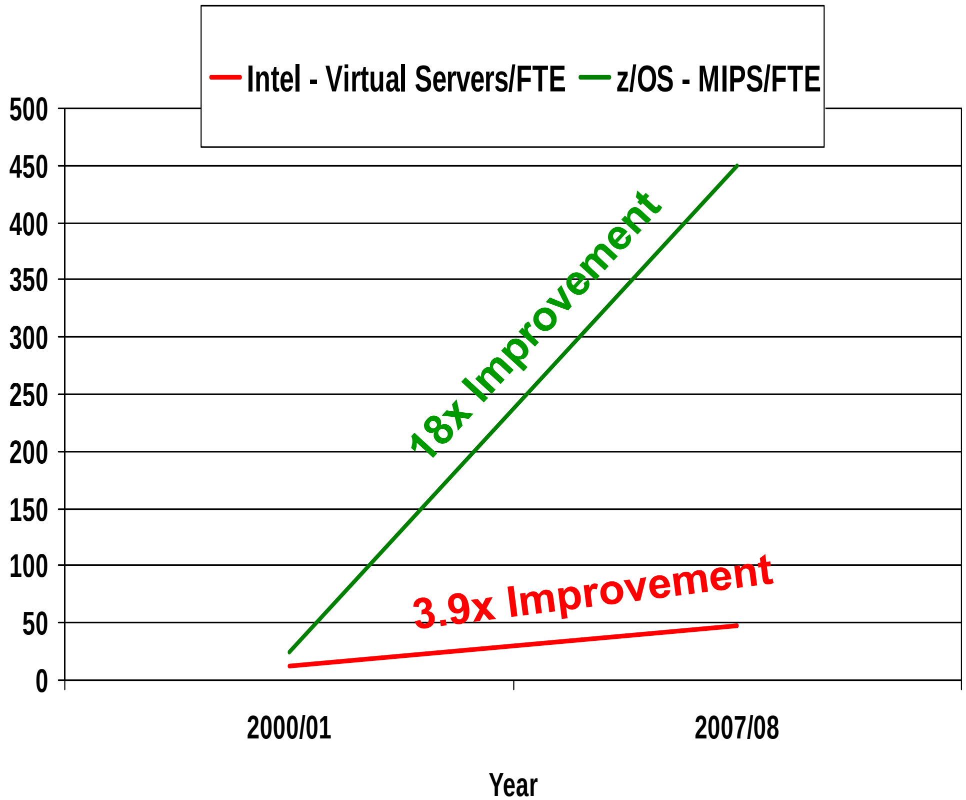
Unit Cost (3yr TCA) **$71/RpH**

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

10x performance at 1/10 the cost!

Source: Customer Study running 161,166 concurrent reports. Intermediate and complex reports automatically redirected to IBM DB2 Analytics Accelerator for z/OS. Results may vary based on customer workload profiles/characteristics. Note: Indicative ISAS 9700 pricing only internal to IBM, quotes to customer require a formal pricing request with configurations.

System z Labor Cost Trends Favor A Centralized Approach To Management



Large scale consolidation and structured management practices drive increases in labor productivity

Small scale consolidation achieves lesser gains

The more workloads you consolidate and manage with structured practices…

the lower the management labor cost

Source: IBM Scorpion Studies

Accumulated Field Data For Labor Costs

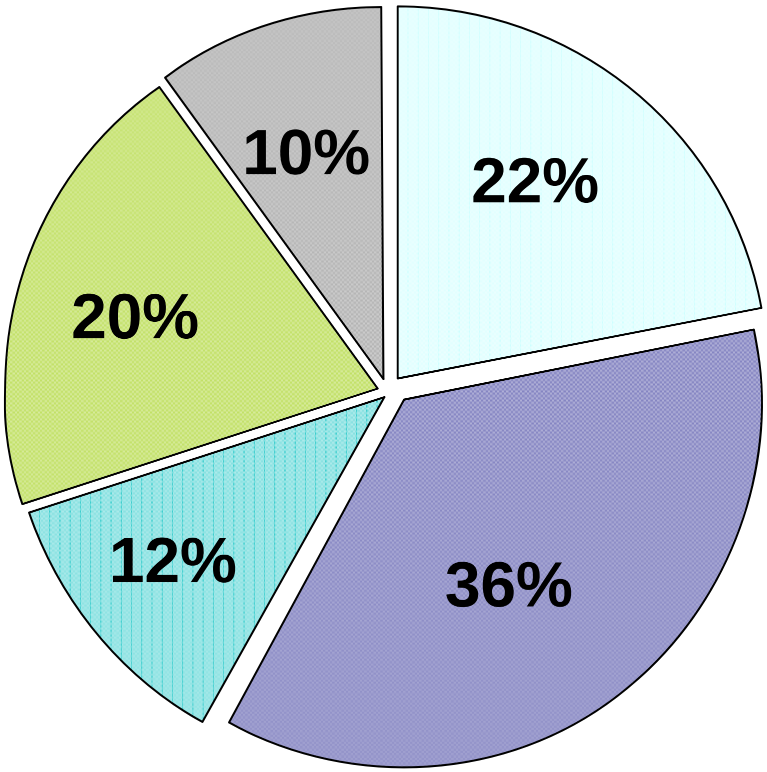
- **Average of quoted infrastructure labor costs**
  - **30.7** servers per FTE (dedicated Intel servers)
    - **67.8** hours per year per server for hardware and software tasks
  - **52.5** Virtual Machines per FTE (virtualized Intel servers)
    - **39.6** hours per year per Virtual Machine for software tasks and amortized hardware tasks
    - Typical 8 Virtual Machines per physical server

- **Best fit data indicates**
  - Hardware tasks are **32** hours per physical server per year
    - Assume this applies to Intel or Power servers
    - Internal IBM studies estimate 320 hours per IFL for zLinux scenarios
  - Software tasks are **36** hours per software image per year
    - Assume this applies to all distributed and zLinux software images

Labor model based on customer data from IBM studies

Five Key IT Processes For Infrastructure Administration

Time spent on each activity



- **Deployment Management**
  - *Hardware set-up and software deployment*
- **Incident/Capacity Management**
  - *Monitor and respond automatically*
- **Asset Management**
  - *Hardware and software asset tracking*
- **Security Management**
  - *Access control*
- **Change Management**
  - *Hardware and software changes*

Allocation based on customer data from IBM study

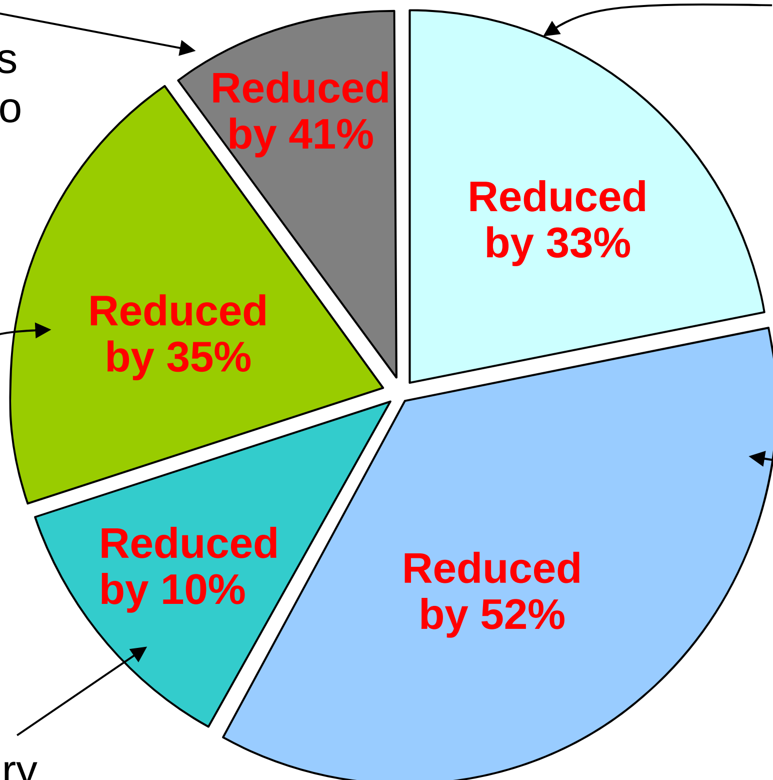
zManager Labor Cost Reduction Benefits Case Study

5032 total hours per year reduced by 38% to 3111 hours per year

Change Management
Standardization of images and firmware, visibility into relationships among resources

Security Management
Centralized fine-grain administrator access control

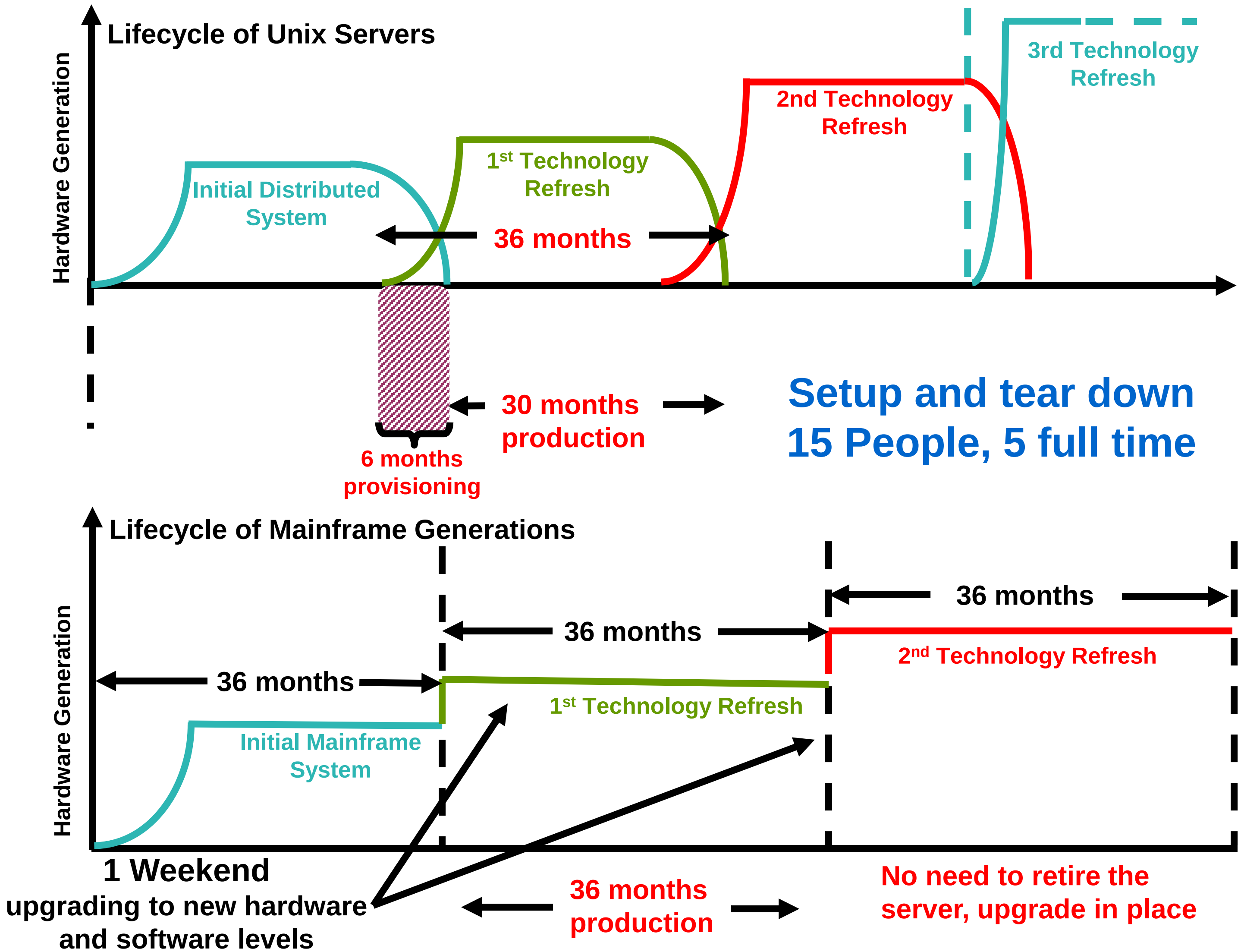
Asset Management
Automated discovery, entitlement management



Deployment Management
Automatic setup and configuration of the hypervisor and out-of-the-box networks

Incident/Capacity Management
Automation to isolate and fix issues

New York Financial Services Company – Useful Lifetime Of 36 Month Lease



Observed at a large financial service customer

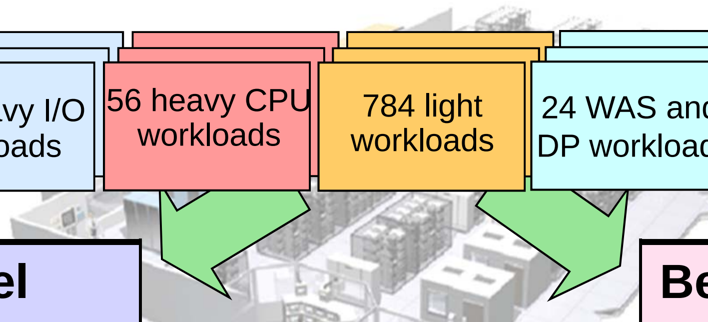
In each 36 month lease there are only 30 months production use

Setup and tear-down time costs 25% more. Plus . . . 41 hours of FTE setup and tear down labor per server = $3,075

Weekend upgrades performed by IBM

Capacity on demand pricing

Fewer Parts to Assemble and Manage



- 40 heavy I/O workloads
- 56 heavy CPU workloads
- 784 light workloads
- 24 WAS and DP workloads
- 20 SAP workloads

Deployed on Intel		Best fit on zEnterprise
183	Servers	1 z196 + 1 zBX (with 105 blades total)
1592	Network (parts)	21
124	Power (KW)	53
19	Administrators	13
70	Storage points	1

Understand The Cost Components

Updated Annual Operations Cost Per Small Server Image

Power, Floor Space	**$1,500**
Annual Hardware Maintenance (prepaid)	**$0**
Annual Connectivity Maintenance	**$240**
Annual Disk Maintenance	**$203**
Annual Software Support	**$10,153**
Annual Enterprise Network	**$1,024**
Annual Sysadmin	**$6,000**
Total Annual Costs	$19,120

Source: IBM Eagle Studies

Save Approx. $10K By Consolidating To z/VM

Updated Annual Operations Cost Per Small Server Image

Power, Floor Space	**$38**
Annual Hardware Maintenance	**$1,500**
Annual Connectivity Maintenance	**$4**
Annual Disk Maintenance	**$203**
Annual Software Support	**$3,626**
Annual Enterprise Network	**$1,024**
Annual Sysadmin	**$3,000**
Total Annual Costs	$9,395

Source: IBM Eagle Studies, IBM ECM project

40

Realize Significant Cost Reductions With Consolidation On Linux For System z

Oracle Consolidations on Linux for System z

Distributed cores to IFLs

Major Transportation Company:
Software costs reduced by 84%, TCO reduced by 50%

46 : 1

Middle East Bank:
Software costs reduced by 76%, TCO reduced by 64%

50 : 1

IBM's 'Big Green' Consolidation Project

Distributed servers to mainframes

Distributed servers running variety of workloads consolidated onto Linux for System z
Average across-the-board reduction in TCO of 70%

130 : 1

Planned ratio for continued consolidation to z196s

200 : 1

Projected ratio for continued consolidation to zNext

290 : 1

Summary

- **Cost per workload is the key metric for the new IT economics**
  - Mainframe cost per work goes down as workload increases

- **Fit for purpose reduces cost of acquisition per workload**

- **zEnterprise's integrated management reduces cost per workload with extreme automation for simplicity**

Thank you

The Savings Are Cumulative

40 heavy I/O workloads | 56 heavy CPU workloads | 784 light workloads | 24 WAS and DP workloads | 20 SAP workloads

Three Year Cost Of	Deployed on Intel	Best fit on zEnterprise
Servers	$46.0M	$26.1M
Network	$0.45M	$0.03M
Power	$0.33M	$0.14M
Labor	$9.02M	$6.09M
Storage	$8.58M	$4.6M
Total	**$64.38M**	**$36.96M**
Total cost per workload	**$70K**	**$40K**

43% less

Results may vary based on customer workload profiles/characteristics. Prices are in US currency. Prices will vary by country

Cost Ratios in all TCO Studies

Average Cost Ratios (z vs Distributed)

		z	Distributed	z vs distributed (%)
Offload	**5-Year TCO**	**$16,351,122**	**$31,916,262**	**51.23%**
	Annual Operating Cost	$2,998,951	$4,405,510	68.07%
	Software	$10,932,610	$16,694,413	65.49%
	Hardware	$3,124,013	$3,732,322	83.70%
	System Support Labor	$3,257,810	$4,429,166	73.55%
	Electricity	$45,435	$206,930	21.96%
	Space	$59,199	$154,065	38.42%
	Migration	$438,082	$10,690,382	4.10%
	DR	$854,266	$2,683,652	31.83%
	Average MIPS	3,954		
	Total MIPS	217,452		
Consolidation	**5-Year TCO**	**$5,896,809**	**$10,371,020**	**56.86%**
	Annual Operating Cost	$716,184	$1,646,252	43.50%
	Software	$2,240,067	$6,689,261	33.49%
	Hardware	$2,150,371	$1,052,925	204.23%
	System Support Labor	$1,766,403	$2,395,693	73.73%
	Electricity	$129,249	$365,793	35.33%
	Space	$84,033	$205,860	40.82%
	Migration	$678,449	$0	
	DR	$354,735	$411,408	86.22%
	Average MIPS	10,821		
	Total MIPS	292,165		