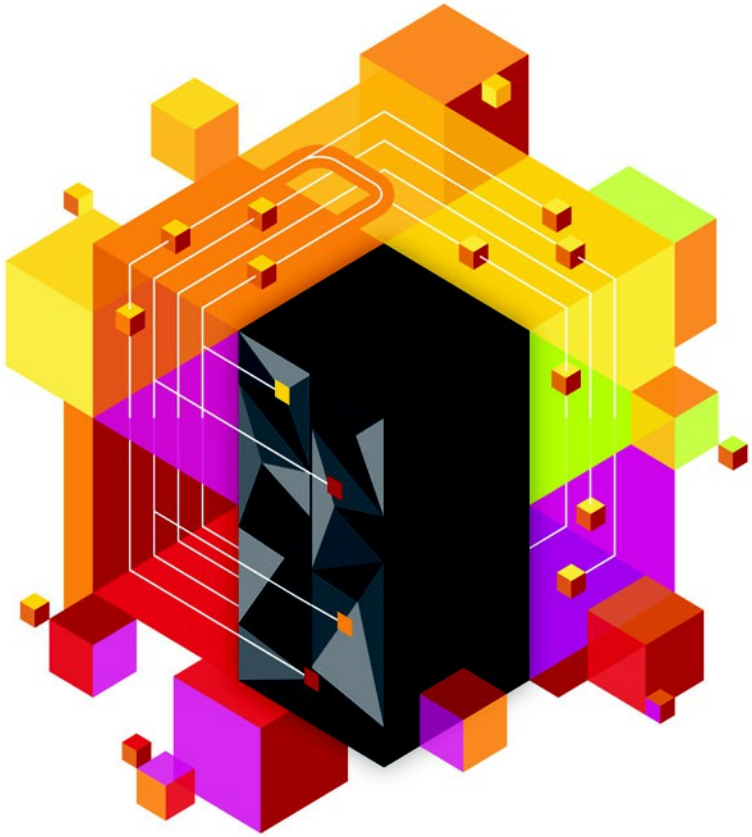
Smarter Computing.

System z Delivers Best TCO – Customer Examples

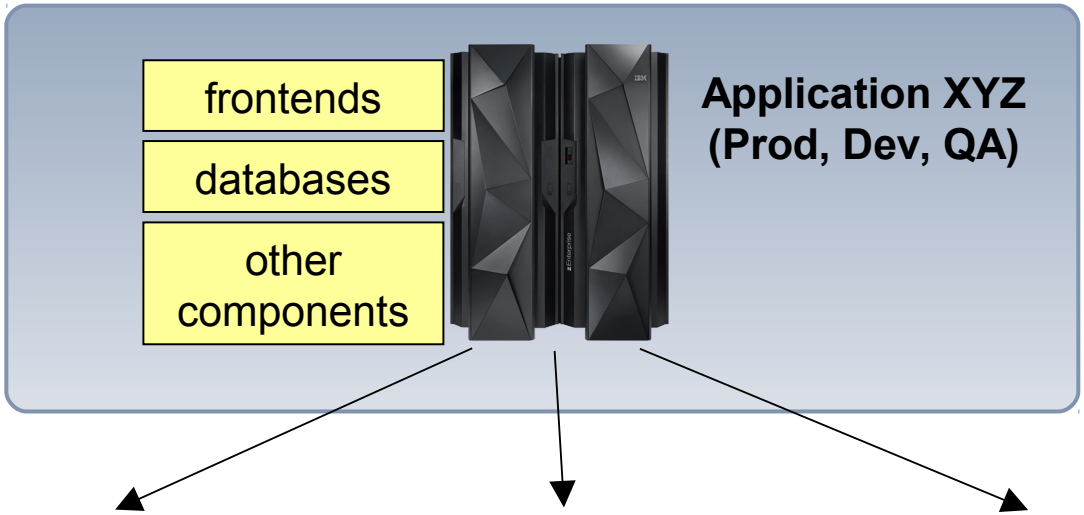


John J Thomas

IBM Competitive Project Office

What Happens In a TCO Study?

Workload identified for analysis



Deployment Choices

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

Key steps in analysis

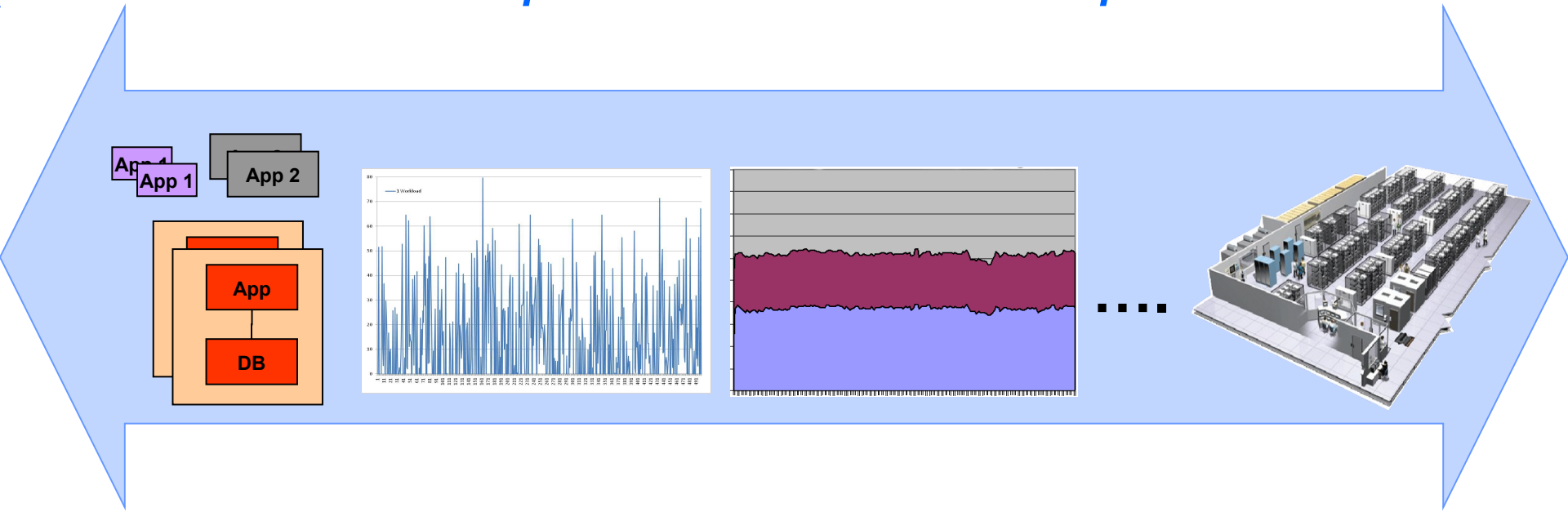
1. **Establish equivalent configurations**
   - Needed to deliver workload
2. **Compare Total Cost of Ownership**
   - TCO looks at different dimensions of cost

Approaches To Establishing Equivalent Configurations

- Bottom up approach
  - Atomic benchmarks
  - Counting cycles, CPI comparisons …
  - IO, memory, cache, co-location effects …
  - Tends to show smaller core expansion factors
- Top down approach
  - “Real world” observations
  - Tends to show much larger core expansion factors
- When atomic benchmarks are assembled to represent “real world”, bottom up numbers approach top down numbers

How Can We Determine Equivalent Configurations?

Real world aspects determine accurate equivalence



Platform factors

GHz, CPI, IO, co-location etc

Variability in demand

Different size servers

Workload Management

Mix workloads with different priorities

....

Top Down approach

What we see in customer environments

Core Proliferation For A Very Large Workload

Configurations for equivalent throughput (10,716 Transactions Per Second)

16x 32-way HP Superdome
App. Production / Dev / Test

8x 48-way HP Superdome
DB Production / Dev /Test

41 GP processors
(38,270 MIPS)

896 processors
(3,668,600 Perf Units)

***22x** more cores!*

zEC12 41-way Production / Dev / Test

Core Proliferation For A Mid-sized Workload

6x 8-way HP DL Production / Dev
2x 64-way p595 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)

***29x** more cores!*

482 Performance Units per MIPS

Core Proliferation For A Small Offload Project

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)

41x more cores

Almost 5 Year Migration

670 Performance Units per MIPS

1 CICS region in production!!
CICS/IDMS migrated to CICS/DB2.
Accessing DB2 thru mapping layer

No Disaster Recovery

Recent (April ’13) x86 Offload

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

2.1 processors
(499 MIPS)

z800 Production /
Dev / Test
(2002 mainframe technology)

60 Linux processors
(383,022 Perf Units)

29x more cores

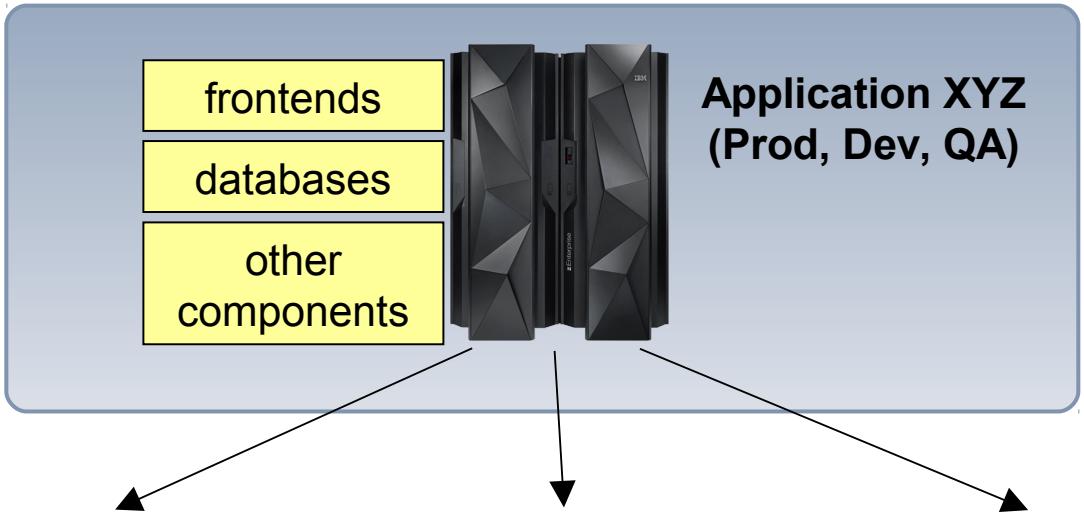
(despite the 9 year technology gap!)

1.5 Year Migration

768 Performance Units per MIPS

What Happens In a TCO Study?

Workload identified for analysis



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

To Understand Total Cost of Ownership
Four Dimensions Of Cost Should Be Considered

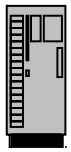
- Cost components
- Environments
- Time Factors
- Non-Functional Requirements / Qualities of Service

Many Cost Components

80:20 rule helps to achieve reasonable results in a short time

Components

Hardware



List vs Discounted
Fully configured vs. basic, Prod. vs. DR
Refresh / upgrade, Solution Edition…

Software

IBM and ISV, OTC and Annual maint (S&S)
MLC, PVU, RVU, ELA, core, system

People

FTE rate, in house vs. contract

Network

Adapters, switches, routers, hubs
Charges, Allocated or apportioned, understood or clueless

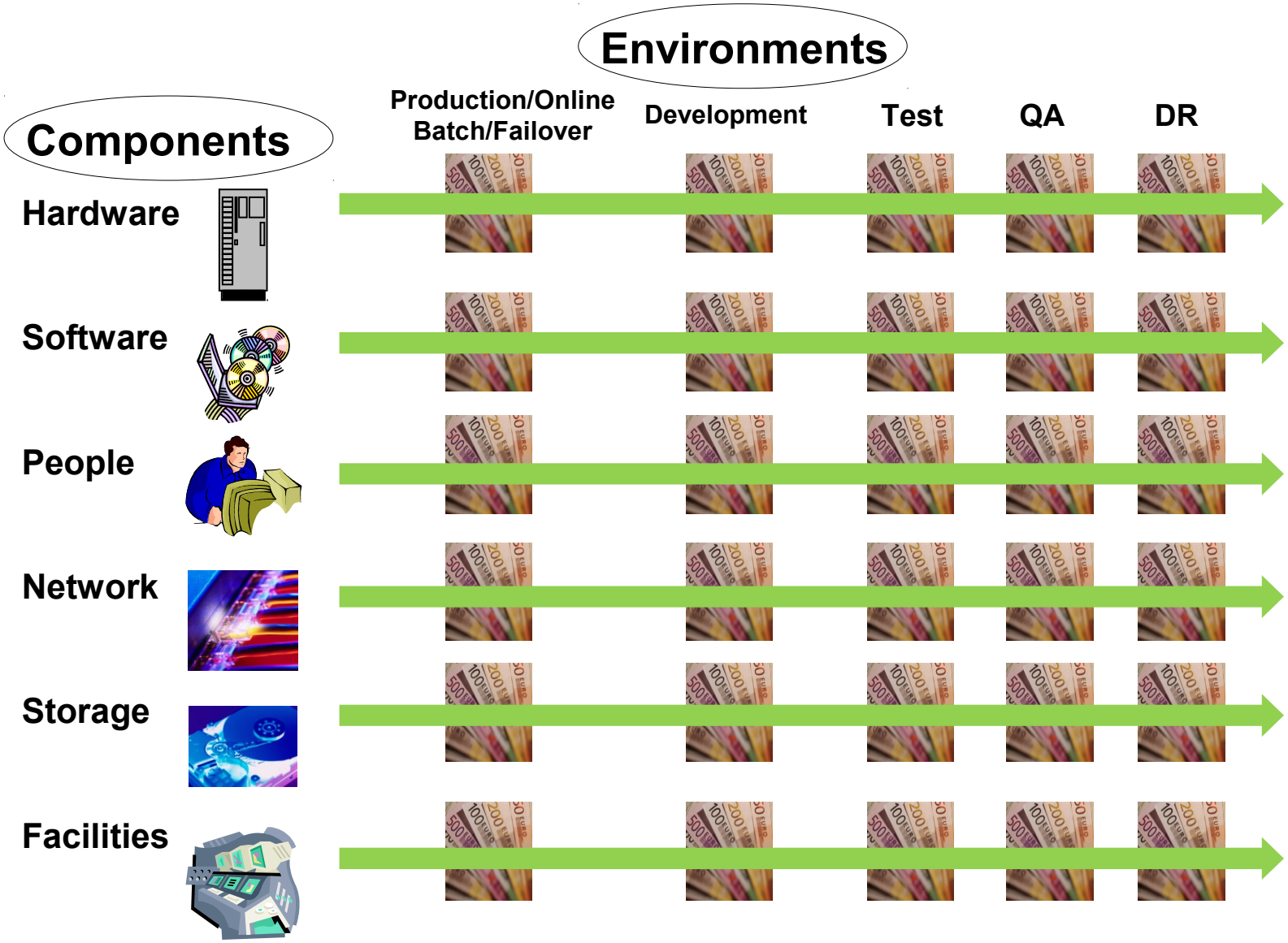
Storage

ECKD, FBA, SAN, Compressed, Primary, secondary
Disk (multiple vendors), tape, Virtual, SSD

Facilities

Space, electricity, air cooling, infrastructure including UPS and generators, alternate site(s), bandwidth

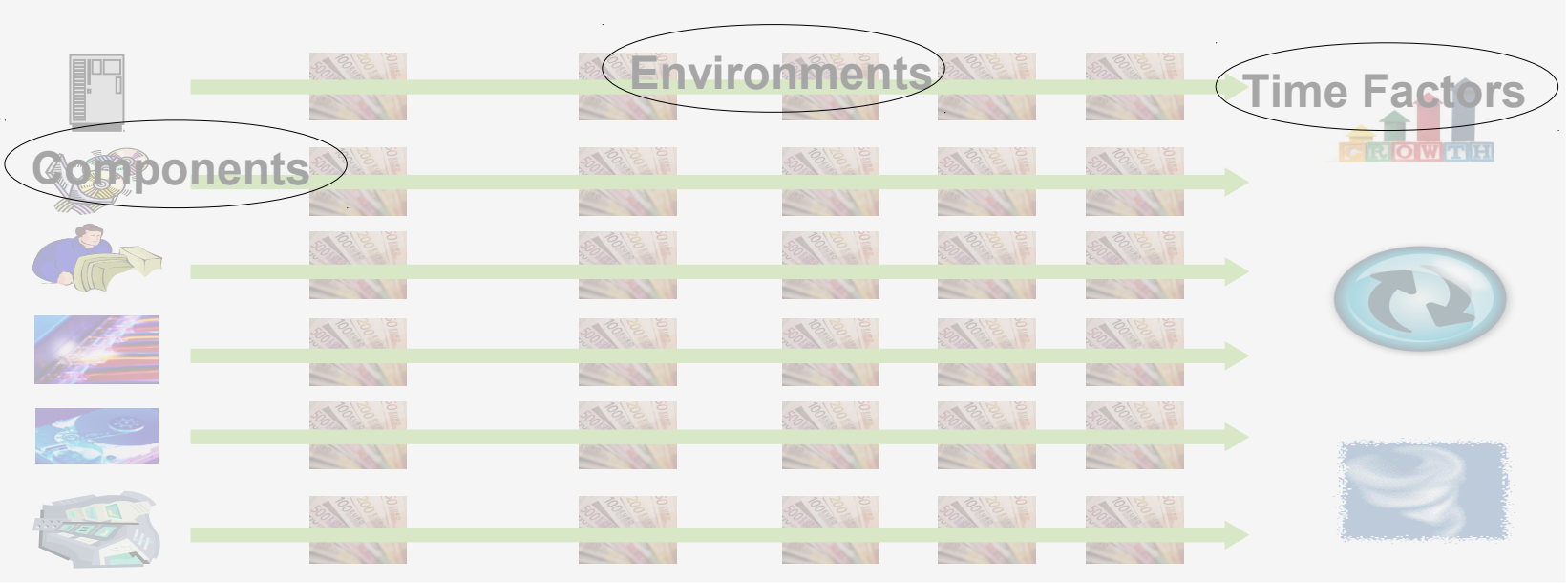
Environments Multiply Components



Time Factors Drive Growth And Cost

- Migration time and effort
- Business organic growth and/or planned business changes affect capacity requirements
  - e.g. Change of access channel or adding a new internet accessible feature can double or triple a components workload
  - Link a business metric (e.g. active customer accounts) to workload (e.g. daily transactions) and then use business inputs to drive the TCO case
- Other periodic changes – hardware refresh or software remediation

Non-Functional Requirements Can Drive Additional Resource Requirements



Availability …Security …Resiliency …Scalability …

Qualities of Service, Non-Functional Requirements

TCO: Understand The Complete Picture

TCA (Total Cost Of Acquisition)

- Cost of acquisition only
- Usually for a period of 3 years
  - But without technology refresh, business growth etc.

- Hardware acquisition and maintenance
- Software acquisition and S&S
  - OS, hypervisor, middleware

- TCA is a subset of TCO
- Represents immediate reality

So What Were The Total Costs In The Core Proliferation Cases We Saw Earlier?

Case	RPE/MIPS	Z Total Cost	Distributed Total Cost	Factor
Large benchmark	95	**$111M** (5 yr. TCA)	**$180M** (5 yr. TCA)	1.62x
Mid size offload	482	**$17.9M** (5 yr. TCO)	**$25.4M** (5 yr. TCO)	1.42x
Small offload	670	**$4.9M** (4 yr. TCO)	**$17.9M** (4 yr. TCO)	3.65x
Even smaller offload	499	**$4.7M** (5 yr. TCO)	**$8.1M** (5 yr. TCO)	1.72x

Lessons Learned Can Be Grouped Into Three Broad Categories

- Always compare to an optimum System z environment
- Look for not-so-obvious distributed platform costs to avoid
- Consider additional platform differences that affect cost

(1) Always Compare To An Optimum System z Environment

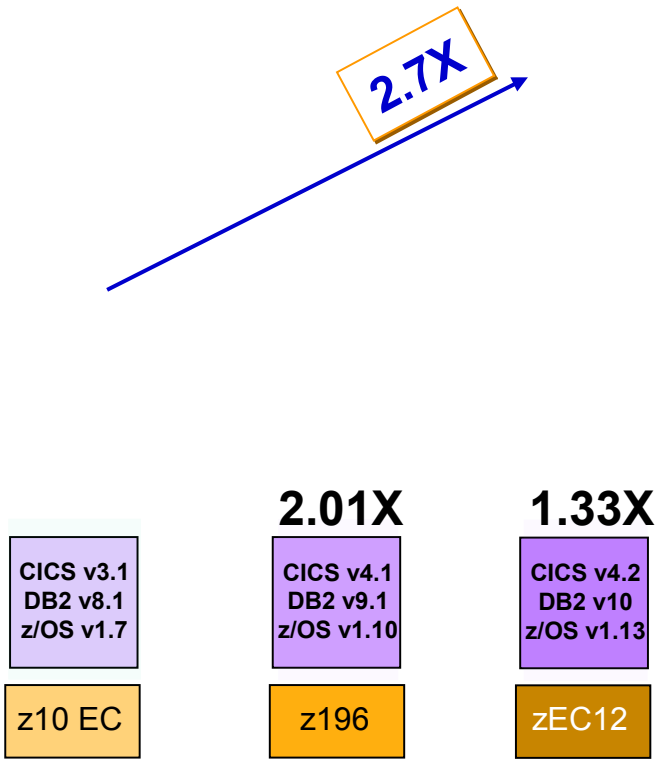
- Updating hardware and software reduces cost
- Sub-capacity may produce free workloads
- Replace ISV software with IBM software
- System z Linux consolidation saves money
- Changing database can impact capacity requirements
- Specialty processors reduce mainframe cost
- Use accelerators when appropriate!

Currency Reduces Cost – Hardware

2 generations,
from z9 to z196

- Typical customer (European bank) hardware refresh scenario
  - 2M investment pays back >1M savings every year – most cases positive in a 3 year period
  - Savings from VWLC->AWLC and specialty processor upgrades
- Comparing latest technology servers to old mainframes is unfair but often done

Performance Improvements Can Lower MLC Costs And Free Up Hardware Capacity



Customer examples:

(1) Large MEA bank

- Delayed upgrade from z/OS 1.6 because of cost concerns
- When finally did upgrade to z/OS 1.8
  - Reduced each LPAR's MIPS by 5%
  - Monthly software cost savings paid for the upgrade almost immediately

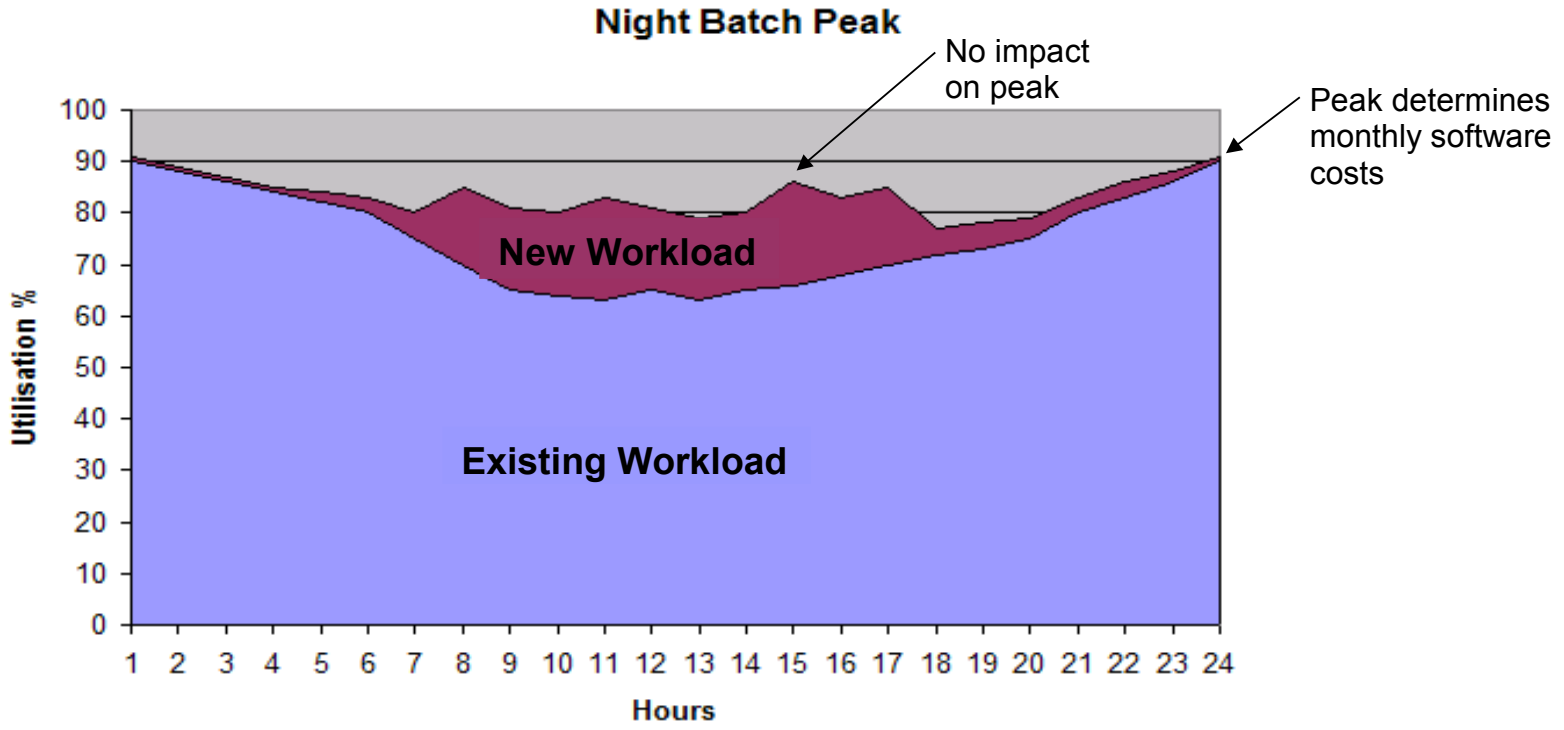
(2) Large European Auto company

- Upgraded to DB2 10
- Realized 38% pathlength reduction for their heavy insert workload
  - Other DB2 10 users saw 5-10% CPU reduction for traditional workloads

Additionally, save costs by moving to newer compilers and tuning

IBM internal core banking workload (Friendly Bank). Results may vary.

Sub-Capacity May Produce Free Workloads



- Standard “overnight batch peak” profile – drives monthly software costs
- Hardware and software are free for new workloads using the same middleware (e.g. DB2, CICS, IMS, WAS, etc.)
- Ensure you exploit any free workload opportunities, and conversely, avoid offloading free applications!

Linux On System z Consolidation Usually Has Lower Costs

Which platform provides the lowest TCA over 3 years?

Oracle DB workload

3 OLTP Database Workloads, each supporting 18K tps

Oracle Enterprise Edition

Oracle Real Application Cluster

3 Oracle RAC clusters
4 server nodes per cluster

12 total HP DL580 servers (192 cores)

$13.2M (3 yr. TCA)

3 Oracle RAC clusters
4 nodes per cluster
Each node is a Linux guest
zEC12 with 27 IFLs

$5.7M (3 yr. TCA)

***Half** the cost*

TCA includes hardware, software, maintenance, support and subscription.
Workload Equivalence derived from a proof-of-concept study conducted at a large Cooperative Bank.

Leverage Accelerators Where Relevant

Standalone Pre-integrated Competitor V3

Quarter Unit

IBM zEnterprise Analytics System 9700

Unit Cost
$51/Reports per Hour

Workload Time	141 mins
Reports per Hour	68,581
Total Cost (3 yr. TCA) (HW+SW+Storage)	$3,530,041

Unit Cost
$17/Reports per Hour

Workload Time	25 mins
Reports per Hour	386,798
Total Cost (3 yr. TCA) (13 GP + 12 zIIP, HW+SW+ Storage + Accelerator V3.1 with PDA N2001-10 hardware)	$6,464,849

***3x** price performance!*

Source: Customer Study on 1TB BIDAY data 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 9700 pricing only internal to IBM, quotes to customer require a formal pricing request with configurations.

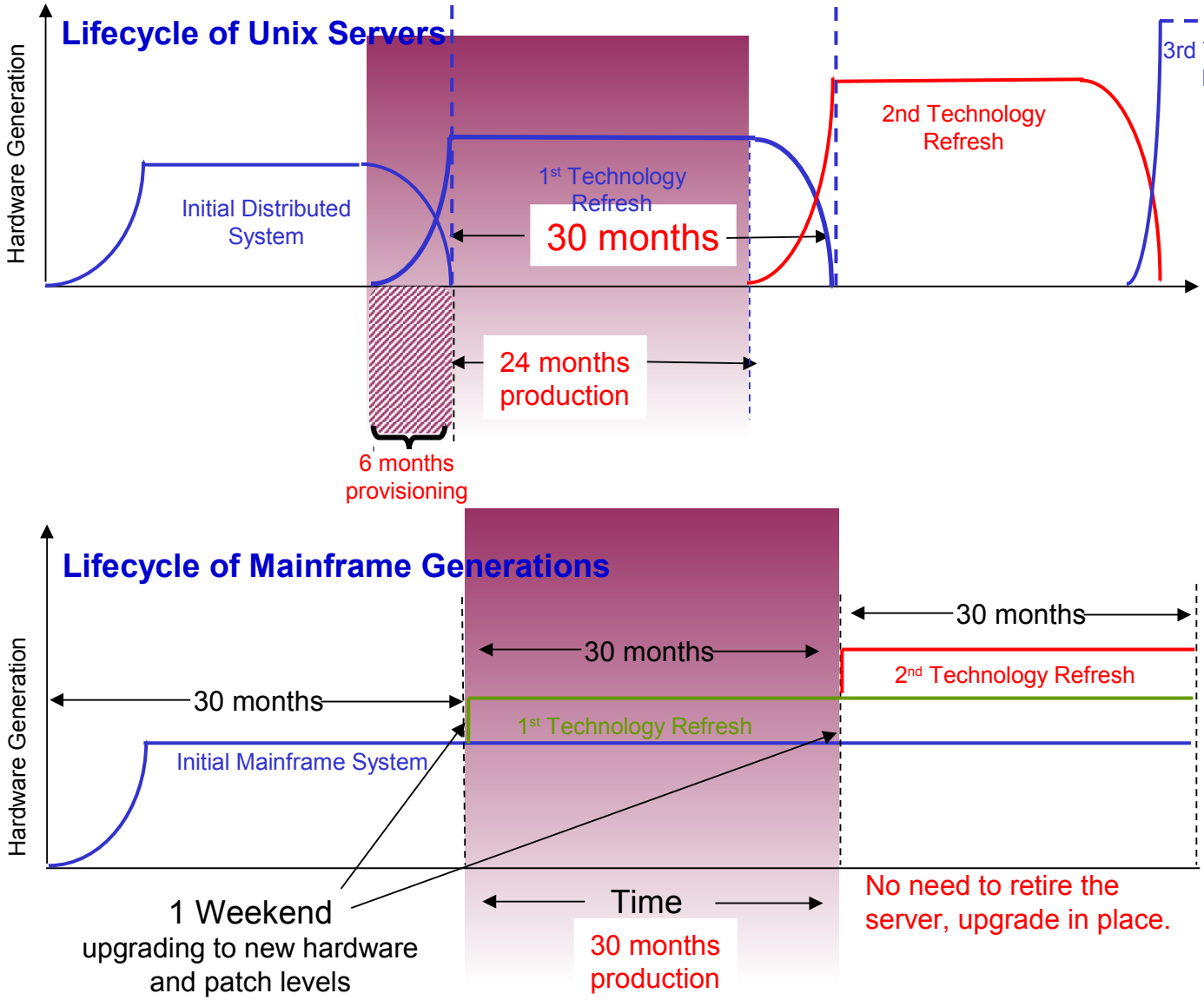
(2) Look For Not-so-obvious Distributed Platform Costs To Avoid

- Distributed servers refresh every 3 to 5 years
- Distributed server disaster recovery is typically at 100%
- Non-production environments require fewer resources on System z
- Customers often overlook significant tools replacement costs

Distributed Servers Need To Be Replaced Every 3 To 5 Years

- IT equipment refreshed 2 – 7 year intervals, normally 3 or 4 years
- Distributed servers re-purchased each time
  - Normally with some additional growth capacity (CPU, memory, I/O and other specialty cards like cryptographic offloads)
- With a growing mainframe, customers normally only have to purchase the additional (new) MIPS capacity
  - Existing MIPS are often carried over to the new hardware
  - Existing memory, I/O facilities and specialty processors / cards are also normally carried over to the new hardware
- Five year studies show this effect, short time periods do not

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

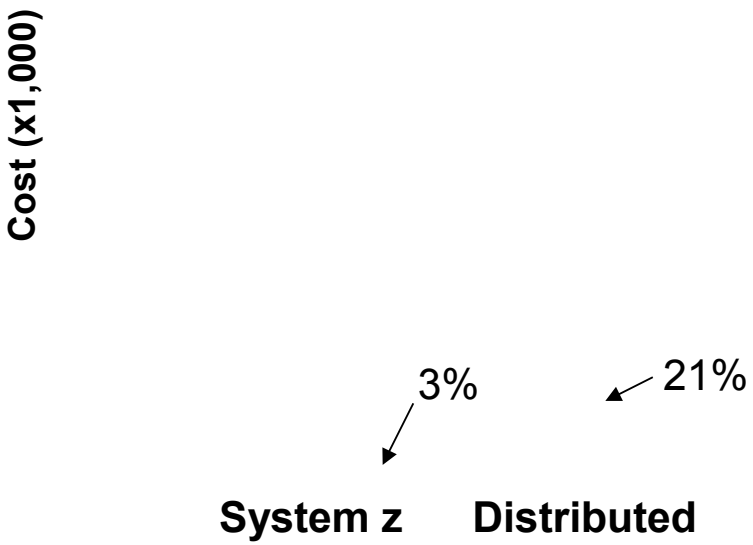
The mainframe by contrast is upgraded over a weekend and is fully productive at all times

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

Disaster Recovery Testing Is Typically More Expensive On Distributed Platforms Too

- A major US hotel chain
  - ~ 200 Distributed Servers (LinTel, Wintel, AIX, and HP-UX)

	Person-hours	*Elapsed days*	*Labor Cost*
Infrastructure Test (7 times)	1,144	7	$89,539
Full Test (4 times)	2,880	13	$225,416
Annual Total – Distributed	14,952*	73	$1,170,281
Mainframe Estimate	2,051*	10	$160,000

* Does not include DR planning and post-test debriefing

- ~~Customer Recovery Time Objective (RTO) estimates:~~
  - Distributed ~ 48 hours to 60 hours
  - Mainframe ~ 2 hours

- Conclusion: Mainframe both simplifies and improves DR testing

Customers Often Overlook Significant Tool Replacement Cost

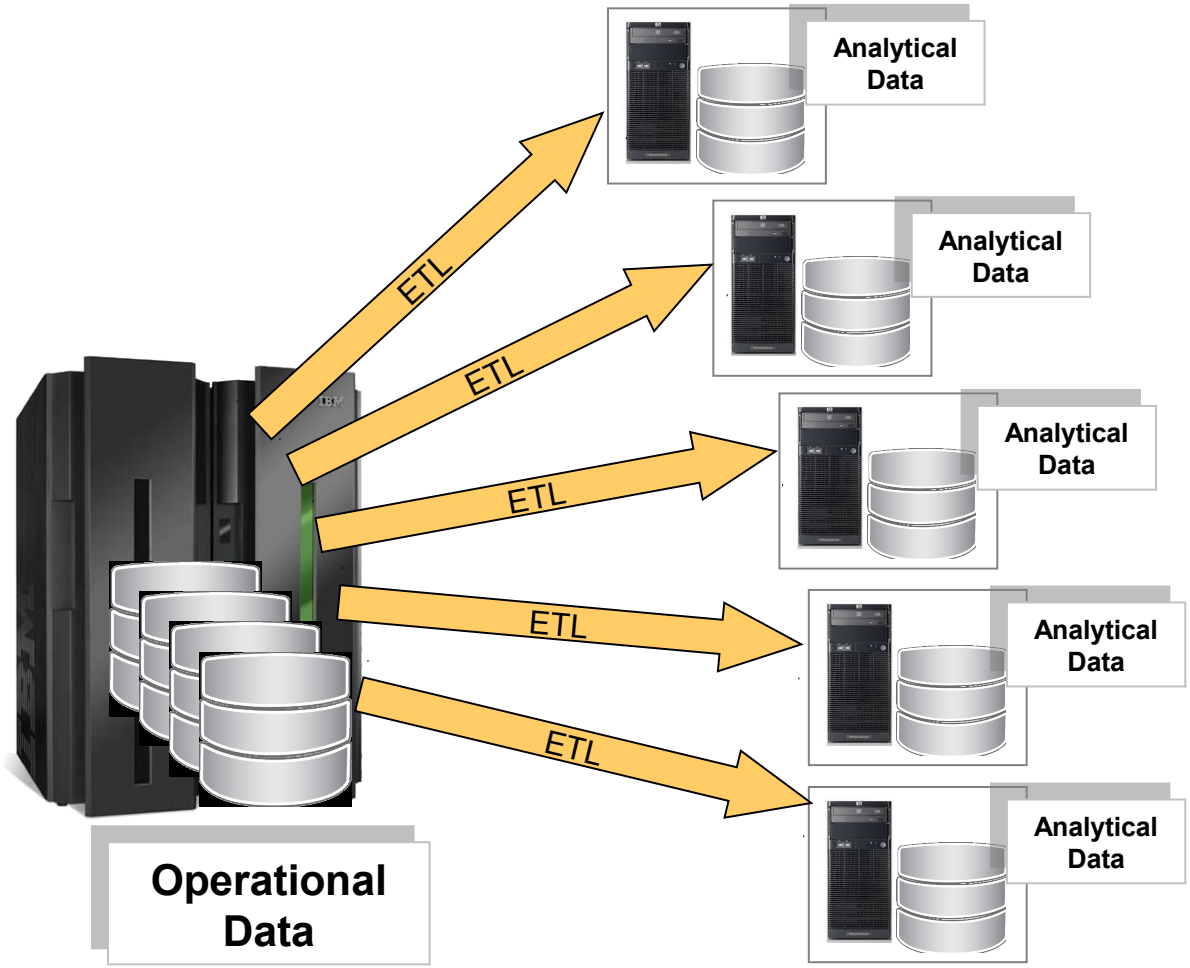
- Customers often struggle to identify all the replacement tools and middleware they will need for an offload
- Straight-line extrapolation of cost from the easily identified subset is often accurate enough

- Customer example: 261 total software products on z/OS
  - 37 product replacements identified in vendor proposal and IBM identified an additional 16 for a total of 53 products of 261 (20%)
- 208 products missing – how to estimate their likely cost, especially given that not all products will end up with one-for-one replacements:
  - Applications may be re-written to not need missing products
  - New code could be written to perform the function from scratch
  - Adding operations labor to manually do the function could be an option

(3) Consider Additional Platform Differences That Affect Cost

- Mainframe blockade effects
- Cost of adding incremental workloads to System z is less than linear
- Offloading chatty applications introduces latency
- Batch challenges non-mainframes
- Cost of administrative labor is lower on System z
- System z responds flexibly to unforeseen business events
- System z cost per unit of work is much lower than distributed

In Some Cases, "Mainframe Blockade" Results in Significant MIPS Burn



A large European bank:

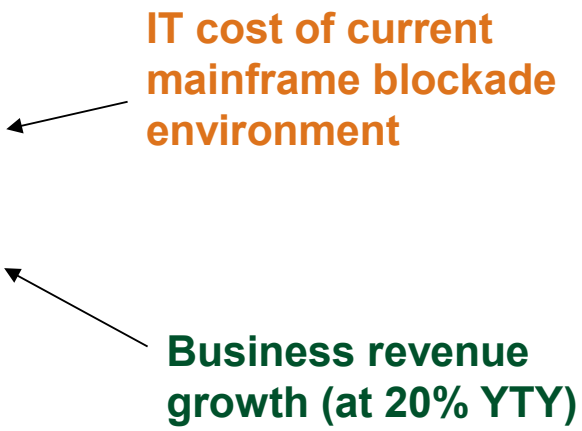
- 120 database images created from bulk data transfers
- 1,000 applications on 750 cores with 14,000 software titles
- ETL consuming 28% of total distributed cores and **16% of total MIPS**

A large Asian bank:

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

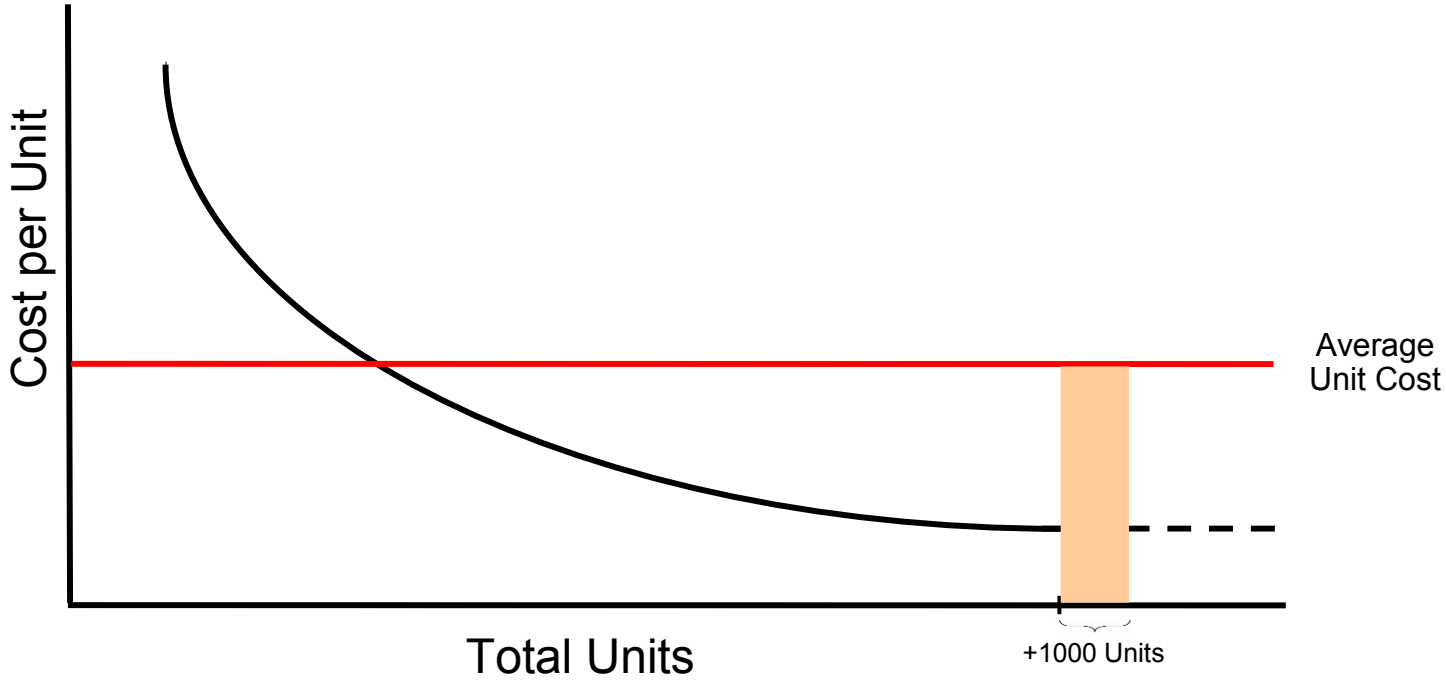
Businesses are Finding the Cost of a "Mainframe Blockade" Strategy is Not Sustainable!

Large European Bank –
Mainframe Blockade Environment
Compared To Business Growth

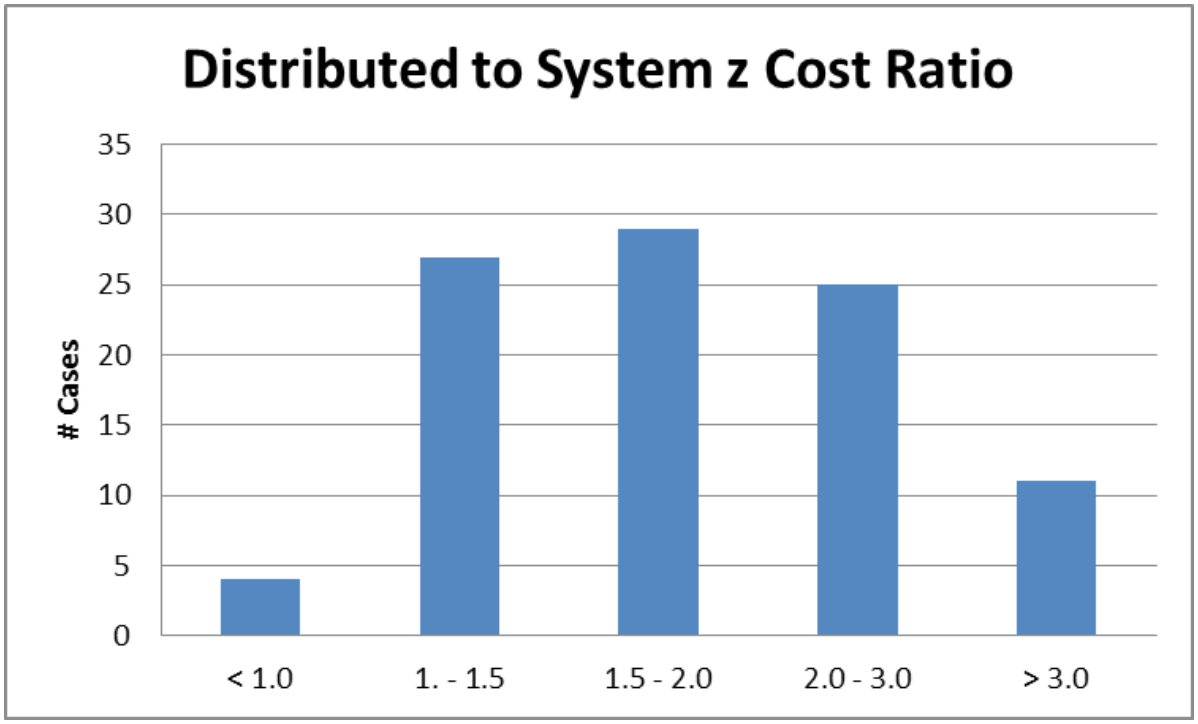


Cost Of Adding Incremental Workloads To System z Is Less Than Linear

- 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



Numerous TCO Studies Prove These Learned Lessons



- 97 “z vs distributed” out of 300+ total customer studies
- Average cost of distributed alternative is 2.2 times greater than System z
- Only 4 out of 97 studies showed lower costs on distributed

Eagle TCO Engagements

- **Free of Charge** total cost of ownership study that helps customers evaluate the lowest cost option among alternative approaches. The study usually requires one day for an on-site visit and is **specifically tailored to a customer's enterprise.**
- The study can be focused on at least one of the areas below :

- We conduct Eagle studies for System z, POWER, and PureSystems accounts
- Over 300 customer studies since the formation of the TCO Eagle team in 2007
- **Engage our Eagle-Eyed TCO Experts!**
  - Start by requesting sending an email to **eagletco@us.ibm.com**