zEnterprise – The Ideal Platform For Smarter Computing

TCO Lessons Learned From Customer Engagements

IBM Eagle Team Helps Customers Find The Lowest Cost Solution

- IBM Eagle Team comprises 8 senior technical IT staff – full time performing Eagle TCO studies
- Performing customer studies in every part of the world
- Over 250 studies completed in the past 5 years

- Eagle TCO studies often fall into the following categories
  - Offload
    - Proposed movement of workloads from a z/OS environment to distributed
  - New z/OS workload
    - Compare growth in distributed and z/OS mainframe environments
  - Consolidation
    - Movement of Linux or UNIX workloads to System z Linux
    - New workloads on Linux for System z

- Contact us via Craig Bender (csbender@us.ibm.com)

All the examples in this module are from actual customer studies

Typical Eagle TCO Study Approach

Approach

- Establish scope of study – applications, platforms, etc.
- Gather information
- Build the cost model
- Review with customer and iterate as needed
- Produce final report

Process

- Meet with customer face-to-face to establish scope and gather information (1 day). **Kick off and set expectations.**
- Complete study (30 day target)
- No charge
  - This does not mean no effort by the customer!

To Understand Total Cost
Four Dimensions Of Cost Should Be Considered

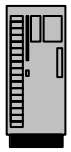
- Components
- Environments
- Time Factors
- Non-Functional Requirements / Qualities of Service

IT Solutions Require Many Cost Components

Components

The Cost Components may come from many sources
80:20 rule helps to achieve reasonable results in a short time

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

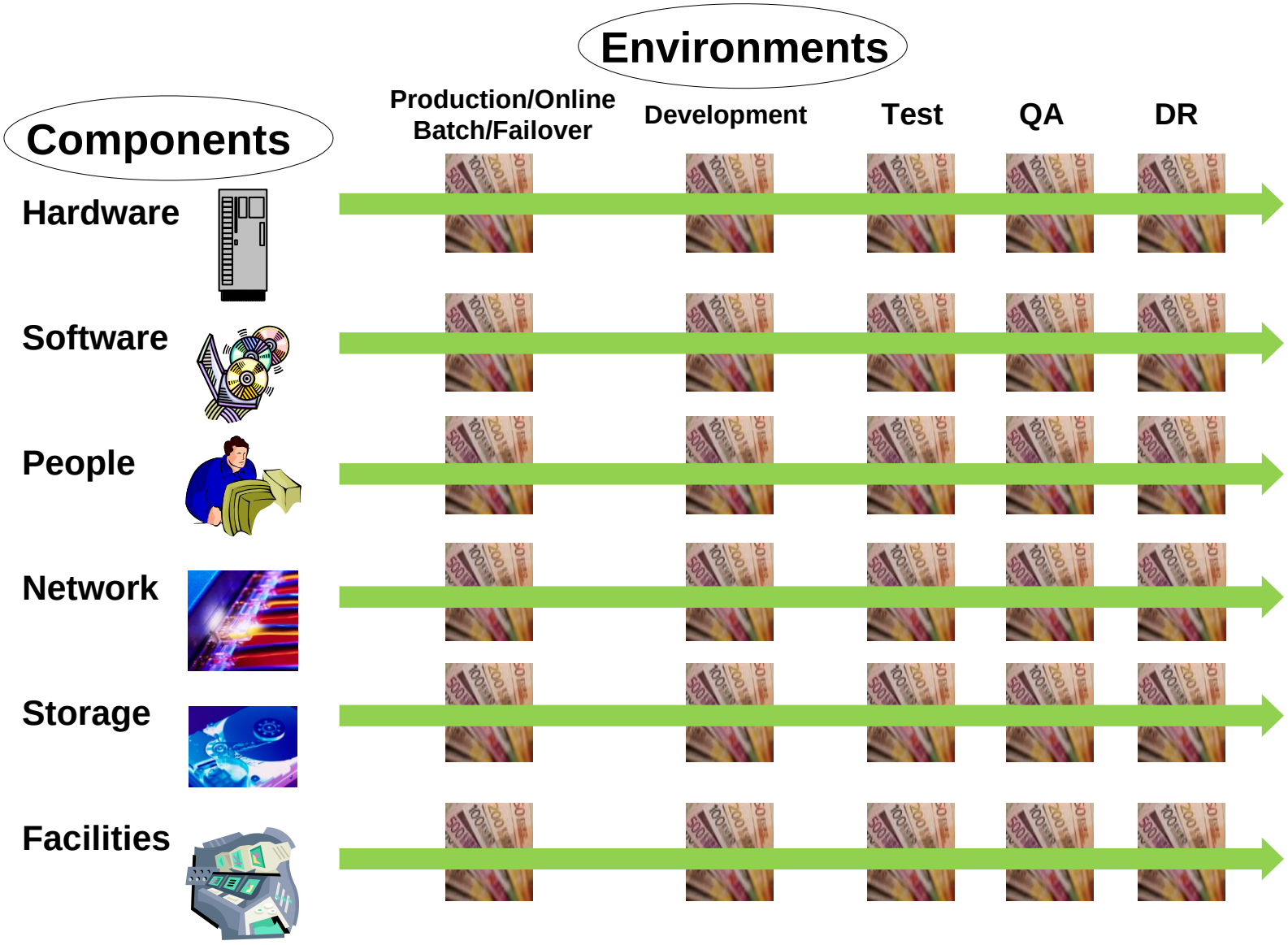
Storage

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

Facilities

Space, electricity, air cooling, Data center, asset life,
Infrastructure including UPS and generators, alternate site,
hot spots, new vs old, measured, managed

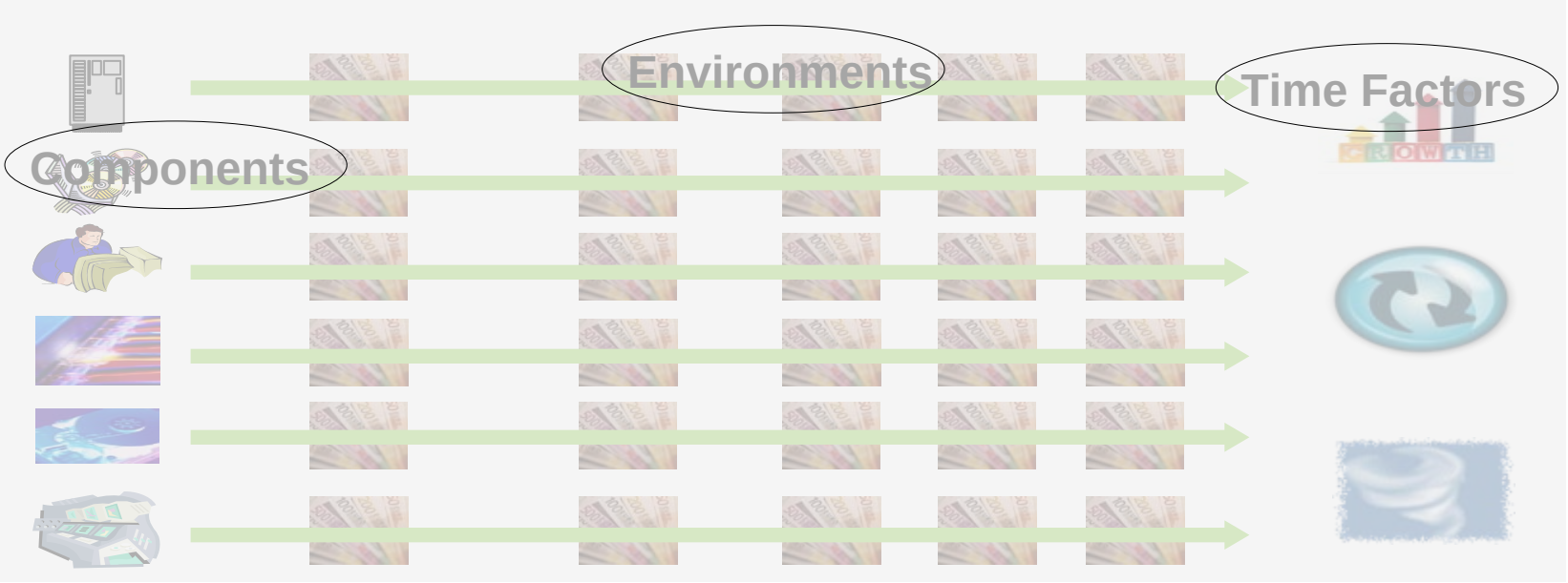
Environments Multiply Components



Time Factors Drive Growth And Cost

- Migration time and effort estimates are important elements of the total cost picture
- Business organic growth and/or planned business changes can dramatically change capacity requirements
  - Something as simple as a change of access channel or adding a new internet accessible feature can double or triple a components workload
  - It can be very useful to link a business metric (eg. active customer accounts) to workload (eg. daily transactions) and then use business inputs to drive the TCO case
- Other periodic changes – like hardware refresh or software remediation – can also be important cost drivers and are often very different cost profiles between platforms

Non-Functional Requirements Can Drive Additional Resource Requirements



Availability …Security …Resiliency …Scalability …

Qualities of Service, Non-Functional Requirements

Understand The Complete Picture

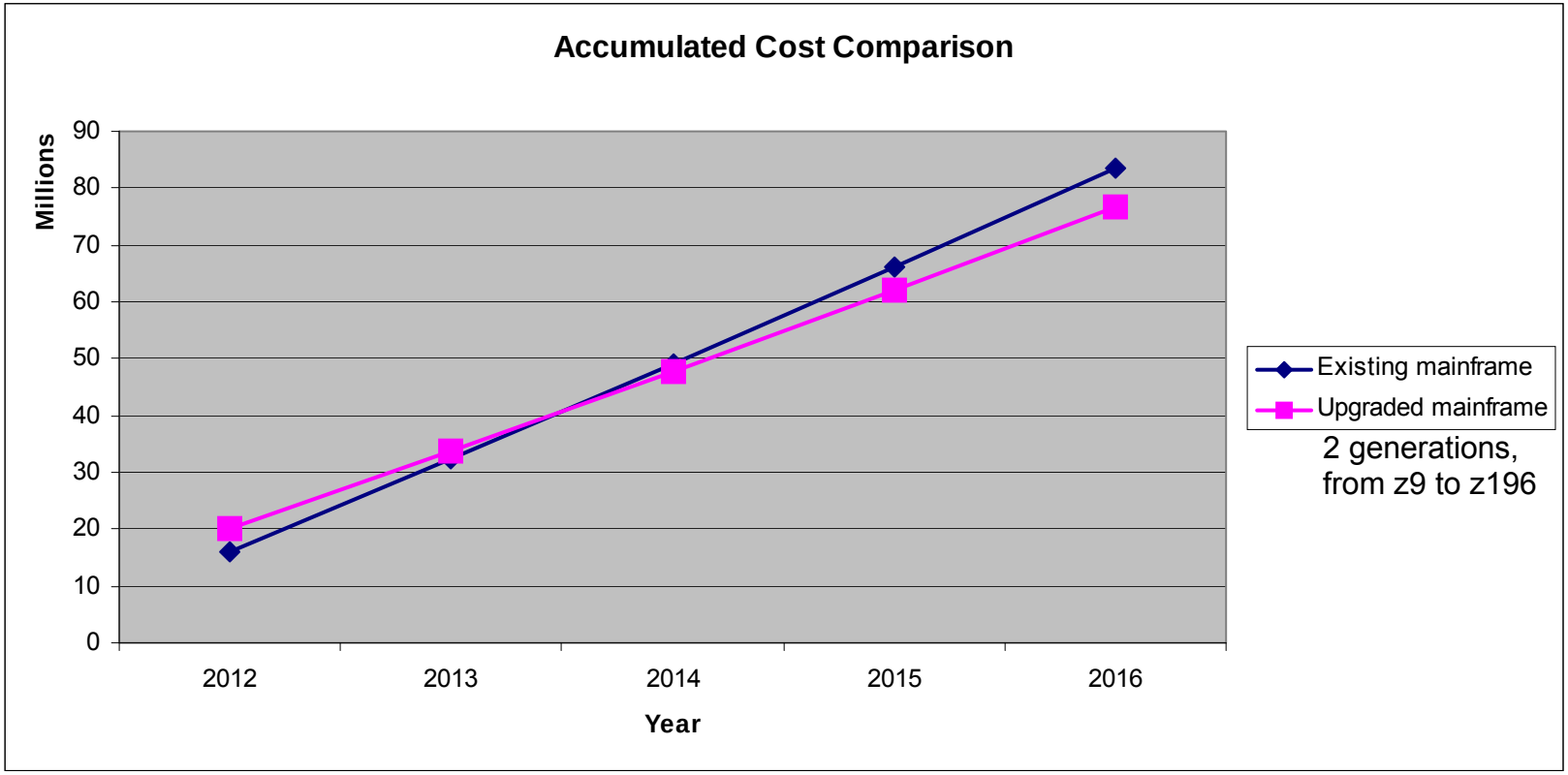
And then compare the alternatives on a level playing field

TCO Lessons Learned

- Make sure you are comparing to an optimum z environment
  - Currency reduces cost
  - Specialty processors and appliances revolutionize mainframe cost
  - Sub-capacity may produce free workloads
  - Replacing ISV software with IBM is a better deal
  - System z Linux consolidation is often a no-brainer

- Don’t forget to consider these platform cost differences
  - Cost of adding incremental workloads to System z is less than linear
  - New mainframe workloads with unbeatable price points via Solution Edition
  - Distributed servers need to be replaced every 3 to 5 years
  - Changing database can have dramatic capacity impacts
  - Disaster Recovery can be more expensive than you might think without a mainframe
  - Chatty applications are not good offload candidates
  - Customers often overlook significant tool replacement cost
  - Security breaches have high costs
  - Mainframe can respond flexibly to unforeseen business events
  - WAS on z/OS has better quality of service at the same price point
  - High offload costs almost never pay back in operational savings
  - Late migration projects extend dual systems cost
  - Non-production environments require fewer resources on the mainframe
  - Mainframe cost per unit of work much lower than distributed

Currency Reduces Cost – Hardware

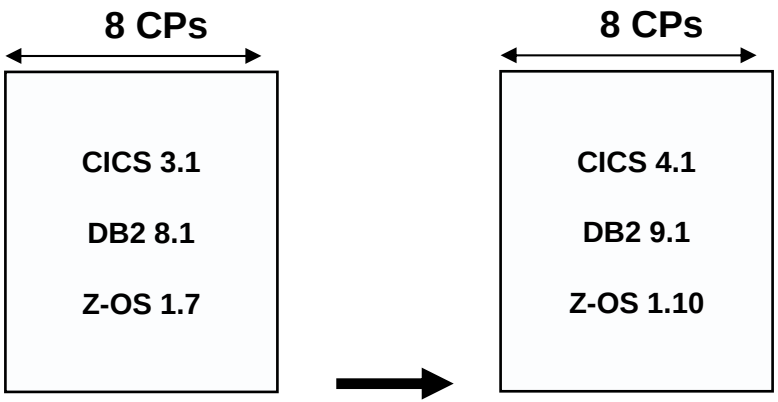


- Typical customer hardware refresh scenario
  - 2M investment pays back >1M savings every year – most cases positive in a 3 year period
  - Majority of the savings are from technology dividends, but also additional specialty processor offload

- Comparing latest technology servers to old mainframes is unfair but often done – e.g. “It would be much cheaper to run this workload using the latest <fill in your favourite server here>”

Currency Reduces Cost – Software

Benchmark – Only Software Changes:



System z — Same Hardware

1,560 tps — **2,330 tps**

Comments:
- Single LPAR configuration with CICS and DB2 collocated
- COBOL code recompiled with later COBOL compiler
- Later software (eg. DB2 v10) will be even better!

A similar real customer example:

- Customer running z/OS 1.6 across tens of LPARs delayed upgrade for 9 months due to "cost of upgrade effort"
- When eventually completed they found that z/OS 1.8 reduced each LPAR's MIPS substantially
  - Mainly improvements in RMF/SMF code inside z/OS
- The monthly software cost savings and hardware upgrade avoidance savings paid for the upgrade effort almost immediately and could have reduced operational spend many months earlier

Conclusion:
Keeping current saves money

Currency Reduces Cost – Don't Forget About Tuning And Newer Compilers

- For core systems of record, raw throughput can affect the scalability of the whole organization

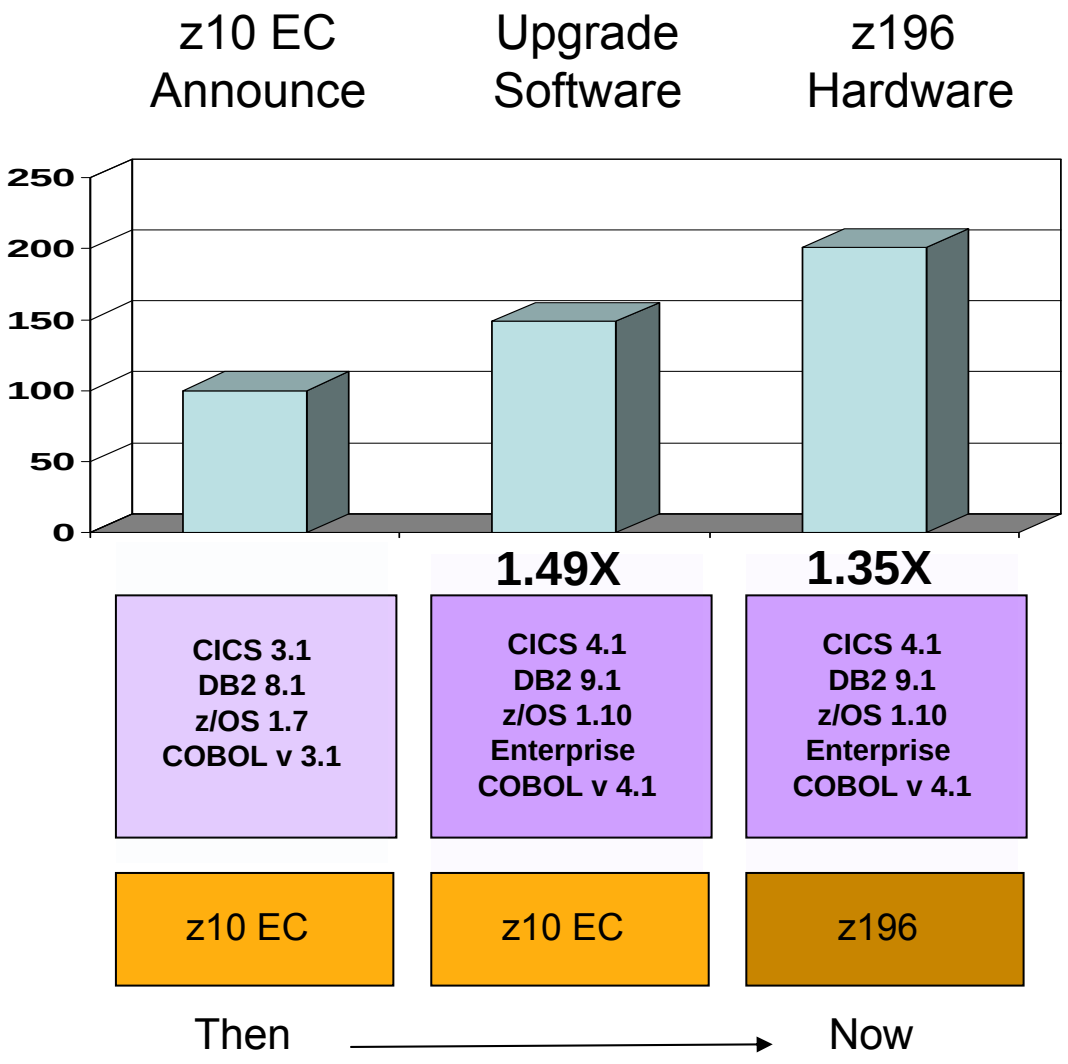
- IBM focus on tuning
  - Targets for release delta is -5% pathlength
  - Continued investment in compiler and JIT performance
  - Seeking out more hardware/software optimization and offload opportunities

- DB2 v10 is a good example
  - Between 5% and 10% CPU reduction for traditional workloads
    - BMW Autos found 38% pathlength reduction for their heavy insert workload
  - 10x number of users by relieving memory constraints
  - IDAA appliance in zEnterprise for DB2 offloads too

Currency Reduces Cost – CICS/DB2

Continued investment to optimize key software for z/OS environment

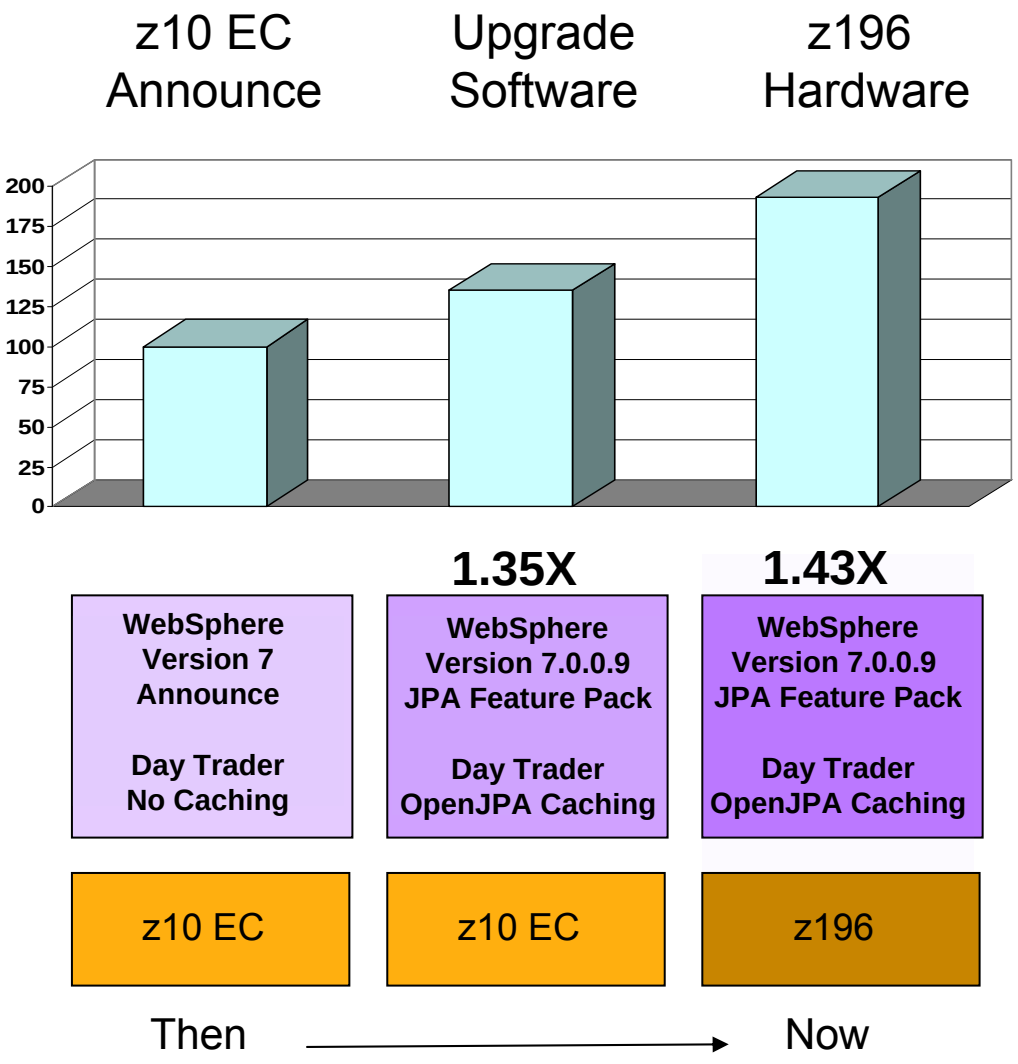
- Upgrade CICS/DB2 stack produces 1.49 times performance improvement on same z10 hardware
- Move to z196 hardware produces 1.35 times performance improvement
- From then to now – **2.01** times performance improvement



Currency Reduces Cost – WebSphere z/OS

Continued investment to optimize WebSphere software for z/OS environment

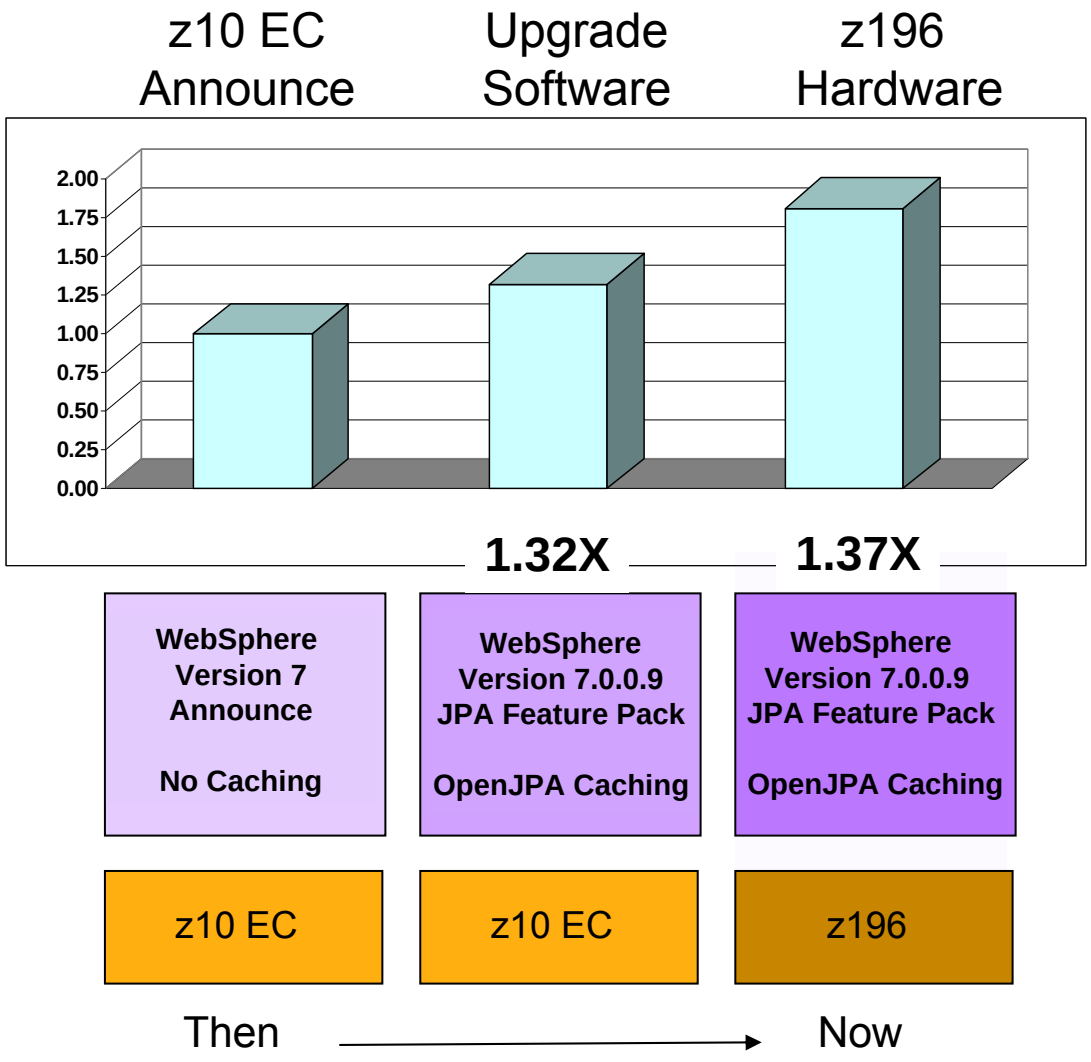
- 1.35 times performance improvement for JPA 2.0 applications that exploit the OpenJPA caching facilities available in the WebSphere Version 7 JPA Feature Pack.
- Move to z196 hardware produces 1.43 times performance improvement
- From then to now – **1.93** times performance improvement



Currency Reduces Cost – WebSphere On System z Linux

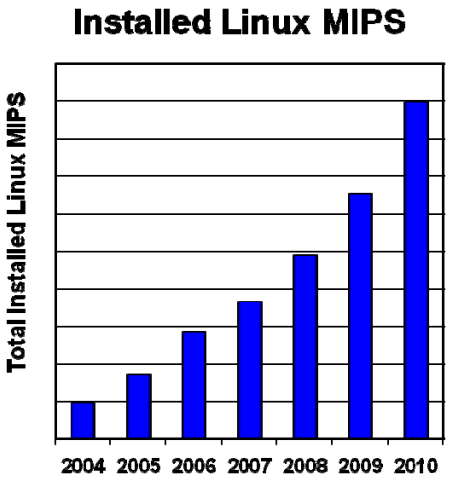
Similar results are achieved for WebSphere software in a Linux for System z environment

- 1.32 times performance improvement for JPA 2.0 applications that exploit the OpenJPA Caching facilities available in the WebSphere Version 7 JPA Feature Pack.
- Move to z196 hardware produces 1.37 times performance improvement
- Combined hardware and software - **1.81** times performance improvement



Specialty Processors And Appliances Revolutionize Mainframe Cost

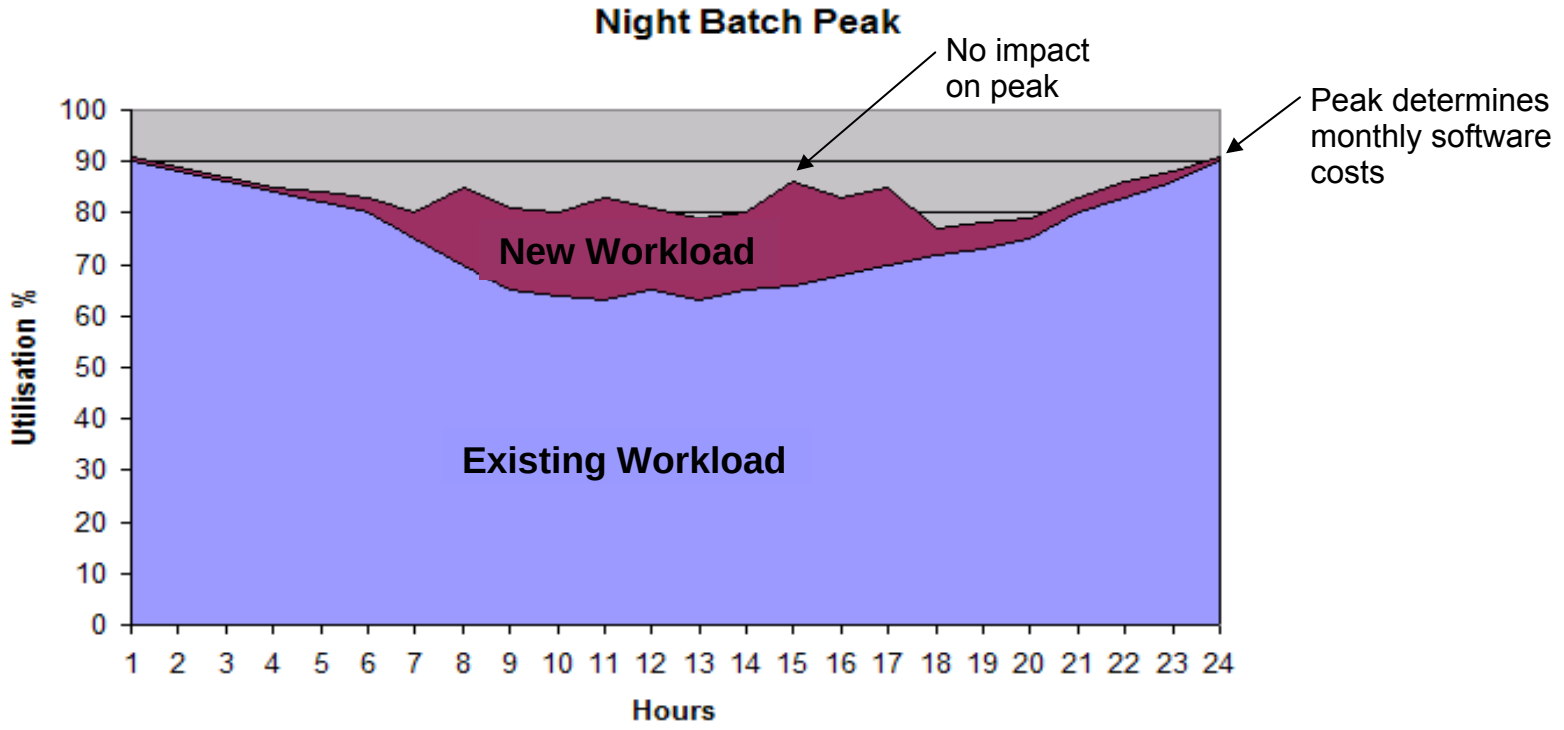
- Special assist processors for System z
  - For Java or XML workloads (zAAP)
  - For selected data, networking and security workloads (zIIP)
  - For Linux workloads (IFL)
- Attractive pricing
  - Hardware is from $35K per processor one time charge
    - Considerably less cost than a general purpose z/OS processor
  - No charge for IBM software running on zAAP/zIIP
  - IBM software on an IFL costs 120 PVU's (less than an Intel dual core)



Linux now 18% of all installed MIPS

- Cost impact of zAAP/zIIP depends on offload percentage
  - Between 40%-95% is customer experience
  - Tuned to task – for example, 3-5x pathlength expansion of Java countered with 80% offload to zAAP
- IFL processors have had a big impact on Enterprise Linux
  - This is now the lowest cost place to run most IBM software
  - Fantastic UNIX-style workload consolidation option, or a good place to co-locate UNIX-style apps with z/OS data and services
  - Over 80% of the top 100 System z clients are running Linux on the mainframe

Sub-Capacity May Produce Free Workloads



- Standard “overnight batch peak” profile – drives monthly software costs
- New workloads using the same middleware (eg. DB2, CICS, IMS, WAS, etc.) are totally free from a hardware and software perspective in the business day (and these are usually the largest cost components)
- Ensure you exploit any free workload opportunities, and conversely, avoid offloading free applications!

Replacing ISV Software With IBM Is A Better Deal

- This medium-sized financial company in Europe started with a typical mixed software environment, matching our average profile shown on the left
- Migrating to IBM tooling changed their software cost profile to that shown on the right

Average Profile (BEFORE)		
Weighted MIPS		8,800
Cost Per MIPS per Year		Profile
IBM Software	1,000.00	24.72%
		0.00%
ISV Software	1,540.00	38.07%
TOTAL SW	2,540.00	

Actuals (AFTER)		
Weighted MIPS		8,900
Cost Per MIPS per Year		Profile
IBM OTC	376.09	13.66%
IBM MLC	1,023.77	37.20%
ISV Software	136.09	4.94%
TOTAL SW	1,535.95	

- IBM software costs are now slightly higher than the average customer but ISV software costs are dramatically lower – saving $1000/MIPS per year from their annual software bill

Replacing ISV Software With IBM Is Also A Better Deal Than Offloading

	Mainframe Offload	Move To IBM Tooling
Investment $, Period	$54M over 2 years	$3M over 1 year
Predicted Annual Cost Savings	$13M from year 3	$6M from year 2
5 Year TCO, Breakeven	$140M, year 7	$101M, year 2
Level Of Risk	Very High	Very Low

- **Finance perspective**
  - **Large, risky and expensive project with distant payback under unlikely assumptions, versus**
  - **A small, low risk and cheap project with instant payback**

System z Linux Consolidation Is Often A No-Brainer

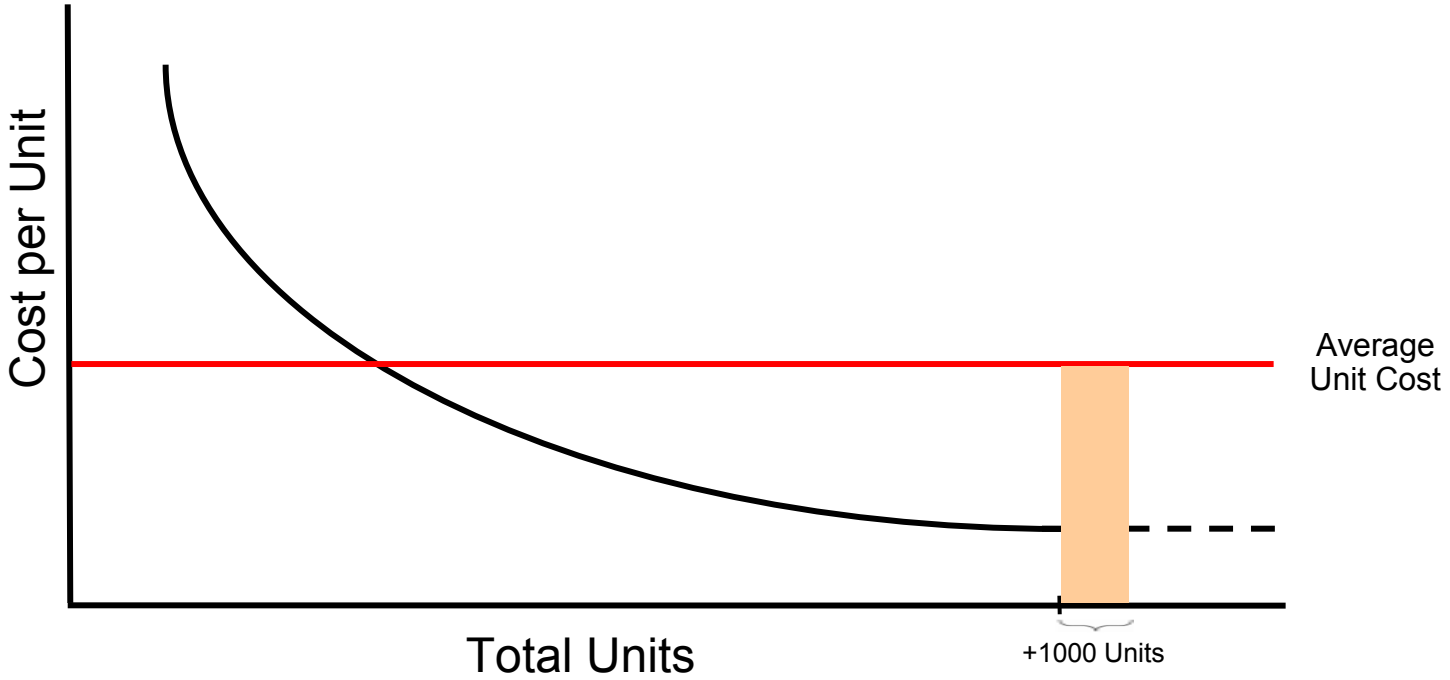
- Large financial services company with a mixture of dedicated (Oracle) and virtualized (WAS products) environments on Intel
  - Scaling out rapidly – up to 172 images on 836 cores

	Refresh existing x86	Consolidate on zLinux	Observation
Software	9.76M	3.70M	Software costs down 62%
Hardware	2.32M	5.22M	Hardware costs up 125%
Chargebacks	3.83M	0.69M	Charge to department down 82%
Migration	0.18M	0.41M	
Cost Avoidance	-	-0.61M	

- 5 Year Savings: 6.6M, with financing delivering savings in year 1 budget
- Despite significant Intel virtualization already, moving to System z Linux still achieves a further 26:1 core consolidation
- This proposal uses IBM Solution Edition pricing to deliver a bottom line price for an all-in 5 year total, in this case including two new mainframes in a pair of data centers

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 – so the average cost to this point is unrepresentative
- In our experience, a doubling of capacity results in as little as a 30% cost growth for software on z/OS for example
- Mainframe accounting often misrepresents incremental mainframe costs because it establishes a linear relationship between capacity usage and cost
  - E.g. $3000/MIPS/yr “rule of thumb” for each MIPS we have



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

- Customer determined that their current deployment of WAS applications would cost a similar amount each year on either the mainframe, or on a distributed platform
- Then they examined the incremental cost of adding one more large WAS application to each platform
- The mainframe demonstrated a clear advantage

	Incremental Mainframe	Incremental Distributed
5 Year TCO	**1.29M** **(657K OTC, 42K Y1, 147K Y2-5)**	**1.56M** **(378K OTC, 192K Y1, 249K Y2-5)**

- Although moving existing WAS applications between platforms is unlikely, future WAS deployments will therefore be targeted to the mainframe

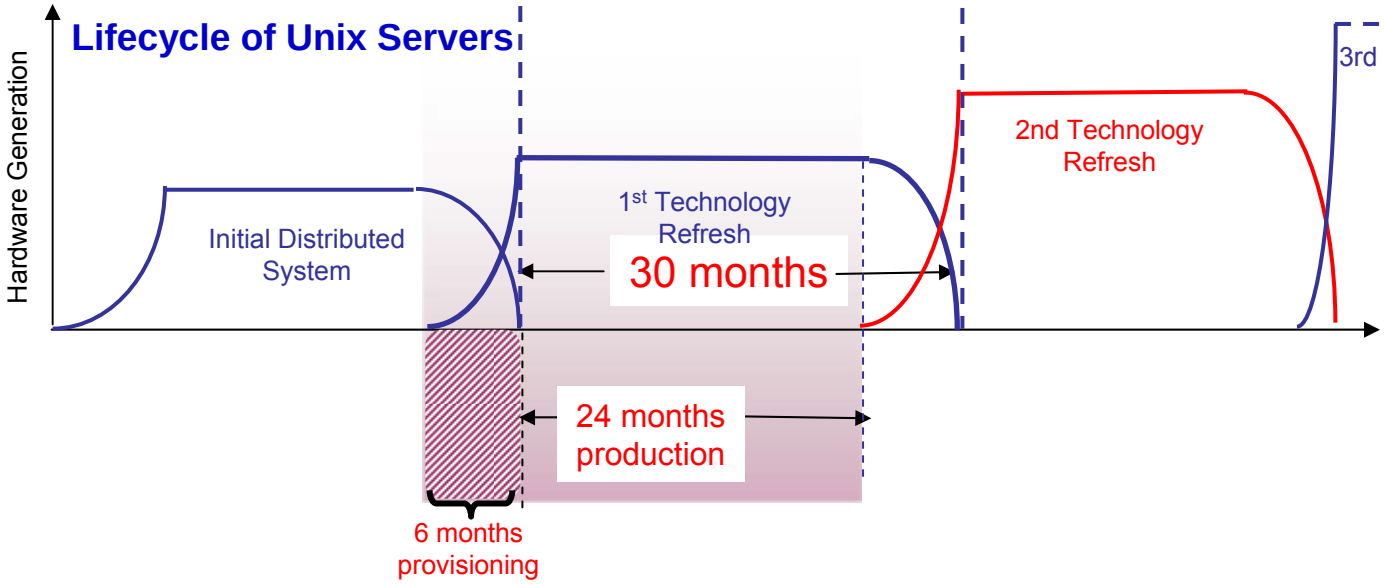
New Mainframe Workloads With Unbeatable Price Points Via Solution Edition

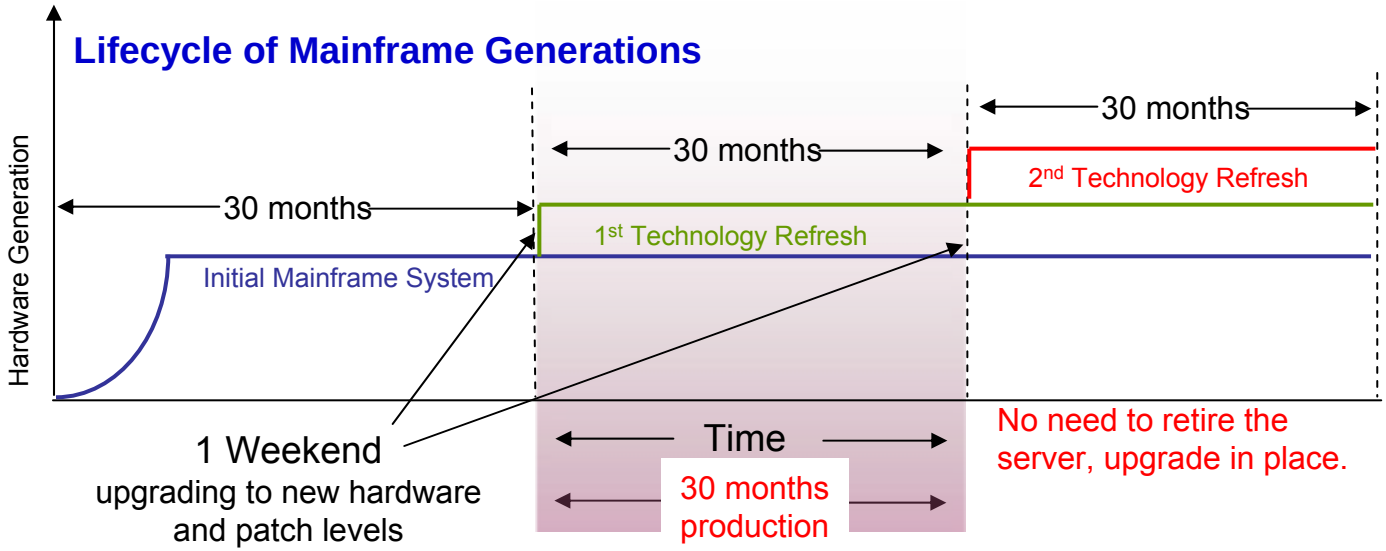
- Bundle of System z **hardware, software** and **maintenance**
  - 3 or 5 year **Best Price**
- Focus: new System z workload opportunities
  - Not for existing workloads
- Solution Editions usually include:
  - System z hardware (new footprint or incremental)
  - Prepaid hardware maintenance
  - Comprehensive middleware stack (including subscription and support)
  - Storage as an option for all Solution Editions
  - Services for some Solution Editions

Distributed Servers Need To Be Replaced Every 3 To 5 Years

- IT equipment is regularly refreshed as technology changes
  - This can range from 2 – 7 year intervals, normally 3 or 4 years
- With non-mainframe servers the existing system capacity is re-purchased each time, normally with some additional growth capacity
  - Applies to CPU capacity but also to memory, I/O cards and other specialty cards like crytographic offloads, etc.
- Conversely 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
  - To be tax compliant the original machine serial number is retained on the refreshed hardware
- This basic difference in technology refresh costs is only visible in a business case that extends for longer than a typical refresh period
  - Short business cases falsely favour non-mainframe systems

Distributed Servers Need To Be Replaced Every 3 To 5 Years (2)





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

Changing Database Can Have Dramatic Capacity Impacts

- IBM Taiwan recently migrated a large IMS DB customer to DB2
  - Capacity (MIPS) requirement for the database doubled (2x)
- A European customer reported a 3x storage requirement increase migrating IMS to Oracle RAC, as well as the usual huge growth in processor cores required when moving from a mainframe to distributed
- A small European customer rehosted a mainframe which once used IDMS but had been migrated to DB2 a few years prior to the rehost
  - This customer broke our record for the highest MIPS-to-RPE (distributed performance units from Ideas International) ratio we've ever seen
  - This suggests that mechanical conversion of code, which rehosting is supposed to be to maintain the "business logic", adds yet more overheads onto any existing overheads already in place – badly written applications get worse still!
  - Application and data modernization in place are faster, safer and more efficient – and normally cheaper to!

Disaster Recovery Can Be More Expensive Than You Might Think Without A Mainframe

- Disaster Recovery has two main costs – provisioning the environment, and the work effort required to do the regular testing
- We examined just the effort a customer expends on frequent DR testing
  - ~ 200 Distributed Servers (LinTel, Wintel, AIX, and HP-UX)

	Person-hours	Elapsed days	Labor Cost
Infrastructure Test (7 times)	1,144	7	$89,539
Of Which 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

* These numbers do not include time for DR planning and post-test debriefing

- In a typical DR scenario the customer believes that the RTO would be:
  - Distributed platforms ~ 48 hours to 60 hours
  - Mainframe platform ~ 20 minutes

- Conclusion: Mainframe both simplifies and improves DR testing

Chatty Applications Are Not Good Offload Candidates

- "Chatty" applications refers to those which contain frequent unpredictable reads and/or updates of one or more data sources
- Customer and offload vendor proposed a hybrid rehost
  - Move high MIPS CICS online workload to Linux on System z while retaining DB2 and some VSAM data on z/OS
- IBM requested input from a customer application architect and found:
  - Most CICS application programs access some VSAM and DB2 data which would need to remain in z/OS
  - Individual programs have a chatty data access profile
  - CICS transactions typically execute <100ms, the mainframe component of an end-to-end user interaction accounts for no more than 20% of the total
- Conclusion: "breaking apart" the CICS transaction programs is not a good idea because of induced latency impacts on SLA

Customers Often Overlook Significant Tool Replacement Cost

- Customers often struggle to identify all the replacement tools and middleware they will need for an offload
- Our experience suggests that a 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

- In this case we extrapolated from the known products cost and a few years later were proven to be very close to the mark

Security Breaches Have High Costs

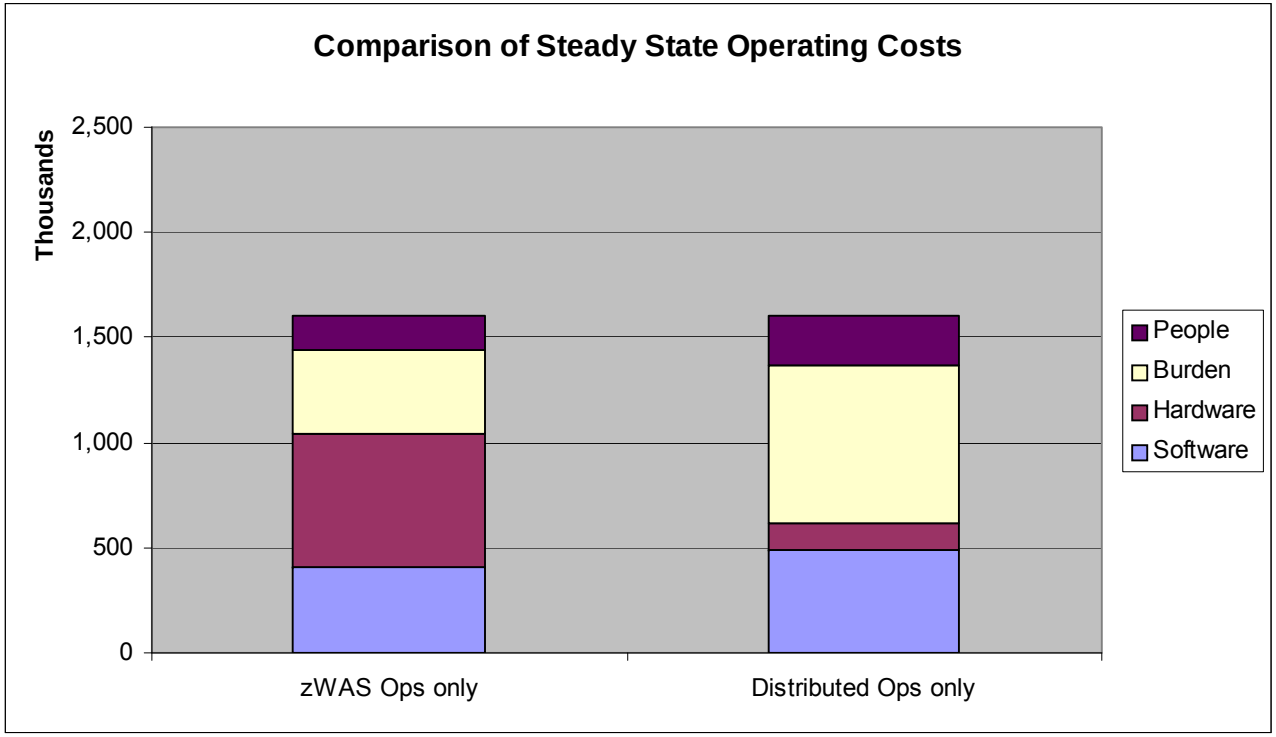
- Our real example security breach was a result of theft of physical hard drives which contained customer and agent data (data, audio and video files)
  - Fifty-seven drives stolen from a data closet in a leased call center office
- Government regulations required:
  - Letters should be sent to customers and agents
  - Various local and national authorities needed to be notified
  - Fines can be levied (but weren't in this case)
- Security breach was conservatively estimated to cost $10M, including:
  - Cost of diagnosing exposure
    - Approximately 700 people worked on identifying "what and who" was breached
  - Cost of restoring lost files from backups
  - Cost of handling phone calls from customers and documenting issues
  - Cost of sending letters to customers and agents (~$2M estimated, >1M letters)
  - Cost of replacement hardware (~$6M, switched to encrypted drives in all machines / SAN, requiring other hardware and software upgrades across the organization)
  - Cost of credit monitoring (volunteered by the organization; not required by law)

Mainframe Can Respond Flexibly To Unforeseen Business Events

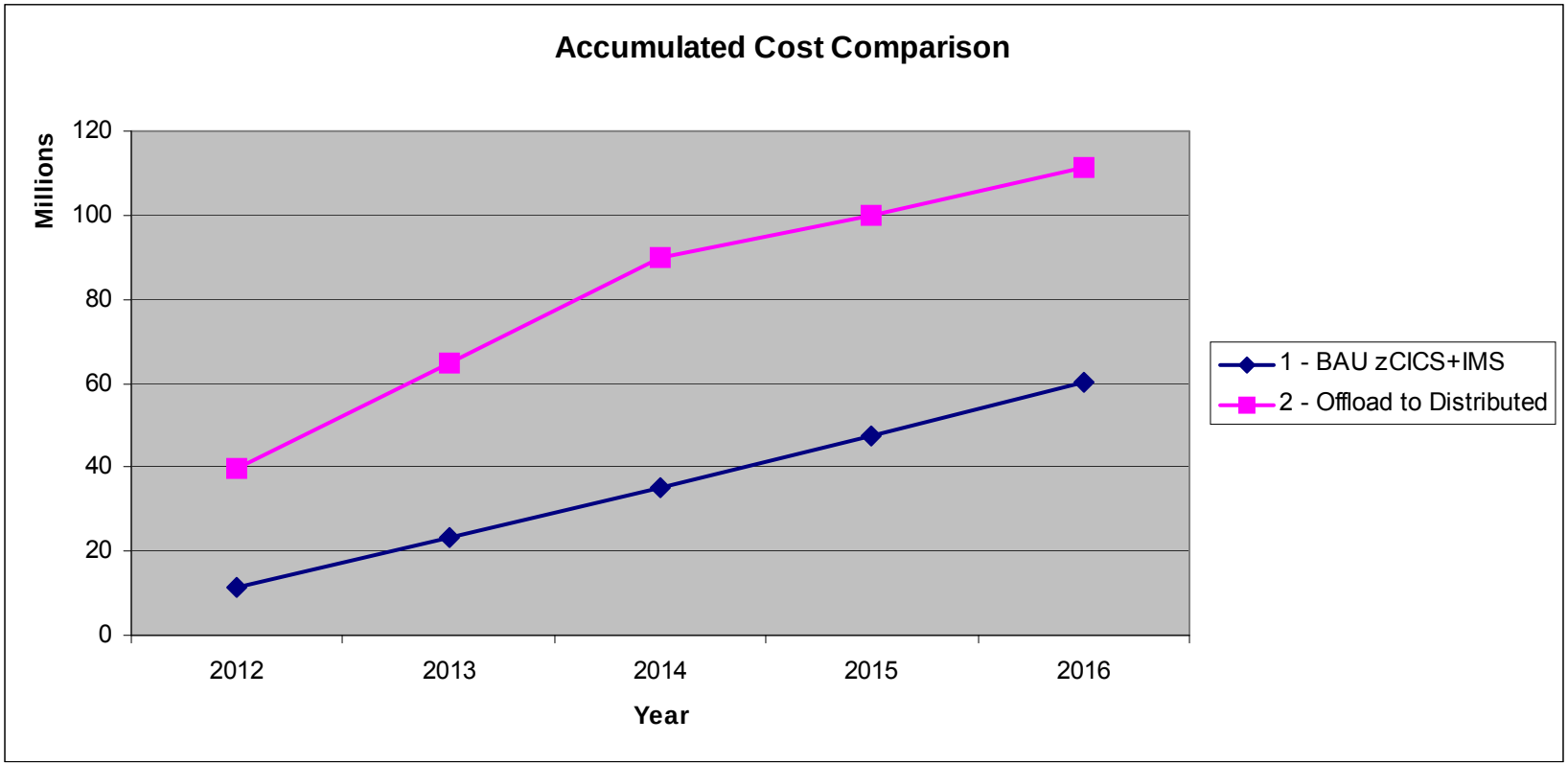
- Transportation company experienced a natural disaster
  - Required them to re-run a whole weeks business while continuing to operate normally
  - Able to turn on double capacity immediately to achieve this
- Customer informed IT department that they would be running a Super Bowl advertisement with very short notice
  - Massive capacity spike
  - Temporarily turned on additional capacity
  - Stress tested their systems prior to the event despite short notice

WAS On z/OS Has Better Quality Of Service At The Same Price Point

- This compares the existing mainframe to the estimated cost of the same workload (many mixed WAS applications) on non-mainframe
  - 2 plexed mainframes with 5 instances of z/OS versus 14 Sun servers running 60 instances of Solaris
- Burden refers to storage (disk and tape), IT Service Management (operations and administration), networking, and datacenter/facilities
- Just average annual steady state operations costs, no acquisition or growth



High Offload Costs Almost Never Pay Back In Operational Savings



- Typical mainframe offload cost profile showing the huge investment cost up front
- Once migration completes the operating cost is quite similar to the current mainframe so there will never be payback on this investment
- This chart compares existing old mainframe to new distributed

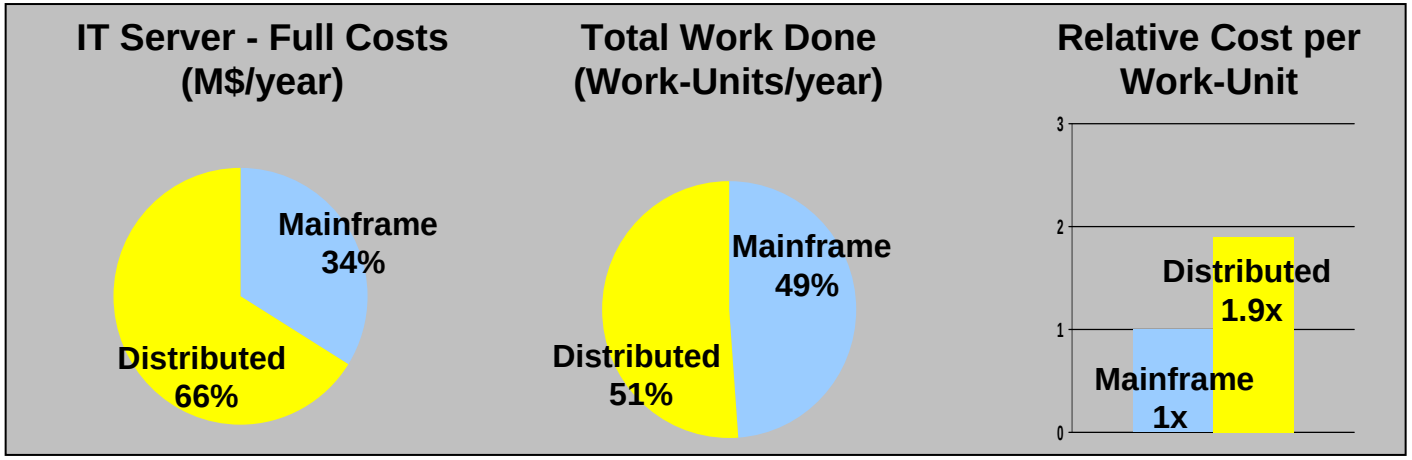
Late Migration Projects Extend Dual Systems Cost

- Customer wanted to move 3,500 MIPS with Micro Focus believing it would only cost $10M and complete within 1 year
- 18 months later they had spent $25M and moved only a 10% of their MIPS
- Additional costs came from
  - Internal staff to cover the overrun
  - Substantial manual steps replaced mainframe automation
  - Needed many additional software products
  - Ended up acquiring additional distributed capacity over initial prediction (just to support the 10% they actually offloaded)
  - Extending the dual-running period of the rehost
- Executive sponsor lost their job

Non-Production Environments Require Fewer Resources On The Mainframe

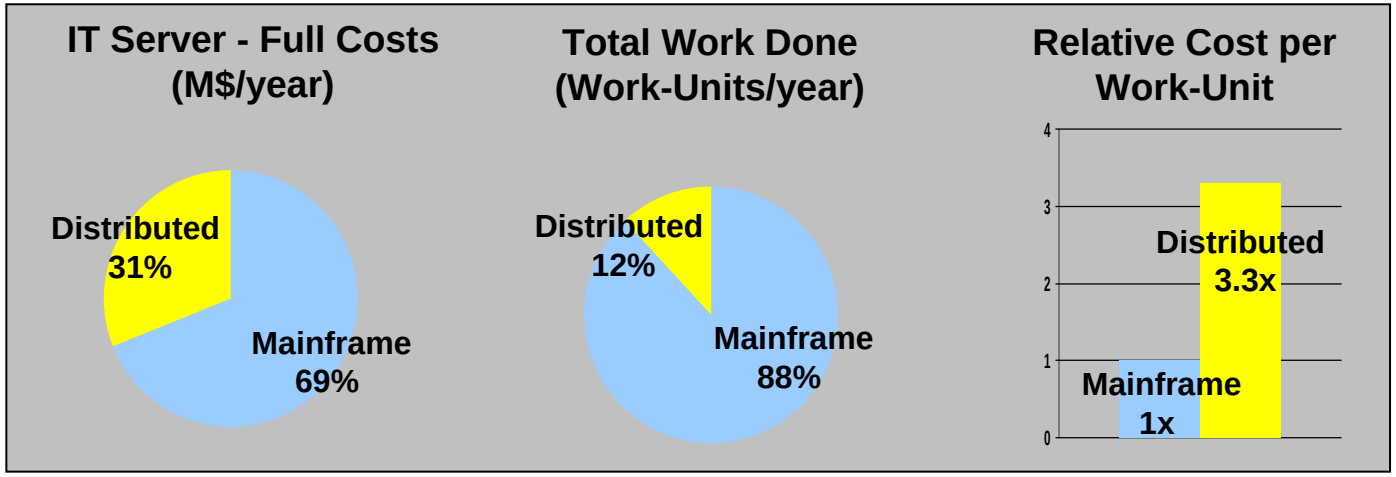
- Development and Test capacity
  - Centralized – Prod+25%
  - Distributed – a range, but around Prod+150%
- High Availability mechanisms for Production
  - Dedicated failover (Prod x 2.5)
  - N+1 clustered (Prod x 2 worst case)
  - Mainframe (usually Prod x 1, sometimes less!)

Mainframe Cost Per Unit Of Work Much Lower Than Distributed



Case 1:
Very large retail bank
total of service delivery plus application development

Mainly UNIX distributed (>5000 servers)



Case 2
Medium retail bank
service delivery only
Mainly mainframe
2500 mips, 13 M txns/day
Limited distributed servers
Windows + some UNIX
(~350 servers, 12% util'n)

Data from 3Q06 Scorpion studies

The distributed IT Total Cost/Work-Unit is approx. 2-3x Mainframe Cost/Work-Unit
The Mainframe typically does more work, Distributed has a lot of supporting infrastructure

Eagle TCO Studies Show Offloading To Distributed Is Far More Expensive

- **IBM Eagle Studies are TCO analyses for customers**
  - Cost and risk analysis of mainframe vs. alternative
  - Tailored to individual customer workloads
    - Cost factors unique to each enterprise
    - Costs evaluated over five-year period
- **94% of IBM Eagle studies concluded that System z offered a better solution than the distributed alternative**
  - System z is 52% the cost of distributed when offloading from z/OS
  - System z is 60% the cost of distributed when consolidating Linux applications
- Contact *Craig Bender* (csbender@us.ibm.com)

Results may vary