G22

TCO: Comparing System z and Distributed Environments – A Customer's View

Marlin Maddy - IBM - Executive IT Consultant

IBM System z Expo
September 17-21, 2007
San Antonio, TX

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

AIX*	System z9
AS/400*	Tivoli*
DB2*	WebSphere*
DB2 Universal Database	z/OS*
IBM*	z/VM*
IBM eServer	zSeries*
IBM logo*	
On Demand Business logo	
pSeries*	
S/390*	

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Intel is a trademark of the Intel Corporation in the United States and other countries.
Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
Java and all Java-related trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc., in the United States and other countries.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
UNIX is a registered trademark of The Open Group in the United States and other countries.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.

Have you heard/made these statements?

" My mainframe cost 2x, 5x, 10x compared to my distributed environment“ Mainframe

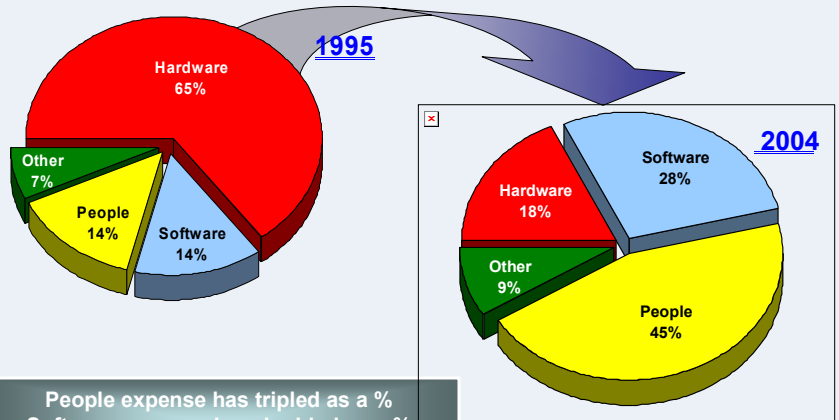
“Mainframe software costs are expensive and are driving me off the platform” Mainframe

"We are on a get off the mainframe strategy“ Mainframe

"We keep adding servers and people“ Distributed

“Our infrastructure can not support our servers” Distributed

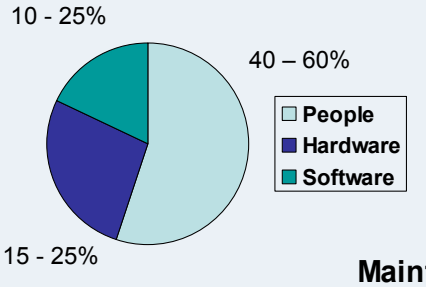
Throughout the past 10 years the cost dynamics of supporting corporate IT infrastructures has changed significantly as has the landscape



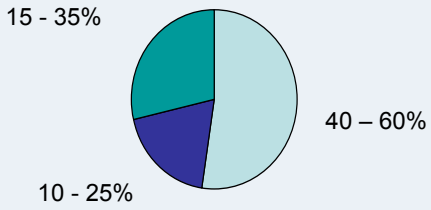
People expense has tripled as a %
Software expense has doubled as a %
Hardware is less than 1/3 of its original %

Cost components

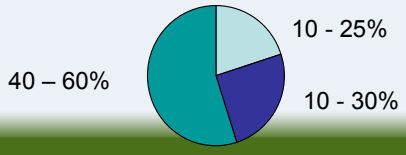
PC servers



Unix servers



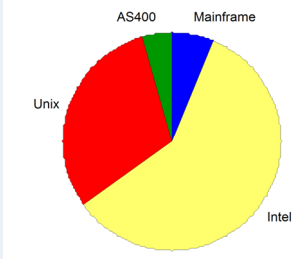
Mainframe servers



Based on IBM Scorpion customer analyses

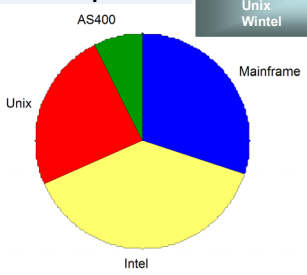
Installed vs. Used capacity

Installed Capacity:
33M tpms*



Used Capacity:
4M tpms*

Typical Utilization	
Mainframe	80 – 90%
Unix	10 – 30%
Wintel	5 – 12%



* system capacity (tpms) is an approximation of the transaction processing capability of each system. It cannot be compared to other commercial ratings or benchmarks and is invalid outside of the context of this IBM study.

Server utilization varies significantly by platform and that needs to be accounted for in the business case. Virtualization and partitioning has improved processor utilizations for distributed platforms but they are still considerably lower than the mainframe environment

Why is utilization low?

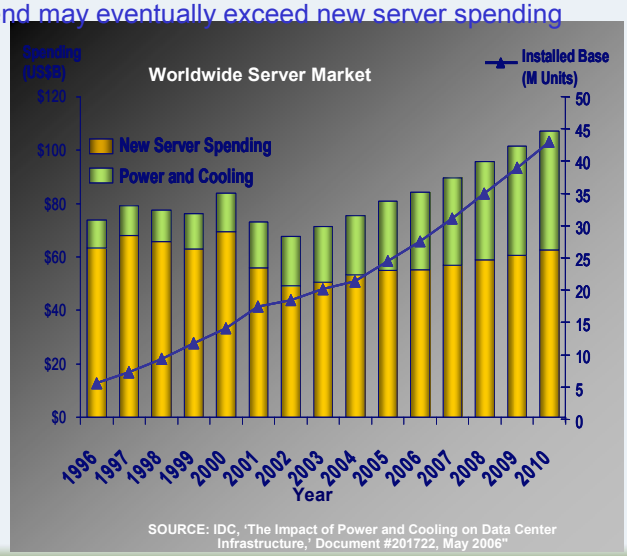
- Use of response time as a measure of capacity
  - **Buy rather than tune**
- Backup, development, test, training and integration servers
- Peaked, spiky workloads on dedicated rather than shared hardware
- I/O Bound workloads, contention
- Utilization controlled to avoid system stress and outages
- Incompatible release levels
- Incompatible maintenance windows

The New Economics of IT; A Paradigm Shift

- Power and cooling spend may eventually exceed new server spending

2000 – Raw processing "horsepower" is the primary goal, while the infrastructure to support it is assumed ready

2006 – Raw processing "horsepower" is a given, but the infrastructure to support deployment is a limiting factor

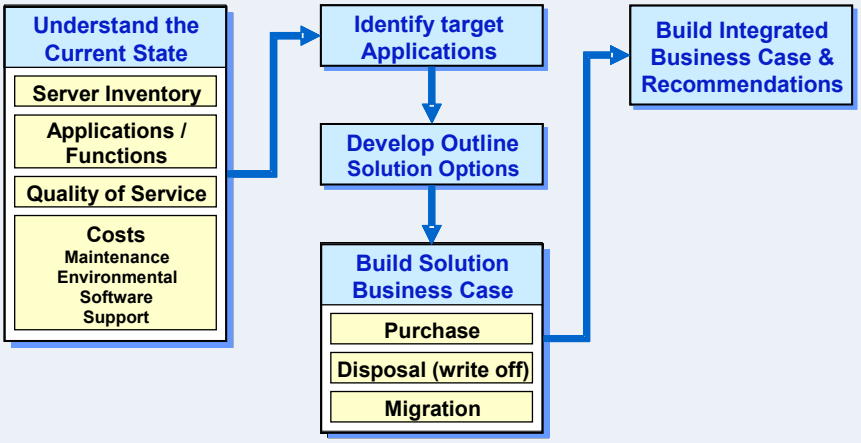


SOURCE: IDC, 'The Impact of Power and Cooling on Data Center Infrastructure,' Document #201722, May 2006"

Building a business case to fairly compare platforms

IBM System z Expo
September 17-21, 2007
San Antonio, TX

Scorpion recommended study flow



Considerations for comparative analysis. How do we level the playing field so we can make a good platform decision?

1. **Select the Type of Analysis**
   - Will we do TCA or TCO or ICO? What is the difference?
2. **Choose the Application(s) – are they strategic, mission critical, infrastructure?**
   - What is the "current environment"? Can we make assumptions about the future?
3. **Include all components - What components does the application need?**
   - How much resource sharing is possible? Are some components in place?
4. **Consider only useable capacity - Know your workload and the target!**
   - Will my environment have to change to do this? How much?
5. **Choose realistic, scalable staffing numbers - Will we have to add staff?**
   - Can history show us the future? Who does what, and will it change?
6. **Build business cases that reflect real costs, not necessarily practices**
   - Chargeback mechanisms? Upgrade versus disposal? Specialty engines
7. **Quantify Quality of Service - in dollars if possible**
   - Cost of outage? Recoverability? Performance & Response time?
8. **What is the impact of this application in the greater context of the enterprise?**
   - There may be positive and negative impacts. What at the intangibles?

And there are more items to consider…

1.Select the type of analysis

1. **Total Cost of Acquisition (TCA)**
   - The easiest and fastest to do
   - Typically considers new hardware and software
   - Well oriented to computer professionals - forward looking, technology based, no financials
   - Vendors love it - oriented to glossy brochures, a functional matrix/cost comparison
   - The least useful to IT department - leads to complexity and duplication

2. **Total Cost of Ownership (TCO)**
   - Not as easy to do, but still project oriented
   - Typically considers new hardware, software, environmentals, staff, and other
   - Many Rules of Thumb in the press - sense of security
   - Consultants love it - easy to add value, difficult to confirm - fishing expeditions
   - Good for cost recovery but less than ideal for making smart business decisions

3. **Incremental Cost of Ownership (ICO)**
   - Most accurate for decision making
   - Most difficult and thus most infrequently used - usually one-of-a-kind
   - Depends on understanding the current environment - implementation context
   - Typically considers current and new hardware, software, environmentals, staff, and other
   - Can include or exclude the cost of strategic change - one project can fund another
   - Is **NOT** the budget $$ divided by MIPS or machines, times the new stuff!

Build business cases that reflect real costs, not practices

- **Business case horizon – 3 years**
  - **Upgrade vs. disposal vs. reuse vs. "free"**
  - **Being "fair" may not be fair**
- **Chargeback**
  - **Mainframe chargeback pools are typically 50 -60% overstated**
    - **Software contracts**
    - **People – Operations and monitoring**
    - **Default bucket – history**
  - **Open systems charged by box, regardless of cost/complexity**
  - **Infrastructure omitted all together**
  - **Incremental cost is 20 -25% of the full chargeback cost**
    - **Hardware price/performance**
    - **Software flat slope, ISVs?**
    - **Do you need to hire additional people? New skills?**

TCO -- Components

of TCO include these over the evaluation period:

- Acquisition Cost (Direct)– Cost of initial acquisition and planned upgrades
- Maintenance Cost (Direct) -- Cost of required maintenance above warranty
- Operating System SW Cost (Direct) – Cost of operating system for servers
- Application and Database SW Cost (Direct) – Cost of application and database SW
- Special Management SW Cost (Direct) – Cost of any specific/unique SW
- Facilities – Power/Space/Cooling (Direct/Indirect) – Cost of facilities
- Networking Cost – (Direct/Indirect) – Cost of communications Components
- Application Admin Cost (Direct) – Cost of SysAdmin personnel
- Operations Cost (Direct/Indirect) – Cost of standard Operations
- Training Cost (Indirect) – Cost of any specialized training
- Conversion Cost (Direct)– Cost of converting existing systems
- Availability & Reliability (Direct/Indirect) – Cost of any additional servers
- Disaster Recovery (Direct/Indirect) – Cost of additional server, Are the alternatives equal?
- Productivity (Direct) – Are there improvements?
- Security costs (Direct/Indirect) – Cost of additional server, Are the alternatives equal?

Full range of TCO factors considerations – often ignored

- *Availability*
  - High availability
  - Hours of operation
- *Backup / Restore / Site Recovery*
  - Backup
  - Disaster Scenario
  - Restore
  - Effort for Complete Site Recovery
  - SAN effort
- *Infrastructure Cost*
  - Space
  - Power
  - Network Infrastructure
  - Storage Infrastructure
- *Additional development and implementation*
  - Investment for one platform – reproduction for others
- *Controlling and Accounting*
  - Analyzing the systems
  - Cost
- *Operations Effort*
  - Monitoring, Operating
  - Problem Determination
  - Server Management Tools
  - Integrated Server Management – Enterprise Wide
- *Security*
  - Authentication / Authorization
  - User Administration
  - Data Security
  - Server and OS Security
  - RACF vs. other solutions
- *Deployment and Support*
  - System Programming
    - Keeping consistent OS and SW Level
    - Database Effort
  - Middleware
    - SW Maintenance
    - SW Distribution (across firewall)
  - Application
    - Technology Upgrade
    - System Release change without interrupts
- *Operating Concept*
  - Development of an operating procedure
  - Feasibility of the developed procedure
  - Automation
- *Resource Utilization and Performance*
  - Mixed Workload / Batch
  - Resource Sharing
    - shared nothing vs. shared everything
  - Parallel Sysplex vs. Other Concepts
  - Response Time
  - Performance Management
  - Peak handling / scalability
- *Integration*
  - Integrated Functionality vs. Functionality to be implemented (possibly with 3rd party tools)
  - Balanced System
  - Integration of / into Standards
- *Further Availability Aspects*
  - Planned outages
  - Unplanned outages
  - Automated Take Over
  - Uninterrupted Take Over (especially for DB)
  - Workload Management across physical borders
  - Business continuity
  - Availability effects for other applications / projects
  - End User Service
  - End User Productivity
  - Virtualization
- *Skills and Resources*
  - Personnel Education
  - Availability of Resources

TCO – Chargeback Considerations

Chargeback models for various hardware and software platforms vary for a myriad of reasons and are likely steeped in the history of the IT department and the IT industry itself.

When comparing solution platforms customers may “fairly”, or “unfairly” burden one solution or another due to real or perceived overhead costs. Be aware of these factors and try to the real cost of the platform options.

Allocation methods will vary by platform and may include: “MIPS”, “MSU’s”, “Servers” “CPUs”, “Cores”, “Processor Value Units”, “Users”, “Transactions”, etc. A wide variety of methods.

Cost factors for hardware and software that may get allocated to your solution might include:
Purchase, Maintenance, Supporting Hardware (power, networks, cooling), headcount related to system admin, physical plant, floor space, operations, systems management, systems management software licensing, backup infrastructure, disaster recovery facilities cost, etc.

TCO – Chargeback

After all the TCO analysis of alternatives has been completed and you have detailed cost of ownership over multiple years your analysis can then be injected into the customers chargeback system.

Chargeback can massively distort the alternatives depending on how they architected their cost pools. This is especially true for many legacy systems where the mainframe absorbed much of the extraneous costs (there is the urban myth of corporate jets ending up in the mainframe cost pool – not all myths are untrue).

You have to be aware of how chargeback will be used. Changing this will be a massive undertaking and likely beyond the scope of the platform alternatives being considered but if you know beforehand you may be able to influence how the alternatives are presented.

In some cases, customers have developed chargeback models so that IT departments can chargeback the cost of running IT to business end user departments, presumably to fairly reflect the cost of doing business.

Chargeback is fine for allocating costs back to the business, but is not a good methodology for making business decisions

Key Points – Distributed Costs

- **The cost of running additional workload on distributed servers may go up linearly in many areas**
  - **Labor is now the highest cost element in distributed environments**
  - **Administrative staff costs increase in proportion to the number of servers**
  - **New workload requires additional servers**
  - **Cost of additional servers is linear**
  - **Cost of software licenses is linear**
  - **Electrical and air conditioning costs also increasing and may require a significant datacenter build out**
- Result – scale out strategies do not dramatically reduce the cost per unit of work as the workload grows

Facility costs

- " The data center utility bill exceeds the cost of acquiring new computers for some companies"
  - **annual electric bill for a high-end server is $8760 - IDC**
- ""The cost of datacenter floor space is inconsequential compared with the cost of operating and cooling a datacenter"
  - **"Modern computing hardware requires 3 sq ft of cooling infrastructure for every sq ft of floor space"**
  - **ex. $20/sq ft for space and $60/sq ft for cooling**
  - **That is 6x the ratio of 10 years ago**
- " The average annual utility cost for a 100,000 sq ft datacenter has reached $5.9M
- "You pay once to power the systems and again to cool them"
  - **And again and again for redundancy**
- "Businesses paid about 20% more last year than in 2004, with rates jumping as much as 40% is some parts of the country"

Source: Information Week 2/27/06

A recent article

The new TCO and the value of the mainframe
Published on: 11 Jan 2007

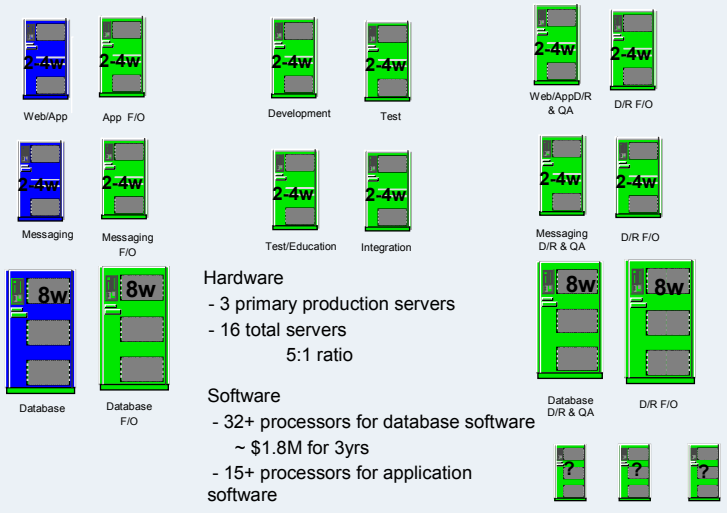
The Mainstream -- January 2007 -- Issue 22

http://www-306.ibm.com/software/swnews/swnews.nsf/n/cres6x3lc8

Server Proliferation

- Describe a current application environment
  - **Production**
    - **Database server? How many?**
    - **Application server? How many?**
    - **Messaging server? How many?**
    - **Failover servers? For each?**
  - **Additional Servers**
    - **Development servers? Multiple levels?**
    - **Test servers? Multiple levels?**
    - **Systems test? Multiple levels?**
    - **Quality Assurance servers?**
    - **Education servers?**
  - **Disaster Recovery**
    - **Do you have a DR site?**
- How many applications/types of workload do you have?

e-business Servers - Complexity and Cost



Hardware
- 3 primary production servers
- 16 total servers
  5:1 ratio

Software
- 32+ processors for database software
  ~ $1.8M for 3yrs
- 15+ processors for application software

Key Examples of Common Business Case Myths (1/2)

Myth #1:
- **Units required:** Comparing total System z to only the operational distributed systems

Reality:
- Z installations include backup, test LPARs; redundancy and other units often triple the actual distributed solution size

1:3 ratio of production to total distributed servers is common

Myth #2:
- **Utilization:** Distributed utilization is low, but virtualization fixes it

Reality:
- Utilization is so low and difficult to optimize that virtualization is only part of the solution; consolidation is a much better alternative

5-30% vs.80-95%

Myth #3:
- **One time costs** discourage mainframe refresh; distributed fits budgets better

Reality:
- System z only charges for additional capacity, vs. throwaway; hidden costs are often missed in distributed, while chargeback schemes exaggerate mainframe TCA

Every 4 years

Myth #4:
- **SW costs** for System z are always excessively higher

Reality:
- Distributed SW costs are often hidden; some processor-based costs can be higher than System z

Distributed software costs may be "higher" for real work

Myth #5:
- **People costs** are hard to estimate and are best left out of TCO

Reality:
- People costs are a large portion of the cost, and distributed systems are labor intensive

40-60% vs. 10-25%

Key Examples of Common Business Case Myths (2/2)

Myth #6:
- **Space and power** are not major factors in TCO

Reality:
- System z consolidation saves space – and can delay a space upgrade for years
- Typical x86 server consumes 40% of TCO in power costs over 4 years!

Can prevent $5M space upgrades and power

Myth #7:
- **Chargeback** pools have been recently corrected

Reality:
- Analyses show significant overstatement in almost every study

Average 50% overstated

Myth #8:
- **Security:** Distributed systems are now "good enough"

Reality:
- Significant losses and embarrassment in banking; System z is unparalleled in security

One management POC vs. hundreds

Myth #9:
- **Flexibility:** changes are slower on mainframes than distributed

Reality:
- Testing is faster on most mainframe apps and downtime reduced

Add capacity on demand

Myth #10:
- **Reliability:** Distributed systems are now good enough for mission critical apps

Reality:
- Workloads are returning to System z for reliability and disaster recovery; distributed systems keep adding more boxes (and staff) as back-up

99.9% vs. ??

Source: IBM, Morgan Stanley

Have a Great Day!

IBM System z Expo
September 17-21, 2007
San Antonio, TX