WebSphere® for z/OS® Prescriptive Use Cases

Providing the Benefits of Co-location on z/OS with Competitive Price Performance

September 30, 2009

z/OS Software Used in This Effort

z/OS V1.9

CTG V7.0

CICS V3.2

DB2 V9.10 for z/OS, PUT0903

WPS for z/OS V6.2.0.1

WAS for z/OS V6.1.0.24

Distributed Software Used in This Effort

AIX V6.1

WAS for AIX V6.1.0.23

WPS for AIX V6.2.0.1

JDBC V3.52.110

VIO Server V2.1.1.10

System z Hardware Used in This Effort

z10 2097 in LPAR mode utilizing two engines

System p Hardware Used in This Effort

Power 595 9119 with a two-processor partition

Table of Contents

Executive Overview

Positive trends in customer adoption of WebSphere® middleware on System z® and z/OS® are driving the need for a more comprehensive evaluation framework to determine the platform on which a new workload runs. One of the most visible components for customers in making a platform decision is price/performance, which is the subject of this white paper. To address price competitiveness, we use the term, “prescriptive use cases” going forward in this document. The WebSphere z/OS prescriptive use cases are intended to provide solid proof points and exhibits that demonstrate the cost benefits of deploying new WebSphere workloads on z/OS for customers who have already made a significant investment in z/OS. The prescriptive use cases will provide our customers with answers to the following two questions:

- Why and when should I deploy WebSphere middleware on System z and z/OS?
- Is the price/performance competitive with alternative deployment options?

The prescriptive use cases effort recognizes the customer’s starting point, in terms of commitment to z/OS, and then illustrates the value of deploying the new WebSphere workloads on z/OS, that is, the achievement of performance equivalent to alternative deployment platforms at a competitive price while building on the customer’s existing z/OS investment in Qualities of Service (QoS).

The ground-breaking results of the prescriptive use cases effort summarized in this document show that deploying new WebSphere workloads on z/OS, co-located on the same system as their traditional batch and online workloads, enables transaction rates that are price competitive with alternatives. This enables customers to place the same WebSphere workloads on System z with local access to z/OS “back ends” like DB2® and CICS®. The z/OS configuration can handle hundreds of thousands of Business Services (such as “open new account”) per hour while providing the additional z/OS QoS such as high availability, security, and scalability that customers have invested in for their traditional online and batch workloads.

Introduction to the WebSphere for z/OS Prescriptive Use Cases

The following introduction outlines the problem statement, hypothesis, and a high-level description of the WebSphere for z/OS prescriptive use cases output. It also explains how and when the information can be used and to whom (which customers) it applies.

Problem Statement

Prior to the WebSphere for z/OS prescriptive use cases effort described in this white paper, for customers who wanted to understand the cost competitiveness of deploying new WebSphere workloads on z/OS, IBM provided limited collateral that exhibited solid proof points to support this assessment. Most information available was anecdotal data related to the QoS provided by WebSphere Application Server or WebSphere Process Server on z/OS. In some situations, expensive, time-consuming comparisons were needed to effectively evaluate the value propositions between a z/OS and distributed implementation, often resulting in unacceptable delays.

Hypothesis

The expectation for the prescriptive use cases effort is that the performance of a single distributed environment connecting to back-end z/OS resources, measured at a set of various utilization rates (or steady state Business Services per hour rates) is price comparable to an all-z/OS environment. Stated differently, the purpose is to prove the following hypothesis:

Co-location of a WPS workload on z/OS along with “traditional z/OS” (online; batch) workloads provides equivalent throughput and cost, as compared to a distributed (System p®/AIX® application that interacts with CICS and DB2 z/OS).

High-Level Description of the Validation Effort Deliverable

The key deliverable from the validation effort is documented price/performance competitiveness based on quantifiable costs, thereby establishing a proof point for continued customer investment and deployment on z/OS.

To Which Customers Does This Apply?

The prescriptive use cases focus on existing System z and z/OS customers who have already deployed CICS, IMS™, and/or DB2 workloads on z/OS; that is, customers who have considerable investment in online and batch processing on z/OS. These customers might have extra capacity to support new workloads and understand the value of z/OS and its inherent QoS, or based on the findings of this document, might elect to add capacity to further exploit the advantages of their System z environment.

A Customer Business Scenario for Investment in New Hardware and Software: Key Steps

First, a customer business challenge is identified. One such challenge, Enterprise Business Transformation (EBT), is necessary for many IBM customers. An example of EBT is when a bank acquires additional banks and must then deploy its existing business processes to the newly acquired banks.

Second, a customer accepts SOA as the infrastructure to deploy the EBT solution. IBM has been very successful in helping customers to understand that IBM has a solution that implements SOA: the WebSphere portfolio.

In the third step, given that IBM understands the business problems and can provide a good architecture, infrastructure, and product set, the next question is raised: *where should it be deployed*? Sometimes, the perception is that it is too expensive to deploy on z/OS.

Without an answer to the question of price competitiveness, the discussion of z/OS as a candidate platform for deployment sometimes stops.

This is when the evidence that z/OS is a price/performance-competitive platform is crucial.

The prescriptive use cases can be applied when price/performance assessments are needed, such as:

- Deployment of new WebSphere applications
- Significant anticipated growth of an existing WebSphere application, that is, organic growth or merger/acquisition
- Optimization of an existing server and WebSphere application infrastructure
- Significant changes to QoS requirements

Applying the “Rule of Four”

Four basic rules can also help to determine whether to deploy on z/OS:

1. Do you have current hardware (z9®, z10)?
2. Does the business application require frequent access to z/OS data (DB2 z/OS or IMS-DB) or transactions (CICS, IMS, MQ)?
3. Are the related workloads highly dynamic, unpredictable, or of high business value?
4. Can the workloads take advantage of specialty processors (zIIP or zAAP)?

Applying the ”Rule of Four” is more of an individual entity experience than a hard and fast set of “rules” but, as stated in Rule 1, if you already have an investment in System z, then you are part of the way there. Rule 2 relates to co-location, that is, co-locating the applications (WPS or WAS) close to the data (DB2 z/OS) and transactions (like CICS, IMS, or WMQ). Having everything under the control of z/OS has many benefits (beyond the scope of this paper) that can provide significant value to your applications and business. Rule 3 is about utilization rates and the application’s importance to the business. This also leads to further platform differentiation and operational benefits, that is, WPS is a WAS application and inherits its System z affinity from WAS. It provides, for example, the WAS z/OS controller/servant architecture and the exploitation of Workload Manager (WLM) within z/OS. Rule 4 relates to zIIPs and zAAPs, (specialty offload processors) which represent significant price/performance benefits for deploying on z/OS. If the application workload represented exploits these processors, System z often delivers a lower implementation cost, as proven by the testing described in this document.

The focus of the prescriptive use cases is on quantifying the second and fourth rules (co-location, and the leveraging of specialty processors) so that when a new workload is placed on the same system as the existing workloads, it provides a distinct and cost-justifiable benefit.

There are numerous value points for deploying on z/OS, one of which is co-location – a key consideration because it can be quantified in terms of throughput, CPU consumption, and (most visibly) cost. The test described in this paper proved the hypothesis of whether a new WebSphere workload can be placed on the same physical configuration, running concurrently alongside existing online and batch workloads, at a comparable price point to a separate, distributed system that “talks to the z/OS back end.”

Description of the Test and Workload

This section explains the elements comprising the test and why this effort is different from numerous other benchmarks that were conducted by IBM and others.

Test Overview

The comparison included:

- **A single scenario**: A "mixed workload" based on the WPS
- **A single proof point**: Business Services per hour (based on Business Services per second). Note that this measurement is described in detail, beginning on page 9 of this document, in the “Proof Point” section.
- **A single topology:** “All” z/OS as compared to a heterogeneous AIX and z/OS configuration

Single Scenario – Mixed Workload

A WPS workload was chosen for this comparison for two reasons:

- A set of configurations and workloads was available for use within IBM.
- Business Process Management is a strategic and growing investment area for many companies.

A mixed, or co-located workload (running other things besides WPS, such as batch and online) was chosen. The workload is a combination of real workloads that have been modeled from some of IBM’s largest WPS on z/OS client test environments and workloads.

Compared to a typical benchmark, which simply involves a single transaction or single application run to high stress levels or transaction rates, this effort involved a mixed application representative of a real-world scenario, in which, different styles of transactions (more than just Java™ transactions) were blended.

About the ERWW Application Used in This Effort

The ERWW production workload is a well-known, cross-industry application centered on the activity of processing orders. It is based on the Transaction Processing Performance Council (TPC-C) specification, developed by numerous organizations. (More information about the TPC-C specification is available at www.tpc.org.) The ERWW production workload is enabled for WAS z/OS. It mimics actual existing IMS, CICS, and DB2 systems. This application was used to create a “composite application” using SOA principles with WPS for z/OS V6.2.0.1.

The ERWW Order Processing System serves as the principle “customer-like” application and workload for validating new releases of WAS z/OS. It is also used cooperatively by many IBM internal development teams to drive z/OS workloads and benchmarks on System z hardware.

The ERWW workload is a typical example of an IBM customer’s mission-critical batch online processes. What is important about the workload is:

- The variation of the Business Services
- The fact that Business Services are run concurrently, from multiple “users”
- The ability to drive the workload efficiently at multiple “Business Services per hour” rates

What are the activities of the ERWW workload?

The ERWW application includes typical activities that most companies conduct, such as:

- Entering a new order
- Processing payments
- Checking the status of the order
- Delivering products or services
- Checking the stock level
- Changing the price of an item
- Giving a price quote
- Performing a customer inquiry

Details of the ERWW Order Processing Core Workload on z/OS

The ERWW Production workload was used to create a “composite application” using SOA principles with WPS for z/OS V6.2.0.1.

Figure 1 illustrates the business logic that is contained in the session beans (purple circles) and the entity beans that encapsulate the data (aqua circles). The back-end database is DB2 for z/OS.

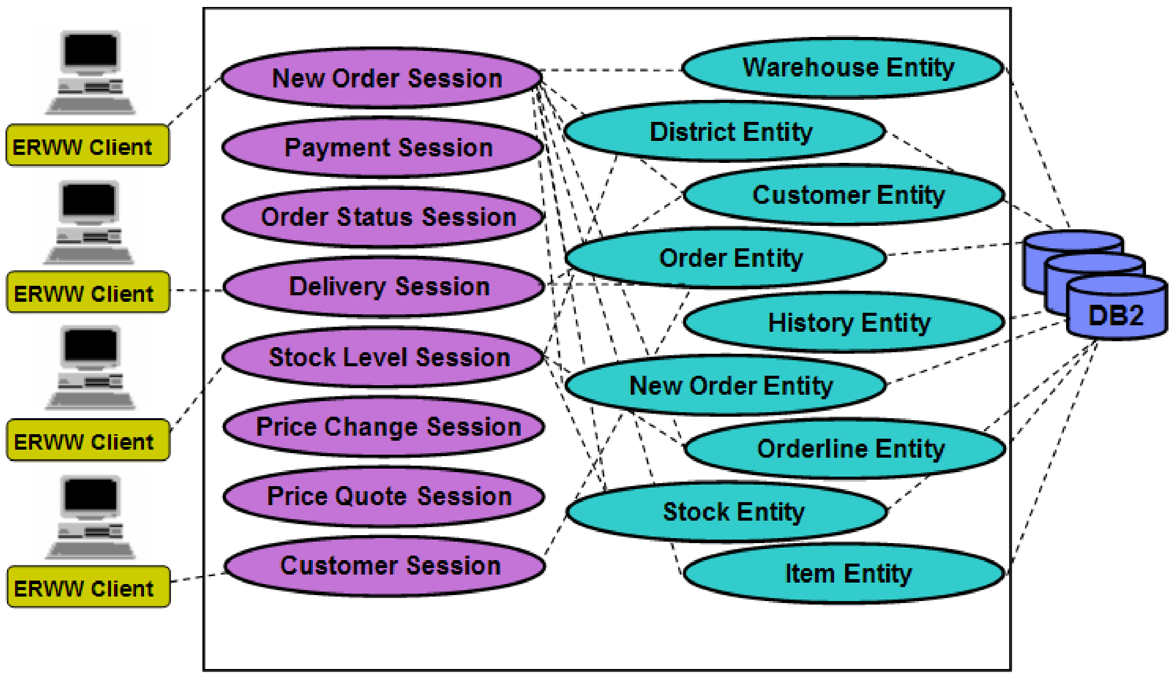

Figure 1 - ERWW J2EE Component Interaction

Proof Point

Business Services per hour was selected as the proof point for this test, with the approach of normalizing the number of Business Services per hour driven from the z/OS configuration, so that an equivalent number of Business Services per hour on a distributed system could be compared.

In this context, a **Business Service is defined as a collection of related, structured activities that together produce a service or product that meets the needs of a client**. These activities are critical to virtually any organization because they can be used to generate revenue and they often represent a significant percentage of a company's operating costs. Business Services represent meaningful business-critical transactions to virtually all of IBM's customers.

Customers can typically relate to these types of Business Services much better than a single business transaction that involves, for example, a back-end system locating a number in a table. This is even more relevant when the number of Business Services is increased exponentially, as in the case of a large bank that runs thousands of processes and acquires additional banks that must run the same processes – without the disruption of service.

Examples of the Business Services That Were Used in This Effort

The following list briefly describes a few of the ERWW transactions and the services that are invoked.

- **Enter a New Order** - A "NewOrder" request message is received and the request is put on a HumanTask work list for processing. When the work list item is selected for processing, it is either "accepted," after which the "NewOrder" SCA service is invoked and a "NewOrder" is entered into the ERWW system, or it is "rejected" and the process bypasses the "NewOrder" SCA service and terminates.
- **Give a Price Quote (or Price Change) -** The user (simulator) enters an item number and quantity into a list. The application generates a price quote for the information entered by invoking the "PriceQuote" transaction. From the price quote reply, the application selects the first item in the list and changes the price for that item by invoking the "PriceChange" transaction. The test case executes a read and update transaction from a process.
- **Process Payments (Credit/Debit) -** The user (simulator) enters funds transfer information into the Web application screen. The application executes a credit transaction and a debit transaction from two separate accounts. This test case executes read and update transactions from a process.

How the Business Service Measurement Differs From Typical Benchmarks

This prescriptive use cases effort was *not* a performance benchmark, that is, it did not involve running the machine at a high utilization to arrive at a performance measurement. Instead, the test was centered on running steady state workloads in the system as a typical customer would do. By design, we kept the disparity between the systems and the workloads to a minimum, that is, the amount of memory available, the types and speeds of the processors, and the actual Business Services executed were relatively consistent across the three measured runs. By doing so, it provided us with a sound and easy-to-understand foundation for making a cross-platform comparison.

Industry standard benchmarks are typically difficult to correlate to "real-world" workloads because they represent a unit of work that is standardized. Measurements against such benchmarks are hard (or impossible) to relate to what a large installation with mission-critical data actually does.

Topology: What This Effort Compared

A single topology, that is, a simple z/OS structure, was compared to a heterogeneous-based AIX and z/OS topology. In this case, the term, “simple,” describes a single z/OS system, with no data sharing, as compared to a distributed system with AIX. This configuration was chosen because it is the comparison that is probably required in 75 percent to 80 percent of engagements.

The chart in Figure 2 illustrates the simple topology for the price/performance comparison.

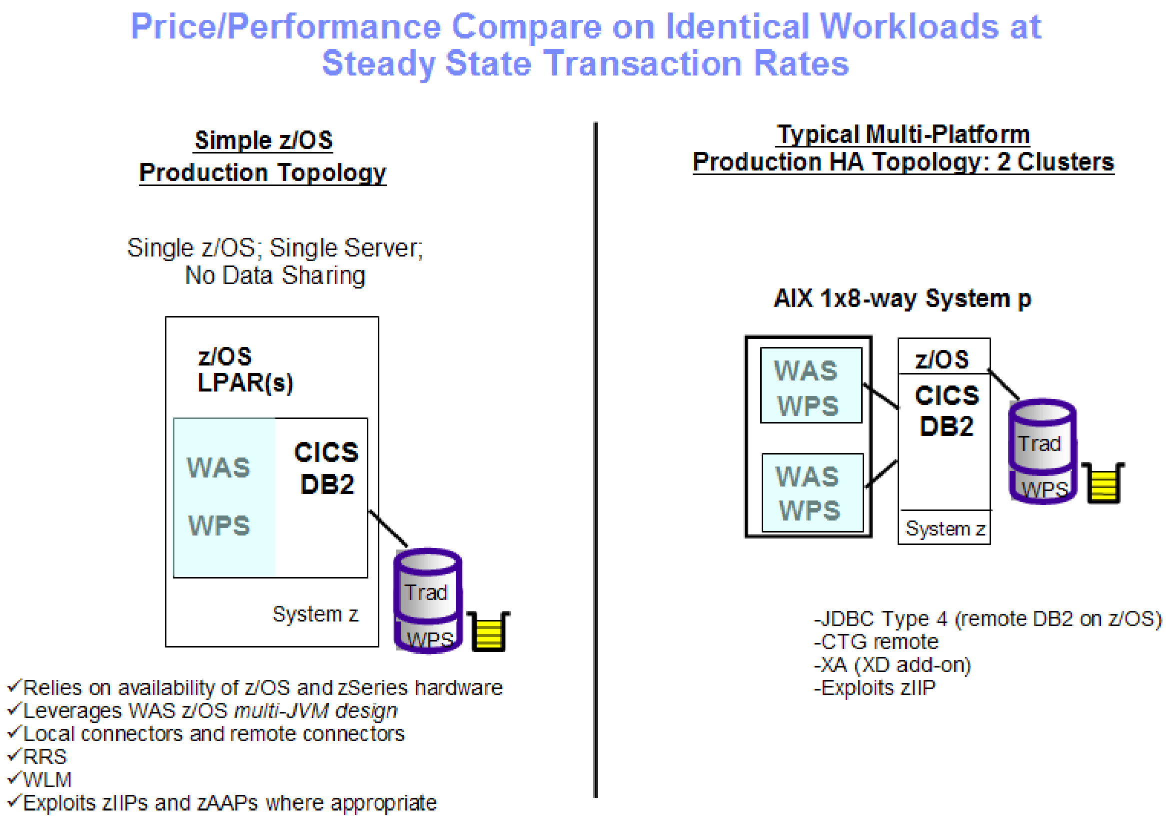

Figure 2 – Topology for Price/Performance Comparison

The recommended configuration is on the left. Using the previously outlined “Rule of Four,” this simple z/OS configuration aligns well. It includes CICS and DB2 running traditional online and batch workloads. The new WebSphere workload (WPS) was added in this configuration (co-located on z/OS).

The right side of the image represents a compilation of the heterogeneous platform, or in our use cases, AIX and System p running duplex, that is, two copies of WAS and WPS. The lines can be viewed as connection points. On the right are the traditional applications, CICS and DB2, that house the key business data for the traditional (batch and online) and for the new WebSphere workload. The new workload communicates through the connection points to the existing online and batch environment, and uses the DB2 data on z/OS.

The following items were configured, "tested," "priced out," and compared in the selected topologies:

- Business Services rates and "cost" for adding a WPS workload to the "traditional z/OS" ("co-located"), compared to
- The cost of attaining similar Business Service rates for the same WPS workload running on a single AIX "box" communicating with the DB2 running on the "traditional" z/OS

The z/OS topology and the configuration of the test environment are very much like those deployed at many companies, and are the same topologies that that we often recommend. This is because they align with the co-location benefits. They have the fewest moving parts and the fewest things that might fail. They also enable us to configure other LPARs as clones to have very easy failover scenarios. One of the big strengths of the z/OS topology we chose for this test is that it is already configured for scalability and high availability of the existing mission-critical batch and online workloads.

Execution of the Workloads

The execution validation effort included measured and "priced" deployment of the co-located z/OS topology and contrasted it to the distributed AIX topology:

- **Co-located topology:** Runs to determine Business Services per hour at multiple utilization rates with the new workload running in a single z/OS image
- **Distributed topology**: Runs at the same utilization rates as the co-located topology to attain similar throughput
- **"Pricing out"** of the above topologies and assessment of their cost comparison

Results and Impact of the Validation

As the following chart in Figure 3 shows, the 3-year hardware and software costs are comparable for the data points measured.

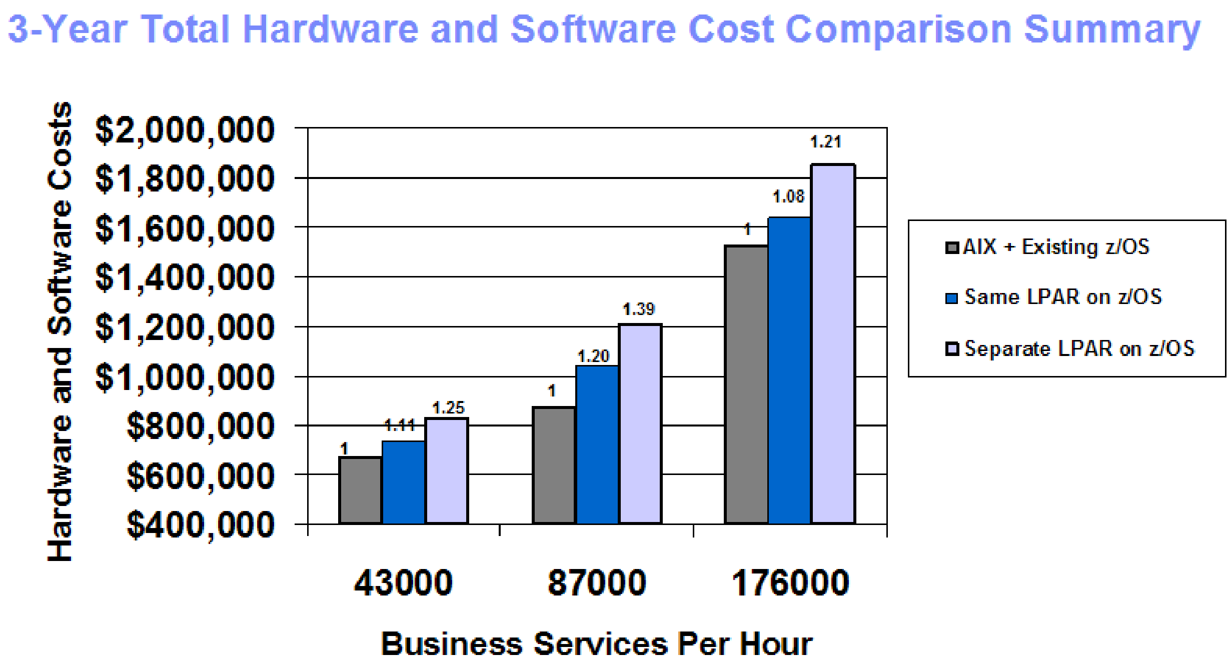

Figure 3 – Summary: 3-year software and hardware costs for deployment on z/OS are competitive.

Note: Each of the bars in the above chart includes:

- **Hardware costs,** such as GPs, zIIPs, zAAPs, memory, and maintenance
- **Software costs,** such as WPS or distributed operating system licensing, 3-year WPS or distributed operating system support and subscription, the cost of the z/OS operating system, and additional licensing costs for other middleware (for example, CICS and DB2)

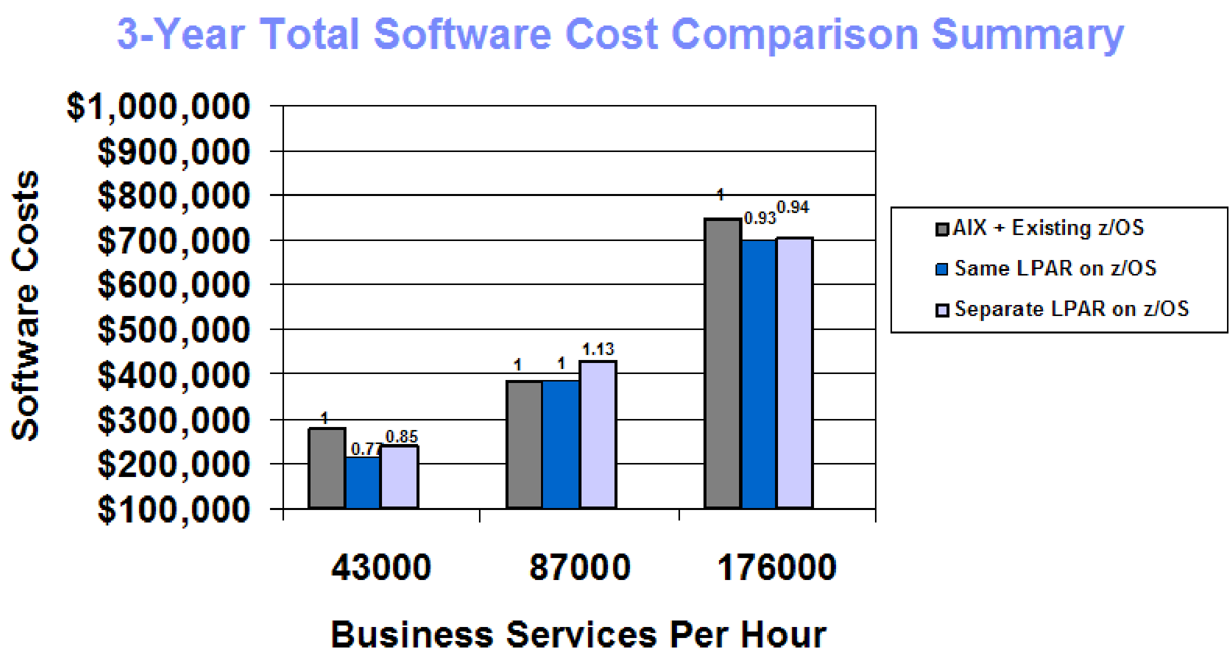

Figure 4 – Summary: 3-year software costs for deployment on z/OS are competitive.

These results refute the perception that software always costs less when deployed on a distributed system as compared to being deployed in a running production topology deployed on z/OS.

Impact of the Results

The reason why the comparison of the price/performance between the workloads on z/OS and AIX is important is that, with tests based on real-world customer scenarios (not on optimized benchmarks designed to highlight an individual performance aspect of the system), the findings showed the z/OS configuration is capable of handling a significantly large number of concurrent Business Services (such as “open new account”) per hour – at a comparable cost to distributed -- while also providing z/OS non-functional value, for example, high availability, security, and scalability. (There is additional discussion about this subject in the next section of this white paper.)

The performance of the machines is relatively comparable, that is, there was no major disparity for the type of workload and the amount of work put through the system, and there were many congruencies. The response time rates and transaction volumes were consistent. As in the real world, a determined amount of work is done with a certain set of qualities and, from a cost point of view, the results are neutral regardless of whether the new workload is in the same LPAR as the existing online and batch workloads, or in a separate LPAR.

Given the neutrality of the price/performance results, most customers can now base their deployment topology or configuration decision on other factors such as non-functional application requirements, for example, QoS, disaster recovery, and so on.

Context

As the results show, price/performance in topologies like the ones tested in this effort should no longer be a significant factor in deciding where to run new workloads. Why, then, would an organization choose to place a new workload on z/OS instead of a distributed system?

First, note that **WebSphere Process Server** (**WPS) and WebSphere Application Server (WAS) are cross-platform offerings, and WebSphere is the same "*above* the specification interface line."** It is important to note that the WAS/WPS differentiation is *not* in the open standard specification support offered (which is common across platforms).

Characteristics of Workloads Based on Open Standards

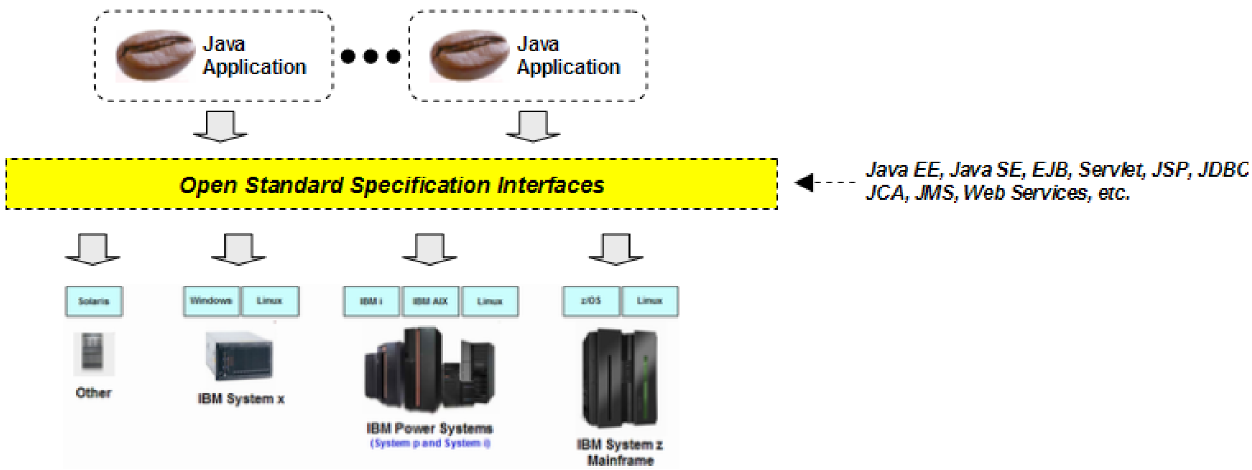

Figure 5 – Differentiation is not in the open standard specification support offered.

Alignment across platforms offers some key benefits:

- It provides a **common application environment** across the enterprise.
- It provides the **ability to promote applications "up the ladder"** without causing concern about the loss of interface function.
- It provides the **ability to architect application designs that span multiple platforms** without having to make sacrifices based on the platform.
- It provides the **ability to settle on a common set of application tools** across virtually all platforms.
- It provides the **ability to have an essentially common management interface** across virtually all platforms.

Differentiation Value for Deploying Java Workloads on z/OS

WPS/WAS z/OS differentiation and customer value occurs *below* the open specification line. How it is implemented (its features, functions, attributes, and QoS) is dependent on the platform. Some notable examples of WPS/WAS active exploitation of System z and z/OS are:

- **Workload Manager (WLM) and the Controller/Servant architecture** for intelligent dynamic capacity expansion, flow control, and routing provide ***availability and scalability.*** No other WebSphere platform has this design.
- **Resource Recovery Services (RRS) and two-phase commit processing** offer ***reliability***. Distributed implementations can implement XA, but it is less efficient than RRS in a parallel sysplex environment. RRS is ready to exploit from the start.
- **EAL4+ Certification** through logical partitions (LPAR), cryptography for Clear Key and Secure Key, and z/OS Security Server, including RACF provide ***multi-level security.***
- **WAS z/OS V7,** announced and available since 3Q 2008, introduced features (such as Fast Response Cache Accelerator [FRCA]) that offer significantly improved ***performance***, and enhancements (such as thread hang recovery) for ***availability***. These features are offered *only* in WebSphere for z/OS. For workloads that can benefit from FRCA, improvements of 40-70 percent have been measured.

The Operational Value of “Co-location” of New Workloads on z/OS

Workload deployment on z/OS offers applications the operational benefits of co-location, that is, the ability to co-locate the application layer with the data layer in the same z/OS operating system instance. This leverages many of the inherent advantages of the System z hardware and z/OS operating system that can be exploited to business advantage in the areas of performance, efficiency, and operational benefits, specifically:

- Reduced security overhead (eliminates the need for encryption and SSL)
- Easier workload management
- Enablement of multiple security options
- Reduced complexity with fewer points of failure, fewer monitoring points, and reduced overall time to resolution

The three major advantages stated above (differentiation coupled with single-tier operational benefits, such as disaster recovery, scalability, and high availability, along with the ground-breaking and now proven co-location price/performance benchmarks) give z/OS a distinct competitive advantage in the selection of where to run new workloads.

Additional information about the technical advantages and the resulting business value of WAS on z/OS is available in the *Why WebSphere Application Server for z/OS?* white paper (available on ibm.com/support/techdocs).

An Example

A major bank provides an example of an IBM customer who uses WAS for z/OS in a large implementation to provide support for multiple delivery channels and to reduce IT maintenance and support costs associated with duplicate business logic. The bank currently has over 27 production WAS for z/OS business applications, including online banking, mortgage, and brokerage services.

Some benefits include:

- Less than 100 ms response time with 350 Transactions per Second (TPS) that scales up to 1,000 during peak business
- EAL Security Certification Level 4
- Reuse of existing CICS and IMS assets
- Lowest unit cost in the enterprise (lowered costs by 63 percent)
- 99.997 percent availability since inception, with zero unplanned outages
- Integration with enterprise data and transactions - allows comprehensive utilization of resources, from hardware to IT personnel, through the use of open standards and integration with System z processes and procedures
- J2C tight coupling to back-end transaction subsystems
- Single point of security, logging, and monitoring
- System-level dynamic work load management provided by WLM
- Common transaction coordination across subsystems (CICS, IMS, DB2, MQ, WAS), and enhanced transaction integrity provided by RRS
- Support for next generation technologies - JEE compliant with support for key Web services, network deployment consistent with the distributed platform

Summary

For those who are already running workloads on z/OS, there is now proof that the price/performance of running new WebSphere workloads on z/OS – co-located with existing online and batch workloads and their associated data -- is comparable to running the new WebSphere workloads on a distributed platform. Therefore, companies can realize the additional benefits provided by z/OS and co-location (disaster recovery, scalability, high availability, and others) without incurring significant additional costs for these benefits.

Contact your local IBM WebSphere Sales representative for additional information or questions you might have after reading this document.