



HRG Assessment: Enhancing Business Agility - SOA, WebSphere and iDataPlex

The current generation of Web Services known as Web 2.0 facilitate collaboration and sharing among people and communities on-line. This on-line collaboration results in the creation of new applications, one example is a process known as mash up. Mash up refers to creating new browser based applications that provide user experiences that more closely resemble current generation desk top applications than previous generation static web page content. This new generation of Web 2.0 service based applications can have significant potential impact on the performance and capacity attributes of the web facing servers that underlie today's SOA solutions. These solutions are in large part focused on delivering data, information, and content to web based customers. While Web 2.0 developed applications are essentially free the compute infrastructure of those organizations whose data, content, and application logic are used by those applications is not without cost. There are real and potential business benefits to be derived from supporting Web 2.0 however it is incumbent on organizations making this choice that they plan their infrastructure accordingly. IBM's WebSphere Extended Deployment (XD), WebSphere MQ, and IBM System x iDataPlex are well positioned to help today's internet focused businesses stand up to this challenge.

Service-Oriented Architectures (SOAs) are receiving a great deal of attention as enablers of collaboration and business agility. Today, many companies are considering or implementing SOA, but are also dealing with IT complexity in the forms of multiple technologies, languages, and infrastructures that may or may not be standards-based. In addition, heterogeneous mixtures of production environments can make even the most straightforward application and data integration tasks daunting and unwieldy. IT success today is increasingly tied to enabling organizations to fully leverage all of their key IT assets for competitive advantage.

IBM helps its customers address these challenges with products designed to extend the functionality of a typical application server environment. These products can help users create messaging-centric, cross-environment implementations with high levels of performance, reliability, and security. They also support the virtualization of both workloads and information assets to allow more effective use of existing systems and networks. The result is lower costs and increased flexibility to meet continuously changing resource demands. Perhaps most importantly, all of these products support "mixed server" environments, with support for a wide variety of application servers and other deployment options.

Today with increased pressure to drive cost out of business/IT operations, run greener ecology friendly operations, and seamlessly scale capacity to meet Internet speed customer requirements CXOs increasingly look to architectures based on SOA and Web 2.0 best practices for solutions. WebSphere in conjunction with the recent introduction of IBM's iDataPlex offers a combined hardware, software, and services solution that is eco friendly, flexible, responsive, and positioned to address the evolving requirements of Internet Scale Data Centers.

Business Agility

A variety of factors have empowered organizations to leverage their IT assets more aggressively to gain and maintain competitive advantage. IT resources are now tied more directly than ever before to the bottom line performance of an enterprise. However, complexity and heterogeneity have increased, business environments have grown more competitive, and the end of the internet “boom” has forced organizations to put greater emphasis once again on traditional business metrics and modes of accountability. Intensifying cost control pressures fly in the face of the growing resource levels required to support IT’s greater role in business decisions and operations.

Service-oriented architecture (SOA) is gaining strong customer traction as the architectural context for building flexible business systems. The standards and methodologies established via the Web services model provide a basis for more effective integration of disparate systems. Yet today, many organizations are by necessity focused on dealing with a mixture of heterogeneous servers and deployment environments across which applications must interoperate and communicate in order to effectively support business objectives.

The high level IT challenges for organizations today are to achieve the levels of agility, flexibility, and energy efficiency required to meet committed business goals in a secure and cost-effective manner. While approaches that address these overall challenges may vary across organizations, a number of common requirements and metrics for success apply:

- ***Providing support for mixed server environments.*** Customers need to support all of their IT assets to meet their business goals. They may need to run applications across a wide array of systems and application servers. Merger/acquisition activities, as well as internal consolidation of IT assets, often serve to intensify this requirement. Customers must therefore place deployment platform independence high on their lists of requirements.
- ***Building a flexible infrastructure that can easily and quickly adapt to changing business needs.*** This infrastructure must allow for a more cost-effective and efficient utilization of existing resources to keep costs and new expenditures low. Perhaps more importantly, customer and business partner satisfaction levels depend on the ability to allocate compute power where and when it is needed to meet dynamic utilization demands with sufficient performance. Virtualization of a variety of IT asset types is a key enabler of this flexibility, including the ability to virtualize assets across a variety of deployment environments.
- ***Meeting quality of service (QoS) levels for transaction-based systems.*** Agile businesses require support for handling large and often unpredictable transaction loads in order to meet service level requirements of customers and business partners. Standards-based messaging systems are a viable solution for increasing interoperability across disparate systems and providing persistence mechanisms to ensure transactional integrity and reliability.
- ***Providing Robust Security.*** High volumes of external transactions serve to accentuate the security challenges already facing IT managers today. Organizations need robust intrusion protection to prevent the exposure of sensitive information through improper access or other means.
- ***Providing capacity on demand.*** Customer-facing web portal based applications and support for SOA and Web 2.0 requirements are driving enterprises to compete based on their ability to scale out in order to meet customer SLA and QoS expectations. Increasingly customers whose expectations are not met will seek out alternative providers who have better mastered the challenge of delivering on the promise of customer satisfaction in the face of extreme internet driven scale-out requirements.
- ***Being Green.*** Today there is a real value being placed on energy efficiency or “being green” with energy becoming more costly and concerns over climate change on the rise companies and organizations of all sizes are focusing on reducing their energy consumption footprint. Now whether planning new data centers or

refitting existing data centers CXOs pay keen attention to environmental impacts as well as energy and real-estate costs. Today being green is synonymous with being efficient operationally and from a cost perspective.

While many vendors offer products designed to address some or all of the above requirements, the specific messaging and cross-environment challenges inherent to these requirements necessitate a coordinated and focused approach. IBM is unique in its breadth and depth of solutions specifically tailored to meet these challenges.

WebSphere

IBM has recognized that in order to more effectively address the requirements outlined above, it needed to provide products that address QoS, flexibility, and security issues. IBM has established itself as a leading provider of flexible application infrastructures. The foundation of this leadership, IBM's WebSphere software platform, provides Application and transaction infrastructure, Application integration, Application transformation, Business process management, Commerce, Mobil and Speech middleware, Portals, and Express Middleware. Through its SOA and "Smart SOA" related announcements made at their April 2008 IMPACT customer convention IBM has clearly restated and demonstrated its commitment to SOA and enhancing the comprehensive nature of its infrastructure solutions. IBM has also made these products available on non-WebSphere platforms, making it possible for those organizations that use application servers from a wide range of vendors and sources (or who deploy without application servers) to leverage their capabilities. In doing so, IBM is offering its customers the ability to extend their existing application server environments to more effectively meet the service level and interoperability challenges they face.

WebSphere Extended Deployment (XD)

It is very often difficult to meet QoS requirements when the dynamic nature of transaction levels and interactions with various parties make workloads unpredictable. Upgrading production applications can also degrade service levels when systems must be brought down. It can be expensive and time-consuming to maintain non-production environments for staging new application versions. Finally, network bottlenecks can also cause poor service levels; the bottlenecks arise when applications repeatedly access data and information stores.

Virtualization is often utilized to dynamically allocate servers and storage platforms in order to apply these compute resources where and when needed. However, applying virtualization requires that issues relating to potential loss of service be directly addressed. As an example, sensing and responding to hardware platform problems in order to drive virtualization activities is an important way to maintain QoS levels. Doing so means monitoring all systems and applying policies reflective of QoS requirements. These policies need to either manually or automatically take action to prevent service outages.

A potential way to better address QoS issues is to extend the virtualization concept to include both applications and information sources. Any solution designed to do so, however, must incorporate the ability to virtualize assets across a heterogeneous collection of platforms on which they may be deployed. This "mixed server" approach ensures that all deployment scenarios (including applications servers) can be extended as part of a virtualization plan that optimizes asset usage and performance.

IBM's WebSphere Extended Deployment (XD) is designed to help users implement effective workload and service virtualization that complements hardware virtualization. WebSphere XD's dynamic response to changes in complex workloads reallocates virtual resources, allowing organizations to be more flexible and therefore to more consistently meet required service levels. Complex workloads that can most effectively leverage WebSphere XD's virtualization capabilities are those which contain multiple applications whose peak periods of usage can be substantially different, or those that employ large monolithic applications utilizing resources across many servers in order to address high transaction volumes.

WebSphere XD can also provide a centralized point for workload management and operational control. The service and health policies of all application servers can be managed by the WebSphere XD environment in which they operate. This capability enables the management of service and health policies for all application servers included in the environment served by WebSphere XD. For instance, WebSphere XD could be used to front-end a server farm to which it distributes work requests to application servers (WebSphere and non-WebSphere). Applications running on non-WebSphere application servers can run as-is and would not need to be migrated to a WebSphere deployment environment. In addition, a feature of WebSphere XD would run on non-WebSphere application servers that would communicate workload status back to WebSphere XD as input to service and health policy decisions. In this way, an existing, heterogeneous application server environment can be extended in ways that permit more effective workload management.

WebSphere XD is therefore focused on the virtualization of an organization's software infrastructure with the goal of consistently meeting service levels, regardless of how activity levels may fluctuate. Two types of virtualization are possible with WebSphere XD:

1. *Workload Virtualization*. WebSphere XD can dynamically allocate and manage a pool of application server resources based on the types and volume of workloads using 3 different approaches.

Service Level Management - Users can create policies that represent service level goals. These policies are then used by WebSphere XD to classify and prioritize workloads and route them intelligently to the proper resources.

Application Edition Management – This extends the notion of application management to include multiple versions of applications and services. Deploying new application versions in traditional environments can cause loss of service. While administrators can get around this by creating duplicate staging systems and custom routing algorithms to transition these new versions into production, this solution is not ideal. WebSphere XD's Application Edition Manager enables different application editions to coordinate their request routing and other activities. In this way, multiple deployments of a given application can be supported, thus making it possible to keep multiple versions in production at the same time. Requests can be allocated to different versions based on pre-defined routing rules. Updates can be rolled out seamlessly. It is also possible to orchestrate reversion to older versions when necessary.

Health Management – This permits administrators to identify and mitigate server health problems before they cause production outages. Health policies are defined and conditions monitored, and problems can be handled in three ways. Actions in response to problems can take place “manually”, in “supervised mode” and “autonomically”.

- *Manual mode* - An operator is informed of the problem via an alert that identifies which application server needs to be restarted.
- *Supervised mode* - WebSphere XD recommends corrective action and asks an operator for approval to take the action.
- *Autonomic mode* - WebSphere XD informs an operator it has taken a specific action after the fact.

2. *Information Virtualization*. This unique capability recognizes that even when application resource allocations are optimized, access bottlenecks to shared information resources can severely limit application performance. One way to alleviate the effects of this problem is to persist application sessions in order to limit the need to access information sources. WebSphere XD includes a feature called **ObjectGrid**, a distributed caching framework for storing and sharing Java object data by multiple applications. Each server can have its own ObjectGrid, thus making the framework highly scalable. Because many or all servers in an ObjectGrid can cache objects redundantly, the effects of a specific server's failure on service levels can be minimized or eliminated. ObjectGrid can be used for HTTP session management, allowing the storage of short-lived session data within the application tier. WebSphere XD also includes a **Partitioning Facility**, which permits the design of

applications that can separate logic and data into partitions. These partitions can then be mapped to available servers. Users can utilize “Partition-aware routing” to route requests to the right servers, allowing information to be cached much more efficiently. This approach can result in performance improvements and optimized scalability as the transaction load increases because server resources are used more efficiently.

WebSphere XD’s infrastructure virtualization capabilities are designed to enable service levels to be met consistently and predictably. This is accomplished by permitting the fast adaptation of application server resources (including non-WebSphere application servers) to changing workload demands, reducing deployment complexity, and relieving backend data store access loads. In addition, as a result of the service level and health management features as well as the distributed and redundant information caching capabilities of WebSphere XD, scaling of the application environment can be highly efficient.

WebSphere XD provides broad service and health policy support for virtualized workload management and, customizable health policy and action features, to non-WebSphere servers. These servers include (but are not limited to) J2EE-based and .NET based application servers, PHP Server, Apache Tomcat, Ruby on Rails, JBoss Application Server, and BEA WebLogic Server.

WebSphere MQ

Effective and reliable messaging is a critical requirement to support the transaction-based messaging on which many businesses depend. Messaging systems, by design, focus on effective data and information exchange. A couple of factors have combined to intensify the challenges associated with this process. One is the growing requirement to create interoperability among disparate applications that support critical business processes. The second is the increasing emphasis on systems that meet more rigorous compliance requirements. While messaging standards can help in this effort, the application of such standards must first take into account the potential variety of languages and technologies that already exist within an organization’s application portfolios.

WebSphere MQ has long been IBM’s messaging infrastructure to support the effective exchange of data and information among disparate systems and platforms. IBM continues to work to make WebSphere MQ a versatile “messaging backbone”. The goal is to enable organizations to focus on their core competencies and business goals and not on modifying their applications to accommodate messaging requirements. To illustrate this, it is helpful to note that in today’s programming environments, where the use of Java for application programming is widespread, the belief may exist that messaging needs must be met through adherence to the Java Message Service (JMS) standard. However, while JMS is probably the most widely used API for messaging, a distinction must be drawn between messaging APIs and the messaging engine that implements them. While a typical JMS engine is adequate to implement messages based on this standard and to achieve required quality of service levels, such an engine is not sufficient in the following cases:

1. Where non-Java applications exist
2. Where other languages and programming styles are preferred
3. Where different quality of service levels may be required by differing application environments.

For this purpose, and to complement the native programming interface for WebSphere MQ (known as MQI), IBM has created **IBM Message Service Clients** (also known as **XMS**), a set of messaging clients and APIs that permit the rendering of JMS in a variety of different languages (C, C++, C#, Visual Basic, .NET, in addition to Java). Consistent with IBM’s focus on enabling its products on non-IBM platforms, XMS permits users who have standardized on other technologies, such as .NET, to leverage WebSphere MQ messaging functionality or to integrate non-Java applications with J2EE applications and services without needing to learn specific WebSphere MQ APIs. WebSphere MQ’s support of XMS and JMS in addition to MQI make it attractive for situations where applications must run across a variety of deployment environments (such as .NET, SAP, Oracle, and any combination of J2EE application servers).

WebSphere addresses the issue of transactional reliability through a mechanism that ensures that all operations associated with a transaction are completed as an integral part of delivering a message to its target system. Very often,

transactions will involve a single unit of work that must include updates to multiple resources and operations (which in turn may be tied to “suboperations” that need to be executed as well). WebSphere MQ has been designed to support transactions “natively”, i.e., a message can be defined as a transaction that includes all operations and updates to resources such as databases. In this way, all operations associated with a transaction are completed before a message is finally committed on the target system. This capability is provided through a combination of the control provided by WebSphere MQ’s built-in transaction manager, as well as the asynchronous nature of the messaging environment itself. Messages are routed between queues by the product’s Queue Manager. Many Queue Managers are typically deployed across an organization, which could then be linked together, if required, to form a single “virtual” Queue Manager. This “clustering” of Queue Managers can increase overall reliability and contribute to the scalability of a server which must handle large numbers of messages.

Writing messages to non-volatile storage such as hard drives – “persisting” them - provides a means of ensuring that messages are not lost in case systems fail with transactions only partially completed. While a vast majority of messages do not have to be persisted (because they are transient and therefore non-critical in nature), in some cases it is not appropriate for a query to be reissued. In other cases, the message may involve a specific update or key information which must not be lost. To satisfy this need, WebSphere MQ is capable of writing messages out to a queue persistently (via hard disk storage). It is important to note that IBM distinguishes the different methods by which WebSphere MQ persists messages from other approaches available today. For example, the JMS specification permits implementations to perform what are called “lazy writes”. These involve the persisting of a message to memory before letting the message go, without confirming first that the message had been persisted to disk. While this can provide a net performance boost over a more robust persistence approach, it introduces risk if the system fails. WebSphere MQ allows for this mechanism to be selected, but in those situations where message and transaction reliability are non-negotiable (such as with many financial service transactions), full, disk-based persistence is required, and is natively supported by WebSphere MQ.

The requirement to keep messages secure is also becoming more important over time as organizations use SOAs and other designs to leverage their IT assets for use beyond their office walls. WebSphere MQ has always been able to encrypt messages for transmission to another Queue Manager. This is important, because when a message is sent to a destination system, it sits on a queue before being read by an application and must therefore be protected from unauthorized access. Today, an edition of the product, WebSphere MQ Extended Security Edition, combines WebSphere MQ and Tivoli Access Manager for Business Integration into a single package. This combination allows messages to be encrypted end-to-end between applications, rather than just between Queue Managers. This adds greater security and privacy for messages throughout the business. Additional features include centralized, browser-based, remote security policy administration, and the use of a message-level audit function with records that can support compliance efforts.

The ability for WebSphere MQ to work with a variety of IT infrastructures cannot be overstated. In fact, one of the major benefits of WebSphere MQ is that it can support environments that include multiple application server brands as well as other internally-developed or commercially-available deployment platforms. While many of these have their own messaging infrastructures, WebSphere MQ can provide the cross-environment functionality required to enable messaging and data sharing among, and independent of, all systems and platforms. WebSphere MQ supports over 80 platform configurations, and is not dependent on any one language or technology in order to create or execute messaging calls from any application.

IBM System x iDataPlex

IBM System x iDataPlex is IBM's entry to the massive scale out computing market segment. This new highly scalable (scale-out) rack mounted system is initially targeted at Web 2.0, Video Rendering, Financial Modeling, High Performance Computing and other deployment scenarios requiring extreme scale-out. Media companies that are streaming content over the web in support of cable TV or internet TV have similar real time and scale-out requirements. The IP TV phenomenon is picking up. These customers want the lowest cost most efficient hardware and depend on software, not hardware, for resiliency.

Requirements - massively scalable solutions

In the summer of 2006 IBM made a strategic decision to focus on developing and delivering a solution specifically targeted at evolving Internet scale data centers. An internet scale data center is a data center with 1000s of servers in a single data center with massive scale. The challenges that these Web customers are facing related to power, cooling, density, and scale, are very similar to the challenges that enterprise customers are facing. These customers have deployed their availability or redundancy of their entire data center or the applications they run somewhere in the network layer or in their software and so they do not depend on hardware redundancy for availability. These are web 2.0 companies where they have always bought the cheapest servers available or have their servers assembled themselves. Typical Web 2.0 customers do not rely on hardware for availability, and merely discard a non working server and plug in a new one.

WebSphere and iDataPlex

IBM's WebSphere Extended Deployment (XD), WebSphere MQ, and WebSphere Portal are examples of a few software solutions that are a good fit for iDataPlex as are other applications where the redundancy and availability requirements are met within the software layer rather than through more traditional highly available systems. Companies like Google, Yahoo, and other Internet super stars require scale out capacity on demand in an energy- and cost-efficient configuration. iDataPlex has been specifically architected and engineered by IBM to meet the unique requirements of the Internet Scale Data Centers that comprise this market segment.

iDataPlex is a practical platform for Web 2.0 collaborative applications as well as more traditional Business Intelligence, Operational Business Intelligence, and Risk Analytics Web centric SOA based applications. HRG recommends that customers evaluate iDataPlex as part of their overall datacenter mix to run Web based, Grid (compute grid / storage grid), or HPC based applications on. Using Web 2.0 mash-ups customers will be able to further manipulate and analyze the data that is made available through the iDataPlex resident Portal and underlying software. This puts demands on the processing horsepower that is available through the Web Sphere portal and also puts demands on the underlying SOA based applications and functionality.

SOAs implemented through IBM's WebSphere Extended Deployment, and WebSphere MQ in combination with selected offerings from Tivoli, Rational, and Lotus running iDataPlex provide a means to effectively integrate an enterprise across vertical applications and divisional stacks, and provide highly scalable access (utilizing iDataPlex as the access point platform) to web based communities.

In future it is unlikely that data centers based on current technology will be able to meet the evolving demand for compute power as current generation data center technologies will not scale-out to meet next generation requirements due to current constraints on power, cooling, and available floor space. Today's datacenters will have to be completely redesigned in order to meet customer demand and government regulatory requirements regarding cooling, power consumption, performance/throughput, and security. This realization was one of the key motivations for IBM to develop iDataPlex. In designing iDataPlex IBM focused on creating the ideal server and rack to address future data center requirements:

- highly efficient servers that from the perspective of power
- highly efficient racks from the perspective of cooling
- more compute power within the same or less space = greater compute density.

As a direct result of Web 2.0 companies supporting social networking applications, search engines, and on-line auctions they will need to redesign their datacenters. With the impact of web 2.0 driving rich media and interpersonal interactions over the web (Social networking) this will drive a demand for extreme scale-out in support of continually ramping user demands for processing power, storage to support user requirements, and SOA based data integration.

Next Generation Data Center

iDataPlex is designed for the "next generation" data center, where issues related to power and cooling efficiencies, acquisition and maintenance costs, and manageability of massive scale-out environments are and will continue to be key concerns.

Looking Forward

The requirements to meet and exceed customer expectations, drive productivity, and cost effectively create profitable revenue are some of the key factors that will drive market adoption of iDataPlex. iDataPlex is an essential technology in support of future massive scale-out data centers and customers facing high consumer expectations and demands for continually improving SLAs and QoS. If they want to meet these demands, stay competitive in these markets, sustain the required levels of growth these customers will require an energy efficiency, density, and affordable hardware platform, and iDataPlex is certainly an option that customers should be evaluating based on their application and data center requirements. The reality is that SOA, Web 2.0, and other emerging technologies will keep pushing the envelope for such demands, and will have direct influence on the iDataPlex vision and future development.

Reduced Power consumption

When iDataPlex Flex nodes are compared to more traditional nodes they are slightly more than half the depth which provides for better use of available real estate as well as greater cooling efficiency. Additionally, with the unique design of iDataPlex significant benefit is delivered through the combination of systems, storage, and networking connectivity, with innovative shared cooling and shared power. iDataPlex is based on industry standard components, assembled as a complete solution in IBM manufacturing, and shipped to customer sites. This allows the customer quick installation with minimal time required for set up, installation, and configuration, and also reduces the amount of packaging that has to be put in a local land fill.

Cooling

One key goal for IBM in developing iDataPlex was to eliminate the requirements for Air Conditioning and with the design of the optional iDataPlex liquid rear door cooling unit this goal has been met. By using this rear door cooling system a data center can eliminate the requirement for Computer Room Air Conditioning (CRAC units). The efficiency of the iDataPlex rear door water cooling system, based on data provided by IBM, is more than 100%. When the rear door heat exchanger is used, the air that comes out the back of the rack is cooler than the air going into the rack. Using this system requires that the data center have plumbing and chillers as part of the data center capabilities.

Green Value

From what HRG has learned it is safe to say that iDataPlex delivers significant Green Value with up to a 40% reduction in power used to cool the data center. Again, according to IBM, iDataPlex has better compute density than traditional rack mount systems and requires roughly ½ the floor space. IBM estimates the energy savings on a per rack basis at \$10,148 per rack giving IBM a significant TCO and ROI advantage.

Conclusion

Customer-facing web portal based applications and support for Web 2.0 and SOA requirements are driving enterprises to compete based on their ability to scale out in order to meet customer SLA and Quality of Service (QoS) expectations. IBM's new System x iDataPlex servers offer a solution that is eco friendly, flexible, responsive, and positioned to address the evolving requirements of Internet Scale Data Centers.

Many organizations today are facing increasing application integration demands, as well as the need to scale capacity in a secure, reliable, and cost effective way. A number of vendors provide comprehensive suites of software that address many or all of the functional requirements for developing, deploying, and managing SOA-based systems. WebSphere offerings have been specifically designed and enhanced to meet these requirements. IBM has recognized the need to extend its software offerings with a number of products that are specifically focused on enabling application integration more easily and reliably, and with greater performance, using standards that make the resulting solutions forward-looking in terms of conforming to evolving best practices.

While Web 2.0 developed applications are essentially free the compute infrastructure of those organizations whose data, content, and application logic are used by those applications is not without cost. There are real and potential business benefits to be derived from supporting Web 2.0 however it is incumbent on organizations making this choice that they plan their infrastructure accordingly. IBM's WebSphere Extended Deployment (XD), WebSphere MQ, and IBM System x iDataPlex are well positioned to help today's internet focused businesses stand up to this challenge.

Harvard Research Group is an information technology market research and consulting company. The company provides highly focused market research and consulting services to vendors, resellers, and users of computer hardware, software, and services. For more information please contact Harvard Research Group:

Harvard Research Group™

PO Box 297
Harvard, MA 01451 USA

Tel. (978) 456-3939

Tel. (978) 925-5187

e-mail: hrg@hrgresearch.com

<http://www.hrgresearch.com>