Issue Date: IBM edition for UNIX
October | November 2004, Posted On: 10/1/2004

Balancing the Challenges of an On Demand World

Jaqui Lynch

The buzz surrounding on demand solutions is strong. Customers want 24-7 access to online ordering and support staff. Company executives want this done efficiently and without major budget hits. How do you move forward and satisfy all parties?

To evaluate and attempt to solve the major challenges of implementing on demand solutions, it's important to understand what this term means and what comprises the various components of an on demand environment. This article defines on demand computing, outlines some key technologies and details the top five challenges in implementing such solutions.

What is On Demand?

In the April 9, 2003, launch of IBM* eServer pSeries* on demand solutions, IBM defined an on demand business as: "An enterprise whose business processes—integrated end-to-end across the company and with key partners, suppliers and customers—can respond with speed to any customer demand, market opportunity or external threat." The on demand business world is a combination of hardware, software, technologies and expertise, and involves a change in the business mindset. It's not something that can be purchased off the shelf. To run an on demand business, a company needs a computing infrastructure that can support it—IBM calls this the on demand operating environment.

Today's businesses must be agile in order to interact with business partners and keep their costs as low as possible. The business world is extremely competitive, and customers expect to pay less for more, so there's additional pressure to make services and products more competitive and flexible in customer offerings. Additionally, businesses must be able to handle spikes in business cycles in a more transparent and less expensive manner. These are all critical concepts in the on demand world.

One could think of on demand as being the move from traditional structured infrastructures to more of a matrix management-type infrastructure, where everything and everyone is interconnected and interdependent and where businesses have the capability to move quickly as customers' needs change.

The on demand operating environment is a modular environment where components can be added or removed to adapt. This modular approach allows a company to gradually migrate its technology to on demand with the option of mixing and matching components to take advantage of best of breed. The intent is to allow companies to build infrastructures that are integrated and more manageable than those in the past. IBM has also introduced its IBM Virtualization Engine* Suite to enable monitoring and help support business-process change in the IT environment.

Characteristics in Detail

At IBM's 2003 launch, the four key characteristics of on demand businesses were listed as: responsive, variable, resilient and focused. From a technology perspective, these were tied to the following OS characteristics: open, autonomic, integrated and virtualized.

Open—An open environment is one where the user has a choice of best-of-breed applications and where they can intermix solutions to get the desired results. These applications must be standards-based and interoperate seamlessly. A common component is the requirement for using low-cost networks, which generally means the Internet. The capability to take advantage of open standards can help reduce costs that may occur with vendor lock-in and also allows customers to be more responsive and add new applications as needed. The adoption of POSIX standards and the implementation of Linux* Affinity in AIX* 5L have improved the openness—both perceived and real—of the AIX OS and have improved the portability of software and applications.

Autonomic—Autonomic systems allow for automation to reduce the reliance on hard-to-find, expensive, skilled personnel. Autonomic systems are also designed to provide reliability at previously unheard of levels. In the IBM world, autonomic refers to self-configuring, self-healing, self-optimizing and self-protecting technologies. These systems tend to have features that allow them to do a larger proportion of self-management and tune themselves to adapt to customers' changing needs. The two key goals are to automate systems management and maximize system availability.

Integrated—For an environment to truly respond to on demand needs, it should also provide complete business solutions in an integrated manner. It should be easy to manage and include all of the necessary forms of networking. On demand businesses must have applications and processes that are integrated end-to-end across not only the company, but also with key suppliers, customers and other partners. Flexible middleware and a commitment to open standards are major components of the on demand approach. These standards include technologies such as Java*, SOAP, XML, UNIX*, Lightweight Directory Access Protocol (LDAP), SNMP and Linux.

Virtualized—Finally, the environment should be virtualized so it can be more responsive. Resources must be shareable and have the capability to redeploy rapidly and somewhat transparently. Virtualization allows resources to be shared between systems and also allows for better resource utilization. The capability to dynamically shrink and grow resources with changes in demand is vital. Virtualization provides the facilities to dynamically provision server, storage and application resources as the workload demands. In AIX, this is done by taking advantage of features such as dynamic LPAR, Capacity Upgrade on Demand (CUoD) computing and workload management. These strategies help ensure integration, performance and the capability to move quickly when workloads change.

The Challenges

Now that I've covered what on demand computing is, let's examine the challenges in its implementation. There are a few, but I'll detail five of them: tradeoffs, security, budgets, understanding and planning, and changing the IT thought process.

Tradeoffs—Tradeoffs are a reality with any technology. With on demand it's important to note that decisions must be made regarding security versus performance. How much encryption is enough? How do I best protect data? How do I get the best performance from my storage area network (SAN)? Additionally, open technologies such as XML aren't always as efficient as the equivalent application written in some other language. However, they're far more portable. At what point does this affect the ability to move forward? The key to dealing with tradeoffs is to understand critical business needs and design the modules of the on demand environment to best fit those needs, adopting the best tools in each instance.

Security—Security is a major issue for many reasons. First, data is now leaving and entering the enterprise across networks that the company has no control over (i.e., the Internet). The need to protect data and validate the identity of those accessing the systems is critical. It doesn't help that XML messages are easily read. These issues are being addressed by both technologies and regulations. An additional issue involves the sharing of data between companies. Part of being an on demand business is end-to-end integration with business partners, suppliers and, when needed, customers. This tends to involve a two-way data flow. This data must be carefully validated, and sensitive information encrypted.

Budgets—An on demand world is dynamic and systems must be able to change rapidly to handle ever-changing workloads. This has a major impact on budgeting and planning. Most companies have an annual budget-planning process where funds are put aside to buy more processing power at set intervals. These are typically charged back to the business units based on a million instructions per second (MIPS), rPerf or some other measure of processor power. As companies move into an on demand mode, planning and budgeting become a little more difficult. Processors are now dynamically managed—the system decides which application gets what resources and just moves them. A system may be using anywhere from one tenth of a processor to a number of them, and this can change every couple minutes. Since customers are charged based on usage, this can make the budgeting and planning processes challenging, particularly when on demand technologies are initially implemented. This can have a significant impact on business unit budgets.

Understanding and planning—The on demand concept introduces several acronyms and new technologies. It's important to understand these technologies and where they might best fit into the business model. There are many cool technologies that aren't necessarily the best fit for a certain problem. Planning an on demand environment requires not only understanding the business issues, but also knowing which technologies to apply them to. Specifically, it requires an understanding of virtualization, workload management, partitioning and capacity planning. Those planning the environment must also be able to clearly articulate their technical recommendations to the business side of the house so that the technologies' benefits are understood along with any cost savings or increases. Planning becomes a critical component of the on demand environment—flexibility and openness don't just happen—they're designed into the infrastructure. On demand environments require careful design and a clear understanding of the technologies.

Changing the IT thought process—The hardest challenge to overcome is the cultural one. Those of us in the IT world understand technology and use acronyms in one sentence more often than most people use in a month's worth of conversation. We also like to be very hands-on. As we move away from technology silos to a more integrated business approach where IT is closely aligned with the business, we struggle with several issues. The first is giving up control; it's

difficult to let the machine make decisions that we're used to making. In an on demand world, the business needs drive the IT needs, and the systems tend to be self-managing and highly automated. When they need more resources, they get them if they're available.

Many in IT also tend to be reactive rather than proactive. An on demand business requires a proactive approach with a great deal of planning, primarily to design an environment that can react quickly as business needs change and adapt to integrate new technologies that may be beneficial to the company.

The Keys to Success

As I mentioned, an on demand business is one that's integrated end-to-end with all of its suppliers, customers and partners. On demand isn't a set of products; it's a combination of technologies, products and a way of conducting business. The environment must be designed to be flexible so the business can move rapidly when it needs to, while still maintaining the critical infrastructure that keeps it stable. The keys to helping ensure the success of an on demand project involve understanding and overcoming the five challenges. It's vital to understand the tradeoffs that must be made, and the security issues that must be addressed.

Planning is a critical component, regardless of whether it's for budgets or how to best integrate technologies into the business. Finally, people are a critical component. Those of us in the IT world must change the way we look at computing resources and become more proactive and align ourselves better with the business units so we can more quickly meet their needs. The "not invented here" syndrome tends to combine with the "don't touch it, it's working" complex to affect our ability to accept change. On demand is all about change—everything changes.