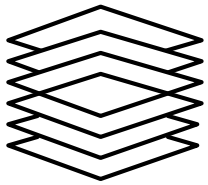

Server Consolidation with IBM eServer xSeries

Bill Moran, EVP and Research Director

EXECUTIVE SUMMARY

Server consolidation is a key issue for businesses of all sizes. Many solutions now provide an optimal platform for server consolidation using a variety of system architectures. This paper reviews the server consolidation capabilities of IBM's eServer xSeries architecture, which is based on Intel processors and runs either Windows or Linux. Over the course of the next several months, as the x440 scales to sixteen way, the xSeries architecture will offer a platform that is suitable for virtually all classes of server consolidation. This architecture will provide the necessary functional capabilities for employing a variety of server consolidation tactics (see Table 1).

IBM's eServer family consists of discrete series' and models that are capable of delivering specific functions to meet targeted business needs. While the product lines share a common name, each class of system has been developed and implemented to meet the specific performance, price, and feature needs of different sets of customers with different hardware and software. Each eServer family member supports IBM's strategic software and middleware, including the Linux operating system, Apache web server, WebSphere web application tools, and the DB2 database server. This commonality masks the architectural differences between platforms so that users do not have to care about the underlying incompatibilities of hardware and systems software, any more than a motorist cares that the pistons in his or her Chevy will not run in a Mitsubishi. IBM's server consolidation strategy spans all major eServer platforms, while its methodology for server consolidation has been in place for more than five years. This methodology encompasses planning and requirements definition, sizing, user issues, and best practices that are customized for each member of the eServer family.

In today's challenging economic environment, few projects will be undertaken without a solid financial justification. Hence, various IBM cost-cutting services can play a critical role in the consolidation process. In fact, despite its powerful rewards, many customers have trouble identifying the most promising candidates for consolidations and they may also have difficulty in building the business case for consolidation. IBM provides specific tools that will help in the Intel space. IBM has developed a data collection tool that will help identify

candidates for consolidation and an ROA (Return on Availability) tool that will help with the business case. IBM also offers services that specifically target the industry standard server environment. Beyond this, IBM can provide the skills to actually do the consolidation.

One attractive feature of the xSeries is that it offers customers a choice of operating system. Many customers will choose Windows, which offers strongly integrated solutions for database, e-mail, and web server solutions. Others will choose Linux, which allows them to assemble low-cost, best-of-breed solutions, both open source and for purchase. VMware's ESX server will prove a key tool for many customers because of its virtual machine functions. Finally, legacy Netware solutions can be accommodated on the xSeries. The key point about the xSeries is that customers can choose solutions that meet their business needs and address their consolidation requirements.

Better workload management is a powerful tool. Vertical scalability allows a single SMP server to support the workloads that were divided up among many smaller servers. The "building block" approach of the xSeries equipped with the x440 offers excellent vertical scalability economically. Customers can expand by ordering four CPUs at a time. Today's xSeries scales to eight processors and sixteen will be offered by the end of the year. It is likely that additional processors will be supported during 2003. In addition, the latest Xeon processors are offered in the xSeries, while the systems that are being consolidated are likely based on older and slower processors. IT managers and their financial counterparts will also note that the Xeon contributes to a performance boost for the xSeries.

Workload management tools allow the xSeries to handle diverse workloads. Two key workload management tools are available. The first is system partitions, which allow administrators to run multiple instances of possibly different operating systems within a single server. For example, both Linux and Windows can be run in a single system simultaneously. Partitioning technology allows the processors, memory, and I/O to be divided between the two operating systems. Partitions are key complements to the other workload management tools offered on the xSeries. The second key tool is software – ARMtech, an Aurema product. ARMtech allows administrators to control the assignment of resources within a single partition or in systems where partitions have not been defined. ARMtech can guarantee that a job gets its share of the system resources or alternately that no job hogs the system.

For applications that can "scale out," i.e., can be segmented to run in parallel on multiple machines simultaneously, clustering proves a powerful tool for server consolidation. The consolidated servers can continue to run their workloads on physically separate nodes in the cluster, while being managed logically as a single server. Both Windows and Linux offer clustering choices.

TABLE 1: Server Consolidation Functional Requirements vs. xSeries Capabilities

Function	Importance for Server Consolidation	xSeries Capabilities
Services	Assist clients to identify and quantify consolidation benefits. Supplement the client staff to do the actual consolidation.	• ALIGN Methodology • IBM Services
Choice of Operating System	Allows applications to be consolidated on best-suited operating system. VMware allows Linux and Windows to run on the same system. It can also allow different versions of the operating system to share the same system.	• Multiple Linux Distributions • Windows Server, Advanced Server, and Datacenter Edition • Netware
Vertical Scalability	Allows many small servers to be replaced with a single, larger server.	• Enterprise X-Architecture Technology • Eight Intel XEON processors (Sixteen-Way by Year End 2002) • 32 GB Maximum Memory Today, 64 GB in Future
Workload Management Tools	Allows multiple, dominant applications to run simultaneously on large servers.	• Partitions • VMware • Aurema Armtech
Horizontal Scalability	Allows multiple servers to be consolidated into clusters that are easily managed.	• SMP Expansion Ports • High Density Rack • Microsoft Cluster Services • VERITAS and SteelEye • Beowulf/Linux
System Management Tools	Allows users to continue managing their consolidated servers.	• IBM Director • Cluster Systems Management (CSM) • Integrated Systems Management Processor
Reliability Features	Reduce the risk of introducing single points of failure when consolidating.	• Chipkill ECC • Memory Protection • Hot-Swap Memory • Memory Mirroring • Predictive Failure Analysis • Active Diagnostics
I/O Scalability	Move PCI slots out of the main system cabinet to allow for greater I/O connectivity (3X slots per processor) RXE-100.	• Remote I/O
Virtual Machine Facility	Allow multiple operating systems to share single system for testing, management, availability.	• VMware: ESX and GSX Server
Autonomic Computing	Reduces the need for human intervention in managing systems and software.	• Auto-Hardware Detection • RAS Features
Windows Consolidation Support	Allows LAN infrastructure, file- and print-sharing services, and other applications to be hosted with better reliability.	• Windows Datacenter Edition • Resource Management Tools
Linux Consolidation Support	Allows consolidation of Linux workloads onto larger servers.	• Resource Management Tools

Note: Features in the table are generally taken from the xSeries 440. The Appendix of this paper compares this system with the smaller xSeries 360.

System management tools help the users to manage their consolidated servers. IBM provides the IBM Director that will assist in managing the system. The Director provides both hardware management such as monitoring the status of the system as well as support functions like real time diagnostics. In addition, other auxiliary functions like Cluster Systems Management and Software Rejuvenation are delivered as extensions to the Director. The Director can easily integrate with the various Framework products like Tivoli, HP Openview, BMC, or CA Unicenter. All in all, IBM delivers under the umbrella of the Director a very rich set of systems management functions.

To protect the servers from failure after they have been consolidated, xSeries offers a number of reliability, availability, and serviceability (RAS) features that can sharply reduce incidents that may affect availability. For example, the xSeries offers extensive technology to guarantee that the system is protected from memory failures. A key technology here is memory mirroring. This means that backup is provided by a duplicate memory. Obviously, not every system will need or can justify this feature. However, when a memory failure in a mission-critical application may cause many hours of downtime and lost revenue, there will be systems that will find memory mirroring to be well justified.

When servers are consolidated, the data associated with each server also needs to be consolidated. Clearly, the combined server must access all the data that the individual servers controlled. The X-Architecture provides an elegant way to do this. The I/O interfaces can be moved outside the chassis of the system and collected into an external device. This will allow far more slots per processor and will also make it possible for different servers to share I/O.

The virtual machine facility in the VMware on the xSeries may be new to many installations. A virtual machine means that multiple, different operating systems can run in a single server. Thus, testing and development work can proceed in parallel. In addition, work that may be running on different versions of an operating system can be consolidated far more easily onto a single system. The virtual machine facility will complement the system's partitions.

BUSINESS BENEFITS

The strong functional capabilities provided by xSeries, the X-Architecture, and IBM support translate into a number of business benefits that are directly associated with IBM's ability to enable server consolidation. These benefits include,

- IBM Services can define the plan for consolidation and help quantify the savings that will be achieved.
- The choice of operating system means that the necessity to port applications to another operating system is greatly reduced. This will normally translate to savings in time and cost.
- The vertical scalability enabled by the xSeries' SMP range allows many small servers to be replaced with a single large server. The reduction of the number

of servers and operating system images results in a correspondingly lower system administrator headcount. Also, larger SMP servers require fewer physical resources (space and power) than multiple smaller systems.

- The horizontal scalability enabled by the xSeries clustering functions allows multiple servers to be consolidated into clusters that can be managed with a single console. This also reduces the number of administrators required to manage the infrastructure.
- The workload management tools allow multiple, dominant applications to be run simultaneously on large xSeries servers without compromising their response time. This improves the productivity of internal users, and protects the brand for external users, who will not perceive unexpected slowdowns when accessing a website.
- The strong availability options for xSeries reduce the risk of the entire consolidated system being down. Improved uptime maintains productivity for internal users, and makes it more likely that external web users will have a successful session on the company's website.

IBM's eServer product family provides multiple platforms that are suitable for performing server consolidation, depending on the circumstances of a particular organization. The xSeries architecture offers an Intel-based solution with a good balance between vertical and horizontal scalability, including very competitive single-system functions, i.e., SMP and workload management capabilities, and multi-system functions, such as clustering for performance, availability, and manageability. For users committed to industry standard technology, xSeries offers an impressive portfolio of tools that provides the flexibility needed to execute a variety of server consolidation tactics covering sites, servers, data, or applications.

TABLE OF CONTENTS

SERVER CONSOLIDATION BENEFITS

Server consolidation provides a major opportunity for businesses of all sizes. For example, the rapid implementation of e-business websites has caused servers to proliferate at an astonishing rate. This follows the legacy of LANs in the 1980s and 90s, which resulted in an enormous number of industry standard servers installed simply to maintain network infrastructures. Today's server farms for handling web traffic of all kinds continue to grow rapidly, potentially introducing yet another requirement for server consolidation in the future. Beyond these patterns, the popularity of both Windows and Linux servers have led to an explosion in servers.

Managing large numbers of servers introduces a variety of costs that can have a profound impact on total cost of ownership (TCO). Some of these costs are overt, due to the proliferation of staff inherently required to manage large numbers of servers. Others are hidden, deriving from the inefficiency of administering servers that are physically dispersed, each of which potentially involves its own standards for management. Complications can also arise due to the increased risks to security and data integrity when large numbers of servers are involved, each of which represents a potential vulnerability. Finally, overall system availability can become unpredictable when smooth functioning depends on large numbers of servers.

Thus, common sense and traditional IT management practices dictate that server consolidation should be a routine, business-as-usual undertaking. Unfortunately, the term "server consolidation" is not well defined and carries substantially different meanings to individual vendors and customers. As a result, the requirements, wishes, claims, and counter-claims related to server consolidation vary widely depending on the viewpoint of the parties and architectures involved. This paper stresses the D.H. Brown Associates, Inc. (DHBA) view of consolidation, which is not necessarily identical to IBM's view.

Generally, server consolidation implies that administrators gain a single point of control, while maintaining flexibility and prioritized service levels for users. Specific server consolidation tactics can be grouped into the following categories:

- *Site Consolidation*, i.e., reducing the number of datacenters in an organization. As a result of acquisitions and other business factors, large organizations often depend on multiple, geographically dispersed datacenters for their operations. This introduces inefficiencies due to facility costs for power, cooling, floor space, etc. These costs can be reduced through consolidating the operations of some datacenters into others, which increases economies of scale. At minimum, the administration of these datacenters can be centralized, even if the servers physically remain in different locations. In a single site, consolidation usually means finding the dispersed servers and, so far as possible, collecting them so that the IT department can manage them efficiently.

- *Hardware Consolidation*, i.e., replacing multiple small servers with fewer, larger servers. This approach reduces costs through simplification, since it eliminates most management overhead related to power and network cables, and applies economies of scale to configuration of computing resources such as processors, memory, and disk. These resources must be managed separately on each individual server in a traditional network. But through server consolidation, only one computer must effectively be managed in a single consolidated system. This potentially reduces the personnel required to service multiple servers. Physical consolidation also reduces physical resource consumption, i.e., floor space, power, and cooling.
- *Data Consolidation*, i.e., merging data from multiple locations into a single system that can be accessed through some standardized method. This approach reduces costs by eliminating the overhead for keeping multiple copies of data consistent. Data consolidation can be performed at different levels, i.e., database consolidation, or storage consolidation through disk virtualization.
- *Application Consolidation*, i.e., combining the operations of multiple similar applications onto a single server. This approach reduces costs by eliminating inefficiencies due to multiple departments purchasing the same application functionality independently, and by improving resource utilization through greater economies of scale.

IT experience shows that many companies developing a server consolidation plan will begin with site consolidation. The first task is to locate all of the servers at the site and identify their workloads. Usually, companies will choose to collect these servers into a few central locations. It is amusing that servers will be found scattered around in the most unusual places, even under desks and in closets. Companies have told DHBA that IBM via the ALIGN process has been a major help to them in simply locating all the servers. IBM was also able to determine the actual growth rate of the server population.

A second step in consolidation is hardware consolidation. It is probably wise to delay this step until the first phase has been completed. In this phase of consolidation, the workloads from a number of smaller servers are combined into a single larger SMP server. Partitioning tools and the virtual machine capabilities of the xSeries will come into play here. When multiple jobs need to run together in a single partition or system, the facilities of Aurema's ARMtech will allow the resource allocation to each job to be controlled.

In parallel with the consolidation of the workloads, the data from the several systems needs to be combined so that the consolidated system can access all required information.

The X-Architecture offers expanded I/O capabilities that will make this task easier. Many installations will also want to consider Storage Area Networks (SANs) or Network Attached Storage (NAS) solutions to the consolidated storage

problem. As a part of this consolidation, high-speed backups should be considered to decrease backup time and increase availability.

Typically, companies find that their data is scattered among numerous databases. This causes increased costs from inconsistent data and difficulty in accessing data to support web access by customers. The solution is to consolidate databases. One part of the solution is easy but another part is hard. Many companies have told DHBA that consolidating multiple Oracle databases into a single system is easy. Of course, all the Oracles need to be at the same level so there will be some effort involved with these upgrades. Thus, combining multiple Oracle, SQLServer, or DB2 databases may be relatively easy. However, combining different databases is a far more challenging task. For example, combining a mixture of Oracle and SQLServer databases into a single database may be an expensive and lengthy process. Such projects should not be undertaken without a careful study of the risks and costs. IBM's Global Services can help to size and offer assistance with these projects. Generally, companies have found large paybacks in database consolidation and with care, such paybacks may be relatively easy to achieve.

The final phase of consolidation concerns application consolidation. This involves replacing multiple instances of an application with a smaller number. Software license fees and systems administrator time can be cut. E-mail systems like Microsoft Exchange or Lotus Notes are generally prime candidates for this kind of consolidation as well as file and print servers. IBM can help to identify applications where such consolidation will be relatively easy and provide an immediate payback.

Achieving an immediate payback needs to be focused on when planning consolidation. DHBA has interviewed customers who planned enormous and multi-year projects. Generally, this is a bad idea. The staff working on the project becomes discouraged because success always seems to lie in the future. More important, senior management in today's economic climate will likely become unhappy with a project that sucks up resources and is unable to show any quick benefits. The answer is to use IBM's help along with the company's financial staff to identify projects that will show tangible financial results in as short a time as possible.

SERVER CONSOLIDATION TOOLS

The demand for server consolidation has resurrected centralized computing practices long associated with mainframes. In response, mainframe architecture has once again become fashionable, introducing demand for mainframe-like functions far outside of their traditional proprietary domain, such as in environments like Linux and Windows. The most effective server consolidation platforms offer capabilities that are drawn from traditional mainframe functions, including some of the following:

- *Services:* For some years, IBM has offered its mainframe customers a methodology called ALIGN that helps identify and justify consolidation opportunities. Now the xSeries offers a slimmed-down version of ALIGN (called ALIGN-lite) targeted at industry standard servers. IBM has also invested in tools to support this process.
- *Vertical Scalability Support:* The ability to support environments that "scale up" on large, single servers containing many processors in SMP configurations, along with large amounts of memory and storage. Vertical scalability support is required to replace many small servers with a single, large server. Usually, Windows or Linux servers will run at very low utilization rates, perhaps 5% – 10%. Combining these workloads proves very cost efficient.
- *Workload Management Tools:* The ability to manage multiple, dominant applications running simultaneously on large servers, in a way that prevents them from interfering with each other's operation. In this manner, they provide consistent levels of responsiveness after being consolidated onto a single server.
- *Horizontal Scalability Support:* Support for appropriate applications to "scale out" on clusters using management tools that allow the cluster to be treated as a single system. Such tools help the administration of multiple servers to be consolidated with a single point of control.
- *System Management Tools:* The IBM Director will deliver the key systems management functions on the consolidated xSeries platform.
- *RAS Options:* The X-Architecture features a strong set of RAS functions. Many were originally developed for other systems and have now been implemented in Intel systems.
- *Virtual Machine Facility:* This is delivered by VMware's GSX and ESX products. These products will allow multiple copies of Windows or Linux to share the same physical system.
- *Autonomic Computing:* This is an IBM program to make systems self-healing and self-managing. Many functions available on the xSeries already fit into the autonomic computing framework.
- *Windows Consolidation Support:* The ability to host Windows applications of all types on the consolidated platform.

- *Linux Consolidation Support:* This enables the consolidation of workloads running on Linux servers, which are rapidly proliferating around the edge of the network, and used for selected departmental applications.

There are many possible ways to implement server consolidation, using a variety of system architectures. This paper reviews the server consolidation capabilities of IBM's eServer X-Architecture, which is based on Intel processors and runs either the Windows or Linux operating system. Today, the xSeries architecture offers a platform that is suitable for virtually all classes of server consolidation, and provides the functional capabilities for employing a variety of server consolidation tactics. The discussion below covers some key xSeries functions, highlighting their ability to support server consolidation.

IBM SERVICES

IBM has developed a methodology to identify the requirements for consolidation. This methodology called ALIGN can handle an environment containing mainframes, various types of UNIX servers, industry standard servers, and other legacy systems.

However, many companies that only have industry standard servers may not need all the capabilities that the full ALIGN methodology includes. For these customers the xSeries group has developed a stripped down version of ALIGN. This version called ALIGN-lite focuses on the industry standard servers that a customer (or perhaps a department in a larger company) may have. It also offers the merit of taking less time than the full ALIGN process and therefore it is less expensive.

The ALIGN-lite process can normally be completed in three or four weeks.[1] It will identify the servers to be consolidated. IBM can also assist in building the business case for the consolidation. However, companies considering consolidation should consider involving their own financial team to sanction the benefits that are projected. Companies have the option of engaging IBM to do the actual consolidation or of doing it themselves. For companies that lack the skills or the resources to do the work themselves it may be reasonable to engage IBM for this project. IBM has made the investment to train people very widely across the areas that they cover to ensure that talent is available to assist with xSeries consolidations.

Many studies have shown that a high percentage of IT projects fail either by being hopelessly late or over budget. A significant percentage of IT projects are abandoned and never completed. This is a key reason to get help from IBM on a consolidation project.

[1] Depending on the circumstance, i.e., how much IBM equipment the customer plans to purchase, the ALIGN-lite study may be free.

IBM can supplement a client's inexperienced staff and drastically reduce the risk of failure. Of course, this will not eliminate all failures because successful consolidations usually depend on the cooperation of the business units involved as well as other business considerations. However, IBM can help identify these risks and suggest ways to deal with them.

Finally, IBM has created tools that assess the workload of different servers and help project how the combined workloads will behave after consolidation. Their data collection tool (CDAT) will be particularly valuable in gathering information on target systems for consolidation.

CHOICE OF OPERATING SYSTEM

Today the xSeries offers two major operating system choices: Windows and Linux. The features of Windows and Linux are well known and it is outside the scope of this paper to compare them.

VMWARE

The flexibility of the xSeries is a significant factor. Both system partitions and VMware products allow multiple copies of operating systems to run in parallel on a single SMP. VMware will allow older versions of Windows including even applications running under DOS to be consolidated. This can mean significant savings in time because otherwise, these applications would need to be converted to run on a more up-to-date version of Windows. Use of VMware means that the applications can be consolidated without this effort and an earlier completion of the consolidation can be achieved. This results in an earlier payback. In addition, the partitioning feature means that all three of these operating systems can run in the same system and the installation will gain some flexibility in moving resources among the partitions.[2]

VERTICAL SCALABILITY SUPPORT

Consolidating multiple workloads onto a single large server requires specific functions that enable individual servers to "scale up" to large configurations. These include support for high-end shared memory multiprocessing (SMP), with many processors in a single system, and support for large amounts of memory. The xSeries provides these features.

IBM significantly improved the vertical scalability of the xSeries platform with the introduction of the Model x440, which offers a building block approach. Customers can grow their systems by adding four processor nodes.

[2] Today the partitions are static and changes will require that the operating systems be reloaded. Thus the frequency of changes may be limited. In the future, when the various operating systems support dynamic partitioning, changes will be easier.

WORKLOAD MANAGEMENT TOOLS

Trying to run multiple dominant applications on a single server, when each of which expects to consume all available resources (i.e., database servers), challenges administrators to ensure consistent responsiveness for all applications. Workload management tools can help to overcome this problem by allowing large numbers of resource-intensive applications to run simultaneously on a single server. This is accomplished through flexible scheduling policies, and makes these tools a key enabler for a variety of server consolidation tactics. Ideally, systems should support two classes of workload management tools: *partitions* and *resource management tools*. Partitions allow administrators to run multiple instances of an operating system within a single server. With partitions, each instance behaves as if it were running on a standalone machine and is totally isolated from other instances. Resource management tools effectively manage multiple applications with constantly changing behavior running within a single operating system instance.

Partitions and resource management tools have long been available on mainframes, but have recently entered the Intel space as well. Indeed, IBM has embraced a number of high-end features that were long reserved for its mainframes in the X-Architecture. The xSeries 440 introduces system partitions, a rigorous form of workload management that allows administrators to deploy multiple instances of an operating system within a single server. Each instance behaves as if it were running on a standalone machine. "Bullet-proof" barriers between the different environments maintain overall system robustness, so that even the most extreme application failure or operating system crash in one partition leaves the others unaffected. The entire environment, i.e., all partitions, can be managed from a single point.

Currently, the boundaries between partitions are static, meaning that the operating system instance running within a partition must be restarted if processors or memory are added or removed. As operating system support for partitions arrives, the xSeries will be able to support dynamic partitions. This ability will greatly increase the flexibility and business value of the partitioning. For example, a single server could host multiple departmental systems during regular business hours. This server could dynamically shift its resources into the central database server's partition, which should not be shut down at any time, to run large batch jobs at night. The partitions could then be reverted to their regular configuration when business resumes.

xSeries also provides resource management tools for controlling the behavior of applications that need to coexist after they have been consolidated. Aurema's ARMtech allows administrators to control the allocation of system resources, including CPU time, memory, and disk I/O, on a fine-grained basis. This is a critical feature for server consolidation, as it ensures that high-priority applications will never be delayed by lower-priority processes running at the same time.

HORIZONTAL SCALABILITY SUPPORT

Some types of applications can "scale out," i.e., they can be segmented to run in parallel on multiple machines simultaneously. For these applications, clustering tools can be used to increase system capacity, including performance and storage. To run on a cluster, applications work in concert with clustering software to partition their workloads into subtasks. The clustering software then distributes these subtasks across the clustered servers. Typically, the clustering software provides sophisticated tools that allow the cluster to be managed as one system. These tools prove to be very useful for server consolidation, since the consolidated servers can continue to run their workloads physically on different nodes in the cluster, while being managed logically as a single server. Windows, Linux, and VMware all support various forms of clustering to meet different requirements. Linux, for example, provides Beowulf clustering, which targets clusters for high-performance technical applications. IBM also offers products from VERITAS and SteelEye that deliver cluster functionality. Both of these products will allow the failover of applications to a backup system in the event of a failure, thus ensuring a higher level of availability than can be achieved in a single system. Deciding which of these products will best meet a customer's requirements is beyond the scope of this document, but IBM can assist. In addition, DHBA has published other documents on the topic. (See the DHBA website, www.dhbrown.com, for selected abstracts.)

HIGH-DENSITY RACK

The xSeries rack model adds value to the horizontal scalability consolidation scenario through its very dense and highly efficient packaging. This high-density rack consolidates significant power into a small 8U (fourteen inches) form factor that supports a fully configured sixteen-way server. The eight-way server fits in a 4U (seven inches) form factor, 75% smaller than most competitive servers. The Rack Manager configures and manages the rack via drag-and-drop depictions on a realistic graphical portrayal of the rack. It also provides health status details of the rack and its components.

SYSTEM MANAGEMENT

A key challenge involved with successful server consolidation is to continue providing users with acceptable levels of control over their environment, even as IT organizations assume some of the management of that environment, at least from a physical standpoint. Systems management tools assist the IT organization in configuring the system to meet user needs. The most important systems management functions are in the IBM Director product.

IBM DIRECTOR

Effective server consolidation solutions must provide leading-edge system management capability. The IBM Director provides a suite of "lifecycle" tools and utilities that manage a large number of Intel servers. Consisting of a management server, the management console and agent, and optional Server

Extensions, it embraces systems management industry standards and integrates into enterprise systems from vendors including IBM Tivoli, Computer Associates, Microsoft, HP, and Intel.

Universal Manageability (UM) Services are Tivoli-ready and provide the agent for IBM Director. These services communicate with physical and logical devices to surface data and to monitor status. Available data covers the operating system, devices, ports, memory and network adapters, power management, system shutdown, event logs, and system monitors.

The IBM Director provides a single, consistent manageability interface for both workgroup and enterprise management. It provides an Event Management task that identifies, categorizes, and automatically initiates actions in response to network events. It supports the creation of an Event Action Plan that automatically pages an administrator when a threshold is exceeded. Other tasks included in the IBM Director are the File Transfer Task, Inventory Management, Process Management, Remote Control, Resource Monitoring, Task Scheduler, and SNMP Manager.

Extensions in the Director added to support the X-Architecture technology include,

- System partition manager, which provides a graphical interface for creating and managing a running partition. It uses the network link to the system management processor to set the relationships among nodes.
- Cluster Systems Management (CSM) allows a single point of control for multiple Linux or pSeries UNIX systems. CSM eases the management of the cluster.
- Active PCI Slot Manager optimizes I/O performance by guiding installation of PCI and PCI-X adapters into the best slots. It also helps decide whether adapters should be placed into the server or into a remote I/O expansion box.

RAS FEATURES

IBM has leveraged its years of experience in building systems into the X-Architecture. This experience shows up clearly in the many reliability features that are available in the system. In this paper, RAS features related to consolidation are highlighted.

HOT-SWAP/REDUNDANT SYSTEM COMPONENTS

The X-Architecture allows failing components to be replaced while the server continues to run, a feature called "hot swap." Currently, hot swap of drives, adaptors, power supplies, and fans is possible. The architecture has now added support for memory. Some components are also replicated so that the redundant items can continue if a failure occurs. Power supplies and fans are backed up in

this fashion and online spare redundant component options are supported for memory, hard drives, and network adapters on xSeries servers.

MEMORY RELIABILITY

The large amounts of memory offered on the xSeries make it feasible to have entire databases in memory. However, with all of the database in memory, a failure leading to a system crash could involve lengthy restart time. Thus, the X-Architecture places great emphasis on memory reliability and availability enhancements. Memory ProteXion, an enhancement borrowed from IBM mainframes, protects against unplanned outages from memory errors. Servers protected with Memory ProteXion have demonstrated reliability in excess of 200 times[3] that of servers using standard error checking and correction (ECC) memory.

Standard ECC memory detects two-bit errors but can correct only single-bit errors. If multiple bits fail at once, the server can crash with reduced memory capacity. Memory ProteXion uses previously unused bits in the data in memory nodule (DIMM) to correct up to eight sequential bit failures per memory controller. This substantially reduces potential server downtime.

Chipkill memory takes reliability even further. If a server experiences so many errors that Memory ProteXion cannot handle them, Chipkill memory corrects up to eight bits per memory controller whether on a single chip or on multiple chips. Together, these functions should reduce uncorrected memory errors to an extremely rare occurrence.

The third line of memory error protection is memory mirroring. With memory mirroring, data is simultaneously written to two independent, identically configured memory cards. If memory failures cause an entire module to fail, the server automatically shifts to using the mirrored memory card. Memory mirroring also enables hot-swap and hot-add memory, which allows failing parts to be replaced while the server is running. Memory errors result in no downtime. However, memory mirroring comes with a cost: half of the memory in the system is not used except for backup.

PREDICTIVE FAILURE ANALYSIS (PFA)

PFA sends a warning message that enables timely replacement of a "failing" component based on predefined threshold values. The IBM Director sets up an alert to a system administrator for an impending failure on processors, chipset (XceL4) cache, memory, fans, power supplies, and hard drives.

[3] This translates to one memory failure for each 26,042 servers per year.

CPU FAILURE RECOVERY

Most Intel-servers today feature Automatic Server Recovery (ASR), which auto-reboots the server when it is not responding. But when a processor fails, most Intel-servers will not reboot and demand immediate attentions from system administrators. IBM features CPU Failure Recovery on its X-Architecture based Intel servers, which will force the failed processor offline, reboot, generate an alert, and continue to operate with the failed processor deconfigured.

REAL-TIME DIAGNOSTICS

Real-Time Diagnostics, another tool unique to the xSeries, allows an administrator to run diagnostics on system resources while users are working. This removes another reason for downtime and increases customer levels of availability. The combination of Real-Time Diagnostics with PFA and hot-swap components provides a highly available server environment.

SOFTWARE REJUVENATION

The Software Rejuvenation tool predicts pending application and operating system failures that could cause downtime by watching for common failures like memory leaks. It automatically refreshes the software for reliable operation, preventing up to 80% of software-related downtime. In the case of failure it will automatically reboot the system.

I/O SCALABILITY

The X-Architecture provides another exclusive technology, remote I/O, inherited from IBM's high-end pSeries servers. Industry standard PCI and PCI-X[4] specifications require that all adapter slots reside in the main system cabinet. As business expands, this means making a tradeoff between the size of the server box and the number of adapters the system can provide. One solution, moving the adapters outside the server box, is not supported by PCI-X. Remote I/O, implemented in some xSeries servers, corrects this deficiency and extends the PCI-X I/O subsystem to allow external attachment of adapter slots.

This technology supports dozens of PCI and PCI-X slots via an external I/O expansion box, the RXE-100,[5] connected to one or several servers. As more I/O slots are needed, plugging in an external I/O expansion unit doubles or triples the number of available slots. In addition to the obviously increased scalability, this expansion box adds reliability by enclosing separate power supplies and fans within the expansion box.

---

4 PCI-X is an enhanced bus that extends the useful life of PCI until next generation I/O architectures become available. PCI-X increases the available I/O bandwidth to 1 GB of data per second.

5 The Remote Expansion Enclosure (RXE) ships with six PCI-X adapter slots and expands to 12 via an optional "six-pack" upgrade.

AUTONOMIC COMPUTING

Autonomic computing is IBM's initiative to develop self-managing eServer computers and related system components. Autonomic computing is based on the notion that systems should be self-managing, self-learning, self-optimizing, and self-healing. Introduced a year ago, autonomic computing[6] is primarily based on an articulation of IBM's vision for autonomic systems. Such systems perform the "self" activities much in the manner of their biological counterparts from which the term "autonomic" stems. IBM points out that autonomic computing has the potential to produce savings of up to 20% in certain IT budgets.

Self-managing, self-learning, self-optimizing, and self-healing systems developments are all part of the IBM effort to make systems and the IT infrastructure more highly available. The more systems that can perform these four tasks, and the more quickly and efficiently they can do them, the more systems are able to run applications when asked, in less time. This is the essence of "high availability." IBM's autonomic computing initiatives include end-to-end automation and end-to-end system management; workload and cluster management; disaster recovery; security; and virtual server hosting.

These initiatives apply to all parts of an IT environment including the client side, firewalls, load balancing, Internet services, Domain Name Server (DNS) services, web application and data servers, directory and security services, storage, SANs, and more. The environment will be multi-vendor. Moreover, the environment will ultimately range from the customer datacenters and commercial IT infrastructures (the first instantiations and current focus), to a grid computing-based environment connecting multiple sites owned by different parties with some commonality of interest.

IBM will support autonomic functions on all members of its eServer product family, including the xSeries. The xSeries already provides capabilities that IBM fits into the autonomic rubric. For example, the Chipkill and ECC memory as well as the self-configuring PCI capability are autonomic functions. With this initiative, IBM is organizing functions such as these into an overall architecture that could even span products from multiple vendors in the future. Moreover, IBM is making autonomic computing a major marketing and technology thrust, which adds credibility to the autonomic computing concept for all servers. As a result of all this, major benefits will accrue to IT administrators and users as these technologies evolve.

WINDOWS CONSOLIDATION SUPPORT

Windows has been extraordinarily successful at enabling small networks to be set up quickly and simply for tasks such as file- and print-sharing. As a result, organizations have become burdened with large numbers of Windows NT servers dedicated to maintaining various networks. Due to Windows NT's limited SMP

[6] This is IBM's new term for what was called "project eLiza."

scalability, most servers it runs on are configured with four or fewer processors, which means that more servers have to be deployed for handling a particular workload. Also, many organizations have installed a large numbers of Windows NT servers simply to serve as Primary Domain Controllers (PDCs), in order to authenticate users on the network. Because PDCs have widely fluctuating resource consumption, administrators usually try to avoid deploying PDCs on servers that run critical applications. This results in a profusion of servers dedicated solely to hosting the PDCs. A variety of tools can be used to rehost these Windows NT functions on larger SMP systems, which potentially offer better economies of scale to provide the same services.

LINUX CONSOLIDATION SUPPORT

Linux has become a mainstream operating environment for infrastructure solutions, "edge of network" applications, development platforms, and technical computing. Because Linux runs extraordinarily well on low-end Intel X86 servers, a proliferation of Linux servers for a variety of applications is starting to emerge, similar to the spread of Windows NT servers over the past decade.

APPENDIX

TABLE A1: xSeries 440 and xSeries 360 Features

	IBM xSeries 440	IBM xSeries 360
Form Factor	• 4U Rack-Mount per Node (Up to Four Nodes)	• 3U Rack-Mount
Processor	• Up to Sixteen Intel Xeon MP (Two, Four, Eight CPUs per Node)	• One to Four Intel Xeon MP
Chipset	• IBM EXA XA-32 Chipset • Up to Four EXA-Cores (SMP Expansion Modules), Linked via Scalability Ports (Up to Two SMP Expansion Modules per Node)	• IBM EXA XA-32 Chipset • One EXA-Core
Chipset Cache	• 32 MB XceL4 Server Accelerator (Level 4) Cache per SMP Expansion Module	• None
Memory	• Dual-port Memory Controller • Four-Way Interleaved 133 MHz ECC SDRAM • 6.4 GB/Sec. Memory Bandwidth • Sixteen DIMM Slots per SMP Expansion Module	• Single-Port Memory Controller • Two-Way Interleaved 266 MHz DDR ECC SDRAM • 3.2 GB/Sec. Memory Bandwidth • Eight DIMM Slots
I/O	• Dual PCI-X Host-bridge Controller • Six Internal PCI-X Slots (Two x 133 MHz, Two x 100 MHz, Two x 66 MHz) per Node • Up to Twelve Additional PCI-X Slots with Optional Remote I/O	• Single PCI-X Host-Bridge Controller • Six Internal PCI-X Slots (Two x 100 MHz, Four x 66 MHz) • Up to Twelve Additional PCI-X Slots with Optional Remote I/O
Drive Bays	• Two Disks • Diskette • CD-ROM	• Three Disks • Diskette • CD-ROM
Integrated Controllers	• Dual-Channel Ultra3 SCSI Storage Controller • 10/100/1000 Ethernet Controller	• Single-Channel Ultra3 SCSI Storage Controller • 10/100 Ethernet Controller
RAS Features	• Hot-Swap/Redundant Power and Fans • Optional Internal RAID 0, 1 • Chipkill Memory, Memory ProteXion, and Memory Mirroring	• Hot-Swap/Redundant Power (Optional) and Fans • Optional Internal RAID 0, 1, 5 • Chipkill Memory
TPC-C Performance	• Four-Way: 55,138.60 tpmC @ $6.98/tpmC • Eight-Way: 92,398.49 tpmC @ $7.70/tpmC	• Two-Way: 23,027.66 tpmC @ $4.41/tpmC • Four-Way: 45,230.03 tpmC @ $4.52/tpmC