Enterprise X-Architecture Remote I/O and IBM @server xSeries Servers

by Mark T. Chapman
IBM Server Group

Executive Summary

One of the dilemmas confronting IT managers is how much system capacity to buy initially versus later. If you buy more server than you need—to ensure future growth capacity—you will have tied up capital for months or years that could have been better utilized elsewhere; you may even have spent money on hardware you will never fully use. On the other hand, if you buy too little up front you may find yourself buying additional servers to make up the shortfall, and then facing the challenges and costs associated with adding, housing and administering a series of small servers rather than one large one.

To address this problem, IBM developed Enterprise X-Architecture™ technology, which features modular, scalable servers and remote I/O, allowing you to "pay as you grow" so you don't need to buy excess system capacity up front. You buy only what you need, when you need it. IBM X-Architecture[1] technology is a blueprint for extending the benefits of advanced mainframe technologies to Intel® processor-based servers. These benefits are in the areas of availability, scalability, performance, systems management and service and support. The Enterprise X-Architecture design builds on this blueprint by adding a number of new enterprise-class capabilities to IBM @server xSeries systems.

This paper focuses on Enterprise X-Architecture features that are available today in new xSeries systems such as the x360; how these technologies work and how they can help you better meet today's business and IT challenges. These features include Active™ PCI-X I/O support, remote I/O and Real Time Diagnostics.

Introducing Remote I/O on IBM @server **xSeries Systems**

If you buy a server today and later your business needs require that you expand the capabilities of your server, what do you do? Typically your server will provide the option to add more processors, memory and disk capacity. If your need for disk space is great enough you might even resort to some sort of external storage, whether a direct-attached Ultra160 SCSI expansion unit, a Fibre Channel-based storage area network (SAN) or a network-attached storage (NAS) device. But what if your needs include additional *adapter slots?*

More slots may be required in the future than in the past because of the demanding, often business-critical types of environments that industry-standard servers now support. Consider the need for adapter slots driven by redundant Fibre Channel or SCSI RAID adapters with failover capability, SCSI adapters for tape or optical libraries, multiple Ethernet cards (often redundant), high-speed cluster interconnect adapters (often redundant with failover), perhaps InfiniBand™ interface adapters in the near future, and so on. These can quickly consume the limited number of adapter slots inside a typical compact, rack-optimized server.

Plus, as your business grows and the number of users increases more adapter slots are required. Slots can quickly become occupied by Ethernet adapters as more users are added to the network; typically an additional adapter (or multiple adapters for redundancy/failover and load-balancing) is needed for approximately every 250 e-mail users. In a terminal server environment, another adapter is required for every 50-75 users that are added. These examples illustrate the need for investment protection through I/O scalability. Until now your only recourse for more slots has been to purchase additional servers, even though your existing servers are sufficient in all other respects.

With the revolutionary advances of IBM Enterprise X-Architecture[2] technology, your servers no longer have to be limited by the number of adapter slots inside. With Enterprise X-Architecture

---

[1] See the white paper entitled "IBM X-Architecture Technology" at http://**ibm.com**/eserver/xseries for more information. From the xSeries home page, select **Library** for links to the different types of documentation available.

[2] Read "Introducing Enterprise X-Architecture Technology" at http://**ibm.com**/eserver/enterprise_xarchitecture for more

technology, if you need more I/O slots you simply plug in an external I/O expansion unit to double or even triple the number of available slots. Remote I/O is an extension of the PCI-X I/O subsystem implemented in selected xSeries servers such as the x360. In keeping with the IBM @server goals, remote I/O offers new ways to manage end-to-end growth, risks and costs, and provides unmatched scalability through innovative technology.

PCI-X I/O Subsystem

The PCI bus has done an admirable job of keeping up with the I/O bandwidth needs of servers. PCI started out with a theoretical limit of 132MBps (mega*bytes* per second)—although the early systems could deliver only 30-40 MBps. The current state of the art in PCI buses allows multiple 64-bit 66MHz bus segments capable of delivering 400-500 MBps each. But even this bandwidth is not sufficient for the emerging world of 10Gbps (giga*bits* per second)—or higher—I/O, including Fibre Channel, Ethernet and InfiniBand. Without another performance boost, PCI would soon be the bottleneck that keeps these high-speed networks from connecting to servers at the networks' maximum speed. I/O bottlenecks already prevent industry-standard servers from achieving the balanced performance that is characteristic of high-end systems.

Recognizing that the PCI bus is "running out of steam" as an I/O architecture and becoming a limitation to server performance, the industry developed an enhanced bus, called PCI-X, to extend the useful life of PCI until next-generation I/O architectures, such as InfiniBand, are ready to take over. PCI-X extends the useful life of PCI by increasing the available I/O bandwidth considerably.

IBM has been instrumental for many years in establishing industry standards, including the definition of the PCI-X specification. PCI-X provides a new generation of capabilities for the PCI bus, including more efficient data transfers, more adapters per bus segment and faster bus speeds for server systems. PCI-X enhances the PCI standard by doubling the throughput capability and providing new adapter-performance options while maintaining backward compatibility with PCI adapters. PCI-X technology allows all current 66MHz PCI adapters—either 32-bit or 64-bit—to operate normally on the PCI-X bus. PCI-X adapters take advantage of the new **100MHz** and **133MHz** bus speeds, which allow a single 64-bit adapter to move as much as **1GB** of data per second. (The next PCI-X specification [2.0] will support bus speeds of up to **266MHz**.) Additionally, PCI-X supports twice as many 66MHz 64-bit adapters in a single bus as PCI does.

Active PCI-X

Building upon the success of IBM Active™ PCI, Enterprise X-Architecture technology introduces Active PCI-X in selected xSeries servers such as the x360, with support for both PCI and PCI-X adapters. Active PCI-X helps to increase total server availability by letting you add or replace adapters while the system is running. Active PCI-X features can be categorized as follows:

- **Hot-swap** — Allows you to replace adapters without having to shut down and restart the server.

- **Hot-add** — Provides easy upgradeability, by allowing you to add new adapters to the server while it's running. (IBM was the first in the industry to offer this capability.)

- **Failover** — Allows a second—backup—adapter to pick up the workload automatically if the primary adapter fails.

Remote I/O

Until now, a limitation to the industry specifications for PCI and PCI-X has been that all adapter slots must be contained in the main system cabinet. This required a trade-off in server implementation. On one hand, designers want to maximize the number of adapter slots in a box. At the same time, they want to shrink the size of the server as much as possible as components

information.

become ever smaller, to minimize the rack space required. However, reducing the size of the system beyond a certain point means sacrificing adapter slots. (This is why server vendors offer a choice of models with many slots and few drive bays or many bays but few slots. It's either that or increase the overall size of the server—which causes other problems in a crowded data center.)

One obvious solution is to limit the number of slots inside the server chassis while extending additional bus segments to external "card cages," or I/O expansion units, to hold adapters. By moving adapter slots outside of the main system cabinet the server can be made much smaller. Unfortunately, the PCI-X bus specification doesn't explicitly support external I/O. As a result, no industry-standard server vendor has introduced this capability—until now.

Today, using IBM Enterprise X-Architecture remote I/O support, it is possible to add *dozens* of PCI/PCI-X adapter slots—via external I/O expansion boxes—to a single server (subject to the hardware limitations of the specific server), providing incredible I/O scalability at a very economical cost. Modular remote I/O is a feature of IBM XpandOnDemand™ scalability, which allows you to "pay as you grow." Buy only what you need, when you need it. It allows IBM to continue to shrink server cabinets while increasing I/O scalability via external expansion units.

Remote I/O not only permits more adapters to be used by a server (some in an external remote expansion enclosure), but it also offers potentially greater reliability. (Utilizing separate cabinets with their own power supplies and fans provides redundancy and may result in less heat build-up than having many adapters, processors and hard disk drives installed in the main unit.) Plus, using IBM Director, a remote I/O solution can be administered and managed as if all of the adapter slots are built into one server cabinet—something that cannot be done with a multiple server solution. Another key advantage of the IBM remote I/O enclosure on select servers is that *it may be shared by more than one server*[3].

And, these modular remote I/O enclosures connect to xSeries servers using the same time-tested, high-performance interconnect technology used on IBM @server pSeries and iSeries (formerly IBM RS/6000® and AS/400®, respectively) servers. Remote I/O allows you to expand your systems in an orderly, controlled fashion through the use of mainframe-inspired technologies now available on industry-standard xSeries servers.

IBM Enterprise X-Architecture technology's XpandOnDemand scalability allows you start out small and inexpensively with a single server, such as the x360. Later you can add an IBM RXE-100 Remote Expansion Enclosure[4] when needed (and as your budget allows) or set up a SAN. The RXE-100 ships with six PCI-X adapter slots and is expandable to 12 slots via an optional "six-pack" upgrade. (See *Figure 1* for configuration examples.)

This provides you with unmatched I/O flexibility. Look at the I/O configuration choices you have using an x360 server and an RXE-100 expansion unit:

- One stand-alone x360 server with **six PCI-X adapter slots**

- One **six**-slot RXE-100 attached to one **six-slot** x360 server (this effectively gives you one server with **12** PCI-X slots)

- One **12**-slot RXE-100 attached to one **six-slot** x360 server (producing one server with **18** PCI-X slots)

---

[3] Not all xSeries servers that support remote I/O also support the sharing of the remote I/O enclosures.

[4] For more information about the x360 server and the RXE-100 Remote Expansion Enclosure, refer to the individual spec sheets for those products.
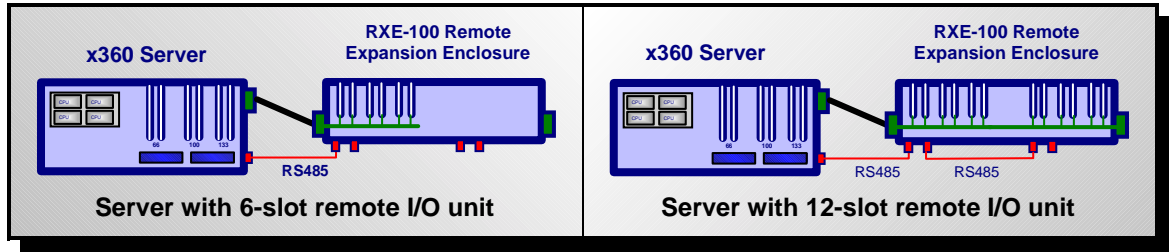
Figure 1. *Remote I/O configuration examples*

(**Note:** The heavy black lines shown in the illustrations are the remote I/O cables. These carry the data traffic between the server and the RXE-100. The RS485 cables are used to link the I/O six-packs to the systems management controller—the IBM Remote Supervisor Adapter—in the x360. The RXE-100 receives power-on and power-off signals from the x360 and returns hardware event signals to the Remote Supervisor Adapter.)

The x360 server is the *first* 4-way system offered in a slender **3U** rack-optimized form factor, and the RXE-100 is an equally rack-dense **3U**. Because you can start small, with just a single x360 server, you have the flexibility to expand considerably in a single 42U rack. For example, to describe just a few of the possible combinations, you could have:

- **Fourteen** x360 servers (giving you **84** adapter slots, **56** processors and **3.08TB**[5] of disk space in a rack)

- **Nine** servers and **five** RXE-100 Remote Expansion Enclosures (yielding up to **114** slots, **36** processors and **1.98TB** of disk storage)

- **Seven** servers and **seven** I/O expansion units (resulting in up to **126** slots, **28** processors and **1.54TB** of storage)

If you need additional disk capacity, the SAN-optimized x360 allows you to add one or more external modular hard disk drive storage units, (for example, IBM EXP300 Storage Expansion Units capable of holding **1.03TB** of Ultra160 SCSI storage per **3U** unit, or the various IBM FAStT [Fibre Array Storage Technology] SAN solutions[6]) and create a SAN for even more flexibility and growth potential:

- **Thirteen** x360 servers and **one** EXP300 (giving you **78** adapter slots, **52** processors and **3.89TB** of disk space in a rack)

- **Seven** servers, **five** RXE-100 Remote Expansion Enclosures and **two** EXP300 units (yielding up to **102** slots, **28** processors and **3.6TB** of disk storage)

- **Five** servers, **five** I/O expansion units and **four** EXP300 units (resulting in up to **90** slots, **20** processors and **5.21TB** of storage)

Because the RXE-100 offers a choice of **3.5m** (11.5 ft.) or **8m** (26.25 ft.) cables[7] to connect to the x360 servers, the RXE-100 units can be housed in a separate rack adjacent to the one containing the disk expansion enclosures and servers[8]. This provides you with even more flexibility in terms of rack configurability. The x360, the RXE-100 and the EXP300 are all 3U units, so you can mix and match any combination of 14 of these boxes in a rack.

---

[5] GB means 1,000,000,000 bytes and TB means 1,000,000,000,000 bytes when referring to storage capacity. Accessible capacity is less. Maximum hard disk drive capacity may require the replacement of the standard drives with the largest supported drive available.

[6] Go to http://**ibm.com**/pc/ww/eserver/xseries/fa_san.html for more information about IBM FAStT Fibre Channel storage products.

[7] A 3.5m cable ships with the RXE-100; the 8m cable is optional.

[8] This is the main purpose of the 8m cable.

These and other possible configurations give you tremendous capability to tailor your processor, I/O and disk capacity to your needs, and the flexibility to easily alter the configuration if your needs change. It also affords you the ability to buy only as much as you need today, then grow incrementally as your requirements increase. Contrast this with a high-end system where you must spend a lot of money up front for a single monolithic unit capable of holding a large, fixed number of adapters (assuming you could find one with enough slots to meet your future needs).

(*Figure 2a* shows a rear view of an RXE-100 with both six-packs installed atop an x360—note the terra cotta-colored levers signifying that the adapter slots support hot-add and hot-swap capability. *Figure 2b* focuses on the simple cable [with blue handles] connecting the two units. These design features promote simplicity, serviceability and flexibility.)
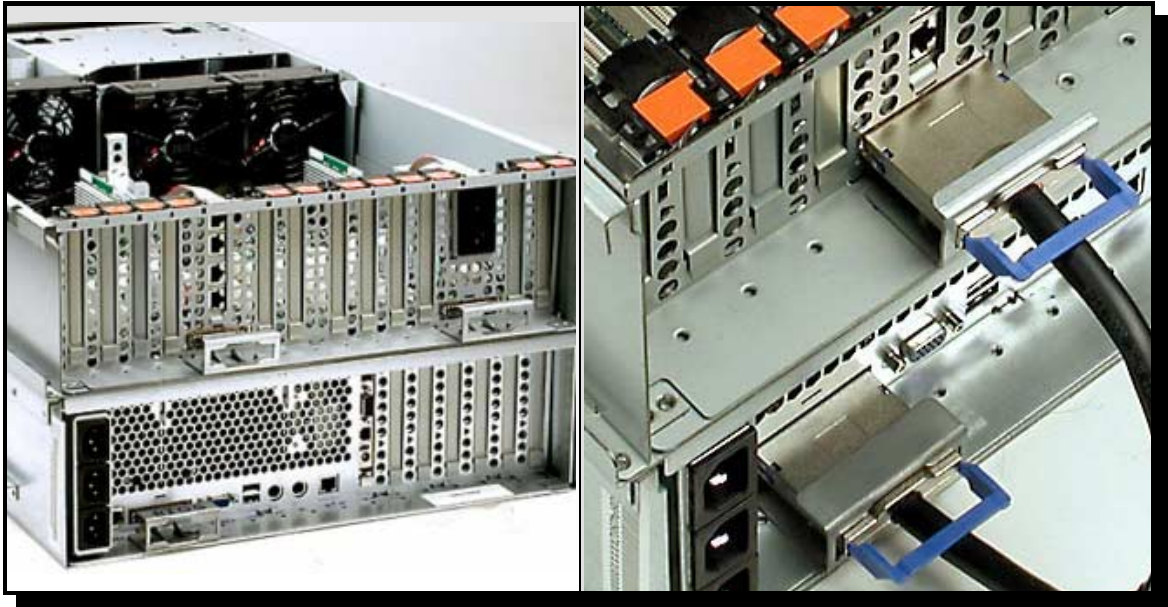
Figures 2a and 2b. *Details of RXE-100 (top) and x360 server (bottom)*

Not only does Enterprise X-Architecture remote I/O give you great expandability, it comes with amazing performance as well. The RXE-100 Remote Expansion Enclosure connects to the server using a dedicated high-speed I/O port. How fast is the port? Up to **2GBps** in bidirectional mode. (That's giga*bytes*, not giga*bits*.) 2GBps is roughly equivalent to **20 giga*bits*** per second (assuming eight bits per byte plus two bits for transmission overhead)—well beyond what even bidirectional Gigabit Ethernet offers today.

The x360 servers each include one remote I/O port (called an RXE Expansion Port), which means that each x360 can support one RXE-100 unit. Future servers may include two or more such ports, allowing not only the ability to attach multiple RXE-100 boxes but even to create a redundant configuration using one RXE-100 unit and one server. (This would entail connecting two remote I/O ports on a single server to two ports on one RXE-100 unit with both six-packs installed. One six-pack would provide failover backup for the other.)

Outstanding expandability isn't the whole story regarding the x360 rack-optimized server. The x360 utilizes the latest in server processor technology, the Intel Xeon™ Processor MP, with speeds up to 1.6 GHz. (In fact, Intel is using x360 servers as the validation platform for its new Xeon Processor MP to verify that the new processor can be deployed effectively in enterprise-class systems.) Along with the blazing processor speed, the x360 also implements a "quad-pumped" 400MHz front side bus (FSB). Before the x360, the fastest industry-standard servers had a front side bus speed of 133MHz (many servers still use a 100MHz FSB). This determines how fast the processor can access main memory. By contrast, Enterprise X-Architecture technology enables

servers such as the x360 to implement the 400MHz FSB of the Intel Xeon Processor MP. The x360 even supports one 133MHz 64-bit adapter—the first industry-standard server to do so. Having this capability provides investment protection. Many of the upcoming communications and I/O standards will require 133MHz adapter slots for maximum performance. Some of these technologies include: iSCSI and TCP/IP communications over 10Gb Ethernet, InfiniBand and next-generation Fibre Channel. But what if you need more than one 133MHz adapter? Once again, the RXE-100 comes to the rescue. Each *pair* of adapter slots in the first six-pack (plus one pair in the x360) supports either two 100MHz (or lower) adapters or one 133MHz card. If you need more 133MHz slots, add the second six-pack for a maximum of seven 133MHz slots. Combining the x360 with the RXE-100 gives you flexibility that no other industry-standard server in the world can offer today: either 14 100MHz slots and four 66MHz slots (in the x360), or seven 133MHz slots and four 66MHz slots or various combinations in between.

In addition, the x360 delivers on the Enterprise X-Architecture strategy of offering a significantly improved memory subsystem for higher reliability and performance. The x360 offers third generation IBM Chipkill™ memory, which stores the advanced error correction logic in the memory controller, allowing the use of low-cost, standard ECC DIMMs, yet provides as much as **100 times** the error correction of standard ECC memory. Furthermore, the x360 also supports up to 8GB of high-speed, 200MHz double data rate (DDR) main memory, for even higher performance.

For maximum reliability and serviceability, the x360 comes standard with an integrated IBM Remote Supervisor Adapter that works with IBM Director systems management software and IBM Predictive Failure Analysis®. Together they proactively monitor server components and the overall system environment. These advanced tools alert the administrator of impending hardware issues up to 48 hours in advance of actual component failure. Once an alert has been received, IBM Director triggers the Advanced Light Path Diagnostics™ feature, which leads service personnel directly to the component that requires service—without requiring anyone to pull the system out of the rack and remove the cover simply to identify the component. The use of many hot-add/hot-swap components means you may not even have to power down the system to service it. Two more innovations add to the overall system reliability: 1) Quad-pumped cooling units keep the processors from overheating: uniquely designed thermal pumps that cool the new Intel processors using self-contained liquid refrigerants in place of traditional heat sinks. 2) Zone system cooling keeps the rest of the system components cool: two fans for the memory and PCI-X slots, four to cool the processors and hard disk drive backplane, and an integrated cooling device in each hot-swap fan. These seemingly minor advancements are actually very important to a server. Heat is a major cause of component failure; less internal heat means greater reliability of system components.

In addition to the expandability, performance, reliability and serviceability benefits, there is also the potential cost savings. Say you buy a 4-way server with six adapter slots in it. One day you find you need more slots. You can't add more slots to that server, so you have to buy another server to accommodate more adapters, doubling your initial cost. Contrast this with the IBM XpandOnDemand remote I/O solution. First you buy a 4-way server, then when you need more slots, instead of buying a second server with all the processors, disk drives and other features you don't need (and shouldn't have to pay for), you buy a modular RXE-100 for thousands of dollars less—doubling and even tripling the adapter slots available to a single server.

Real Time Diagnostics

Expandability, performance and economy are all important features in a server, but equally important is your ability to prevent or minimize server downtime. A technology new to industry-standard servers that helps in that regard is IBM Real Time Diagnostics, another Enterprise X-Architecture feature first implemented in the x360 server. Based on the Distributed Management Task Force (DMTF) Common Information Model (CIM), Real Time Diagnostics allows an administrator to run diagnostics on system resources while users are still working, thereby increasing system uptime and taking IBM customers closer to OnForever™ levels of

availability. This is a capability that has been sorely lacking in the industry-standard server market, and this void is being filled with Enterprise X-Architecture technology.

IBM, working with Intel and PC-Doctor, introduced industry-standard extensions to CIM to support concurrent diagnostics in 1999. This Common Diagnostic Model (CDM) standardizes how to interface concurrent diagnostics (also called OS-Present Diagnostics or On-Line Diagnostics) with the operating system, making the diagnostic routines available to all CIM-aware diagnostic applications. With the diagnostic extensions to CIM now defined, IBM is working with independent hardware vendors to define methods and create concurrent diagnostic tools and utilities to keep xSeries servers up and running.

IBM Real Time Diagnostics is our implementation of CDM. The combination of IBM Predictive Failure Analysis (system health monitoring), Real Time Diagnostics and hot-swap components means that you may never again have to shut down your xSeries server to run diagnostics or replace hot-swap components. Real Time Diagnostics can be operated through the IBM Director management software to provide a consistent, easy to use management interface for controlling many system functions.

IBM Director

IBM Director[9] is a powerful suite of tools and utilities included with xSeries servers. Built upon industry standards, IBM Director is designed to manage servers in the Intel environment and support a variety of operating systems, including Microsoft® Windows®, IBM OS/2®, Novell NetWare, Linux® and SCO UnixWare. IBM Director supports a variety of industry standards such as DMI, CIM, WBEM, WMI, SNMP, MPM, TCP/IP, IPX, SNA, NetBIOS, SLIP, XML and HTTP, among others. IBM Director is being updated to support Enterprise X-Architecture capabilities, including remote I/O. IBM Director will gain the ability to configure and manage the RXE-100 Remote Expansion Enclosure. IBM Director consists of three major components: a management server, the management console and agent, plus an optional fourth component, Server Extensions. (*Figure 3* shows how the new IBM Director 3.1 console may look with the Enterprise X-Architecture enhancements.)
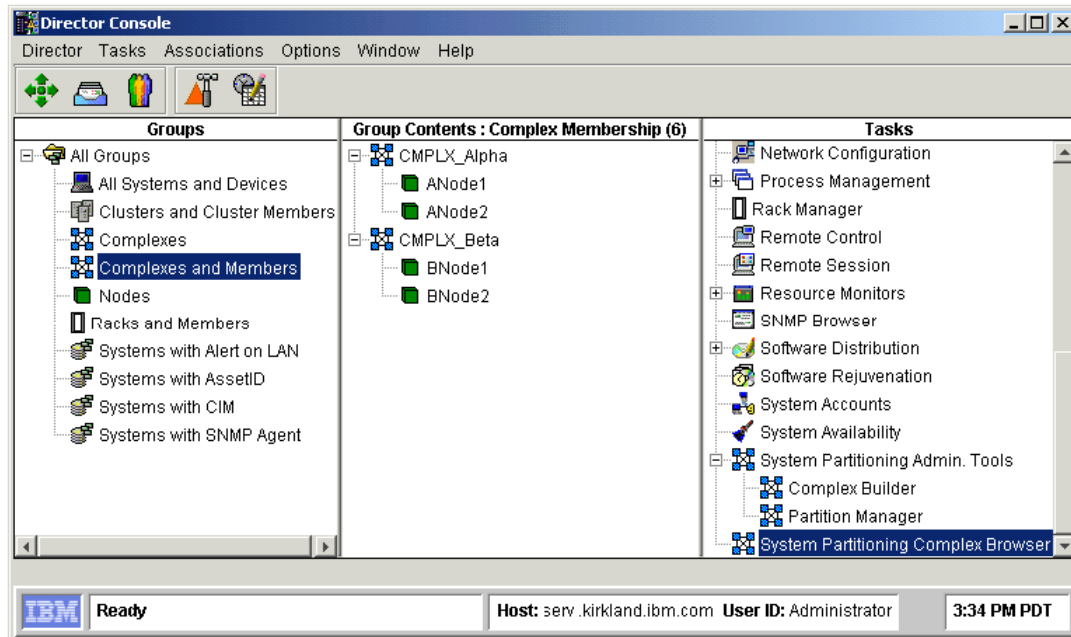

Figure 3. *IBM Director console*

---

9 For more on IBM Director, read the white paper "IBM Director with UM Services™" at
http://www5.pc.ibm.com/ww/me.nsf/Technical+Information%5cxSeries+and+Netfinity?OpenView&WW .

The IBM Director Server Extensions are often referred to as life cycle tools. These no-charge programs, included with xSeries systems, extend the manageability of your server hardware throughout its life cycle to help administrators configure, deploy, manage and maintain IBM @server xSeries servers efficiently. IBM Director Server Extensions include:

- Advanced Systems Management
- Capacity Manager
- Cluster Manager
- Rack Manager

- RAID Manager
- Software Rejuvenation[10]
- System Availability

How Enterprise X-Architecture Technologies Fit into Project eLiza

Real Time Diagnostics and IBM Director are just two of the technologies that go into the IBM Project eLiza initiative. Project eLiza is a multibillion dollar program aimed at creating self-managing servers that require little or no human interaction, with the goal of making computing networks as easy to manage as today's kitchen appliances. Four major categories serve as the foundation for Project eLiza technologies: self-optimizing, self-healing, self-configuring and self-protecting.

Project eLiza incorporates many IBM X-Architecture features, including the other high-availability aspects of the IBM OnForever initiative. Utilizing IBM @server xSeries servers, businesses now have access to a number of critical self-managing capabilities that help predict, identify and repair problems before availability is affected.

Visit http://**ibm.com**/servers/eserver/introducing/eliza for more information about Project eLiza.

Conclusion

Today, IBM Enterprise X-Architecture remote I/O technology offers unprecedented capabilities for pay-as-you-grow expandability using industry-standard servers.

Start small with an inexpensive server such as the IBM @server xSeries 360. As your I/O needs increase, add an IBM RXE-100 Remote Expansion Enclosure with six more adapter slots, doubling the total to 12 slots. If you eventually need more I/O, add an optional six-pack to increase your I/O capability to 18 PCI or new advanced PCI-X adapters. Where else but IBM can you find an industry-standard server capable of this sort of expansion?

If your disk storage needs also increase beyond the 220.2GB internal capacity of the x360, use one or more of those I/O slots to add Ultra160 SCSI RAID or Fibre Channel host bus adapters. From there, attach IBM EXP300 Storage Expansion Units, capable of holding more than a terabyte of disk storage apiece, or use an IBM FAStT Fibre Channel storage solution to connect as much as 16TB of storage.

Combining remote I/O, PCI-X support, a 400MHz front side bus and Real Time Diagnostics with the other capabilities offered by the IBM X-Architecture design (including Predictive Failure Analysis, Light Path Diagnostics and Advanced System Management processors, to name but a few) provides you with exceptional levels of reliability, availability, serviceability, performance, flexibility and scalability. And all at an affordable price. Buy only what you need today and add more later as the need arises.

---

[10] For more about IBM Director Software Rejuvenation, read the white paper by that name at http://www5.pc.**ibm.com**/ww/me.nsf/Technical+Information%5cxSeries+and+Netfinity?OpenView&WW.

These technologies all work together with other high-availability features of the IBM OnForever and new Project eLiza initiatives to help reduce downtime and decrease the costs of servicing and managing your servers.

Affordable, reliable servers with the kind of expandability previously only found in much more expensive (and often proprietary) servers—what more could you ask for?

Additional Information

Visit our Web site at http://**ibm.com**/eserver/xseries (or call **1-888-SHOPIBM**) for more information on IBM $@$server xSeries servers direction, products and services, including part numbers and prices for x360 servers, RXE-100 Remote Expansion Enclosures, IBM EXP300 Storage Expansion Units and IBM FAStT SAN products. From the xSeries home page, select **Library** for a list of the types of documentation available.