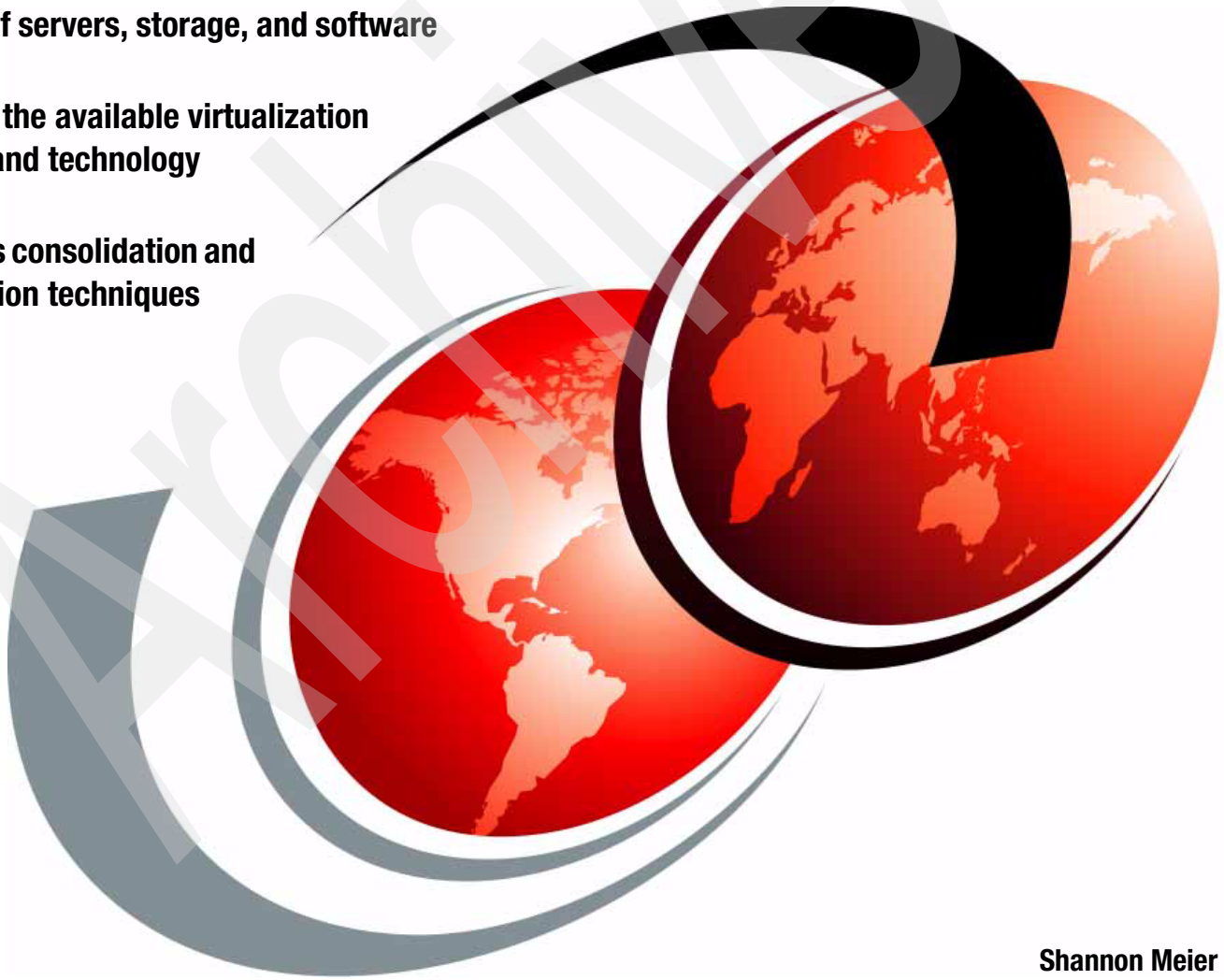
IBM

IBM Systems Virtualization: Servers, Storage, and Software

- Covers the complete IBM Systems platform of servers, storage, and software
- Describes the available virtualization products and technology
- Introduces consolidation and virtualization techniques



Shannon Meier

Redpaper

ibm.com/redbooks

International Technical Support Organization

IBM Systems Virtualization: Servers, Storage, and Software

April 2008

REDP-4396-00

Note: Before using this information and the product it supports, read the information in “Notices” on page vii.

First Edition (April 2008)

This edition applies to virtualization products and technologies on IBM System z, System p, System i, System x and BladeCenter servers, as well as IBM System storage and management software available at the time of writing.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

Redbooks (logo) ®
i5/OS®
z/Architecture®
z/OS®
z/VM®
z/VSE™
z9™
Active Memory™
AIX 5L™
AIX®
BladeCenter®
Calibrated Vectored Cooling™
Chipkill™
Cool Blue™
DB2®
DS4000™
DS6000™
DS8000™
Enterprise Storage Server®
Enterprise Workload Manager™
ESCON®
FlashCopy®
FICON®
Geographically Dispersed Parallel Sysplex™
GDPS®
HiperSockets™
IBM®
Micro-Partitioning™
NetView®
Parallel Sysplex®
PowerExecutive™
PowerVM™
Predictive Failure Analysis®
Processor Resource/Systems Manager™
POWER™
POWER Hypervisor™
POWER4™
POWER5™
POWER5+™
POWER6™
PR/SM™
Redbooks®
ServerProven®
ServicePac®
System i™
System i5™
System p™
System p5™
System x™
System z™
System z9®
System Storage™
System/360™
Tivoli®
TotalStorage®
Virtualization Engine™
WebSphere®
X-Architecture®

The following terms are trademarks of other companies:

SAP, and SAP logos are trademarks or registered trademarks of SAP AG in Germany and in several other countries.

Network Appliance, and the Network Appliance logo are trademarks or registered trademarks of Network Appliance, Inc. in the U.S. and other countries.

InfiniBand, and the InfiniBand design marks are trademarks and/or service marks of the InfiniBand Trade Association.

AMD, AMD Opteron, AMD-V, Direct Connect, HyperTransport, the AMD Arrow logo, and combinations thereof, are trademarks of Advanced Micro Devices, Inc.

Java, Solaris, Sun, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Hyper-V, Microsoft, SQL Server, Windows Server, Windows Vista, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Itanium, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Preface

Businesses are moving forward with server and storage virtualization projects and realizing the benefits. IBM® offers the industry's broadest set of virtualization capabilities. The cross-platform virtualization, automation, and systems management solutions available from IBM enable customers to access and manage resources simply and dynamically for better asset utilization and reduced operating costs.

This paper serves as both an introduction to virtualization, as well as an overview of pertinent IBM hardware and software virtualization offerings. We first introduce the concepts of virtualization and the benefits of virtualizing your systems. We then describe virtualization options for each of the IBM Systems platforms as well as software and storage technologies that are used to implement virtualization.

This paper is suitable for people who want to expand their knowledge of virtualization and what IBM can offer with its systems and software.

The team that wrote this paper

This paper was produced by a team of specialists from IBM.

Shannon Meier is a Systems Architect in the East Region supporting the IBM Communications Sector. Shannon has been working on cross-platform hardware solutions as a Systems Architect for four years and was previously a mainframe hardware technical specialist.

Additional contributions by:

Bill Virun is a member of the IBM Tivoli® Business Development team. He has 20+ years of engineering management experience in IBM Research and Development in both hardware and software development engineering.

Joshua Blumert is a System x™ Technical Specialist in IBM East Region, focusing on Linux® on x86 and blade servers. Josh has six years of experience at IBM and 12 years at Silicon Graphics previously.

M. Tim Jones is an embedded software architect and the author of several books, including *GNU/Linux Application Programming*. His engineering background ranges from the development of kernels for geosynchronous spacecraft to embedded systems architecture and networking protocols development. Tim is a Consultant Engineer for Emulex Corp. in Longmont, Colorado.

Thanks to the following people for their contributions to this project:

- Jack Alford
- Catherine Brohaugh
- Mike Ebbers
- Bob Haimowitz
- Bart Jacob
- Maria Jenny
- Kevin Lambert
- Annette Miller

- Tony Pearson
- Michael Schuckman
- Scott Vetter
- David Watts
- Lynn Winkelbauer

Become a published author

Join us for a two- to six-week residency program! Help write a book dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You will have the opportunity to team with IBM technical professionals, Business Partners, and Clients.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you will develop a network of contacts in IBM development labs, and increase your productivity and marketability.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this paper or other IBM Redbooks® in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an e-mail to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

1

Executive overview

This paper serves as both an introduction to virtualization, as well as an overview of pertinent IBM hardware and software virtualization offerings.

This paper is organized into three main areas:

- An introduction to virtualization and its benefits
- An overview of the hardware technologies that use virtualization
- An overview of the software technologies that enable and manage virtualization

Business and IT management professionals want to deliver new and enhanced business services in less time and with greater agility. The goal of providing rapid near real-time response to customer requirements is in part being achieved by taking advantage of *virtualization*. Virtualization provides many benefits, including improved physical resource utilization, improved hardware efficiency, and reduced power and cooling expenses; however, it introduces another set of systems (*virtual machines*) that have to be controlled, provisioned, managed, updated, patched, and retired.

Management of this new virtualized environment should be considered carefully to help reduce the burden of the operations team. IT management is embracing virtualization to deliver greater value to the business. Virtualization of the environment delivers benefits that enable IT management to do more with less. However, the virtualized IT environment can become more complex in the process and must be managed.

Consolidation of resources obtained through virtualization offers compelling return on investments for the business. Businesses are moving forward with server and storage virtualization projects and are realizing the financial benefits. Virtualization should be a key initiative for many reasons, including:

- Environmental concerns

  Many customers are facing similar issues. The organic growth of application environments has caused servers to proliferate from one or two servers to dozens or even hundreds of servers and beyond. This proliferation creates a real challenge in the data center when dealing with the power, cooling, and physical space. Virtualization allows for consolidation and a more effective use of resources.

- Infrastructure simplification

  Server sprawl can create an unwieldy environment. Virtualization enables you to simplify your physical infrastructure. Virtualizing enables you to get more use out of existing assets by sharing assets that were previously dedicated, and utilizing previously wasted white space. This makes them more effective and easier to manage.

  For example, being able to share I/O adapters between logical partitions on a server means that you need fewer adapters, which means that you need fewer cables, which also means that you need fewer ports on your switches. Virtualization allows you to consolidate your existing environment and also to contain and manage future growth.

- Improved manageability

  When implemented correctly, virtualization (recommended in conjunction with automation) can increase the manageability of your infrastructure. IBM has been developing cross-platform tools using IBM Systems Director and IBM Tivoli software to help manage heterogeneous infrastructures with a common set of tools. The use of cross-platform tools begins the process of building a single pane of glass view into your environment.

  An easier to manage environment can affect the number of people that are required to run the environment. Customers can see a reduction in the support staff that is required, control the trend of increasing support staff to manage the growing physical environment, and free up people to work on other projects (such as those that drive business revenue versus expenses).

- Improved responsiveness

  Virtualization can enable you to improve business responsiveness and operational speed by re-allocating resources dynamically to applications as needed. This re-allocation helps better match changing business cycles or handle unexpected surge in demand. This is possible because you have decoupled the physical from the virtual, thus providing an abstraction layer and flexibility. Faster application deployments and flexibility can be seen as a business advantage where, for example, being first to market with a new service can be key to success.

- Business resiliency

  Virtualization can help you meet your business challenges through an available and resilient infrastructure—from mitigating planned or unplanned outages to full disaster recovery. Virtualization can help:

  - Increase the availability of your applications and workloads
  - Insulate users from system failures
  - Lower the cost of disaster recovery

- Total cost of ownership

  Virtualization enables you to simplify IT infrastructure and management and help workloads become independent of hardware resources. This enables clients to make business-driven policies to deliver resources based on time, cost, and service-level requirements. Making better use of IT assets in support of business goals enables you to reduce total cost of ownership.

For a virtualization solution to be effective, it needs to be flexible, dynamic, and secure, and it needs to integrate easily within a heterogeneous infrastructure. An effective virtualization solution allows you to maximize the utilization of your hardware, reduce the cost of your hardware through shared resources, and deploy new instances quickly without necessarily requiring additional hardware.

IBM recognized the importance of virtualization in the 1960s with the development of the System/360™ Model 67 mainframe and has continued to innovate ever since. Where possible, IBM takes advantage of proven investments in advanced server technologies and

skills. This strategy allows greater capability for across server lines without having to reinvent technologies and methods. By using the proven capabilities of other server families, IBM shortens development cycles, reduces development costs, and increases interoperability for your complex IT environments. When IBM designs a server, the goal is always to get as close as possible to a platform that maintains continuous operation.

The cost of managing the IT infrastructure has become the largest and fastest-growing component of overall IT spending for many organizations. Virtualization helps address this cost through the consolidation of physical resources. However, companies must be careful not to add complexity by creating a sharp increase in the number of managed virtual resources.

IBM Systems Director helps address these needs by unifying its industry-leading server and storage management products under one family—IBM Systems Director and TotalStorage® Productivity Center. The IBM Systems Director family provides a modular, open standards-based set of solutions that can be tailored easily to fit the requirements of any size business and can be integrated seamlessly into enterprise management solutions from IBM Tivoli and others. With capabilities that include configuration, discovery, health and status monitoring, automated response, and power and virtualization management, the IBM Systems Director family gives IT professionals what they need to manage both physical and virtual systems throughout multiple IT environments. IBM Systems Director can work in partnership with software management solutions from IBM, provided through the IBM Tivoli suite of offerings.

IBM Tivoli Service Management enables clients to better manage their infrastructure, operations and IT processes, to deliver services aligned to business goals more effectively. IBM software addresses the management challenges of a virtualized IT environment to provide the ability to embrace virtualization and strive to achieve your business' various service level agreements. Business service level agreement breaches are minimized using IBM best practices and IBM Tivoli service management solutions that are designed to manage the complexities introduced by virtualization.

The IBM model not only simplifies the way an IT organization views resources, it also affects the way resources are allocated and added in response to change. The ability to plan system resources in flexible, granular increments combined with the notion that capacity is a variable object that can be altered in real time rather than a fixed object, changes how customers can respond to business needs.

IBM is uniquely positioned to offer both a breadth and depth of virtualization offerings. No other vendor can bring together the virtualization solutions (infrastructure simplification, rapid application deployment, and business resiliency), server, storage and application virtualization, and cross-platform management in a heterogeneous environment. IBM virtualization balances the needs for flexibility and isolation in a virtualized environment, while simplifying management so that customers can manage the ever-increasing server and storage requirements while maintaining or reducing current staff.

2

Virtualization benefits

Virtualization helps you take back control of your infrastructure. Virtualization enables you to see and manage your computing resources in ways that offer more flexibility because you are not restricted by implementation, location, or physical packaging. With virtualization, you have a *logical* rather than a physical view of data, computing power, storage capacity, and other resources. By gaining greater control of your infrastructure, you can improve cost management.

Businesses are pursuing financial savings through both server and storage consolidation. The consolidation is achieved using virtualization. Virtualization is the abstraction of the physical resource into a virtual resource that is decoupled from the underlying hardware. Consolidation of server and storage hardware using virtualization offers return on investment for the business.

While cost saving is a primary driver for initial virtualization deployment, the full value of virtualization lies in its ability to offer:

- Improved total cost of ownership (TCO)

  By decreasing management costs and increasing asset utilization, you can experience a rapid return on investment (ROI) with virtualization. In addition, by virtualizing resources and making them easier to migrate or fail over to other physical devices or locations, you can enhance system availability and help lower the cost and complexity of disaster-recovery solutions.

- Increased flexibility

  Virtualization supports the pooling of resources that can be managed centrally through an enterprise hub to better support changing business requirements dynamically.

- Enabled access through shared infrastructure

  Virtualization provides a resilient foundation and shared infrastructure that enables better access to infrastructure and information in support of business applications and service-oriented architectures (SOA).

Companies of all sizes are aggressively adopting IBM virtualization solutions to help with:

- Infrastructure simplification

  Virtualization can help control infrastructure sprawl through the deployment of virtual servers and storage that run securely across a shared hardware environment. Virtualization not only helps with server consolidation, but also server containment when deploying new systems. Consolidating to a virtual infrastructure can enable you to increase server utilization rates from 5% to 15% to over 70%, thus helping improve ROI. In addition, a simplified infrastructure can help lower management costs with a common management platform and tooling.

- Rapid application deployment

  Virtualization can help enable rapid infrastructure provisioning (ex. minutes compared to days). It can help developers to speed application test and deployment, enhance collaboration and improve access to the infrastructure. The ease and flexibility of creating and reconfiguring guest operating systems means that development and test environments can realize significant benefits from virtualization. For example, you can use Dynamic Logical Partitions (LPARs) on System p™ in a shared development environment where applications can reside in logically separate OS environments, but on shared hardware. Partitions can be expanded dynamically for load testing and contracted dynamically when testing is complete. Thus, you can maximize the investment in the environment and quickly make changes based on demands and business priorities.

- Business resiliency

  Virtualization can help IT managers secure and isolate application workloads and data within virtual servers and storage devices for easier replication and restoration. This added resiliency can provide IT managers with greater flexibility to maintain a highly available infrastructure while performing planned maintenance, and to configure low-cost disaster-recovery solutions.

  Virtualization technologies solve many traditional backup issues, because they decouple the bindings between the operating system (with the application and data) and the underlying hardware. For example, you could have a different hardware topology in the recovery site, both in terms of numbers of servers and configuration of those, and still be able to boot all your guests on the two different data centers.

- Managing a virtualized infrastructure

  The IBM Systems Director and IBM Tivoli Service Management offers the right systems management platform and common tools to support both virtual and physical devices. IT managers can address configuration, deployment, monitoring, workload management and additional management functions in a consistent and common way across their infrastructures. This can help simplify problem determination, increase productivity and lower management costs.

IBM takes a holistic approach to virtualization by working across all resource types, taking advantage of decades of mainframe experience, embracing diversity of resources and integrating the virtual and physical worlds. We can help you virtualize not only your broad physical infrastructure, but your application workloads and information as well.

Virtualization allows you to freely mix and match technologies through common management tools for managing distributed heterogeneous resources. This added freedom offers capabilities to lower switching costs, add flexibility and freedom of choice, and mask complexity. Not having to manage each computer or resource separately—but instead managing them together virtually—allows for significant improvements in utilization and administrative costs.

Virtualization allows for easier maintenance, more reliability, and improved problem determination. Because you can now create virtual machines by simply copying a disk from a given template into a production disk, reliability and ease of maintenance of that guest operating system increases. In the past, administrators were used to installing operating systems on different hardware platforms with different disk controllers, different network cards, and so forth, thus increasing the complexity of software stack and drivers being installed. Because one of the key advantages of virtualization is that of shielding the real hardware being used, it provides a consistent set of virtual hardware interfaces that get exposed to the guest operating system.

3

Consolidation using virtualization

Businesses are pursuing financial savings through both server and storage consolidation. This consolidation is achieved using *virtualization*. The distributed computing environment in many data centers is comprised of numerous under-utilized servers. The ability to reduce floor space, power consumption, heat generation, air conditioning costs, and the total number of physical servers that are managed easily supports the server and storage consolidation business plans.

This chapter discusses the concept of *consolidation* using virtualization. In this chapter, we discuss the following topics:

3.1 Server consolidation

The process of *server consolidation* requires up front planning to decide which physical servers along with their associated workloads can be consolidated together. The workload performance characteristics are considered for consolidation compatibility along with business considerations. The physical servers are transformed into virtual machines, and then compatible virtual server machines are consolidated onto a reduced number of physical servers hosting the virtual environment for consolidation.

The utilization of the physical server hosting the consolidated virtual machines increases as additional virtual machines are hosted on it. The former physical servers can be turned off and removed from the data center, delivering on the business goals of the server consolidation effort. This new environment will need to be managed as more and more physical servers are virtualized and consolidated.

The need to discover and inventory the virtual machine assets in the virtual environment will emerge as the data center progresses forward with physical to virtual server conversions. Full visibility into the business services, including batch and online workloads, provided by the underlying IT infrastructure that supports these services, is important. To manage these services properly, it is critical to see and understand the make up of these business services. This is where IBM Tivoli Service Management comes into play, which we discuss in Chapter 9, “Systems and virtualization management” on page 67.

You must consider the capabilities of the platform that you have chosen for consolidation. Where possible, IBM takes advantage of proven investments in advanced server technologies and skills. This strategy allows greater capability across server lines without having to reinvent technologies and methods. By using the proven capabilities of other server families, IBM shortens development cycles, reduces development costs, and increases interoperability for your complex IT environments.

For example, the mainframe team that developed logical partitions (LPARs) in the 1980s on the mainframe also worked on the POWER™ Hypervisor that you see on IBM Power Systems platform today. The mainframe team works on key hardware designs across IBM platforms. For example, the team that worked on the mainframe cooling systems also designed the Calibrated Vectored Cooling™ feature that you see in IBM System x and IBM BladeCenter® today. This innovative teamwork allows IBM to bring more efficient, higher quality products to market.

Running more of your applications on shared hardware makes reliability, availability, and serviceability (RAS) even more important. IBM uses skills and technologies from the mainframe and incorporates them across product lines to increase RAS. IBM Power Servers are measured at world class levels of hardware availability. Smart system design contributes greatly to this. For example:

- The POWER Hypervisor™ runs as part of the firmware. Moving and restricting virtualization capabilities to the hardware allows you to improve efficiency as well as security.
- IBM System x and IBM BladeCenter incorporate mainframe RAS capabilities such IBM Chipkill™ memory, which allows correction of multiple, single-bit errors using DIMMs. These types of capabilities are critical when supporting real-time applications in a shared environment.

When IBM designs a server, the goal is always to get as close as possible to a platform that maintains continuous operation. Building highly available servers takes more than just using reliable components or redundant disk drives, fans, and power supplies. Naturally, these

things are done; however, it is important to address both causes of downtime-scheduled and unscheduled outages-to truly improve operational efficiency. That is why serviceability (for example, performing real time nondisruptive service actions such as installing and replacing components) is also an important factor. For example, IBM System p servers offer dynamic processor sparing so that a processor failure is transparent to the application.

3.2 Storage consolidation

The amount of data and information being generated by businesses continues to grow. The IT data center manager must deal with this high rate of growth, while at the same time look for ways to reduce costs. *Storage consolidation* helps the data center manager deal with the rapid growth and costs concerns. Increasing the utilization of the storage hardware, similar to what was discussed for the server hardware, is cost-effective and helps meet the growing demand. Storage consolidation is the allocation or provisioning of shared storage resources. Shared storage is connected to the servers using Fibre Channel or IP-based networks. This type of storage is called *Storage Area Networks* (SAN).

Storage virtualization software, similar in concept to server virtualization, abstracts the storage hardware volumes of data into a logical or virtual view of the volume. Using SAN hardware with storage virtualization gives the data center a method to support storage provisioning in a manner that is independent of the underlying storage hardware. Storage virtualization can enable data sharing, data tiering, improved storage hardware utilization, improved availability, and disaster recovery capabilities. Storage virtualization software separates the representation of the storage to the operating system from the physical device. Utilization rates of storage are likely to be improved when moving towards network based storage that is virtualized.

3.3 Network virtualization

If physical server farms are consolidated into virtual server farms, then parts of the physical network can be replaced by a virtual network, saving money and reducing management complexity. Network performance and bandwidth between the servers is increased, enabling new data-intensive applications. We do not discuss *network virtualization* in detail in this paper, but in this section, we give a brief overview of the topic along with a highlight of the various technologies within the platform-specific topics.

There are business-critical application requirements to manage and utilize network resources more efficiently with regard to performance, resource usage, people cost, availability, and security. Network virtualization includes the ability to manage and control portions of a network that can even be shared among different enterprises, as individual or virtual networks, while maintaining isolation of traffic and resource utilization.

Network virtualization includes technologies such as Virtual Private Networks (VPNs), HiperSockets™, Virtual Networks, and VLANs. It also includes the ability to prioritize traffic across the network to ensure the best performance for business-critical applications and processes. Instrumentation of network resources and operations (such as SNMP), which can be abstracted across the server and networking devices, are key enablers for on demand behavior.

Network virtualization technologies include:

- Virtual IP address (VIPA) takeover

  Addresses abstract the physical connection of servers to networks. VIPA takeover allows for automatic recovery of network connections between different servers.

- HiperSockets

  A mainframe-based technology that allow any-to-any TCP/IP network connection between virtual servers. This technology provides secure IP communication at memory speed between virtual servers, thus creating a base for closer integration of applications and the implementation of new data-intensive applications.

- Virtual Ethernet

  Power technology that enables internal TCP/IP communication between partitions.

- Virtual Fabric Architecture

  In 2004, IBM and Cisco announced VFrame on BladeCenter. VFrame virtualizes Fibre Channel SAN and Ethernet network on a single physical interface using InfiniBand® technology.

- Virtual LANs (VLANs)

  A commonly used standards-based technique in which physical networks are shared in a secure way between multiple applications or user groups.

- Virtual Private Networks (VPNs)

  A commonly used standards-based technique that encrypts data between two TCP/IP endpoints to provide end-to-end physical security for the transport.

It is important to note that much of the virtualization is at the platform layer, requiring support in the Hypervisors and, in some cases, in the firmware, to enable sharing between different operating systems (in addition to the sharing that is provided in the operating system for functions, such as VLAN, QoS, and VPNs). Also note that integrating the management of these networking resources in the context of the applications they support and the servers on which the applications run begins to reduce the complexity of managing servers and networks separately. Examples include IBM Systems Director support for Cisco and Nortel blades, and IBM Tivoli Provisioning Manager support for configuring server network resources, such as adapters, IP addresses, and VLANs.

The platform is also involved in provisioning and configuration, performance management, security, and availability services. Functions such as Intrusion Detection filters can also be integrated in some operating systems. The use of policy to abstract the management of these services is an element of the network virtualization support, which is being enhanced to better integrate with the other resources (server/storage) and exploited by more of the service-level manager disciplines.

What we are driving toward is the notion of life cycle management of resources, where the resources are provisioned, monitored, and managed according to business goals using standards-based instrumentation and operations to reduce the complexity for customers in managing these disparate sets of resources from different vendors. Abstraction of physical interfaces and the support for sharing them across different virtual servers are the key concepts which network virtualization addresses.

3.4 Application virtualization

We do not discuss *application virtualization* in detail in this paper, but we give a brief overview here on IBM WebSphere® Extended Deployment.[1]

The need to apply the principles of virtualization to software management led IBM to create WebSphere Extended Deployment (WebSphere XD), a powerful tool rooted in application-focused virtualization. WebSphere XD works at the business application layer in an array of heterogeneous servers or a server farm of homogeneous servers and allows the user to view infrastructure resources as a single, consistent entity.

WebSphere XD reduces the potential for bottlenecks when various server types are in use and manages the workload at the application layer dynamically so that applications are started and stopped at the proper time and in the desired priority. In this way, the strong workload management capabilities of WebSphere XD enable the business to meet its established service policies.

From the business perspective, WebSphere XD presents a compelling value proposition by accomplishing three important goals:

- Business service level objectives are met and workload throughput is increased.
- Utilization of hardware resources is maximized, often resulting in a consolidation of servers.
- The IT infrastructure can support new, innovative applications and types of workloads.

WebSphere Extended Deployment contains the following components, which can be purchased together as a single solution or as individual components. The WebSphere Extended Deployment, Version 6.1 components include:

- Operations Optimization

  Provides application virtualization, centralized workload management and administration for Java™ applications running on WebSphere and other middleware application servers
- Data Grid

  Contains high-performance features enabling applications to process massive data volumes efficiently
- Compute Grid

  Enables the scheduling, execution and monitoring of batch type jobs with enhanced service policy and workload management

Through centralized workload management, application virtualization, and management of large data volumes WebSphere Extended Deployment delivers enhanced qualities of service. WebSphere XD provides a more flexible infrastructure for your SOA. Application resources are allocated dynamically in response to changes in workload demand, based on customer-defined business goals. WebSphere Extended Deployment enables more efficient resource utilization as well as improved application performance and scalability.

[1] For more information about WebSphere XD 6.1, see Announcement Letter 207-088, “IBM WebSphere Extended Deployment V6.1 delivers workload management, application virtualization, and innovative application pattern support”, which is available at:
http://www.ibm.com/isource/cgi-bin/goto?it=usa_annred&on=207-088

4

Server virtualization techniques

With server virtualization, you can create multiple virtual servers on a single physical server. Each virtual server has its own set of virtual hardware on which operating systems and applications are loaded. IBM systems with virtualization can prioritize system resources and allocate them dynamically to the virtual servers that need them most at any given time—all based on business priorities.

Virtualization was first introduced by IBM in the 1960s to allow the partitioning of large mainframe environments. IBM has continued to innovate around server virtualization and has extended it from the mainframe to the IBM Power Systems, IBM System p, and IBM System i™ product lines. In the industry-standard environment, VMware, Microsoft® Virtual Server, and Xen offerings are available for IBM System x and IBM BladeCenter systems. Today, IBM server virtualization technologies are at the forefront in helping businesses with consolidation, cost management, and business resiliency.

IBM recognized the importance of virtualization with the development of the System/360 Model 67 mainframe. The Model 67 virtualized all of the hardware interfaces through the Virtual Machine Monitor, or VMM. In the early days of computing, the operating system was called the *supervisor*. With the ability to run operating systems on other operating systems, the term *hypervisor* resulted (a term coined in the 1970s). Logical partitioning has been available on the mainframe since the 1980s. The Power team began taking advantage of the mainframe partitioning skills and knowledge about 10 years ago and brought forth Dynamic LPARs with POWER4™ and then Advanced POWER Virtualization with POWER5™ in 2004 (which was re-branded to PowerVM™ in 2008).

There are several types of virtualization.[1] In this chapter, we describe them in order to position the relative strengths of each and relate them to the systems virtualization offerings from IBM and IBM Business Partners.

[1] This section includes virtualization techniques as described by Tim Young in IBM DeveloperWorks article “Virtual Linux: An overview of virtualization methods, architectures, and implementations”, December 2006.

4.1 Full virtualization

Full virtualization, otherwise known as *native virtualization*, is an interesting method of virtualization. This model uses a virtual machine that mediates between the guest operating systems and the native hardware (see Figure 4-1). *Mediate* is the key word here, because the VMM mediates between the guest operating systems and the bare hardware. Certain protected instructions must be trapped and handled within the hypervisor because the underlying hardware is not owned by an operating system but, instead, is shared by it through the hypervisor.

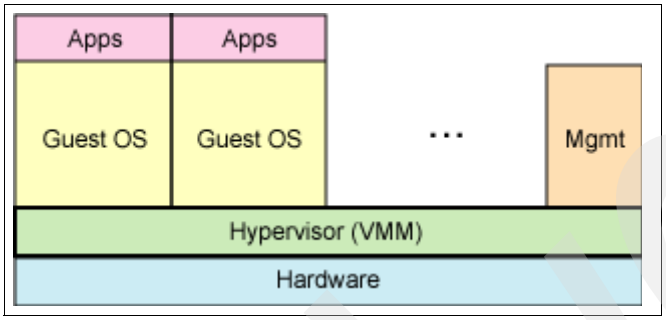


Figure 4-1 Full virtualization uses a hypervisor to share the underlying hardware

Full virtualization is faster than hardware emulation, but performance is less than bare hardware because of the hypervisor mediation. The biggest advantage of full virtualization is that an operating system can run unmodified. The only constraint is that the operating system must support the underlying hardware.

Some hardware create problems for the full method of virtualization. For example, certain sensitive instructions that need to be handled by the VMM do not trap. Therefore, hypervisors must scan and trap privileged-mode code dynamically to handle this problem. VMware ESX Server and Microsoft Virtual Server utilize full virtualization. Overhead is more substantial in this type of environment because it requires translation, trapping, and emulation of certain instructions. Intel® and AMD™ are starting to provide hardware support for virtualization through Intel VT and AMD-V™, thereby reducing the overhead.

4.2 Paravirtualization

Paravirtualization is another popular technique that has some similarities to full virtualization. This method uses a hypervisor for shared access to the underlying hardware but integrates virtualization-aware code into the operating system itself (see Figure 4-2). This approach obviates the need for any recompilation or trapping because the operating systems themselves cooperate in the virtualization process.

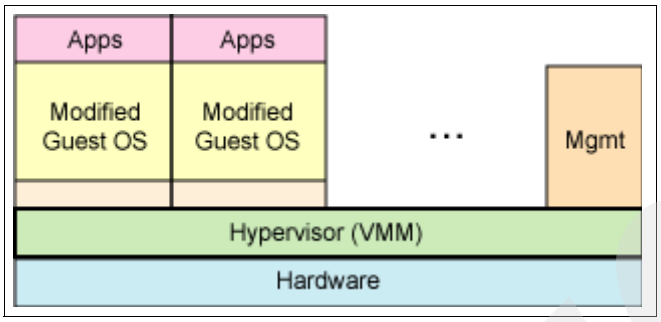


Figure 4-2 Paravirtualization shares the process with the guest operating system

As mentioned, paravirtualization requires the guest operating systems to be modified for the hypervisor. This is a disadvantage for operating systems that do not already incorporate those changes. Power architecture utilizes paravirtualization, and the necessary changes have been made to the AIX® and Linux (SUSE Linux Enterprise Server and Red Hat Enterprise Linux) operating systems. Xen also takes advantage of paravirtualization. Paravirtualization offers performance near that of an unvirtualized system. Similar to full virtualization, multiple different operating systems can be supported concurrently.

4.3 Operating system-level virtualization

Operating system-level virtualization uses a different technique than those covered so far. This technique virtualizes servers on top of the operating system itself. This method supports a single operating system and simply isolates the independent servers from one another (see Figure 4-3).

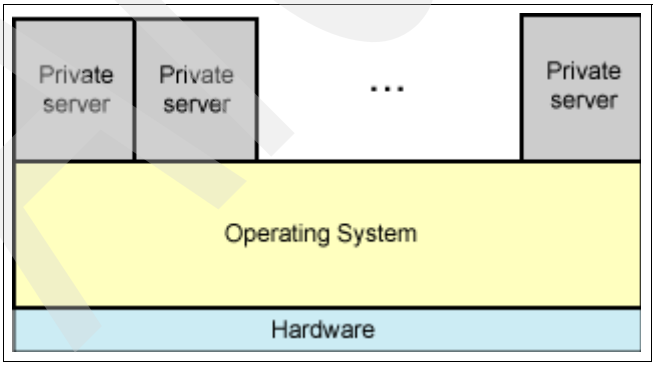


Figure 4-3 Operating system-level virtualization isolates servers

This approach is available through a few methods. IBM WebSphere Application Server Extended Deployment offers this type of environment. Other vendors also provide this type of environment through an application container approach or through a fat client where virtual operating system instances reside under a master operating system instance.

Operating system-level virtualization requires changes to the operating system kernel, but the advantage is native performance. The disadvantage to this approach is that each of these virtual instances is tied to the master operating system instance. That is, they have to be at the same operating system and patch levels, and if the master operating system is brought down, all of the virtual environments come down with it. The fat client approach can also add a significant amount of overhead to the system.

5

System z virtualization

From simple multitasking to logical partitions to complete simulation of virtual hardware, the virtualization capabilities of the IBM mainframe continues to represent some of the most mature and sophisticated virtualization technology in the industry today.

During spikes in demand, the IBM mainframe's ability to redistribute system resources quickly and scale up or out can make the difference between flawless execution and the cost of slow response times or system crashes. For example, a single IBM System z™ mainframe can scale up to millions of transactions per day or scale out to manage tens to hundreds of virtual servers. It can also redistribute system resources dynamically to manage varying server demands on the systems resources autonomically.

We discuss the following topics in this chapter:

5.1 Introduction

The System z platforms are designed to handle the demanding requirements of business critical core business systems. Business critical core business systems are characterized by the need to:

- Support high volume, unpredictable, dynamically changing, and competing workloads for multiple application systems
- Provide the maximum in high volume, complex, and data intensive I/O performance and capacity
- Provide the maximum in system availability, disaster recovery, and security

System z architecture is designed with a larger cache, and more advanced reliability, availability, and serviceability (RAS) features than other platforms in order to support these demanding high volume or virtualized workloads. System z multi-dimensional virtualization technology consists of both hardware and software that work together seamlessly as part of the System z architecture. System z hardware provides a robust and reliable foundation which supports a powerful and flexible software layer.

IBM introduced the first virtualization technology, the virtual machine (VM) Hypervisor, in 1968. Today's version, z/VM®, provides the ability to create tens to hundreds of virtual machines or virtual systems and continues to evolve. In 1988, IBM introduced Processor Resource/System Manager (PR/SM™) technology, which lets you create multiple logical partitions (LPARs), each of which, like z/VM, allows a full-scale OS to operate concurrently on the same physical mainframe footprint. PR/SM is fundamental to all System z mainframes, providing a highly stable and security-rich, firmware-encapsulated virtualization technology adjunct to z/VM, which is a full-scale hypervisor.

IBM mainframes have a history of hardware and software innovation, coordination, integration, and reliability designed to help meet evolving on demand business requirements. System z allows you to:

- Host multiple virtual servers running mission-critical applications
- Virtualize or partition and share physical system resources dynamically to support each concurrently executing server instance operating in either a z/VM virtual machine or an LPAR
- Virtualize processor, communication, memory, storage, I/O, and networking resources with advanced Hypervisor technology
- Direct system resources to business-critical applications with hardware- and software-enabled dynamic resource allocation
- Use mainframe system resources to achieve high system utilization for efficiency and scaling
- Utilize specialty engines such as the Integrated Facility for Linux (IFL), System z Integrated Information Processor (zIIP), and System z Application Assist Processor (zAAP) to add full capacity processing engines for new workloads that help manage hardware and software licensing costs.
- Perform flexible, cost-efficient, highly secure interserver communications at near memory-speed (through Virtual Switches, Guest LANs, and HiperSockets)
- Use world-class, software-based advanced systems management, administration, and accounting tools
- Help ensure high availability, reliability, resiliency and system-wide disaster recovery (Parallel Sysplex® and Geographically Dispersed Parallel Sysplex™ (GDPS®))

5.2 Mainframe virtualization technologies

System z advanced hardware lets you logically partition the machine, share CPU, memory and I/O (channels and associated devices), add or remove computing Capacity on Demand, and provide high-speed communications among partitions.

Collectively, System z hypervisor technologies provide the ability to support and dispatch multiple LPARs and virtual machines efficiently. The PR/SM layer supports an ever-growing number of LPARs. The latest System z technology can support up to 60 distinct partitions or servers. z/VM runs in an LPAR and provides a virtualization layer designed to allow the capability to run hundreds to thousands of virtual server images for larger deployments.

5.2.1 Logical partitioning and PR/SM

Processor Resource/Systems Manager™ (PR/SM) is System z hardware technology that allows multiple operating systems to run on the same physical processor, with each operating system running in its own LPAR. System z PR/SM can currently support up to 60 of these logical partitions.

There are two types of partitions:

- A *dedicated* partition runs on the same dedicated physical processors at all times, which means the processors are not available for use by other partitions even if the operating system running on that partition has no work to do. This eliminates the need for PR/SM to get involved with swapping out one guest and dispatching another.
- *Shared* partitions, alternatively, can run on all remaining processors (those that are not being used by dedicated partitions). This allows idle systems to be replaced by systems with real work to do at the expense of PR/SM overhead incurred by the dispatching of the operating systems. In contrast with dedicated partitions, shared partitions provide increased processor utilization, but at the potential expense of performance for a single operating system.

5.2.2 Address spaces within LPARS

The z/OS® operating system employs address spaces and storage-protect keys to further protect key programs within an LPAR from corrupting each other's private storage or data areas. Storage-protect keys prevent programs from altering another's storage by matching the program's storage protect key against the storage protect key of the storage used by other programs in the LPAR. Additionally, extensive system locking and serialization techniques coordinate system events and actions. The combination of address spaces, storage protect keys, and system locking all help ensure z/OS provides the maximum in data integrity and availability.

5.2.3 HiperSockets

HiperSockets introduced the ability for virtual servers to communicate at near memory speed. This System z feature offers several advantages. From the hardware side, you can benefit from efficiency by eliminating much of the latency often associated with external physical networks. Implementing HiperSockets can also result in cost savings by reducing the amount of hardware (cabling, routers, switches, and hubs) that would otherwise be required if physical servers, rather than virtual ones, were connected together through an external network. Reducing the amount of physical components is one of the first steps towards infrastructure simplification which can translate to reductions in maintenance effort and staffing required for

systems operation. HiperSockets can also provide a security advantage because inter-server communication is contained within the System z server and, therefore, eliminates exposures that might arise when separate physical servers communicate with each other.

5.2.4 z/OS Workload Manager

System z Workload Manager (WLM) is the foundation for effective resource sharing on System z. Unlike other workload managers, System z Workload Manager is service goal or Service Level Agreement (SLA) oriented. Instead of telling the system how to manage its workloads (as in the UNIX® environment), you simply tell System z what your service goals or SLAs are and WLM determines the optimum way to achieve those goals automatically.

WLM requirement specifications can be very fine grained. You can specify requirements right down to transaction type, time of day, and user. WLM then adjusts what it has to do continuously with each new workload, in increments as small as 1/000th of a second or less, carefully balancing prioritized workloads so the most mission-critical workloads always get done. While this fine level execution time granularity only applies to z/OS workloads, and not to Linux workloads, WLM does help in Linux load balancing. System z Workload Manager manages z/OS workloads not just across logical partitions, but also across all System z mainframes within the Parallel Sysplex (which can extend to 24 separate System z mainframes), giving you a true system-wide view and system-wide management.

5.2.5 Intelligent Resource Director

The System z Intelligent Resource Director (IRD) complements System z Workload Manager. Where Workload Manager focuses on ensuring that the most important jobs or transactions get their required execution priorities, Intelligent Resource Director ensures that those tasks always get all the resources they need to do the job.

For example, when IRD sees a job is not getting the data as fast as it needs to, it can add additional data channels from a reserve pool of channels automatically, or take away underutilized channels from lower priority jobs, to meet the new need dynamically. IRD can also work with the Logical Channel subsystems to ensure that I/O requests from higher priority jobs always get put at the head of the queue so they are serviced first. Similarly, IRD works with DS8000™ Enterprise disk systems to ensure that the same kind of priority I/O queuing is done across different System z mainframes within a Parallel Sysplex. System z Work Load Manager and Intelligent Resource Director work together to balance workload requirements against available resources in order to achieve service level agreement requirements.

5.2.6 System z Parallel Sysplex

While System z hardware, operating systems, and middleware have long supported multiple applications on a single server, Parallel Sysplex clustering allows multiple applications to communicate across servers-and it can even support one large application spanning multiple servers, resulting in optimal availability for that application.

With Parallel Sysplex clustering and its ability to support data sharing across servers, IT architects can design and develop applications that have one integrated view of a shared data store. This eliminates the need to partition databases, which in non-System z environments typically creates workload skews requiring lengthy and disruptive database repartitioning. This often means downtime for the application—as well as extra costs, lost revenue and missed opportunities for your business. Also, ensuring data integrity with non-System z

partitioned databases often requires application-level locking, which in high-volume-transaction environments could lead to service-level agreements not being met.

Data sharing with Parallel Sysplex conveys the unique advantage of enabling nondisruptive database growth with automatic load rebalancing. System z shared database architectures also provide inherent locking services for the ultimate in data integrity. Parallel Sysplex data sharing capabilities help to prevent obstacles to application availability caused by partitioned database architectures. Single-view database simplicity helps remove unnecessary complication in the IT infrastructure. A simpler IT infrastructure in turn reduces the likelihood of unexpected outages and the impact of planned outages.

Geographically Dispersed Parallel Sysplex (GDPS) is a multi-site or single-site end-to-end application availability solution that provides the capability to manage remote copy configuration and storage subsystems (including the IBM System Storage™ Enterprise Storage Server®), to automate Parallel Sysplex operation tasks and perform failure recovery from a single point of control.

GDPS supports IBM Metro Mirror (PPRC), z/OS Global Mirror (XRC), or Global Mirror disk replication architectures, and gives you the capability to manage your remote copy configuration and storage subsystems. GPDS includes the automation of Sysplex operational tasks, and can include z/OS and Open System data. GPDS is an open architecture, and is application independent.

5.2.7 Capacity Upgrade on Demand

IBM provides the capability to quickly and nondisruptively activate “extra” processor capacity that is built directly into System z products, IBM Capacity Upgrade on Demand (CUoD) for a permanent increase of processing capability, and IBM On/Off Capacity on Demand (On/Off CoD) for a temporary capacity increase that lets you revert to your previous processing level whenever you want.

Types of Capacity on Demand for System z servers include:

- Permanent capacity for nondisruptive growth:

  Capacity Upgrade on Demand provides a means of planned growth for customers who know they will need increased capacity but are not sure when they will need it.

- Temporary capacity for fluctuating workloads:

  On/Off Capacity on Demand provides for planned and unplanned short-term growth driven by temporary processing requirements such as seasonal activity, period-end requirements or special promotions.

- Interim capacity for continued operation:

  Capacity BackUp temporarily supplements backup servers with the extra processing capacity they need in case of an unforeseen production server outage.

Project Big Green from IBM Spurs Global Shift to Linux on Mainframe[a]

01 August 2007—In one of the most significant transformations of its worldwide data centers in a generation, IBM announced that it consolidated about 3900 computer servers onto about 30 System z mainframes running the Linux operating system. The company anticipates that the new server environment will consume approximately 80% less energy than the current set up and expects significant savings over five years in energy, software and system support costs.

At the same time, this transformation makes the IT infrastructure from IBM more flexible to evolving business needs. The initiative is part of *Project Big Green*, a broad commitment that IBM announced in May 2007 to sharply reduce data center energy consumption for IBM and its clients.

"As one of the world's largest technology providers, IBM consistently assesses how our systems can be maximized to support our employees and clients," said Mark Hennessy, vice president and chief information officer, IBM. "A global account consolidation truly demonstrates that IBM is committed to driving stronger energy and technology optimization, and cost savings."

"The mainframe is the single most powerful instrument to drive better economics and energy conservation at the data center today," said James Stallings, general manager, IBM System z mainframe. "By moving globally onto the mainframe platform, IBM is creating a technology platform that saves energy while positioning our IT assets for flexibility and growth."

a. IBM press release 01 August 2007

5.3 Linux on IBM System z

The rise of Linux in the IT world, from an interesting academic exercise to a popular platform for hosting enterprise applications, is changing the way enterprises think about their computing models. Linux on the IBM System z platform can help with infrastructure simplification and existing modernization by taking advantage of the outstanding IBM mainframe and z/VM capabilities.

Linux on System z is an attractive platform that brings the strengths of the mainframe to new workloads, offering a higher level of uptake, integration with existing data and core applications available with z/OS and z/VSE™, less complex manageability, and dynamic workload management capabilities that can help to optimize your cost base. An effective way to grow your Linux workload capacity is to add more Linux guests to a VM system. By consolidating Linux servers onto one platform, hundreds or thousands of Linux instances on a single server requires less energy, cooling, and floor space.

Linux runs on the mainframe two ways:

- Natively in a logical partition
- As a virtual machine running under z/VM

IBM provides the capability to add specialty engines to existing (or net new) mainframes which help users expand the use of the mainframe for new workloads, while helping to lower cost of ownership.

5.3.1 Integrated Facility for Linux

The Integrated Facility for Linux (IFL) is a central processor (CP) that is dedicated to Linux workloads. The attractively priced IFL processor enables you to purchase additional processing capacity exclusively for Linux workloads, without affecting the MSU rating or the IBM System z model designation. This means that an IFL will not increase charges for System z software running on general purpose (standard) processors in the server.

IFLs are managed by PR/SM in logical partition with dedicated or shared processors CPs. The implementation of an IFL requires a Logical Partition (LPAR) definition, following normal LPAR activation procedure, and an LPAR defined with an IFL cannot be shared with a general purpose processor. IFLs are supported by z/VM, the Linux operating system and Linux applications, and cannot run other IBM operating systems.

Integrated Facility for Linux highlights include:

- Linux workload on the IFL does not result in any increased IBM software charges for the traditional System z operating systems and middleware
- An IFL has the functionality of a general purpose System z processor and operates on full capacity
- The consistent IFL price for the System z generations, beside the lower IFL price for z9™ BC, can deliver price/performance gains with each successive generation of technology
- z/VM V4/V5 and most IBM middleware products that run on Linux on System z are priced per processor
- Many software vendors have adopted the IBM pricing model for the traditional and Linux workloads

Figure 5-1 shows that adding IFL capacity to your mainframe does not increase the IBM software licensing costs for the traditional mainframe environment (for example z/OS).

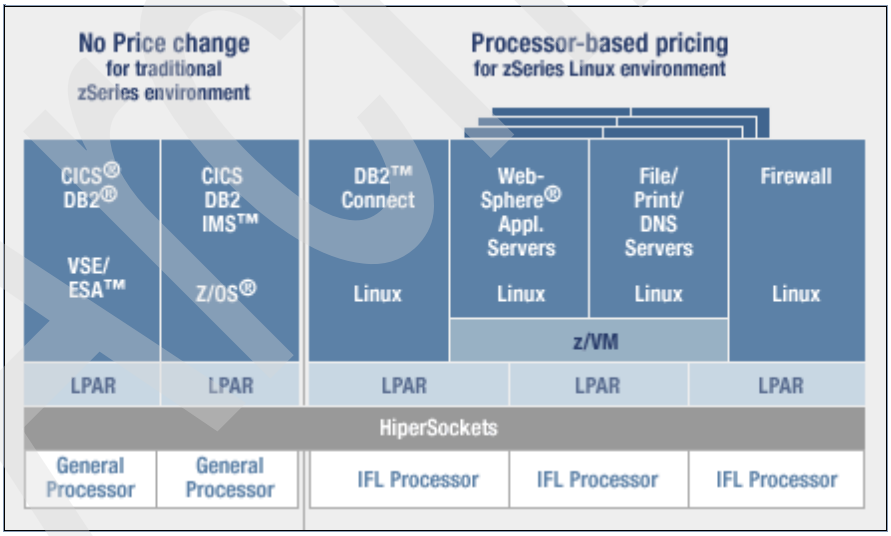


Figure 5-1 Adding IFL capacity to mainframe does not increase the IBM software licensing costs

5.3.2 Benefits of running Linux on the mainframe

Linux running on the IBM System z platform can reduce costs and over time the savings can be considerable. Potential savings can be derived from several areas. For example:

- Unified IT infrastructures based on the consolidation of distributed servers to virtual Linux servers on System z can enable higher utilization rates and reduce software licensing costs, minimize complexity, reduce maintenance effort with intelligent workload management, streamline the network and provide a more security-rich environment.
- Linux running on mainframes provides high qualities of service for your Linux applications utilizing the historical strengths of the System z environment and the established business processes and disciplines for disaster recovery and business resiliency.
- You can take advantage of your existing z/OS or z/VSE programs and data while deploying new applications rapidly with Linux on the mainframe.
- Linux-based solutions in the System z environment are easily deployed, extending the existing business applications and the open standards support of System z and enlarging the application developer and skill base for consistently high qualities of service to the user.
- Linux is platform agnostic, thereby allowing you to utilize the same tools and deploy across a range of architectures with a simple recompile. The mainframe with z/VM provides an excellent development environment. New virtual machines can be deployed in minutes on existing capacity, and if the virtual machine is not in use, its capacity is freed up for use by other virtual guests.

5.4 z/VM

z/VM is key to the software side of virtualization on the mainframe. The z/VM hypervisor is designed to help clients extend the business value of mainframe technology across the enterprise by integrating applications and data while providing exceptional levels of availability, security, and operational ease. z/VM virtualization technology is designed to allow the capability for clients to run hundreds to thousands of Linux servers on a single mainframe running with other System z operating systems, such as z/OS, or as a large-scale Linux-only enterprise server solution. z/VM V5.3 can also help to improve productivity by hosting non-Linux workloads such as z/OS, z/VSE, and z/TPF.

z/VM V5 is the result of nearly 40 years of innovation and refinement and can provide users with the ability to respond to rapidly changing market requirements more quickly and easily than with discrete single operating system servers. Unlike distributed hardware-based solutions, virtualization technology allows customers to virtualize processor, communications, storage, I/O, and networking resources to help reduce the need to duplicate hardware, programming and data resources.

Contrasted with a discrete server implementation, z/VM-based System z solutions are designed to provide significant savings, which can help lower total cost of ownership (TCO) for deploying new business and enterprise application workloads on a mainframe.

z/VM delivers support for hardware technologies such as FICON® channels, high-speed communication adapters and advanced storage solutions. z/VM also takes advantage of 64-bit real and virtual memory on System z servers. z/VM, supporting up to thousands of Linux guests, offers an ideal platform for consolidating select UNIX and Linux workloads on a single System z server.

Running Linux as a guest of z/VM is designed to provide the capability of running hundreds to thousands of Linux images while benefiting from the reliability, availability, scalability, security and serviceability characteristics of System z servers. At the same time, it allows customers to exploit the exceptional capabilities of z/VM virtualization.

5.4.1 z/VM helps control costs with virtualization and server consolidation

One IBM System z server running z/VM Version 5 can do the job of many distributed servers scattered across the enterprise by hosting a variety of platforms such as Linux on System z, z/OS, z/OS.e, z/VM, z/VSE, and z/TPF. With Linux on System z, the combination of the z/VM hypervisor and IBM mainframes can address infrastructure simplification issues faced by many large enterprises while benefiting from greater availability, scalability, virtualization, security, and reliability. System z environments with z/VM provide flexibility and management characteristics that can make it possible for customers to satisfy the requirements of an On Demand Business by deploying new Linux servers in minutes.

The complexity of maintaining large numbers of distributed servers can be relieved with a single IBM mainframe and can help to reduce costs by requiring less floor space, less energy consumption, and less person power. Simplification of the network by using HiperSockets can provide savings and reduce cabling, hubs, switches, and routers, as well as help to reduce maintenance effort. For many companies, critical enterprise data and applications are entrusted to System z environments. Using Linux on System z can enhance the value of customer applications and data by allowing Linux applications to communicate with other System z applications and access critical data where it resides. This can help to improve responsiveness and reduce unnecessary duplication of data.

Multiple Linux systems on System z servers can be easily created and managed with z/VM. Linux server images can share physical resources as well as programs and data and internal high-speed communications. z/VM V5 supports Integrated Facility for Linux (IFL) processors, the attractively priced hardware feature for Linux workloads available for System z. Linux on System z supports the IBM z/Architecture® (64-bit) on the System z platforms.

5.4.2 How z/VM virtualization works

The heart of z/VM is a multi-programming, multi-processing operating system kernel known as the Control Program (CP).[1] CP is the component of z/VM that creates and dispatches virtual machines on the real System z hardware. CP supports hundreds of commands, including the configuration of the virtual machine, and it lets customers change virtual machine configurations nearly at will.

System z virtualization depends on a hardware instruction called Start Interpretive Execution (SIE, pronounced *sigh*). Start Interpretive Execution is a special instruction that z/VM (and PR/SM) use to run virtual servers in their very efficient native mode except when z/VM or PR/SM needs to get control back for handling special situations, such as page faults, I/O channel program translation, privileged instructions, and so on. The result is that for most of the time these logical/virtual servers run without any emulation overhead, making z/VM and PR/SM highly efficient virtualization machines.

Virtualization of processors

Because z/Architecture defines that a System z data processing system can house one to 64 processors, each virtual machine z/VM creates can have one to 64 virtual processors. z/VM provides control over processor resources by letting a system administrator assign a share

[1] z/VM virtualization details provided from “Basics of z/VM Virtualization” by Bill Bitner and Brian Wade in the z/Journal February/March 2008 edition.

value to each virtual machine. This share value typically sets the minimum amount of processor resource a virtual machine can expect CP to allocate to it. z/VM also lets the system administrator define a maximum share value to prevent a guest from excessively consuming processor resource. The z/VM system administrator or system operator can adjust share settings while virtual machines are running.

Virtualization of memory

z/VM lets virtual machines share memory, which helps reduce memory requirements. All guest memory is virtual. CP overcommits physical memory by keeping resident only those guest pages that appear to have been needed in the recent past. When physical memory is scarce, CP moves stagnant guest pages first to expanded storage (a high-speed page storage buffer) and eventually to disk. CP brings these pages back to memory if the guest ever needs them again.

Virtualization of I/O devices

z/VM uses various methods to provide devices to virtual machines. CP can dedicate, or attach, a real device to a virtual machine. This gives the virtual machine exclusive use of the entire real device. Tape drives are attached typically to virtual machines. CP also can virtualize a device, which means it gives a guest a portion of a real device. This can be a portion in time, such as of a processor, or a portion of the device's storage capacity, such as of a disk drive.

z/VM provides disks to guests in various ways. While CP can dedicate entire disk volumes to virtual machines, more common is for CP to divide real disk volumes into disjoint, contiguous cylinder or block ranges called minidisks, thereby letting many guests each use some fraction of a real volume's storage capacity. Minidisks can be exclusive to virtual machines, or many virtual machines can use a single minidisk simultaneously, thereby sharing data. With Temporary Disk (TDISK), a system administrator can assign CP a pool of disk volumes it can use to instantiate minidisks users need for only a short time. When a user defines a minidisk, CP finds some free space in the pool and uses it to create the minidisk. When the user no longer needs the minidisk, he issues a command to dispose of it, and CP clears the space and returns it to the pool.

Because CP mediates access to minidisks it can use memory to improve their performance. Central to z/VM's minidisk strategy is the CP Minidisk Cache (MDC). With MDC, CP uses real memory or expanded memory to cache recently read portions of minidisks. This greatly improves performance for minidisks that are frequently read, such as those containing object code libraries or frequently used binaries.

Network connectivity is an important concern in many environments. z/VM meets customers' network needs by offering several networking options. The Control Program can dedicate network devices to virtual machines. The dedicated device can be a channel-to-channel adapter, an IBM Open Systems Adapter (OSA) that provides Ethernet connectivity, or a HiperSockets device, a kind of network adapter that connects one LPAR to another. z/VM also has its own TCP/IP stack, which guests can use as though it were an IP router.

A common network option used today is the virtual switch. Here, CP equips each virtual machine with a simulated IBM OSA and connects all those simulated OSAs to a simulated LAN segment called a guest LAN. Also connected to the guest LAN is a real OSA that CP manages. With this configuration established, CP can provide packet- or frame-switching functions for the guests, just as a real switch would in a real external network. In this way, the guest LAN becomes an extension of a real external LAN segment.

5.5 Why IBM System z

System z is designed with a processor book architecture. The business-class systems have a one book architecture, while the enterprise-class systems have a four book architecture. Each book is a complete System z processing system, including:

- A multichip module (MCM) housing the System z processor infrastructure
- Memory cards providing the complete memory system for the book
- An I/O connectivity infrastructure consisting of hot pluggable Memory Bus Adaptors and Self Timed Interface Fan-Out cards to connect the book to the external I/O systems.

All books are interconnected through a high speed, 2-way ring structure that provides each System z processor access to all memory storage and all caches across the System z. While System z will normally keep all processing within a single book for maximum efficiency, where necessary it can assign any available processor in any book to support a given task, and that processor can access any memory and cache in any book.

The book architecture provides System z with uniquely flexible and dynamic configuration options. Of the multiple processor units (PUs) that come in each book (which varies depending on which generation of machine that you have), only three to four have predefined roles. Two of these are defined as *System Assist Processors* (SAPs) and are devoted exclusively to supporting disk and tape I/O operations. Based upon the System z model, one or two of the processors are designated as hot spares that are invoked automatically to replace any failing processor.

The remaining units are defined as user characterizable, which means you can individually define each processor unit to perform any of several specialized processing roles (for example, Linux, Java, and DB2®). As a result, you can tailor and optimize System z to meet the unique processing characteristics of your particular environment.

The System z book structure is also the basis for cost effectively tailoring System z in terms of the processing power you require. This is because while each book comes with all processors enabled, you can disable all but one of the user characterizable processors if you initially do not require the full capacity. You only pay for the activated processors. Then, as your processing needs grow, you can simply activate any of the deactivated processors to meet the growing need.

The mainframe also provides the capability to run processors at subcapacity levels. This is controlled through microcode, and with today's increasingly powerful engines, enables customer to purchase hardware (and license software accordingly) in flexible increments.

System z processors are superscalar, meaning they can execute multiple instructions on each processor cycle, resulting in more processing parallelism and higher performance. System z architecture means extremely high processing accuracy. Each processing unit contains dual processors, and each processor has its own Instruction and Execution units, as well as its own floating point functions. This means that each instruction is executed twice, once on each processor, and the results are compared automatically to ensure simplified but highly accurate error detection.

Each System z PU also has a Compression unit on the chip, providing excellent hardware compression performance. Similarly, each PU also has a CP Assist for Cryptographic Function (CPACF) on the chip to provide high performance hardware encryption and decryption support for clear key operations. The Compression and CPACF units are integrated and as a result both units benefit from combining their buffers and interfaces.

5.5.1 Qualities of Service

System z is built for high availability. For example, should a processor fail, the status of the application running on the processor will be preserved and the application will continue on a newly assigned spare processor without user intervention. If no spare processor exists, the application state and workload will be passed to another active processor, after which System z will recover the failing task and enable it to resume normal operations. System z also has Redundant I/O Interconnect, which means it has dual connections to each I/O subsystem so that each book can get to each I/O device even if one of the I/O connections fails, or the other book is offline for maintenance purposes.

System z technology delivers the high level of availability required in today's global networked environment. Linux running in a System z environment can take advantage of this high level of availability for its workload. Even in a single footprint, the System z platform is designed to avoid or recover from failures to minimize business disruptions. High availability is realized through component reliability, redundancy, and design features that assist in providing fault avoidance and tolerance as well as permitting concurrent maintenance and repair.

Another aspect of availability is nondisruptive growth, in most cases enabled by IBM Capacity Upgrade on Demand. System z servers have the capability to add server capacity and virtual servers non-disruptively and to install FICON, ESCON®, and OSA-Express features without bringing the system down. Some upgrades can be initiated by customers over the Internet on select servers.

Figure 5-2 shows a typical System z configuration including operating systems, standard CPs, ICF (Integrated Coupling Facility), and Specialty Engines: zAAP, zIIP, IFLs (Integrated Facility for Linux).

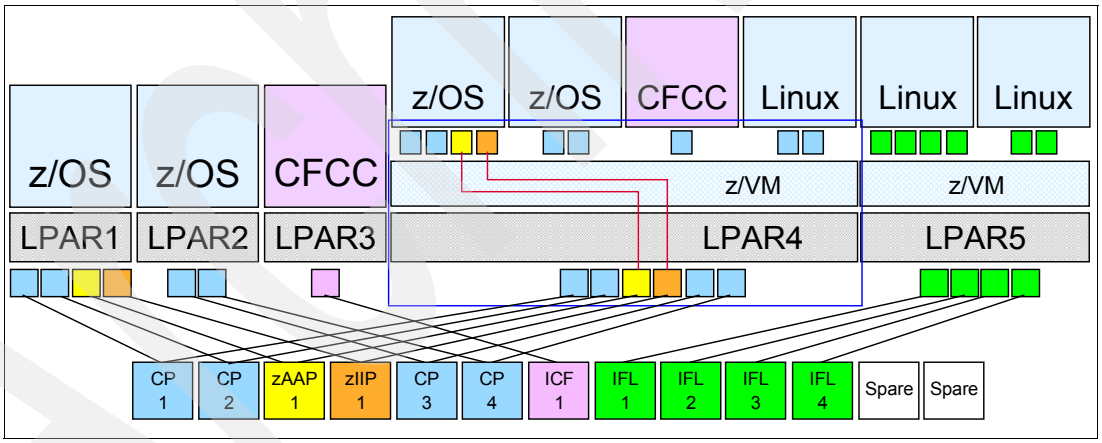


Figure 5-2 System z virtualized configuration

5.5.2 I/O capabilities

System z has a unique channel architecture which is designed to provide powerful and flexible support for the most demanding I/O performance and high volume workloads. This is managed through the foundation technology called Logical Channel Subsystems (LCSS). Each System z can have up to four of these logical channel subsystems, each capable of supporting 15 logical partitions, and each capable of addressing 256 data channels and 64,000 I/O devices. As a result, a single System z today can handle over 1,000 data channels and a quarter million I/O devices.

The Logical Channel Subsystem's architecture is further enhanced by the System z Multiple Image Facility. This Multiple Image Facility allows all 15 logical partitions that share a common Logical Channel Subsystem to directly access each I/O device without having to forward the request through an intermediate partition, as is the case with UNIX architectures. The direct result of a shorter path length is improved I/O performance.

System z Channel Spanning allows each device (disk or tape units attached using Fibre Channel technology), to appear to be on any Logical Channel subsystem. This means it can be accessed by any partition on the mainframe. The direct result is greater I/O flexibility and simplified operational management.

5.5.3 Specialty engines

System z offers the unique capability to add specialty processors to help users expand the use of the mainframe for new workloads, while helping to lower cost of ownership. The IBM System z specialty engines can run independently or complement each other.

These specialty engines include:

- Integrated Facility for Linux (IFL)

  The attractively priced IFL processor enables you to purchase additional processing capacity exclusively for Linux workloads, without affecting the MSU rating or the IBM System z model designation.

- Integrated Information Processor (zIIP)

  zIIP is designed to help free-up general computing capacity and lower overall total cost of computing for select data and transaction processing workloads for business intelligence (BI), ERP and CRM, and select network encryption workloads on the mainframe. The zIIP's execution environment accepts eligible work from z/OS, which manages and directs the work between the general purpose processor and the zIIP. IBM will not impose software charges on zIIP capacity.[2]

- Application Assist Processors (zAAPs)

  zAAPs are attractively priced specialized processing engines that provide a strategic z/OS Java execution and XML parsing environment. IBM does not impose software charges on zAAP capacity.

- Internal Coupling Facility (ICF)

  The coupling facility allows multiple z/OS LPARs to share, cache, update and balance data access. This hardware is a major component of a Parallel Sysplex. ICFs also allow Internal Coupling (IC) links to help eliminate the requirements for external CF links. There is no associated software cost for the ICF.

- System Assist Processor (SAP®)

  The SAP is standard on IBM System z servers and is a dedicated I/O processor to help improve efficiencies and reduce the overhead of I/O processing of every IBM System z logical partition regardless of the operating system (z/OS, z/VM, Linux, z/VSE and z/TPF). For very high I/O intensive workloads, an additional one or more SAPs can be purchased (model dependent). There is no software cost for the SAP.

Taking advantage of all these exceptional System z capabilities with Linux provides these attractive attributes to the Linux on System z environment. With the portability of Linux, applications can be moved quickly and easily to System z servers. The capacity of the

[2] Additional IBM software charges apply when additional general purpose CP capacity is used. The amount of general-purpose processor savings varies based on the amount of workload executed by the zIIP, among other factors.

System z platform makes it ideal to help simplify your operation and reduce your costs by consolidating the number of servers in your business. z/VM offers advanced technology to help achieve these results.

Macro 4 boosts speed-to-market with IBM System z9® Business Class[a]

By implementing an IBM System z9 Business Class server with zAAP, zIIP and IFL engines, Macro 4 has improved its ability to run multiple independent virtual server environments simultaneously.

Business need: Prepare for deployment of new mainframe technologies at customer sites; manage rapid growth in demand for processing resources; improve time-to-market for new products; keep total cost of ownership low

Solution: Worked with IBM and Thesaurus to install a new IBM System z9 Business Class (z9 BC) platform with two standard engines, one Integrated Facility for Linux (IFL) engine, one IBM System z Application Assist Processor (zAAP) and one System z Integrated Information Processor (zIIP)

Benefits: Ability to run dozens of virtual test and development environments side-by-side and to run multiple operating systems on a single physical server; ability to increase capacity 800 percent without requiring disruption to live systems; anticipated return on investment within 12 months

a. Case study available at http://ibm.com/systems/z/advantages/virtualization

5.5.4 Total cost of ownership

Taking advantage of the strengths of the mainframe and virtualization can enable you to see savings on:

- Hardware

  Is your server farm growing out of control? Linux running on System z can help you simplify your infrastructure. The System z virtualization capabilities can allow you to consolidate hundreds of Linux servers onto a single server and deploy these virtual servers in a matter of minutes. Fewer servers can also mean additional savings in physical space and heating, air conditioning, and electricity costs.

- Software

  Because Linux is an open operating system, you might be able to implement your core business applications on Linux on System z, helping you from being locked into costly licenses for proprietary software. Consider also that Linux software is usually priced on a per-engine basis. Because a single System z server can run multiple Linux applications on a single engine, your licensing costs can be reduced by consolidating servers onto a System z server.

- Staffing

  Consolidating multiple servers onto a single System z server running multiple virtual Linux servers can mean less labor required for system management and maintenance. The centralized system management and autonomic computing capabilities of Linux on System z can also help cut down on the errors and workload-balancing tasks that otherwise can eat up countless IT staff hours.

- Business continuity

  Linux on System z can give you the ability to scale up and out—on demand—to meet spikes in server activity, helping minimize costly transaction delays and potentially

devastating system crashes. With a suite of built-in features, System z servers can rapidly respond to, or even anticipate, threats to system health, helping prevent budget-sapping downtime.

6

Power Systems virtualization

IBM Power Systems platforms (System i and System p) provides a virtualization feature named PowerVM (formerly *Advanced POWER Virtualization* or *APV*) that is based on paravirtualization. The supported operating systems (AIX, Linux, and i5/OS®) have been modified to support and optimize the underlying hardware virtualization capabilities. Word class POWER virtualization balances the need for flexible resource management and isolation within a dynamic environment.

Key benefits of deploying PowerVM Editions and IBM Power Systems include:

- Go green and save, by potentially cutting energy costs up to 65% through server consolidation[1]
- Potentially reduce the cost of existing infrastructure by up to 72% [2]
- Manage growth, complexity and risk, by potentially reducing server sprawl and administration while improving availability and application performance
- Take advantage of 40 years of leadership in virtualization from IBM

We discuss the following topics in this chapter:

[1] Based on an IBM study, data available at http://www.ibm.com/systems/power/software/notices/index.html#1

[2] From the paper “Impact of IBM System p Server Virtualization,” Transforming the IT Value Equation with POWER6™ Architecture. International Technology Group, May 2007.

6.1 PowerVM virtualization features

PowerVM comes in three editions. You choose the right offering for your business needs:

- PowerVM Express Edition

  PowerVM Express Edition is offered exclusively on the p550 and p520 Express servers and is designed for users looking for an introduction to more advanced virtualization features at a highly affordable price. With PowerVM Express Edition, users can create up to three partitions on the server, take advantage of virtualized disk and optical devices VIOS, and even try out the Shared Processor Pool.

- PowerVM Standard Edition

  For users ready to get the full value out of their server, IBM offers PowerVM Standard Edition providing the most complete virtualization functionality for UNIX and Linux in the industry. PowerVM Standard Edition available on all POWER5 and POWER6 processor-based servers and includes features designed to allow businesses to increase system utilization, while helping applications continue to get the resources they need.

- PowerVM Enterprise Edition

  PowerVM Enterprise Edition is offered exclusively on POWER6 processor-based servers and includes all the features of PowerVM Standard Edition plus an exciting new capability called Live Partition Mobility. Live Partition Mobility allows for the movement of a running partition from one POWER6 server to another with no application downtime resulting in better system utilization, improved application availability and energy savings. With Live Partition Mobility, planned application downtime due to regular server maintenance can be a thing of the past.

PowerVM provides key benefits by helping:

- Improve business responsiveness and operational speed by re-allocating resources to applications dynamically as needed to better match changing business cycles or handle unexpected surge in demand
- Reduce total cost of ownership and make better use of IT assets by improving server use significantly and by sharing I/O resources
- Simplify IT infrastructure management by making workloads independent of hardware resources, thereby enabling clients to make business-driven policies to deliver resources based on time, cost, and service-level requirements

Table 6-1 IBM Power Systems PowerVM Editions

	Express	**Standard**	**Enterprise**
Maximum LPARs	3 per server	10 per core	10 per core
Management	IVM	IVM, HMC	IVM, HMC
Virtual I/O Server	✓	✓	✓
Lx86	✓	✓	✓
Shared Dedicated Capacity[a]	✓	✓	✓
Multiple Shared Processor Pools[a]		✓	✓
Live Partition Mobility[a]			✓

a. POWER6 processor-based models only

Table 6-2 Availability of PowerVM features by server offering

	JS12	JS22	520	520	520	550	550	570	570	575	595		
	7998-60X	7998-61X	8203-EA4	9407-M15	9408-M25	8204-EA8	9409-M50	9117-MMA	9406-MMA	9125-F2A	9119-FHA	System p POWER5/5+	System i POWER5/5+
Express Edition			✓			✓							
Standard Edition	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Enterprise Edition	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓		
Micro-Partitioning	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
VIOS	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓[a]
IVM	✓	✓	✓			✓						✓[b]	
Lx86	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	
Live Partition Mobility	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓		
Multiple Shared Processor Pools			✓	✓	✓	✓	✓	✓	✓	✓	✓		
HMC			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

a. VIO supports AIX and Linux operating system partitions
b. IVM is supported on POWER5/5+ System p5 560Q servers and below, and the BladeCenter JS21

6.1.1 Micro-Partitioning Technology and Shared Processor LPARs

Shared processor logical partitioning (LPAR) allows clients to *slice up* a machine into virtual partitions and provides the flexibility to change the allocation of system resources dynamically for those environments. The Micro-Partitioning™ feature provides the capability to create multiple virtual partitions within a processor to a granularity of 1/100th of a CPU, with a 1/10th of a CPU partition minimum.

Any of the virtual servers can run on any of the physical processors. Thus, the processor resources are fully shared, making it possible to run the physical server at very high utilization levels.

6.1.2 Integrated Virtualization Manager

The Integrated Virtualization Manager (IVM) literally allows you to point, click, and consolidate workloads with its easy-to-use, browser-based interface. The IVM lowers the cost of entry into POWER5 virtualization because it does not require the use of a hardware management console for single system partitioning.

With IVM, you can partition a single system, including the creation of LPARs and management of virtual storage and Ethernet. Best of all, IVM is included at no additional cost with the optional purchase of PowerVM on most IBM System p entry to mid-range models, as well as the BladeCenter JS21, JS21, and select Power Systems servers. It is packaged as part of the Virtual I/O Server.

6.1.3 Virtual I/O Server

The Virtual I/O Server is a special-purpose partition, called the *Hosting Partition*, which provides virtual I/O resources to client partitions. The Virtual I/O Server owns the resources

that are shared with clients. A physical adapter that is assigned to a partition can be shared by one or more partitions, enabling administrators to minimize the number of physical adapters they require for individual clients.

The Virtual I/O Server is thus designed to reduce costs by eliminating the need for dedicated network adapters, disk adapters, and disk drives. Unlike other virtualization techniques, PowerVM does not require all devices to be virtualized. Devices can be:

- A mixture of dedicated devices assigned to partitions for maximum performance
- Used in the Virtual I/O Hosting Partition to be shared by multiple partitions to provide higher efficiency of resources and adapters

6.1.4 Live Partition Mobility

Part of the PowerVM Enterprise Edition offering, Live Partition Mobility allows clients to move a running partition from one physical System p POWER6 server to another System p POWER6 server without application downtime, helping clients to avoid application interruption for planned system maintenance, provisioning, and workload management. Live Partition Mobility is supported on IBM System p POWER6 servers.

The migration operation, which takes just a few seconds, maintains complete system transactional integrity.[3] The migration transfers the entire system environment, including processor state, memory, attached virtual devices, and connected users.

Live Partition Mobility allows you to move partitions around such that you can perform previously disruptive operations on the machine when it best suits you, rather than when it causes the least inconvenience to the users. Live Partition Mobility helps you meet continuously increasingly stringent service-level agreements (SLAs), because it allows you to move running partitions and applications proactively from one server to another.

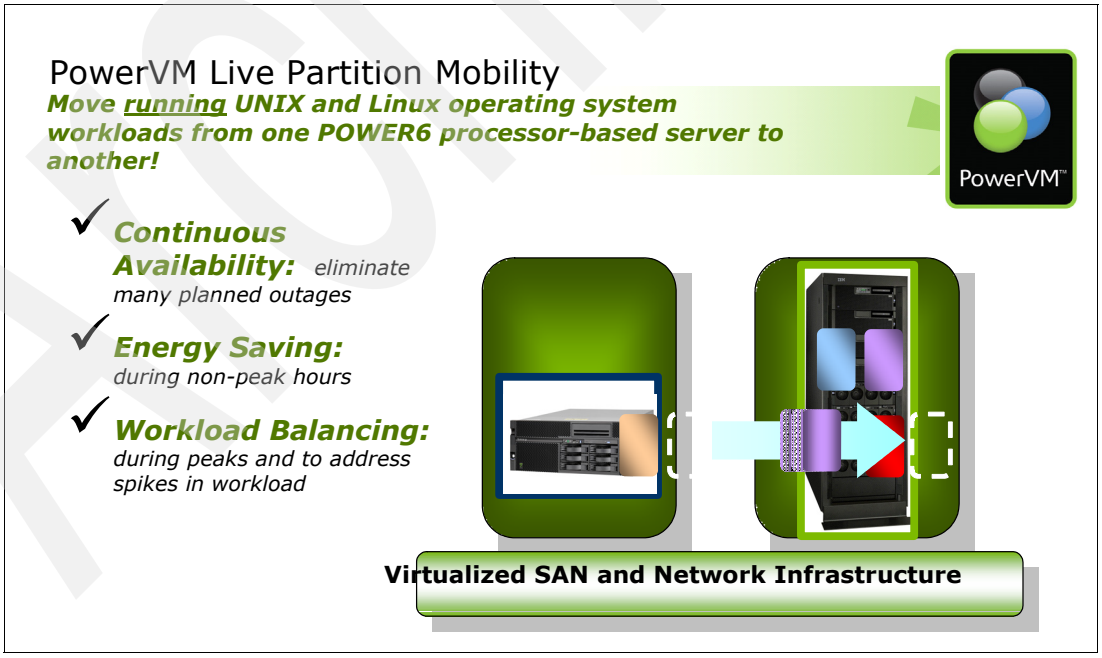


Figure 6-1 Benefits of PowerVM Live Partition Mobility

[3] For more information about Live Partition Mobility, see *PowerVM Live Partition Mobility on IBM System p*, SG24-7460.

The ability to move running partitions from one server to another offers the ability to balance workloads and resources. Should a key application's resource requirements peak unexpectedly to a point where there is contention for server resources, you might move it to a larger server or move other, less critical, partitions to different servers and use the freed up resources to absorb the peak.

Live Partition Mobility can also be used as a mechanism for server consolidation, because it provides an easy path to move applications from individual, stand-alone servers to consolidation servers. If you have partitions with workloads that have widely fluctuating resource requirements over time (for example, with a peak workload at the end of the month or the end of the quarter) you can use Live Partition Mobility to consolidate partitions to a single server during the off-peak period, allowing you to power-off unused servers. Then move the partitions to their own, adequately configured servers, just prior to the peak. This approach also offers energy savings by reducing the power to run machines and the power to keep them cool during off-peak periods.

Note: At the time of writing, Live Partition Mobility functions are available on supported systems using the same management interface. A machine managed by IVM, can migrate to another system managed by IVM. The same is true for systems managed by the HMC. However, neither IVM to HMC, nor HMC to IVM partition migrations are supported.

6.1.5 Shared dedicated capacity

A new feature for POWER6, Shared Dedicated Capacity allows for the "donation" of spare CPU cycles for dedicated processor partitions to be utilized by the shared pool, thus increasing overall system performance. The dedicated partition maintains absolute priority for dedicated CPU cycles, and sharing will only occur when the dedicated partition has not consumed all its resources. Shared Dedicated Capacity is supported on IBM System p POWER6 servers.

6.1.6 Multiple shared processor pools

Multiple shared processor pools allows for automatic nondisruptive balancing of processing power between partitions assigned to the shared pools-resulting in increased throughput and the potential to reduce processor-based software licensing costs.

6.1.7 PowerVM Lx86

Run x86 Linux applications on POWER. This feature enables the dynamic execution of x86 Linux instructions by mapping them to instructions on a POWER processor-based system and caching the mapped instructions to optimize performance. PowerVM Lx86 software is designed with features that enable users to easily install and run a wide range of x86 Linux applications on Power Systems platforms with a Linux on POWER operating system.

6.1.8 Unique capabilities

Unlike other UNIX virtualization solutions:

- Virtualization capability is available on every model across the System p line.
- PowerVM provides Live Partition Mobility for UNIX systems
- PowerVM allows for far better granularity than hardware partitioning techniques as your partitions do not have to be carved in increments of system boards.

- LPARs provide complete security isolation in both the kernel and application spaces. The kernel level is still exposed when you run in an application container type solution.
- PowerVM allows you to run multiple isolated OS instances on the Hypervisor, so you are not tied to a host OS (such as you would be using an application container approach). This allows you to run different patches and OS levels in the same physical machine, as well as not cause an outage for the entire system if a partition needs to be rebooted.
- System p resources can be reconfigured dynamically between logical partitions, that is no reboots are required for the changes to take effect.
- There is not a limit to how large your logical partitions can be. LPARs can grow to the full size of the machine (which could be important for test LPARs for example). Other solutions might allow you to have a partition of only four processors or less.
- System p memory can be reallocated dynamically.
- System p I/O can be virtualized, is reconfigurable, and a mix of dedicated and shared I/O is supported across partitions.

PowerVM is designed to enable you to aggregate and manage resources through a consolidated, logical view. (System p was awarded the Best Virtualization Solution at LinuxWorld 2006.) Key benefits of deploying System p Virtualization include:

- Lowering the cost of existing infrastructure by up to 62%[4]
- Increasing business flexibility, allowing you to meet anticipated and unanticipated needs
- Reducing the complexity to grow your infrastructure

6.1.9 PowerVM usage scenarios

Virtualization has multiple usage scenarios to attain these benefits and to assist clients improve their IT efficiencies, effectiveness, and management. The following scenarios are only a sample of the capabilities that are available using PowerVM:

- Server consolidation of several environments including:
  - Underutilized servers enabled by PowerVM fine tuning granularity capabilities.
  - Mixed workloads enabled by PowerVM LPAR, dynamic LPAR, and Shared Pool flexibility features.
  - Applications with varied and dynamic resource requirements enabled by PowerVM partition load manager and on-demand capabilities.
- Rapid deployment of new workloads enabled by PowerVM profiles and management processes:
  - Greater exploitation and fault avoidance of I/O devices enabled by PowerVM Virtual I/O capabilities.
  - Lower cost, higher performance interconnect capabilities enabled by PowerVM Virtual LAN support.
- Application development and testing in secure domains enabled by PowerVM partition isolation and security features.
- Support multiple operating system environments for testing, migration, or application optimization with PowerVM multi-operating system support:

[4] Business Case for IBM System p5™ Virtualization, Economic Benefits of IT Simplification. International Technology Group, 10 Feb 2006. Study methodology: Companies in financial services, manufacturing and retail with $15 Billion+ revenues and total 200 000+ employees focusing on UNIX large enterprise environments with multiple, broad-ranging applications.

– Virtual BladeCenter to provide a single server "blade-like" implementation of parallel workloads with dynamics to support on-demand application needs enabled by PowerVM dynamic logical partitioning and partition load manager.
– Low cost approach to selected business continuity solutions with simplified management enabled by the combined features of PowerVM.

6.1.10 Client feedback

The following quotations tell you what IBM clients are saying.

By leveraging the advanced virtualization capabilities of our IBM infrastructure, we're able to squeeze every last dime out of the environment that's on the floor, with no impact on its stability. It gives us a powerful tool to maximize the efficiency of our IT spending.

Dave Russo Assistant Vice President, Web Infrastructure and Design of MetLife

To stay competitive with other companies, including the larger ones, we have to take a positive approach to more advanced technology and break into new areas. The virtualization and introduction of the p5-550Q are helping us achieve this goal.

Mr. Adachi, Plala Networks, Inc.

The business can now request the computing capacity that it needs for a new application, and we simply carve out a new virtual server in a partition on one of the System p5 servers.

Abe Boersma, Product Manager System p/AIX, Rabobank

6.2 AIX workload partitions

AIX 6.1 extends the capabilities of the AIX OS to include new virtualization approaches including the ability to relocate applications between systems without restarting the application, new security features to improve and simplify security administration, new availability features inspired by IBM history in enterprise-class systems and numerous features designed to make the AIX OS easier and less expensive to manage. This AIX release underscores the firm commitment by IBM to long-term UNIX innovations that deliver business value.

AIX 6.1 introduces a new mechanism for virtualization: *workload partitions* (WPARs). In contrast to LPARs, which are created and managed at the server's firmware level, AIX WPARs are software partitions that are created from, and share the resources of, a single instance of the AIX operating system. This means that you must have AIX 6.1 to create WPARs, but you can create WPARs on any System p hardware that supports AIX 6.1, including POWER4, POWER5, and POWER6 hardware. You do not need an HMC or IVM to create or manage WPARs.

There are two kinds of workload partitions, System WPARs and Application WPARs.[5]

- System WPARs

  System WPARs are autonomous virtual system environments that have their own private file systems, users and groups, login, network space, and administrative domain. To users

and applications, a system WPAR appears almost exactly like a full AIX system. Operating system services, such as telnet, are supported, so if network information has been configured, users can telnet into a system WPAR as root or any other defined user, issue commands, and run applications as they would on any other AIX system.

There are many uses for System WPARs. For example, if you are setting up an application development or test environment, you can create it in a system WPAR. You do not have to acquire a dedicated server or LPAR but, if the new environment should have serious problems, any adverse effects will be confined to the WPAR.

- Application WPARs

  Application WPARs provide an environment for isolation of applications and their resources to enable checkpoint, restart, and relocation at the application level. An application WPAR is essentially a wrapper around a running application or process for the purposes of isolation and mobility. It lacks some of the system services provided by system WPARs-for example, it is not possible to log in or telnet into an application WPAR. When the application running in an application WPAR terminates, the WPAR also ceases to exist. Application WPARs are most useful when you want to enable Live Application Mobility-that is, when you want to be able to move a running application from one AIX system to another. You might want to relocate applications to avoid downtime resulting from scheduled maintenance or to improve performance by moving an application to a more powerful server.

Figure 6-2 shows LPARs and AIX Workload Partitions are complementary technologies and can be used together. Each square on this diagram represents a processor or micro-partitioned processor, while each dotted line square represents a WPAR.

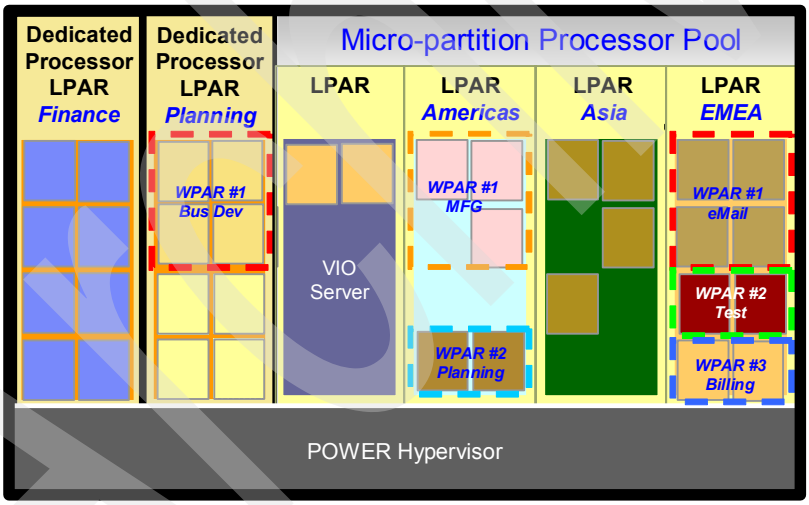


Figure 6-2 LPARS and Workload Partitions together on a system

6.2.1 Comparing WPARs with LPARs

As noted earlier, you do not need access to the HMC or IVM to create WPARs as you do for LPARs. WPARs are lightweight and quicker to install, because they share many of the file systems and resources of the global AIX system in which they reside. While using an LPAR requires you to install an entire operating system, creation of system WPARs only installs private copies of a few file systems, and application WPARs share even more of the global

[5] WPAR description provided by Jack Alford from the Developer Works article "AIX 6.1 Workload Partitions: Basic management of Workload Partitions in AIX" available at: http://www.ibm.com/DeveloperWorks/aix/library/au-workload/

system's resources. As a result, a WPAR can be created in just a few minutes without installation media.

Ongoing administration and maintenance of WPARs should be simpler-fewer AIX licenses might be required, and you do not have to install fixes and updates on so many virtual systems. There is a command for synchronizing the filesets of a WPAR with the corresponding filesets on the global system, so you have the choice of propagating AIX fixes to WPARs or continuing to run with the current versions of system files.

While LPARs offer a significantly higher degree of workload isolation, WPARs might provide enough isolation for your particular workloads, especially temporary ones such as development or test environments. Similarly, with LPARs, you can achieve a greater degree of control over the usage of resources-by allocating entire processors or precise fractions of processors to an LPAR, for example. With WPARs, you do not have such fine control over resource allocations, but you can allocate target shares or percentages of CPU utilization to a WPAR (if have used the AIX Workload Manager, you will find the share and percentage resource allocation scheme familiar). Similar differences exist for the allocation of memory, number of processes, and other resources.

6.3 Why IBM System p

In today's business climate, companies can gain a significant competitive advantage by adapting quickly to shifting conditions-embracing uncertainty rather than resisting it. Through flexible, responsive IT platforms designed to anticipate and adapt to constantly changing business landscapes, companies can simplify complex infrastructures, safeguard key assets, reduce migration risks and help achieve on demand computing without compromising system affordability.

With IBM System p servers, featuring IBM POWER6 and POWER5+™ processors, you no longer need to trade function, reliability, and system utilization for price These servers extend IBM leading-edge capabilities to help enable highly adaptive, efficient operations in an extremely competitive marketplace-helping businesses simplify their IT infrastructures, safeguard critical data and mitigate the potential risks of platform migration without compromising affordability, responsiveness or flexibility.

Reasons to look at IBM System p

When you need to control costs, while improving overall performance, client satisfaction and the ability to respond quickly to changing business challenges-consider the IBM System p5 family of servers and the wide variety of solutions supported for both AIX 6, AIX 5L™ (UNIX) and Linux operating system environments. The following capabilities and the stability of the IBM POWER architecture make System p servers the right platform for a UNIX and Linux migration that can support your infrastructure for years to come and upon which to achieve the lower TCO benefits that can come from server consolidation:

- System p servers set dramatic new industry performance records

  IBM System p servers, now with POWER6 processors, demonstrate continued commitment by IBM to organizations of all sizes. In addition to energy conservation and virtualization technologies, System p servers offer jaw-dropping performance. In fact, System p servers took the lead in over 70 key computing performance benchmarks[6].

- IBM System p Virtualization technologies help make System p servers unique to the UNIX/Linux world

[6] IBM performance benchmark information is available at http://ibm.com/systems/p/hardware/benchmarks

You can achieve greater overall business value than server consolidation alone with the combination of unique advanced virtualization technologies-available on System p servers-that enable you to do more with a single system than ever before.

- Partition migration is enabled between physical UNIX servers through Live Partition Mobility.
- Dynamic logical partitioning increases the flexibility of partitioned systems by enabling administrators to add, remove or move system resources (such as memory, PCI adapters and processors) between partitions without the need to reboot each partition.
- Micro-Partitioning technology supports both server consolidation and mixed workloads by allowing system resource allocation to be fine-tuned to changing business requirements through the activation of multiple partitions in each CPU in units as small as 1/10th of a processor.
- Virtual I/O enables the sharing of physical LAN adapters and disk storage devices on a system, which can reduce the number of physical devices needed and help achieve greater utilization of those already installed.
- Virtual LAN provides high-speed secure partition-to-partition communications-simplifying system configurations helping to reduce costs.
- Cross-partition workload management allows dynamic workload adjustments to optimize throughput by continuously monitoring application workloads.
- IBM Virtualization and Capacity on Demand technologies help increase flexibility and resource utilization while helping lower systems and administration costs.

- Proven IBM mainframe-inspired reliability features.

  Leading-edge mainframe-inspired reliability, availability and serviceability features contribute to ease of management and improved availability. IBM mainframes are the gold standard for reliability and availability. Many of these unique characteristics have been adapted for the System p family. Isolated logical partitions are designed to shield each application and advanced autonomic self-healing and self-configuring security features are embedded in the OS.

- AIX is designed to automatically handle day-to-day activities and empower you to make better decisions

  Workload management features enable you to free staff from repetitive activities. Advanced accounting features are designed to give you key data necessary for improved resource utilization and planning. AIX Version 6.1 is designed with new features for virtualization, security, availability, and manageability designed to make AIX Version 6 even more flexible, secure, and available than previous versions.

- Use the right OS for your business, on the same system, on the same processor, at the same time

  You can simultaneously run AIX Version 6.1, AIX 5L V5.3 and AIX 5L V5.2 as well as multiple versions of Linux using logical partitioning. And, System p servers offer a single binary compatible architecture, so this revolutionary technology is new and innovative-but not disruptive.

6.4 Why IBM System i

System i is virtualized to manage multiple applications and processes. One of the key factors to the i5/OS environment's efficiency is the ability to run multiple business processes and applications reliably and securely together. In a study of large enterprises using multiple operating systems, IBM found utilization rates on i5/OS-based servers were over 10 times

that compared to Intel-processor based servers and over twice as high as UNIX and other mid-range based systems. The high rate of i5/OS utilization is achieved through the use of a variety of proven virtualization technologies, such as subsystems (multiple workloads managed in a single operating system image) and logical partitions (multiple workloads managed in independent operating system images).

Subsystems are independent operating environments within an i5/OS instance, through which the system coordinates and automatically manages work flow and resource use for jobs, processes and applications. The system can contain many subsystems, each of which can be assigned defined system resources such as memory pools and processor priority. i5/OS subsystems are routinely used to separate multiple Web, batch and transaction processing application components. Subsystems can be tuned manually for specific workloads, but most companies let i5/OS automatically handle routine subsystem prioritization and workload balancing.

Virtualization by micro-partitioning enables multiple images of i5/OS, IBM AIX or Linux to be run on the same system with resources automatically balanced by the IBM POWER Hypervisor. Unlike most industry virtualization implementations on UNIX and Intel processor-based systems, the POWER-processor based Micro-partitioning is directly based the proven IBM mainframe hypervisor architecture.

The POWER Hypervisor implementation ensures true separation of operating systems functionality from the performance-optimized firmware layer that handles management of system hardware resources. The POWER Hypervisor ensures each operating system partition-either i5/OS, AIX or Linux-is completely independent and secure. Up to ten partitions can be defined per processor, with dynamic or automatic balancing of processor resources between the partitions. Companies deploying i5/OS have routinely deployed their business applications using logical partitioning to optimize their IT operations over the past decade.

6.4.1 VMware ESX Server on integrated servers

System i integration with BladeCenter and System x allows businesses to operate heterogeneous environments that include Intel-compatible servers running VMware ESX Server in addition to core business applications on the System i platform.

Figure 6-3 shows the integration of System x with System i using an adapter.

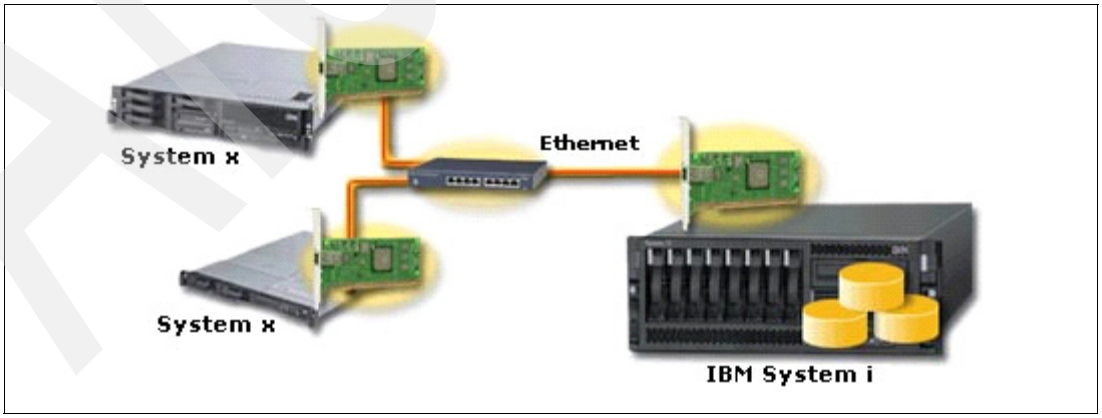


Figure 6-3 Integration of System x with System i using an adapter

VMware ESX Server is supported on BladeCenter and System x models that are attached to the System i platform through the iSCSI host bus adapter (iSCSI HBA). Support varies by i5/OS release.

T-Mobile Direct cuts complexity with System i virtualization[a]

The Challenge: Support extremely rapid growth in number of retail outlets by increasing the capacity and performance of core business systems; gain greater flexibility and ease of management.

The Solution: Working with Triangle, an IBM Premier Business Partner, T-Mobile Direct upgraded to an IBM System i5™ 570 with six IBM POWER5 microprocessors running i5/OS and Linux. The i570 also manages a number of business-critical Windows®-based systems; now deploying IBM Virtualization Engine™ technologies to optimize server performance and automate management.

The Benefit: Greater capacity and performance. Enhanced options for integrating Windows-based systems into System i5 environment. Improved resilience and performance for Windows systems; potential for significant savings in software licensing costs.

a. Case study available at http://www.ibm.com/systems/i/success/

7

System x and BladeCenter Virtualization

In addition to contributing to open source virtualization initiatives like Xen, IBM has chosen to work with IBM partners to provide virtualization in the x86 environment. For example, IBM is the number one reseller of VMware products.

Until recently, all virtualization on x86 architecture was implemented in software. However, Intel and AMD have developed hardware virtualization technology that is designed to:

- Allow guest operating systems, VMMs, and applications to run at their standard privilege levels.
- Eliminate the need for binary translation and paravirtualization.
- Provide more reliability and security.

The first phase of this iterative development endeavor was implemented in the processors. Intel named their hardware virtualization technology *VT-x*. AMD named their hardware virtualization technology *Pacifica, al*. Additional phases of hardware virtualization technology are being developed for I/O and memory.

The end result of these virtualization hardware assists is to reduce the complexity of the hypervisor, which can reduce overhead and improve performance significantly. IBM System x Servers support both Intel and AMD virtualization hardware assists.

We discuss the following topics in this chapter:

7.1 Benefits to virtualization

There are many benefits to virtualizing your x86 infrastructure. Potentially every guest can have an identical hardware stack as the virtualization layer pretends to have all guests running on the same hardware. For example, in the x86 space this means:

- Easier maintenance

  You no longer need to maintain different sets of drivers or keep track of Windows and Linux hardware configurations, thus increasing the ease of management of the software stack.

- More reliability

  You no longer need to validate or certify any given type of hardware and device drivers for your Windows and Linux platforms. Linux and Windows hosts can potentially all have a single (virtual) standard SCSI adapter, Ethernet interface, and so forth. Having a consistent and thoroughly tested low-level standardization of your Linux and Windows devices will likely improve reliability.

- Easier problem determination

  During difficult problem determination steps to debug your application, you no longer need to take into account potential incompatibilities of hardware and device drivers. Physical deployments typically require a deep investigation of the actual hardware and software configuration to determine the root cause of the problem.

7.2 Hardware support for full virtualization and paravirtualization

Hardware support for virtualization on System x and BladeCenter systems with Intel or AMD processors is a relatively new offering compared to other virtualization capabilities.

The x86 architecture creates some issues when it comes to virtualization. Certain privileged-mode instructions do not trap, and can return different results based upon the mode. For example, the x86 STR instruction retrieves the security state, but the value returned is based upon the particular requester's privilege level. This is problematic when attempting to virtualize different operating systems at different levels. For example, the x86 supports four rings of protection, where level 0 (the highest privilege) typically runs the operating system, levels 1 and 2 support operating system services, and level 3 (the lowest level) supports applications. Hardware vendors have recognized this shortcoming (and others), and have produced new designs that support and accelerate virtualization.

Intel has produced a new virtualization technology that supports hypervisors for both the x86 (VT-x) and Itanium® (VT-i) architectures. Intel Virtualization Technology is comprised of a set of processor enhancements that improve traditional software-based virtualization solutions. These integrated features give virtualization software the ability to take advantage of offloading workload to the system hardware, enabling more streamlined virtualization software stacks and *near native* performance characteristics.

The VT-x supports two new forms of operation, one for the VMM (root) and one for guest operating systems (non-root). The root form is fully privileged, while the non-root form is deprivileged (even for ring 0). The architecture also supports flexibility in defining the instructions that cause a VM (guest operating system) to exit to the VMM and store off processor state.

AMD has produced hardware-assisted virtualization technology called *AMD-V*. Rapid Virtualization Indexing, an enhancement to AMD-V technology in quad-core AMD Opteron™

processors, is designed to increase the performance of virtualized application while enabling faster switching between virtual machines (VMs), so you can host more VMs per server and maximize the benefits of virtualization. AMD's Direct Connect™ Architecture includes an on-die memory controller for optimum memory management—a key performance enabler—while HyperTransport™ technology increases platform scalability and throughput.

Among other things, AMD-V maintains a control block for guest operating systems that are saved on execution of special instructions. The VMRUN instruction allows a virtual machine (and its associated guest operating system) to run until the VMM regains control (which is also configurable). The configurability allows the VMM to customize the privileges for each of the guests. AMD-V also amends address translation with host and guest memory management unit (MMU) tables.

These new technologies can be used by a number of virtualization techniques including Xen, VMware, User-mode Linux, and others.

7.3 VMware ESX Server and VirtualCenter

IBM and VMware can help optimize your IT environment with hardware and software designed for virtualization. With IBM System x, BladeCenter and System Storage offerings running VMware software, applications and services can be deployed in highly reliable and secure virtual machines, which can be provisioned, consolidated and managed centrally, simplifying the IT infrastructure and driving down total cost of ownership so your business can realize innovation.

VMware Infrastructure Version 3.5 includes:

- Support for virtual machines with 64 GB of RAM and physical machines with up to 256 GB of memory
- Storage VMotion enables live migration of virtual machine disks from one data storage system to another without disruption or downtime
- Update Manager automates patches and update management for ESX Server hosts and virtual machines

One of the benefits of virtualization is *stateless* hardware. State-full information and applications can be tied to a virtual machine. The hardware can run just the hypervisor. Therefore, if a system fails, the virtual machine can be migrated to a new hypervisor. VMware ESX 3i is a light weight version of ESX and can be booted from flash memory or the network.[1] This makes the hardware easier to maintain and deploy.

VMware has explained the benefits of having headroom in your servers (the IBM x3850 M2 and x3950 being ideal examples of vertically scalable servers). VMware discusses the benefits of CPU dense ESX server hosts saying:

> *The chance that the scheduler can find room for a particular workload without much reshuffling of virtual machines will always be better when the scheduler has more CPUs across which it can search for idle time. For this reason, it will generally be better to purchase two four-way ESX Server licenses than to purchase four two-way machines. Similarly, two eight-way servers will provide more scheduling flexibility than four 4-way servers.*[2]

[1] Linux-based hypervisors, such as XEN or KVM, can be run stateless as well.

[2] From “Tips and Techniques for Implementing Infrastructure Services on ESX Server” which is available at:
http://www.vmware.com/vmtn/resources/409

Table 7-1 shows that scheduling opportunities scale exponentially rather than linearly when there are more cores available.

Table 7-1 Scheduling opportunities scale exponentially when there are more cores available

ESX host system	Number of cores	Scheduling opportunities (VM = 2 vCPUs)
4P dual core	8	28
8P dual core	16	120
8P quad core	32	496

VMware comes in several editions including:

- VMware Infrastructure Enterprise Edition

  Enterprise Edition contains the entire array of virtual infrastructure capabilities for resource management, workload mobility, and high availability. Includes VMware ESX Server, VMware ESX Server 3i, VMware Consolidated Backup, VMware Update Manager VMware VMotion, VMware Storage VMotion, VMware DRS with Distributed Power Management (DPM), VMware HA.

- VMware Infrastructure Standard Edition

  Standard Edition is designed to bring higher levels of resiliency to IT environments at greater value. Includes VMware HA, VMware ESX Server, VMware ESX Server 3i, VMware Consolidated Backup and VMware Update Manager.

- VMware Infrastructure Foundation Edition

  Unlike the previous VMware Infrastructure 3 Starter Edition, VMware Infrastructure Foundation will have no restrictions on shared storage connectivity, memory utilization or number of CPUs of the physical server. Includes VMware ESX Server, VMware ESX Server 3i, VMware Consolidated Backup and VMware Update Manager.

New features such as VMware High Availability (VMware HA, formerly known as Distributed Availability Services or DAS), Distributed Resource Scheduler (DRS), and Consolidated Backup will provide higher availability, guaranteed service level agreements, and quicker recovery from failures than was ever possible before and coming close to the availability you get from more expensive and complicated alternatives such as physically clustered servers. The System x3850 M2 and x3950 server with their scale-up abilities are uniquely positioned to take advantage of the larger workloads now available to be virtualized.

7.3.1 VMware VirtualCenter

VirtualCenter allows you to rapidly provision virtual machines and monitor performance of physical servers and virtual machines. VirtualCenter intelligently optimizes resources, ensures high availability to all applications in virtual machines and makes your IT environment more responsive with virtualization-based distributed services such as VMware DRS, VMware High Availability (HA) and VMware VMotion.

Case Study: AISO.net generates big cost savings by harnessing the power of data center virtualization.[a] As designed by Sirius, the new solution's key attributes included a marked consolidation in hardware—with nearly 100 of its stand-alone servers replaced by four System x servers—as well as a major improvement in the infrastructure's overall flexibility and resiliency. Four IBM x3650s utilizing VMware Infrastructure 3 with VMotion were used.

Business need: While AISO.net had established itself as a leader among "Green" service providers, its rapid growth forced it to confront the challenges of rising costs and maxed-out capacity. The company needed to transform its infrastructure strategy to have more control of its resources and maintain its superior performance.

Solution: With the help of IBM Business Partner Sirius Computer Solutions, AISO.net achieved a massive consolidation of its data center resources through the use of advanced virtualization technologies deployed in an innovative fashion.

Benefits:

- 60% reduction in power and cooling costs through data center consolidation
- Reduced need to invest in server hardware to accommodate customer growth
- 99.9% availability through automated failover capabilities
- Estimated 50% increase in average server utilization levels

a. Case study available at http://www.ibm.com/systems/x/solutions/infrastructure/virtualization/

7.3.2 IBM Software Support Services: Support Line and VMware ServicePac

IBM provides affordable, unlimited telephone support with highly-skilled technicians who can help with installation and configuration questions, how-to usage, and software defects. You can purchase VMware offerings just as you purchase IBM System x and BladeCenter servers. Simply choose the offering that best meets your needs as you configure your server. When purchasing VMware Infrastructure, you are required to order both the license to the software and the subscription for the software.[3]

7.4 Other x86 Virtualization Software Offerings

VMware has garnered a large share of the x86 virtualization market, and has arguably the most mature x86 virtualization offering. Other vendors have been working to compete in this space, and it will be interesting to see how their products evolve.

There has been a lot of work to introduce virtualization to Linux in the x 86 markets using hypervisor technology. Advantages to a Linux-based hypervisor include:

- The hypervisor has the advantage of contributions from the entire open source communities, not just one vendor (Open Source Solution).
- Linux supports a very large base of hardware platforms, so it is not limited to just the platforms certified by a single vendor. Also, as new technologies are developed, a Linux based hypervisor can take advantage of said technologies, such as iSCSI, InfiniBand, 10 Gig-Ethernet, and so on.

We include comments on a few of the x86 virtualization offerings in this section, but this is by no means a complete list.

[3] IBM sells the purchaser entitlement to the VMware products and the products are distributed directly from VMware to the purchaser.

7.4.1 Xen

Xen is an open source virtual machine monitor for x86-compatible computers. Xen makes it possible for multiple guest operating systems to run on a single computer by using a software layer called a hypervisor to mediate access to the real hardware.

Xen originated as a research project at the University of Cambridge, led by Ian Pratt, senior lecturer at Cambridge and founder of XenSource, Inc. The Xen hypervisor is a unique open source technology, developed collaboratively by the world's best engineers at over 20 of the most innovative data center solution vendors, including Intel, AMD, Cisco, Dell, Egenera, HP, IBM, Mellanox, Network Appliance™, Novell, Red Hat, SGI, Sun™, Unisys, Veritas, Voltaire, and XenSource.

The first public release of Xen was made available in 2003. XenSource, Inc. was acquired by Citrix Systems in October 2007. XenSource's products have subsequently been renamed under the Citrix brand:

- XenExpress was renamed *XenServer Express Edition* and *XenServer OEM Edition* (embedded hypervisor)
- XenServer was renamed *XenServer Standard Edition*
- XenEnterprise was renamed *XenServer Enterprise Edition*

Citrix Systems completed its acquisition of XenSource[4] on 22 October 2007, and the Xen project moved to:

http://xen.org

XenSource Inc. and Virtual Iron Software Inc. are promoting Xen as the primary open source competitor to commercial virtualization products such as VMware. Red Hat Inc. includes the Xen hypervisor as part of Red Hat Enterprise Linux (RHEL) software, describing this combination as *integrated virtualization*.

7.4.2 Microsoft Virtual Server

Microsoft Virtual Server is a virtualization solution that facilitates the creation of virtual machines on the Windows XP and Windows Server® 2003 operating systems. Originally developed by Connectix, it was acquired by Microsoft prior to release.

The current version is Microsoft Virtual Server 2005 R2 SP1. New features in R2 SP1 include Linux guest operating system support, Virtual Disk Precompactor, SMP (but not for the Guest OS), x64 Host OS support (but not Guest OS support), the ability to mount virtual hard drives on the host OS and additional operating systems including Windows Vista®. It also provides a Volume Shadow Copy writer that enables live backups of the Guest OS on a Windows Server 2003 or Windows Server 2008 Host. A utility to mount VHD images is also included since SP1.[5]

Virtual Server 2005 has some limitations. Although Virtual Server 2005 can run on hosts with x64 processors, it cannot run guests that require x64 processors (guests cannot be 64-bit). It also makes use of SMP, but does not virtualize it (it does not currently allow guests to use more than one CPU each)[6].

[4] Press release available at http://www.citrix.com/English/NE/news/news.asp?newsID=683171
[5] See http://en.wikipedia.org/wiki/Microsoft_Virtual_Server
[6] Virtual Server 2005 R2 SP1 Download Frequently Asked Questions found at: http://www.microsoft.com/technet/virtualserver/software/faq.mspx

Microsoft has announced the launch of their enterprise-level Hyper-V™ hypervisor, which is a key feature of Microsoft Windows Server 2008. Microsoft says that like Virtual Server, Windows Server 2008 Hyper-V provides server virtualization. Support for hardware virtualization will be available in Windows itself. This new approach provides a hypervisor that runs directly on the hardware. One or more partitions can then be created on top of the hypervisor, each providing a VM. The low-level support provided by the Windows hypervisor lets virtualization be done more efficiently, providing better performance.

The final version of Hyper-V is scheduled to ship shortly after the release of the new operating system. It will be available for all three 64-bit editions of this new operating system: Standard, Enterprise, and Data Center.

7.4.3 Virtual Iron

Virtual Iron is a Xen-based virtualization solution that enables server partitioning for single and multi-server configuration, no downtime virtual server migration (LiveMigrate), advanced management capabilities for rapid provisioning, high availability and disaster recovery (LiveRecovery) and capacity management (LiveCapacity).

Virtual Iron's latest release is Version 4.2 (released in December 2007), with enhancements that focus on speed, continuity, and overall ease of use. There is now support for 32-bit and 64-bit versions of both SUSE Linux Enterprise Server 10 and Red Hat Enterprise Linux 5.

Major enhancements include:

- Multi-pathing for virtual server Ethernet and Fibre Channel networks to support business continuity and redundancy.
- LiveSnapshot: virtual server snapshots for hot backup and patch management. These capabilities enable off-loaded, space efficient and no-downtime backups on live virtual machines running in production environments and also reduce the time for virtual machine patching in development and test processes.
- The ability to increase the size of both disk groups and virtual disks dynamically, providing increased storage on demand.

Virtual Iron Version 4.2 comes in three versions:

- A free version of the software supports up to 12 virtual machines on one physical machine
- An Enterprise Edition, at a per socket cost
- An Extended Enterprise at a per socket cost

Virtual Iron is currently pricing its software to be extremely competitive. Virtual Iron might not have all the features that VMware does, but its price points can make it very attractive to customers, particularly in the small and medium business (SMB) space.

7.4.4 KVM

Kernel-based Virtual Machine (KVM) is a full virtualization solution for Linux on x86 hardware containing virtualization extensions (Intel VT or AMD-V). It consists of a loadable kernel module, kvm.ko, that provides the core virtualization infrastructure and a processor specific module, kvm-intel.ko or kvm-amd.ko. This allows a virtual machine to run as a standard Linux process. KVM uses the processor virtualization technology (Intel VT or AMD SVM) to virtualize the hardware. It also provides a user space component to emulate PC hardware.

KVM was introduced to the Linux kernel as of 2.6.20, but it can be run in earlier kernels because it is a kernel module. The advantage of KVM is that it is a kernel module and, therefore, supports any platform that the host OS supports.

7.5 The benefits of choosing IBM System x and BladeCenter

You need easier ways to manage and optimize IT systems. A virtualized environment can help save money, make IT staff more productive, and simplify system resource allocation. IBM System x and BladeCenter servers provide a robust and very affordable platform designed for virtualization, enabling you to run interactive applications by day, maintenance, inventory or computational programs at night.

Highlights include:

- Increase productivity and utilization with IBM System x and BladeCenter scale-up and scale-out virtualization solutions.
- Deliver more reliable performance in less physical space with BladeCenter technology, providing real cost savings in the data center.
- High performance System x servers aim to provide breakthrough 64-bit performance, business-critical reliability, and improved scalability.
- System x and BladeCenter hardware provides the manageability that helps solve real client problems as they relate to a server over its lifetime.

7.5.1 X-Architecture

These X-Architecture® technologies differentiate IBM System x from other Intel servers, in many ways, including the following:

- Active Memory™ with features such as Memory ProteXion, hot-add memory, and Chipkill memory provide a level of reliability and availability that helps reduce downtime and maintain data integrity. Chipkill memory (based on memory used in the NASA Pathfinder's 1997 mission to Mars) helps keep the server up and running by tolerating a full chip failure.
- Two times the memory availability than previous generations with 32 DIMM slots, running DDR II PC2-5300 creates a more balanced total system design.
- Active PCI allows for installation or replacement of PCI Express adapters without taking the system down.
- Light-path diagnostics provides rapid identification of failing or failed components, helping you minimize downtime.
- Extensive Predictive Failure Analysis® (PFA) support helps predict failures on fans, VRM, and power supplies.

The ability of application servers to deliver high performance varies greatly, depending on the specifics of any given customer's workload. IBM raises the bar for scalable, high-performance application servers with the introduction of IBM eX4 Architecture, the ground-breaking fourth-generation of the IBM Enterprise X-Architecture technology. Incorporating this chipset,

the IBM System x3950 M2 and x3850 M2 deliver the performance, availability, expandability, and manageability required for the next generation of industry-standard servers. The result is demonstrated by recently published record-setting benchmark results.[7]

At the heart of IBM eX4 architecture is the *Hurricane 4* memory controller, with 60% front-side bus speed improvement and up to 60% faster memory. Separate front-side bus for each processor socket means a more balanced system design. Figure 7-1 shows the architecture of the x3950 M2.

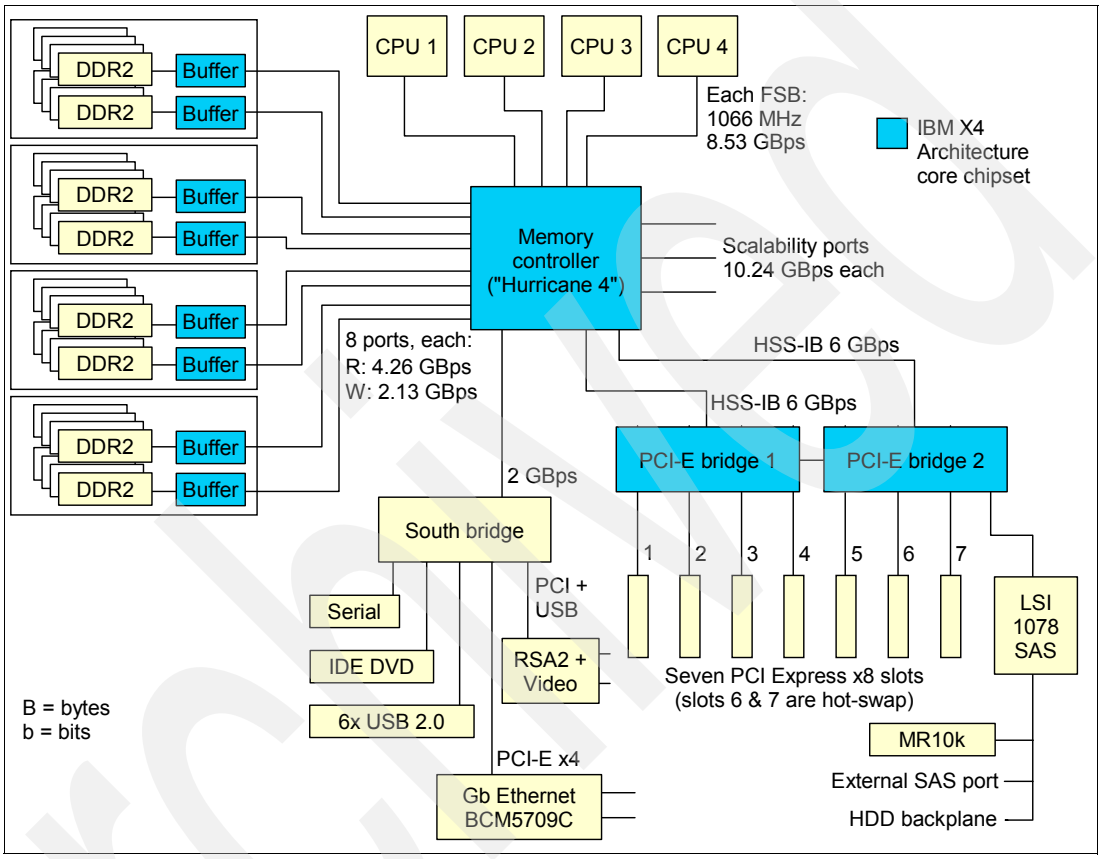


Figure 7-1 Block diagram of the System x3950 M2

The reduction front-side bus contention is achieved through a *snoop filter* in the Hurricane controller that stores the memory addresses for data contained in every processor's L2 cache. To compare, for the competitor's architecture, all reads that are initiated by any processor (due to a miss in the processors cache) must check to see if that data is in any other processors' caches before it can obtain and use it from main memory. These cache *snoops* contribute to a significant amount of traffic on the FSB, because for each processor cache miss, the memory controller has to initiate transactions to and from all of the other processors. On a system under load, this level of traffic can have drastic effects on FSB transaction latencies due to the number of snoops performed, and the amount of queuing of these snoops that occurs on the memory controller.

Looking at the same situation using eX4 Architecture, a processor cache miss will invoke a lookup into the memory controller's snoop filter while at the same time requesting the data from the main system memory. If the memory address is found in the snoop filter, the memory controller will request the data from the processor that contains the address. However, more often than not, the memory address is not found in the snoop filter, meaning no additional

[7] See http://www.ibm.com/systems/x/resources/benchmarks/

traffic is placed on the FSB, and the memory controller will use the data returned from the main system memory. This dramatically reduces the amount of traffic on the FSB, freeing the processors from excess snoop traffic, and greatly reduces transaction queue depth (and, thus, the latencies) on a memory access.

Each x3950 M2 and x3850 M2 can have up to 4 processors installed (each with up to four cores), and up to 256 GB of memory installed. These servers can be also be connected together (the x3850 M2 first requires the ScaleXpander Option Kit) to form a much larger single-image server, up to four systems or *nodes*. With four nodes connected together, the single complex has 16 processors (64 cores), and up to 1 TB of RAM. This type of system is suited to applications such as massive database engines.

7.5.2 IBM BladeCenter

IBM BladeCenter can be likened to physical partitioning. BladeCenter allows you to centralize computing power for physically separate environments while sharing infrastructure components like power and switches, and centralizing management. BladeCenter supports a variety of environments including Windows, Linux, Solaris™, AIX, and other environments all within the same chassis.

BladeCenter also supports virtualization. VMware support for BladeCenter can be found on ServerProven®.[8] BladeCenter is the best of both worlds, because it incorporates this physical partitioning along with traditional virtualization (using VMware, Microsoft Virtual Server, Xen, and so forth), which contributes to its flexibility as a solution.

BladeCenter is innovative technology that can help reduce the cost of ownership with a more efficient use of valuable floor space due to its modular design. Its simplified management can help speed up such tasks as deploying, re-provisioning, updating, and troubleshooting hundreds of blade servers with relative ease. All this can be done remotely with one graphical console using IBM Systems Director systems management tools. In addition, blade servers can provide improved performance by potentially doubling current rack density, thereby permitting you to integrate your resources and share key components. This can result in cost reductions and increased availability.

Employing blade servers can help you use available rack and floor space more efficiently. For example:

- A single rack can hold six BladeCenter chassis and 84 dual processor blade servers, giving you a total of 168 CPUs per rack, which is twice the current density of comparable 1U non-blade server.
- Blade servers support a minimum of one 4 port 1 GB Ethernet switch for up to 14 blade servers. You can use a total of four switch modules within the BladeCenter, combining of Fibre Channel or Ethernet

The underlying hardware forms an important base for virtualization (for example, I/O bandwidth). Virtualization means more applications per server, which puts more demands on the system. The BladeCenter H and Cisco VFrame Solution offers:

- Fabric sharing and I/O consolidation; other solutions do not necessarily offer fabric consolidation.
- The IBM offering is high-speed, including 10 Gb Ethernet and 4X InfiniBand, which also supports 1 Gb Ethernet, 4 Gb Fibre Channel, SAS, and iSCSI.

[8] See http://www.ibm.com/servers/eserver/serverproven/compat/us/nos/vmwaree.html

- The solution from IBM is based on industry standards and is compatible across all BladeCenter chassis and many OEM switches. Other vendors designs can be proprietary and limited to only certain chassis and switches.
- The IBM solution scales to 512 servers.
- The offering from IBM is simple to use and virtualizes the server, workload, and network connectivity, offering policy-driven automated blade failover.

Another offering, IBM BladeCenter Open Fabric Manager, allows for the I/O virtualization of Ethernet and Fibre Channel connections within a system by providing users the option to assign Ethernet and Fibre Channel port addresses used by their server blades through software as an alternative to the addresses that are burned in to the hardware during manufacturing. This will be discussed in greater detail in next section.

7.5.3 Hardware address virtualization: BladeCenter Open Fabric Manager

Blade technology has enabled a new form of virtualization, *hardware address virtualization*. This technology is the ability to modify and control the hardware addresses of network and storage controllers. For NICs, these addresses would be the MAC addresses, and for SAN network controllers, these addresses would be the world wide names (WWNs). Traditionally, these addresses are unique addresses that are used to identify every hardware device in the network or SAN. This allows servers to be uniquely identified on the SAN or Network and have access to SAN and Network resources. If a server is replaced or upgraded, new MAC addresses and WWNs are used, and the fabrics need to be update for these new addresses.

With hardware address virtualization, these address no longer have to be tied to a specific hardware device. They can be maintained in a management device and assigned to servers as needed. A replacement of a failed server no longer requires the propagation of new addresses, because the existing addresses can be maintained. Maintaining the existing address is very important on the SAN, because access to storage resources is normally controller through WWNs. SAN administrators have to collect servers' WWNs and create tables in their SAN storage to assign which servers can access which storage. If a server or HBA is replaced, these tables need to be updated after gathering the new WWNs, which is often a manual process. If the WWNs can be reused, then the transition to a replacement server on the SAN can be automated. This action is significant if a server is booting from the SAN.

With SAN boot, the server no longer has its operating system on a local hard drive. The server's "personality" can be maintained on the SAN fabric. Therefore, if a server is replaced, its role can be assigned quickly to the replacement server, enabling a new form of high availability in servers, with failover taking place at the hardware layer. Pools of standby servers can be created to take over for a failed server with very simple reassignments of hardware addresses. Also, server replacement and upgrades can be simplified to plug and play.

This capability is all possible with blade technology, as blade servers share resources in a common infrastructure, and are controlled by a common management controller. The MAC addresses and WWNs can be stored in the Management controller and assigned to blades. This process can be automated to create plug and play and hot standby servers. This feature is implemented in the IBM BladeCenter with Open Fabric Manager (BOFM). BOFM is implemented differently than other solutions as it stores the hardware addresses in the management controller, whereas other solutions can store the hardware addresses in each individual blade switch module. BOFM can operate independent of fabric switches, versus requiring specific fabric devices.

IBM BladeCenter Open Fabric Manager benefits include:

- Management of over a thousand servers

  A single Advanced Management Module with Open Fabric Manager can allow a LAN and SAN administrator to pre-assign MAC and WWN addresses for up to 100 chassis or up to 1400 blade servers.

- Reduced deployment times

  BOFM can help cut deployment times from days to hours or from hours to minutes, depending on the number of servers in your domain, allowing you to go into production much faster than in a rack environment. Simply plug the server in your chassis, and it inherits the addresses that you preassigned.

- Simplified management

  The Advanced Management Module (AMM) provides a Web-based interface to control the Open Fabric Manager functions. From the AMM the administrator can create addresses for the blades, save the addresses to a file, deploy the addresses to the blade slots in the same chassis or in up to 100 different chassis. This administration can all be done without any servers installed in the chassis.

- Failover capability

  BOFM offers an advanced utility that provides automatic blade failover allowing for the configuration of a standby blade pool. The standby blade pool can exist within the chassis or within another chassis in the 100 chassis domain. You are also able to create an event action plan for the individual blades that you want to monitor for failures. BOFM detects if a blade has failed or has been removed automatically and without intervention.

7.5.4 Systems management

IBM Systems Director is a comprehensive systems-management solution that is designed to help reduce costs and improve productivity.

IBM Systems Director:

- Delivers advanced System x and BladeCenter manageability with a portfolio of exclusive, predictive server tools. Based on industry standards, it is supported across IBM platforms.
- Automates many of the processes that are required to manage systems proactively, including capacity planning, asset tracking, preventive maintenance, diagnostic monitoring, troubleshooting, and more. It has a graphical user interface that provides easy access to local and remote systems.
- Can be used in environments with multiple operating systems (heterogeneous environments) and integrated with robust workgroup and enterprise management software from IBM (such as Tivoli software), Computer Associates, Hewlett-Packard, Microsoft, NetIQ, and BMC Software.
- Is included with the purchase of System x and BladeCenter systems and is offered for sale to help manage select non-IBM systems. Optional, fee-based extensions to IBM Systems Director are available if you want more advanced management capabilities.

8

Storage virtualization

Storage area networks enable you to share homogeneous storage resources across the enterprise. For many companies, however, information resources are spread over various locations and storage environments with products from different vendors. The best solution takes advantage of the investment already made and provides growth when needed.

We discuss the following topics in this chapter:

8.1 SAN Volume Controller

The IBM System Storage SAN Volume Controller (SVC) helps manage the complexity and costs of SAN based storage. Based on virtualization technology, SVC supports a virtualized pool of storage from the storage systems attached to a SAN. This storage pool helps you tap unused storage capacity and make your businesses more efficient and resilient. SVC helps simplify storage management by presenting a single view of storage volumes. Similarly, SVC is an integrated solution supporting high performance and continuous availability in open systems environments.

The solution runs on clustered storage engines, based on System x servers and open-standards-based technology. Industry-standard host bus adapters (HBAs) interface with the SAN fabric. SVC represents storage to applications as virtual disks, created from the pool of managed disks residing behind the storage engines. Your storage administrators can scale performance by adding storage engines and capacity by adding disks to the managed storage pool.

8.1.1 Features and benefits

The SAN Volume Controller has the following features and benefits:

- Manage storage volumes from your SANs and combine the capacity of multiple storage controllers, including storage controllers from other vendors, into a single resource with a single view of the volumes
- Migrate data from one device to another without taking the storage offline and reallocate, scale, upgrade, and back up storage capacity without disrupting applications
- Increase storage capacity utilization and uptime, as well as administrator productivity and efficiency, while taking advantage of existing storage investments through virtualization and centralization of management
- Support advanced copy services across all attached storage, regardless of the intelligence of the underlying controllers
- Make better use of existing storage using virtual disks and management easier with a comprehensive easy-to-use graphical interface that incorporates the Storage Management Initiative Specification (SMI-S) application programming interface (API)
- Supports local area network (LAN) free and server-free backups and the IBM SubsystemDevice Driver (SDD) multi-pathing software and a variety of multi-pathing drivers
- Manage up to eight petabytes (PB) of total usable storage capacity and higher by adding storage engine pairs to the initial configuration
- Apply copy services across disparate storage devices within the network because advanced copy services, such as FlashCopy® and Remote Mirror and Copy, are supported across the managed storage
- Respond with flexibility and speed

SVC is designed to deliver significant value to organizations facing the challenges of today's explosive growth in information. SVC helps combine capacity from different storage systems, helps provide common copy functions and enable data movement without server disruption, and is designed to support management of diverse storage from a single point. Refer to Figure 8-1.

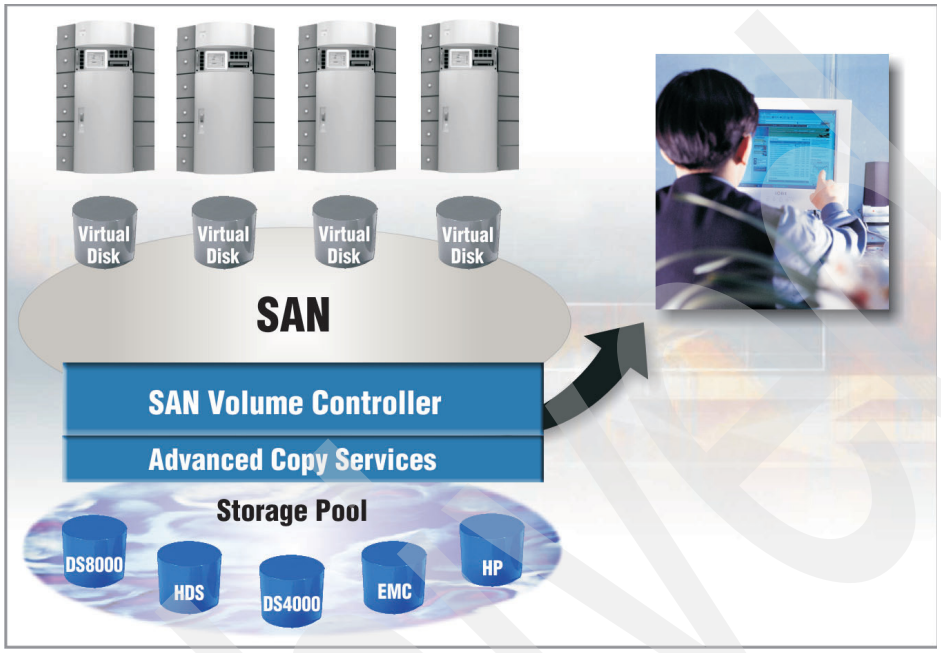


Figure 8-1 SAN Volume Controller

In just over four years to 2007, IBM has installed 10 000 SVC engines at more than 3400 customer installations. IBM is the leader in the industry for Block Level Virtualization. SVC is designed to support up to world class levels of availability and has one of the fastest industry-standard storage performance benchmarks recorded for any disk controller.[1] SVC supports the virtualization of non-IBM as well as IBM storage (Including EMC, HP, HDS, Sun, and Dell).

Helps keep your applications running

Organizations often have to take their storage offline to migrate data or change the storage infrastructure. These actions interrupt applications accessing their data, which increases application downtime. To migrate data (for example from one vendor's array to another), you might need to obtain specialized services-at high rates because of the skills required.

SVC is designed to help you avoid the need to take your storage offline. SVC's data migration capabilities support continued access to data while data migration and storage infrastructure change is occurring. Continuous availability of data for applications can help you to avoid the cost and impact of downtime. In addition, SVC is designed to work with IBM TotalStorage Productivity Center to help automatically allocate more capacity to an application that needs it, again helping your business to run without interruption.

[1] See http://www.storageperformance.org/results/

High availability

SVC has the following high availability features:

- Each SVC node is covered by a pair of uninterruptible power supply.
- Each node in a node pair mirrors it cache to its partner nodes, and I/O complete is not returned to the server until it has been written in each of the two nodes of a I/O group.
- The SVC code can be upgraded nondisruptively.
- The SVC nodes use System x servers with fault-tolerant features including Chipkill memory, multiple HBAs and the HBAs use CRC checking on the data flowing in and out of the system.

All of this comes to together to deliver a very highly reliable Enterprise class Storage controller package.

Helps reduce the cost and complexity of the storage environment

SVC can help you take advantage of the cost savings of midrange storage while achieving the benefits of enterprise storage. Because the cost of midrange storage can be up to 50% less than enterprise storage, this can be a significant amount. SVC allows you to create tiered classes of storage to help you use the most efficient type of storage for the types of data you are storing. SVC is designed to provide robust attributes of enterprise storage across the entire virtualized environment.

In many environments, organizations have different types of storage that are managed in different ways and that have different functions (such as EMC TimeFinder, Hitachi ShadowImage, or IBM FlashCopy). To manage these different systems requires developing many sets of skills. In addition, organizations can develop automation based on one of these functions. If they do, they might be unable to move data from one system to another because the applications using that data become *tied* to the functions on one system, reducing choices among vendors and reducing flexibility in using storage systems.

SVC is designed to provide a single place to manage multiple, different disk systems. It also provides a common set of functions that are designed to work in the same way regardless of the disk system type. In this way, SVC can help avoid vendor lock-in and reduce management complexity, both of which can help you reduce storage costs.

Case Study: CINECA, Italy's largest high-performance computing (HPC) research center used SVC and DS4500 storage to reduce their storage costs 10% in their first year.[a]

a. Case study available by searching `http://ibm.com/software/success` for CINECA.

Helps make people more productive

SVC creates a virtualized pool of your heterogeneous storage environment. By executing storage management tasks for the entire pool from a central point, SVC can help storage administrators become much more productive. They only have to learn one interface, and there are fewer tasks to execute because the action can be taken across the virtualized pool. Together, these can generate additional operational savings.

Case Study: The Natural History Museum virtualized their data storage with IBM Business Partner Tectrade and IBM SAN Volume Controller.

SVC gives us a single interface through which we can manage the entire storage environment, however large it grows," adds Paul Richards, Head of Information and Communication Technology. "We can expand easily and seamlessly without increasing the administrative workload, and even add other forms of storage like optical media if needed.[a]

a. Case study available by searching the following Web site for *Natural History Museum*: http://ibm.com/software/success

Helps utilize storage assets more efficiently

It is often estimated that many organizations are only using their storage assets at about 25% to 50% capacity. SVC is designed to help organizations significantly improve their utilization by combining the storage capacity from many disk arrays into a single storage resource, which storage administrators can manage from a central point. Administrators no longer need to keep spare space available for each individual server in case it is needed. Rather, they have a shared pool of space that is shared among all servers and so it can be smaller.

SVC also applies copy services, such as point-in-time copies and replication across disparate storage arrays to further help you increase the utilization of your assets.

Circuit City boosts IT infrastructure effectiveness with storage virtualization[a]

Business need: Build a scalable storage infrastructure that can keep pace with business growth; improve administrator productivity as well as application and server availability

Solution: A virtualized storage environment based on IBM System Storage SAN Volume Controller, IBM TotalStorage Productivity Center software and IBM System Storage DS4500 and DS8100 disk storage systems

Benefits: The use of tiered storage architecture helps to reduce the overall cost of storage; nondisruptive data migrations enabled by SVC reduce maintenance windows from hours to minutes, which helps improve application availability; the multi-tier storage environment helps enable

a. Case study available by searching the following Web site for *Circuit City*: http://ibm.com/software/success

8.2 Virtualization Engine TS7520: Virtualization for open systems

The IBM Virtualization Engine TS7520 combines hardware and software into an integrated tiered solution designed to provide tape virtualization for open systems servers connecting over Fibre Channel and iSCSI physical connections. When combined with physical tape resources for longer term data storage, the TS7520 Virtualization Engine is designed to provide an increased level of operational simplicity and energy efficiency, support a low cost of ownership and increase reliability to provide significant operational efficiencies.

One of the biggest issues with backup planning today is that the amount of data that is backed up is growing but the time allotted for a backup (the backup window) is shrinking or remaining static.

With backup windows shrinking, tolerance for hardware failure has virtually disappeared. The TS7520 Virtualization Engine is designed to help address these issues by reducing tape mechanical delays and providing fault tolerant architecture options supporting high availability.

8.2.1 Key functions of the TS7520

Key functional features of the TS7520 include:

- Designed to manage growth to 1.3 PB
- Enhanced Caching provides the backup application direct access to data whether in cache or on physical tape and a more intelligent policy-based data migration
- Encryption, with network or backstore tape to help protect sensitive customer data
- Hardware assisted compression designed to improve system performance with replication, compression, and encryption
- Network Data Management Protocol (NDMP) provides NAS tie-ins for data movement over the network
- Control Path Failover and Data Path Failover can help provide higher availability over the control path and SAN
- Hosted Backup allows you to run supported backup applications on the TS7520
- iSCSI provides latest open systems standard
- Virtual support of LTO 2 and LTO 3 Tape Drives and 3592 Tape Drives Model J1A/E05
- Virtual support of a 3584 Tape Library
- Physical direct attach support for 3494, TS3310, TS3100/3200, and TS3500 Tape Libraries
- Configuration of two and four TS7520 Virtualization Engines as an active-active cluster
- Support for real-time compression of data, reducing disk storage requirements
- On demand allocation of disk storage to help maximize storage use with virtual cartridges, with static allocation also supported for customized environments
- Interaction with the TS7520 Cache Controllers to perform transparent failover/failback from path (HBA, port, switch, channel) or storage controller failure to minimize disruption to backup or restore activities

8.2.2 TS7520 benefits

The TS7520 offers clients many benefits, including:

- Remote Copy: Allows for the copied tape to reside on a remote TS7520 or vault
- Scalable Configuration: Allows for extensive data growth, 9.75 TB to 1.3 PB uncompressed, depending on the configuration
- Hardware Assisted Compression: Designed to provide increased performance during backup/restore, helps alleviate CPU cycles in compression mode
- Active Failover: Designed to provide automated policies for TS7520 node high availability
- Control Path and Data Path Failover: Allows the TS7520 to provide higher availability to the servers it supports

- TS7520 Supports 1, 2, or 4 servers: Designed to provide high availability options, allows for the TS7520 Virtualization Engine to be shared, designed to provide higher performance options
- 4096 Virtual Drives: Allows for faster backup due to virtual resources allocated versus physical tape drives
- Encryption Support: Designed to provide data security over the network and when the data is at rest on the tape cartridge using the IBM TS1120 Tape Drive

8.3 Virtualization Engine TS7700: Mainframe virtual-tape

The IBM Virtualization Engine TS7700 is a mainframe virtual-tape solution that is designed to optimize tape processing. Through the implementation of a fully integrated tiered storage hierarchy of disk and tape, the benefits of both technologies can be used to help enhance performance and provide the capacity needed for today's tape processing requirements. Deploying this innovative subsystem can help reduce batch processing time, total cost of ownership, and management overhead.

The TS7700 Virtualization Engine is designed to reduce or eliminate a number of bottlenecks that might be present in any given tape environment. Any reduction in bottlenecks will vary depending on the characteristics of installed equipment and workload, but reducing bottlenecks can help reduce batch processing time.

For example, if the batch process is constrained, the TS7700 Virtualization Engine's processing power, cache capacity and support for 4 Gbps IBM FICON attachment can help alleviate bottlenecks and reduce the batch window.

8.3.1 Key functions of the TS7700

The functions of the TS7700 include:

- Fewer "real" tape mounts, because most mount requests are satisfied from the Tape Volume Cache (TVC)
- Fewer physical tape cartridges required due to higher use of cartridge capacity
- Less floor space required to support the tape process as a result of fewer physical resources
- Improved performance due to the elimination of most of the physical movement of tape
- Reduced operating costs such as power, maintenance, operations, and support staff
- Enhanced reclamation policies to help migrate data between pools and provide efficient storage of long term data
- Advanced Policy Management features:
  - Cache management provides the ability to manage virtual volume retention or deletion from the tape volume cache.
  - Volume Pooling provides the ability to group selected logical volumes on physically separate cartridges or cartridge pools.
  - Selected Logical Volumes Dual Copy provides the ability to create a second copy of the volume on a separate cartridge.
- Cross Server Data Sharing
- Remote Dual Copy for use with disaster backup and recovery or remote tape vaulting

8.3.2 TS7700 benefits

The TS7700 offers clients many benefits, including:

- Tiered hierarchy of disk and tape storage: Helps improve performance and capacity to help achieve a low total cost of ownership for tape processing.
- Maximum of 768 virtual drives and 1,000,000 virtual volumes: Helps reduce or eliminate bottlenecks in your tape environment, supports the re-reference of volumes without the physical delays typical to tape I/O, and helps increase performance of tape processes.
- Advanced policy management: Cache management for volume retention and deletion, volume pooling that allows grouping of logical volumes on physically separate cartridges or cartridge pools.
- Grid capability: Two or three TS7700s can be configured to participate in a GRID environment. This configuration helps to support high availability requirements, helps to enhance availability during planned maintenance, service, or system upgrades and avoids the transportation of cartridges in the event of a disaster.
- Supports TS1120 tape drive-based encryption: Helps secure sensitive at-rest data.
- Manageability: The IBM TS3000 System Console allows IBM Technical Service to download new microcode and remotely monitor the installation and can automatically dispatch a service representative when required

9

Systems and virtualization management

Management is key to the success of a virtualization solution. There are several tools that are available to help simplify and manage virtual systems. IBM Systems Director takes a hardware virtualization focus, whereas the IBM Tivoli suite of management products focuses more at the service level and from an enterprise view.

IBM Systems Director unifies its industry-leading server and storage management products—IBM Systems Director and IBM TotalStorage Productivity Center—with enhanced virtualization management support. The IBM Systems Director family provides a modular, open standards-based set of solutions that can be tailored easily to fit the requirements of any size business and can be integrated seamlessly into enterprise management solutions from IBM Tivoli.

IBM Tivoli Service Management enables clients to better manage their infrastructure, operations, and IT processes, to more effectively deliver services aligned to business goals. Systems Director can take information from various systems and platforms and feed it to the Tivoli Enterprise management products

We discuss the following topics in this chapter:

9.1 Managing a virtualized IT infrastructure with IBM Software

IBM IT management solutions deliver operational management products to visualize, control, and automate the management of the virtual environment. Together these technologies and products allow the business to increase workload velocity, increase utilizations, respond to changing market conditions faster, and adapt to customer requirements. The IBM IT operational management products today provide significant management capabilities of virtual environments. The flexibility provided by the IBM systems management products allows the IT infrastructure to be multivendor, heterogeneous, and still remain manageable by IBM.

Managing virtualized environments creates new and unique requirements. The ability to take a virtual machine and resize the memory or CPU power dynamically brings new capabilities to the business that can be exploited to deliver higher efficiencies. The ability to move virtual machines from physical host to physical host while the virtual machines remain operational is another compelling capability. The ability to fully exploit virtual and physical resources to optimize and adapt workload velocity and resource usage yields reduced operational complexity and improves return on assets. As a result, systems management software is adapting to accommodate the new and evolving abilities of the mixed virtual and physical IT infrastructure.

One of the key values that IBM systems management software can provide is to mask the complexities that are introduced by virtualization. Businesses are embracing virtualization because it brings value and it enhances capabilities for business continuity and disaster recovery. The ability to use business policy based process automation for orchestrating, provisioning, workload, and service level management all in line with business goals will drive higher levels of virtualization adaptation. As the use of virtualization continues to be driven deeper into the data center, system management vendors must update existing tools to handle this dynamic infrastructure by adapting these tools to work with both virtual and non-virtual machines.

9.2 IBM Systems Director

IBM Systems Director is the platform management family that provides IT professionals with the tools they need to better coordinate and manage virtual and physical resources in the data center.

The cost of managing the IT infrastructure has become the largest and fastest-growing component of overall IT spending for many organizations. Virtualization helps address this cost through the consolidation of physical resources. However, it also adds complexity by creating a sharp increase in the number of managed virtual resources. IT professionals are seeking more advanced capabilities and tools for managing both their physical and virtual systems across multiple architectures and environments.

IBM Systems Director helps address these needs by unifying under one family its industry-leading server and storage management products—IBM Systems Director and IBM TotalStorage Productivity Center—with newly enhanced virtualization management support. With capabilities that include configuration, discovery, health and status monitoring, automated response, and power and virtualization management, the IBM Systems Director family gives IT professionals what they need for managing both physical and virtual systems across multiple IT environments.

As virtualization becomes reality in today's IT infrastructures, the IBM Systems Director family can help businesses realize the full potential by providing a unified approach to platform management designed to lower IT operational costs and increase productivity.

IBM Systems Director is a cross-platform hardware management solution that is designed to deliver superior hardware manageability, enable maximum system availability, and help lower IT costs. IBM Systems Director helps you get started faster and be more productive. Enhancements improve ease of use and deliver a more open, integrated toolset. Its industry-standard foundation enables heterogeneous hardware support and works with a variety of operating systems and network protocols. Taking advantage of industry standards allows for easy integration with other systems' management tools and applications.

This comprehensive hardware management solution includes:

- Inventory of hardware features and settings
- System health notification
- Proactive and automated systems management

IBM Systems Director delivers comprehensive, remote hardware management with:

- New, powerful user interface that works the way you do
- Lighter product footprint that offers more choice about the amount of hardware management you want
- Open, standards-based design that facilitates management of heterogeneous hardware environments
- Enhanced integration with higher-level management products, including Microsoft SMS and Microsoft Operations Manager (MOM)
- Cross-platform hardware management solution; serves as the common thread across IBM Systems Director
- Self-managing, smart tools
- Easy installation and setup
- Comprehensive BladeCenter support, with easy, single point of configuration, deployment, and management

IBM Systems Director is included with the purchase of IBM System p, System x, and BladeCenter systems and is offered for sale to help manage select non-IBM systems. Optional, fee-based extensions to IBM Systems Director are available if you want more advanced management capabilities.

Use IBM Systems Director stand-alone or with existing enterprise or workgroup management environments to access and manage physically dispersed IT assets more efficiently. Flexible capabilities help you realize maximum system availability and lower IT costs. With IBM Systems Director, IT administrators can view and track the hardware configuration of remote systems in detail and monitor the usage and performance of critical components, such as processors, disks, and memory.

In addition to the improvements to IBM Systems Director, enhanced extensions are also available. These tools are tightly integrated with IBM Systems Director for consistent management from a single console.

9.2.1 Features

IBM Systems Director has the following features:

- An easy-to-use, integrated toolset helps you get started faster, accomplish more, save time, and manage more systems per technician.
- Easy installation and setup
- Support for non-IBM hardware: innovative use of industry standards from Common Information Model (CIM) to SNMP enables heterogeneous hardware management, protecting your existing IT investment
- Seamless integration protects your investments in other management packages with more extensive hardware manageability.
- Integrated, centralized SQL database: internal database makes system-related data available, even when the specific system is not directly available.
- Multiple operating system support: IBM Systems Director smoothly handles a variety of operating systems.

Integration with and complementary function for your existing workgroup or enterprise systems management applications, including:

- IBM Tivoli Enterprise Framework
- Tivoli NetView® NT
- Computer Associates Unicenter TNG
- Microsoft SMS and Microsoft Operations Manager
- HP OpenView Network Node Manager
- BMC Patrol
- NetIQ

You can potentially lower the total cost of managing and supporting your networked systems. By deploying IBM Systems Director, you can achieve reductions in ownership costs through:

- Reduced downtime
- Increased productivity of IT personnel and end users
- Reduced service and support costs

9.2.2 IBM Systems Director extensions

The IBM Systems Director family provides optional management extensions that support a broad set of capabilities ranging from virtualization management, to IT optimization, to remote management and much more. System Director extensions are designed to be modular, allowing IT professionals to tailor their management capabilities to their specific needs and environment.

You can extend IBM Systems Director for greater management and optimization with optional tools that integrate into the IBM Systems Director interface for consistent, single point of management.[1] There are several areas of management that the Systems Director extensions address (including Advanced System Monitoring and Replication Management), but the focus of the paper is virtualization and optimization.

[1] IBM continues to add new features and functionality, but not all functionality is available across all platforms today. Platform support for particular features is available upon request.

IBM Systems Director for Virtualization

IBM System Director can be extended to provide the following virtualization functions:

- Discover, visualize, and manage both physical and virtual systems from a single console

  The IBM Virtualization Manager extension simplifies management of VMware, Xen Microsoft and POWER based Virtual Server environments. Virtualization Manager also integrates with and complements VMware VirtualCenter, linking together management for physical and virtual resources. Virtualization Manager also interoperates with VMware and Vmotion to create a high availability solution in the VMware environment.

- Manage availability

  IBM Systems Director Virtual Availability Management allows customers using Xen-based virtualization to create a high availability farm to help manage and respond to planned and unplanned outages as well as simplifying maintenance and migration tasks.

- Manage image

  IBM Systems Director Virtual Image Management provides a single, unified view of all system templates and server images to help customers manage and deploy their systems. It allows customers to easily deploy new physical and virtual servers based on system templates and images. The templates can be used to create, customize and clone virtual and physical images on x86, AMD and POWER-based systems.

- Know what IT costs

  IBM Usage and Accounting Manager (formerly *CIMS software* from CIMS Lab) allows alignment of IT costs with company priorities and the ability to account for individual departments use of key applications, servers, and other IT resources by providing an extremely flexible, end-to-end tool that helps you improve IT cost management. With it, you can truly understand your costs. In addition, you can track, allocate, and invoice based on actual resource use by department, user and many additional criteria.

- Deploy mainframe virtual systems

  The IBM Systems Director z/VM Center extension introduces a standardized way to deploy new z/VM virtual LINUX systems, but has the easy to use Director interface that does not require specific z/VM knowledge to operate. It uses the MAP interface to z/VM that offers a Common Information Model (CIM) based interface for z/VM system management functions. Its concept of virtual server templates and operating system templates allows one to repeat the creation of z/VM virtual guests and the deployment of Linux into these guests easily in a customized way.

More information about IBM Virtualization Manager

Features at a glance for IBM Virtualization Manager include:

- Visualize and navigate virtual to physical resources and relationships, with information in both tree and graphical topology views
- Review status of physical hosts and virtual machines
- See events that have occurred on the physical host or virtual machine
- For VMware, invoke VMotion directly from user interface or through automated policies
- Statically migrate virtual machines on VMware, Microsoft Virtual Server, and Xen (see product documentation for details)
- Capabilities can be used both with and independently of VMware VirtualCenter (VMware VMotion requires VirtualCenter)
- Start, shutdown, suspend, resume, and restart virtual machines
- Add and remove physical host to and from farms (for VMware)

- Invoke the VMware or Microsoft Virtual Server user interface
- Display of physical resource utilization
- Create and delete virtual machines

Benefits of IBM Virtualization Manager include:

- Simplified management
  - Single console manages both physical and virtual systems
  - Administer multiple virtualization technologies from a single console
  - Easy to install and use
  - Understand resource linkages using the built-in topology viewer
- Managed availability:
  - Proactive virtual machine migration based on predictive alerting helps avoid downtime
  - Drive VMware VMotion using physical hardware status information
  - Helps increase serviceability by migration of virtual machines to a standby server during a service window
- Tool integration:
  - Integration into IBM Systems Director provides a feature rich set of management tools
  - Allows you to take advantage of IBM Systems Director's existing facilities for alerting
  - Event action plans, security and system health
  - Consolidates management across IBM System x, BladeCenter, System p, and System i
- Cost:
  - Base Virtualization Manager is available at no additional cost to IBM customers

IBM Systems Director for Optimization

IBM System Director can be extended to provide the following optimization functions:

- Monitor and manage your energy utilization

  Active Energy Manager Version 3.1 is the next generation product to IBM PowerExecutive™ which was previously available from IBM for x86 systems only. It now supports multiple IBM platforms and provides new capabilities that build upon the functions previously available with IBM PowerExecutive V2.10.

  Active Energy Manager will measure, monitor, and manage the energy components built into IBM systems enabling a cross-platform management solution. It currently supports IBM BladeCenter, POWER, and System x servers. It also provides a source of energy management data that can be exploited by Tivoli enterprise solutions such as IBM Tivoli Monitoring and IBM Tivoli Usage and Accounting Manager.

  It is a key component of the IBM Cool Blue™ portfolio within Project Big Green. This solution helps customers monitor energy consumption to allow better utilization of available energy resources. The application software enables customers to trend actual energy consumption and corresponding thermal loading of IBM Systems running in their environment with their applications.

- Manage your capacity

  Capacity Manager tracks resource use, identifies multiple levels of existing or potential bottlenecks, and makes recommendations to improve performance. By predicting future server bottlenecks and proactively alerting IBM Systems Director, capacity manager performs automated corrective actions to minimize downtime. Capacity manager generates capacity and performance reports in XML format.

- Deploy systems remotely

  Remote Deployment Manager enables remote, unattended installation of new and existing systems. It helps automate deployment tasks, such as initial operating system installation, BIOS updates, and disposal of retired systems. You can do all these tasks without visiting the remote system, thereby reducing travel and labor costs.

- Build, manage and expand cluster environments efficiently:

  IBM offers a complete portfolio of cluster software for IBM Cluster System Management to help organizations using IBM System p servers running AIX or Linux, IBM System x servers running Linux or a combination. Cluster-ready software from IBM enables collections of IBM servers to behave like a single high-performance system for end users and system administrators.

IBM unveils plan to combat data center energy crisis; Allocates $1 Billion to advance *green* technology and services

IBM is going *green* and capabilities such as *Active Energy Manager* among many others support that goal. IBM plans to double computing capacity in its own data centers and will use new technologies to avoid five billion kilowatt hours of new energy use.

In May 2007, IBM announced it would be devoting $1 billion per year across its businesses, mobilizing IBM to radically increase the level of energy efficiency in IT. The initiative outlines a plan including new products and services to sharply reduce data center energy consumption and its impact on the environment for IBM and its clients that run the world's business and public technology infrastructures.

Called *Project Big Green*, the initiative from IBM is targeted at combating the energy crisis in data centers where energy demand and energy costs are dramatically increasing. The initiative includes new technologies and services designed to deliver energy savings which will be deployed by a new global *green team* of more than 850 energy efficiency specialists from across IBM. The savings are substantial—for an average 25 000 square foot data center, clients should be able to achieve 42% energy savings. This savings equates to 7439 tons of carbon emissions saved per year.

Energy efficiency in the data center has quickly become a critical issue as businesses rely increasingly more on IT innovation. Today, according to IDC, roughly 50 cents is spent on energy for every dollar of computer hardware. This is expected to increase by 54% over the next coming years.[a]

a. Source: IDC, Worldwide Server Power and Cooling Expense 2006–2010 Forecast, Doc #203598, September 2006

9.3 IBM Tivoli Service Management

Systems management tools need to be simple and intuitive to use, and must be able to be deployed in a straightforward manner. An easy to understand, high-level view of virtual and physical servers and the workloads running across those servers and their relationships is the first step to begin managing the environment.

Being able to visually see and understand the physical and virtual topology of the servers, associated infrastructure and composite workloads is critical. The ability to monitor and drill down into the details for problem analysis and isolation is the next important aspect that the tools must provide. Problem analysis is aided by the comprehensive and cohesive view of the

operation, along with the ability to seamlessly move between various tools which launch in context with single sign-on cross tool authentication.

Discovery

The IBM Tivoli Application Dependency Discovery Manager (TADDM) product provides visibility of interdependencies between application, computer systems and networking devices, using agent-less and credential-free discovery and automated application maps.

TADDM provides automatic discovery of the cross-tier infrastructure and creates a top-down tier map of components. TADDM can discover hosts and operating systems such as:

- IBM System p including (AIX and Logical Partitions)
- HP-UX
- Linux
- Solaris
- Microsoft Windows Server
- VMware ESX Server (parent-child relationships between hosts and the VMs)
- Citrix
- IBM z/OS

TADDM helps the IT operations team understand what they have and how the business services relates to the physical and virtual infrastructure.

Integrating data

Running an efficient IT infrastructure requires that the management tools provide data to the IT operations staff visually, and in an easy to understand manner. The need to integrate data from multiple sources efficiently and quickly is important to keeping the IT staff effective. A comprehensive and cohesive view of the operational state of the heterogeneous, virtual and physical IT datacenter is critical for rapid problem analysis and isolation.

The IBM Tivoli Change and Configuration Management Database (CCMDB) is a data store that is used to track IT assets, their relationship, their configuration and changes in the IT infrastructure. The CCMDB can be populated by TADDM automatically. Using the automated discovery with the CCMDB enables an organization to maintain an accurate view of the infrastructure. TADDM's automated application map provides the visibility between physical host servers and the virtual machines that they having running on them.

The understanding of where the virtual machines are running and their dependencies on the physical hosts is important to understand when performing problem analysis activities. The visualization provided by TADDM facilitates the visual mapping of the business service topology, providing the IT data center management with an understanding of correlation between the business services and the IT infrastructure, including the virtual resources.

Virtual machine monitoring

IBM Tivoli Monitoring for Virtual Servers is another important tool used by the data center staff after they have consolidated servers and are using virtual machines in addition to their physical servers. Tivoli Monitoring for Virtual Servers is able to drill down, identify and resolve virtual server availability and performance issues.

Tivoli Monitoring provides superior visualization to help identify trends, see impacts, and take action, all from a centralized, easy to use portal. It also provides an automated solution based on best practices for monitoring Citrix Access Suite, VMware ESX Server, and Microsoft Virtual Server. Using ITM for Virtual Servers, physical and virtual servers can be monitored and measured for availability.

The introduction of virtual machines into the IT infrastructure is often accomplished by the conversion of existing physical servers to virtual servers. This physical-to-virtual conversion creates a virtual machine of the physical server. The converted virtual machines are then hosted together (consolidated) onto a new physical host that is running virtualization layer software, such as VMware ESX Server. However, there are times when the creation of a virtual server is needed to support a new business service, and does not currently exist.

Provisioning

IBM Tivoli Provisioning Manager (TPM) is adept at provisioning a new virtual server or if need be, the traditional physical server. TPM provides a single integrated solution and includes an extensive set of inventory, OS provisioning, software distribution and patch management capabilities on a service oriented architecture (SOA).

In combination with the TPM task automation engine, the administrator can fully automate custom procedures that might require additional configuration changes to network, storage or virtual server resources. TPM can automate key process steps such as release targeting, testing and deployment. Automating operational tasks helps optimize the efficiency and accuracy of deployments.

A new automation package for creating and managing virtual machines using VMware ESX Server 3.0 and VMware VirtualCenter 2.0 is provided with Tivoli Provisioning Manager version 5.1.0.2 (Fix Pack 2).

Workload scheduling and automation

IBM Tivoli Workload Automation, including IBM Tivoli Workload Scheduler, IBM Tivoli Dynamic Workload Broker, and IBM Enterprise Workload Manager™, provides visibility and control of composite workloads across mixed application and virtual and physical resource environments.

Through a single point of control, Tivoli Workload Automation can help automate key service execution steps such as planning and modeling event-driven and time-driven composite workloads across virtualized environments, dynamically brokering workloads to best available virtual resources while resolving cross-workload and physical resource dependencies, and dynamically adapting workload execution to incidences, problems and configuration changes in the underlying virtual and physical IT infrastructure.

The cost of managing the IT infrastructure is one of the largest components of overall IT spending for many organizations. Server virtualization addresses this cost through the consolidation of physical server resources. IT data center managers are seeking systems management tools for managing both their physical and virtual systems across multiple architectures and environments.

Servers that power your important applications often run at a fraction of actual capacity. Virtualizing those operations to power multiple applications on fewer servers can help minimize your hardware, related maintenance and labor costs. It also helps you create pools of shared resources that your entire organization can draw on, as needed. The under-utilized physical server is many times running a single purpose application, owned by a department, and supports a business service. The cost accounting is straight forward for this single server, single application. As the IT data center consolidates these low utilization servers the cost accounting becomes more complex. The cost of the physical host is now shared among several virtual machines and applications.

Cost management

The IBM Tivoli Usage and Accounting manager product is the tool needed to assist with this cost allocation problem. Tivoli Usage and Accounting manager helps solve this problem by

collecting, analyzing, reporting, and billing based on usage and costs of shared of these computing resources:

- Windows
- UNIX (AIX, HP/UX, Sun Solaris)
- Linux (Red Hat and Novell SUSE)
- i5/OS
- VMware

Tivoli Usage and Accounting manager helps you improve IT cost management. With it you can understand your costs and track, allocate, and invoice based on actual resource use by department, user, and many additional criteria. The solutions enables you to consolidate a wide variety of usage data with data collectors associated with operating systems, databases, internet infrastructure, e-mail systems, network and printing, and customized usage data import collection from any application or system.

Backup and recovery

The IT data center has many physical and virtual servers and associated information or data that must be protected against loss. The IBM Tivoli Storage Manager family of products is designed to help provide a comprehensive data protection solution focused on the key data protection activities of backup, archive, recovery, space management, and disaster recovery planning.

IBM Tivoli Storage Manager (TSM) minimizes the data needed for backup by only backing up new and changed files. To protect information in a VMware virtual server environment, Tivoli Storage Manager offers two approaches to backing up your data:

- TSM can be used with VMware's Consolidated Backup (VCB)
- The TSM backup/archive client can be installed on the guest operating system in the virtual machine.

TSM installed on a virtual server machine (guest machine) is used to provide file level restore and more granular restore for applications (for example, Microsoft Exchange and SQL Server®). TSM supports VMware ESX Server, AIX partitions, and Microsoft Virtual Server with IBM standard product running unchanged in guest systems or partitions.

9.4 Consolidated storage management with IBM software

IBM System Storage SAN Volume Controller (SVC) is an essential technology for the IT data centers that are virtualizing storage. SVC can help simplify the deployment and administration, while increasing utilization of the physical storage devices. SVC provides, a single point of control for heterogeneous storage resources, and pools the storage capacity of multiple storage systems on a SAN. SVC provides improved storage resource utilization and is able to present a homogenous storage view of heterogeneous storage.

SVC provides significant value to the IT Data center that is experiencing significant growth in online information storage needs. SVC creates a virtualized pool of the storage environment and can manage different types of storage, allowing the storage operations team to use SVC as the single interface, saving storage management time.

Storage discovery for the CCMDB is accomplished using IBM TotalStorage Productivity Center Limited Edition V3.1.2. The Productivity Center Limited Edition provides configuration information about storage arrays and SAN fabric switches for both IBM and heterogeneous storage. It also includes a health monitor and topology viewer of the SAN as well as basic

configuration management of storage systems. Integration between Productivity Center and CCMDB is through the discovery library capability.

9.4.1 IBM TotalStorage Productivity Center

IBM TotalStorage Productivity Center can manage all of your storage assets with a single, comprehensive management suite. TotalStorage Productivity Center brings together management of the SAN, storage devices, and host resources, including databases and file systems, into a single control point. It provides a modular, integrated set of products that can be purchased individually or in different combinations.

TotalStorage Productivity Center Standard Edition includes the following as a single package:

- IBM TotalStorage Productivity Center for Disk
- IBM TotalStorage Productivity Center for Data
- IBM TotalStorage Productivity Center for Fabric

The IBM TotalStorage Productivity Center suite of products includes the products listed in Table 9-1.

Table 9-1 IBM TotalStorage Productivity Center suite of products

Component	Use
IBM TotalStorage Productivity Center for Fabric	To manage, monitor and control your SAN fabric
IBM TotalStorage Productivity Center for Data	To manage the capacity utilization of storage systems, file systems and databases and to automate file-system capacity provisioning
IBM TotalStorage Productivity Center for Disk	To perform device configuration and management of multiple devices from a single user interface, tune and proactively manage the performance of storage devices on the Storage Area Network (SAN)
IBM TotalStorage Productivity Center for Replication	To control and monitor copy services operations (like FlashCopy, Metro Mirror and Global Mirror capabilities) and use data replication for data protection and disaster recovery
IBM TotalStorage Productivity Center Basic Edition	To perform basic device management services for IBM System Storage DS4000™, DS6000™ and DS8000 products, IBM SAN Volume Controller and heterogeneous storage environments

9.4.2 IBM System Storage Productivity Center

IBM System Storage Productivity Center helps provide extended device configuration capabilities for heterogeneous devices while also consolidating management to a centralized platform. The System Storage Productivity Center software is designed to allow you to manage a storage system in context of the broader storage environment.

The System Storage Productivity Center now integrates enterprise SAN management with the low-level device configuration functions that were previously only available in standalone consoles for DS8000 and SVC. With this integration, DS8000 and SVC administrators now have access to an expanded toolset.

Highlights:

- Designed as a new management console for IBM System Storage DS8000 and IBM System Storage SAN Volume Controller.
- Management from a single pane of glass with integrated DS8000 GUI.
- Through additional integration testing System Storage Productivity Center is intended to reduce complexity and improve interoperability.
- The storage topology viewer is a dynamic graphical map of the overall SAN, including device relationships, health and detailed configuration.
- The pre-installed software can be licensed to provide advanced capabilities, such as disk performance, collection storage resource management, storage provisioning and advanced copy services management.
- Intended to extend basic device configuration utilities.

The IBM System Storage Productivity Center includes the following pre-installed (separately purchased) software:

- IBM TotalStorage Productivity Center Basic Edition
- IBM System Storage SAN Volume Controller (CIM Agent and GUI)

9.5 Summary

Providing a consolidated view of the IT infrastructure and workloads being managed has significant value, as the number of machines, physical and virtual being managed expands dramatically. A new server can be provisioned in minutes using virtualization technologies, when it took many days, if not weeks to procure and provision a new physical server. The consolidated view of the IT infrastructure being managed also provides a consistent view and improves operational staff efficiencies. The staff can use this common interface that has the same look and feel to manage multiple resources.

Systems management, when done well, provides the data center with the ability to take full advantage of the capabilities offered from virtualization by hiding and making manageable the additional complexities that it brings along with the benefits. IBM systems management software is the able to handle today the complexities of multivendor, heterogeneous, mixed physical and virtual resource environments. Utilizing IBM systems management software allows the business focus on providing business services and not on the complexity associated with the IT infrastructure supporting the critical business services.

One of the key goals of systems management is to help manage the complexity and increase the efficiency and effectiveness of the IT staff regardless of the technologies being deployed in the data center. Systems management software from IBM can deliver valuable capabilities now and will continue to adapt and improve as the virtual IT infrastructure continues to evolve.

10

Summary

Many customers report that energy efficiency and IT simplification are very important to them. Companies of all sizes are aggressively adopting IBM virtualization solutions to help with infrastructure simplification, rapid application deployment, business resiliency, and management of a virtualized infrastructure. When it is time to act on an emerging opportunity or react to a competitive threat, the essence of an on demand business is getting the right information, to the right people, at the right time to create value or mitigate risk.

IBM virtualization enables you to:

- Address environmental concerns through more efficient shared infrastructure
- Simplify your existing infrastructure while managing future growth
- Improve the management of your infrastructure and your operators
- Improve flexibility and responsiveness to business demands
- Increase business resiliency
- Improve total cost of ownership (TCO)

IBM takes a holistic approach to virtualization by working throughout all resource types, taking advantage of decades of mainframe experience, embracing diversity of resources, and integrating the virtual and physical worlds.

Virtualization helps you take control of your infrastructure by enabling you to see and manage your computing resources in ways that offer more flexibility. You are not restricted by implementation, location, or physical packaging. With virtualization, you have a logical rather than physical view of data, computing power, storage capacity, and other resources. By gaining greater control of your infrastructure, you can improve cost management.

The goal is to be able carry out server consolidation, raise resource utilization rates, and leave the operations staff capable of dynamically provisioning capacity on demand. Companies are increasingly turning to virtualization as a way to simplify their infrastructure, improve their total cost of ownership, and increase their flexibility. Virtualization is available across the IBM hardware line, along with cross-platform tools to create a single pane of glass view into your infrastructure.

IBM is uniquely positioned to offer both a breadth and depth of virtualization offerings. No other vendor can bring together the virtualization solutions (infrastructure simplification, rapid application deployment, and business resiliency), server, storage and application virtualization, and cross platform systems management in a heterogeneous environment. IBM virtualization balances the needs for flexibility and isolation in a virtualized environment, while simplifying management so that customers are able to manage the ever-increasing server and storage requirements while maintaining or reducing current staff.

Related publications

We consider the publications that we list in this section particularly suitable for a more detailed discussion of the topics that we cover in this paper.

IBM Redbooks publications

For information about ordering these publications, see “How to get Redbooks” on page 82. Note that some of the documents referenced here might be available in softcopy only.

- *Integrated Virtualization Manager on IBM System p5*, REDP-4061
- *Introducing N_Port Identifier Virtualization for IBM System z9*, REDP-4125
- *IBM System p Advanced POWER Virtualization (PowerVM) Best Practices*, REDP-4194
- *Advanced POWER Virtualization on IBM System p Virtual I/O Server Deployment Examples*, REDP-4224
- *z/VM and Linux on IBM System z: The Virtualization Cookbook for SLES9*, SG24-6695
- *IBM Virtualization Engine Platform Version 2 Technical Presentation Guide*, SG24-7112
- *IBM Virtualization Engine TS7510: Tape Virtualization for Open Systems Servers*, SG24-7189
- *Virtualization on the IBM System x3950 Server*, SG24-7190
- *Planning, Installing, and Using the IBM Virtualization Engine Version 2.1*, SG24-7213
- *A Virtualization Experience: IBM Worldwide Grid Implementation*, SG24-7229
- *IBM z/VM and Linux on IBM System z: Virtualization Cookbook for Red Hat Enterprise Linux 4*, SG24-7272
- *IBM Virtualization Engine V2.1 for System z*, SG24-7276
- *IBM Virtualization Engine TS7700: Tape Virtualization for System z Servers*, SG24-7312
- *Virtualization and Clustering Best Practices Using IBM System p Servers*, SG24-7349
- *The IBM Virtualization Engine TS7510: Getting Started with i5/OS and Backup Recovery and Media Services*, SG24-7510
- *IBM TS7520 Virtualization Engine: Planning, Implementation, and Usage Guide*, SG24-7520
- *PowerVM Virtualization on IBM System p Managing and Monitoring*, SG24-7590
- *PowerVM Virtualization on IBM System p Introduction and Configuration Fourth Edition*, SG24-7940

Online resources

These Web sites are also relevant as further information sources:

- IBM System z: About Virtualization

  http://ibm.com/systems/z/advantages/virtualization
- AIX 6.1 Workload Partitions

  http://www.ibm.com/DeveloperWorks/aix/library/au-workload/
- Tips and Tricks for Implementing Infrastructure Services on ESX Server

  http://www.vmware.com/vmtn/resources/409
- Citrix Completes Acquisition of XenSource

  http://www.citrix.com/English/NE/news/news.asp?newsID=683171
- Xen

  http://xen.org
- Microsoft Virtual Server on Wikipedia

  http://en.wikipedia.org/wiki/Microsoft_Virtual_Server
- Virtual Server 2005 R2 SP1 Download Frequently Asked Questions

  http://www.microsoft.com/technet/virtualserver/software/faq.mspx
- IBM System x Benchmarks

  http://www.ibm.com/systems/x/resources/benchmarks/
- ServerProven: Support for VMware ESX Server

  http://www.ibm.com/servers/eserver/serverproven/compat/us/nos/vmwaree.html
- IBM Software success stories

  http://ibm.com/software/success
- Storage Performance Council benchmark results

  http://www.storageperformance.org/results/

How to get Redbooks

You can search for, view, or download Redbooks, Redpapers, Technotes, draft publications and Additional materials, as well as order hardcopy Redbooks, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

IBM Systems Virtualization: Servers, Storage, and Software

Covers the complete IBM Systems platform of servers, storage, and software

Describes the available virtualization products and technology

Introduces consolidation and virtualization techniques

Businesses are moving forward with server and storage virtualization projects and realizing the benefits. IBM offers the industry's broadest set of virtualization capabilities. The cross-platform virtualization, automation, and systems management solutions available from IBM enable customers to access and manage resources simply and dynamically for better asset utilization and reduced operating costs.

This paper serves as both an introduction to virtualization, as well as an overview of pertinent IBM hardware and software virtualization offerings. We first introduce the concepts of virtualization and the benefits of virtualizing your systems. We then describe virtualization options for each of the IBM Systems platforms as well as software and storage technologies that are used to implement virtualization.

This paper is suitable for people who want to expand their knowledge of virtualization and what IBM can offer with its systems and software.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

REDP-4396-00