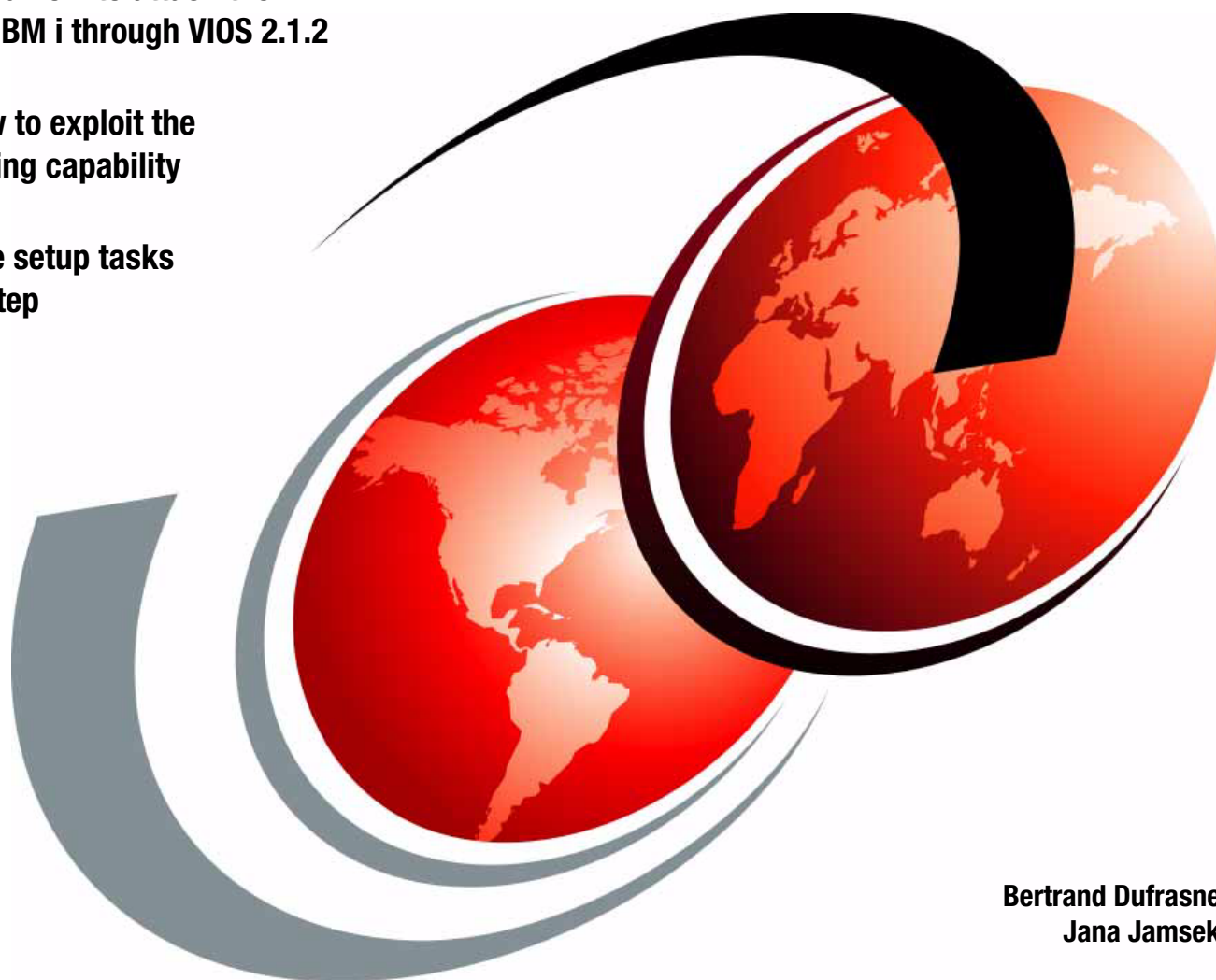


IBM XIV Storage System with the Virtual I/O Server and IBM i

Understand how to attach the XIV
server to IBM i through VIOS 2.1.2

Learn how to exploit the
multipathing capability

Follow the setup tasks
step-by-step



Bertrand Dufrasne
Jana Jamsek



International Technical Support Organization

**IBM XIV Storage System with the
Virtual I/O Server and IBM i**

February 2010

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

First Edition (February 2010)

This edition applies to Version 10, Release 2, of the IBM XIV Storage System software when attaching the XIV Storage System server to an IBM i client through a Virtual I/O Server 2.1.2 partition.

This document created or updated on February 23, 2010.

© Copyright International Business Machines Corporation 2010. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices	v
Trademarks	vi
Preface	vii
The team who wrote this paper	vii
Become a published author	viii
Comments welcome	viii
Chapter 1. Overview of the IBM XIV Storage System server	1
1.1 Key design features	2
1.2 Architecture overview	3
1.2.1 Hardware elements	3
1.2.2 Parallelism	5
1.2.3 Full storage virtualization	5
1.2.4 System usable capacity	10
1.2.5 Capacity allocation and thin provisioning	10
1.3 Reliability, availability, and serviceability	14
1.3.1 High availability	14
1.3.2 Consistent performance	14
1.3.3 Write path redundancy	15
1.3.4 System quiesce and graceful shutdown	16
1.3.5 Rebuild and redistribution	17
1.4 Physical architecture and components	18
1.4.1 IBM XIV Storage System models 2810-A14 and 2812-A14	18
1.4.2 IBM XIV Storage System hardware components	20
1.5 IBM XIV Storage Management software	23
1.5.1 The XIV Storage Management GUI	25
1.5.2 Using the XCLI	30
Chapter 2. Overview of the Virtual I/O Server with IBM i	33
2.1 Introduction to IBM PowerVM	34
2.1.1 IBM PowerVM overview	34
2.1.2 Virtual I/O Server	35
2.1.3 Node Port ID Virtualization	36
2.2 PowerVM client connectivity to the IBM XIV Storage System server	37
Chapter 3. Planning for the IBM XIV Storage System server with IBM i	39
3.1 Requirements	40
3.2 SAN connectivity	41
3.3 Planning for the Virtual I/O Server	42
3.3.1 Physical Fibre Channel adapters and virtual SCSI adapters	42
3.3.2 Queue depth in the IBM i operating system and Virtual I/O Server	42
3.3.3 Multipath with two Virtual I/O Servers	43
3.4 Best practices	43
3.4.1 Distributing connectivity	43
3.4.2 Zoning SAN switches	43
3.4.3 Queue depth	44
3.4.4 Number of application threads	44
3.5 Planning for capacity	44

3.5.1 Net usable capacity	45
3.5.2 Storage pool capacity	45
3.5.3 Capacity reduction for an IBM i client	46
Chapter 4. Implementing the IBM XIV Storage System server with IBM i	47
4.1 Connecting a PowerVM client to the IBM XIV Storage System server	48
4.1.1 Creating the Virtual I/O Server and IBM i partitions	48
4.1.2 Installing the Virtual I/O Server	51
4.1.3 IBM i multipath capability with two Virtual I/O Servers	54
4.1.4 Connecting with virtual SCSI adapters in multipath with two Virtual I/O Servers	54
4.2 Configuring XIV storage to connect to IBM i by using the Virtual I/O Server	56
4.2.1 Creating a storage pool	57
4.2.2 Defining the volumes	60
4.2.3 Connecting the volumes to the Virtual I/O Server	64
4.3 Mapping the volumes in the Virtual I/O Server	67
4.3.1 Using the HMC to map volumes to an IBM i client	69
4.4 Installing the IBM i client	72
Related publications	79
IBM Redbooks	79
Other publications	79
Online resources	79
How to get Redbooks	80
Help from IBM	80
Index	81

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

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>

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

AIX®
BladeCenter®
DS8000®
i5/OS®
IBM®
Micro-Partitioning™

Power Systems™
POWER6®
PowerVM™
POWER®
Redbooks®
Redpaper™

Redbooks (logo) ®
System i®
System Storage™
XIV®

The following terms are trademarks of other companies:

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

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

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

In this IBM® Redpaper™ publication, we discuss and explain how you can connect the IBM XIV® Storage System server to the IBM i operating system through the Virtual I/O Server (VIOS). A connection through the VIOS is especially interesting for IT centers that have many small IBM i partitions. When using the VIOS, the Fibre Channel host adapters can be installed in the VIOS and shared by many IBM i clients by using virtual connectivity to the VIOS.

The team who wrote this paper

This paper was produced by a team of specialists from around the world working at the International Technical Support Organization (ITSO) in San Jose, CA.

Bertrand Dufrasne is an IBM Certified Consulting I/T Specialist and Project Leader for IBM System Storage™ disk products at the ITSO in San Jose, CA. He has worked at IBM in various areas of IT. He has authored many IBM Redbooks® publications and has developed and taught technical workshops. Before joining the ITSO, he worked for IBM Global Services as an Application Architect. Bertrand holds a Master of Electrical Engineering degree from the Polytechnic Faculty of Mons, Belgium.

Jana Jamsek is an IT Specialist for IBM Slovenia. She works in Storage Advanced Technical Support for Europe as a specialist for IBM Storage Systems and the IBM i (i5/OS®) operating system. Jana has eight years of experience in working with the IBM System i® platform and its predecessor models, as well as eight years of experience in working with storage. She has a master degree in computer science and a degree in mathematics from the University of Ljubljana in Slovenia.

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

John Bynum
Robert Gagliardi
Gary Kruesel
Christina Lara
Lisa Martinez
Vess Natchev
Aviad Offer
Christopher Sansone
Wesley Varela
IBM U.S.

Ingo Dimmer
IBM Germany

Haim Helman
IBM Israel

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 publications 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



Overview of the IBM XIV Storage System server

The IBM XIV Storage System architecture is designed to deliver performance, scalability, and ease of management while harnessing the high capacity and cost benefits of Serial Advanced Technology Attachment (SATA) drives. The system employs off-the-shelf products as opposed to traditional offerings that use proprietary designs and, therefore, require more expensive components.

In this chapter, we provide an overview of the XIV concepts and architecture. We summarize, and in some cases repeat for convenience, the information that is in the IBM Redbooks publication *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

Tip: If you are already familiar with the XIV architecture, you can skip this chapter.

1.1 Key design features

In this section, we describe the key design features of the XIV architecture.

Massive parallelism

The system architecture ensures full exploitation of all system components. Any I/O activity involving a specific logical volume in the system is always inherently handled by all spindles. The system harnesses all storage capacity and all internal bandwidth. It also takes advantage of all available processing power, which is true for both host-initiated I/O activity and system-initiated activity, such as rebuild processes and snapshot generation. All disks, CPUs, switches, and other components of the system contribute to the performance of the system at all times.

Workload balancing

The workload is evenly distributed over all hardware components at all times. All disks and modules are used equally, regardless of access patterns. Despite the fact that applications might access certain volumes more frequently than other volumes, or access certain parts of a volume more frequently than other parts, the load on the physical disks and modules will be balanced perfectly.

Pseudo-random distribution ensures consistent load balancing after adding, deleting, or resizing volumes, and after adding or removing hardware. The balancing of all data on all system components eliminates the possibility of a hot spot being created.

Self-healing

Protection against a double disk failure is provided by an efficient rebuild process that brings the system back to full redundancy in minutes. In addition, the XIV system extends the self-healing concept, resuming redundancy even after failures in components other than disks.

True virtualization

Unlike other system architectures, storage virtualization is inherent to the basic principles of the XIV design. Physical drives and their locations are completely hidden from the user, which dramatically simplifies storage configuration so that the system can lay out the user's volume in the optimal way. The automatic layout maximizes the system's performance by using system resources for each volume, regardless of the user's access patterns.

Thin provisioning

The system supports *thin provisioning*, which is the capability to allocate actual storage to applications in a just-in-time and as needed basis. Thin provisioning allows the most efficient use of available space and, as a result, significant cost savings compared to traditional provisioning techniques. Thin provisioning is achieved by defining a logical capacity that is larger than the physical capacity and by using space based on what is consumed rather than what is allocated.

Processing power

The XIV open architecture uses the latest processor technologies and is more scalable than solutions that are based on a closed architecture. The XIV system avoids sacrificing the performance of one volume over another and, therefore, requires little to no tuning.

1.2 Architecture overview

The XIV architecture incorporates a variety of features that are designed to uniformly distribute data across internal resources.

1.2.1 Hardware elements

The primary components of the XIV system are known as *modules*. Modules provide processing, cache, and host interfaces and are based on “off-the-shelf” Intel®-based systems. They are redundantly connected to one another through an internal switched Ethernet network, as shown in Figure 1-1. All of the modules work together concurrently as elements of a grid architecture. Therefore, the system harnesses the powerful parallelism that is inherent in such a distributed computing environment. We discuss the grid architecture in 1.2.2, “Parallelism” on page 5.

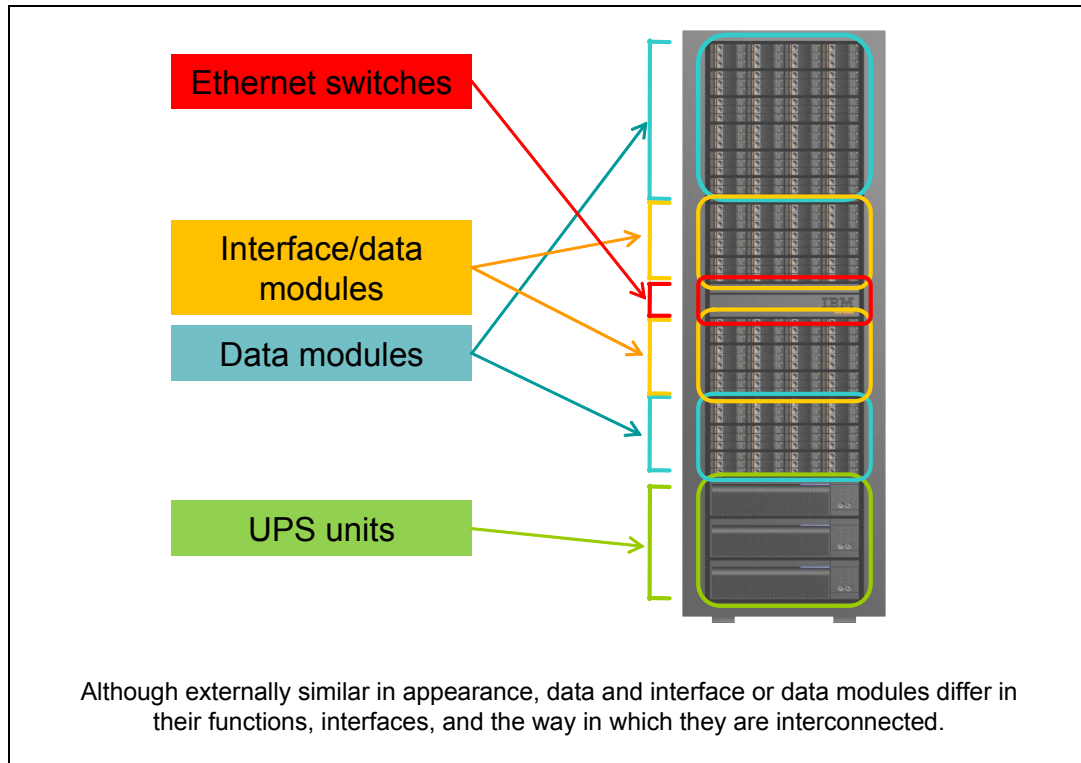


Figure 1-1 Major hardware elements of the XIV system

Data modules

At a conceptual level, the data modules function as the elementary “building blocks” of the system. They provide storage capacity, processing power, and caching, in addition to advanced system-managed services. The data module’s ability to share and manage system software and services are key elements of the physical architecture, as shown in Figure 1-2.

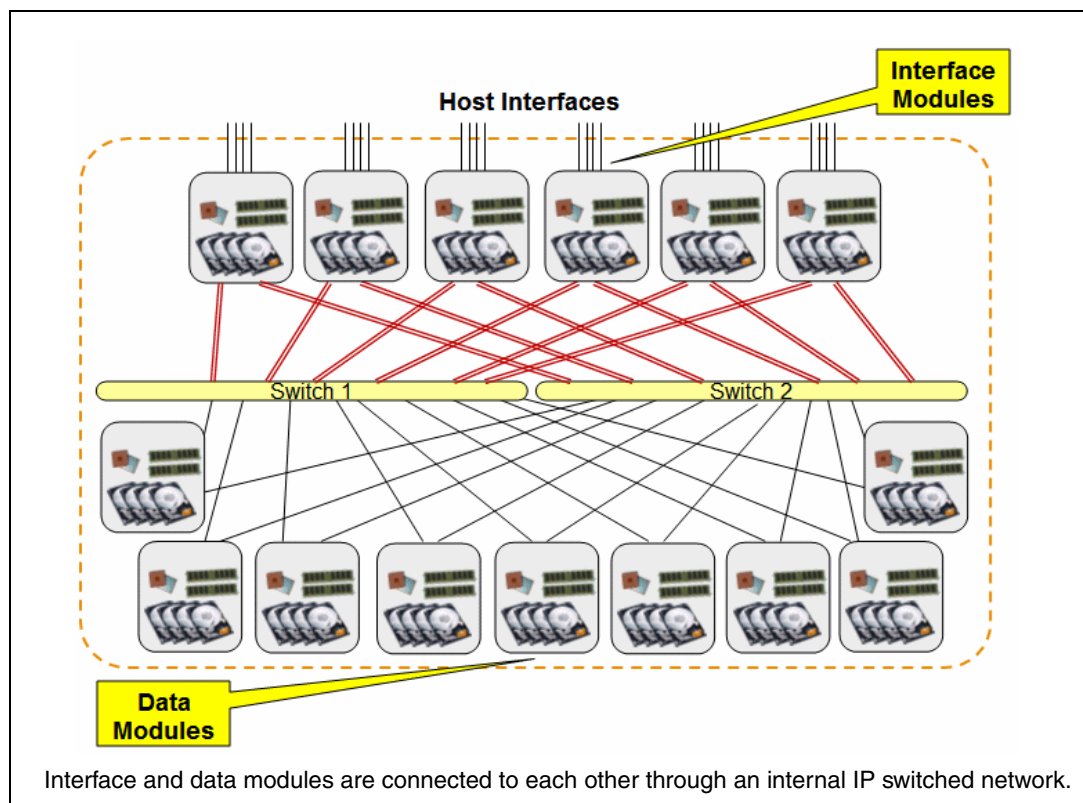


Figure 1-2 Architectural overview

Note: Figure 1-2 depicts the conceptual architecture only. Do not misinterpret the number of connections or other elements of this diagram as a precise hardware layout.

Interface modules

Interface modules are equivalent to data modules in all aspects, with the following exceptions:

- ▶ In addition to disk, cache, and processing resources, interface modules include both Fibre Channel and iSCSI interfaces for host system connectivity, remote mirroring, and data migration activities. Figure 1-2 conceptually illustrates the placement of interface modules within the topology of the XIV IBM Storage System architecture.
- ▶ The system services and software functionality associated with managing external I/O resides exclusively on the interface modules.

Ethernet switches

The XIV system contains a redundant switched Ethernet network that transmits both data and metadata traffic between the modules. Traffic can flow in any of the following ways:

- ▶ Between two interface modules
- ▶ Between two data modules
- ▶ Between an interface module and a data module

1.2.2 Parallelism

The concept of *parallelism* pervades all aspects of the XIV architecture by means of a balanced, redundant data distribution scheme in conjunction with a pool of distributed (or grid) computing resources. In addition to hardware parallelism, the XIV system also employs sophisticated algorithms to achieve optimal parallelism.

Data is distributed across all drives in a *pseudo-random* fashion. The patented algorithms provide a uniform yet random spreading of data across all available disks to maintain data resilience and redundancy. Figure 1-3 on page 6 provides a conceptual representation of the pseudo-random data distribution within the XIV system.

1.2.3 Full storage virtualization

The data distribution algorithms employed by the XIV system are innovative in that they are deeply integrated into the system architecture, instead of at the host or storage area network level. The XIV system is unique in that it is based on an innovative implementation of full storage virtualization within the system.

Traditional subsystems rely on storage administrators to carefully plan the relationship between logical structures, such as arrays and volumes, and physical resources, such as disk packs and drives, to strategically balance workloads, meet capacity demands, eliminate hot spots, and provide adequate performance. The implementation of full storage virtualization employed by the XIV system eliminates many of the potential operational drawbacks that can be present with conventional storage subsystems, while maximizing the overall usefulness of the subsystem.

In this section, we elaborate on the logical system concepts, which form the basis for the system full storage virtualization.

Logical constructs

The logical architecture of the XIV system incorporates constructs that underlie the storage virtualization and distribution of data, which are integral to its design. The logical structure of the system ensures that there is optimum granularity in the mapping of logical elements to both modules and individual physical disks, thereby guaranteeing an equal distribution of data across all physical resources.

Partitions

The fundamental building block of logical volumes is called a *partition*. Partitions have the following characteristics on the XIV system:

- ▶ All partitions are 1 MB (1024 KB) in size.
- ▶ A partition contains either a primary copy or secondary copy of data:
 - Each partition is mapped to a single physical disk:
 - This mapping is dynamically managed by the system through innovative data distribution algorithms to preserve data redundancy and equilibrium. For more information about the topic of data distribution, see “Logical volume layout on physical disks” on page 9.
 - The storage administrator has no control or knowledge of the specific mapping of partitions to drives.
 - Secondary copy partitions are always placed in a different *module* than the one that contains the primary copy partition.

Important: In the context of the logical architecture of the XIV system, a partition consists of 1 MB (1024 KB) of data. Do not confuse this definition with other definitions of the term *partition*.

Figure 1-3 illustrates that data is uniformly, yet randomly distributed over all disks. Each 1 MB of data is duplicated in a primary and secondary partition. For the same data, the system ensures that the primary partition and its corresponding secondary partition are not in the same module.

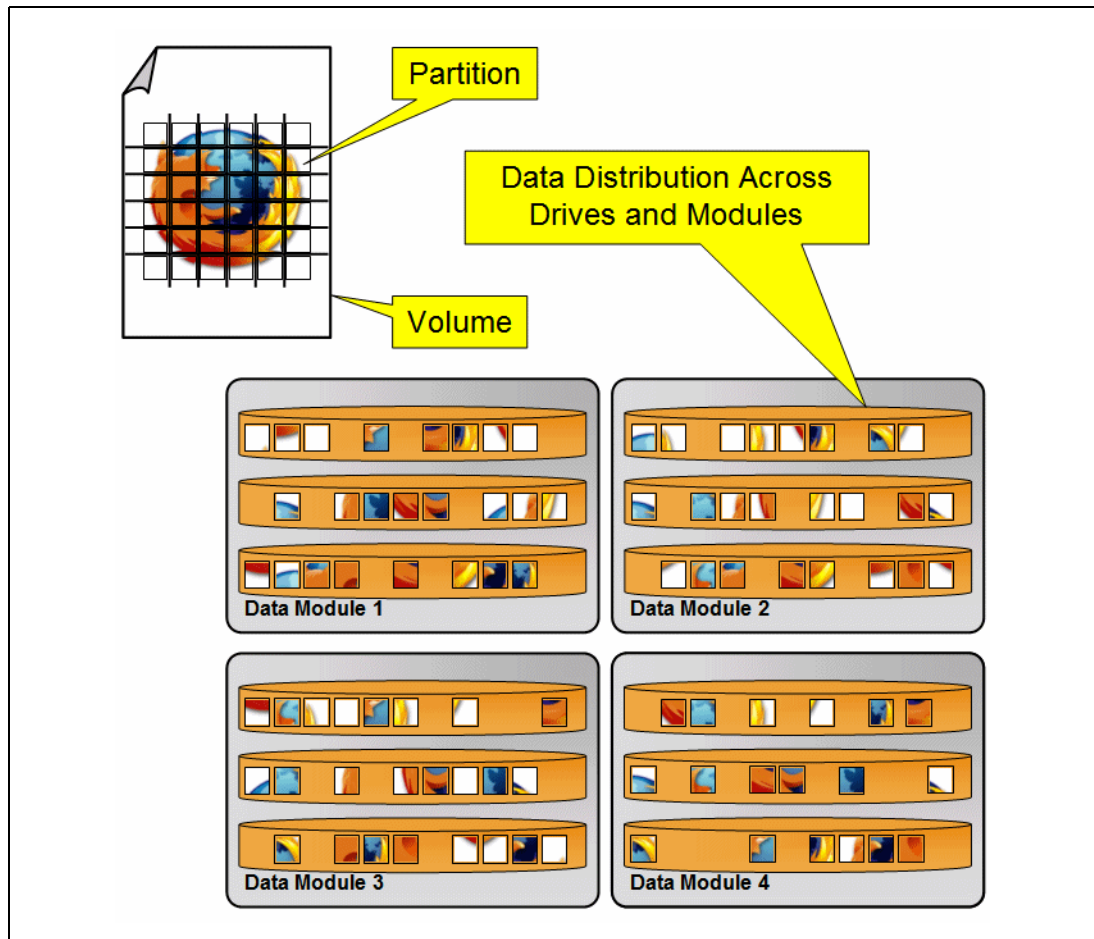


Figure 1-3 Pseudo-random data distribution¹

Logical volumes

The XIV system presents logical volumes to hosts in the same manner as conventional subsystems. However, both the granularity of logical volumes and the mapping of logical volumes to physical disks differ in the following ways:

- ▶ Every logical volume is comprised of 1 MB (1024 KB) constructs of data called *partitions*.
- ▶ The physical capacity associated with a logical volume is *always* a multiple of 17 GB (decimal).

¹ Copyright 2005-2008 Mozilla. All Rights Reserved. All rights in the names, trademarks, and logos of the Mozilla Foundation, including without limitation, Mozilla, Firefox, as well as the Firefox logo, are owned exclusively by the Mozilla Foundation.

Therefore, although it is possible to present a block-designated logical volume to a host that is *not* a multiple of 17 GB, the *actual physical space* that is allocated for the volume is always the sum of the minimum number of 17 GB increments that are needed to meet the block-designated capacity.

Note: The initial physical capacity that is allocated by the system upon volume creation can be less than this amount, as discussed in “Logical and actual volume sizes” on page 10.

- ▶ The maximum number of volumes that can be concurrently defined on the system is limited by the following factors:
 - The logical address space limit
 - The logical address range of the system permits up to 16 377 volumes, although this constraint is purely logical, and therefore, is not normally a practical consideration.
 - The same address space is used for both volumes and snapshots.
 - The limit imposed by the logical and physical topology of the system for the minimum volume size

The physical capacity of the system, based on 180 drives with 1 TB of capacity per drive and assuming the minimum volume size of 17 GB, limits the maximum volume count to 4 605 volumes.

Important: The logical address limit is ordinarily not a practical consideration during planning, because under most conditions, this limit is not reached. It is intended to exceed the adequate number of volumes for all conceivable circumstances.

Storage pools

Storage pools are administrative boundaries that enable storage administrators to manage relationships between volumes and snapshots and to define separate capacity provisioning and snapshot requirements for such uses as separate applications or departments. Storage pools are not tied in any way to physical resources, nor are they part of the data distribution scheme.

A *logical volume* is defined within the context of one storage pool. Because storage pools are logical constructs, a volume and any snapshots associated with it can be moved to any other storage pool, as long as there is sufficient space.

As a benefit of the system virtualization, there are no limitations on the size of storage pools or on the associations between logical volumes and storage pools. In fact, manipulation of storage pools consists exclusively of metadata transactions and does not trigger any copying of data. Therefore, changes are completed instantly and without any system overhead or performance degradation.

Storage pools have the following characteristics:

- ▶ The size of a storage pool can range from 17 GB (the minimum size that can be assigned to a logical volume) to the capacity of the entire system.
- ▶ *Snapshot reserve capacity* is defined within each storage pool and is effectively *maintained separately* from logical, or master, volume capacity. The same principles apply for thinly provisioned storage pools, which are discussed in 1.2.5, “Capacity allocation and thin provisioning” on page 10, with the exception that space is not *guaranteed* to be available for snapshots because of the potential for hard space depletion.

- Snapshots are structured in the same manner as logical, or master, volumes.

Snapshot reserve: The snapshot reserve must be a minimum of 34 GB. The system preemptively deletes snapshots if the snapshots fully consume the allocated available space.

- Snapshots are only deleted automatically when inadequate physical capacity is available within the context of each storage pool. This process is managed by a snapshot deletion priority scheme. Therefore, when the capacity of a storage pool is exhausted, only the snapshots that reside in the affected storage pool are deleted in order of the deletion priority.
- ▶ The space allocated for a storage pool can be dynamically changed by the storage administrator:
 - The storage pool can be increased in size. It is limited only by the unallocated space on the system.
 - The storage pool can be decreased in size. It is limited only by the space that is consumed by the volumes and snapshots that are defined within that storage pool.
- ▶ The designation of a storage pool as a regular pool or a thinly provisioned pool can be dynamically changed even for existing storage pools. For an in-depth discussion of thin provisioning, see 1.2.5, “Capacity allocation and thin provisioning” on page 10.
- ▶ The storage administrator can relocate logical volumes between storage pools without any limitations, provided that sufficient free space is available in the target storage pool.

Notes:

- ▶ When moving a volume into a storage pool, the size of the storage pool is not automatically increased by the size of the volume. Likewise, when removing a volume from a storage pool, the size of the storage pool does not decrease by the size of the volume.
- ▶ The system defines capacity by using decimal metrics. With decimal metrics, 1 GB is 1 000 000 000 bytes. With binary metrics, 1 GB is 1 073 741 824 bytes.

Snapshots

A *snapshot* represents a point-in-time copy of a volume. Snapshots are similar to volumes, except that snapshots incorporate dependent relationships with their source volumes, which can be either logical volumes or other snapshots. Because they are not independent entities, a given snapshot does not necessarily wholly consist of partitions that are unique to that snapshot. Conversely, a snapshot image does not share all of its partitions with its source volume if updates to the source occur after the snapshot was created.

- ▶ *Snapshot reserve capacity* is defined within each storage pool and is effectively *maintained separately* from logical, or master, volume capacity.

Snapshot reserve: The snapshot reserve must be a minimum of 34 GB. The system preemptively deletes snapshots if the snapshots fully consume the allocated available space.

Snapshots are automatically deleted only when inadequate physical capacity is available within the context of each storage pool. This process is managed by a snapshot deletion priority scheme. Therefore, when the capacity of a storage pool is exhausted, only the snapshots that reside in the affected storage pool are deleted in order of the deletion priority.

- A logical volume can have multiple independent snapshots. This logical volume is also known as a *master volume*.

Snapshot consistency groups

A *consistency group* is a group of volumes of which a snapshot can be made at the same point in time, thus ensuring a consistent image of all volumes within the group at that time. The consistency between the volumes in the group is paramount to maintaining data integrity from the application perspective. By first grouping the application volumes into a consistency group, it is possible to later capture a consistent state of all volumes within that group at a given point-in-time by using a special snapshot command for consistency groups.

Issuing a snapshot command results in the following process:

1. Complete and destage writes across the constituent volumes.
2. Instantaneously suspend I/O activity simultaneously across all volumes in the consistency group.
3. Create the snapshots.
4. Resume normal I/O activity across all volumes.

The XIV system manages these suspend and resume activities for all volumes within the consistency group.

Logical volume layout on physical disks

The XIV system manages the distribution of logical volumes over physical disks and modules by means of a dynamic relationship between primary data partitions, secondary data partitions, and physical disks. This virtualization of resources in the XIV system is governed by the data distribution algorithms.

The distribution algorithms preserve the equality of access among all physical disks under all conceivable conditions and volume access patterns. Essentially, although not truly random in nature, the distribution algorithms in combination with the system architecture preclude the occurrence of hot spots:

- A fully configured XIV system contains 180 disks, and each volume is allocated across at least 17 GB (decimal) of capacity that is distributed evenly across all disks.
- Each logically adjacent partition on a volume is distributed across a different disk. Partitions are *not* combined into groups before they are spread across the disks.
- The pseudo-random distribution ensures that logically adjacent partitions are never striped sequentially across physically adjacent disks. Each disk has its data mirrored across all other disks, *excluding the disks in the same module*.
- Each disk holds approximately one percent of any other disk in other modules.
- Disks have an equal probability of being accessed regardless of aggregate workload access patterns.
- When the system is scaled out through the addition of modules, a new *goal distribution* is created whereby just a minimum number of partitions is moved to the newly allocated capacity to arrive at the new distribution table.

The new capacity is fully utilized within several hours and with no need for any administrative intervention. Thus, the system automatically returns to a state of equilibrium among all resources.

- Upon the failure or phase-out of a drive or a module, a new goal distribution is created whereby data in non-redundant partitions is copied and redistributed across the remaining modules and drives.

The system rapidly returns to a state in which all partitions are again redundant, because all disks and modules participate in achieving the new goal distribution.

1.2.4 System usable capacity

The IBM XIV Storage System server reserves physical disk capacity for the following instances:

- ▶ Global spare capacity
- ▶ Metadata, including statistics and traces
- ▶ Mirrored copies of data

Global spare capacity

The dynamically balanced distribution of data across all physical resources by definition obviates the inclusion of dedicated spare drives that are necessary with conventional RAID technologies. Instead, the XIV system reserves capacity on each disk to provide adequate space for the redistribution or rebuilding of redundant data in the event of a hardware failure.

The global reserved space includes sufficient capacity to withstand the failure of a full module and an additional three disks, and still allows the system to execute a new *goal distribution* and return to full redundancy.

Important: The system will tolerate multiple hardware failures, including up to an entire module in addition to three subsequent drive failures *outside of the failed module*, provided that a new goal distribution is fully executed before a subsequent failure occurs. If the system is less than 100% full, it can sustain more subsequent failures based on the amount of unused disk space that will be allocated at the event of failure as spare capacity.

Note: The XIV system does not manage a global reserved space for snapshots.

Metadata and system reserve

The system reserves roughly 4% of the physical capacity for statistics, traces, and the distribution table.

Net usable capacity

The calculation of the *net usable capacity* of the system consists of the total disk count, minus the disk space reserved for sparing (which is the equivalent of one module plus three more disks), multiplied by the amount of capacity on each disk that is dedicated to data (that is 96% because of metadata and system reserve), and finally reduced by a factor of 50% to account for data mirroring achieved by using the secondary copy of data.

1.2.5 Capacity allocation and thin provisioning

Thin provisioning is a central theme of the virtualized design of the IBM XIV Storage System server, because it uncouples the virtual, or apparent, allocation of a resource from the underlying hardware allocation.

Logical and actual volume sizes

The physical capacity that is assigned to traditional volumes is equivalent to the logical capacity presented to hosts, which does not have to be the case with the XIV system. For a given logical volume, there are effectively two associated sizes. The physical capacity allocated for the volume is not static, but it increases as host writes fill the volume.

Logical volume size

The *logical volume size* is the size of the logical volume that is observed by the host, as defined upon volume creation or as a result of a resizing command. The storage administrator specifies the volume size in the same manner regardless of whether the storage pool will be a thin pool or a regular pool. The volume size is specified in one of two ways, depending on units:

- ▶ In terms of GB, the system allocates the soft volume size as the minimum number of discrete 17 GB increments that are needed to meet the requested volume size.
- ▶ In terms of blocks, the capacity is indicated as a discrete number of 512 byte blocks. The system still allocates the soft volume size consumed within the storage pool as the minimum number of discrete 17 GB increments that are needed to meet the requested size (specified in 512 byte blocks). However, the size that is reported to hosts is equivalent to the precise number of blocks that are defined.

Incidentally, the snapshot reserve capacity associated with each storage pool is a soft capacity limit. It is specified by the storage administrator, although it effectively limits the hard capacity that is consumed collectively by snapshots.

Tip: Defining logical volumes in terms of blocks is useful when you must precisely match the size of an existing logical volume that resides on another system.

Actual volume size

The actual volume size reflects the total size of volume areas that were written by hosts. The actual volume size is not controlled directly by the user and depends only on the application behavior. It starts from zero at volume creation or formatting and can reach the logical volume size when the entire volume has been written. Resizing of the volume affects the logical volume size, but does not affect the actual volume size.

The actual volume size reflects the physical space that is used in the volume as a result of host writes. It is discretely and dynamically provisioned by the system, not the storage administrator.

The discrete additions to the actual volume size can be measured in two different ways, by considering either the allocated space or the consumed space. The *allocated space* reflects the physical space used by the volume in 17 GB increments. The *consumed space* reflects the physical space used by the volume in 1 MB partitions. In both cases, the upper limit of this provisioning is determined by the logical size that is assigned to the volume.

The following elements also come into account for the actual volume size:

- ▶ Capacity is allocated to volumes by the system in increments of 17 GB because of the underlying logical and physical architecture. There is no smaller degree of granularity than 17 GB.
- ▶ Application write access patterns determine the rate at which the allocated hard volume capacity is consumed and subsequently the rate at which the system allocates additional increments of 17 GB up to the limit defined by the logical volume size. As a result, the storage administrator has no direct control over the actual capacity allocated to the volume by the system at any given point in time.
- ▶ During volume creation, or when a volume has been formatted, *zero physical capacity is assigned to the volume*. As application writes accumulate in new areas of the volume, the physical capacity that is allocated to the volume grows in increments of 17 GB and can ultimately reach the full logical volume size.
- ▶ Increasing the logical volume size *does not* affect the actual volume size.

Thinly provisioned storage pools

While volumes are effectively thinly provisioned automatically by the system, storage pools can be defined by the storage administrator (when using the GUI) as either *regular* or *thinly provisioned*. When using the Extended Command Line Interface (XCLI), no specific parameter exists to indicate thin provisioning for a storage pool. You indirectly and implicitly create a storage pool as thinly provisioned by specifying a pool soft size that is greater than its hard size.

With a regular pool, the “host-apparent” capacity is *guaranteed* to be equal to the physical capacity that is reserved for the pool. The total physical capacity that is *allocated to* the constituent individual volumes and collective snapshots at any given time within a regular pool will reflect the current usage by hosts, because the capacity is dynamically consumed as required. However, the remaining unallocated space within the pool remains reserved for the pool and cannot be used by other storage pools.

In contrast, a thinly provisioned storage pool is not fully backed by hard capacity, meaning that the entirety of the logical space within the pool cannot be physically provisioned unless the pool is transformed first into a regular pool. However, benefits can be realized when physical space consumption is less than the logical space that is assigned, because the amount of logical capacity assigned to the pool that is not covered by physical capacity is available for use by other storage pools.

When a storage pool is created by using thin provisioning, that pool is defined in terms of both a soft size and a hard size independently, as opposed to a regular storage pool in which these sizes are by definition equivalent. A hard pool size and soft pool size are defined and used as explained in the following sections.

Hard pool size

The hard pool size is the maximum actual capacity that can be used by all the volumes and snapshots in the pool.

Thin provisioning of the storage pool maximizes capacity utilization in the context of a group of volumes, wherein the aggregate “host-apparent,” or soft, capacity assigned to all volumes surpasses the underlying physical, or hard, capacity that is allocated to them. This utilization requires that the aggregate space available to be allocated to hosts within a thinly provisioned storage pool must be defined independently of the physical, or hard, space allocated within the system for that pool. Thus, the storage pool hard size that is defined by the storage administrator limits the physical capacity that is available collectively to volumes and snapshots within a thinly provisioned storage pool. The aggregate space that is assignable to host operating systems is specified by the storage pool’s soft size.

Regular storage pools effectively segregate the hard space that is reserved for volumes from the hard space that is consumed by snapshots by limiting the soft space that is allocated to volumes. However, thinly provisioned storage pools permit the totality of the hard space to be consumed by volumes with no guarantee of preserving any hard space for snapshots. Logical volumes take precedence over snapshots and might be allowed to overwrite snapshots if necessary as hard space is consumed. The hard space that is allocated to the storage pool that is unused (or the incremental difference between the aggregate logical and actual *volume* sizes) can, however, be used by snapshots in the same storage pool.

Careful management is critical to prevent hard space for both logical volumes and snapshots from being exhausted. Ideally, hard capacity utilization must be maintained under a certain threshold by increasing the pool hard size as needed in advance.

Notes:

- ▶ Storage pools control when and which snapshots are deleted when insufficient space is assigned within the pool for snapshots.
- ▶ The soft snapshot reserve capacity and the hard space allocated to the storage pool are consumed only as changes occur to the master volumes or the snapshots themselves, not as snapshots are created.

Soft pool size

The soft pool size is the maximum logical capacity that can be assigned to all the volumes and snapshots in the pool.

Thin provisioning is managed for each storage pool independently of all other storage pools:

- ▶ Regardless of any unused capacity that might reside in other storage pools, snapshots within a given storage pool are deleted by the system according to corresponding snapshot preset priority if the hard pool size contains insufficient space to create an additional volume or increase the size of an existing volume. (Snapshots are only deleted when a write occurs under those conditions, and not when allocating more space.)
- ▶ As discussed previously, the storage administrator defines both the soft size and the hard size of thinly provisioned storage pools and allocates resources to volumes within a given storage pool without any limitations imposed by other storage pools.

The designation of a storage pool as a regular pool or a thinly provisioned pool can be dynamically changed by the storage administrator:

- ▶ When a regular pool must be converted to a thinly provisioned pool, the soft pool size parameter must be explicitly set in addition to the hard pool size, which remains unchanged unless updated.
- ▶ When a thinly provisioned pool must be converted to a regular pool, the soft pool size is automatically reduced to match the current hard pool size. If the combined allocation of soft capacity for existing volumes in the pool exceeds the pool hard size, the storage pool cannot be converted. Of course, this situation can be resolved if individual volumes are selectively resized or deleted to reduce the soft space consumed.

System-level thin provisioning

The definitions of hard size and soft size naturally apply at the subsystem level as well. By extension, it is necessary to permit the full system to be defined in terms of thin provisioning to achieve the full potential benefit that was previously described, which is namely, the ability to defer deployment of additional capacity on an as-needed basis.

With the architecture of the XIV system, the global system capacity can be defined in terms of both a hard system size and a soft system size. When thin provisioning is not activated at the system level, these two sizes are equal to the system's physical capacity.

With thin provisioning, these concepts have the meanings as discussed in the following sections.

Hard system size

The hard system size represents the physical disk capacity that is available within the XIV system. Obviously, the system's hard capacity is the upper limit of the aggregate hard capacity of all the volumes and snapshots. It can only be increased by installing new hardware components in the form of individual modules (and associated disks) or groups of modules. Certain conditions can *temporarily* reduce the system's hard limit.

Soft system size

The soft system size is the total, “global,” logical space that is available for all storage pools in the system. When the soft system size exceeds the hard system size, it is possible to logically provision more space than is physically available, thereby allowing the aggregate benefits of thin provisioning of storage pools and volumes to be realized at the system level.

The soft system size obviously limits the soft size of all volumes in the system and has the following attributes:

- The soft system size is not related to any direct system attribute and can be defined to be larger than the hard system size if thin provisioning is implemented. Keep in mind that the storage administrator *cannot* set the soft system size.

Note: If the storage pools within the system are thinly provisioned, but the soft system size does not exceed the hard system size, the total system hard capacity cannot be filled until all storage pools are regularly provisioned. Therefore, define all storage pools in a non-thinly provisioned system as regular storage pools.

- The soft system size is a purely logical limit. However, you must exercise care when the soft system size is set to a value greater than the maximum potential hard system size. Obviously, it must be possible to upgrade the system’s hard size to be equal to the soft size. Therefore, defining an unreasonably high system soft size can result in full capacity depletion. For this reason, defining the soft system size is not within the scope of the storage administrator role.

Certain conditions can *temporarily* reduce the system’s soft limit.

1.3 Reliability, availability, and serviceability

The unique modular design and logical topology of the IBM XIV Storage System server fundamentally differentiate it from traditional monolithic systems. This architectural divergence extends to the exceptional reliability, availability, and serviceability aspects of the system. In addition, the XIV system incorporates autonomic, proactive monitoring and self-healing features that are capable of transparently and automatically restoring the system to full redundancy within minutes of a hardware failure and taking preventive measures to preserve data redundancy before a component malfunction actually occurs.

1.3.1 High availability

The rapid restoration of redundant data across all available drives and modules in the system during hardware failures, and the equilibrium resulting from the automatic redistribution of data across all newly installed hardware, are fundamental characteristics of the IBM XIV Storage System architecture. These characteristics minimize exposure to cascading failures and the associated loss of access to data.

1.3.2 Consistent performance

The IBM XIV Storage System server is capable of adapting to the loss of an individual drive or module efficiently and with relatively minor impact compared to monolithic architectures. While traditional monolithic systems employ an N+1 hardware redundancy scheme, the XIV system harnesses the resiliency of the grid topology, in terms of the ability to sustain a component failure and by maximizing consistency and transparency from the perspective of

attached hosts. The potential impact of a component failure is vastly reduced, because each module in the system is responsible for a relatively small percentage of the system's operation. Simply put, a controller failure in a typical N+1 system likely results in a dramatic (up to 50%) reduction of available cache, processing power, and internal bandwidth, where the loss of a module in the XIV system translates to only one-fifteenth of the system resources and does not compromise performance nearly as much as the same failure with a typical architecture.

Additionally, the XIV system incorporates innovative provisions to mitigate isolated disk-level performance anomalies through *redundancy-supported reaction*. For more information about redundancy-supported reaction, see the Redbooks publication *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

1.3.3 Write path redundancy

Data that arrives from the hosts is temporarily placed in two separate caches before it is permanently written to disk drives located in separate modules. This design guarantees that the data is always protected against possible failure of individual modules, even before the data has been written to the disk drives.

Figure 1-4 on page 16 illustrates the path that is taken by a write request as it travels through the system. The diagram is intended to be viewed as a conceptual topology. Therefore, do not interpret the specific numbers of connections and so forth as literal depictions. Also, for purposes of this discussion, the interface modules are depicted on a separate level from the data modules. However, in reality the interface modules also function as data modules. The following numbers correspond to the numbers in Figure 1-4:

1. A host sends a write request to the system. Any of the interface modules that are connected to the host can service the request, because the modules work in an active-active capacity. The IBM XIV Storage System server does not perform load balancing of the requests. Load balancing must be implemented by storage administrators to equally distribute the host requests among all interface modules.
2. The interface module uses the system configuration information to determine the location of the primary module that houses the referenced data, which can be either an interface module, including the interface module that received the write request, or a data module. The data is written only to the local cache of the primary module.
3. The primary module uses the system configuration information to determine the location of the secondary module that houses the copy of the referenced data. Again, this module can be either an interface module or a data module, but it cannot be the same as the primary module. The data is redundantly written to the local cache of the secondary module.

After the data is written to cache in both the primary and secondary modules, the host receives an acknowledgement that the I/O is complete, which occurs independently of copies of either cached or dirty data that is being destaged to physical disk.

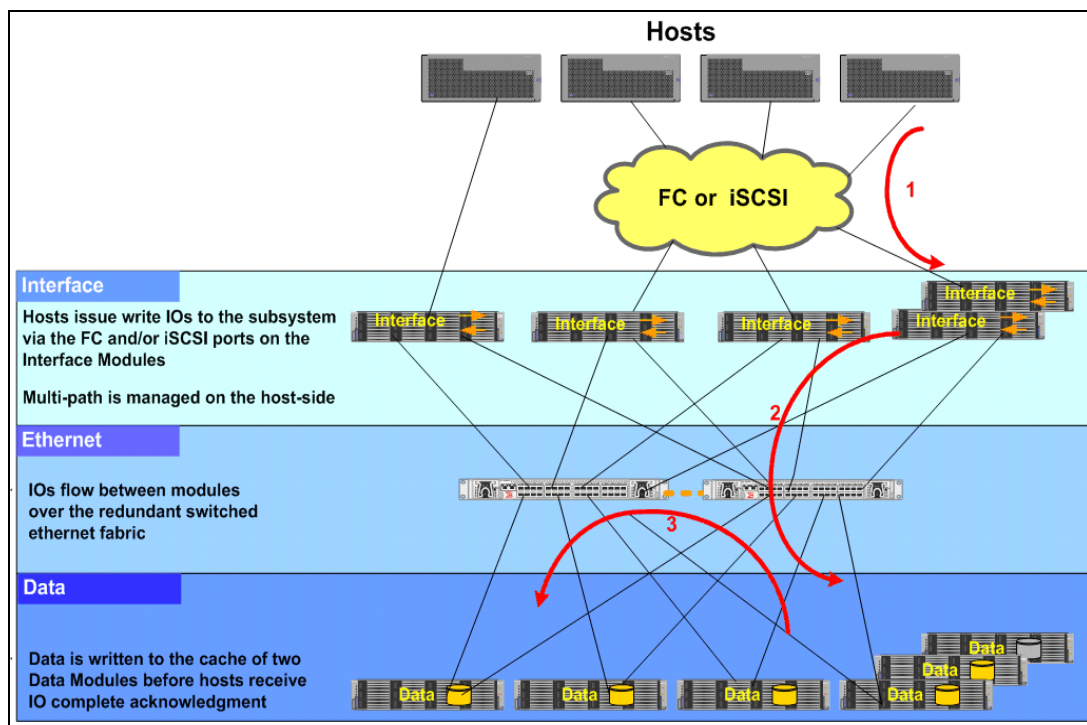


Figure 1-4 Write path

1.3.4 System quiesce and graceful shutdown

When an event occurs that compromises both sources of power to the redundant uninterruptible power supplies of the IBM XIV Storage System server, the system executes the graceful shutdown sequence. Full battery power is guaranteed during this event, because the system monitors available battery charge at all times and takes proactive measures to prevent the possibility of conducting write operations when battery conditions are non-optimal.

Because of the grid topology of the XIV system, a system quiesce event entails the graceful shutdown of all modules within the system. Each module can be thought of as an independent entity that is responsible for managing the destaging of *dirty data*, that is, written data that has not yet been destaged to physical disk. The dirty data within each module consists of equal parts primary and secondary copies of data, *but will never* contain both primary and secondary copies of the same data.

Each module in the XIV system contains a local, independent space that is reserved for caching operations within its system memory. In addition, each module contains 8 GB of high speed volatile *memory* (a total of 120 GB), from which 5.5 GB (and 82.5 GB overall) are dedicated for caching data.

Note: The system does *not* contain non-volatile memory space that is reserved for write operations. However, the close proximity of the cache and the drives, in conjunction with the enforcement of an upper limit for dirty, or non-destaged, data on a per-drive basis, ensures that the full destage will occur while operating under battery power.

1.3.5 Rebuild and redistribution

The IBM XIV Storage System server dynamically maintains the pseudo-random distribution of data across all modules and disks while ensuring that two copies of data exist at all times when the system reports *Full Redundancy*. Obviously, when there is a change to the hardware infrastructure as a result of a failed component, data must be restored to redundancy and distributed, or when a component is added, or *phased-in*, a new data distribution must accommodate the change.

Goal distribution

The process of achieving a new goal distribution while simultaneously restoring data redundancy because of the loss of a disk or module is called a *rebuild*. Because a rebuild occurs as a result of a component failure that consists full data redundancy, a period exists during which the *non-redundant data* is both restored to full redundancy and homogeneously redistributed over the remaining disks.

The process of achieving a new goal distribution (only occurring when redundancy exists) is known as *redistribution*, during which all data in the system (including both primary and secondary copies) is redistributed, when it is a result of the following events:

- ▶ The replacement of a failed disk or module following a rebuild, also called a *phase-in*
- ▶ When one or more modules are added to the system, called a *scale-out upgrade*

Following either of these occurrences, the XIV system immediately initiates the following sequence of events:

1. The XIV distribution algorithms calculate which partitions must be relocated and copied. The resultant distribution table is called the *goal distribution*.
2. The data modules and interface modules begin concurrently redistributing and copying (in the case of a rebuild) the partitions according to the goal distribution:
 - This process occurs in a parallel, any-to-any fashion concurrently among all modules and drives in the background, with complete host transparency.
 - The priority that is associated with achieving the new goal distribution is internally determined by the system. The priority cannot be adjusted by the storage administrator:
 - Rebuilds have the highest priority. However, the transactional load is homogeneously distributed over all the remaining disks in the system resulting in a low density of system-generated transactions.
 - Phase-outs (caused by the XIV technician removing and replacing a failed module) have a lower priority than rebuilds, because at least two copies of all data exist at all times during the phase-out.
 - Redistributions have the lowest priority, because there is neither a lack of data redundancy nor has the system detected the potential for an impending failure.
3. The system reports *Full Redundancy* after the goal distribution has been met.

Following the completion of goal distribution resulting from a rebuild or phase-out, a subsequent redistribution must occur when the system hardware is fully restored through a phase-in.

Goal distribution: The goal distribution is transparent to storage administrators and cannot be changed. In addition, the goal distribution has many determinants depending on the precise state of the system.

Important: Never perform a phase-in to replace a failed disk or module until after the rebuild process has completed. These operations must be performed by the XIV technician.

Preserving data redundancy

Conventional storage systems maintain a static relationship between RAID arrays and logical volumes by preserving data redundancy only across a subset of disks that are defined in the context of a particular RAID array. However, the XIV system dynamically and fluidly restores redundancy and equilibrium across all disks and modules in the system during the rebuild and phase-out operations.

1.4 Physical architecture and components

In this section, we describe the IBM XIV Storage System hardware components.

1.4.1 IBM XIV Storage System models 2810-A14 and 2812-A14

Figure 1-5 shows the XIV system.



Figure 1-5 Front and rear views of the XIV system

The IBM XIV Storage System family consists of two machine types, the 2810-A14 and the 2812-A14. The 2812 machine type comes standard with a three-year manufacturer warranty. The 2810 machine type is delivered with a one-year standard warranty. Most of the hardware features are the same for both machine types. Table 1-1 lists the major differences.

Table 1-1 Machine type comparisons

Machine type	2810-A14	2812-A14
Warranty	1 year	3 years
CPUs per interface module	1 or 2 years	2 years
CPUs per data module	1 year	1 year

Figure 1-6 summarizes the main hardware characteristics of the XIV models 2810-A14 and 2812-A14.

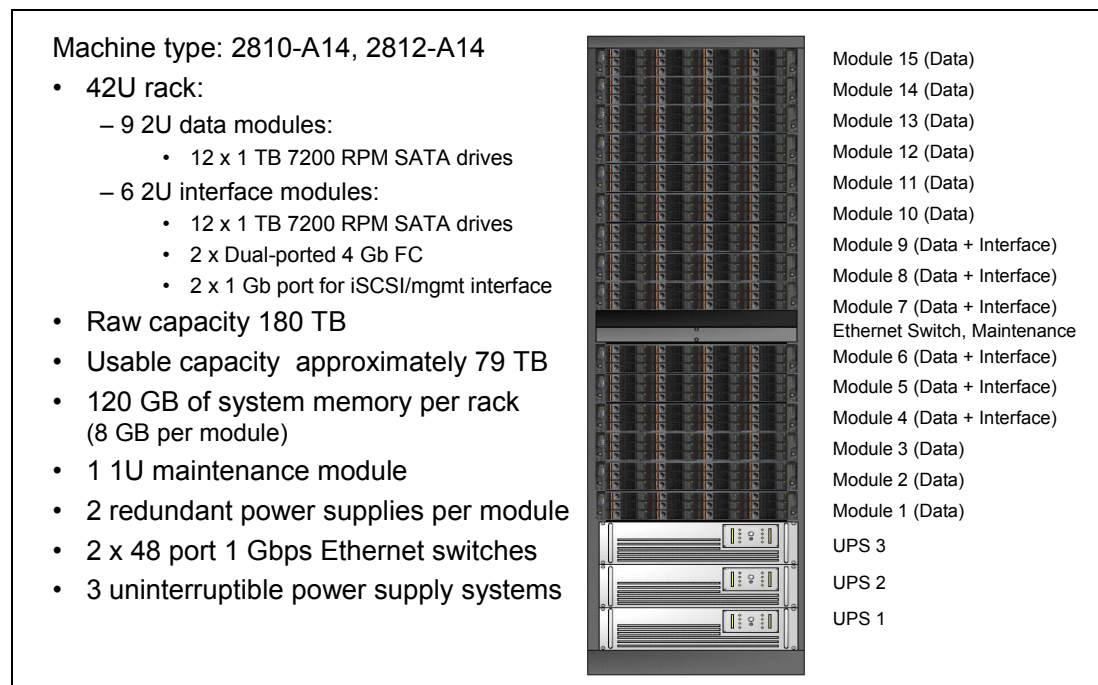


Figure 1-6 Hardware overview for the 2810-A14 and 2812-A14

All XIV hardware components come pre-installed in a standard 19-inch data center class rack. At the bottom of the rack, an uninterruptible power supply module complex, which is made up of three redundant uninterruptible power supply units, is installed and provides power to the various system components.

Fully populated configurations

A fully populated rack contains 9 data modules and 6 interface modules for a total of 15 modules. Each module is equipped with the following connectivity adapters:

- ▶ USB ports
- ▶ Serial ports
- ▶ Ethernet adapters
- ▶ Fibre Channel adapters (interface modules only)

Each module also contains twelve 1 TB SATA disk drives. This design translates into a total usable capacity of 79 TB (180 TB raw) for the complete system. For information about *usable capacity*, see 1.2.3, “Full storage virtualization” on page 5.

Partially populated configurations

The 2810-A14 and 2812-A14 models are also available in partially configured racks, allowing for more granular capacity configurations. These partially configured racks are available in usable capacities ranging from 27 to 79 TB.

Figure 1-7 shows details for these configuration options and the various capacities, drives, ports, and memory.

Total modules	6	9	10	11	12	13	14	15
Usable capacity (TB)	27	43	50	54	61	66	73	79
Interface modules (Feature #1100)	3	6	6	6	6	6	6	6
Data modules (Feature #1105)	3	3	4	5	6	7	8	9
Disk drives	72	108	120	132	144	156	168	180
Fibre Channel ports	8	16	16	20	20	24	24	24
iSCSI ports	0	4	4	6	6	6	6	6
Memory (GB)	48	72	80	88	96	104	112	120

Figure 1-7 XIV partial configurations

1.4.2 IBM XIV Storage System hardware components

At a minimum, for all configurations, the following components are supplied with the system:

- ▶ 3 uninterruptible power supply units
- ▶ 2 Ethernet switches
- ▶ 1 Ethernet switch redundant power supply
- ▶ 1 maintenance module
- ▶ 1 Automatic Transfer Switch (ATS)
- ▶ 1 modem
- ▶ 8-24 Fibre Channel ports
- ▶ 0-6 iSCSI ports
- ▶ Complete set of internal cabling

Note: For the same reason that the system is not dependent on specially developed parts, there might be differences in the actual hardware components that are used in your particular system compared with those components described in the following sections.

In the following sections, we describe some of these hardware components.

Rack

The XIV hardware components are installed in a standard 482.6 mm (19-inch) rack XIV Model 2810 and 2812 redesigned door. Adequate space is provided to house all components and to properly route all cables. The rack door and side panels are locked with a key to prevent unauthorized access to the installed components.

Uninterruptible power supply module complex

The uninterruptible power supply module complex consists of three uninterruptible power supply units. Each unit maintains an internal 30 seconds of storage of system power, for use in the event of any temporary failure of the external power supply to protect the system from failure. In case of an extended external power failure or outage, the uninterruptible power supply module complex maintains battery power long enough to allow a safe and orderly shutdown of the XIV system. The complex can sustain the failure of one uninterruptible power supply unit, while protecting against external power outages.

Data module

The fully populated rack hosts 9 data modules (modules 1-3 and modules 10-15). The only difference between data modules and interface modules (see “Interface module” on page 22) is the additional host adapters and Gigabit Ethernet adapters in the interface modules and the option of a dual CPU configuration for the newer interface modules.

Each data module contains four redundant Gigabit Ethernet ports. These ports together with the two switches form the internal network, which is the communication path for data and metadata between all modules. One Dual Gigabit Ethernet adapter is integrated in the System Planar (port 1 and 2). The remaining two ports (3 and 4) are on an additional Dual Gigabit Ethernet adapter installed in a Peripheral Component Interconnect Express (PCIe) slot as seen in Figure 1-8.

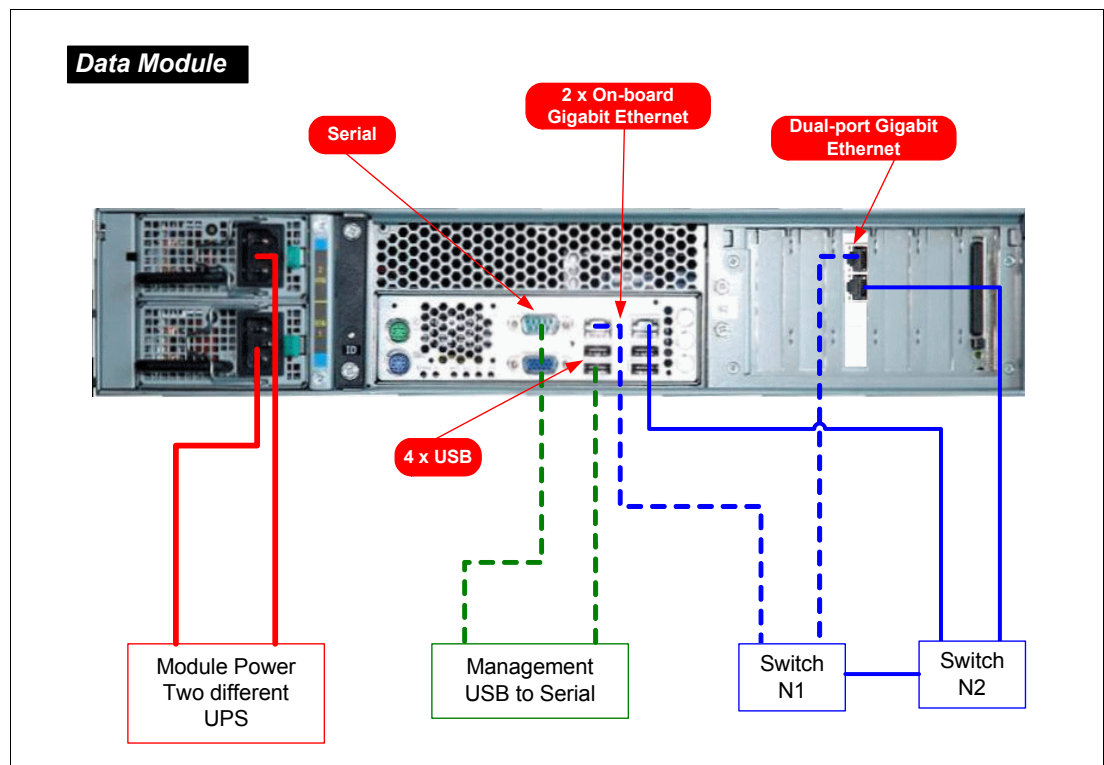


Figure 1-8 Data module connections

Interface module

The interface module is similar to the data module. The only difference is that it contains iSCSI and Fibre Channel ports through which hosts can attach to the XIV system.

Four Fibre Channel ports are available in each interface module for a total of 24 Fibre Channel ports. They support 1, 2, and 4 Gbps full-duplex data transfer over short wave fibre links, using a 50 micron multi-mode cable. They also support new end-to-end error detection through a Cyclic Redundancy Check (CRC) for improved data integrity during reads and writes.

In each module, the Fibre Channel ports are allocated in the following manner:

- ▶ Ports 1 and 3 are allocated for host connectivity.
- ▶ Ports 2 and 4 are allocated for additional host connectivity or remote mirror and data migration connectivity.

Note: Using more than 12 Fibre Channel ports for host connectivity does not necessarily provide more bandwidth. Use enough ports to support multipathing, without overburdening the host with too many paths to manage.

Six iSCSI ports (two ports per interface modules 7 through 9) are available for iSCSI over IP/Ethernet services. These ports support a 1 Gbps Ethernet network connection, connect to the user's IP network through the Patch Panel, and provide connectivity to the iSCSI hosts.

SATA disk drives

The SATA disk drives, used in the XIV system, are 1 TB, 7200 rpm hard drives designed for high-capacity storage in enterprise environments. These drives are manufactured to a higher set of standards than typical off-the-shelf SATA disk drives to ensure a longer life and increased mean time between failure (MTBF).

The XIV system was engineered with substantial protection against data corruption and data loss. Several features and functions implemented in the disk drive also increase reliability and performance:

- ▶ SAS interface

The disk drive features a 3 Gbps SAS interface supporting key features in the SATA specification, including Native Command Queuing (NCQ) and staggered spin-up and hot-swap capability.

- ▶ 32 MB cache buffer

The internal 32 MB cache buffer enhances the data transfer performance.

- ▶ Rotation Vibration Safeguard (RVS)

In multi-drive environments, rotational vibration, which results from the vibration of neighboring drives in a system, can degrade hard drive performance. To aid in maintaining high performance, the disk drive incorporates the enhanced RVS technology, providing up to a 50% improvement over the previous generation against performance degradation, and therefore, leading the industry.

- ▶ Advanced magnetic recording heads and media

There is an excellent soft error rate for improved reliability and performance.

- ▶ Self-Protection Throttling (SPT)

SPT monitors and manages I/O to maximize reliability and performance.

- ▶ Thermal Fly-height Control (TFC)
TFC provides a better soft error rate for improved reliability and performance.
- ▶ Fluid Dynamic Bearing (FDB) Motor
The FDB Motor improves acoustics and positional accuracy.
- ▶ Load/unload ramp
The read/write heads are placed outside the data area to protect user data when the power is removed.

All XIV disks are installed in the front of the modules, twelve disks per module. Each SATA disk is installed in a disk tray, which connects the disk to the backplane and includes the disk indicators on the front. If a disk is failing, it can be replaced easily from the front of the rack. The complete disk tray is one FRU, which is latched in its position by a mechanical handle.

Important: Never swap SATA disks in the XIV system within a module nor place them in another module because of internal tracing and logging data that they maintain.

Interconnection and switches

The internal network is based on two redundant 48-port Gigabit Ethernet switches. Each of the modules (data or interface) is directly attached to each of the switches with multiple connections, and the switches are also linked to each other. This network topology enables maximum bandwidth utilization, because the switches are used in an active-active configuration, while being tolerant to any failure of the following individual network components:

- ▶ Ports
- ▶ Links
- ▶ Switches

The Gigabit Ethernet Layer 3-Switch contains 48 copper and 4 fiber ports (small form-factor pluggable (SFP) capable of one of 3 speeds, 10/100/10000 Mbps), robust stacking, and 10 Gigabit Ethernet uplink capability. The switches are powered by redundant power supplies to eliminate any single point of failure.

1.5 IBM XIV Storage Management software

The IBM XIV Storage Management software is used to communicate with the IBM XIV Storage System software, which in turn interacts with the XIV hardware. You can install it on a Microsoft® Windows®, Linux®, AIX®, HPUX, or Solaris workstation that then acts as the management console for the XIV system.

The XIV Storage Management software is provided at the time of installation. Optionally you can download it from the following Web address:

<http://www.ibm.com/systems/support/storage/XIV>

For detailed information about the compatibility of the XIV Storage Management software, see the XIV interoperability matrix or the System Storage Interoperability Center (SSIC) at the following address:

<http://www.ibm.com/systems/support/storage/config/ssic/index.jsp>

For instructions about how to install the XIV Storage Management software, see Chapter 4, “Configuration,” in the Redbooks publication *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

The IBM XIV Storage Management software includes a user-friendly and intuitive GUI application, as well as an XCLI component that offers a comprehensive set of commands to configure and monitor the system. The XCLI is a powerful text-based, command line-based tool that enables an administrator to issue simple commands to configure, manage, or maintain the system. This tool includes the definitions that are required to connect to hosts and applications. The XCLI tool can be used in an XCLI session environment to interactively configure the system or as part of a script to perform lengthy and more complex tasks.

The basic storage system configuration sequence that is followed in this chapter includes the initial installation steps, followed by the disk space definition and management. Figure 1-9 shows an overview of the configuration flow. Keep in mind that, because the XIV Storage Management GUI is extremely intuitive, you can easily and quickly achieve most configuration tasks.

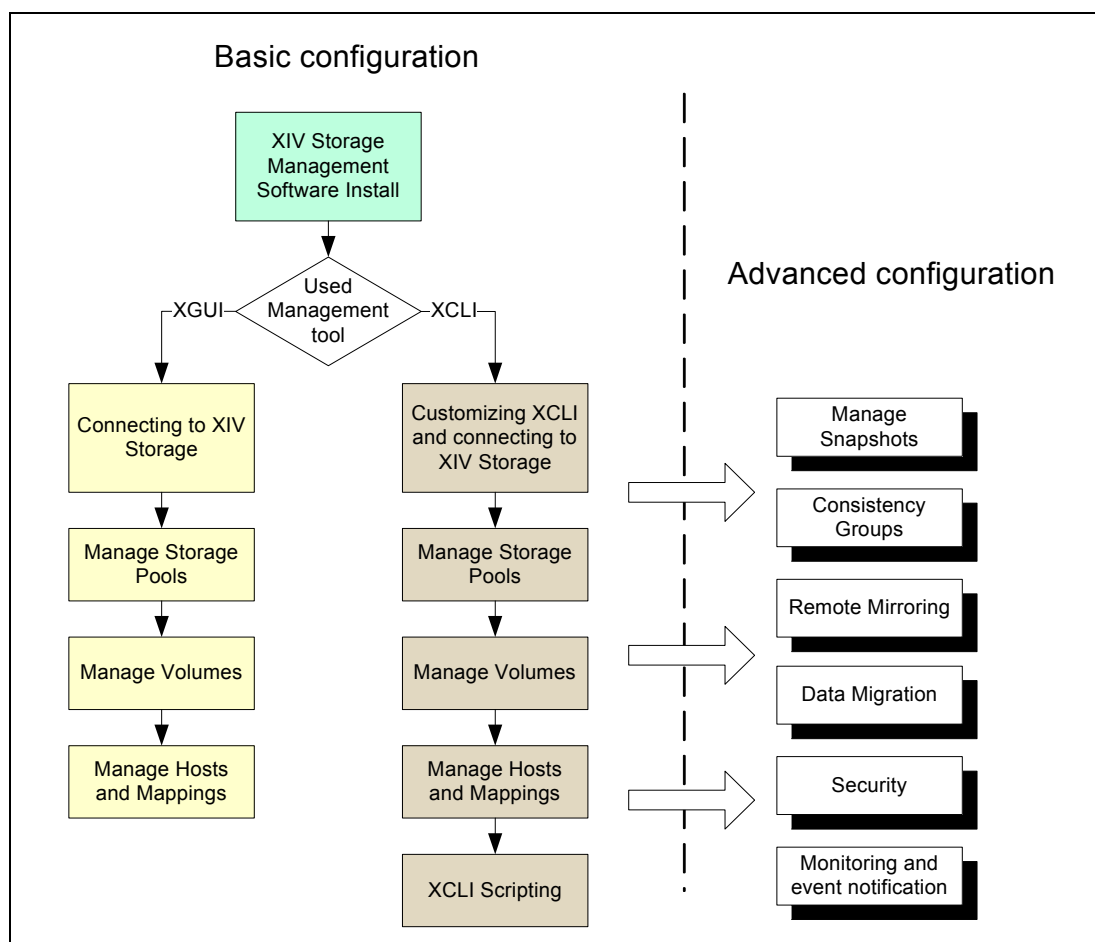


Figure 1-9 Basic configuration sequence and advanced tasks

1.5.1 The XIV Storage Management GUI

In this section, we take you through the XIV Storage Management GUI. We explain how to start it and present tips for navigating its options and features.

Launching the XIV Storage Management GUI

When you launch the XIV Storage Management GUI application, a login window prompts you for a user name and the corresponding password to grant access to the XIV system. The default user is *admin*, and the default corresponding password is *adminadmin*, as shown in Figure 1-10.

Important: You must change the default passwords to properly secure your system.

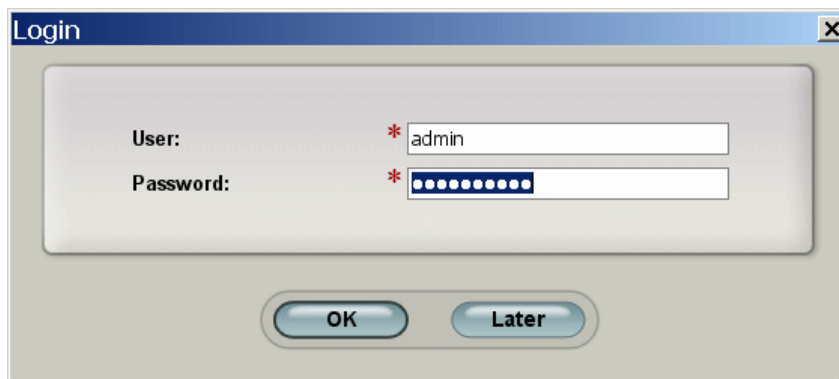


Figure 1-10 Login window with default access

The default admin user comes with storage administrator (storageadmin) rights. The XIV system offers role-based user access management.

To connect to an XIV system, you must initially add the system to make it visible in the GUI by specifying its IP addresses.

To add the system:

1. Make sure that the management workstation is set up to have access to the LAN subnet where the XIV system resides. Verify the connection by pinging the IP address of the XIV system.

If this is the first time you start the GUI on this management workstation and no XIV system was previously defined to the GUI, the Add System Management window (Figure 1-11 on page 26) automatically opens and you perform either of the following actions:

- If the default IP address of the XIV system was not changed, select **Connect Directly**, which populates the IP/DNS Address1 field with the default IP address. Click **Add** to effectively add the system to the GUI.
- If the default IP address was changed to a client-specified IP address (or set of IP addresses, for redundancy), you *must* enter those addresses in the IP/DNS Address fields. Click **Add** to effectively add the system to the GUI.

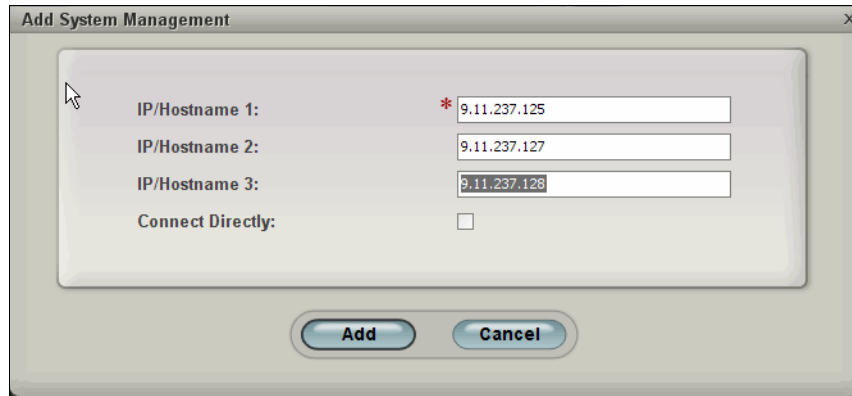


Figure 1-11 Add System Management window

2. After you return to the main XIV Storage Management window, wait until the system is displayed and indicates a status of *enabled*. Under normal circumstances, the system shows a status of *Full Redundancy* in a green label box.
3. Move the mouse pointer over the image of the XIV system and click to open the XIV Storage System Management main window (Figure 1-12).

XIV Storage Management GUI main menu

The XIV Storage Management GUI is mostly self-explanatory with a well organized structure and simple navigation (Figure 1-12).

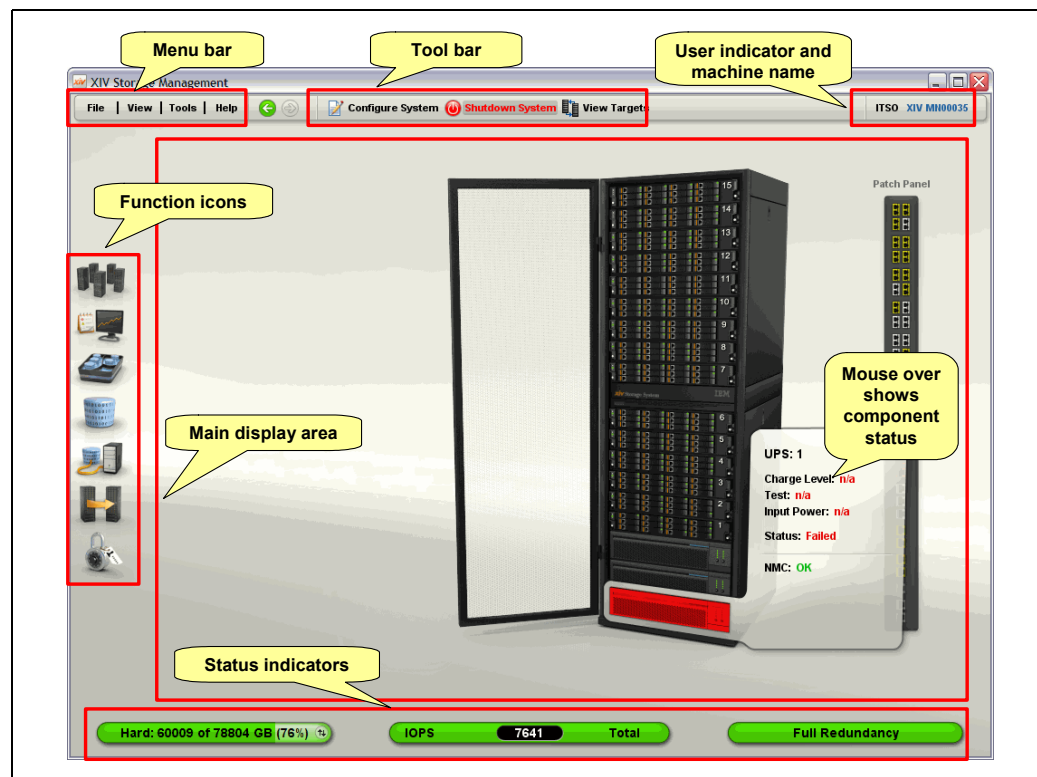


Figure 1-12 XIV Storage Management main window: System view

The main window of the XIV Storage Management software is divided into several areas:

- **Function icons**

This set of vertically stacked icons, on the left side of the main window, is used to navigate between the functions of the GUI based on the icon that is selected. Moving the mouse pointer over an icon opens a corresponding option menu. Figure 1-13 shows the various menu options that are available from the function icons.

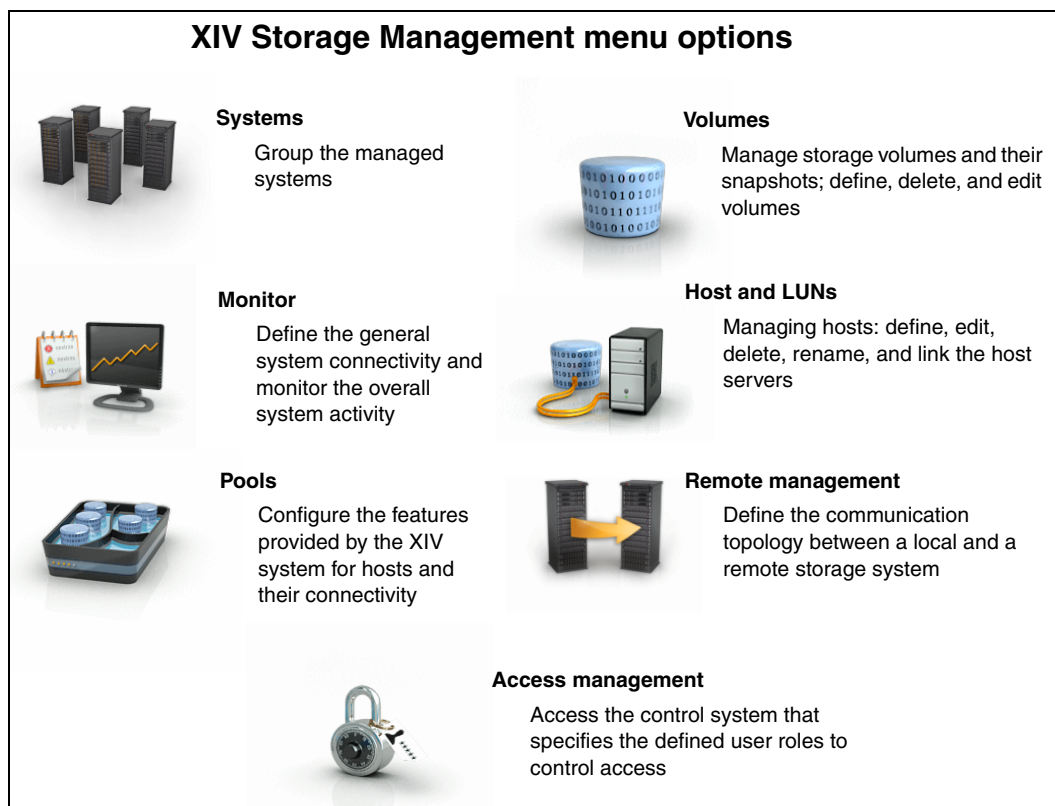


Figure 1-13 Menu items in the XIV Storage Management software

- **Main display**

This area occupies the major part of the window and provides graphical representation of the XIV system. Moving the mouse pointer over the graphical representation of a specific hardware component (module, disk, and uninterruptible power supply unit) open a status callout. When a specific function is selected, the main display shows a tabular representation of that function.

- **Menu bar**

The menu bar is used to configure the system. It is also used as an alternative to the Function icons for accessing the various functions of the XIV system.

- **Toolbar**

The toolbar is used to access a range of specific actions that are linked to the individual functions of the system.

► Status bar

The status bar is at the bottom of the window and indicates the overall operational status of the XIV system:

- The first indicator on the left shows the amount of soft or hard storage capacity that is currently allocated to storage pools. It also provides alerts when certain capacity thresholds are reached. As the physical, or hard, capacity consumed by volumes within a storage pool passes certain thresholds, the color of this meter indicates that additional hard capacity might need to be added to one or more storage pools.
- The second indicator (in the middle) displays the number of I/O operations per second (IOPS).
- The third indicator on the far right shows the general system status and will, for example, indicate when a redistribution is underway.

In addition, an Uncleared Event indicator is visible when events occur for which a repetitive notification (called an *Alerting Event*) was defined that has not yet been cleared in the GUI.

Tip: The configuration information regarding the connected systems and the GUI is stored in various files under the user's home directory.

As a useful and convenient feature, all the commands issued from the GUI are saved in a log in the format of the XCLI syntax. This syntax includes quoted strings. However, the quotation marks are needed only if the value that is specified contains blanks.

The default location is in the Documents and Setting folder of the Microsoft Windows current user, for example:

```
%HOMEDRIVE%%HOMEPATH%\Application Data\XIV\GUI10\logs\guiCommands*.log
```

XIV Storage Management GUI features

Several features, as we describe in this section, enhance the usability of the GUI menus. These features enable you to quickly and easily execute tasks.

As shown in Figure 1-14, commands can be executed against multiple selected objects.

Multiple objects commands

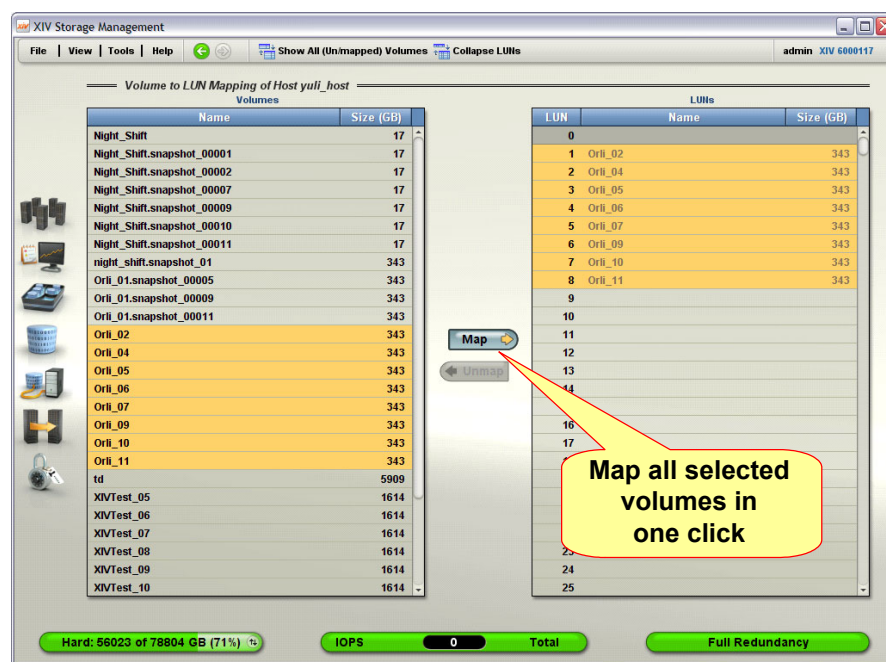


Figure 1-14 Multiple objects commands

As shown in Figure 1-15, menu tips are displayed when the mouse pointer is placed over disabled menu items. The tips explain why the item is not selectable in a given context.

Menu tool tips

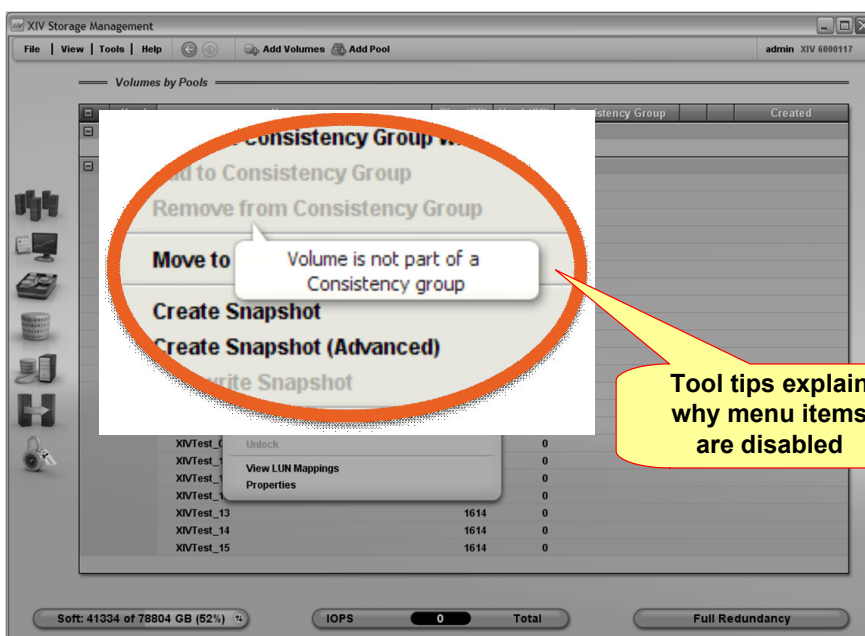


Figure 1-15 Menu tool tips

Figure 1-16 illustrates the keyboard navigation shortcuts.

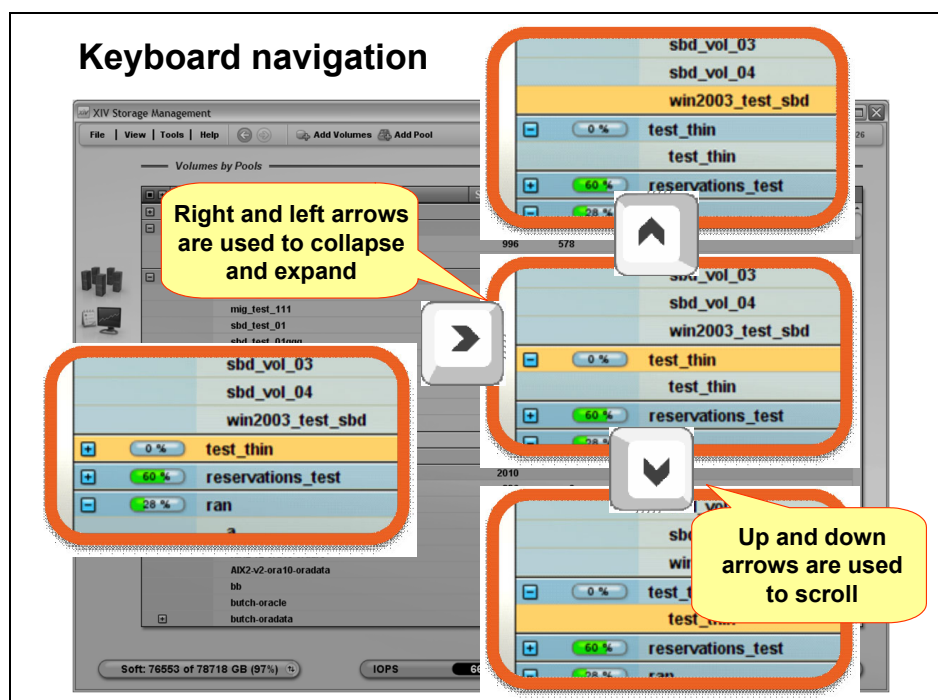


Figure 1-16 Keyboard navigation

For more information about the XIV Storage Management GUI, see Chapter 4, “Configuration,” in *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

1.5.2 Using the XCLI

After you install the XIV Storage Management software, the XCLI session link is added to the desktop. Usage of the XCLI session environment is simple and has interactive help for command syntax. We provide examples of the XCLI command in this section.

To start an XCLI session, click the desktop link or click **Launch XCLI** from the systems menu in the GUI (as shown in Figure 1-17). Starting from the GUI automatically provides the current user ID and password and connects you to the selected system. Otherwise you are prompted for user information and the IP address of the system.

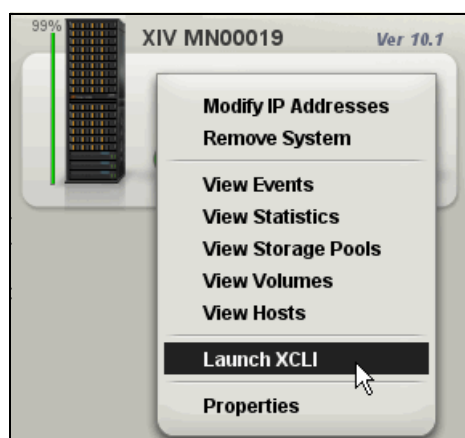


Figure 1-17 Clicking Launch XCLI from the GUI

You can define a *script* to specify the name and path to the commands file. (The lists of commands are executed in User Mode only).

XCLI session features

The XCLI session offers command and argument completions, along with possible values for the arguments. There is no need to enter user information or IP addresses for each command:

- ▶ Executing a command: Simply type the command.
- ▶ Command completion: Type part of a command and press Tab to see possible valid commands.
- ▶ Command argument completion: Type a command and press Tab to see a list of values for the command argument (Figure 1-18).

```
>> user_ <TAB>
user_define      user_delete      user_group_add_user  user_group_create
user_group_delete user_group_list  user_group_remove_user user_group_rename
user_group_update user_list        user_rename          user_update

>> user_list <TAB>
show_users=      user=

>> user_list user= <TAB>
xiv_development  xiv_maintenance  admin    technician
smis_user        ITS0

>> user_list user=admin
Name  Category    Group Active Email Address Area Code Phone Number Access All
admin storageadmin      yes
```

Figure 1-18 XCLI session example

Help with the XCLI commands

For help about the usage and commands, type the help commands as shown in Example 1-1.

Example 1-1 XCLI help commands

```
xcli
xcli -c "XIV MN00035" help
xcli -c "XIV MN00035" help command=user_list format=full
```

The first command prints the usage of **xcli**. The second command prints all the commands that can be used by the user in that particular system. The third command one shows the usage of the **user_list** command with all the parameters.

There are various parameters to get the result of a command in a predefined format. The default is a user readable format. You can specify the **-s** parameter to get it in a comma-separated format or specify the **-x** parameter to obtain it in an XML format.

Note: The XML format contains all the fields of a particular command. The user and the comma-separated formats provide the default fields as a result.

To list the field names for a specific `xcli` command output, use the `-t` parameter as shown in Example 1-2.

Example 1-2 XCLI field names

```
xcli -c "XIV MN00035" -t name,fields help command=user_list
```

For more information about using XCLI, see Chapter 4, "Configuration," in *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

For complete and detailed documentation of the IBM XIV Storage Management software, see the *XCLI Reference Guide*, GC27-2213, and the *XIV Session User Guide*. You can find both of these documents in the IBM XIV Storage System Information Center at the following address:

<http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp>



Overview of the Virtual I/O Server with IBM i

In this chapter, we provide an overview of IBM PowerVM™ and especially the Virtual I/O Server (VIOS). We describe different Fibre Channel connectivity configurations in the VIOS and recommend how to best set up the VIOS for connecting the IBM XIV Storage System server to an IBM i client.

2.1 Introduction to IBM PowerVM

Virtualization on IBM Power Systems™ servers can provide a rapid and cost-effective response to many business needs. Virtualization capabilities are becoming an important element in planning for IT floorspace and servers. With growing commercial and environmental concerns, there is pressure to reduce the power footprint of servers. Also with the escalating cost of powering and cooling servers, consolidation and efficient utilization of the servers is becoming critical.

Virtualization on Power Systems servers enables an efficient utilization of servers by reducing the following areas:

- ▶ Server management and administration costs, because there are fewer physical servers
- ▶ Power and cooling costs with increased utilization of existing servers
- ▶ Time to market, because virtual resources can be deployed immediately

2.1.1 IBM PowerVM overview

IBM PowerVM is a special software appliance tied to IBM Power Systems, that is, the converged IBM i and IBM p server platforms. It is licensed on a Power Systems processor basis. PowerVM is a virtualization technology for AIX, IBM i, and Linux environments on IBM POWER® processor-based systems.

PowerVM offers a secure virtualization environment with the following major features and benefits:

- ▶ Consolidates diverse sets of applications that are built for multiple operating systems (AIX, IBM i, and Linux) on a single server
- ▶ Virtualizes processor, memory, and I/O resources to increase asset utilization and reduce infrastructure costs
- ▶ Dynamically adjusts server capability to meet changing workload demands
- ▶ Moves running workloads between servers to maximize availability and avoid planned downtime

Virtualization technology is offered in three editions on Power Systems:

- ▶ PowerVM Express Edition
- ▶ PowerVM Standard Edition
- ▶ PowerVM Enterprise Edition

They provide logical partitioning technology by using either the Hardware Management Console (HMC) or the Integrated Virtualization Manager (IVM), dynamic logical partition (LPAR) operations, Micro-Partitioning™ and VIOS capabilities, and Node Port ID Virtualization (NPIV).

PowerVM Express Edition

PowerVM Express Edition is available only on the IBM Power 520 and Power 550 servers. It is designed for clients who want an introduction to advanced virtualization features at an affordable price.

With PowerVM Express Edition, clients can create up to three partitions on a server (two client partitions and one for the VIOS and IVM). They can use virtualized disk and optical devices, as well as try the shared processor pool. All virtualization features, such as

Micro-Partitioning, shared processor pool, VIOS, PowerVM LX86, shared dedicated capacity, NPIV, and virtual tape, can be managed by using the IVM.

PowerVM Standard Edition

For clients who are ready to gain the full value from their server, IBM offers the PowerVM Standard Edition. This edition provides the most complete virtualization functionality for UNIX® and Linux in the industry and is available for all IBM Power Systems servers.

With PowerVM Standard Edition, clients can create up to 254 partitions on a server. They can use virtualized disk and optical devices and try out the shared processor pool. All virtualization features, such as Micro-Partitioning, Shared Processor Pool, Virtual I/O Server, PowerVM Lx86, Shared Dedicated Capacity, NPIV, and Virtual Tape, can be managed by using an Hardware Management Console or the IVM.

PowerVM Enterprise Edition

PowerVM Enterprise Edition is offered exclusively on IBM POWER6® servers. It includes all the features of the PowerVM Standard Edition, plus the *PowerVM Live Partition Mobility* capability.

With PowerVM Live Partition Mobility, you can move a running partition from one POWER6 technology-based server to another with no application downtime. This capability results in better system utilization, improved application availability, and energy savings. With PowerVM Live Partition Mobility, planned application downtime because of regular server maintenance is no longer necessary.

2.1.2 Virtual I/O Server

The VIOS is virtualization software that runs in a separate partition of the POWER system. VIOS provides virtual storage and networking resources to one or more client partitions.

The VIOS owns the physical I/O resources such as Ethernet and SCSI/FC adapters. It virtualizes those resources for its client LPARs to share them remotely by using the built-in hypervisor services. These client LPARs can be created quickly, typically owning only real memory and shares of CPUs without any physical disks or physical Ethernet adapters.

With Virtual SCSI support, VIOS client partitions can share disk storage that is physically assigned to the VIOS LPAR. This virtual SCSI support of VIOS is used to make storage devices, such as the IBM XIV Storage System server, that do not support the IBM i proprietary 520-byte/sectors format that is available to IBM i clients of VIOS.

VIOS owns the physical adapters, such as the Fibre Channel storage adapters that are connected to the XIV system. The logical unit numbers (LUNs) of the physical storage devices that are detected by VIOS are mapped to VIOS virtual SCSI (VSCSI) server adapters that are created as part of its partition profile.

The client partition with its corresponding VSCSI client adapters defined in its partition profile connects to the VIOS VSCSI server adapters by using the hypervisor. VIOS performs SCSI emulation and acts as the SCSI target for the IBM i operating system.

Figure 2-1 shows an example of the VIOS owning the physical disk devices and its virtual SCSI connections to two client partitions.

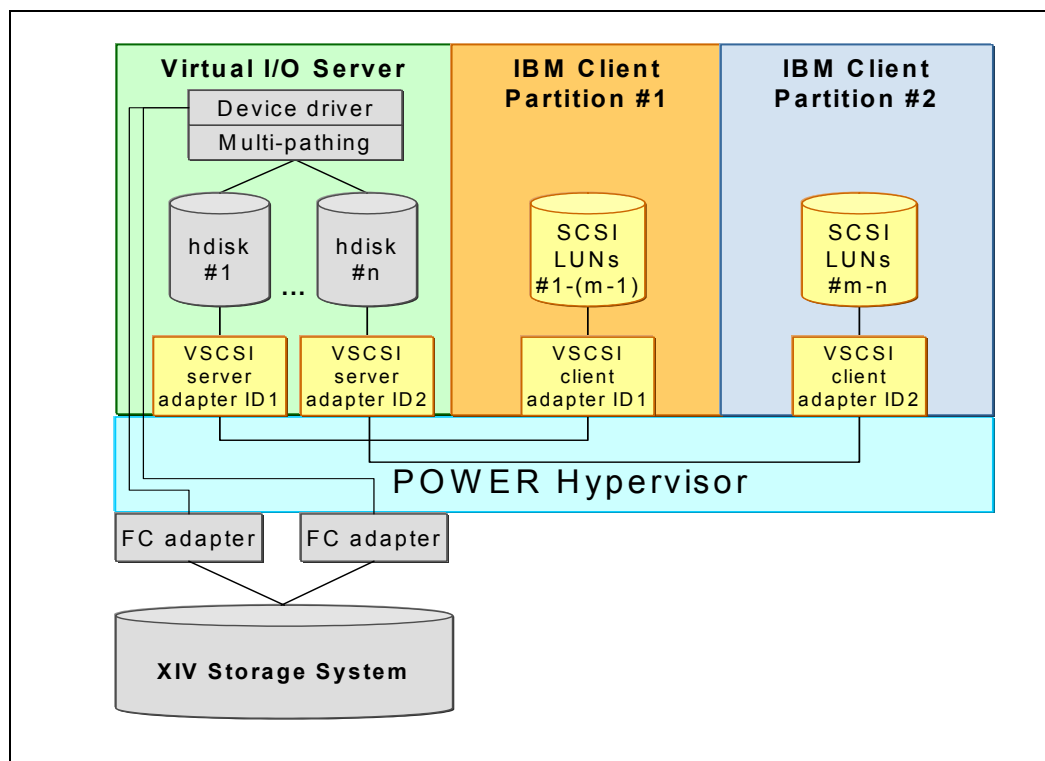


Figure 2-1 VIOS virtual SCSI support

2.1.3 Node Port ID Virtualization

The VIOS technology has been enhanced to boost the flexibility of IBM Power Systems servers with support for NPIV. NPIV simplifies the management and improves performance of Fibre Channel SAN environments by standardizing a method for Fibre Channel ports to virtualize a physical node port ID into multiple virtual node port IDs. The VIOS takes advantage of this feature and can export the virtual node port IDs to multiple virtual clients. The virtual clients see this node port ID and can discover devices as though the physical port was attached to the virtual client.

The VIOS does not do any device discovery on ports by using NPIV. Thus no devices are shown in the VIOS connected to NPIV adapters. The discovery is left for the virtual client and all the devices found during discovery are detected only by the virtual client. This way, the virtual client can use FC SAN storage specific multipathing software on the client to discover and manage devices.

For more information about PowerVM virtualization management, see the IBM Redbooks publication *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590.

Note: Connection through VIOS NPIV to an IBM i client is possible only for storage devices that can attach natively to the IBM i operating system, such as the IBM System Storage DS8000®. To connect other storage devices, use VIOS with virtual SCSI adapters.

2.2 PowerVM client connectivity to the IBM XIV Storage System server

The XIV system can be connected to an IBM i partition through VIOS. See 4.1, “Connecting a PowerVM client to the IBM XIV Storage System server” on page 48, for detailed instructions about how to set up the environment on an IBM POWER6 system to connect the XIV system to an IBM i client with multipath through two VIOS partitions.



Planning for the IBM XIV Storage System server with IBM i

In this chapter, we provide information to assist you in properly planning for the use of an IBM XIV Storage System server in an IBM i environment.

Note: In this paper, the IBM i operating system resides in a logical partition (LPAR) on an IBM Power Systems server or Power Blade platform.

3.1 Requirements

When connecting the IBM XIV Storage System server to an IBM i operating system by using the Virtual I/O Server (VIOS), you must have the following requirements in place:

- ▶ The IBM i partition and the VIOS partition must reside on a POWER6 processor-based IBM Power Systems server or POWER6 processor-based IBM Power Blade servers.

POWER6 processor-based IBM Power Systems include the following models:

- 8203-E4A (IBM Power 520 Express)
- 8261-E4S (IBM Smart Cube)
- 9406-MMA
- 9407-M15 (IBM Power 520 Express)
- 9408-M25 (IBM Power 520 Express)
- 8204-E8A (IBM Power 550 Express)
- 9409-M50 (IBM Power 550 Express)
- 8234-EMA (IBM Power 560 Express)
- 9117-MMA (IBM Power 570)
- 9125-F2A (IBM Power 575)
- 9119-FHA (IBM Power 595)

The following servers are POWER6 processor-based IBM Power Blade servers:

- IBM BladeCenter® JS12 Express
- IBM BladeCenter JS22 Express
- IBM BladeCenter JS23
- IBM BladeCenter JS43

- ▶ You must have one of the following PowerVM editions:

- PowerVM Express Edition, 5765-PVX
- PowerVM Standard Edition, 5765-PVS
- PowerVM Enterprise Edition, 5765-PVE

- ▶ You must have Virtual I/O Server Version 2.1.1 or later.

Virtual I/O Server is delivered as part of PowerVM.

- ▶ You must have IBM i, 5761-SS1, Release 6.1 or later.

- ▶ You must have one of the following Fibre Channel (FC) adapters supported to connect the XIV system to the VIOS partition in POWER6 processor-based IBM Power Systems server:

- 2 Gbps PCI-X 1-port Fibre Channel adapter, feature number 1957
- 2 Gbps PCI-X 1-port Fibre Channel adapter, feature number 1977
- 2 Gbps PCI-X 1-port Fibre Channel adapter, feature number 5716
- 2 Gbps PCI-X Fibre Channel adapter, feature number 6239
- 4 Gbps PCI-X 1-port Fibre Channel adapter, feature number 5758
- 4 Gbps PCI-X 2-port Fibre Channel adapter, feature number 5759
- 4 Gbps PCIe 1-port Fibre Channel adapter, feature number 5773
- 4 Gbps PCIe 2-port Fibre Channel adapter, feature number 5774
- 4 Gbps PCI-X 1-port Fibre Channel adapter, feature number 1905
- 4 Gbps PCI-X 2-port Fibre Channel adapter, feature number 1910
- 8 Gbps PCIe 2-port Fibre Channel adapter, feature number 5735

Note: Not all listed Fibre Channel adapters are supported in every POWER6 server listed in the first point. For more information about which FC adapter is supported with which server, see the IBM Redbooks publication *IBM Power 520 and Power 550 (POWER6) System Builder*, SG24-7765, and the IBM Redpaper publication *IBM Power 570 and IBM Power 595 (POWER6) System Builder*, REDP-4439.

The following Fibre Channel host bus adapters (HBAs) are supported to connect the XIV system to a VIOS partition on IBM Power Blade servers JS12 and JS22:

- LP1105-BCv - 4 Gbps Fibre PCI-X Fibre Channel Host Bus Adapter, P/N 43W6859
- IBM SANblade QMI3472 PCIe Fibre Channel Host Bus Adapter, P/N 39Y9306
- IBM 4 Gb PCI-X Fibre Channel Host Bus Adapter, P/N 41Y8527

The following Fibre Channel HBAs are supported to connect the XIV system to a VIOS partition on IBM Power Blade servers JS23 and JS43:

- IBM SANblade QMI3472 PCIe Fibre Channel Host Bus Adapter, P/N 39Y9306
 - IBM 44X1940 QLOGIC ENET & 8Gbps Fibre Channel Expansion Card for BladeCenter
 - IBM 44X1945 QMI3572 QLOGIC 8Gbps Fibre Channel Expansion Card for BladeCenter
 - IBM 46M6065 QMI2572 QLogic 4 Gbps Fibre Channel Expansion Card for BladeCenter
 - IBM 46M6140 Emulex 8Gb Fibre Channel Expansion Card for BladeCenter
 - IBM SANblade QMI3472 PCIe Fibre Channel Host Bus Adapter, P/N 39Y9306
- You must have IBM XIV Storage System firmware 10.0.1b and later.

3.2 SAN connectivity

In this section, we help you plan your IBM XIV Storage System SAN connectivity to an IBM i client.

Fibre Channel protocols, speed, distance

The Fibre Channel ports that are available in each XIV interface module support 1, 2, or 4 Gbps full duplex data transfer over short Fibre Channel links using a 50 micron multi-mode cable. The transfer rate depends on the capability of the device that is connected and is auto-negotiated with the device. If multiple devices and switches with different speeds are connected to the XIV Fibre Channel port, the XIV system automatically detects the lowest connection speed and adjusts to it.

The XIV system can be connected to the VIOS at a maximum distance of 500m (the maximum distance for short wave Fibre Channel links using 50 micron optical cables).

Supported SAN switches

For the list of supported SAN switches when connecting the XIV system to the IBM i operating system, see the System Storage Interoperation Center at the following address:

http://www-03.ibm.com/systems/support/storage/config/ssic/displaysssearchwithoutjs.wss?start_over=yes

Switch zoning

For redundancy, set up at least two connections from two Fibre Channel adapters in the VIOS, through two SAN switches, to the Fibre Channel interfaces of the XIV system. Each host connects to ports from at least two interface modules in the XIV system. See Figure 3-1 on page 44 for an example.

You must zone the switches so that each zone contains one FC Adapter on the host system (initiator) and multiple FC ports in the XIV system (target).

3.3 Planning for the Virtual I/O Server

In this section, we provide planning information for the VIOS partitions.

3.3.1 Physical Fibre Channel adapters and virtual SCSI adapters

It is possible to connect up to 4,095 logical unit numbers (LUNs) per target and up to 510 targets per port on a VIOS physical FC adapter. Because you can assign up to 16 LUNs to one virtual SCSI (VSCSI) adapter, you can use the number of LUNs to determine the number of virtual adapters that you need.

Note: When the IBM i operating system and VIOS reside on an IBM Power Blade server, you can define only one VSCSI adapter in the VIOS to assign to an IBM i client. Consequently the number of LUNs to connect to the IBM i operating system is limited to 16.

3.3.2 Queue depth in the IBM i operating system and Virtual I/O Server

When connecting the IBM XIV Storage System server to an IBM i client through the VIOS, consider the following types of queue depths:

- ▶ The IBM i queue depth to a virtual LUN
SCSI command tag queuing in the IBM i operating system enables up to 32 I/O operations to one LUN at the same time.
- ▶ The queue depth per physical disk (hdisk) in the VIOS
This queue depth indicates the maximum number of I/O requests that can be outstanding on a physical disk in the VIOS at a given time.
- ▶ The queue depth per physical adapter in the VIOS
This queue depth indicates the maximum number of I/O requests that can be outstanding on a physical adapter in the VIOS at a given time.

The IBM i operating system has a fixed queue depth of 32, which is not changeable. However, the queue depths in the VIOS can be set up by a user. The default setting in the VIOS varies based on the type of connected storage, type of physical adapter, and type of multipath driver or Host Attachment kit that is used. Typically for the XIV system, the queue depth per physical disk is 32, the queue depth per 4 Gbps FC adapter is 200, and the queue depth per 8 Gbps FC adapter is 500.

Check the queue depth on physical disks by entering the following VIOS command:

```
lsdev -dev hdiskxx -attr queue_depth
```

If needed, set the queue depth to 32 by using the following command:

```
chdev -dev hdiskxx -attr queue_depth=32
```

This command ensures that the queue depth in the VIOS matches the IBM i queue depth for an XIV LUN.

3.3.3 Multipath with two Virtual I/O Servers

The IBM XIV Storage System server is connected to an IBM i client partition through the VIOS. For redundancy, you connect the XIV system to an IBM i client with two or more VIOS partitions, with one VSCSI adapter in the IBM i operating system assigned to a VSCSI adapter in each VIOS. The IBM i operating system then establishes multipath to an XIV LUN, with each path using one different VIOS. For XIV attachment to the VIOS, the VIOS integrated native MPIO multipath driver is used. Up to eight VIOS partitions can be used in such a multipath connection. However, most installations might do multipath by using two VIOS partitions.

See 4.1.3, “IBM i multipath capability with two Virtual I/O Servers” on page 54, for more information.

3.4 Best practices

In this section, we present general best practices for IBM XIV Storage System servers that are connected to a host server. These practices also apply to the IBM i operating system.

With the grid architecture and massive parallelism inherent to XIV system, the recommended approach is to maximize the utilization of all the XIV resources at all times.

3.4.1 Distributing connectivity

The goal for host connectivity is to create a balance of the resources in the IBM XIV Storage System server. Balance is achieved by distributing the physical connections across the interface modules. A host usually manages multiple physical connections to the storage device for redundancy purposes by using a SAN connected switch. The ideal is to distribute these connections across each of the interface modules. This way, the host uses the full resources of each module to which it connects for maximum performance. It is *not* necessary for each host instance to connect to each interface module. However, when the host has more than one physical connection, it is beneficial to have the connections (cabling) spread across the different interface modules.

Similarly, if multiple hosts have multiple connections, you must distribute the connections evenly across the interface modules.

3.4.2 Zoning SAN switches

To maximize balancing and distribution of host connections to and IBM XIV Storage System server, create a zone for the SAN switches such that each host adapter connects to each XIV interface module and through each SAN switch. Figure 3-1 on page 44 illustrates this type of connection.

Note: Create a separate zone for each host adapter that connects to a switch, for each zone that contains the connection from the host adapter, and for all connections to the XIV system.

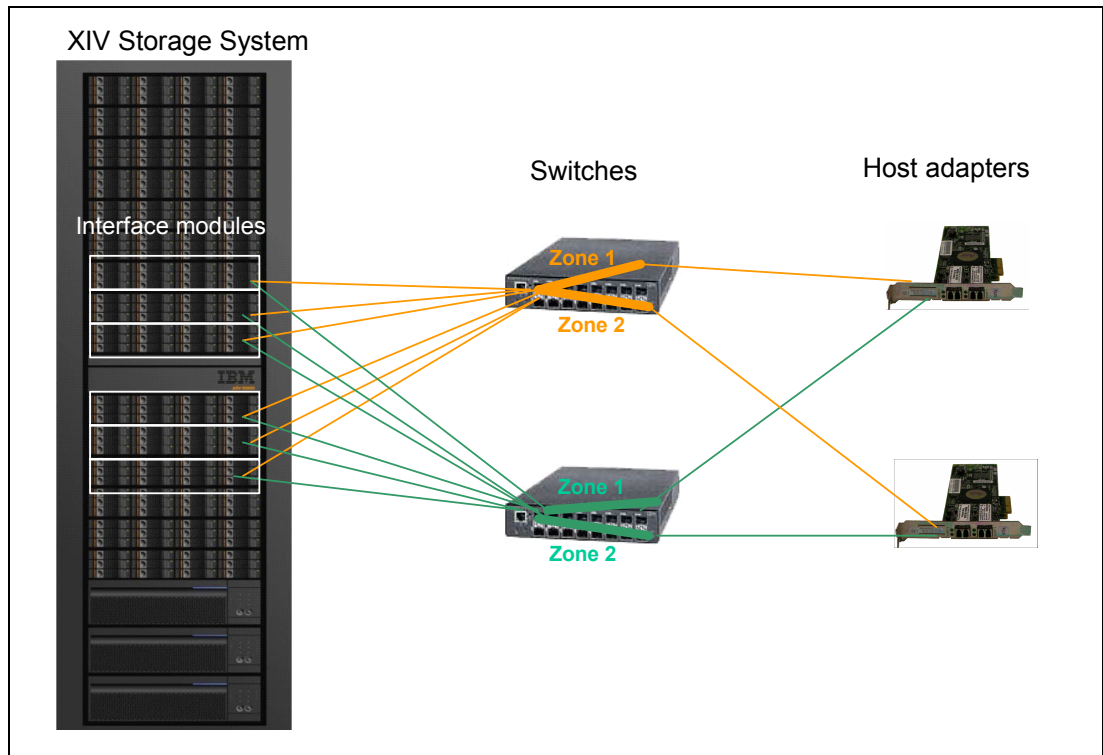


Figure 3-1 Zoning host connections

3.4.3 Queue depth

SCSI command tag queuing for LUNs on the IBM XIV Storage System server enables multiple I/O operations to one LUN at the same time. The LUN queue depth indicates the number of I/O operations that can be done simultaneously to a LUN.

The XIV architecture eliminates the existing storage concept of a large central cache. Instead, each module in the XIV grid has its own dedicated cache. The XIV algorithms that stage data between disk and cache work most efficiently when multiple I/O requests are coming in parallel. This is where the host queue depth becomes an important factor in maximizing XIV I/O performance. Therefore, configure the host HBA queue depths as large as possible.

3.4.4 Number of application threads

The overall design of the IBM XIV Storage System grid architecture excels with applications that employ threads to handle the parallel execution of I/Os. The multi-threaded applications will profit the most from XIV performance.

3.5 Planning for capacity

In this section, we summarize the storage capacity for different IBM XIV Storage System configurations. We also describe the sizes of storage pools and LUNs, as well as explain capacity reduction for an IBM i client.

3.5.1 Net usable capacity

Table 3-1 and Table 3-2 show the number of interface or data modules, dedicated data modules, number of disk units, spare capacity, and net usable capacity of the IBM XIV Storage System configurations.

Table 3-1 XIV configurations and storage capacity - Part 1

	Full	Partial	Partial	Partial
Number of modules	15	14	13	12
Number of interface or data modules	6	6	5	5
Number of dedicated data modules	9	8	8	7
Number of disks	180	168	156	144
Spare capacity (in disk units)	15	15	15	15
Net decimal capacity, rounded down in full TB	79 TB	73 TB	67 TB	61 TB

Table 3-2 XIV configurations and storage capacity - Part 2

	Partial	Partial	Partial	Partial
Number of modules	11	10	9	6
Number of Interface/Data modules	5	4	4	2
Number of dedicated data modules	6	6	5	4
Number of disks	132	120	108	72
Spare capacity (in disk units)	18	15	18	15
Net decimal capacity, rounded down in full TB	54 TB	50 TB	43 TB	27 TB

For more information about usable capacity, see the Redbooks publication *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

3.5.2 Storage pool capacity

When a regular storage pool is created, its size is rounded up to the multiple of 2^{34} bytes. The size is then presented as the integer part of the capacity expressed in decimal GB.

Consider an example where, when creating a regular storage pool, we define capacity at 3000 GB. The capacity is rounded up to the next multiple of 2^{34} . This translates into the following equation:

$$175 \times 17179869184 \text{ bytes} = 3006477107200 \text{ bytes} = 3006.477107200 \text{ GB (decimal)}$$

The size of storage pool is presented as an integer of a decimal in GB. Therefore, it is shown as 3006 GB.

Similarly, the size of a LUN in the regular storage pool is rounded up to the multiple of 2^{34} bytes and is presented as an integer part of the decimal in GB. When a thinly provisioned storage pool is created, both its hard size and soft size are expressed as a multiple of 2^{34} bytes. When a LUN is created in a thinly provisioned storage pool, its actual size and logical size are the multiple of 2^{34} bytes and are presented as an integer part of the decimal in GB.

3.5.3 Capacity reduction for an IBM i client

After a LUN is connected to an IBM i partition through the VIOS, the capacity of the LUN is reduced by about 11%.

The IBM i storage uses disks that support 520 bytes per sector. The IBM XIV Storage System server only supports 512 bytes per sector. Therefore, it is necessary to convert a 520 bytes per sector data layout to a 512 byte per sector when connecting XIV LUNs to an IBM i client.

Basically, to do the conversion, for every page in the IBM i client (8 x 512-byte sectors), an extra 512-byte sector is allocated. The extra sectors contain the information that was previously stored in the 8-byte sector headers. Because of this process, you must allocate nine sectors of XIV storage or midrange storage for every eight sectors in the IBM i client. From a capacity point of view, the usable capacity is multiplied by 8 divided by 9 or app.0.89, which means that it is reduced by about 11%, when the logical volumes report to an IBM i client.

In the XIV system, we define a LUN with as size of 150 GB. The capacity is rounded up to the next multiple of (2^{34}), which is the following result:

$$9 \times 2^{34} \text{ bytes} = 154618822656 \text{ bytes}$$

This capacity is presented in the XIV Storage Management GUI as 154 GB. After the LUN is connected to the IBM i client through the VIOS, its usable capacity for the IBM i client, is reduced in the following way:

$$(8 / 9) \times 154618822656 \text{ bytes} = 137438953472 \text{ bytes}$$

The capacity of such a LUN, as reported in IBM i system service tools (SST), is 137.438 GB.

In the following example, we calculate the usable capacity that is available to an IBM i client from a nine-module XIV system connected through the VIOS, using one regular storage pool and no snapshots, and defining the LUNs with a size of 100 GB.

In this example, the net capacity of a nine-module XIV system is 43.087 TB, as shown in Table 3-2 on page 45. From this capacity, we can define a regular storage pool with the following capacity:

$$2507 \times 2^{34} \text{ bytes} = 43.07 \text{ TB}$$

After creating the LUNs and connecting them to the IBM i client by using the VIOS, the capacity is reduced to 1 divided by 9. Therefore, the approximate usable capacity for the IBM i operating system is calculated as follows:

$$43.07 \times (8 / 9) \text{ TB} = 38.28 \text{ TB}$$



Implementing the IBM XIV Storage System server with IBM i

In this chapter, we use examples to illustrate the tasks for configuring the IBM XIV Storage System server for use with the IBM i operating system. In the examples, all of the IBM i disk space, including the load source disk (boot disk), consists of the XIV logical unit numbers (LUNs).

The LUNs are connected to the IBM i partition in multipath with two Virtual I/O Servers (VIOS). We explain this setup in 4.1, “Connecting a PowerVM client to the IBM XIV Storage System server” on page 48.

In addition, we describe the following tasks:

- ▶ Creating a storage pool on the XIV system
- ▶ Defining XIV volumes (LUNs) for the IBM i operating system and connecting them to both VIOS partitions
- ▶ Assigning the corresponding VIOS disk devices to virtual SCSI (VSCSI) adapters in the IBM i client partition
- ▶ Installing the IBM i operating system on XIV volumes

4.1 Connecting a PowerVM client to the IBM XIV Storage System server

The XIV system can be connected to an IBM i partition through the VIOS. In the following sections, we explain how to set up the environment on a POWER6 system to connect the XIV system to an IBM i client with multipath through two VIOS partitions.

4.1.1 Creating the Virtual I/O Server and IBM i partitions

In this section, we describe the steps to create a VIOS partition and an IBM i partition through the POWER6 Hardware Management Console (HMC). We also explain how to create VSCSI adapters in the VIOS and the IBM i partition and how to assign them so that the IBM i partition can work as a client of the VIOS.

For more information about how to create the VIOS and IBM i client partitions in the POWER6 server, see 6.2.1, “Creating the VIOS LPAR,” and 6.2.2, “Creating the IBM i LPAR,” in the IBM Redbooks publication *IBM i and Midrange External Storage*, SG24-7668.

Creating a Virtual I/O Server partition in a POWER6 server

To create a POWER6 logical partition (LPAR) for the VIOS:

1. Insert the PowerVM activation code in the HMC if you have not already done so. Select **Tasks** → **Capacity on Demand (CoD)** → **Advanced POWER Virtualization** → **Enter Activation Code**.
2. Create the new partition. In the left pane, select **Systems Management** → **Servers**. In the right pane, select the server to use for creating the new VIOS partition. Then select **Tasks** → **Configuration** → **Create Logical Partition** → **VIO Server** (Figure 4-1).

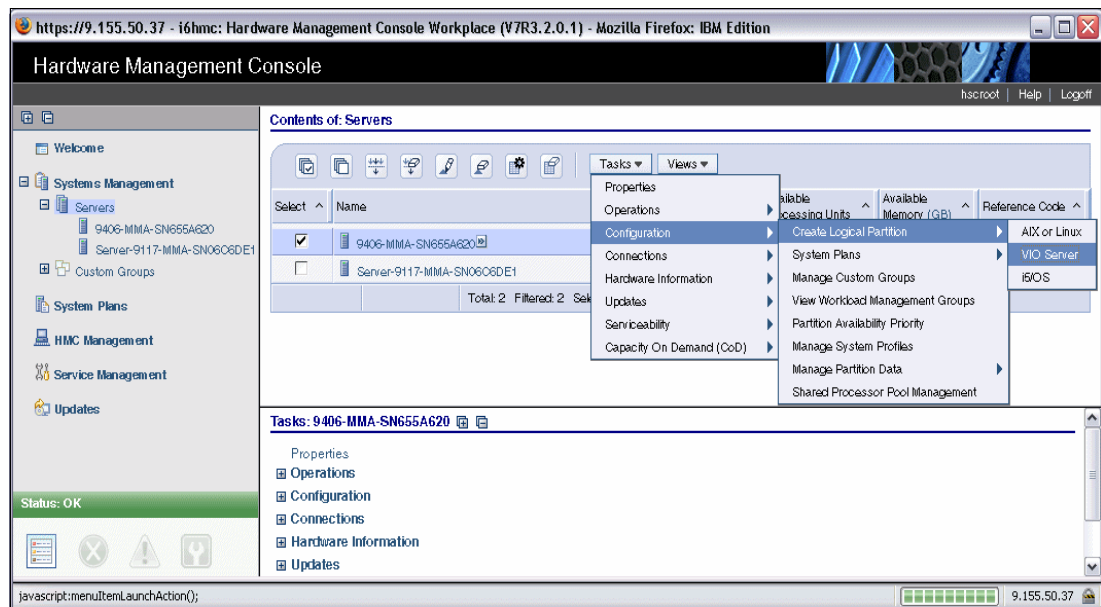


Figure 4-1 Creating the VIOS partition

3. In the Create LPAR wizard:
 - a. Type the partition ID and name.
 - b. Type the partition profile name.
 - c. Select whether the processors in the LPAR will be dedicated or shared. We recommend that you select **Dedicated**.
 - d. Specify the minimum, desired, and maximum number of processors for the partition.
 - e. Specify the minimum, desired, and maximum amount of memory in the partition.
4. In the I/O panel (Figure 4-2), select the I/O devices to include in the new LPAR. In our example, we include the RAID controller to attach the internal SAS drive for the VIOS boot disk and DVD_RAM drive. We include the physical Fibre Channel (FC) adapters to connect to the XIV server. As shown in Figure 4-2, we add them as *Required*.

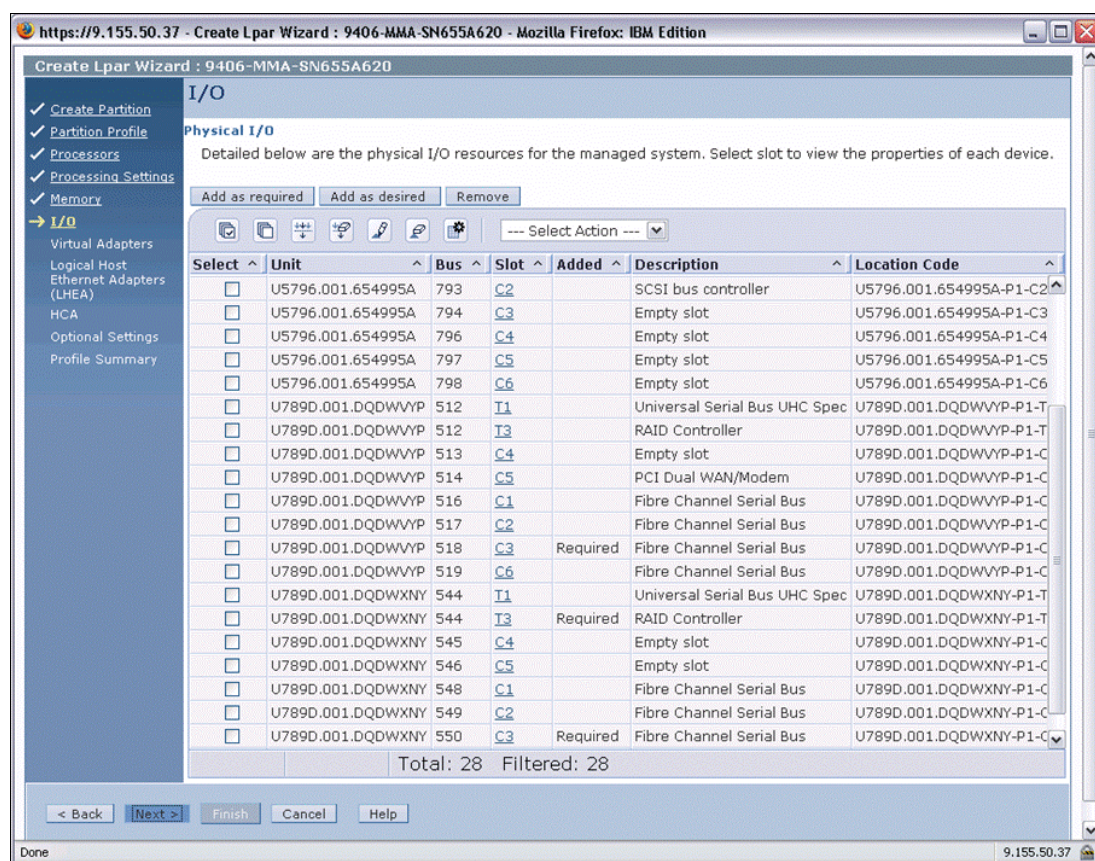


Figure 4-2 Adding the I/O devices to the VIOS partition

5. In the Virtual Adapters panel, create an Ethernet adapter by selecting **Actions** → **Create** → **Ethernet Adapter**. Mark it as **Required**.
6. Create the VSCSI adapters that will be assigned to the virtual adapters in the IBM i client:
 - a. Select **Actions** → **Create** → **SCSI Adapter**.
 - b. In the next window, either leave the **Any Client partition can connect** selected or limit the adapter to a particular client.

If DVD-RAM will be virtualized to the IBM i client, you might want to create another VSCSI adapter for DVD-RAM.

7. Configure the logical host Ethernet adapter:
 - a. Select the logical host Ethernet adapter from the list.
 - b. In the next window, click **Configure**.
 - c. Verify that the selected logical host Ethernet adapter is not selected by any other partitions, and select **Allow all VLAN IDs**.
8. In the Profile Summary panel, review the information, and click **Finish** to create the LPAR.

Creating an IBM i partition in the POWER6 server

To create an IBM i partition (that will be the VIOS client):

1. From the HMC, select **Systems Management** → **Servers**. In the right panel, select the server in which you want to create the partition. Then select **Tasks** → **Configuration** → **Create Logical Partition** → **i5/OS**.
2. In the Create LPAR wizard:
 - a. Insert the Partition ID and name.
 - b. Insert partition Profile name.
 - c. Select whether the processors in the LPAR will be dedicated or shared. We recommend that you select **Dedicated**.
 - d. Specify the minimum, desired, and maximum number of processors for the partition.
 - e. Specify the minimum, desired, and maximum amount of memory in the partition.
 - f. In the I/O panel, if the IBM i client partition is not supposed to own any physical I/O hardware, click **Next**.
3. In the Virtual Adapters panel, select **Actions** → **Create** → **Ethernet Adapter** to create a virtual Ethernet adapter.
4. In the Create Virtual Ethernet Adapter panel, accept the suggested adapter ID and the VLAN ID. Select **This adapter is required for partition activation** and click **OK** to continue.
5. Still in the Virtual Adapters panel, select **Actions** → **Create** → **SCSI Adapter** to create the VSCSI client adapters on the IBM i client partition that is used for connecting to the corresponding VIOS.
6. For the VSCSI client adapter ID, specify the ID of the adapter:
 - a. For the type of adapter, select **Client**.
 - b. Select **Mark this adapter is required for partition activation**.
 - c. Select the VIOS partition for the IBM i client.
 - d. Enter the server adapter ID to which you want to connect the client adapter.
 - e. Click **OK**.

If necessary, you can repeat this step to create another VSCSI client adapter to connect to the VIOS VSCSI server adapter that is used for virtualizing the DVD-RAM.
7. Configure the logical host Ethernet adapter:
 - a. Select the logical host Ethernet adapter from the drop-down list and click **Configure**.
 - b. In the next panel, ensure that no other partitions have selected the adapter and select **Allow all VLAN IDs**.
8. In the OptiConnect Settings panel, if OptiConnect is not used in IBM i, click **Next**.

9. In the Load Source Device panel, if the connected XIV system will be used to boot from a storage area network (SAN), select the virtual adapter that connects to the VIOS.

Note: The IBM i Load Source device resides on an XIV volume.

10. In the Alternate Restart Device panel, if the virtual DVD-RAM device will be used in the IBM i client, select the corresponding virtual adapter.
11. In the Console Selection panel, select the default of HMC for the console device. Click **OK**.
12. Depending on the planned configuration, click **Next** in the three panels that follow until you reach the Profile Summary panel.
13. In the Profile Summary panel, check the specified configuration and click **Finish** to create the IBM i LPAR.

4.1.2 Installing the Virtual I/O Server

In this section, we explain how to install the VIOS. For information about how to install the VIOS in a partition of the POWER6 server, see the Redbooks publication *IBM i and Midrange External Storage*, SG24-7668.

If the disks that you are going to use for the VIOS installation were previously used by an IBM i partition, you must reformat them for 512 bytes per sector.

Installing the Virtual I/O Server software

To install the VIOS software:

1. Insert the VIOS installation disk into the DVD drive of the VIOS partition.
2. Activate the VIOS partition:
 - a. In the left pane, select **Systems Management** → **Servers** and select the server.
 - b. Select the VIOS partition and then select **Operations** → **Activate** from the pop-up menu.
3. In the Activate Logical Partition window (Figure 4-3 on page 52):
 - a. Select the partition profile.
 - b. Select the **Open a terminal window or console session** check box.
 - c. Click **Advanced**.

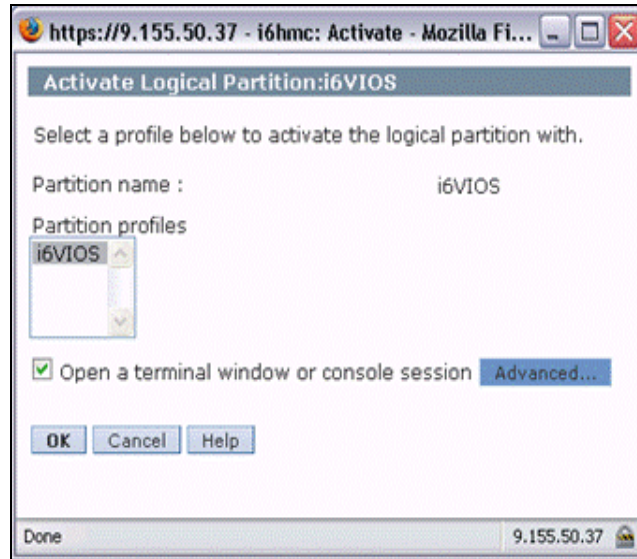


Figure 4-3 Activate LPAR window

4. In the next window, choose boot mode **SMS** and click **OK**. Then, click **OK** to activate the VIOS partition in SMS boot mode while a terminal session window is open.

Console window: Perform the following steps in the console window.

5. In the SMS main menu (Figure 4-4), select option **5. Select Boot Options**.

```
PowerPC Firmware
Version EM320_031
SMS 1.7 (c) Copyright IBM Corp. 2000,2007 All rights reserved.
-----
Main Menu
1.  Select Language
2.  Setup Remote IPL (Initial Program Load)
3.  Change SCSI Settings
4.  Select Console
5.  Select Boot Options
-----
Navigation Keys:

X = eXit System Management Services
-----
Type menu item number and press Enter or select Navigation key:5
```

Figure 4-4 SMS main menu

6. From the Multiboot menu, select option **1. Select Install/Boot Device**.
7. From the Select Device Type menu, select option **3. CD/DVD**.
8. From the Select Media Type menu, select option **5. SATA**.

9. From the Select Media Adapter menu, select the media adapter.
10. From the Select Device menu, select the device **SATA CD-ROM**.
11. From the Select Task menu, select option **2. Normal boot mode**.
12. In the next console window, when prompted by the question “Are you sure you want to exit System Management Services?”, select option **1. Yes**.

The VIOS partition is automatically rebooted in normal boot mode.

13. In the VIOS installation panel that opens, at the prompt for the system console, type 1 to use this terminal as the system console.
14. At the installation language prompt, type 1 to use English as the installation language.
15. In the VIOS welcome panel, select option **1 Start Install Now with Default Settings**.
16. From the VIOS System Backup Installation and Settings menu, select option **0 Install with the settings listed above**.

Installation of the VIOS starts, and its progress is shown in the Installing Base Operation System panel.

17. After successful installation of the VIOS, log in with the prime administrator user ID padmin. Enter a new password and type a to accept the software terms and conditions.
18. Before you run any VIOS command other than the **chlang** command to change the language setting, accept the software license terms by entering the following command:

```
license -accept
```

To view the license before you accept it, enter the following command:

```
license -view
```

Using LVM mirroring for the Virtual I/O Server

Set up LVM mirroring to mirror the VIOS root volume group (rootvg). In our example, we mirror it across two RAID0 arrays of hdisk0 and hdisk1 to help protect the VIOS from potential CEC internal SAS disk drive failures.

Configuring Virtual I/O Server network connectivity

To set up network connectivity in the VIOS:

1. In the HMC terminal window, logged in as padmin, enter the following command:

```
lsdev -type adapter | grep ent
```

Look for the logical host Ethernet adapter resources. In our example, it is ent1, as shown in Figure 4-5.

```
$ lsdev -type adapter | grep ent
ent0          Available   Virtual I/O Ethernet Adapter (1-lan)
ent1         Available   Logical Host Ethernet Port (1p-hea)
```

Figure 4-5 Available logical host Ethernet port

2. Configure TCP/IP for the logical Ethernet adapter entX by using the **mktcpip** command syntax and specifying the corresponding interface resource enX.
3. Verify the created TCP/IP connection by pinging the IP address that you specified in the **mktcpip** command.

Upgrading the Virtual I/O Server to the latest fix pack

As the last step of the installation, upgrade the VIOS to the latest fix pack.

4.1.3 IBM i multipath capability with two Virtual I/O Servers

The IBM i operating system provides multipath capability for allowing access to an XIV volume (LUN) through multiple connections. One path is established through each connection. Up to eight paths to the same LUN or set of LUNs are supported. Multipath provides redundancy in case a connection fails, and it increases performance by using all available paths for I/O operations to the LUNs.

With Virtual I/O Server release 2.1.2 or later, and IBM i release 6.1.1 or later, it is possible to establish multipath to a set of LUNs, with each path using a connection through a different VIOS. This topology provides redundancy in case either a connection or the VIOS fails. Up to eight multipath connections can be implemented to the same set of LUNs, each through a different VIOS. However, we expect that most IT centers will establish no more than two such connections.

4.1.4 Connecting with virtual SCSI adapters in multipath with two Virtual I/O Servers

In our setup, we use two VIOS and two VSCSI adapters in the IBM i partition, where each adapter is assigned to a virtual adapter in one VIOS. We connect the same set of XIV LUNs to each VIOS through two physical FC adapters in the VIOS multipath and map them to VSCSI adapters serving IBM i partition. This way, the IBM i partition sees the LUNs through two paths, each path by using one VIOS. Therefore, multipathing is started for the LUNs. Figure 4-6 on page 55 shows our setup.

For our testing, we did not use separate switches as shown in Figure 4-6, but rather used separate blades in the same SAN Director. In a real, production environment, use separate switches as shown in Figure 4-6.

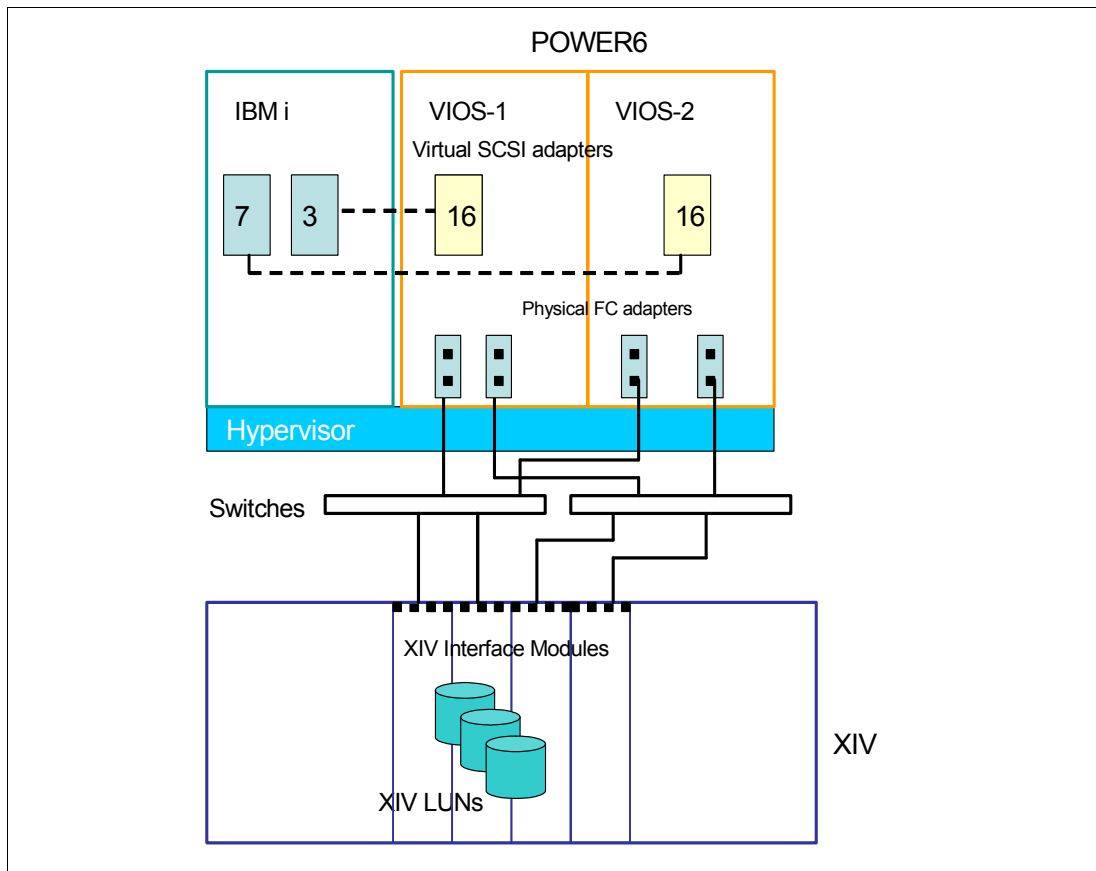


Figure 4-6 Setup for multipath with two VIOS

To connect XIV LUNs to an IBM i client partition in multipath with two VIOS:

Important: Perform steps 1 through 5 in each of the two VIOS partitions.

1. After the LUNs are created in the XIV system, use the XIV Storage Management GUI or Extended Command Line Interface (XCLI) to map the LUNs to the VIOS host as shown in 4.2, “Configuring XIV storage to connect to IBM i by using the Virtual I/O Server” on page 56.
2. Log in to VIOS as administrator. In our example, we use PUTTY to log in as described in 6.5, “Configuring VIOS virtual devices,” of the Redbooks publication *IBM i and Midrange External Storage*, SG24-7668.

Type the **cfgdev** command so that the VIOS can recognize newly attached LUNs.

3. In the VIOS, remove the SCSI reservation attribute from the LUNs (hdisks) that will be connected through two VIOS by entering the following command for each hdisk that will connect to the IBM i operating system in multipath:

```
chdev -dev hdiskX -attr reserve_policy=no_reserve
```

4. Set the attributes of Fibre Channel adapters in the VIOS to `fc_err_recov=fast_fail` and `dyntrk=yes`. When the attributes are set to these values, the error handling in FC adapter allows faster transfer to the alternate paths in case of problems with one FC path. To make multipath within one VIOS work more efficiently, specify these values by entering the following command:

```
chdev -dev fscsi0 -attr fc_err_recov=fast_fail dyntrk=yes-perm
```

5. To get more bandwidth by using multiple paths, enter the following command for each hdisk (hdiskX):

```
chdev -dev hdiskX -perm -attr algorithm=round_robin;
```
6. Map the disks that correspond to the XIV LUNs to the VSCSI adapters that are assigned to the IBM i client. First, check the IDs of assigned virtual adapters. Then complete the following steps:
 - a. In the HMC, open the partition profile of the IBM i LPAR, click the **Virtual Adapters** tab, and observe the corresponding VSCSI adapters in the VIOS.
 - b. in the VIOS, look for the device name of the virtual adapter that is connected to the IBM i client. You can use the command **lsmap -all** to view the virtual adapters.
 - c. Map the disk devices to the SCSI virtual adapter that is assigned to the SCSI virtual adapter in the IBM i partition by entering the following command:

```
mkvdev -vdev hdiskxx -vadapter vhostx
```

Upon completing these steps, in each VIOS partition, the XIV LUNs report in the IBM i client partition by using two paths. The resource name of disk unit that represents the XIV LUN starts with DMPxxx, which indicates that the LUN is connected in multipath.

4.2 Configuring XIV storage to connect to IBM i by using the Virtual I/O Server

In this section, we explain how to create an XIV storage pool, define LUNs, and map those LUNs to the two VIOS partitions to which the IBM i client is connected. While we use the XIV Storage Management GUI to perform these actions, you can also use the XIV Storage Management XCLI to complete them. For more information about the XIV Storage management GUI and XCLI, see 1.5, “IBM XIV Storage Management software” on page 23, and the Redbooks publication *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659.

4.2.1 Creating a storage pool

To create an IBM XIV Storage System storage pool:

1. Set up the XIV Storage Management GUI and connect to the XIV system as explained in 1.5.1, “The XIV Storage Management GUI” on page 25.
2. In the main GUI, click the **Pools** icon and select **Storage Pools** from the option menu as shown in Figure 4-7.



Figure 4-7 Selecting the Storage Pools option

3. In the Storage Pools window (Figure 4-8), click the **Add Pool** button from the toolbar.

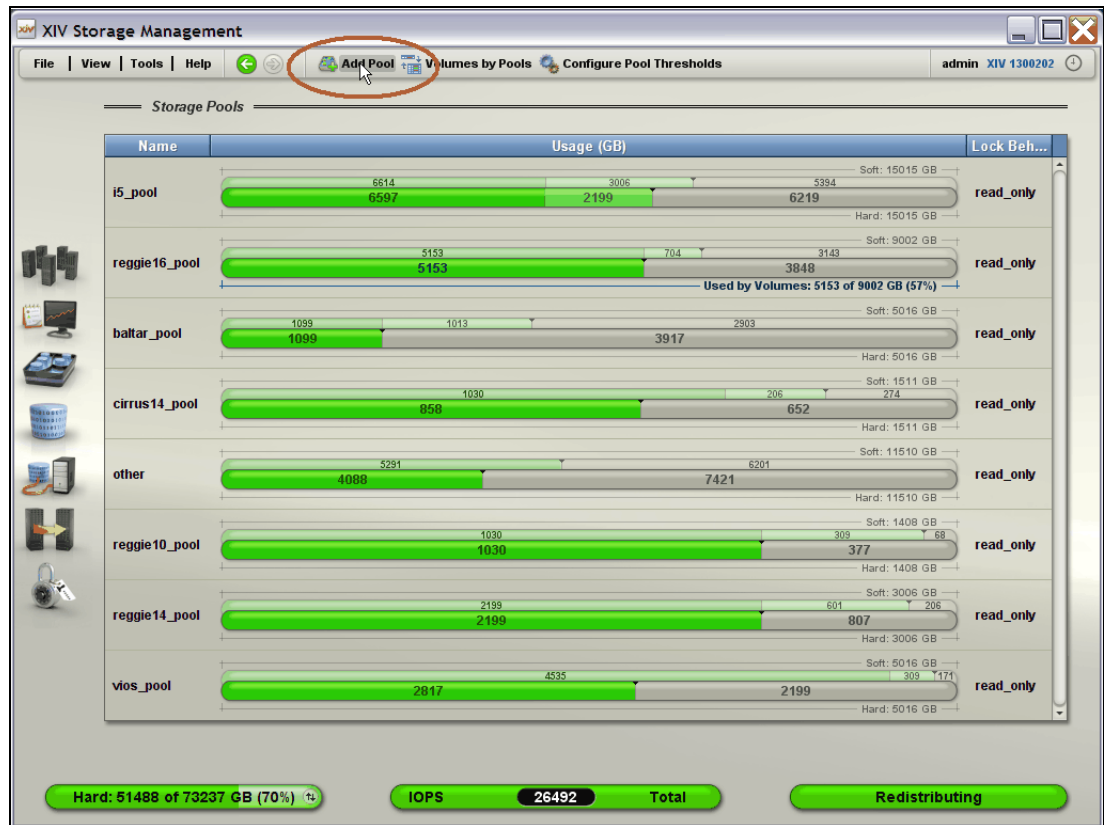


Figure 4-8 Adding a storage pool

4. In the Add Pool panel (Figure 4-9), specify the type of storage pool being created (Regular or Thinly provisioned), its size, the size to reserve for snapshots in the pool, and the name of the pool. Then click **Add**.

Select Type: Regular Pool

Total System Capacity: 73237 GB

Allocated: 51488 GB | Pool Size: 3006 GB | Free: 18743 GB

Pool Size: 3006 GB

Snapshots Size: 1511 GB

Pool Name: * IBM I Pool

Add Cancel

Figure 4-9 Defining the new storage pool

Although we specified 3000 GB for Pool Size, the pool that is created has a size of 3006 GB, because the size of a storage pool in the XIV system must be a multiple of 2^{34} bytes. Therefore, the specified 3000 GB was rounded up to the next multiple of 2^{34} bytes, of which the integer part, 3006 GB, is shown in the GUI. For more information about the capacity of storage pools, see 3.5, “Planning for capacity” on page 44.

The newly created storage pool is highlighted in the Storage Pools GUI window (Figure 4-10).

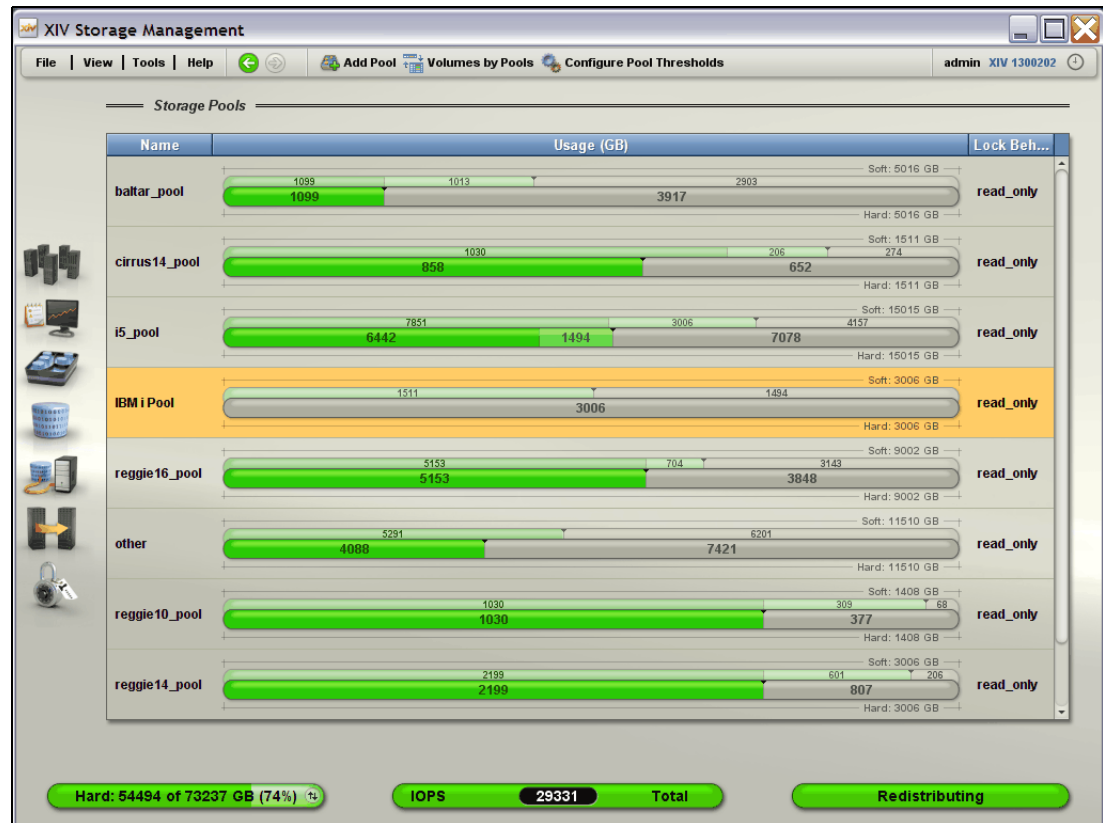


Figure 4-10 New storage pool

4.2.2 Defining the volumes

The next step is to create volumes in the IBM XIV Storage System storage pool.

Note: In the XIV system, we create the same type of volumes for all the host servers, except for the Hewlett-Packard servers.

1. In the Storage Pools window (Figure 4-11), right-click the storage pool in which you want to create volumes, and select **Volumes and Snapshots**.

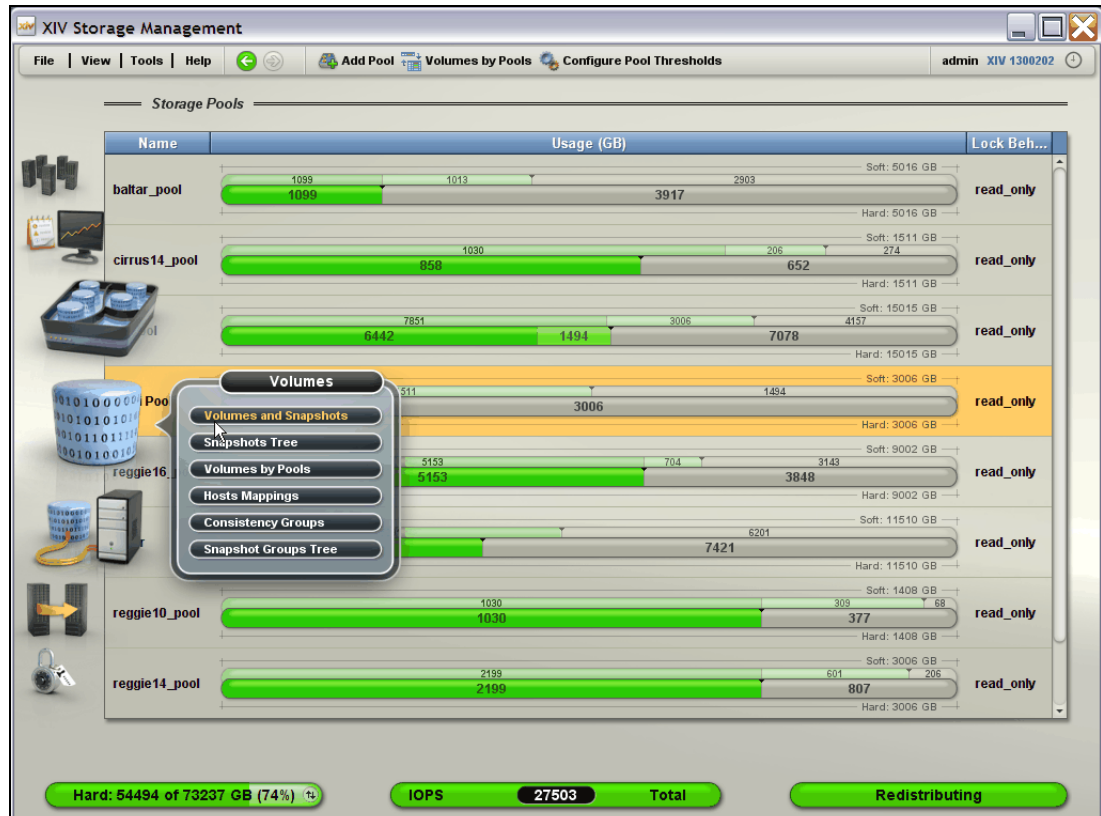


Figure 4-11 Storage Pools window

2. In the Volumes and Snapshots window (Figure 4-12), which lists all the volumes that are presently defined in the XIV system, from the toolbar, click **Add Volumes**.

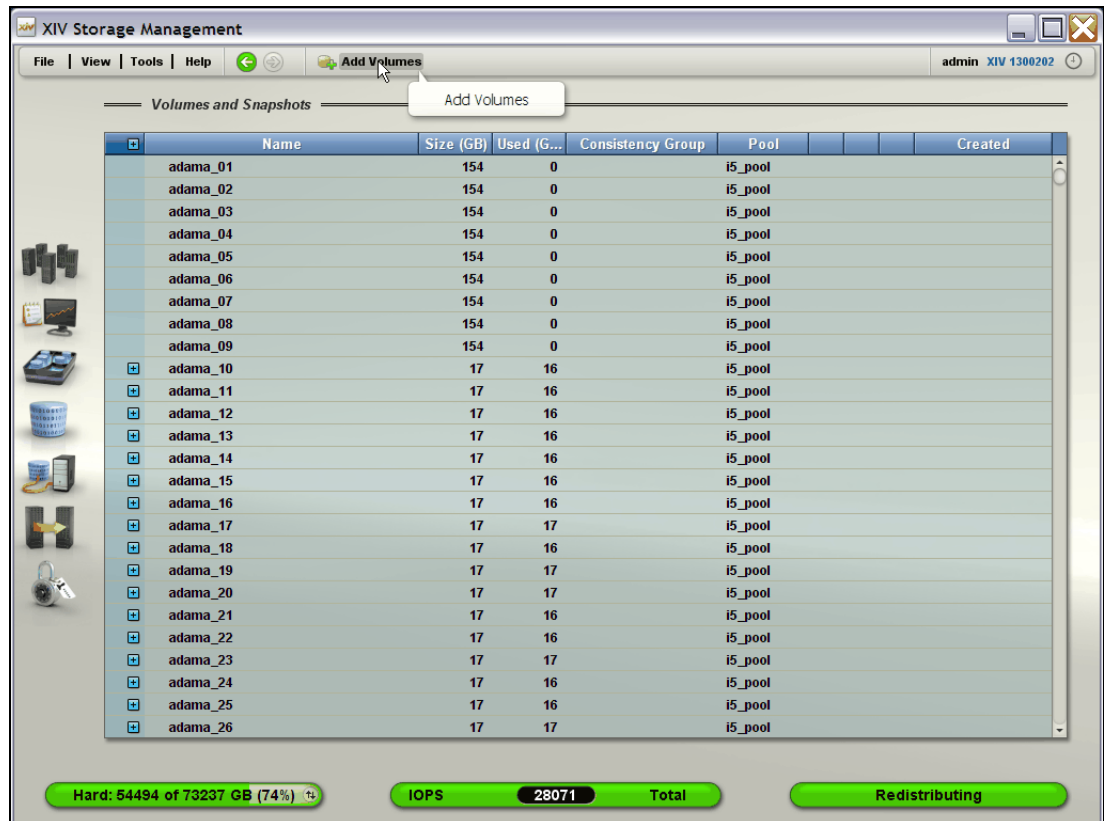


Figure 4-12 Volumes and Snapshots window

3. In the Create Volumes window (Figure 4-13), make sure that the correct storage pool is selected and specify the number of volumes to create, their size, and a name. Click **Create**. When creating multiple volumes, the corresponding suffix is automatically appended at the end of the specified volume name.

In our example, we created nine new LUNs, with a size of 150 GB, in the storage pool IBM i Pool as shown in Figure 4-13. The specified size of 150 GB was rounded up to the next multiple of 2^{34} bytes, of which the integer part, 154 GB, is shown in the GUI. For more information about LUN capacity, see 3.5, “Planning for capacity” on page 44.

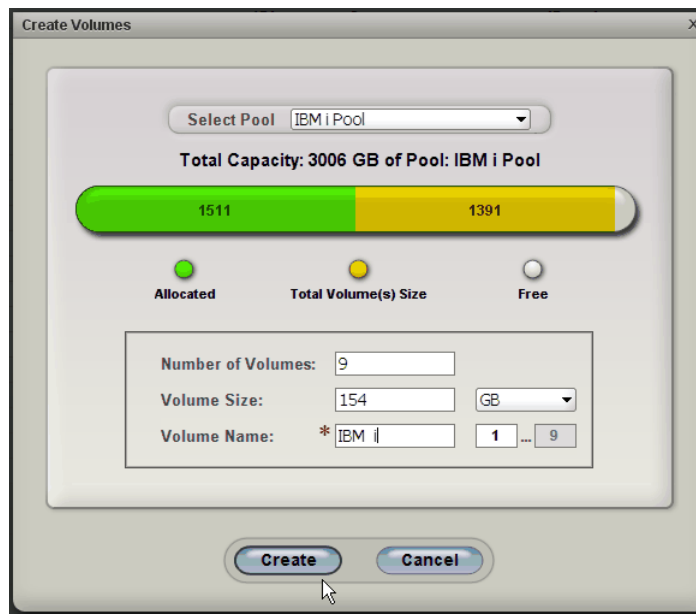


Figure 4-13 Creating new volumes

While the LUNs are being created, the XIV Storage Management GUI shows a progress indicator (Figure 4-14).

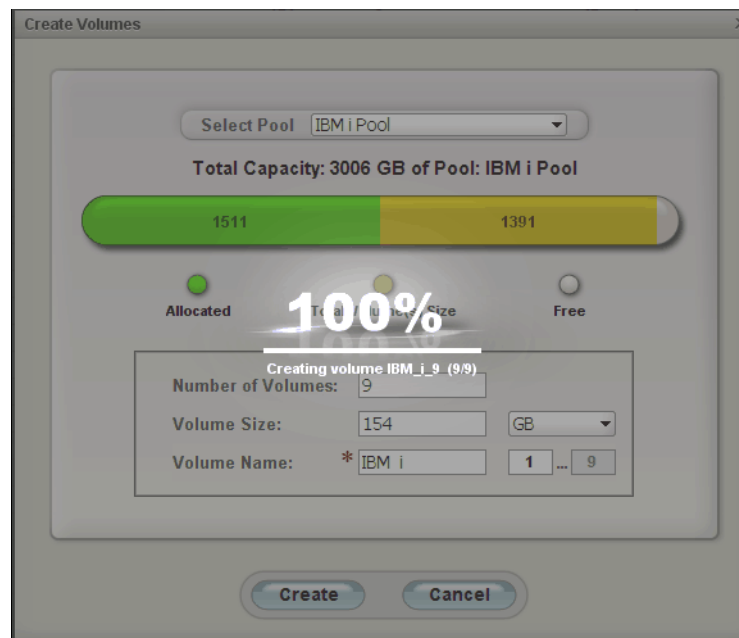


Figure 4-14 Progress of creating the volumes

After the volumes are all created, they are listed and highlighted in the Volumes and Snapshots window (Figure 4-15).

Tools | Help | Add Volumes admin XIV 1300202

Volumes and Snapshots

	Name	Size (GB)	Used (G...	Consistency Group	Pool				Created
	hoopla_20	34	0		other				
	IBM_i_1	154	0		IBM i Pool				
	IBM_i_2	154	0		IBM i Pool				
	IBM_i_3	154	0		IBM i Pool				
	IBM_i_4	154	0		IBM i Pool				
	IBM_i_5	154	0		IBM i Pool				
	IBM_i_6	154	0		IBM i Pool				
	IBM_i_7	154	0		IBM i Pool				
	IBM_i_8	154	0		IBM i Pool				
	IBM_i_9	154	0		IBM i Pool				
	linuxd_01	103	76		other				
	linuxd_02	103	103		other				
	linuxd_03	103	76		other				
	linuxd_04	103	103		other				
	linuxd_05	103	75		other				
	linuxd_06	103	103		other				
	linuxd_07	103	75		other				
	linuxd_08	103	103		other				
	linuxd_09	103	75		other				
	linuxd_10	103	103		other				
	linuxd_11	103	75		other				
	linuxd_12	103	103		other				
	linuxd_13	103	75		other				
	linuxd_14	103	103		other				
	linuxd_15	103	75		other				
	linuxd_16	103	103		other				

54494 of 73237 GB (74%) IOPS 26984 Total Redistributing

Figure 4-15 Listed new LUNs

4.2.3 Connecting the volumes to the Virtual I/O Server

To connect the volumes to the VIOS partitions:

1. In the XIV Storage Management GUI (Figure 4-16), click the **Hosts and Clusters** icon and select **Hosts Connectivity** from the option menu.



Figure 4-16 Selecting Host Connectivity

2. In the Host Connectivity window (Figure 4-17), select the VIOS partition to which you want to connect the XIV volumes (LUNs), right-click, and select **Modify LUN Mapping**.

Note: Selecting the VIOS partition (not a particular worldwide port name (WWPN) in the VIOS) assigns LUNs to both WWPNs of the FC adapters that are available in the VIOS.

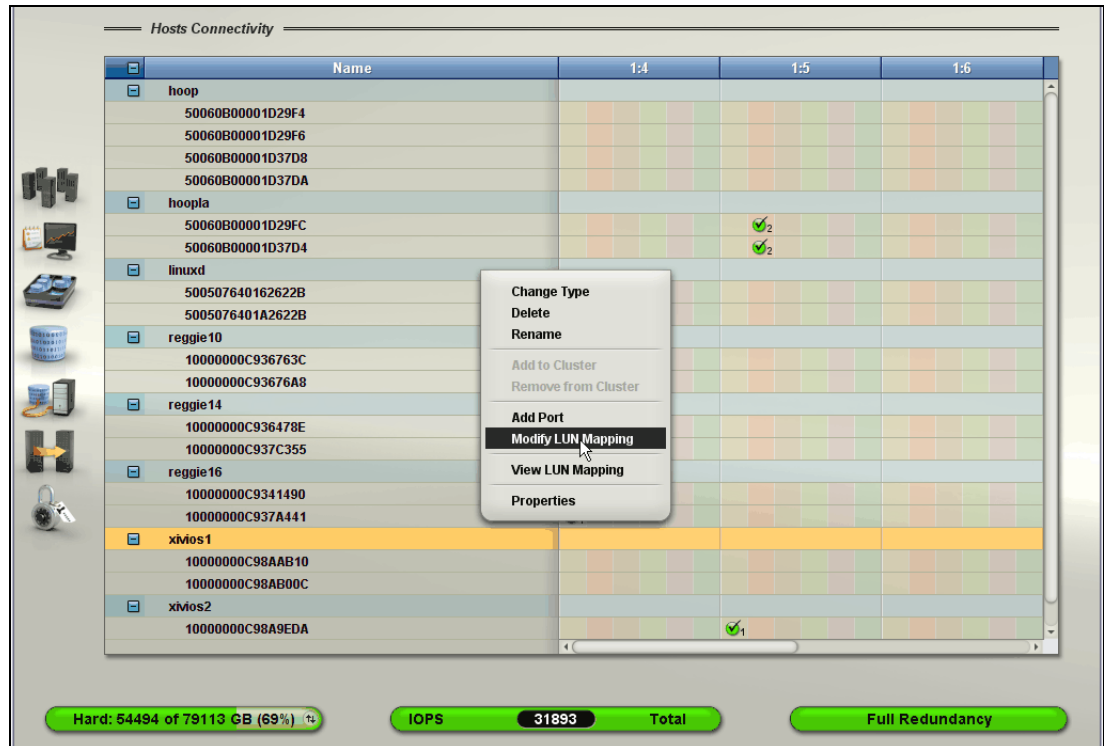


Figure 4-17 Selecting the Modify LUN Mapping option

3. In the Volume to LUN Mapping window (Figure 4-18), which shows two panes, in the left pane, select the LUN that you want to add to the host, and then click **Map** to map the LUN to the VIOS. The LUN that was just mapped is added to the list shown in the right pane.

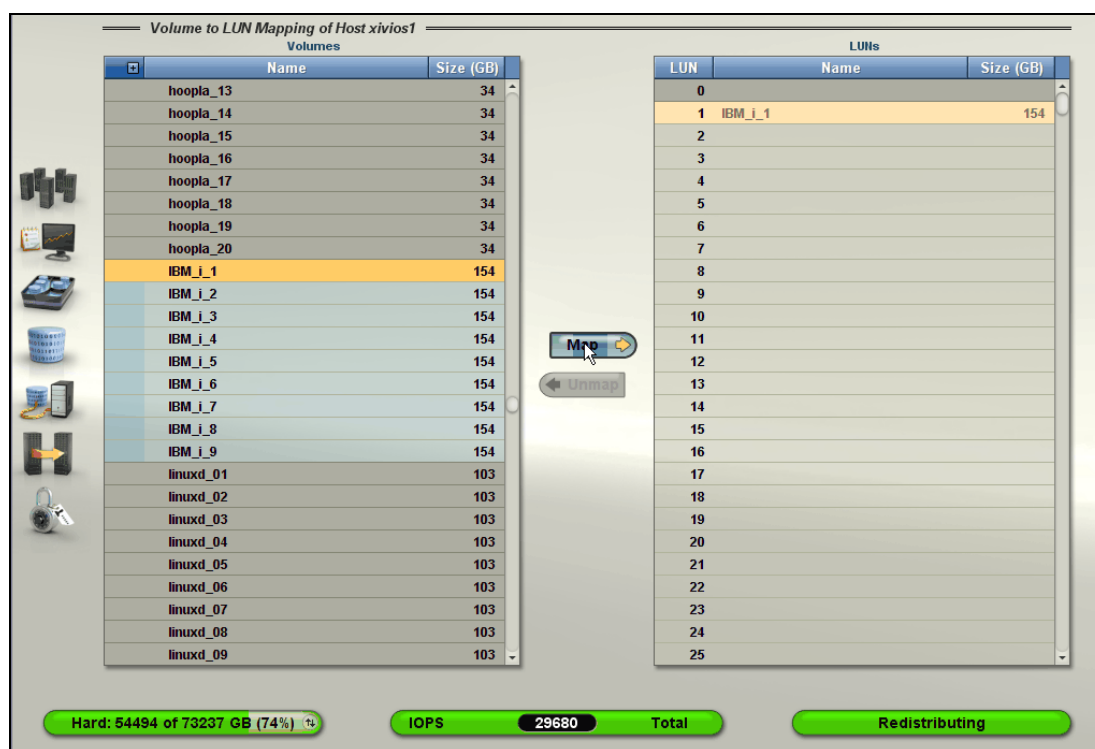


Figure 4-18 Mapping LUNs to the VIOS

4. To establish IBM i multipathing using two VIOS partitions, map the same LUNs to the second VIOS. Therefore, perform steps 2 on page 65 and 3 on page 66 for the other VIOS partition.

While adding the LUNs to the second VIOS partition, you see the warning message that the LUNs are already mapped to another VIOS (Figure 4-19). Click **OK** to close the message.

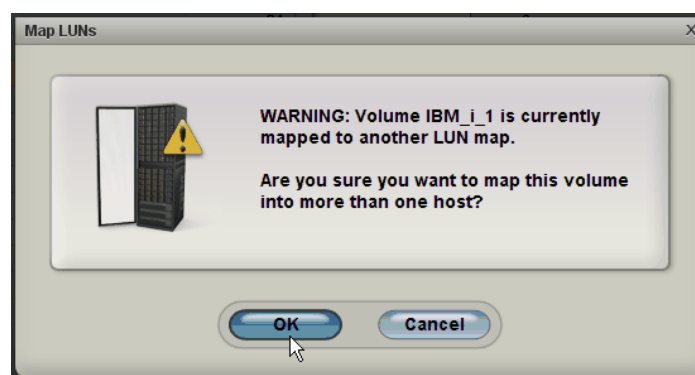


Figure 4-19 Warning message when adding LUNs to the second VIOS

In our example, we add only one LUN to both VIOS partitions. This LUN is the IBM i Load Source disk unit.

In each VIOS, we must map this LUN to the VSCSI adapter that is assigned to the IBM i client and start IBM i installation. After installing the IBM i Licensed Internal Code, we can add the other LUNs to both VIOS, map them for each VIOS to the VSCSI adapters for the IBM i client, and finally add them to IBM i in Dedicated Service Tools (DST) environment.

4.3 Mapping the volumes in the Virtual I/O Server

The XIV volumes are now added to both VIOS partitions. To make them available for the IBM i client, perform the following tasks in each VIOS:

1. Connect to the VIOS partition. In our example, we use a PUTTY session to connect.
2. in the VIOS, enter the `cfgdev` command to discover the newly added XIV LUNs, which makes the LUNs available as disk devices (hdisks) in the VIOS. In our example, the LUNs that we added correspond to hdisk132 - hdisk140, as shown in Figure 4-20.

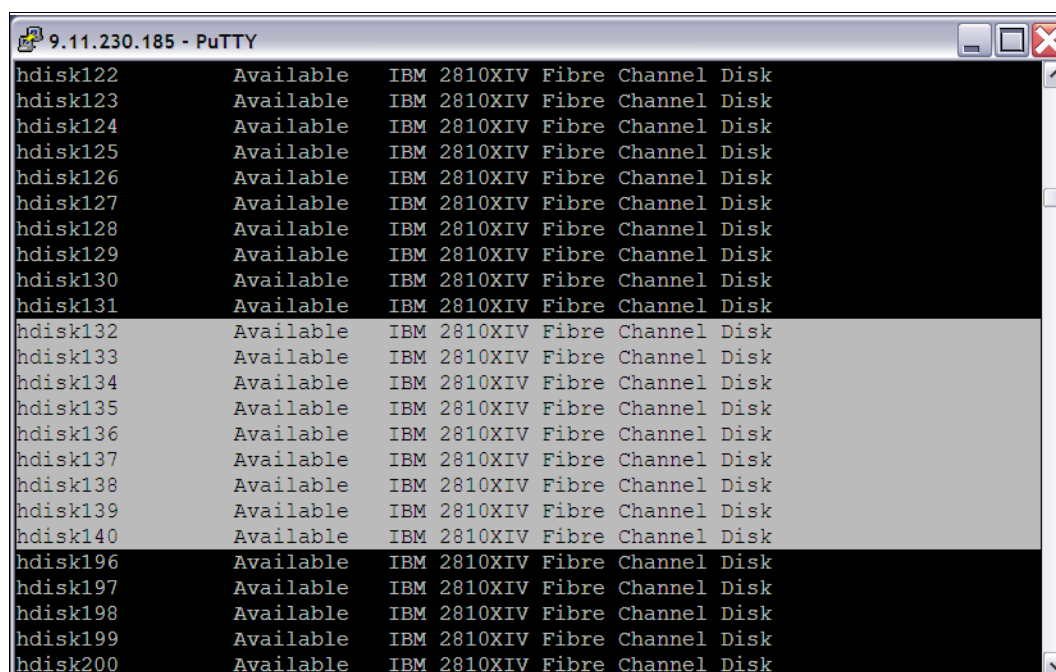


Figure 4-20 The hdisks of the added XIV volumes

As we previously explained, to realize a multipath setup for IBM i, we connected (mapped) each XIV LUN to both VIOS partitions. Before assigning these LUNs (from any of the VIOS partitions) to the IBM i client, make sure that the volume is not SCSI reserved.

3. Because a SCSI reservation is the default in the VIOS, change the reservation attribute of the LUNs to *non-reserved*. First, check the current reserve policy by entering the following command:

```
lsdev -dev hdiskx -attr reserve_policy
```

Here *hdiskx* represents the XIV LUN.

If the reserve policy is not *no_reserve*, change it to *no_reserve* by entering the following command:

```
chdev -dev hdiskX -attr reserve_policy=no_reserve
```

4. Before mapping hdisks to a VSCSI adapter, check whether the adapter is assigned to the client VSCSI adapter in IBM i and whether any other devices are mapped to it.

- a. Enter the following command to display the virtual slot of the adapter and see any other devices assigned to it:

```
lsmmap -vadapter <name>
```

In our setup, no other devices are assigned to the adapter, and the relevant slot is C16 (Figure 4-21).

```
$ lsmmap -vadapter vhost5
SVSA          Physloc          C1
-----
vhost5        U9117.MMA.10AF384-V2-C16      0
VTD           NO VIRTUAL TARGET DEVICE FOUND
```

Figure 4-21 Virtual SCSI adapter in the VIOS

- b. From the HMC, edit the profile of the IBM i partition. Select the partition and choose **Configuration** → **Manage Profiles**. Then select the profile and click **Actions** → **Edit**.
- c. In the partition profile, click the **Virtual Adapters** tab and make sure that a client VSCSI adapter is assigned to the server adapter with the same ID as the virtual slot number. In our example, client adapter 3 is assigned to server adapter 16 (thus matching the virtual slot C16) as shown in Figure 4-22.

Processors

Memory

I/O

Virtual Adapters

Power Controlling

Settings

Logical Host Ethernet Adapters (LHEA)

Tagged I/O

S

resources allow for the sharing of physical hardware between logical partitions. The adapter settings are listed below.

Number of virtual adapters : 15

Number of virtual adapters : 12



--- Select Action ---

Type	Adapter ID	Connecting Partition	Connecting Adapter	Rec
Client SCSI	10	xivios-2(3)	19	No
Client SCSI	11	xivios-1(2)	15	Yes
Client SCSI	2	xivios-1(2)	11	Yes
Client SCSI	3	xivios-1(2)	16	Yes
Client SCSI	4	xivios-1(2)	17	No
Client SCSI	5	xivios-1(2)	18	No
Client SCSI	6	xivios-1(2)	19	No
Client SCSI	7	xivios-2(3)	16	Yes
Client SCSI	8	xivios-2(3)	17	No
Client SCSI	9	xivios-2(3)	18	No
Server Serial	0	Any Partition	Any Partition Slot	Yes

Figure 4-22 Assigned virtual adapters

5. Map the relevant hdisks to the VSCSI adapter by entering the following command:

```
mkvdev -vdev hdiskx -vadapter <name>
```

In our example, we map the XIV LUNs to the adapter vhost5, and we give to each LUN the virtual device name by using the -dev parameter as shown in Figure 4-23.

```
$ mkvdev -vdev hdisk132 -vadapter vhost5 -dev vadamaboot132
vadamaboot132 Available
```

Figure 4-23 Mapping the LUNs in the VIOS

After completing these steps for each VIOS, the LUNs are available to the IBM i client in multipath (one path through each VIOS).

4.3.1 Using the HMC to map volumes to an IBM i client

As an alternative to mapping volumes by using the VIOS commands, you can assign the volumes to the IBM i partition from the POWER6 HMC. This function is available on HMC level EM340-095 from May 2009 or later.

To map the volumes to an IBM i client:

1. Connect to VIOS, discover volumes, and change the reserve_policy attribute of the LUNs. Note that you *cannot* use HMC to do this.
2. In the HMC window, expand **Systems Management** and click **Servers**. In the right pane, select the server on which the VIOS and IBM i client partitions reside, expand the pop-up menu at the server, and choose **Configuration** → **Virtual Resources** → **Virtual Storage Management** (Figure 4-24).

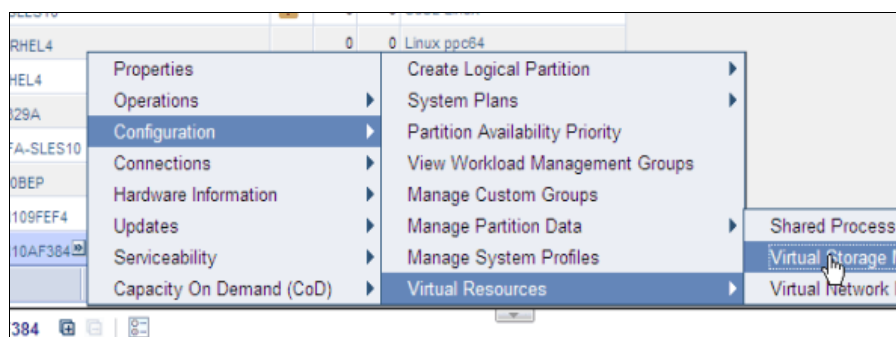


Figure 4-24 HMC: Selecting Virtual Storage Management for the selected server

3. In the Virtual Storage Management window (Figure 4-25), select the VIOS to which to assign volumes. Click **Query VIOS**.

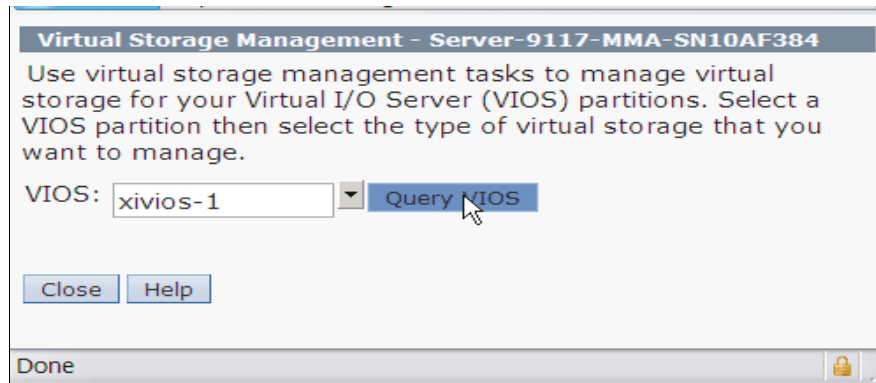


Figure 4-25 HMC: Clicking Query VIOS

4. After the Query VIOS results are displayed, click the **Physical Volumes** tab to view the disk devices in the VIOS (Figure 4-26).
 - a. Select the hdisk that you want to assign to the IBM i client and click **Modify assignment**.

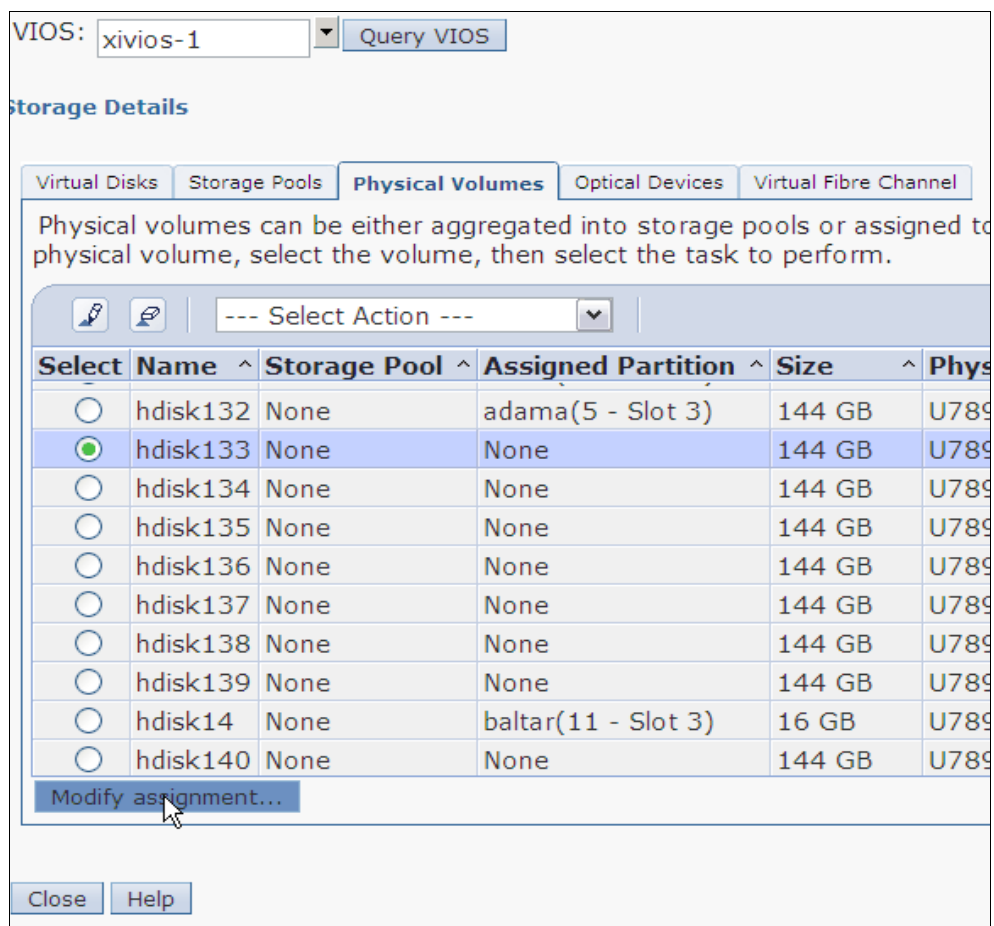


Figure 4-26 HMC: Modifying the assignment of physical volumes in the VIOS

- b. In the Modify Physical Disk Partition Assignments window (Figure 4-27), from the pull-down list, choose the IBM i client to which you want to assign the volumes.

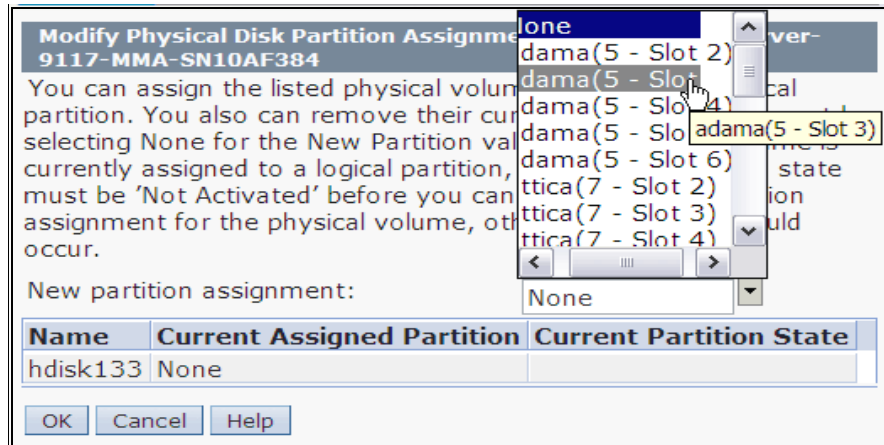


Figure 4-27 HMC: Selecting the partition to which to assign LUNs

- c. Click **OK** to confirm the selection for the selected partition (Figure 4-28).

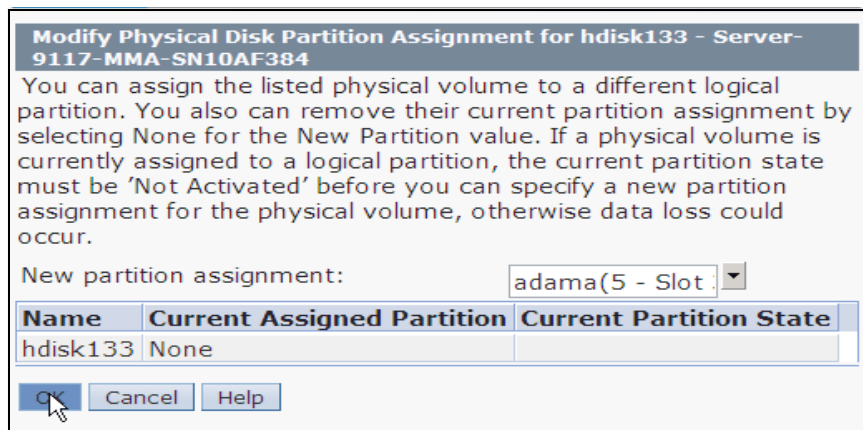


Figure 4-28 HMC: Confirming to assign to the selected partition

You now see the information message indicating that the hdisk is being reassigned (Figure 4-29). The Query VIOS function starts automatically again.

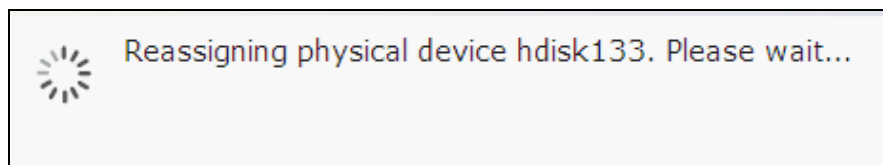


Figure 4-29 HMC: Re-assigning the physical device

The newly assigned volumes are displayed in the query results.

Important: To have multipath to the IBM i client LUNs through two VIOS, you must perform steps 1 through 4 for each VIOS.

4.4 Installing the IBM i client

To install the IBM i client:

1. Insert the installation media *Base_01 6.1.1* in the DVD drive of the client IBM i partition.

Installing IBM i: To install IBM i, you can use the DVD drive that is dedicated to the IBM i client partition or the virtual DVD drive that is assigned in the VIOS. If you are using the virtual DVD drive, insert the installation media in the corresponding physical drive in the VIOS.

2. In the IBM i partition, make sure that the tagged adapter for load source disk points to the adapter in the VIOS to which the XIV LUNs are assigned. Also, ensure that the tagged adapter for Alternate restart device points to the server VSCSI adapter with the assigned optical drive or to the physical adapter with DVD-RAM.

To check the tagged adapter:

- a. Select the partition in HMC, and select **Configuration** → **Manage Profiles** from the pull-down menu.
- b. Select the profile and select **Actions** → **Edit**.
- c. In the partition profile, click the **Tagged I/O** tab.

As shown in Figure 4-30, we use client adapter 3 for the load source and client adapter 11 for the alternate installation device.

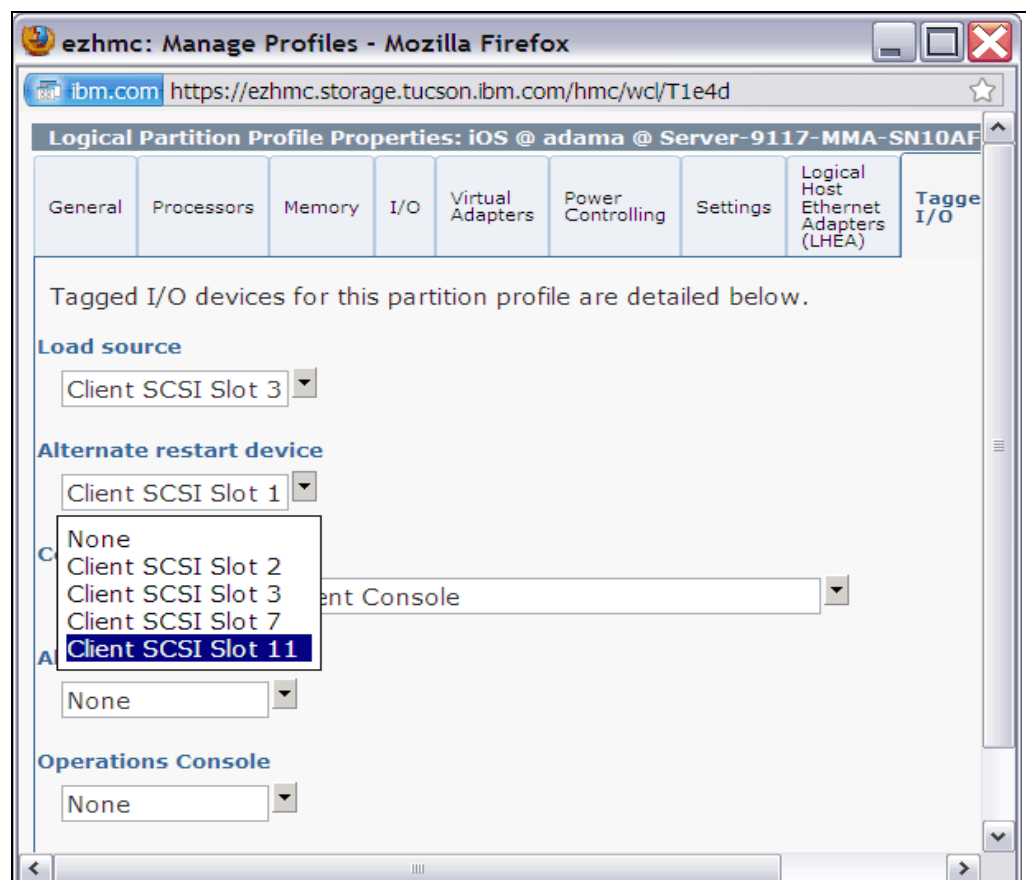


Figure 4-30 Tagged adapters in the IBM i client

- d. Still in the partition profile, click the **Virtual Adapters** tab, and verify that the corresponding server virtual adapters are the ones with the assigned volumes and DVD drive. See Figure 4-22 on page 68.
3. In the IBM i client partition, make sure that IPL source is set to **D** and that Keylock position is set to **Manual**.

To verify these settings, select the IBM i client partition in HMC, choose **Properties**, and click the **Settings** tab to see the currently selected values. Figure 4-31 shows the settings in the client partition used for this example.

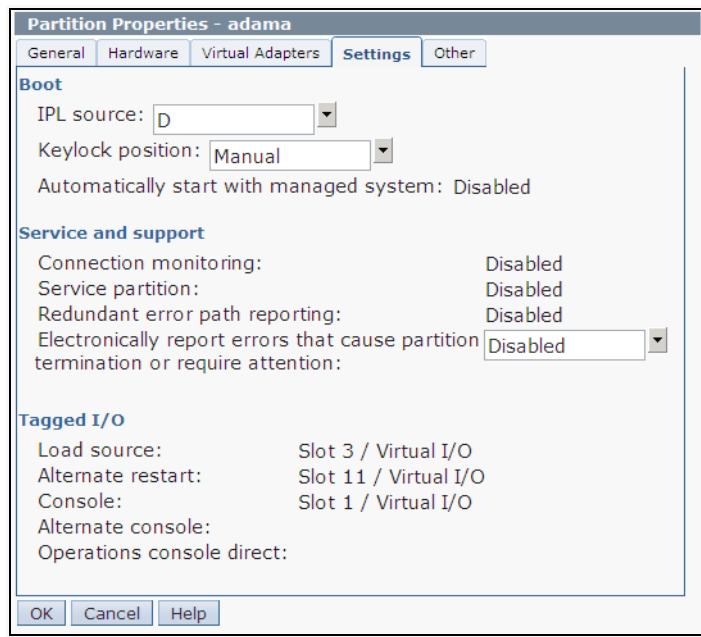


Figure 4-31 Settings in the client partition

4. To activate the IBM i client partition, select this partition in the HMC and select **Activate** from the pull-down menu. You can open the console window in the HMC.

Alternatively with the IBM Personal Communications tool, you can use Telnet to connect to HMC port 2300 (Figure 4-32).

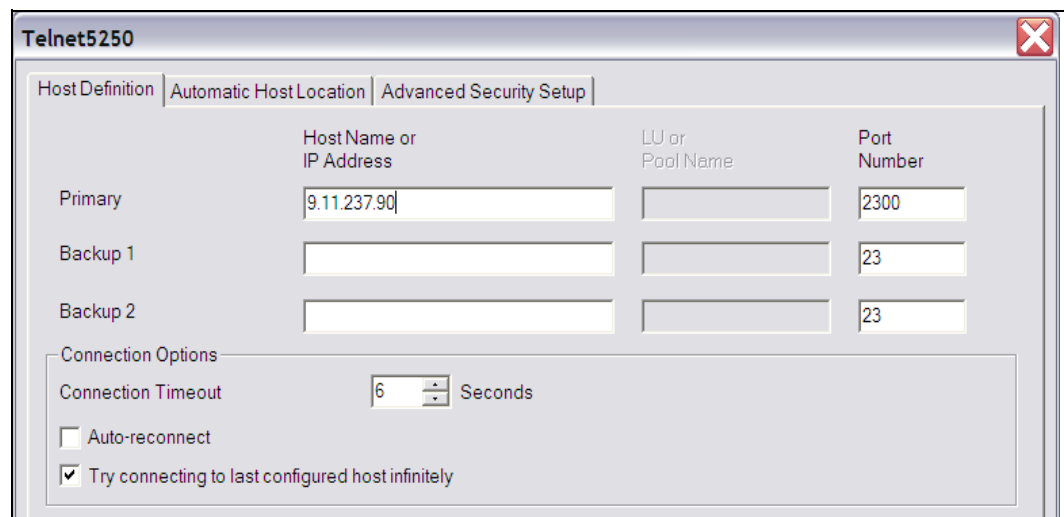


Figure 4-32 Using Telnet to connect to the HMC

Select the appropriate language, server, and partition (Figure 4-33).

```

HMC Remote 5250 Console System Selection
HMC: ezhmc

Select one of the following and press Enter:

Option  System Name      Type      Model      Serial      State
  1      Fibonacci-SN1003A 9123      710        1003A0A     Operating
  2      Cauchy-SN10020EA- 9123      710        10020EA     Operating
  3      Newton-SN10D95AC- 9123      710        10D95AC     Operating
  4      Dirac-710-SN10D95 9123      710        10D957C     Operating
  5      Bessel-SN100212A- 9123      710        100212A     Operating
  6      Fourier-SN06EAB2B 9123      710        06EAB2B     Operating
  7      Server-9117-MMA-S 9117      MMA        109FEF4     Operating
  8      Cramer-SN10D958C- 9123      710        10D958C     Operating
  9      Server-9117-MMA-S 9117      MMA        10AF384     Operating
 10      Server-9119-FHA-S 9119      FHA        029F634     Operating
 11      Euler-710-SN06EAB 9123      710        06EAB1B     Operating
 12      Descartes-SN10D92 9123      710        10D921C     Operating
                                     more...

System: 9_

```

Figure 4-33 Selecting the server

Open the console window (Figure 4-34).

```

Type option, press Enter.
1=Connect dedicated
2=Connect shared
3=Show Details

Option  Partition      Partition State  Reference Use  Console Status
  2      5: adama           Running         00000000      2      Connected
  _      6: boxey           Running         00000000      1      Connected
  _      7: attica          Running         00000000      1      Connected

```

Figure 4-34 Open console window of a partition

5. In the first console panel when installing the IBM i, select a language for the IBM i operating system (Figure 4-35).

```

Language feature . . . . . : 2924
System: E10AF384

Press Enter to confirm your choice for language feature.
Installing the system will continue.

```

Figure 4-35 IBM i installation: Language selection

6. In the next console panel (Figure 4-36), select **1. Install Licensed Internal Code**.

```

Install Licensed Internal Code
System: E10AF384

Select one of the following:

1. Install Licensed Internal Code
2. Work with Dedicated Service Tools (DST)
3. Define alternate installation device

Selection
1

```

Figure 4-36 IBM i installation - Selecting Install Licensed Internal Code

7. Select the device for the load source unit. In our installation, we initially assigned only one XIV LUN to the tagged VSCSI adapter in IBM i, and we assigned the other LUNs later in System i DST. Therefore, we only have one unit to select as shown in Figure 4-37.

```

Select Load Source Device

Type 1 to select, press Enter.

Opt  Serial Number  Type  Model  Sys  Sys  I/O  I/O
   1  YH8EJU4FEBQY  6B22  050   255   3    0    0    1    0

```

Figure 4-37 IBM i installation - Selecting the load source unit

8. In the Install Licensed Internal Code (LIC) panel, select option **2. Install Licensed Internal Code and Initialize System** (Figure 4-38).

```

Install Licensed Internal Code (LIC)

Disk selected to write the Licensed Internal Code to:
Serial Number  Type  Model  I/O Bus  Controller  Device
YH8EJU4FEBQY  6B22  050    0      1      0

Select one of the following:

1. Restore Licensed Internal Code
2. Install Licensed Internal Code and Initialize system
3. Install Licensed Internal Code and Recover Configuration
4. Install Licensed Internal Code and Restore Disk Unit Data
5. Install Licensed Internal Code and Upgrade Load Source

Selection
2

```

Figure 4-38 Installing the Licensed Internal Code and initializing the system

9. When the warning is displayed, press F10 to accept it for the installation to continue. The installation starts, and you can observe the progress as shown in Figure 4-39.

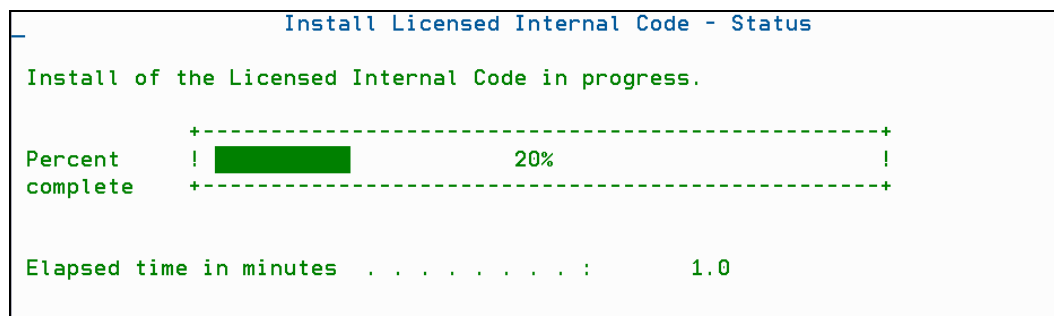


Figure 4-39 Progress of installing the Licensed Internal Code

10. After the installation of the Licensed Internal Code is completed, access the DST and add more disk units (LUNs) to the System i auxiliary storage pools (ASPs). Select the following options in DST as prompted by the panels:
 - a. Select option **4. Work with disk units.**
 - b. Select option **1. Work with disk configuration.**
 - c. Select option **3. Work with ASP configuration.**
 - d. Select option **3. Add units to ASPs.**
 - e. Choose an ASP to add disks to or use.
 - f. Select option **3. Add units to existing ASPs.**
 - g. In the Specify ASPs to Add Units to window, select the disk units to add to the ASP by specifying the ASP number for each disk unit, and press Enter.

In our example, we connected eight additional 154 GB LUNs from the XIV system and added them to ASP1. Figure 4-40 shows the load source disk and all the disk units added in ASP1. The disk unit names start with DMP, which indicates that they are connected in multipath.

ASP	Unit	Serial Number	Type	Model	Resource Name	Status	Hot Spare Protection
1						Unprotected	
	1	YH8EJU4FEBQY	6B22	050	DMP001	Configured	N
	2	YV9N8BP2R3Q3	6B22	050	DMP017	Configured	N
	3	YVQ9LKTH24PL	6B22	050	DMP007	Configured	N
	4	YBDGZKX8ES6S	6B22	050	DMP014	Configured	N
	5	YPCZSKB9P95K	6B22	050	DMP003	Configured	N
	6	Y97Y883RM53N	6B22	050	DMP015	Configured	N
	7	Y642ZRE5YHGM	6B22	050	DMP009	Configured	N
	8	YGRU6SYMMLMU	6B22	050	DMP005	Configured	N
	9	YAP3Z6SY6ENR	6B22	050	DMP011	Configured	N

Press Enter to continue.

Figure 4-40 LUNs in the System i ASP1

11.Exit the DST and choose option **2. Install the operating system** (Figure 4-41).

```

                                IPL or Install the System
                                System:  E10AF384

Select one of the following:

    1. Perform an IPL
    2. Install the operating system
    3. Use Dedicated Service Tools (DST)
    4. Perform automatic installation of the operating system
    5. Save Licensed Internal Code

Selection
    2

Licensed Internal Code - Property of IBM 5761-999 Licensed
Internal Code (c) Copyright IBM Corp. 1980, 2009. All
rights reserved. US Government Users Restricted Rights -
Use duplication or disclosure restricted by GSA ADP schedule
Contract with IBM Corp.

```

Figure 4-41 Installing the IBM i operating system

12.In the next console window, select the installation device type and confirm the language selection. The system starts an IPL from the newly installed load source disk (Figure 4-42).

```

                                Licensed Internal Code IPL in Progress
                                11/05/09 16:22

IPL:
  Type . . . . . : Attended
  Start date and time . . . . . : 11/05/09 16:21:51
  Previous system end . . . . . : Abnormal
  Current step / total . . . . . : 10 16
  Reference code detail . . . . . : C6004057

IPL step                                Time Elapsed    Time Remaining
Journal Recovery                        00:00:01        00:00:00
IFS Initialization                      00:00:01        00:00:00
>Data Base Recovery
Journal Synchronization
Commit Recovery

Item:
  Current / Total . . . . . :

Sub Item:
  Identifier . . . . . :
  Current / Total . . . . . :

```

Figure 4-42 Licensed Internal Code IPL when installing the operating system

13. Upon completion of the IPL, when the system prompts you to load the next installation media (Figure 4-43), insert the media B2924_01 into the DVD drive. For the message at the console, type G.

Figure 4-43 Loading the next installation media volume

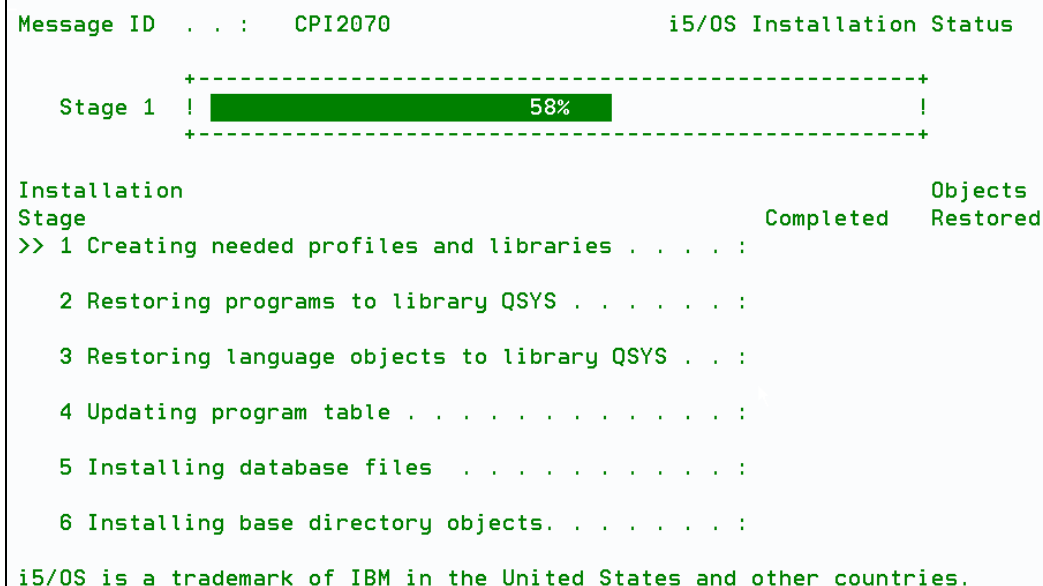


Figure 4-44 Progress of the IBM i installation

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this paper.

IBM Redbooks

For information about ordering these publications, see “How to get Redbooks” on page 80. Note that some of the documents referenced here may be available in softcopy only.

- ▶ *IBM i and IBM System Storage: A Guide to Implementing External Disks on IBM i*, SG24-7120
- ▶ *IBM i and Midrange External Storage*, SG24-7668
- ▶ *IBM PowerVM Virtualization Managing and Monitoring*, SG24-7590
- ▶ *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659

Other publications

These publications are also relevant as further information sources:

- ▶ *IBM XIV Storage System Host System*, GC27-2215
- ▶ *IBM XIV Storage System Installation and Service Manual*, GA32-0590
- ▶ *IBM XIV Storage System Introduction and Theory of Operations*, GC27-2214
- ▶ *IBM XIV Storage System Model 2810 Installation Planning Guide*, GC52-1327-01
- ▶ *IBM XIV Storage System Pre-Installation Network Planning Guide for Customer Configuration*, GC52-1328-01
- ▶ *IBM XIV Storage System XCLI Manual*, GC27-2213
- ▶ *XCLI Reference Guide*, GC27-2213

Online resources

These Web sites are also relevant as further information sources:

- ▶ IBM XIV Storage System overview
<http://www.ibm.com/systems/storage/disk/xiv/index.html>
- ▶ IBM XIV Storage System Information Center
<http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp>
- ▶ System Storage Interoperability Center (SSIC)
<http://www.ibm.com/systems/support/storage/config/ssic/index.jsp>

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 publications, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

Numerics

2810-A14 model 18, 20

2812-A14 model 18, 20

A

address space 7

AIX 23

Alerting Event 28

allocated space 11

ATS (Automatic Transfer Switch) 20

Automatic Transfer Switch (ATS) 20

B

bandwidth 23

battery 21

block 7

block-designated capacity 7

buffer 22

C

cache 3–4

buffer 22

capacity

allocation 10

reduction 44, 46

unallocated 12

client partition 43, 47, 69

comma-separated format 31

configuration flow 24

connectivity adapters 19

consistency group 9

application volumes 9

special snapshot command 9

consumed space 11

CRC (Cyclic Redundancy Check) 22

Cyclic Redundancy Check (CRC) 22

D

data integrity 9

data module 4, 15, 17, 19, 21–22

Gigabit Ethernet adapter 21

only difference 21

separate level 15

data redundancy 5, 14, 17–18

full 17

non-redundant data 17

default IP address 25

deletion priority 8

destage 16

detailed information 23

dirty data 16

disk drive 15, 22

disk unit 45, 76

distribution 17

DVD drive 51, 72–73

E

Ethernet

fabric 3–4

network 4

port 21

switch 4, 23

exp 34 45–46

Extended Command Line Interface (XCLI) 12, 24, 28, 31

command example 30

session 24, 30–31

environment 30

link 30

F

FDB (Fluid Dynamic Bearing) 23

Fibre Channel 4, 19–20

adapter 40–42

connectivity configuration 33

port 22, 36, 41

protocol 41

SAN environment 36

Fluid Dynamic Bearing (FDB) 23

full redundancy 17

function icons 27

G

Gigabit Ethernet adapter 21

global spare capacity 10

goal distribution 9–10, 17

full redundancy 10

priority 17

graceful shutdown 16

grid architecture 3

grid topology 14, 16

H

hard pool size 12–13

hard space depletion 7

hard storage capacity 28

hard system size 13

Hardware Management Console (HMC) 34, 68–69

HMC (Hardware Management Console) 34, 68–69

host connectivity 22

hot spot 9

HPUX 23

I

I/O activity 2

- I/O operation 28, 42, 44
- IBM 2810-A14 20
- IBM 2812-A14 20
- IBM i
 - best practices 43
 - queue depth 42
 - requirements for connecting to XIV Storage System 40
- IBM XIV Storage Management software 24, 32
- Integrated Virtualization Manager (IVM) 34–35
- interface module 4, 15, 17, 19, 22, 41, 43
 - dual CPU configuration 21
 - Gigabit Ethernet adapters 21
 - host requests 15
- IP address 25, 31
- iSCSI port 20, 22
 - interface module 22
- IVM (Integrated Virtualization Manager) 34–35

L

- LAN subnet 25
- Linux 23
- logical volume 2, 5–7, 9–10, 46
 - hard space 12
 - layout on physical disks 9
 - size 11
- LUN 42, 44–45

M

- machine type 19
- main display 27
- management workstation 25
- mapping 5
- master volume 9
- maximum number 7
- maximum volume count 7
- mean time between failure (MTBF) 22
- menu bar 27
- metadata 4, 10, 21
- Microsoft Windows 23
- modules 3
- MTBF (mean time between failure) 22
- multipath 43

N

- Native Command Queuing (NCQ) 22
- NCQ (Native Command Queuing) 22
- net usable capacity 10
- Node Port ID Virtualization (NPIV) 34, 36
- non-redundant data 17
- NPIV (Node Port ID Virtualization) 34, 36

P

- parallelism 3, 5
- partition 5
- patch panel 22
 - IP network 22
- phase-in 17

- phase-out 17
- physical architecture 4, 18
- physical capacity 2, 4, 6–8, 10, 12–13
 - zero 11
- physical disk 2, 5–6, 42
 - logical volume layout 9
 - queue depth 42
- pool size 12–13
- pool soft size 12
- Power Blade servers 40
- power supply 19–21
- PowerVM 34
 - Enterprise Edition 35
 - Express Edition 34
 - Standard Edition 35
- PowerVM Live Partition Mobility 35
- primary partition 6
- pseudo-random distribution algorithm 5

Q

- Query VIOS
 - function 71
 - result 70
- queue depth 42, 44

R

- rack 19
- rack door 21
- rebuild 17
- Redbooks Web site 80
 - Contact us viii
- redistribution 17
- redundancy 5, 14
- redundancy-supported reaction 15
- Redundant Power Supply (RPS) 20, 23
- regular pool 12–13
- regular storage pool 12
- remote mirroring 4
- reserve capacity 11, 13
- Rotation Vibration Safeguard (RVS) 22
- RPS (Redundant Power Supply) 20, 23
- RVS (Rotation Vibration Safeguard) 22

S

- SAN
 - connectivity 41
 - switches 41
- SAS (Serial-Attached SCSI) 22
- SATA (Serial Advanced Technology Attachment) 1, 22
 - disk 23
 - disk drives 22
- scale-out upgrade 17
- script 24, 31
- self-healing 2, 14
- Self-Protection Throttling (SPT) 22
- Serial Advanced Technology Attachment (SATA) 1, 22
 - disk 23
 - disk drives 22

- Serial-Attached SCSI (SAS) 22
- serviceability 14
- SFP (small form-factor pluggable) 23
- shutdown sequence 16
- small form-factor pluggable (SFP) 23
- snapshot 8
 - consistency group 9
 - reserve 8
- snapshot reserve capacity 7–8, 11, 13
- soft pool size 13
- soft storage capacity 28
- soft system size 13–14
- Solaris workstation 23
- space limit 7
- spare capacity 10
- SPT (Self-Protection Throttling) 22
- SSIC (System Storage Interoperability Center) 23
- status bar 28
- storage administrator 7–8, 11–13, 15, 17
- storage capacity 28
- storage pool 5, 7–8, 11–14, 28, 47, 57
 - capacity 11, 45
 - characteristics 7
 - logical volumes 8
 - thinly provisioned 12
- storage virtualization 2, 5
 - innovative implementation 5
- switch zoning 41
- system quiesce 16
- system reserve 10
- system size
 - hard 13
 - soft 14
- System Storage Interoperability Center (SSIC) 23
- system usable capacity 10
- system-initiated activity 2
- system-level thin provisioning 13

T

- TFC (Thermal Fly-height Control) 23
- Thermal Fly-height Control (TFC) 23
- thin provisioning 2, 10
 - storage pool 12
 - system level 13
- toolbar 27

U

- unallocated capacity 12
- uninterruptible power supply 16, 21, 27
 - module 19, 21
- usable capacity 20

V

- VIOS (Virtual I/O Server) 33, 35–36, 40, 47
 - logical partition 35
 - multipath 43
 - partition 40–41, 47–48, 55
 - planning 42

- queue depth 42
- Version 2.1.1 40
- Virtual I/O Server (VIOS) 33, 35–36, 40, 47
 - logical partition 35
 - multipath 43
 - partition 40–41, 47–48, 55
 - planning 42
 - queue depth 42
 - Version 2.1.1 40
- virtual SCSI adapter 36, 42, 47–49
- virtualization management (VM) 34, 36–37
- VM (virtualization management) 34, 36–37
- volume count 7
- volume size 7, 11
 - actual 11

X

- XCLI (Extended Command Line Interface) 12, 24, 28, 31
 - command example 30
 - session 24, 30–31
 - environment 30
 - link 30
- XIV Storage Management GUI 25, 30
 - features 28
 - launching 25
 - main menu 26
- XIV Storage Management software 23–24, 27, 30
 - compatibility 23
 - documentation 32
 - function icons 27
 - menu options 27
- XIV Storage System
 - 2810-A14 model 18
 - 2812-A14 model 18
 - architecture 3, 5–6, 14–15, 44
 - data distribution 17
 - design 2
 - distribution algorithm 17
 - family 19
 - hardware 3, 5, 19, 23, 27
 - hardware component 20
 - I/O performance 44
 - Information Center 32
 - key design features 2
 - logical architecture 5
 - logical hierarchy 6
 - logical volume layout on physical disks 9
 - LUN 42, 56, 67, 75
 - queue depth 42
 - open architecture 2
 - overview 1
 - planning for an IBM i environment 39
 - PowerVM client connectivity 37
 - reliability 14
 - requirements 40
 - SATA disks 23
 - software 23
 - system usable capacity 10
 - technician 18
 - virtualization 5

volume 47

Z

zero physical capacity 11



IBM XIV Storage System with the Virtual I/O Server and IBM i



**Understand how to
attach the XIV server
to IBM i through VIOS
2.1.2**

**Learn how to exploit
the multipathing
capability**

**Follow the setup
tasks step-by-step**

In this IBM Redpaper publication, we discuss and explain how you can connect the IBM XIV Storage System server to the IBM i operating system through the Virtual I/O Server (VIOS). A connection through the VIOS is especially interesting for IT centers that have many small IBM i partitions. When using the VIOS, the Fibre Channel host adapters can be installed in the VIOS and shared by many IBM i clients by using virtual connectivity to the VIOS.

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