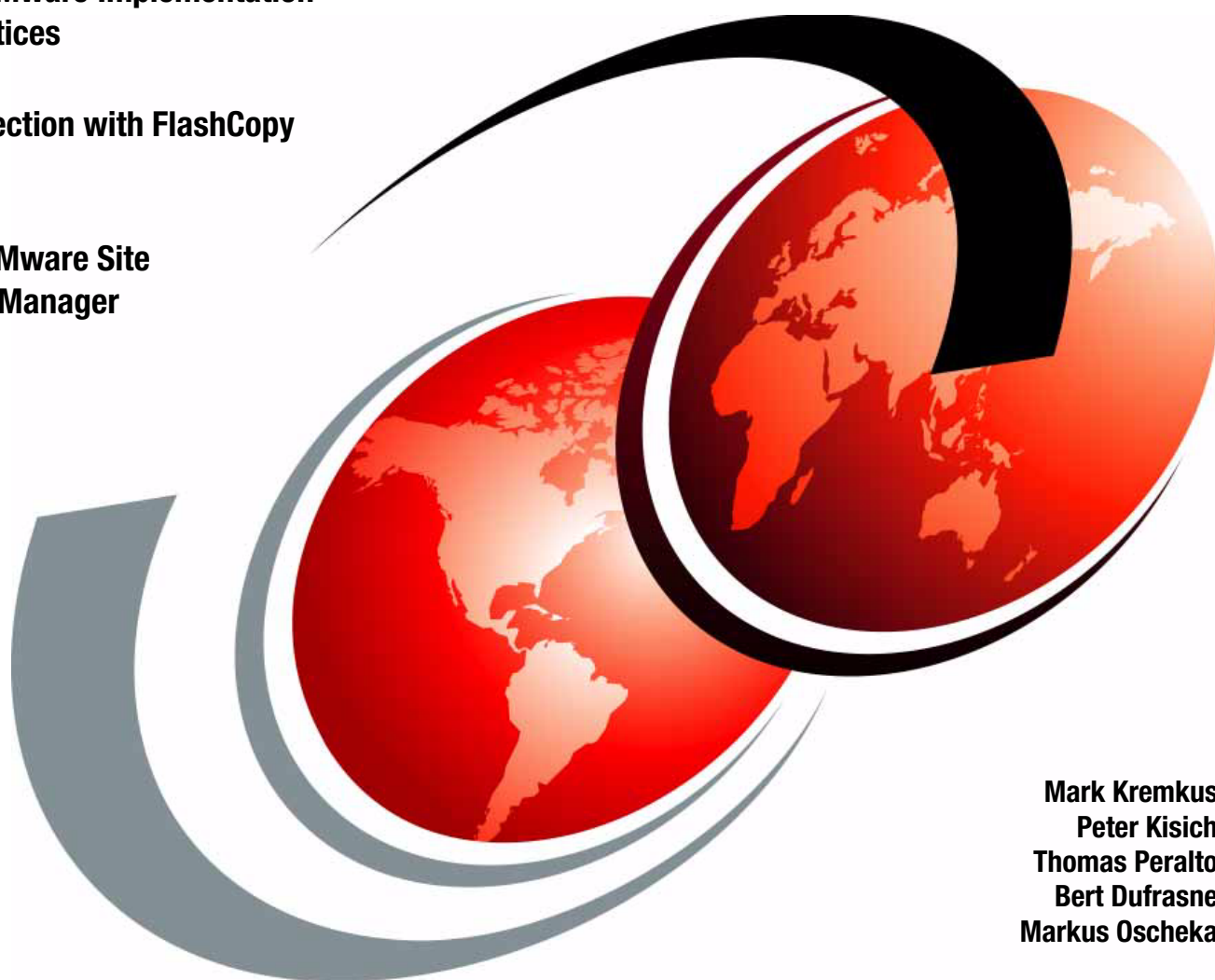


XIV Storage System in a VMware Environment

XIV and VMware Implementation
Best Practices

Data Protection with FlashCopy
Manager

XIV and VMware Site
Recovery Manager



Mark Kremkus
Peter Kisich
Thomas Peralto
Bert Dufrasne
Markus Oscheka



International Technical Support Organization

XIV Storage System in a VMware environment

March 2013

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

First Edition (March 2013)

This edition applies to the IBM XIV Storage System (Machine types 2812-114 and 2810-114) with XIV system software Version 11.1.1.

© Copyright International Business Machines Corporation 2013. 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	viii
Now you can become a published author, too!	ix
Comments welcome	ix
Stay connected to IBM Redbooks	x
Chapter 1. XIV Storage System with VMware	1
1.1 Introduction	2
1.1.1 IBM XIV Storage System integration with XIV	4
1.1.2 VAAI Support	4
1.1.3 vSphere deployment flexibility with the IBM XIV Storage System	4
Chapter 2. XIV and VMware Integration	9
2.1 Integration concepts and implementation Best Practices	10
2.1.1 vSphere storage architectural overview	11
2.1.2 XIV and VMware Storage Distributed Resource Scheduler	12
2.1.3 XIV and VMware LUN and datastore provisioning best practices	15
2.1.4 XIV and VMware general connectivity best practices	16
2.1.5 XIV and VMware thin provisioning	18
2.1.6 XIV and VMware best practices for Quality of Service	24
Chapter 3. VMware vStorage APIs Array Integration	27
3.1 VAAI overview	28
3.1.1 Software prerequisites to use VAAI	32
3.1.2 Installing the IBM VAAI device driver on an ESXi 4.1 server	32
3.1.3 Confirming VAAI Hardware Acceleration is detected	34
3.1.4 Disabling and enabling VAAI on the XIV on a per volume basis	37
3.1.5 Testing VAAI	38
Chapter 4. Attaching VMware ESX to XIV	41
4.1 VMware ESX 3.5 and XIV	42
4.1.1 Installing HBA drivers	42
4.1.2 Scanning for new LUNs	42
4.1.3 Assigning paths from an ESX 3.5 host to XIV	45
4.2 VMware Multi-Pathing architecture overview	50
4.3 VMware ESX and ESXi 4.x and XIV	52
4.3.1 Installing HBA drivers	52
4.3.2 Identifying ESX host port WWN	52
4.3.3 Scanning for new LUNs	53
4.3.4 Attaching an ESX/ESXi 4.x host to XIV	55
4.3.5 Configuring ESX/ESXi 4.x host for multipathing with XIV	57
4.3.6 Performance tuning tips for ESX/ESXi 4.x hosts with XIV	64
4.3.7 VMware vStorage API Array Integration (VAAI)	68
4.4 VMware ESXi 5.0/5.1 and XIV	68
4.4.1 ESXi 5.0/5.1 Fibre Channel configuration	68
4.4.2 Performance tuning tips for ESXi 5 hosts with XIV	69

4.4.3	Creating datastores that are larger than 2 TiB	73
Chapter 5.	IBM Storage Management Console for VMware vCenter	75
5.1	The IBM Storage Management Console for VMware vCenter	76
5.1.1	Installation	76
5.1.2	Customizing the plug-in	77
5.1.3	Adding IBM Storage to the plug-in	80
5.1.4	Checking and matching XIV Volumes	83
5.1.5	Creating a datastore	83
5.1.6	Using a read-only user	85
5.1.7	Locating the user guide and release notes	85
5.1.8	Troubleshooting	85
Chapter 6.	Data Protection in vSphere environments with XIV	87
6.1	Data Protection in vSphere environments with XIV	88
6.1.1	Data Protection solution concepts and terminology	88
6.1.2	vSphere Data Protection solution components	89
6.1.3	Tivoli Storage FlashCopy Manager for VMware	91
6.1.4	Comprehensive data protection for vSphere	93
Chapter 7.	XIV Storage System and VMware Site Recovery Manager	111
7.1	XIV Storage System and VMware Site Recovery Manager	112
7.1.1	Overview of XIV Remote Mirroring	113
7.1.2	VMware Site Recovery Manager overview	114
7.1.3	Minimum XIV and SRM solution prerequisites	116
7.1.4	XIV and SRM integration with the XIV Storage Replication Adapter	118
7.1.5	Site Recovery Manager operations	118
7.2	Quick install guide for VMware Site Recovery Manager	160
7.2.1	Installing and configuring the database environment	161
7.2.2	Installing the vCenter server	175
7.2.3	Installing and configuring vCenter client	179
7.2.4	Installing the SRM server	184
7.3	Installing the vCenter Site Recovery Manager plug-in	188
7.3.1	Installing XIV Storage Replication Adapter for VMware SRM	189
7.3.2	Configuring the IBM XIV System Storage for VMware SRM	190
Related publications		193
IBM Redbooks		193
Other publications		193
Online resources		194
Help from IBM		194

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 grant 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 websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites 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.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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:

DB2®	Redbooks®	Tivoli®
DS8000®	Redpaper™	XIV®
FlashCopy®	Redbooks (logo)  ®	
IBM®	System Storage®	

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.

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

QLogic, and the QLogic logo are registered trademarks of QLogic Corporation. SANblade is a registered trademark in the United States.

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

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

Preface

The IBM® XIV® Storage System is an excellent choice for VMware storage requirements. XIV achieves consistent high performance by balancing the workload across physical resources. This paper includes information about the following topics:

- ▶ XIV Storage System and VMware Integration Concepts
- ▶ XIV Storage System and VMware Implementation Best Practices
- ▶ XIV Storage System integration harnessing VMware APIs including:
 - vStorage APIs for Array Integration (VAAI)
 - vSphere API for Storage Awareness (VASA)
 - vStorage API for Data Protection (VADP) interfacing with Tivoli® Storage FlashCopy® Manager (FCM) and Tivoli Storage Manager for Virtualization Environments (TSM for VE)
- ▶ Connection for ESX version 3.5, ESX/ESXi version 4.x, and ESXi 5.0/5.1
- ▶ The IBM vCenter plug-in
- ▶ The XIV Storage Replication Adaptor (SRA)
- ▶ VMware Site Recovery Manager (SRM)

This IBM Redpaper™ is intended for those who want to plan, design, and configure an XIV based storage solution in a VMware environment.

The team who wrote this paper

This paper was produced by a team of specialists from around the world.



Mark Kremkus is a Senior Information Technology (IT) Specialist in the Advanced Technical Skills organization. He has 11 years of experience in the design of high-performance, high-availability solutions. Mark achieved Consulting-level certification in Actualizing IT solutions. Mark's areas of expertise are enterprise storage performance analysis with emphasis on using empirical data to perform mathematical modeling of disk storage performance and integrating storage with open systems hypervisors. He writes and presents on these topics. He holds a Bachelor of Science degree in Electrical Engineering from Texas A&M.



Peter Kisich is a Senior Storage Technical specialist with extensive experience in storage area network (SAN), network-attached storage (NAS), VMware, Backup, and the UNIX operating system as they relate to storage.



Thomas Peralto is a principal consultant in the Storage Solutions Engineering group. He has extensive experience in implementing large and complex transport networks and mission-critical data protection throughout the globe. Mr. Peralto also serves as a data replication and data migration expert and speaks both at national and international levels for IBM on the best practices for corporate data protection.



Bert Dufrasne is an IBM Certified Consulting IT Specialist and Project Leader for IBM System Storage® disk products at the International Technical Support Organization, San Jose Center. He has worked at IBM in various IT areas. He also authored many IBM Redbooks® publications and has developed and taught technical workshops. Before joining the IBM International Technical Support Organization (ITSO), he worked for IBM Global Services as an Application Architect. He holds a Master's degree in Electrical Engineering.



Markus Oscheka is an IT Specialist for Proof of Concepts and Benchmarks in the Disk Solution Europe team in Mainz, Germany. His areas of expertise include set up and demonstration of IBM System Storage solutions in open environments. He wrote several IBM Redbooks and acted as the co-project lead for the IBM DS8000® and IBM XIV Storage IBM Redbooks. He holds a degree in Electrical Engineering from the Technical University in Darmstad.

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

Eugene Tsypin, Carlos Lizarralde, Brian Carmody,
IBM

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

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 email 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

Stay connected to IBM Redbooks

- ▶ Find us on Facebook:
<http://www.facebook.com/IBMRedbooks>
- ▶ Follow us on Twitter:
<http://twitter.com/ibmredbooks>
- ▶ Look for us on LinkedIn:
<http://www.linkedin.com/groups?home=&gid=2130806>
- ▶ Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:
<https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm>
- ▶ Stay current on recent Redbooks publications with RSS Feeds:
<http://www.redbooks.ibm.com/rss.html>



XIV Storage System with VMware

This chapter is a high-level introduction to the concepts, features, and business relationships that make the XIV Storage System and VMware, the perfect fit for an optimized storage-server virtualization solution.

1.1 Introduction

Virtualization technology is transforming business. Companies are increasingly virtualizing their environments to meet these goals:

- ▶ Consolidate servers
- ▶ Centralize services
- ▶ Implement disaster recovery
- ▶ Set up remote or thin client desktops
- ▶ Create clouds for optimized resource use

Organizations often deploy server virtualization to gain economies of scale by consolidating underutilized resources to a new platform. Equally crucial to a server virtualization scenario is the storage itself. Implementing server virtualization without taking storage into account can cause challenges, such as uneven resource sharing and performance and reliability degradation.

The IBM XIV Storage System, with its grid architecture, automated load balancing, and ease of management, provides best-in-class virtual enterprise storage for virtual servers. It also provides the following advantages to help meet your enterprise virtualization goals:

- ▶ IBM XIV end-to-end support for VMware solutions, including vSphere and vCenter
- ▶ Provides hotspot-free server-storage performance
- ▶ Optimal resource use
- ▶ An on-demand storage infrastructure that allows simplified growth

IBM collaborates with VMware on the strategic, functional, and engineering levels. IBM XIV system uses this technology partnership to provide robust solutions and release them quickly. The XIV system is installed at the VMware Reference Architecture Lab and other VMware Engineering Development labs. It is used for early testing of new VMware product release features. Among other VMware product projects, IBM XIV took part in the development and testing of VMware ESX 4.1.

IBM XIV engineering teams have ongoing access to VMware co-development programs, such as developer forums. They also have access to a comprehensive set of developer resources including toolkits, source code, and application programming interfaces. This access translates to excellent virtualization value for customers.

For more information, see *A Perfect Fit: IBM XIV Storage System with VMware for Optimized Storage-Server Virtualization*, available at:

ftp://ftp.hddtech.ibm.com/isv/A_Perfect_Fit_IBM_XIV_and_VMware.pdf

VMware offers a comprehensive suite of products for server virtualization:

- ▶ VMware ESX and ESXi server: This production-proven virtualization layer runs on physical servers. It allows processor, memory, storage, and networking resources to be provisioned to multiple virtual machines.
- ▶ VMware Virtual Machine file system (VMFS): A high-performance cluster file system for virtual machines.
- ▶ VMware Virtual symmetric multiprocessing (SMP): Allows a single virtual machine to use multiple physical processors simultaneously.
- ▶ VMware Virtual Machine: A representation of a physical system by software. A virtual machine has its own set of virtual hardware on which an operating system and applications are loaded. The operating system sees a consistent, normalized set of hardware regardless of the actual physical hardware components. VMware virtual

machines contain advanced hardware features, such as 64-bit computing and virtual symmetric multiprocessing.

- ▶ vSphere Client: An interface allowing administrators and users to connect remotely to the VirtualCenter Management Server or individual ESX installations from any Windows PC.
- ▶ VMware vCenter Server: Centrally manages VMware vSphere environments. It gives IT administrators dramatically improved control over the virtual environment compared to other management platforms. Formerly called VMware VirtualCenter.
- ▶ Virtual Infrastructure Web Access: A web interface for virtual machine management and remote consoles access.
- ▶ VMware VMotion: Allows the live migration of running virtual machines from one physical server to another, one datastore to another, or both. This migration has zero downtime, continuous service availability, and complete transaction integrity.
- ▶ VMware Site Recovery Manager (SRM): A business continuity and disaster recovery solution for VMware ESX servers providing VM-aware automation of emergency and planned failover/failback scenarios between data centers incorporating either server or storage-based datastore replication.
- ▶ vStorage APIs for Storage Awareness (VASA): An API that facilitates the awareness of specific storage-centric attributes to vCenter. These functional and non-functional characteristics are automatically surfaced by a VASA-compatible storage subsystem and presented to vCenter to enhance intelligent automation of storage resource management in conjunction with the VMware Profile-Driven Storage resource classification and deployment methodology.
- ▶ VMware Storage Distributed Resource Scheduler (DRS): Facilitates the automated management of initial VMDK placement. It also facilitates continual, dynamic balancing of VMDKs among clustered datastores by identifying the most appropriate resource candidates based on capacity, performance, and functional characteristics that are specific to the requirements of individual virtual machines or clusters. Beginning in vSphere 5.0, VMware Storage DRS can take advantage of VASA-based and administrator-based storage resource classifications to realize simplification of heterogeneous storage management based on the concept of Profile-Drive Storage, which organizes diverse storage resources into profiles meeting specific classification criteria.
- ▶ VMware high availability (HA): Provides easy-to-use, cost-effective high availability for applications running in virtual machines. If a server fails, effected virtual machines are automatically restarted on other production servers that have spare capacity.
- ▶ VMware Consolidated Backup (VCB): Provides an easy-to-use, centralized facility for agent-free backup of virtual machines that simplifies backup administration and reduces the load on ESX installations. VCB is being replaced by VMware vStorage APIs for Data Protection.
- ▶ VMware vStorage APIs for Data Protection: Allows backup software, such as IBM Tivoli Storage Manager for Virtual Environments (version 6.2 or later), optionally in conjunction with Tivoli Storage FlashCopy Manager for VMware (version 3.1 or later), to perform customized, scheduled centralized backups at the granularity of virtual machines, and recovery at the datastore, virtual machine, or file level. You do not have to run backup tasks inside each virtual machine.
- ▶ VMware Infrastructure software development kit (SDK): Provides a standard interface for VMware and third-party solutions to access VMware Infrastructure.

1.1.1 IBM XIV Storage System integration with XIV

IBM XIV provides end-to-end support for VMware with ongoing support for VMware virtualization solutions as they evolve and are developed. Specifically, IBM XIV works in concert with the following VMware products and features:

- ▶ vSphere ESX
- ▶ vSphere Hypervisor (ESXi)
- ▶ vCenter Server using the IBM Storage Management Console for VMware vCenter
- ▶ vStorage APIs for Data Protection (VADP) (using Tivoli Storage FlashCopy Manager and Tivoli Storage Manager for Virtual Environments)
- ▶ vSphere vMotion and Storage vMotion
- ▶ vSphere APIs for Storage Awareness (VASA) in concert with VMware Distributed Resource Scheduler (DRS) and Storage I/O Control (SIOC)
- ▶ vSphere Storage APIs for Array Integration (VAAI)

1.1.2 VAAI Support

ESX/ESXi 4.1 brought a new level of integration with storage systems through the use of vStorage API for Array Integration (VAAI). VAAI helps reduce host usage and increases scalability and the operational performance of storage systems, particularly in densely configured, multi-tenant virtual environments. The traditional ESX operational model with storage systems forced the ESX host to issue many identical commands to complete certain types of operations, including cloning operations. Using VAAI, the same task can be accomplished with far fewer commands, reduced contention, and with the potential to greatly reduce resource consumption at all levels along the I/O path.

For more information, see Chapter 3, “VMware vStorage APIs Array Integration” on page 27.

1.1.3 vSphere deployment flexibility with the IBM XIV Storage System

The minimum implementation of a VMware virtualization environment utilizing XIV Storage System requires the deployment of at least one ESX/ESXi server to host the Virtual Machines and one vCenter server and vCenter client. Also, ensure that VAAI is enabled and that the vCenter plug-in is installed. Finally, you need redundancy at both the network and SAN levels.

You can implement a clustered, high-availability solution in your environment by adding and deploying an additional server (or servers) running under VMware ESX/ESXi in addition to implementing the VMware High Availability option for your ESX/ESXi servers. This feature is designed to operate in synergy with VMware Distributed Resource Scheduler (DRS) and Storage DRS, particularly in a clustered datastore environment. The high-availability feature maintains VM and application availability while optimizing capacity utilization and load balancing for performance. For more information about VMware HA and DRS, refer to 2.1.2, “XIV and VMware Storage Distributed Resource Scheduler” on page 12 and to the paper located at:

<http://www.vmware.com/resources/techresources/402>

Tivoli Storage FlashCopy Manager for VMware seamlessly integrates with the advanced snapshot technology of the XIV Storage System and vStorage APIs for Data Protection to implement robust end-to-end centralized backup at the datastore or datastore cluster level and restore capabilities. With Tivoli Storage FlashCopy Manager for VMware, there is no

need to deploy OS-specific agents on each VM. This solution can be further enhanced by incorporating Tivoli Storage Manager for Virtual Environments to implement off-host incremental data backup (supporting VMware Changed Block Tracking) and archival processes targeting appropriate nearline or lower-tiered media. Refer to the following VMware Knowledge Base article for further details about VADP:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1021175

To further improve the availability of your virtualized environment, simplify business continuity and disaster recovery solution leveraging integrated, automated VM-aware failover and failback, consider implementing ESX servers, vCenter server, and another XIV storage system at the recovery site. Also, install VMware Site Recovery Manager, and use the Storage Replication Adapter to integrate VMware Site Recovery Manager with your XIV storage systems at both sites. The Site Recovery Manager itself can also be implemented as a virtual machine on the ESX server. Of course, both the primary data center and the disaster recovery data center(s) must incorporate redundant networks and SANs.

As part of a comprehensive strategy to meet SLAs and enforce performance protection for business-critical VMs in a highly consolidated virtualized storage environment, consider implementing vSphere Storage I/O Control (SIOC), ideally in conjunction with the QoS performance classes of the XIV Storage System to address prioritization at the LUN level as well. Introduced in vSphere 4.1, SIOC consistently monitors latency and dynamically alleviates the impact of resource contention during peak workload periods by limiting resource consumption by non-critical, performance-hungry VMs to the quotas set by the administrator in favor of improving performance for VMs hosting high-priority workloads, particularly during peak periods. For additional guidance about implementing vSphere SIOC, refer to “vSphere Storage I/O Control (SIOC)” on page 25, and the technical paper at:

<http://www.vmware.com/files/pdf/techpaper/VMW-vSphere41-SIOC.pdf>

Figure 1-1 on page 6 is a modular architectural overview that summarizes the key vSphere and XIV integration points.

Deep XIV storage integration with VMware is provided at no additional charge as part of IBM XIV licensing. Integration works either ready-for-use, as is the case for VAAI with vSphere 5.0, or with simple software plug-ins or drivers.

Optional, separately-licensed Tivoli products provide additional VMware data protection integration.

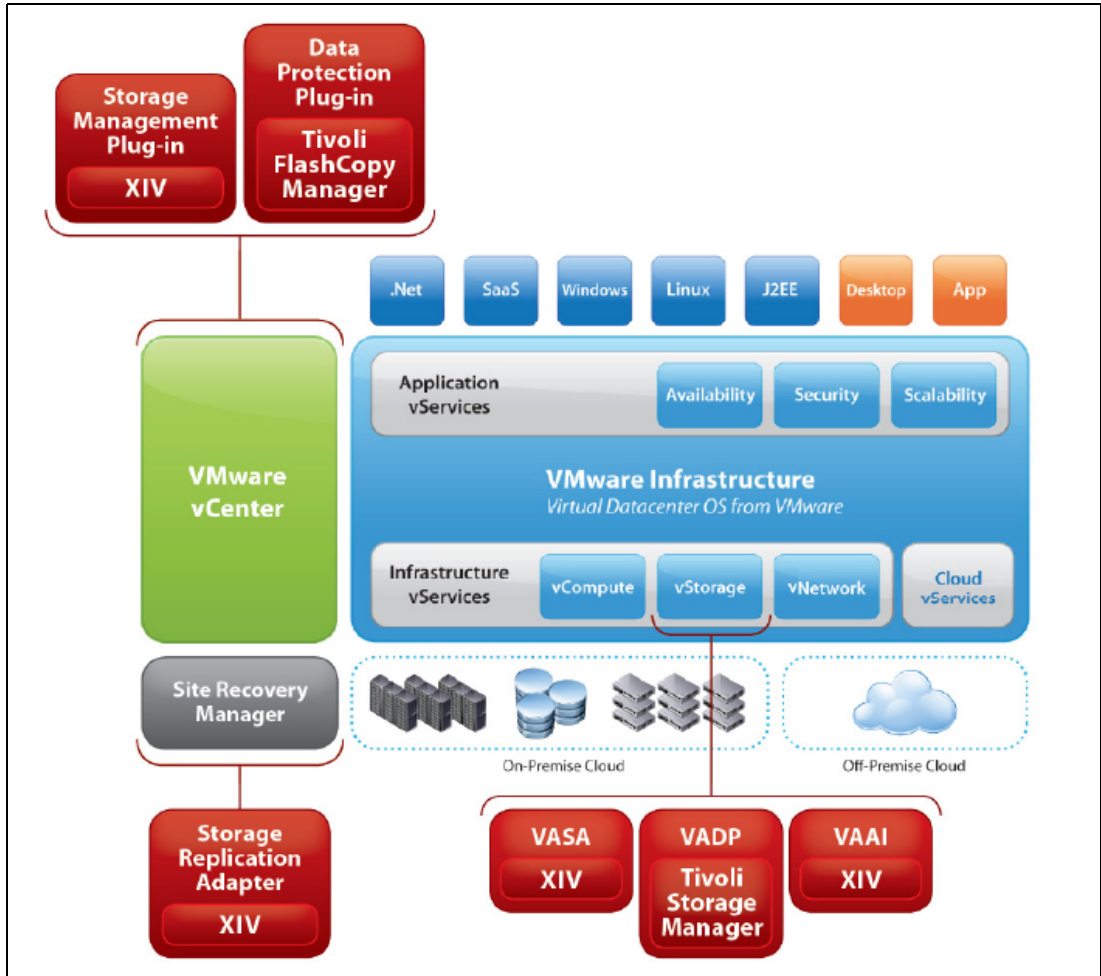


Figure 1-1 XIV Integration Points with VMware

Note: Although the APIs within the categories of VAAI and VASA in a vSphere 5.x environment require no specific tuning or best practices implementation on the part of VMware or XIV administrators, full exploitation of these features in the broader context of the vSphere 5.x ecosystem requires end-to-end awareness of feature interoperability, storage-agnostic interoperability limitations that can be imposed by VMware, and VMware's recommended best practices. This publication does not cover these topics exhaustively; however, you are encouraged to investigate them further using the following resources and to subsequently incorporate these elements into solution design and planning specific to the intended usage of vSphere:

VMware VAAI:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1021976

VMware VASA:

http://kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&docType=kc&externalId=2004098&sliceId=1&docTypeID=DT_KB_1_1&dialogID=455566651&stateId=1%200%20455588236

Figure 1-2 illustrates a full solution including disaster recovery capability.

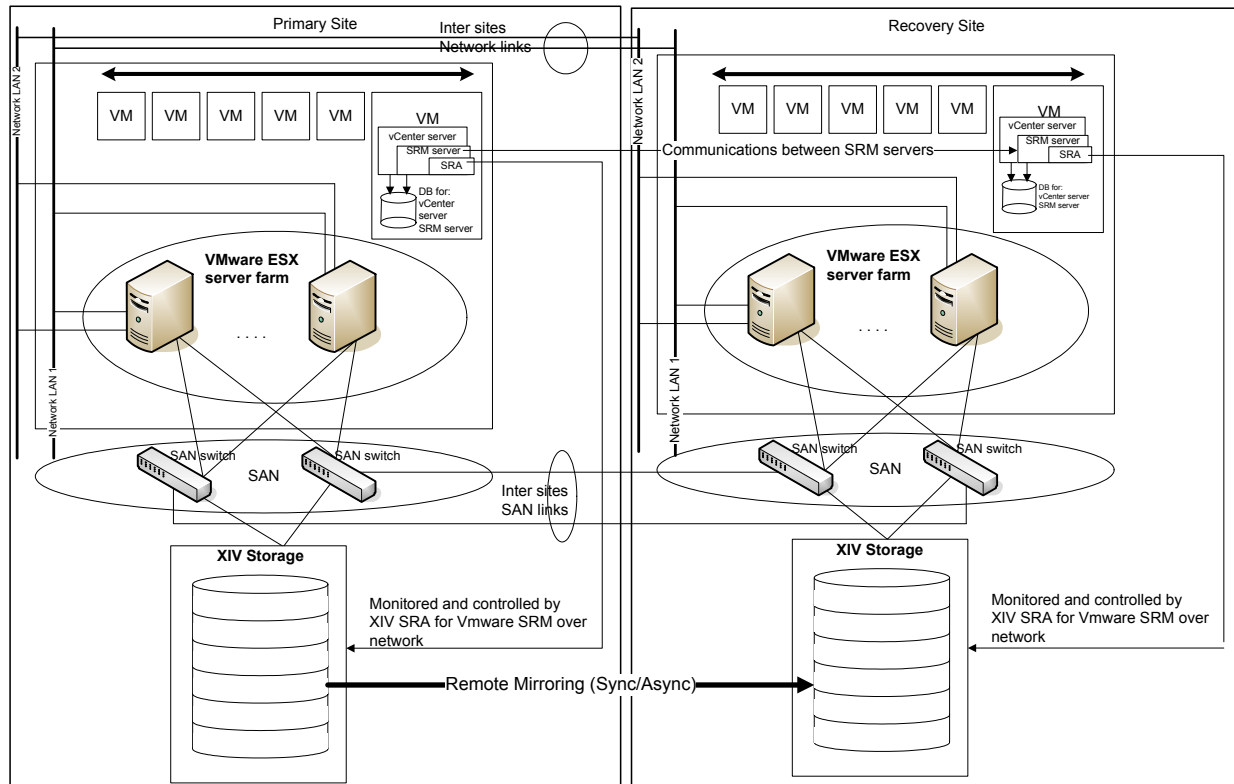


Figure 1-2 Integrated vSphere Disaster Recovery Environment Built on IBM XIV Storage System

The remainder of this chapter is divided into several major sections. The first two sections are dedicated to comprehensively addressing the unique integration concepts underlying the deployment of a powerful, adaptable vSphere environment on the XIV Storage System. The concepts described offer a deeper understanding of the robust synergy and value achievable when deploying vSphere environments harnessing XIV storage with emphasis on implementation best practices. The final three sections address specifics for VMware ESX 3.5, ESX/ ESXi 4.x, and ESXi 5.1:

- ▶ XIV Storage and VMware Integration Concepts and Implementation Best Practices
- ▶ vStorage APIs for Array Integration (VAAI)
- ▶ The IBM Storage Management Console for VMware vCenter
- ▶ The XIV Storage Replication Adapter for VMware Site Recovery Manager
- ▶ Backup and restore leveraging VMware ADP and Tivoli Storage FlashCopy Manager for VMware on the XIV Storage System



XIV and VMware Integration

This chapter addresses IT decision makers, storage administrators, and VMware administrators. It offers a complete overview of XIV storage and VMware integration concepts and general implementation best practices.

2.1 Integration concepts and implementation Best Practices

At a fundamental level, the goal of both the XIV Storage System and VMware's storage features is to significantly reduce the complexity of deploying and managing storage resources. With XIV, storage administrators can provide consistent tier-1 storage performance and quick change-request cycles because they perform little planning and maintenance to keep performance levels high and storage optimally-provisioned.

The underlying strategies devised within the vSphere storage framework to insulate administrators from complex storage management tasks, non-optimal performance, and capacity resource utilization, include:

- ▶ Make storage objects much larger and more scalable, reducing the number that to be managed by the administrator
- ▶ Extend specific storage resource-awareness by attaching features and profiling attributes to the storage objects
- ▶ Help administrators make the correct storage provisioning decision for each Virtual Machine or even fully automate the intelligent deployment of Virtual Machine storage.
- ▶ Remove many time-consuming and repetitive storage-related tasks, including the need for repetitive physical capacity provisioning.

Clearly, vCenter relies upon the storage subsystem to fully support several key integration features to effectively implement these strategies. Appropriately compatible storage, such as XIV, is essential.

To provide contrast, consider traditional storage provisioning in vSphere, which typically tasks the vSphere administrator with the following storage-centric responsibilities:

- ▶ Determine the correct datastore on which to initially place a VM's virtual disk.
- ▶ Continuously monitor datastores for capacity consumption.
- ▶ Continuously monitor datastores for performance/latency.
- ▶ Repetitively deploy physical LUN as capacity consumption grows.
- ▶ Ensure that a VM remains backed by a suitable storage resource throughout its lifecycle.

Additional concerns for vSphere administrators can include:

- ▶ Possible mistrust of Thin Provisioning due to Out-Of-Space situation.
- ▶ Possible mistrust of physical capacity usage reporting of Thin Provisioned LUNs.

The remainder of this chapter addresses each of these hurdles, demonstrating the concepts and operational practices necessary to derive maximal value from the unique vSphere-specific capabilities of the XIV Storage System. As an introduction to essential storage principles in the vSphere environment, a brief overview precedes the discussion of integration principles and best practices.

2.1.1 vSphere storage architectural overview

First, consider the vSphere storage architecture, including physical and logical storage elements shown in Figure 2-1. While not intended to thoroughly explore vSphere storage concepts and terminology, the essential components and their relationships provide the foundational framework necessary to understand the upcoming integration principles.

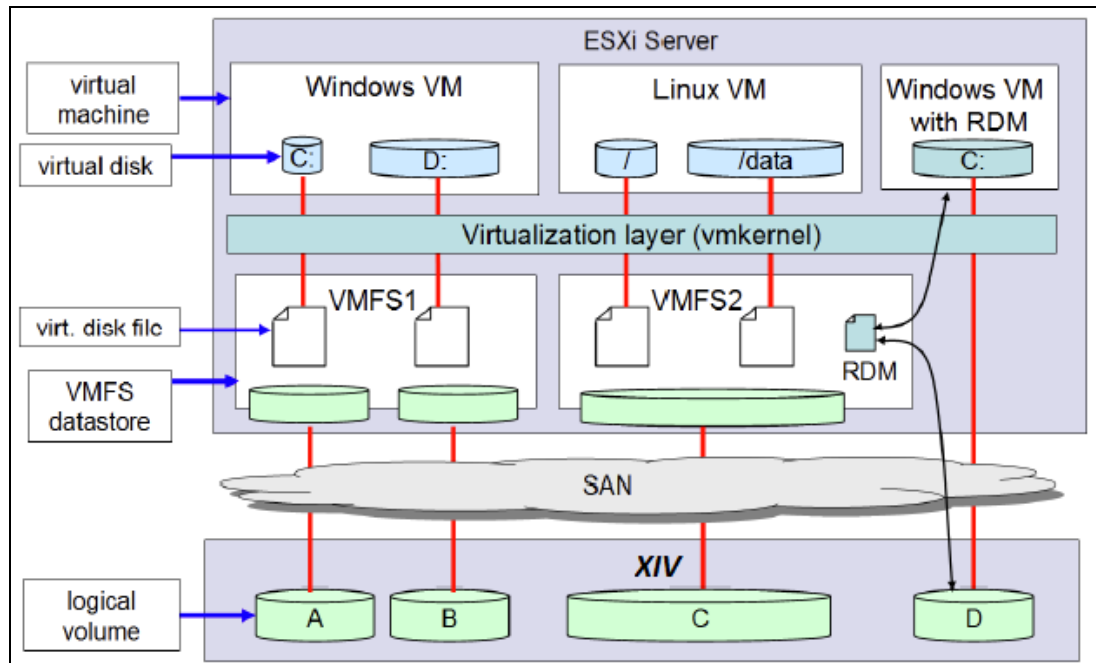


Figure 2-1 ESX/ESXi Basic Storage Elements in the vSphere Infrastructure

The VMware file system (VMFS) is the central abstraction layer that acts as a medium between the storage and the hypervisor layers. The current generation of VMFS evolved to include the following distinguishing attributes, among others:

- ▶ *Clustered file system*: Purpose-built, high performance clustered file system for storing virtual machine files on shared storage (Fibre Channel and iSCSI). The primary goal of VMFS's design consists of functioning as an abstraction layer between the VMs and the storage to efficiently pool and manage storage as a unified, multi-tenant resource.
- ▶ *Shared data file system*: Enables multiple vSphere hosts to read and write from the same datastore concurrently.
- ▶ *Online insertion or deletion of nodes*: Adds or removes vSphere hosts from VMFS volume with no impact to adjacent hosts or VMs.
- ▶ *On-disk disk file locking*: Ensure that the same virtual machine is not accessed by multiple vSphere hosts concurrently. This topic is fully addressed in Chapter 3, "VMware vStorage APIs Array Integration" on page 27.

What follows examines concepts and best practices crucial to building an adaptable, efficient, high-performance vSphere infrastructure with the XIV Storage System's inherently cloud-optimized design and deep vSphere integration capabilities at its foundation.

2.1.2 XIV and VMware Storage Distributed Resource Scheduler

VMware Storage Distributed Resource Scheduler (SDRS) balances the VM virtual disk (VMDK) placement dynamically across datastores by identifying the most appropriate resource candidates based on capacity, performance, and functional characteristics that are specific to the requirements of individual virtual machines and clusters. However, because all storage created for vSphere on the XIV Storage System is balanced across all processors, ports, spindles, and cache modules in the XIV Storage System, XIV already effectively addresses many of the complex performance management challenges that SDRS load-balancing is designed to alleviate. By definition, the XIV Storage System's grid architecture achieves an optimal balance without hitting thresholds and triggers that instigate redistribution of disk resources.

VASA and Profile Driven Storage

vStorage APIs for Storage Awareness (VASA) enables vSphere 5.0 or higher to incorporate storage resource attributes dynamically surfaced by VASA-compatible storage subsystems into the persistent monitoring of storage resources for purposes of invoking intelligence-driven, storage-centric operations, for example through Storage Distributed Resource Scheduler (SDRS). vSphere 5.0 introduces another feature called *Profile Driven Storage* that serves as a resource classification and policy enforcement mechanism that works in synergy with VASA and consequently enhances SDRS functionality even further. Figure 2-2 on page 13 illustrates the functional relationships between all of these elements and their integration into the VMware infrastructure.

VASA and Profile Driven Storage synergy enable the following benefits:

- ▶ Profile Driven Storage helps make the initial placement of a VM error-free by allowing administrators to create profiles that contain storage characteristics and then map the VM's and datastore resources to these profiles:
 - These storage resource characteristics can be surfaced through VASA or can be user-defined business tags (for example, gold, silver, bronze).
- ▶ VASA's dynamic reporting also enables VMs to remain compliant with their predefined storage requirements based on classifications set forth using Profile Driven Storage:
 - The status of a VM can change from compliant to non-compliant based on transient conditions, and appropriate actions can be taken by administrators or by SDRS to restore and maintain compliance.

In summary, adding vSphere 5 VASA support delivers actionable storage insights, such as availability, alerts, and events, to both enhance SDRS and provide VMware administrators with enhanced storage infrastructure management and decision-making capabilities.

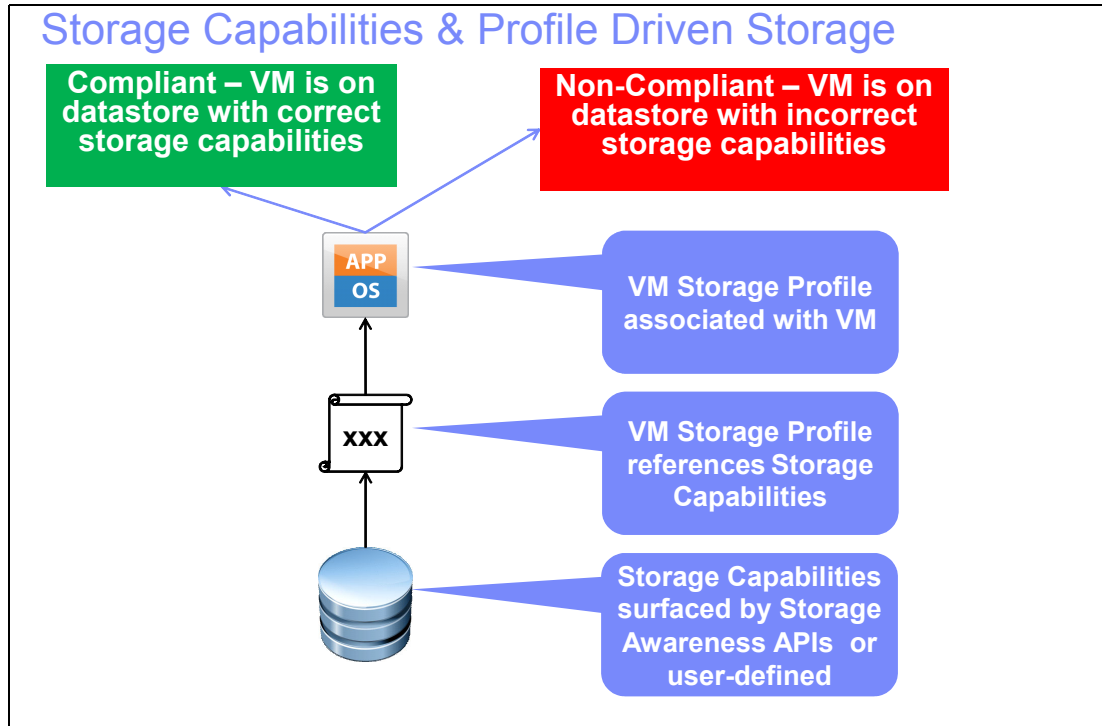


Figure 2-2 Conceptual View of VASA and PDS Integration

IBM Storage Provider for VMware VASA

The IBM Storage Provider for VMware VASA, referred to as “VASA Provider,” improves the ability to monitor and automate storage-related operations in VMware vSphere.

The VASA Provider supports multiple vCenter consoles and multiple XIV Storage Systems, and constitutes an alternative tool for viewing information about the XIV Storage System within vCenter, including:

- ▶ Real-time disk status
- ▶ Real-time alerts and events

Installed on Microsoft Windows Server, the VASA provider functions as a standard vSphere management plug-in for each VMware vCenter server while interacting with vStorage APIs for Storage Awareness (VASA) to deliver information about storage topology, capabilities and state, and events and alerts to VMware vSphere.

For the most recent version of the IBM Storage Provider for VMware VASA and information about installation prerequisites and instructions, refer to:

http://www-933.ibm.com/support/fixcentral/swg/selectFixes?parent=ibm~Storage_Disk&product=ibm/Storage_Disk/XIV+Storage+System+%282810,+2812%29&release=All&platform=All&function=all#vSphere%20APIs%20for%20Storage%20Awareness

Additional instructions about installing the Vasa Provider are available in *IBM Storage Provider for VMware VASA, Version 1.1.1*, located at:

http://pic.dhe.ibm.com/infocenter/strhosts/ic/topic/com.ibm.help.strhosts.doc/PDFs/Storage_Prov_for_VMware_VASA_1.1.1_IG.pdf

After installing and setting up the VASA provider, you can examine the attributes reported to vSphere:

1. Select your VMware ESXi cluster within the vCenter client.
2. Navigate to the Storage Views tab, and select the **Show all SCSI Volumes (LUNs)** view. As shown in Figure 2-3, there are four columns that contain information from VASA:

Committed Capacity This is the hard, or physical, capacity consumed on the XIV Storage Array itself, expressed in binary metrics, for example GiB. Both administrators and SDRS can monitor this metric when making VM/VMDK deployment, cloning, and migration decisions.

Thin Provisioned Specifies a value of “true” or “false.” The example in Figure 9-11 shows a value of “true” for all volumes because XIV thin provisions volumes by virtue of its architecture. It is important to distinguish this metric from the thin provisioning status of the storage pool containing the volume. The value specified does not reflect the provisioning status of the associated storage pool.

Storage Array Provides the name of the storage subsystem.

Identifier on Array Correlates the SCSI ID of a LUN with the actual volume name on the associated XIV Storage System.

SCSI ID	Lun	Datastore	Capacity	Committed	Thin Provisioned	Storage Array	Identifier on Array	Vendor	Volume Name
010001000036...	1	FCM_Primary_Te...	192.00 GB	170.09 GB	true	XIV_01_6000105	FCM_1	IBM	IBM Fibre Ch
010004000036...	4	SRM_Placeholder	16.00 GB	90.00 MB	true	XIV_01_6000105	SRM_Placeholder	IBM	IBM Fibre Ch
020000000060...	0	datastore1	45.63 GB					LSILOGIC	LSILOGIC Se

Figure 2-3 IBM VASA Provider Surfacing XIV Storage Characteristics to vSphere Client

In addition to surfacing LUN attributes to vCenter, VASA also surfaces information about the utilization status of thinly-provisioned LUNs that can trigger warning alarms indicating the impending depletion of available physical capacity. As depicted in the example in Figure 2-4 on page 15, the name of this alarm is “Thin-provisioned volume capacity threshold exceeded,” and it appears under both the Tasks & Events and Alarms tabs in the Datastores and Datastore Clusters interface of the vSphere client. In the example, this alarm is triggered because the FCM_1 LUN backing the FCM_Primary_Test datastore reached a committed capacity exceeding 75% of the provisioned capacity, which is the default threshold.

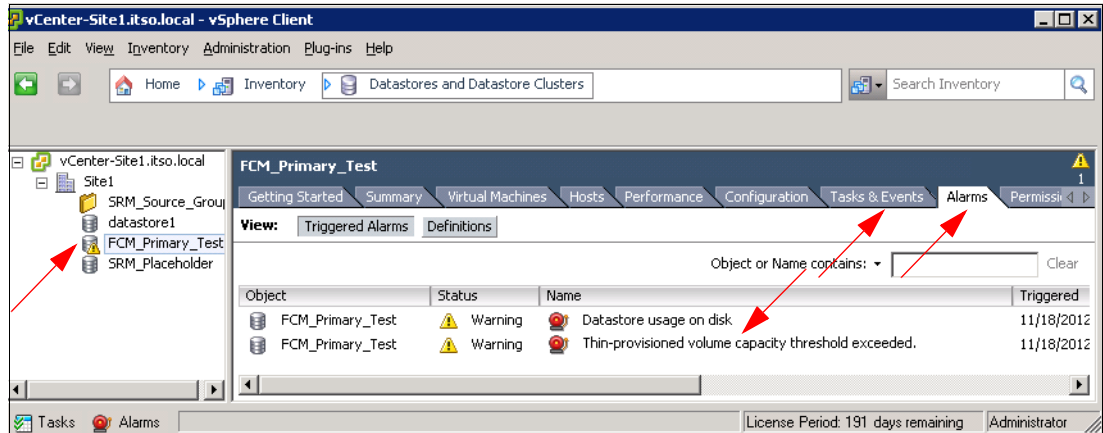


Figure 2-4 IBM VASA Provider Surfacing XIV “Datastore Usage on Disk” Alerts to vCenter Client

Version 1.1.1 of IBM Storage Provider for VMware VASA does not support surfacing capabilities of the XIV Storage System’s LUNs to vSphere client.

2.1.3 XIV and VMware LUN and datastore provisioning best practices

With a key benefit of virtualized infrastructure being server consolidation, the VAAI-capable XIV infrastructure enables customers to apply VMware best practices without compromise, improving VM consolidation without incurring latency or increasing risk. With VAAI-enabled XIV storage, customers can realize the following benefits:

- ▶ Uses larger LUNs to create or expand datastores
- ▶ Uses fewer LUNs to create larger datastores, simplifying storage management
- ▶ Increases the number of VMs in a datastore
- ▶ Potentially increases the number of VMs running on a host or cluster
- ▶ Migrates VMs between data stores without host impact
- ▶ Copies VMs and create templates without host impact
- ▶ Creates new VMs without host impact

In a vSphere environment leveraging XIV storage, the following LUN and datastore sizing recommendations apply:

- ▶ Use large LUNs. Most use cases call for LUNs with a minimum of 1 TB to a maximum of 3.5 TB in size. When selecting an optimal LUN size, consider both of the following caveats and benefits associated with larger LUNs:
 - Factor in the amount of time required for tasks, such as LUN migration, as needed, in the context of business objectives.
 - Consider the capacity overhead resulting from the desired number and life expectancy of concurrent XIV or VMware snapshots. In general, larger LUNs incur a disproportionately larger capacity consumption compared to smaller LUNs as a result of necessitating adherence to the VMDKs with the minimum recovery point objective, increased metadata overhead, and so on.
 - Understand the business implications to failure domains, recovery time objectives, and recovery point objectives as a function of LUN size, particularly for extremely large LUNs that must be engaged in persistent snapshot or remote mirroring relationships.

- The benefits of larger LUNs include:
 - Less orphaned space, which improves both ease of management and overall capacity utilization at the vSphere level.
 - Improved XIV caching for larger LUNs.
- ▶ Avoid using few large volumes to balance workloads and improve performance by increasing queue utilization:
 - Take advantage of all ports on the XIV system. Connectivity and port configuration best practices are described in “XIV and VMware general connectivity best practices” on page 16.
 - Try to balance workload across at least 8 LUNs.
- ▶ LVM extents are supported but not recommended. In vSphere 5.0 with XIV, you can increase the size of a datastore and LUN (up to 64TB) with no disruption of service.
 - vSphere 5.0 drastically increased size of support for datastores (64TB); however, this is not to say that creating LUNs or datastores of this size is a recommended implementation practice.
 - As long as storage pool has capacity/performance, large datastores can be used.
 - Concerns with SCSI-2 reservations to enforce LUN locking are mitigated with VAAI, as described in Chapter 3, “VMware vStorage APIs Array Integration” on page 27.

2.1.4 XIV and VMware general connectivity best practices

When implementing Fibre channel connectivity for the XIV Storage System in a vSphere environment, the configuration must adhere to the following practices:

- ▶ Utilize XIV host cluster groups for LUN assignment
- ▶ Configure single initiator zones
- ▶ At the time of this writing, the VMware specifies that there can be a maximum of 1024 paths and 256 LUNs per ESX/ESXi host, as shown in Table 2-1 on page 17. The following conditions must be simultaneously satisfied to achieve the optimal storage configuration:
 - Effectively balance paths across:
 - Host HBA ports
 - XIV Interface Modules
 - Ensure that the desired minimum number of host paths per LUN and the desired minimum number of LUNs per host can be simultaneously met.
- ▶ Configure the Path Selection Plug-in (PSP) multipathing based on vSphere version:
 - Use Round Robin policy if vSphere version is vSphere 4.0 or higher.
 - Use Fixed Path policy if vSphere version is lower than vSphere 4.0.
 - Do *not* use the Most Recently Used (MRU) policy.

Refer to Figure 2-5 on page 17 and Figure 2-6 on page 18 for suggested configurations to satisfy these criteria.

When implementing iSCSI connectivity for the XIV Storage Systems in a vSphere environment, the configuration must adhere to the following practices:

- ▶ One VMkernel port group per physical Network Interface Card (NIC):
 - VMkernel port is bound to physical NIC port in vSwitch creating a “path”
 - Creates 1-to-1 path for VMware NMP
 - Utilize same PSP as for FC connectivity

- ▶ Enable jumbo frames for throughput intensive workloads (must be done at all layers).
- ▶ Use Round Robin PSP to enable load balancing across all XIV Interface Modules:
 - Each initiator must see a target port on each module.
- ▶ Queue depth can also be changed on the iSCSI software initiator:
 - If more bandwidth is needed, the LUN queue depth can be modified.

Table 2-1 Notable Storage Maximums in vSphere 5.0/5.1

Storage Element Limit	Maximum
Virtual Disk Size	2TB minus 512 bytes
Virtual Disks per Host	2048
LUNs per Host	256
Total Number of Paths per Host	1024
Total Number of Paths to per LUN	32
LUN Size	64TB
Concurrent Storage vMotions per Datastore	8
Concurrent Storage vMotions per Host	2

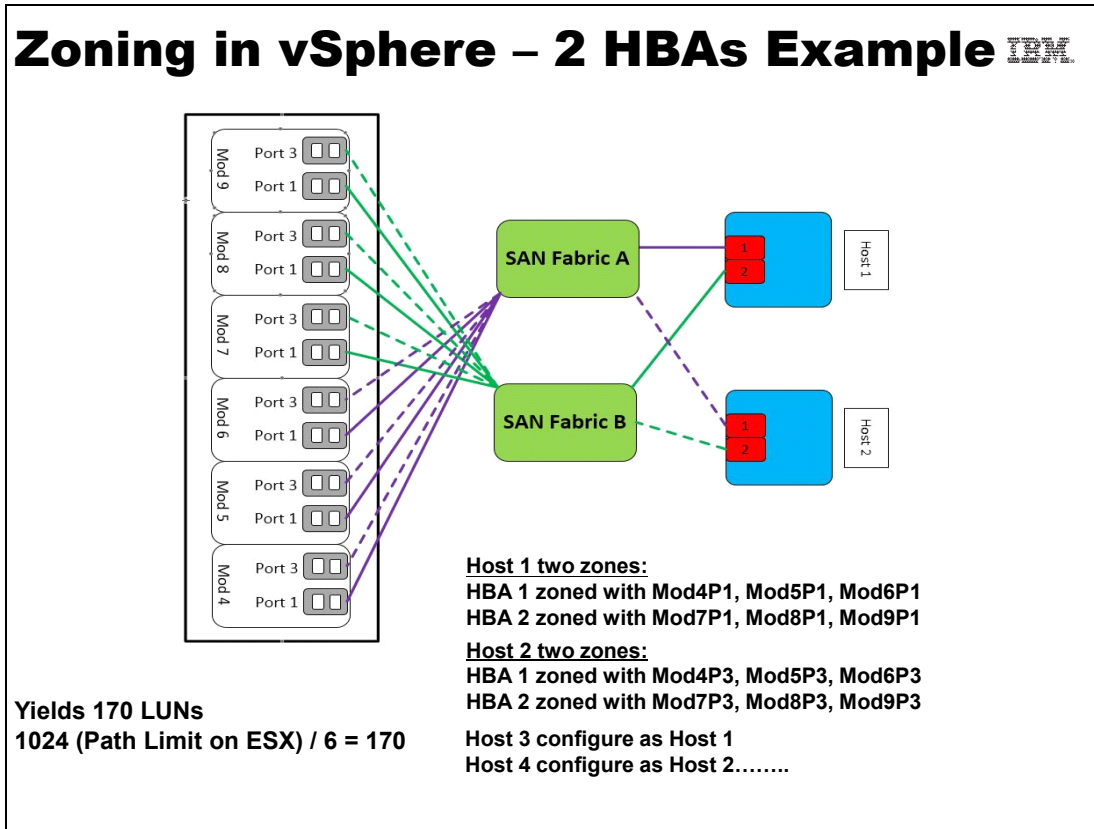


Figure 2-5 Zoning Best Practices - 2 HBAs per ESX Host

Zoning in vSphere – 4 HBAs Example

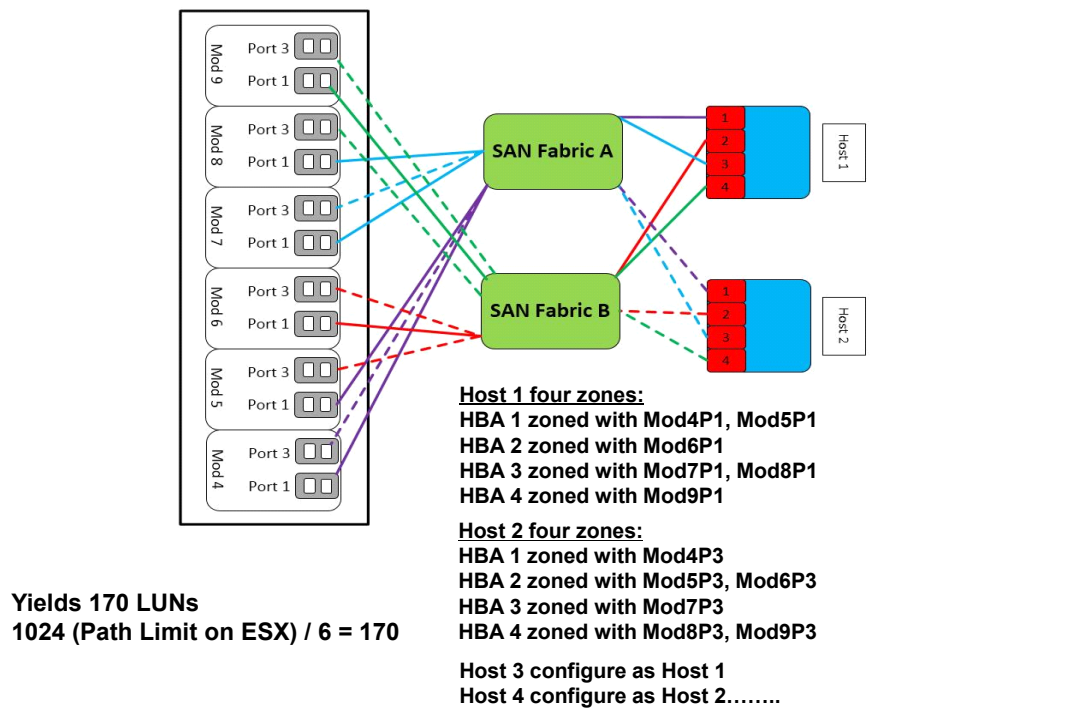


Figure 2-6 Zoning Best Practices - 4 HBAs per ESX Host

2.1.5 XIV and VMware thin provisioning

In this section, the topics of thin provisioning at both the XIV Storage and at the VMware VMFS level are presented conceptually, followed by an examination of best practices necessary to most effectively combine the benefits of thin provisioning in both XIV and VMFS.

XIV thin provisioning conceptual overview

Thin provisioning is a method for optimizing storage utilization by allocating to a volume only the space required to hold its data, deferring space allocation to the time it is actually needed. Fundamentally, thinly provisioned LUNs are unique in that they report a capacity to a host that is not matched by the physical capacity backing the LUN, with the result that the thin provisioning status of the LUN is transparent to the host.

To demonstrate the motivation for thinly provisioned storage, consider a simple example: As a result of the rapid growth of data, a newly deployed application exceeds a capacity utilization threshold that triggers an alert to the storage administrator to plan for expanding the available capacity. Because there is no history of data growth for the new application, the storage administrator must make a decision that involves weighing the potentially complex and time-consuming process of iteratively deploying storage capacity on an “as needed” basis, versus grossly over-provisioning the capacity to minimize effort and risk at the penalty of higher cost and the utilization inefficiencies incurred by deploying large “siloes” storage pools. At a high level, thin provisioning addresses this issue by simplifying the management of storage capacity, reducing cost, and optimizing the utilization of available capacity in the context of a more broadly shared resource pool. Because the capacity that is presented to the server can be larger than the actual capacity (physical capacity) consumed by the server in the shared storage pool, the storage administrator can assign larger thin provisioned volumes

to the server and add the physical capacity only whenever it is necessary. As you can see, there are significant benefits to flexibility and efficiency of deployment when provisioning of capacity at the host-level is decoupled from the provisioning of physical capacity.

Using XIV thin provisioning

The decision to utilize thin provisioning is made at the storage pool level, either regular or thinly provisioned. All volumes in a given storage pool inherit the storage pool type (*regular* or *thin*)

For thin pool, the administrator specifies:

- Pool soft size
- Pool hard size
- Additional parameters specifying behavior regarding snapshots

Changing a resource from *regular* to *thin* constitutes a simple change in designation:

- The volume type might be changed regular to thin by moving from a regular to a thin storage pool. This is a dynamic, immediate change.
- The storage pool type might be changed from regular to thin. This is a dynamic, immediate change.

The following additional changes to the thinly-provisioned resources are possible and occur dynamically and immediately:

- Change storage pool soft size
- Change storage pool hard size
- Move volume into/out of storage pool

There is *zero* performance impact to these actions because XIV volumes are always written thinly.

This topic is explored in depth in the IBM Redbooks Publication *IBM XIV Storage System Gen3 Architecture, Implementation, and Usage*, SG24-7659.

VMFS thin provisioning conceptual overview

Thin provisioning within the context of the file system follows the same basic principles as thinly provisioned volumes within storage pools, except that the provisioned elements and the associated “container” are now the VMFS files and the datastore, respectively. Because the datastore itself might be backed by a thinly provisioned LUN, one more layer of abstraction was added, as has one more opportunity to over-commit real capacity, in this case to the VMs themselves. The following three format options exist for creating a virtual disk within the VMFS file system:

- ▶ Eager Zeroed Thick (EZT): Required for best performance and for VMs classified as Fault Tolerant:
 - Space is reserved in datastore, meaning unused space in the VMDK might *not* be used for other VMDKs in the same datastore.
 - A VMDK is not available until formatted with zeroes, either as a metadata representation in the case of the XIV Storage System or by physically writing to disk in case of storage systems that do not flag this type of activity.
 - With the VAAI WRITE_SAME (Zero Blocks) primitive, the process of zeroing the VMDK is off-loaded to the storage subsystem. This is discussed in further detail in Chapter 3, “VMware vStorage APIs Array Integration” on page 27.

- ▶ Lazy Zeroed Thick (LZT):
 - Unused space in VMDK might *not* be used for other VMDKs in the same datastore.
 - The VMDK is immediately available upon creation. The VMkernel attempts to dynamically initiate the allocation of physical capacity within the storage pool by pre-emptively writing zeroes to the LUN for each VM-generated write targeting new blocks. This is the default provisioning type.
- ▶ Thin:
 - Unused space in VMDK can be used for other VMDKs in the same datastore, which adds another threshold that must be carefully monitored to prevent service interruption as a result of the VMs sharing the datastore and collectively consuming all of the LUN capacity backing the datastore:
 - This is possible because the specified VMDK size represents the *provisioned* size, which is what is presented to the VM itself. However, only the *used* size, or *hard* size in XIV terms, is what is actually subtracted from the datastore's capacity.
 - The capacity utilization percentage at the datastore-level is based on the blocksize and the data previously written for each VMDK co-resident in the datastore.
 - Like the LZT provisioning option, the VMDK is immediately available upon creation. The VMkernel attempts to dynamically initiate the allocation of physical capacity within the storage pool by pre-emptively writing zeroes to the LUN for each VM-generated write targeting new blocks.

Using VMFS thin provisioning

When considering whether to thinly provision VMDKs within a VMFS datastore, weigh the following advantages and disadvantages specific to the vSphere environment being implemented.

Advantages:

- ▶ Unless the administrator effectively synchronizes capacity reclamation between VMFS datastores and the associated LUNs on the XIV Storage System, which is a manual process at the time of this writing, the potential to exploit thin provisioning efficiency at the VMFS level might exceed the thin provisioning efficiency that is possible over time within the XIV Storage System. This is because the VMFS is aware of data that moved or deleted, while the same capacity remains consumed within the XIV Storage System until capacity reclamation can occur. However, if real capacity consumption is not properly managed, the potential benefits achievable by over-representing physically-backed capacity to the virtual machines are greatly reduced.
- ▶ Over-provisioned conditions at the VMFS level can be less frequent and generate fewer alerts because fluctuations in VMDK sizes within a datastore and the associated datastore capacity utilization are dynamically reflected in vCenter due to the awareness of the data consumption within the file system.

Disadvantages:

- ▶ For vSphere releases prior to vSphere 4.1, when a thin provisioned disk grows, the ESX host must make a SCSI reservation to serialize access to an entire LUN backing the datastore. Therefore, the viability of dense VM multi-tenancy is reduced because implementing thinly provisioned VMDKs to increase multi-tenancy incurs the penalty of reducing potential performance by increasing congestion and latency.

- ▶ Compared to storage pool-based thin provisioning within the XIV Storage System, thin provisioning at the VMDK-level has the following drawbacks:
 - There are more objects to monitor and manage because the VMDKs are thinly provisioned; therefore, they must be monitored in conjunction with co-resident VMDKs in the datastore. Furthermore, this must be done for all datastores. In contrast, thin provisioning resource management can be better consolidated at the level of the XIV Storage System, thus providing a global awareness of soft versus hard capacity consumption and facilitating ease of management activities including physical capacity deployment where it really matters—in the storage subsystem itself.
 - Consider the scenario of balancing physical, or hard, capacity among a group of datastores backed by LUNs within a storage pool whose hard capacity cannot be expanded, for example by decreasing the size of a LUN associated with a given datastore in favor of increasing the size of a LUN deemed to have priority. Redistributing capacity among datastores is possible, but cannot be accomplished as a single operation in vSphere as of the time of this writing.

In contrast, by managing the capacity trade-offs among datastores at the XIV level, it is trivial to expand the soft size of both the LUN and the storage pool. The net effect is that the LUN(s) backing the datastore that needs more hard capacity can now effectively borrow that capacity from the pool of unused hard capacity associated collectively with all of the LUNs in the storage pool without the need to contract the soft size of any LUNs. Obviously, if 100% of the physical capacity in the storage pool is already consumed, this requires a coordinated expansion of capacity of the datastore, LUN, and finally the physical capacity in the storage pool. If hard capacity is available in the system, the latter can be accomplished within seconds due to the ease of management in the XIV Storage System; otherwise, it will still necessitate deployment of new XIV modules. Again, capacity monitoring at all levels is of paramount importance to anticipate this condition.

- The scope of potential capacity utilization efficiency is relatively small at the individual datastore level. Leveraging thinly-provisioned LUNs in the XIV Storage System dramatically increases the potential scope of savings by expanding the sphere of capacity provisioning to include all of the datastores co-resident in a storage pool. This is because the potential savings resulting from thin provisioning is effectively proportional to the scale of the capacity pool containing thinly-provisioned resources.

Thin provisioning prerequisites

Successful thin provisioning requires a “thin-friendly” environment at all levels of software in the stack:

- ▶ File system:
 - VMware environments require consideration of the file systems in use by the guest operating systems and the VMFS version.
- ▶ Database
- ▶ Application

Thin-friendly file systems, databases, and applications have the following attributes:

- ▶ Physical locality of data placement: If data is placed randomly across the LUN, the storage system interprets the interspersed free space as being consumed as well.
- ▶ Wherever possible, reuse previously freed-up space: Writes are issued to previously used and subsequently deleted space before being issued to “never-used” space.
- ▶ Provision for the file system to communicate deleted space to the storage subsystem for reclamation.

If these properties are not pervasive across these elements, implementation of thin provisioning might have little benefit and might even incur additional penalties compared to regular provisioning.

In addition, be aware that the following user options and activities might affect the success of thin provisioning:

- ▶ LUN format options.
- ▶ Defrag processes: Swapping algorithms can defeat thin provisioning by touching unused space.
- ▶ “Zero file” utilities can enable space reclamation for storage systems with zero detect or scrubbing capabilities.

Thin provisioning general guidelines

Consider the following guidelines:

1. Ensure that the following classifications of applications are not included as candidates for thin provisioning:
 - Applications that are not thin-friendly
 - Applications that are extremely risk-averse
 - In terms of general storage best practices, highest transaction applications must be excluded from consideration for thin provisioning. However, the sophisticated data distribution characteristics of the XIV Storage System are designed with high transaction applications in mind, so thin provisioning can be effectively utilized for an expansive set of applications.
2. Automate monitoring, reporting, and notifications, and set thresholds according to how quickly your business can respond.
3. Plan procedures in advance for adding space, and decide whether to automate them.
4. Use VAAI and the latest version of VMFS:
 - VAAI ATS primitive limits impact of SCSI2 reservations when thin provisioning is used.
 - Improves performance.

Thin on thin?

In general, the choice of provisioning mode for a given VMDK and datastore combination spans six possibilities determined by three choices at the VMware VMDK level, including *EagerZeroedThick*, *LazyZeroedThick*, and *Thin*, and the standard choice of thick or thin provisioned LUNs within the storage subsystem itself (for simplicity, assume there is a one-to-one mapping between LUN and datastore).

However, the XIV Storage System distinguishes itself in two respects, both of which promote efficiency, maximize resource utilization, and enhance management simplicity:

- ▶ The EZT and LZT options consume the same physical capacity within an XIV Storage System as a result of the EZT zeroing requiring only a logical, metadata implementation as opposed to a physical one.
- ▶ The XIV Storage System inherently implements thin provisioning for all logical volumes as a result of its efficient architecture.

To demonstrate the unique capabilities of the XIV Storage System, the tables in Figure 2-7 and Figure 2-8 illustrate the potential thin provisioning combinations for both a traditional storage system and the XIV Storage System by examining relative risks and benefits using a simplified example scenario:

- ▶ Two 100GB VMDKs provisioned in each of two VMFS datastores
- ▶ Two 200GB VMFS datastores each backed by a single LUN
- ▶ This yields a total of four identical VMDKs:
 - Each VMDK is assumed to consume 10GB physical capacity
- ▶ Both datastores are co-resident within a single storage pool
- ▶ The VMDK and LUN-level thin versus thick provisioning options and associated risk levels are examined in the context of the minimum VMFS datastore capacity allocation and minimum storage pool capacity allocation.

VMware Layer Provisioning Type	Storage Layer Provisioning Type	VMDK - Logical (GB)	VMDK - Physical (Incl. Zeroes) (GB)	VMFS Datastore - Min. Logical (100% Full) (GB)	Datastore - Physical (GB)	Storage Pool - Min. Logical (100% Full) (GB)	Storage Pool - Min. Physical (GB)	Relative Potential Efficiency	VMFS Datastore Potential RISK (Min. Logical Allocations)	Storage Pool Potential RISK (Min. Physical Allocations)
VMDK - EZT	Thick	100	100	200	200	400	400	Low	Low	Low
	Thin	100	100	200	200	400	400	Medium**	Low	Medium**
VMDK - LZT	Thick	100	10	200	20	400	400	Low	Low	Low
	Thin	100	10	200	20	400	40	High	Low	High*
VMDK - Thin	Thick	100	10	20	20	40	40	Medium	High*	High*
	Thin	100	10	20	20	40	40	High	High*	High*

Figure 2-7 Generalized VMDK/LUN Provisioning risk/benefit table

**If co-resident with other thinly-provisioned LUNs in thinly provisioned storage pool

VMware Layer Provisioning Type	Storage Layer Provisioning Type	VMDK - Logical (GB)	VMDK - Physical (Incl. Zeroes) (GB)	VMFS Datastore - Min. Logical (100% Full) (GB)	Datastore - Physical (GB)	XIV Storage Pool - Min. Logical (100% Full) (GB)	XIV Storage Pool - Min. Physical (GB)	Relative Potential Efficiency	VMFS Datastore Potential RISK (Min. Logical Allocations)	XIV Storage Pool Potential RISK (Min. Physical Allocations)
VMDK - EZT/LZT	Thin	100	10	200	20	400	40	High	Low	High*
VMDK - Thin	Thin	100	10	20	20	40	40	High	High*	High*

Figure 2-8 VMDK/LUN Provisioning Risk/Benefit table for XIV

Each table is intended to be used for relative comparison purposes only, and does NOT imply endorsement of implementing the minimum necessary datastore capacity allocation or the minimum necessary physical allocation of storage pool capacity.

As described in Figure 2-7 and Figure 2-8, *risk* is mitigated by:

- ▶ Managing oversubscription with reasonable and acceptable headroom above minimum logical/physical capacity thresholds.
- ▶ Implementing VMware and storage-based alerting at appropriate utilization thresholds
- ▶ Effectively exploiting VAAI integration
- ▶ Using oversubscription *only* with thin-friendly applications and guest operating systems

Based on Figure 2-7 on page 23 and Figure 2-8 on page 23, the following conclusions are drawn to highlight relative differences among thin provisioning options available in a vSphere environment. Generally, the recommendation is against using thin-provisioned VMDKs as a result of the risk against potential gains. The following, however, are universally true:

- ▶ Thinly-provisioned storage pools inherently increase both potential risk and potential benefits. Risk can be mitigated by setting proper monitoring and alert thresholds, and in the case of vSphere, harnessing VAAI integration.
- ▶ Exploit thin provisioning exclusively for platforms, databases, and applications that are known to be “thin friendly.”

For a traditional storage system, the example in the table demonstrates that the optimal combination consists of LZT-provisioned VMDKs within the VMFS, along with thin-provisioned LUNs within the storage system. The caveat is that both the datastores and the storage pools must be carefully monitored and expanded as necessary in advance of VMDK, datastore, and storage pool growth needs.

For the XIV Storage System, the example in the table demonstrates that the optimal combination exploits either EZT or LZT-provisioned VMDKs. This is because EZT/LZT consumes 100% of the specified provisioned capacity within the VMFS datastore, eliminating the need to monitor for the possibility of VM's attempting to write to capacity that has not been provisioned at the VMFS level. At the same time, the same potential capacity utilization benefits can still be realized within the scope of the storage pools, which is the ultimate objective as this layer is where physical capacity is consumed.

In conclusion, the complexity of the joint decision pertaining to the potential thin or thick provisioning combinations is greatly simplified for the XIV Storage System compared to traditional storage systems, as is the subsequent management of the environment.

2.1.6 XIV and VMware best practices for Quality of Service

Deploying VMware datastores on the XIV Storage System complements the functionality of VMware's native Storage I/O Control capability by extending the scope and diversity of storage-centric Quality of Service (QoS) controls available to administrators to meet Service Level Agreements in the vSphere environment. By combining the unique attributes of both XIV and VMware QoS enforcement utilities, the vSphere infrastructure can benefit from improved density of VM multi-tenancy on shared storage resources simultaneously for a diverse range of I/O workloads while maintaining performance thresholds for high-priority processes.

XIV Host-Mapped QoS

With the simple yet effective XIV QoS feature, administrators can specify storage performance thresholds to limit I/O resources at the granularity of individual ESX/ESXi hosts, thereby ensuring service levels and prioritized access for the hosts that are running business-critical VMs and applications. XIV QoS enforces performance prioritization using the two key performance metrics of I/O transactions per second and total bandwidth individually or in parallel.

Administrators can place hosts in one of four QoS groups using a construct called a *Performance Class*. Each group is configured to limit its hosts by an aggregate maximum of IOPS and an aggregate maximum bandwidth (MB per second). Unassigned hosts remain unlimited. This topic is explored in depth in the IBM Redbooks Publication *IBM XIV Storage System Gen3 Architecture, Implementation, and Usage*, SG24-7659.

vSphere Storage I/O Control (SIOC)

Storage I/O Control in vSphere 5.1 enforces preferential access to I/O resources for virtual machines running business critical processes during periods of congestion by throttling I/O access for non-critical VMs occupying shared hosts and datastores. SIOC operates using the concept of shares and limits at both the host and datastore level to distribute I/O resources in a dynamic, relative fashion. SIOC leverages the existing host device queue to control and manage I/O prioritization. Shares translate into ESX/ESXi I/O queue slots to provide a proportion of the total queue slots collectively assigned to the ESX/ESXi host cluster.

I/O queue slot assignment is dynamically determined based on the combination of VM shares and current load:

- ▶ The proportion of I/O queue slots assigned to each host is a function of both the current load and the mixture of current VMs running on the host and the associated shares assigned to them.
- ▶ The maximum number of I/O queue slots that can be used by the virtual machines on a given host cannot exceed the maximum device queue depth for the associated ESX/ESXi host.

In this way, VMs with more shares are allowed to send more I/O transactions concurrently, as are hosts with relatively larger proportions of VM shares.

Because limitations are enforced relative to the dynamic performance capability of the shared storage and host resource instead of using static I/O thresholds, SIOC maintains fairness without leaving I/O resources unused or using unnecessary throttling.

SIOC benefits include:

- ▶ Preventing resource starvation for priority VMs by guaranteeing preferential access dynamically on the basis of storage resource congestion, which is determined by I/O latency monitoring
- ▶ Eliminating the “noisy neighbor” problem and provides proportional fairness of access to shared storage, thus improving the viability of increased VM deployment density
- ▶ Providing granular SLA settings for network traffic

For additional information about SIOC, refer to the white paper at:

<http://www.vmware.com/files/pdf/techpaper/VMW-vSphere41-SIOC.pdf>



VMware vStorage APIs Array Integration

ESX/ESXi 4.1 brought a new level of integration with storage systems through the introduction of vStorage API Array Integration (VAAI), which we discuss in this chapter.

3.1 VAAI overview

VAAI helps reduce host resource utilization overhead while executing common vSphere operations. It also increases scalability and operational performance by offloading certain storage-centric tasks to storage systems that support the relevant commands. In contrast, traditional SCSI commands force the ESX host to issue many repetitive commands to complete certain types of operations. These operations include cloning a virtual machine and creating a new virtual machine with the Thick Provision Eager Zeroed option. For example, the zeroed option writes zeros across the new virtual disk. Using VAAI, the same task can be accomplished with far less effort on the part of the ESX/ESXi server.

IBM XIV with the correct firmware release supports the following T10-compliant SCSI commands (also called primitives) to achieve this new level of integration.

► Hardware Accelerated Move

Hardware Accelerated Move, also known as FULL COPY or XCOPY, offloads copy operations from VMware ESX to the IBM XIV Storage System. This process allows for rapid movement of data when performing copy, move, and VMware snapshot operations within the IBM XIV Storage System. It reduces the processor and HBA workload of the ESX server. Similarly, it reduces the volume of traffic moving through the SAN when performing VM deployment. It does so by synchronizing individual VM level or file system operations, including clone, migration, and snapshot activities, with the physical storage level operations at the granularity of individual blocks on the device(s). The potential scope in the context of the storage is both *within and across* LUNs. This command has the following benefits:

- Expedites copy operations including:
 - Cloning of Virtual Machines
 - Migrating Virtual Machines from one datastore to another (vMotion)
 - Provisioning from template

- Minimizes host processing/resource allocation

Copies data from one XIV LUN to another without reading/writing through the ESX server and network

- Reduces SAN traffic

It is important to note that the Hardware Accelerated Move primitive is utilized by vSphere only when the source and target LUNs are on the same XIV Storage System. For the remaining cases, vSphere implements a standard host-centric data movement process. In this case, the implication is that the SAN, the source and target host(s), and in most cases the network are all again in-band. Figure 3-1 on page 29 provides a conceptual illustration contrasting a copy operation both with and without hardware acceleration.

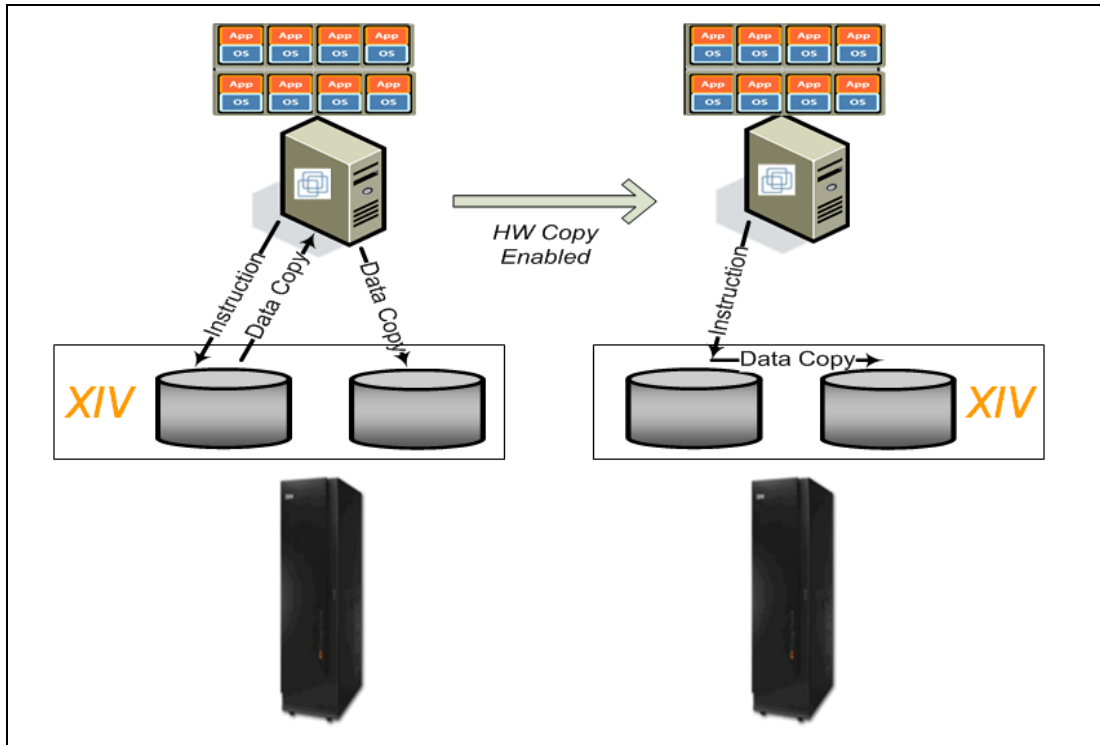


Figure 3-1 Hardware Accelerated Move VAAI

► Hardware Accelerated Initialization

Hardware Accelerated Initialization, or Block Zeroing, exploits the WRITE_SAME command to issue a chain of identical write transactions to the storage system, thus almost entirely eliminating server processor and memory utilization by eliminating the need for the host to execute repetitive identical write transactions. It also reduces the volume of host HBA and SAN traffic when performing repetitive block-level write operations within virtual machine disks to the IBM XIV Storage System. Similarly, it allows the XIV Storage System to minimize its own internal bandwidth consumption. For example, when provisioning a VMDK file with the *eagerzeroedthick* specification, the Zero Block's primitive issues a single WRITE_SAME command that replicates zeroes across the capacity range represented by the difference between the VMDK's provisioned capacity and the capacity consumed by actual data. The alternative requires the ESX host to issue individual writes to fill the VMDK file with zeroes.

The XIV Storage System further augments this benefit by flagging the capacity as having been "zeroed" in metadata without the requirement to physically write zeros to the cache and the disk. The scope of the Zero Blocks primitive is the VMDK creation within a VMFS datastore, and therefore the scope of the primitive is generally within a single LUN on the storage subsystem, but can possibly span LUNs backing multi-extent datastores.

In summary, *Hardware Accelerated Initialization* offers the following benefits:

- Offloads initial formatting of Eager Zero Thick (EZT) VMDKs to XIV
- Assigns zeroes to large areas of storage without writing zeroes from the ESX server
- Speeds creation of new Virtual Machines – EZT VMDKs available immediately
- Reduces elapsed time, server workload, and network workload

Figure 3-2 provides a conceptual illustration contrasting the deployment of an eagerzeroedthick VMDK both with and without Hardware Accelerated Initialization.

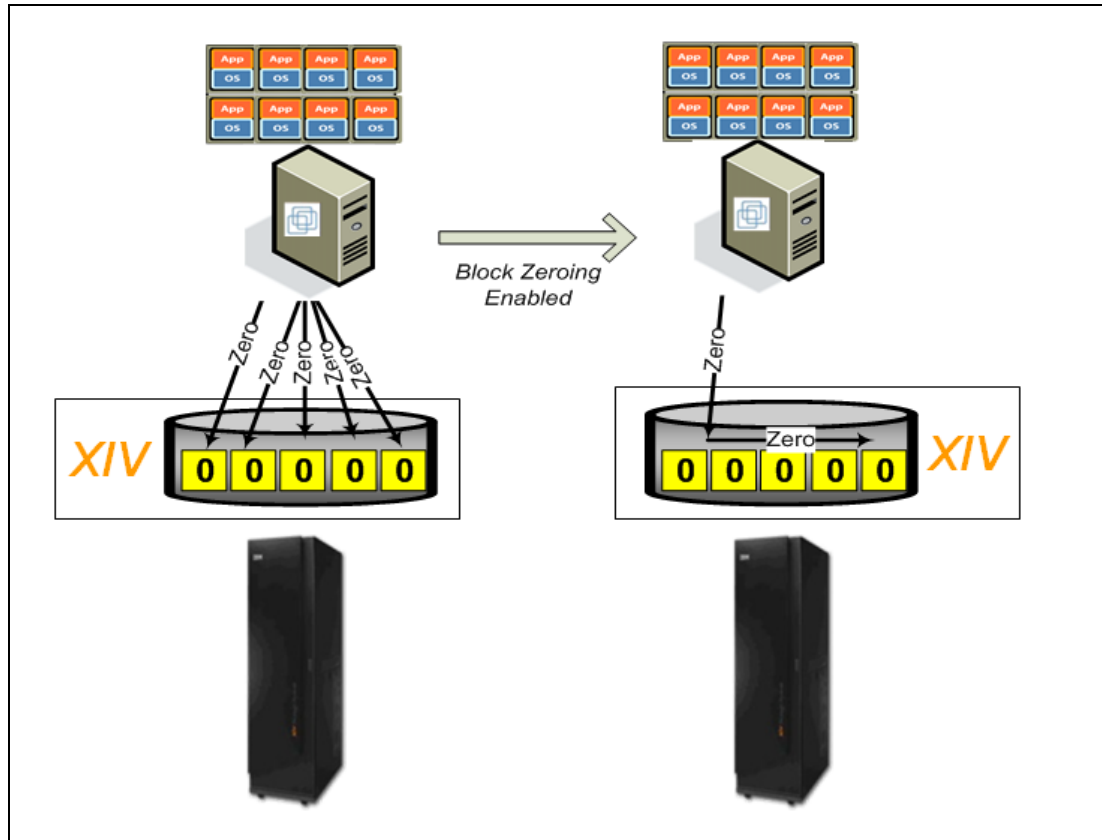


Figure 3-2 Hardware Accelerated Initialization VAAI

► Hardware Assisted Locking

Hardware Assisted Locking, also known as Atomic Test & Set or ATS, intelligently relegates resource access serialization down to the granularity of the block level during VMware metadata updates. It does this instead of using a mature SCSI2 reserve, which serializes access to adjacent ESX hosts with a minimum scope of an entire LUN. An important note is that the VMware File System (VMFS version 3 or higher) exploits ATS in a multi-tenant ESX cluster that shares capacity within a VMFS datastore by serializing access only to the VMFS metadata associated with the VMDK or file update needed through an on-disk locking mechanism. As a result, the functionality of ATS is identical whether implemented to grant exclusive access to a VMDK, another file, or even a Raw Device Mapping (RDM). The ATS primitive has the following advantages, which are obvious in enterprise environments where LUNs are used by multiple applications or processes at one time. In summary, Hardware Assisted Locking offers the following benefits:

- Significantly reduces SCSI reservation contentions by locking a range of blocks within a LUN rather than issuing a SCSI reservation on the entire LUN.
- Enables parallel storage processing.
- Reduces latency for multiple ESX servers accessing the same LUN during common vSphere operations involving VMFS metadata updates, including:
 - VM/VMDK/template creation or deletion
 - VM Snapshot creation/deletion

- Virtual Machine migration and Storage vMotion migration (including when invoked by Distributed Resource Scheduler)
 - Virtual Machine Power on/off
- Increases cluster scalability by greatly extending the number of ESX/ESXi hosts and VMs that can viably co-reside on a VMFS datastore.

Note: The currently implemented VMFS version(s) and the history of VMFS version deployment within a vSphere environment have important implications in the context of the scope that VMFS activities exploit the ATS primitive. More information about this topic is available at the following site:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1021976

Figure 3-3 provides a conceptual illustration contrasting the scope of serialization of access both with and without Hardware Assisted Locking.

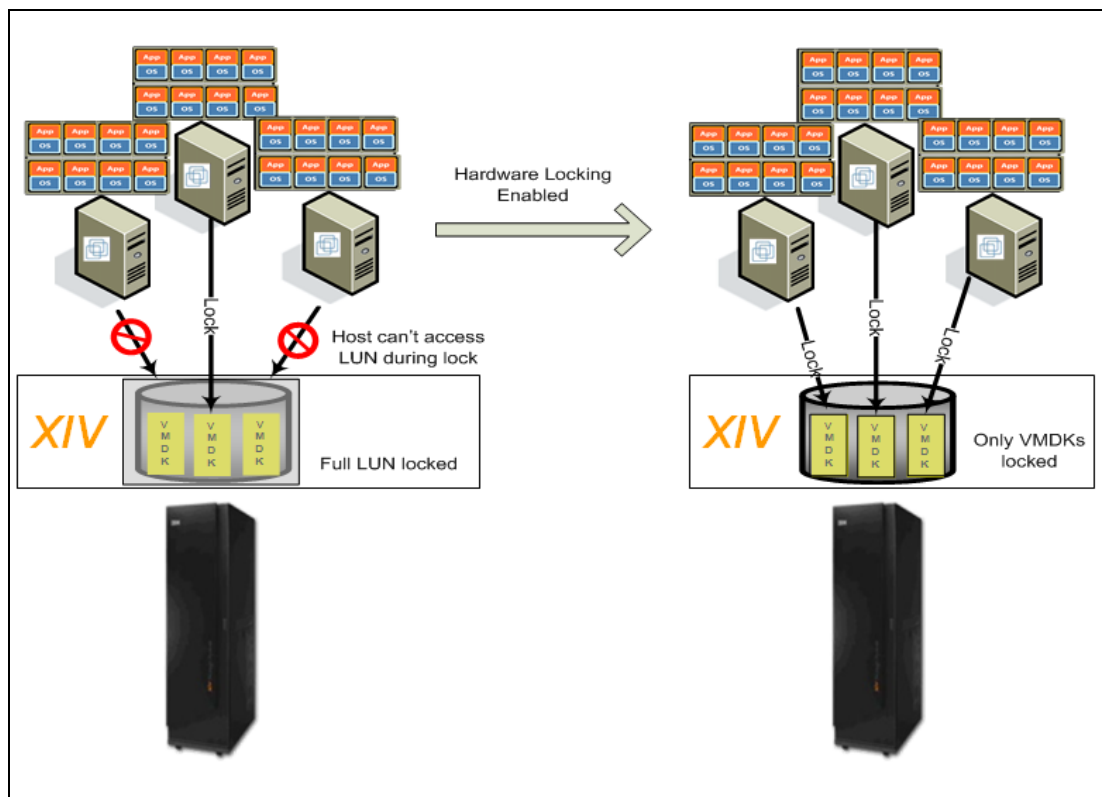


Figure 3-3 Hardware Assisted Locking VAAI

3.1.1 Software prerequisites to use VAAI

The VMware Hardware Compatibility List shows that XIV code version 10.2.4 is required for VAAI support. However, IBM requires that 2nd Generation XIVs run code version 10.2.4a or higher.

XIV Gen 3 needs to be running release 11.0.a or higher, as shown in Table 3-1.

Table 3-1 VAAI support with XIV

	vSphere 4.1	vSphere 5.0
2nd Generation XIV	10.2.4a	10.2.4a
Gen 3 XIV	11.0.a	11.0.a
IBM VAAI plugin	1.1.0.1 or higher	Not required

With ESX/ESXi 4.1, you must install an IBM supplied plug-in on each vSphere server. The initial release was version 1.1.0.1, which supported the XIV. IBM then released version 1.2.0.0, which added support for Storwize V7000 and SAN Volume Controller.

The IBM Storage Device Driver for VMware, the release notes, and the Installation Guide can be downloaded at:

[http://www-933.ibm.com/support/fixcentral/swg/selectFixes?parent=ibm/Storage_Disk&product=ibm/Storage_Disk/XIV+Storage+System+\(2810,+2812\)&release=All&platform=All&function=all](http://www-933.ibm.com/support/fixcentral/swg/selectFixes?parent=ibm/Storage_Disk&product=ibm/Storage_Disk/XIV+Storage+System+(2810,+2812)&release=All&platform=All&function=all)

With vSphere 5.0, you do not need to install a vendor-supplied driver to enable VAAI. It is supported natively.

3.1.2 Installing the IBM VAAI device driver on an ESXi 4.1 server

The IBM Storage device driver for VMware VAAI is a kernel module that allows the VMware VAAI driver to offload certain storage operations to the storage hardware. In this example, they are offloaded to an XIV. The driver needs to be installed on every ESX/ESXi 4.1 server and requires that each server is restarted after installation. Updates to the IBM Storage driver also require that each ESX/ESXi server is rebooted. When combined with server vMotion and vSphere server redundancy, this process usually does not require any guest host outages.

IBM has so far released two versions of the driver that are named as follows:

Version 1.1.0.1 IBM-ibm_vaaip_module-268846-offline_bundle-395553.zip

Version 1.2.0.0 IBM-ibm_vaaip_module-268846-offline_bundle-406056.zip

To confirm if the driver is already installed, use the `vihostupdate.pl` command with the `-query` parameter, as shown in Example 3-1 on page 33. In this example, a version of the driver is already installed. Because only the first 25 characters of the name are shown, it is not clear if it is 1.1.0.1 or 1.2.0.0. If performing this command in the Tech Support Mode shell, use the `esxupdate query` command.

Example 3-1 Checking for the IBM storage device driver

```
vihostupdate.pl --server 9.155.113.136 --username root --password password -query
```

```
-----Bulletin ID-----  -----Installed-----  -----Summary-----  
ESXi410-201101223-UG          2011-01-13T05:09:39 3w-9xxx: scsi driver for VMware ESXi  
ESXi410-201101224-UG          2011-01-13T05:09:39 vxge: net driver for VMware ESXi  
IBM-ibm_vaaip_module-268846  2011-09-15T12:26:51 vmware-esx-ibm-vaaip-module: ESX release
```

Tip: This section involves patching ESXi using the `esxcli`. You can also use the Tech Support mode shell. If you are unsure how to use the shell, consult the plug-in Installation Guide and the following document on the VMware website:

http://www.vmware.com/pdf/vsphere4/r41/vsp_41_esxupdate.pdf

If the driver is already installed and you downloaded the latest version, use the `-scan -bundle` command against the downloaded compressed file. This procedure checks whether you have an older version of the driver. In Example 3-2, the bundle is not installed, indicating that either no driver is installed or only the older version of the driver is installed.

Example 3-2 Checking if the driver is installed or is not at the latest level

```
vihostupdate.pl --server 9.155.113.136 --username root --password password -scan -bundle  
IBM-ibm_vaaip_module-268846-offline_bundle-406056.zip
```

The bulletins which apply to but are not yet installed on this ESX host are listed.

```
-----Bulletin ID-----  -----Summary-----  
IBM-ibm_vaaip_module-268846  vmware-esx-ibm-vaaip-module: ESX release
```

To perform the upgrade or install the driver for the first time, use server vMotion to move all guest operating systems off the server you are upgrading. Install the new driver, place the server in maintenance mode, and reboot it as shown in Example 3-3.

Example 3-3 Installing and then rebooting after installing the new VAAI driver

```
vihostupdate.pl --server 9.155.113.136 --username root --password password --install -bundle  
IBM-ibm_vaaip_module-268846-offline_bundle-406056.zip
```

```
vicfg-hostops.pl --server 9.155.113.136 --username root --password password --operation enter  
Host bc-h-15-b5.mainz.de.ibm.com entered into maintenance mode successfully.
```

```
vicfg-hostops.pl --server 9.155.113.136 --username root --password password --operation reboot  
Host bc-h-15-b5.mainz.de.ibm.com rebooted successfully.
```

When the server reboots, confirm the driver is installed by issuing the `-query` command as shown in Example 3-1. ESXi 4.1 does not have any requirement to claim the storage for VAAI (unlike ESX). More details about claiming IBM storage systems in ESX can be found in the IBM VAAI driver installation guide.

3.1.3 Confirming VAAI Hardware Acceleration is detected

Confirm whether vSphere (ESX/ESXi 4.1 or ESXi 5) detected that the storage hardware is VAAI capable.

Using the vSphere CLI with ESX/ESXi 4.1

Confirm the VAAI status by issuing a command similar to the one shown in Example 3-4. Unlike the ESX/ESXi Tech Support mode console, a Windows operating system does not provide the **egrep** command. However, it can be added by installing a package, such as Cygwin.

To perform the same task using the Tech Support mode shell, run the following command:

```
esxcfg-scsidevs -l | egrep "Display Name:|VAAI Status:"
```

Sample output is shown in Example 3-4.

Example 3-4 Using ESX CLI to confirm VAAI status

```
esxcfg-scsidevs.pl --server 9.155.113.136 --username root --password password -l | egrep  
"Display Name:|VAAI Status:"
```

```
Display Name: IBM Fibre Channel Disk (eui.0017380027820387)  
VAAI Status: supported
```

Using the vSphere CLI with ESX/ESXi 5.0/5.1

In ESXi 5.0/5.1, two tech support mode console commands can be used to confirm VAAI status. In Example 3-5, the **esxcli storage core device list** command is used to list every volume and its capabilities. However, it just reports VAAI is *supported* or *not supported*. Use the **esxcli storage core device vaa status get** command to list the four VAAI functions currently available for each volume. Three of these functions are supported by XIV.

Example 3-5 Using ESXi 5.0/5.1 commands to check VAAI status

```
~ # esxcli storage core device list  
eui.00173800278218b8  
  Display Name: IBM Fibre Channel Disk (eui.00173800278218b8)  
  Has Settable Display Name: true  
  Size: 98466  
  Device Type: Direct-Access  
  Multipath Plugin: NMP  
  Devfs Path: /vmfs/devices/disks/eui.00173800278218b8  
  Vendor: IBM  
  Model: 2810XIV  
  Revision: 0000  
  SCSI Level: 5  
  Is Pseudo: false  
  Status: on  
  Is RDM Capable: true  
  Is Local: false  
  Is Removable: false  
  Is SSD: false  
  Is Offline: false  
  Is Perennially Reserved: false  
  Thin Provisioning Status: unknown
```

```
Attached Filters:
VAAI Status: supported
Other UUIDs: vml.01000300003133303237383231384238323831305849
```

```
~ # esxcli storage core device vaaI status get
eui.00173800278218b8
VAAI Plugin Name:
ATS Status: supported
Clone Status: supported
Zero Status: supported
Delete Status: unsupported
```

Using the vSphere Client

From the vSphere Client, verify whether a datastore volume is VAAI capable by viewing the hardware acceleration status from the Configuration tab (Figure 3-4). Possible states are *Unknown*, *Supported* and *Not Supported*.

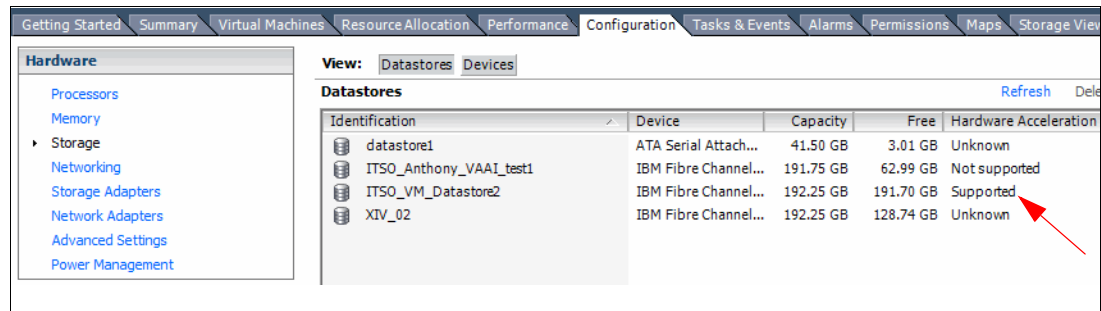


Figure 3-4 Hardware acceleration status

What to do if the Hardware Acceleration status shows as Unknown

ESXi 5.0/5.1 uses an ATS command as soon as it detects a new LUN to determine whether hardware acceleration is possible. For ESX/ESXi 4.1, the initial hardware acceleration status of a datastore or device normally shows as Unknown. The status will change to Supported after ESX/ESXi performs a VAAI offload function. If the attempt by ESX/ESXi to use an offload command fails, the state changes from Unknown to Not Supported. If it succeeds, it changes from Unknown to Supported. One way to prompt this change is to clone a virtual disk that is resident on that datastore. You can also copy a virtual disk to a new file in the relevant datastore in the vSphere Client.

Disabling VAAI globally on a vSphere server

You can disable VAAI entirely in vSphere 4.1 or vSphere 5. From the vSphere Client inventory panel, select the host and then click the Configuration tab. Select **Advanced Settings** in the Software pane. The following options need to be set to 0, which means they are disabled:

DataMover tab DataMover.HardwareAcceleratedMove
DataMover tab DataMover.HardwareAcceleratedInit
VMFS3 tab VMFS3.HardwareAcceleratedLocking

All three options are enabled by default, meaning that the value of each parameter is set to 1, as shown in Figure 3-5 on page 36.

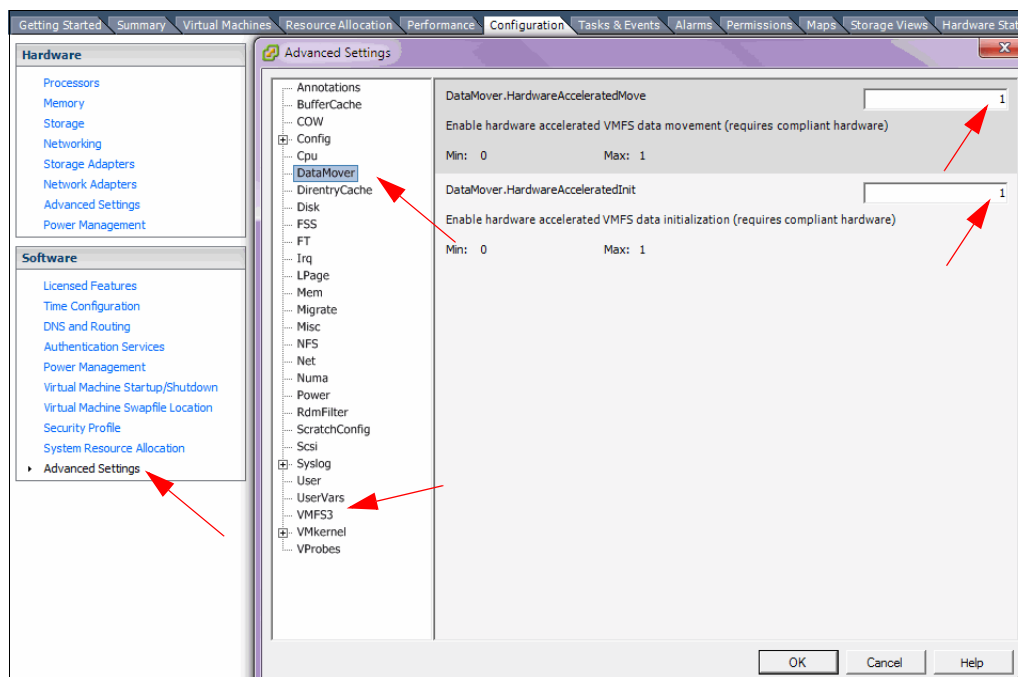


Figure 3-5 Disable VAAI in the vSphere Client

If using the service console to control VAAI, the following commands were tested and found to work on both ESX/ESXi 4.1 and ESXi 5.0/5.1. The first three commands display the status of VAAI. If the value returned for each function is 0, that function is disabled. If the value returned is 1, the function is enabled.

```
esxcfg-advcfg -g /DataMover/HardwareAcceleratedMove
esxcfg-advcfg -g /DataMover/HardwareAcceleratedInit
esxcfg-advcfg -g /VMFS3/HardwareAcceleratedLocking
```

The following commands disable each VAAI function (changing each value to 0):

```
esxcfg-advcfg -s 0 /DataMover/HardwareAcceleratedMove
esxcfg-advcfg -s 0 /DataMover/HardwareAcceleratedInit
esxcfg-advcfg -s 0 /VMFS3/HardwareAcceleratedLocking
```

The following commands enable VAAI (changing each value to 1):

```
esxcfg-advcfg -s 1 /DataMover/HardwareAcceleratedMove
esxcfg-advcfg -s 1 /DataMover/HardwareAcceleratedInit
esxcfg-advcfg -s 1 /VMFS3/HardwareAcceleratedLocking
```

ESXi 5.0/5.1 command syntax

ESXi 5.0/5.1 brings in new syntax that can also be used. Example 3-6 shows the commands to use to confirm status, disable, and enable one of the VAAI functions.

Example 3-6 ESXi VAAI control commands

```
esxcli system settings advanced list -o /DataMover/HardwareAcceleratedMove
esxcli system settings advanced set --int-value 0 --option /DataMover/HardwareAcceleratedMove
esxcli system settings advanced set --int-value 1 --option /DataMover/HardwareAcceleratedMove
```

In addition, the new unmap VAAI command is available in ESXi 5.0/5.1. At time of writing, this command is not supported by the XIV. In Example 3-7, the unmap function is confirmed to be enabled and is then disabled. Finally it is confirmed to be disabled.

Example 3-7 Disabling block delete in ESXi 5.0/5.1

```
~ # esxcli system settings advanced list -o /VMFS3/EnableBlockDelete | grep "Int Value"
  Int Value: 1
  Default Int Value: 1
~ # esxcli system settings advanced set --int-value 0 --option /VMFS3/EnableBlockDelete
~ # esxcli system settings advanced list -o /VMFS3/EnableBlockDelete | grep "Int Value"
  Int Value: 0
  Default Int Value: 1
```

For more information, see this VMWare knowledge base topic:

<http://kb.vmware.com/kb/1021976>

3.1.4 Disabling and enabling VAAI on the XIV on a per volume basis

You can disable or enable VAAI support at the XIV on a per volume basis, although doing so is normally not necessary. The commands are documented here to so that you are aware of how it is done. Generally, do not use these commands unless advised to do so by IBM support.

VAAI management is done using the XCLI. The two relevant commands are `vol_enable_vaa` and `vol_disable_vaa`. If you run these commands without specifying a volume, the command works on all volumes. In Example 3-8, VAAI is disabled for all volumes without the confirmation prompt using the `-y` parameter. VAAI is then enabled for all volumes, again, without confirmation.

Example 3-8 Enabling VAAI for all volumes

```
XIV-02-1310114>>vol_disable_vaa -y
Command executed successfully.
XIV-02-1310114>>vol_enable_vaa -y
Command executed successfully.
```

Example 3-9 displays the VAAI status for an individual volume, disabling VAAI, confirming it is disabled and then enabling it. The `vol_list` command does not show VAAI status by default. Use the `-x` parameter to get the XML output. Because the XML output is long and detailed, only a subset of the output is shown.

Example 3-9 Disabling and enabling VAAI on a per-volume basis

```
XIV_PFE3_7804143>>vol_list vol=ITSO_DataStore1 -x
<XCLIRETURN STATUS="SUCCESS" COMMAND_LINE="vol_list vol=ITSO_DataStore1 -x">
  <OUTPUT> ....
                <enable_VAAI value="yes"/>
                <user_disabled_VAAI value="no"/>
XIV_PFE3_7804143>>vol_disable_vaa vol=ITSO_DataStore1
Command executed successfully.
XIV_PFE3_7804143>>vol_list vol=ITSO_DataStore1 -x
<XCLIRETURN STATUS="SUCCESS" COMMAND_LINE="vol_list vol=ITSO_DataStore1 -x">
  <OUTPUT> ....
```

```
<enable_VAAI value="no"/>
<user_disabled_VAAI value="yes"/>
XIV_PFE3_7804143>>vol_enable_vaa1 vol=ITS0_DataStore1
```

After you enable VAAI for your volume, you need to prompt vSphere to attempt an offload function before hardware acceleration will show as supported. For more information, see 3.1.3, “Confirming VAAI Hardware Acceleration is detected” on page 34.

3.1.5 Testing VAAI

There are two simple tests that you can use on a new datastore to confirm that VAAI offload is working. Testing is best done on a new unused datastore/ Using a new datastore removes the risk that competing I/O confuses your test. Displaying the performance of your selected datastore using XIV Top shows that offload is working.

The hardware accelerated initialization or block zeroing test

This process creates a new virtual machine. You need to run this test twice. Run it the first time with HardwareAcceleratedInit disabled, and the second time with HardwareAcceleratedInit enabled.

To run the test:

1. Create a volume on the XIV and then create a datastore using that volume. This process allows you to run your tests on a datastore that has no competing traffic.
2. Start XIV Top from the XIV GUI, and select the new volume from the volumes column on the left. Hold down the control key and select IOPS and BW (bandwidth in MBps).
3. Disable or enable HardwareAcceleratedInit using the process detailed in “Disabling VAAI globally on a vSphere server” on page 35.
4. From the vSphere Client home page, go to Hosts and Clusters.
5. Right-click your selected ESX/ESXi server, and select **New Virtual Machine**.
6. When prompted to select a configuration, leave it on **Typical**.
7. Give the new virtual machine a name.
8. For a datastore, select the new datastore you created and are monitoring in XIV Top.
9. When prompted for a Guest Operating System, leave it on the default.
10. When prompted to create a disk, leave the default size (40 GB), but select **Supports clustering features such as Fault Tolerance**. If you are using vSphere Client 5.0, select the **Thick Provision Eager Zero** option. This option formats the VMDK with zeros.
11. While the virtual machine is being created, monitor IOPS and throughput in MBps being sent to the datastore in XIV Top. With HardwareAcceleratedInit disabled, you see a large volume of throughput and IOPS. With HardwareAcceleratedInit enabled, you see some IOPS but almost no throughput.

Figure 3-6 shows a virtual machine with HardwareAcceleratedInit disabled. In this test, over 800 IOPS with 700 MBps of throughput are seen for over 60 seconds.

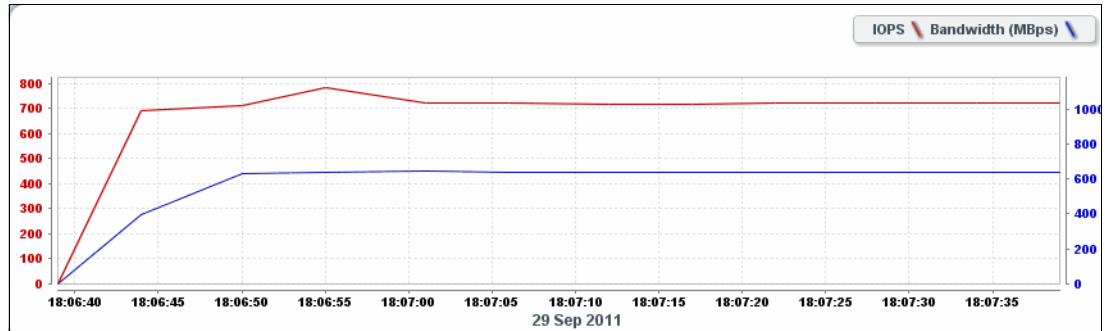


Figure 3-6 Creating a virtual machine with eager zero thick without VAAI enabled

Figure 3-7 shows a virtual machine with HardwareAcceleratedInit enabled. In this test, over 2200 IOPS with 1 MBps of throughput are seen for less than 30 seconds. VAAI reduced the execution time by more than 50% and eliminated nearly all the throughput.

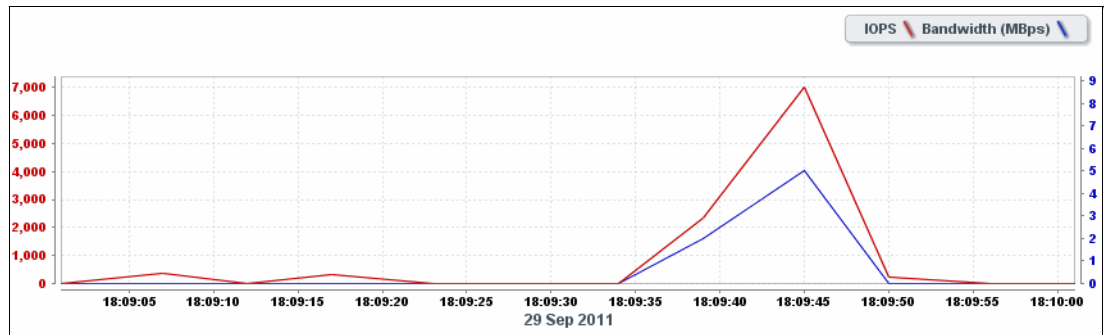


Figure 3-7 Creating a virtual machine with eager zero thick with VAAI enabled

The hardware accelerated move or full copy test

Clone the new virtual machine. You need to clone it twice, one time without VAAI and another time with VAAI. To clone the machine:

1. Disable or enable HardwareAcceleratedMove using the process detailed in “Disabling VAAI globally on a vSphere server” on page 35. In this example, it was disabled for the first test and enabled for the second test.
2. Right-click the new virtual machine, and select **Clone**.
3. Give the new virtual machine a name, and click **Next**.
4. Select an ESX/ESXi server and then select **Next**.
5. Select the same datastore that you created in the previous test and that you are still monitoring in XIV Top.
6. Accept all other defaults to create the clone.
7. While the clone is being created, monitor the throughput and IOPS on the XIV volume in XIV Top.

In Figure 3-8, a virtual machine was cloned with HardwareAcceleratedMove disabled. In this test, over 12000 IOPS with 700 MBps of throughput, were seen in bursts for nearly three minutes. Only one of these bursts is shown.

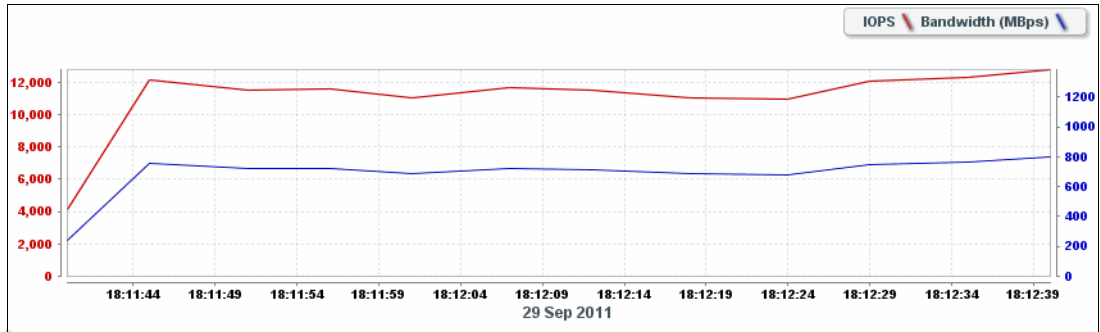


Figure 3-8 Volume cloning without VAAI

In Figure 3-9 a virtual machine was cloned with HardwareAcceleratedMove enabled. No operating system was installed. In this test, the IOPS peaked at 600, with 2 MBps of throughput being seen for less than 20 seconds. This peak means that VAAI reduced the execution time by nearly 90% and eliminated nearly all the throughput and IOPS being sent to the XIV. The affect on server performance was dramatic, as was the reduction in traffic across the SAN.

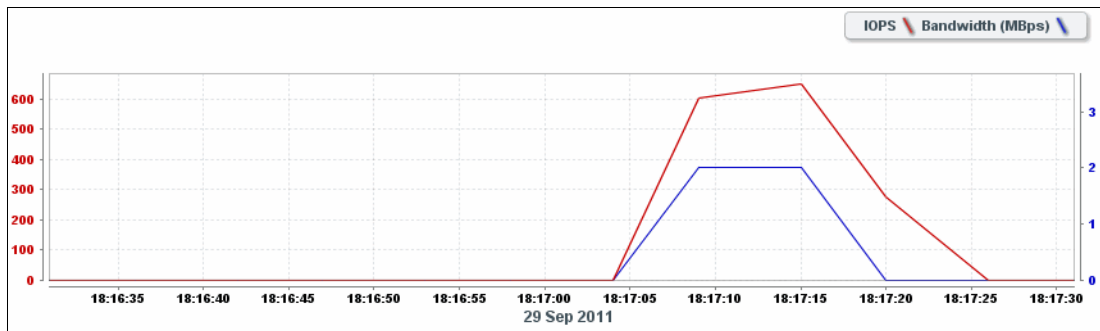


Figure 3-9 Volume cloning with VAAI

Both of these tests were done with server and SAN switch hardware that was less than ideal and no other performance tuning. These results are therefore indicative rather than a benchmark. If your tests do not show any improvement when using VAAI, confirm that Hardware Acceleration shows as Supported. For more information, see 3.1.3, “Confirming VAAI Hardware Acceleration is detected” on page 34.



Attaching VMware ESX to XIV

This chapter describes the considerations and implementation steps involved when attaching an XIV Storage System to a VMware ESX host. The following VMware ESX versions are successively covered: VMware ESX 3.5, VMWare ESX/ESxi 4.x, and VMware ESXi 5.0/5.1.

For each version, we discuss best practices for multipathing and performance tuning.

4.1 VMware ESX 3.5 and XIV

This section describes attaching VMware ESX 3.5 hosts through Fibre Channel.

Details about Fibre Channel configuration on VMware ESX server 3.5 are at:

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_san_cfg.pdf

Refer also to:

http://www.vmware.com/pdf/vi3_san_design_deploy.pdf

Follow these steps to configure the VMware host for FC attachment with multipathing:

1. Check host bus adapters (HBAs) and Fibre Channel (FC) connections from your host to XIV Storage System.
2. Configure the host, volumes, and host mapping in the XIV Storage System.
3. Discover the volumes created on XIV.

4.1.1 Installing HBA drivers

VMware ESX includes drivers for all the HBAs that it supports. VMware strictly controls the driver policy, and only drivers provided by VMware can be used. Any driver updates are normally included in service/update packs.

Supported FC HBAs are available from IBM, Emulex, and QLogic. Further details about HBAs supported by IBM are available from the SSIC website at:

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

Unless otherwise noted in the SSIC, use the firmware and driver versions promoted by VMware in association with the relevant hardware vendor. You can find supported VMware driver versions at:

<http://www.vmware.com/resources/compatibility/search.php?deviceCategory=io>

Tip: If Windows 2003 guests are using LSI Logic drivers, see the following VMware knowledge base topic regarding blocksize: <http://kb.vmware.com/kb/9645697>. Generally, use a maximum block size of 1 MB.

4.1.2 Scanning for new LUNs

Before you can scan for new LUNs on ESX, your host needs to be added and configured on the XIV Storage System.

Group ESX hosts that access the same shared LUNs in a cluster (XIV cluster), as shown in Figure 4-1 on page 43.

Name	Type	Cluster
Standalone Hosts		
itso_esx_cluster		
itso_esx_host1	default	itso_esx_cluster
10000000C92E894D	FC	
10000000C9591A5D	FC	
itso_esx_host2	default	itso_esx_cluster
21000024FF24A426	FC	
21000024FF28C151	FC	

Figure 4-1 ESX host cluster setup in XIV GUI

Assign those LUNs to the cluster, as shown in Figure 4-2.

Name	Size (GB)	LUN	Name
protected_VMFS_1	17.0	0	
protected_VMFS_2	17.0	1	protected_VMFS_1
Quorum	17.0	2	protected_VMFS_2

Figure 4-2 ESX LUN mapping to the cluster

To scan for and configure new LUNs:

1. Complete the host definition and LUN mappings in the XIV Storage System.
2. Click the Configuration tab for your host, and select **Storage Adapters**. Figure 4-3 shows vmhba2 highlighted. However, a rescan accesses all adapters. The adapter numbers might be enumerated differently on the different hosts, but this is not an issue.

Device	Type	SAN Identifier
QLA2340-Single Channel 2Gb Fibre Channel to PCI-X HBA		
vmhba2	Fibre Channel	21:00:00:e0:8b:0a:90:b5
vmhba3	Fibre Channel	21:00:00:e0:8b:0a:90:b5
53c1030 PCI-X Fusion-MPT Dual Ultra320 SCSI		
vmhba0	SCSI	
vmhba1	SCSI	

Details	
vmhba2	
Model:	QLA2340-Single Channel 2Gb Fibre Channel to PCI-X HBA
WWPN:	21:00:00:e0:8b:0a:90:b5
Target:	2
SCSI Target	
Path	Canonical Pa... Type Capacity LUN ID

Figure 4-3 Select storage adapters

3. Select **Scan for New Storage Devices** and **Scan for New VMFS Volumes**. Click **OK**, as shown in Figure 4-4 on page 44.

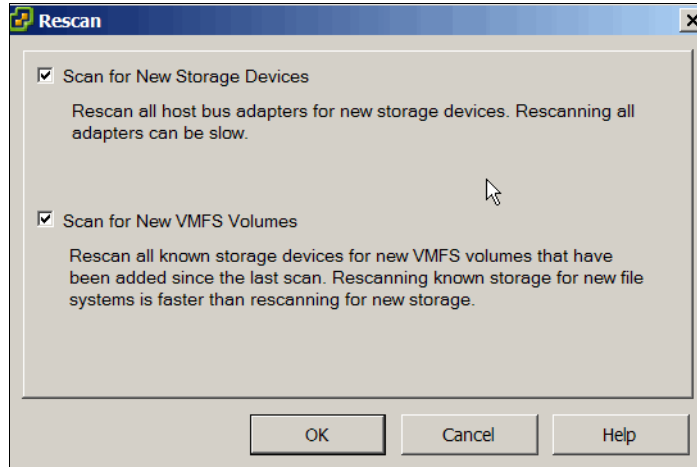


Figure 4-4 Rescan for New Storage Devices

The new LUNs assigned are displayed in the Details window, as shown in Figure 4-5.

Details				
vmhba2				
Model: QLA2340/2340L				
WWPN 21:00:00:e0:8b:0a:90:b5				
Target 2				
SCSI Target				
Path	Canonical Pa...	Type	Capacity	LUN ID
vmhba2:2:0	vmhba2:2:0	disk	32.00 GB	0
vmhba2:2:1	vmhba2:2:1	disk	32.00 GB	1
SCSI Target				
Path	Canonical Pa...	Type	Capacity	LUN ID
vmhba2:3:0	vmhba2:2:0	disk	32.00 GB	0
vmhba2:3:1	vmhba2:2:1	disk	32.00 GB	1

Figure 4-5 FC discovered LUNs on vmhba2

In this example, controller vmhba2 can see two LUNs (LUN 0 and LUN 1) circled in green. These LUNs are visible on two targets (2 and 3) circled in red. The other controllers on the host show the same path and LUN information.

For detailed information about how to use LUNs with virtual machines, see the VMware guides, available at:

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_admin_guide.pdf

http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_3_server_config.pdf

Ensuring common LUN IDs across ESX servers

In Figure 4-2 on page 43, the volumes being mapped to the clustered ESX servers were mapped to a cluster (itso_esx_cluster) defined on the XIV. They were not mapped to each individual ESX server, which were defined to the XIV as hosts (itso_esx_host1 and itso_esx_host2). Map to a cluster because the XIV does not support Network Address Authority (NAA). When multiple ESX servers are accessing the same volume, each ESX server accesses each XIV volume using the same LUN ID. This setup is normal in an ESX cluster using VMFS. The LUN ID is set by the storage administrator when the volume is mapped.

The reason for this requirement is the risk of resigature thrashing related to the LUN ID, not the target. This restriction is described in the topic at <http://kb.vmware.com/kb/1026710>. While the title of the topic refers to ESX 4.x hosts, it also addresses ESX 3.5.

By mapping volumes to the cluster rather than to each host, you ensure that each host accesses each volume using the same LUN ID. Private mappings can be used if necessary.

4.1.3 Assigning paths from an ESX 3.5 host to XIV

All information in this section relates to ESX 3.5 (and not other versions of ESX) unless otherwise specified. The procedures and instructions given here are based on code that was available at the time of writing.

VMware provides its own multipathing I/O driver for ESX. No additional drivers or software are required. As such, the XIV Host Attachment Kit provides only documentation, and no software installation is required.

The ESX 3.5 multipathing supports the following path policies:

- ▶ **Fixed:** Always use the preferred path to the disk. If preferred path is not available, use an alternative path to the disk. When the preferred path is restored, an automatic failback to the preferred path occurs.
- ▶ **Most Recently Used:** Use the path most recently used while the path is available. Whenever a path failure occurs, an alternative path is chosen. There is no automatic failback to the original path. Do not use this option with XIV.
- ▶ **Round-Robin (ESX 3.5 experimental):** Multiple disk paths are used and balanced based on load. Round-Robin is not supported for production use in ESX version 3.5. Do not use this option with ESX 3.5.

ESX Native Multipathing automatically detects IBM XIV and sets the path policy to **Fixed** by default. Do not change this setting. Also, when setting the preferred path or manually assigning LUNs to a specific path, monitor it carefully so you do not overload the IBM XIV storage controller port. Use **esxtop** to monitor outstanding queues pending execution.

XIV is an active/active storage system, and therefore it can serve I/O to all LUNs using every available path. However, the driver with ESX 3.5 cannot perform the same function and by default cannot fully load balance. You can artificially overcome this limitation by confirming the correct pathing policy (correcting if necessary) and distributing the I/O load over the available HBAs and XIV ports. This process is called *manual load balancing*.

To manually balance the load:

1. Set the pathing policy in ESX 3.5 to either **Most Recently Used (MRU)** or **Fixed**. When accessing storage on the XIV, the correct policy is **Fixed**. In the VMware Infrastructure Client, select the server, click the **Configuration** tab, and select **Storage** (Figure 4-6 on page 46).

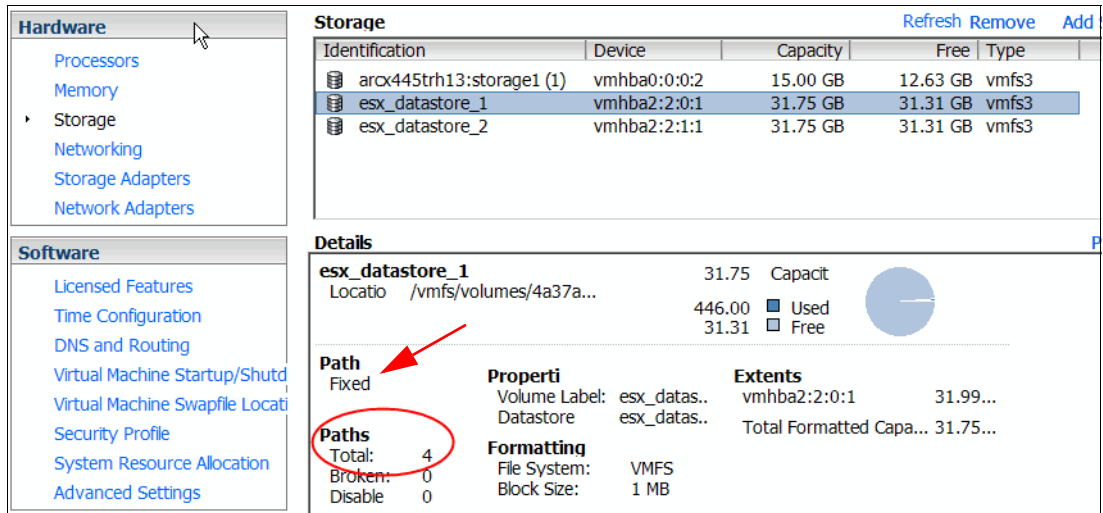


Figure 4-6 Storage paths

In this example, the LUN is highlighted (esx_datastore_1), and the number of paths is 4 (circled in red).

2. Select **Properties** to view more details about the paths. In the Properties window, you can see that the active path is vmhba2:2:0, as shown in Figure 4-7.

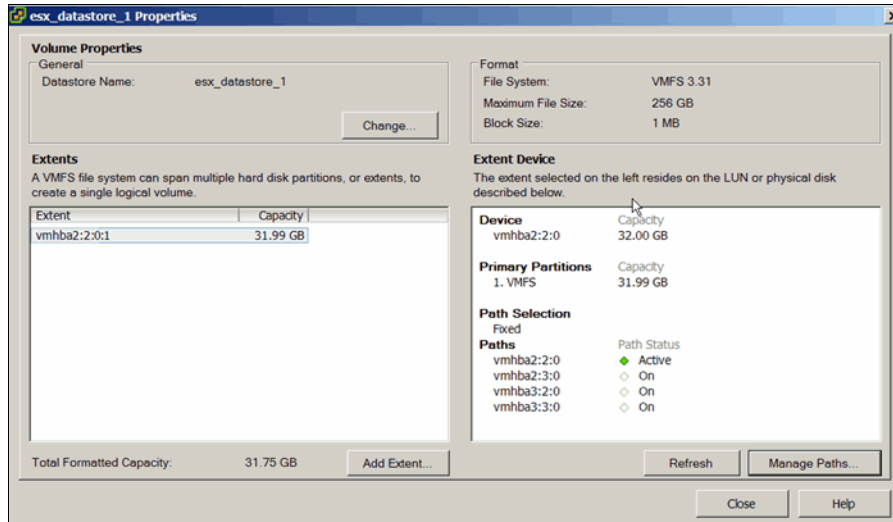


Figure 4-7 Storage path details

- To change the current path, select **Manage Paths**, and the Manage Paths window opens, as shown in Figure 4-8. Set the pathing policy to Fixed if it is not already by selecting **Change** in the Policy window.

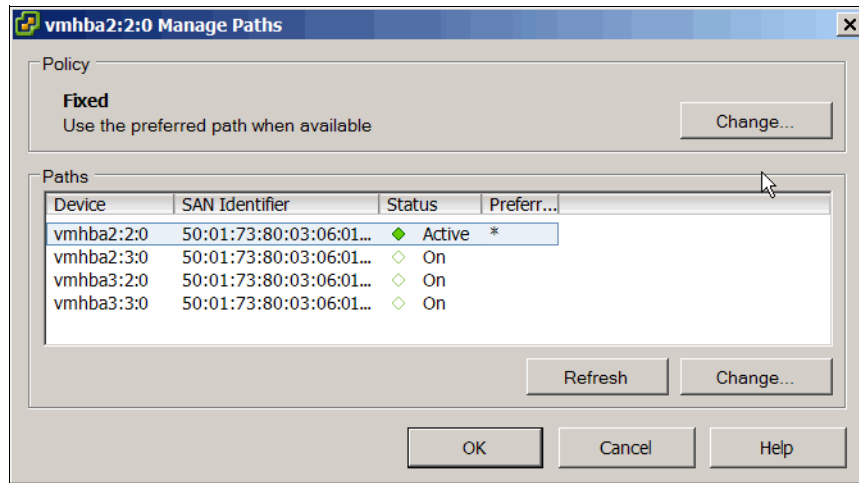


Figure 4-8 Change paths window

- To manually load balance, highlight the preferred path in the Paths pane, and click **Change**. Assign an HBA and target port to the LUN, as shown in Figure 4-9.

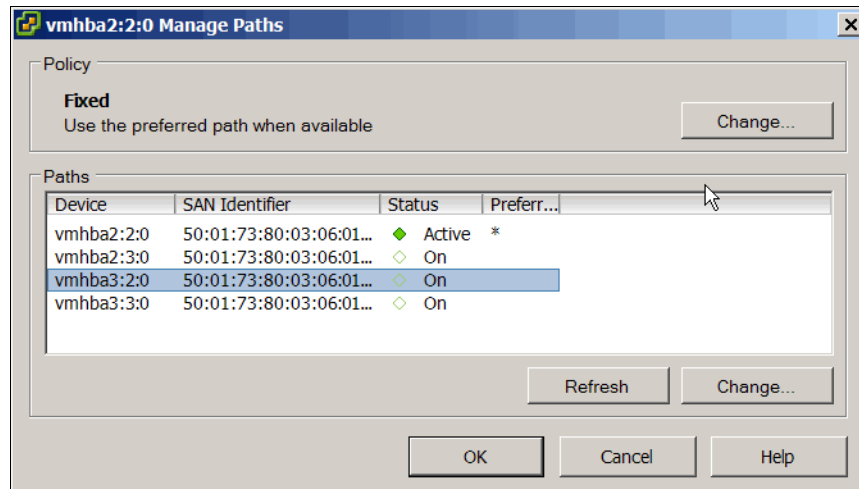


Figure 4-9 Change to new path

Figure 4-10 on page 48 shows setting the preference.

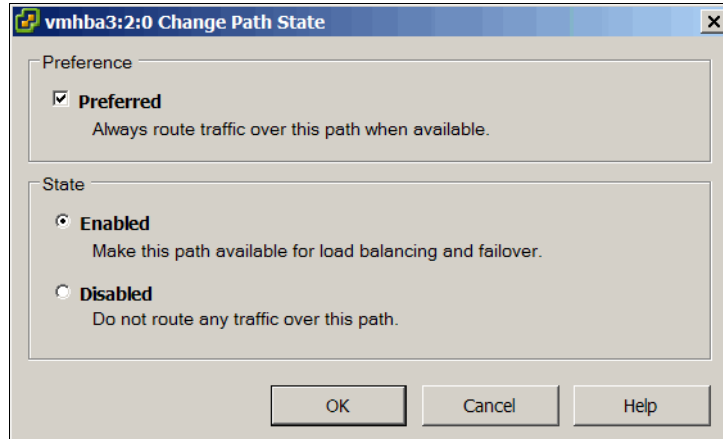


Figure 4-10 Set preferred

Figure 4-11 shows the new preferred path.

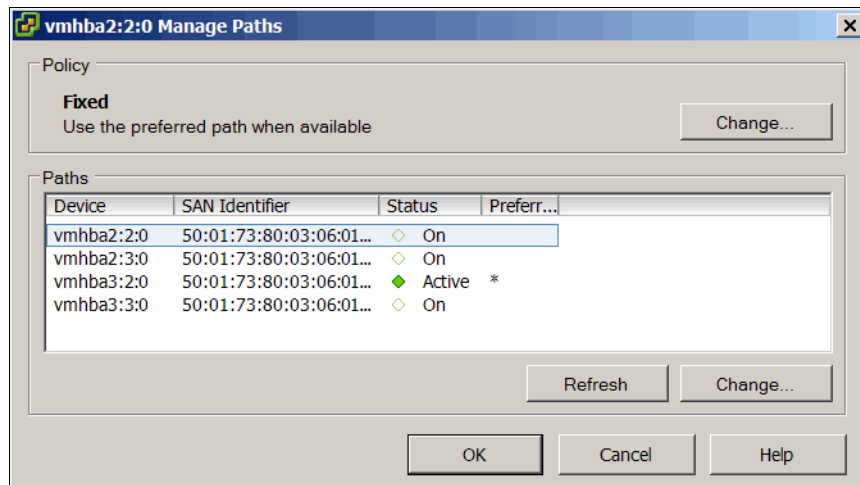


Figure 4-11 New preferred path set

- Repeat steps 1-4 to manually balance I/O across the HBAs and XIV target ports. Because workloads change over time, review the balance periodically.

Example 4-1 and Example 4-2 on page 49 show the results of manually configuring two LUNs on separate preferred paths on two ESX hosts. Only two LUNs are shown for clarity, but this configuration can be applied to all LUNs assigned to the hosts in the ESX datacenter.

Example 4-1 ESX Host 1 preferred path

```
[root@arcx445trh13 root]# esxcfg-mpath -l
Disk vmhba0:0:0 /dev/sda (34715MB) has 1 paths and policy of Fixed
Local 1:3.0 vmhba0:0:0 On active preferred

Disk vmhba2:2:0 /dev/sdb (32768MB) has 4 paths and policy of Fixed
FC 5:4.0 210000e08b0a90b5<->5001738003060140 vmhba2:2:0 On active preferred
FC 5:4.0 210000e08b0a90b5<->5001738003060150 vmhba2:3:0 On
FC 7:3.0 210000e08b0a12b9<->5001738003060140 vmhba3:2:0 On
FC 7:3.0 210000e08b0a12b9<->5001738003060150 vmhba3:3:0 On
```

```
Disk vmhba2:2:1 /dev/sdc (32768MB) has 4 paths and policy of Fixed
FC 5:4.0 210000e08b0a90b5<->5001738003060140 vmhba2:2:1 On
FC 5:4.0 210000e08b0a90b5<->5001738003060150 vmhba2:3:1 On
FC 7:3.0 210000e08b0a12b9<->5001738003060140 vmhba3:2:1 On
FC 7:3.0 210000e08b0a12b9<->5001738003060150 vmhba3:3:1 On active preferred
```

Example 4-2 shows the results of manually configuring two LUNs on separate preferred paths on ESX host 2.

Example 4-2 ESX host 2 preferred path

```
[root@arcx445bvkf5 root]# esxcfg-mpath -l
Disk vmhba0:0:0 /dev/sda (34715MB) has 1 paths and policy of Fixed
Local 1:3.0 vmhba0:0:0 On active preferred

Disk vmhba4:0:0 /dev/sdb (32768MB) has 4 paths and policy of Fixed
FC 7:3.0 10000000c94a0436<->5001738003060140 vmhba4:0:0 On active preferred
FC 7:3.0 10000000c94a0436<->5001738003060150 vmhba4:1:0 On
FC 7:3.1 10000000c94a0437<->5001738003060140 vmhba5:0:0 On
FC 7:3.1 10000000c94a0437<->5001738003060150 vmhba5:1:0 On

Disk vmhba4:0:1 /dev/sdc (32768MB) has 4 paths and policy of Fixed
FC 7:3.0 10000000c94a0436<->5001738003060140 vmhba4:0:1 On
FC 7:3.0 10000000c94a0436<->5001738003060150 vmhba4:1:1 On
FC 7:3.1 10000000c94a0437<->5001738003060140 vmhba5:0:1 On
FC 7:3.1 10000000c94a0437<->5001738003060150 vmhba5:1:1 On active preferred
```

As an alternative to manually setting up paths to load balance, contact your IBM Technical Advisor or pre-sales technical support for a utility called `xivcfg-fixedpath`. This utility can be used to achieve an artificial load balance of XIV LUNs on VMware ESX 3.X and later.

The utility uses standard `esxcfg` and `esxcli` commands, and is run like a script, as shown in Example 4-3. This utility is not available for download through the internet.

Example 4-3 xivcfg-fixedpath utility

```
#!/xivcfg-fixedpath -h
Usage: xivcfg-fixedpath [-L | -T -Y | -V]
    -L #list current preferred paths for XIV devices.
    -T #run in test mode and print out the potentially disruptive commands,
but do not execute them.
    -V #print program version number and exit.
    -Y #do not prompt for confirmation.
```

```
# ./xivcfg-fixedpath -V
xivcfg-fixedpath: version 1.2
```

```
#!/xivcfg-fixedpath
```

```
-----
This program will rescan all FC HBAs, change all XIV disks to Fixed path policy
and reassign the XIV preferred disk path to balance all XIV LUNs across available
paths. This may result in I/O interruption depending on the I/O load and state of
devices and paths. Proceed (y/n)?
```

4.2 VMware Multi-Pathing architecture overview

To provide a functional understanding of the concepts underlying the connectivity of the XIV Storage System to a current generation vSphere environment (at the time of this writing), this section contains a description of the modular multipathing architecture and the associated definitions that were introduced in vSphere 4.0.

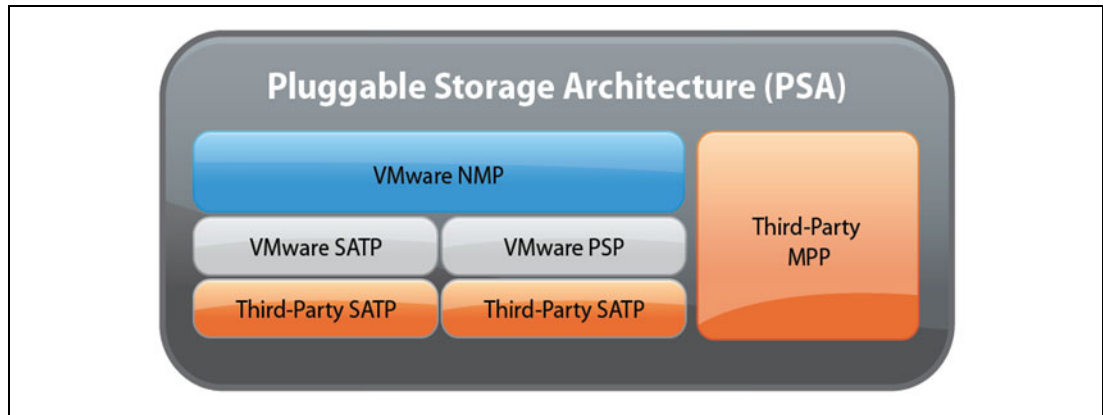


Figure 4-12 VMware Modular Multipathing Conceptual Layout

PSA - Pluggable Storage Architecture: The PSA is a hierarchical construct that enables the VMkernel to incorporate multiple, potentially storage vendor-specific, operational characteristics necessary to maintain path availability and provide optimal load balancing by modularly coordinating the operations of both the native VMware multi-pathing module and other third-party modules simultaneously and in parallel. The PSA performs the following tasks:

- ▶ Monitors both logical and physical path I/O statistics
- ▶ Manages the I/O queuing for both physical paths to the storage and logical devices
- ▶ Allocates logical device bandwidth among multiple virtual machines
- ▶ Manages the discovery and withdrawal of physical paths
- ▶ Absorbs and removes third-party modules as required
- ▶ Masks attributes of virtual machines from specific modules
- ▶ Maps I/O requests targeting a specific logical device to the PSA module that defines access characteristics for that device.

NMP - Native MultiPathing: The NMP is the default multipathing module in VMware, and partially overlaps the functionality of self-contained third-party modules by executing path selection algorithms based on storage subsystem type(s). The NMP is at the top of the granular multipathing hierarchy, and represents the path selection layer that manages the underlying default and storage-vendor specific modules with responsibilities including:

- ▶ Mapping physical paths with logical devices
- ▶ Overseeing the claiming and unclaiming of physical paths
- ▶ Supporting logical device-specific administrative tasks including adding/removing devices and aborts, resets, and so on
- ▶ Managing I/O requests to specific logical devices by selecting the optimal path for a given I/O and executing request retries and contingency actions

MEM (vSphere 4.1+) - Management Extension Modules: Superseding the term “plug-in,” the MEM is a more functionally accurate term that describes the modular elements that integrate into the PSA.

MPP (MPM in vSphere 4.1+) - Multipathing Plug-in (Multipathing MEM in vSphere 4.1+): The MPP overlaps the responsibilities of the NMP described previously, but represents a “self-contained” third-party module provided by the storage vendor to replace both the NMP and its subordinate modules to comprehensively define the multi-pathing dynamics custom to a particular storage subsystem.

SATP (SATM in vSphere 4.1+) - Storage Array Type Plug-in (Storage Array Type MEM in vSphere 4.1+): The NMP delegates the responsibility of defining the path monitoring and failover policies, including activating and deactivating paths, to the SATP/SATM which is a storage subsystem-type aware MEM that accommodates specific characteristics of a class of storage subsystems. The SATP/SATM can consist of either a storage vendor-provided plug-in/module or the default module provided by VMware.

PSP (PSM in vSphere 4.1+) - Path Selection Plug-in (Path Selection MEM in vSphere 4.1+): The PSP/PSM operates in partnership with the SATP/SATM to allow the NMP to choose a specific physical path for I/O requests based on the SATP/SATM, and in turn incorporate the path selection policy to be associated with the physical paths mapped to each logical device. The PSP/PSM can consist of either a storage vendor-provided plug-in/module or the default module provided by VMware.

ALUA - Asymmetric Logical Unit Access: ALUA is a SCSI3 standard adopted by both the ESX/ESXi hypervisor and compatible storage subsystems, and is built into the PSP/PSM. Fundamentally, ALUA standardization facilitates transparency and cooperation between an operating system and the storage subsystem regarding the selection of optimal paths on both a port-by-port and volume-by-volume basis. An optimal path to a given LUN represents the path that incurs the minimal processing overhead for the storage subsystem. Therefore, ALUA enables the seamless, efficient, and dynamic management of paths in the context of both physical ports and logical devices while precluding the need for additional software. The XIV Storage System is ALUA-compatible beginning with XIV software version 10.1.

Example of VMware I/O Flow

To further clarify the mechanics of the NMP and the inter-relationships of the constituent modules, the following chain of events defines the handling of an I/O request originating from either a virtual machine or the hypervisor itself:

1. The NMP identifies and consults the PSP associated with the specific storage subsystem that owns the LUN in question.
2. The PSP/PSM possibly exploiting ALUA, selects the optimal or appropriate physical path on which to issue the I/O request.
3. If the I/O request successfully completes, the NMP reports its completion to the appropriate layer of the hypervisor.
4. If the I/O request results in an error, the NMP calls the appropriate SATP/SATM for the specific storage subsystem.
5. The SATP/SATM interprets the I/O command errors and, when appropriate, activates inactive paths.
6. The PSP/PSM is called again to select a new path upon which to issue the I/O request, since the SATP/SATM process may have since changed the path topology.

4.3 VMware ESX and ESXi 4.x and XIV

This section describes attaching ESX and ESXi 4.x hosts to XIV through Fibre Channel.

4.3.1 Installing HBA drivers

VMware ESX/ESXi includes drivers for all the HBAs that it supports. VMware strictly controls the driver policy, and only drivers provided by VMware can be used. Any driver updates are normally included in service/update packs.

Supported FC HBAs are available from Brocade, Emulex, IBM, and QLogic. Further details about HBAs supported by IBM are available from the SSIC web site:

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

Unless otherwise noted in the SSIC, use the firmware and driver versions promoted by VMware in association with the relevant hardware vendor. You can find supported VMware driver versions at:

<http://www.vmware.com/resources/compatibility/search.php?deviceCategory=io>

Tip: If Windows 2003 guests are using LSI Logic drivers, see the following VMware knowledge base topic regarding block size:

<http://kb.vmware.com/kb/9645697>

Generally, use a maximum block size of 1 MB

4.3.2 Identifying ESX host port WWN

Identify the host port WWN for FC adapters installed in the ESX Servers before you can start defining the ESX cluster and its host members, using the following steps:

1. Run the VMWare vSphere Client.
2. Connect to the ESX Server.
3. In the VMWare vSphere Client, select the server, click the Configuration tab, and then select **Storage Adapters**. Figure 4-13 shows the port WWNs for the installed FC adapters circled in red.

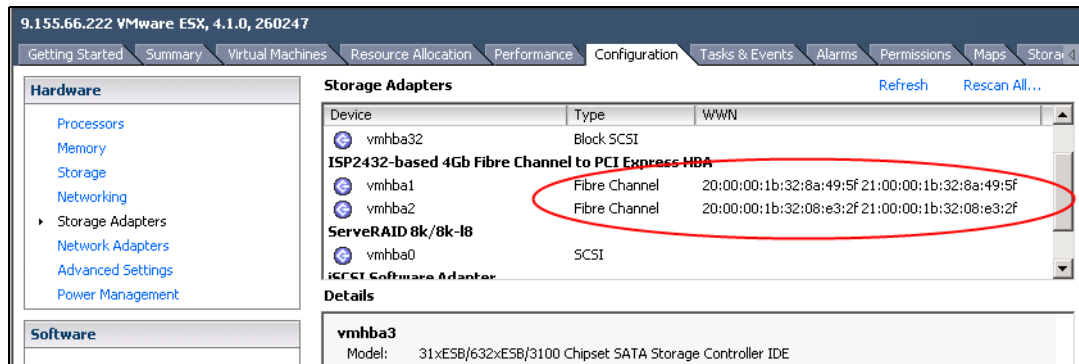


Figure 4-13 ESX host port WWNs

4. Repeat this process for all ESX hosts that you plan to connect to the XIV Storage System.

After identifying the ESX host port WWNs, you are ready to define hosts and clusters for the ESX servers. Create LUNs, and map them to defined ESX clusters and hosts on the XIV Storage System. See Figure 4-1 on page 43 and Figure 4-2 on page 43 for how this configuration might typically be set up.

Tip: Group the ESX hosts that access the same LUNs in a cluster (XIV cluster), and assign the LUNs to that cluster.

Considerations for the size and quantity of volumes

For volumes being mapped to ESX 4.x, the maximum volume size you can create on a second Generation XIV is 2181 GB. The maximum volume size you can create on an XIV Gen3 is 2185 GB.

The following configuration maximums are documented for vSphere 4.1:

- ▶ The maximum number of LUNs per server is 256
- ▶ The maximum number of paths per server is 1024
- ▶ The maximum number of paths per LUN is 32

If each XIV volume can be accessed through 12 fabric paths (which is a large number of paths), the maximum number of volumes is 85. Dropping the path count to six increases the maximum LUN count to 170. For installations with large numbers of raw device mappings, these limits can become a major constraint.

More details are at:

http://www.vmware.com/pdf/vsphere4/r41/vsp_41_config_max.pdf

4.3.3 Scanning for new LUNs

To scan and configure new LUNs:

1. Click the Configuration tab for your host, and select **Storage Adapters**, as shown in Figure 4-14 on page 54.

Here you can see vmhba1 highlighted, but a rescan searches across all adapters. The adapter numbers might be enumerated differently on the different hosts, but this is not an issue.

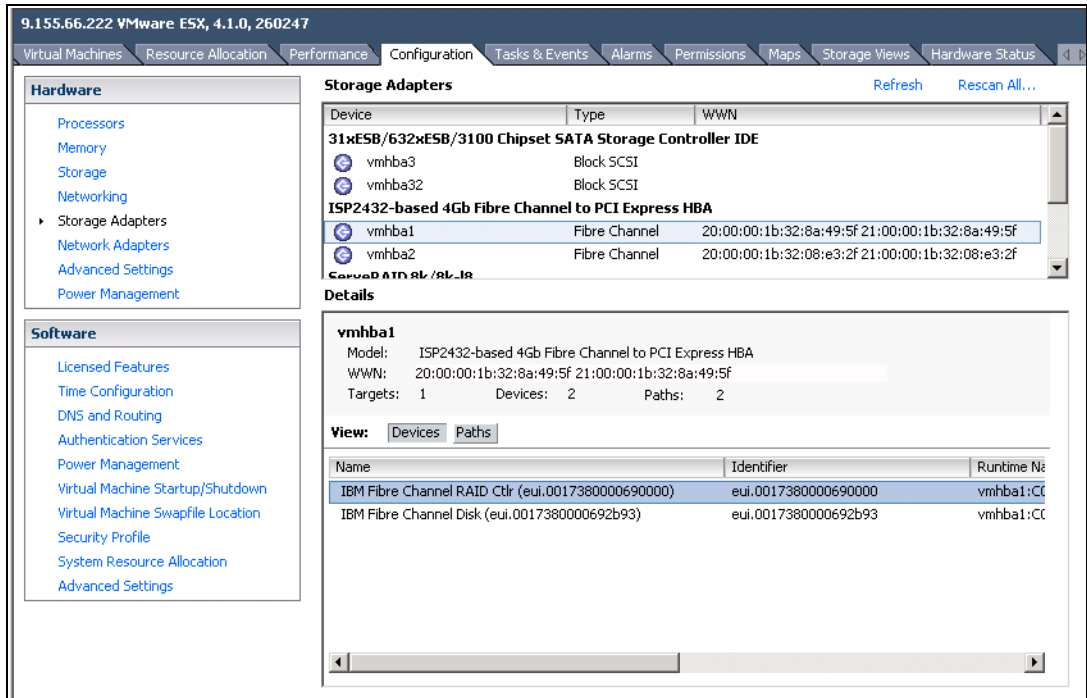


Figure 4-14 Select storage adapters

2. Select **Scan for New Storage Devices** and **Scan for New VMFS Volumes**. Click **OK** to scan for new storage devices, as shown in Figure 4-15.

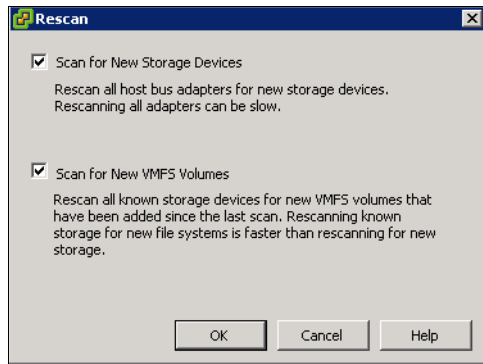


Figure 4-15 Rescan for new storage devices

3. The new LUNs are displayed in the **Details** pane. As shown in Figure 4-16, T controller vmhba1 can see two LUNs (LUN 1 and LUN 2) circled in red. The other controllers on the host show the same path and LUN information.

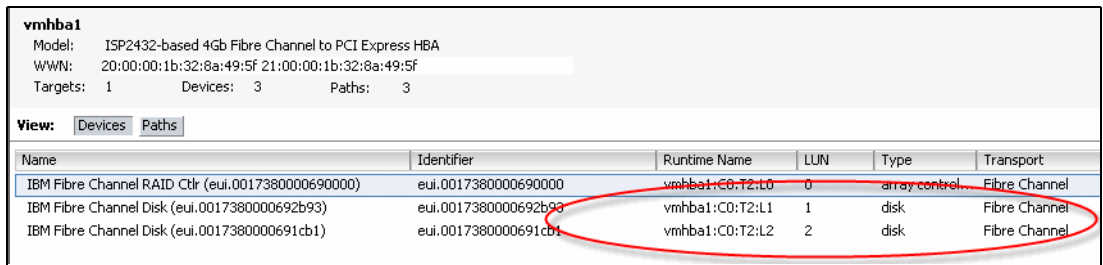


Figure 4-16 FC discovered LUNs on vmhba1

4.3.4 Attaching an ESX/ESXi 4.x host to XIV

This section describes the attachment of ESX/ESXi 4-based hosts to the XIV Storage System. It provides specific instructions for Fibre Channel (FC) and Internet Small Computer System Interface (iSCSI) connections. All the information in this section relates to ESX/ESXi 4 (and not other versions of ESX/ESXi) unless otherwise specified.

The procedures and instructions given here are based on code that was available at the time of writing. For the latest support information, see the Storage System Interoperability Center (SSIC) at:

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

Note: For additional background, an overview of the vSphere Pluggable Storage Architecture hierarchy and its mechanics are in 4.2, “VMware Multi-Pathing architecture overview” on page 50.

By default ESX/ESXi 4 supports the following types of storage arrays:

- ▶ **Active/active storage systems:** Allow access to the LUN simultaneously through all storage ports. Though all the paths are active all the time, the access is further defined as either symmetric or asymmetric depending on the architecture of the storage system. In a symmetric storage system, a given LUN can be owned by more than one storage controller at a time, meaning that there is no need to use another storage controller as a proxy to access the LUN. In contrast, an asymmetric storage subsystem designates a particular storage controller as the exclusive LUN owner on a LUN-by-LUN basis, and therefore accessing the LUN through the non-owning controller (by proxy) potentially incurs the additional overhead associated with involving more than one controller in the I/O path for traditional monolithic storage systems.

The XIV Storage System’s grid architecture harnesses all resources in tandem and represents a notable exception to this phenomenon. Finally, a key attribute of the active/active design is that if one or more ports fail, all of the other available ports continue allowing access from servers to the storage system for all LUNs.

- ▶ **Active/passive storage systems:** Systems where a LUN is accessible over a single storage port. The other storage ports act as backup for the active storage port.
- ▶ **Asymmetrical storage systems (VMW_SATP_DEFAULT_ALUA):** Support asymmetrical logical unit access (ALUA). ALUA-compliant storage systems provide different levels of access on a per-port basis. This configuration allows the SCSI Initiator port to make intelligent decisions in terms of internal bandwidth usage and processing efficiency on the storage subsystem. The host uses some of the active paths as primary and others as secondary. Accessing a LUN using a path that exclusively incorporates the managing controller or processor of a traditional monolithic storage subsystem is referred to as an active-optimized path selection, while accessing a LUN that involves a non-owning controller or processor is referred to as active-non-optimized.

With the release of VMware ESX 4 and VMware ESXi 4, VMware introduced the concept of a Pluggable Storage Architecture (PSA). PSA in turn introduced additional concepts to its Native Multipathing Plug-in (NMP). Refer to 4.2, “VMware Multi-Pathing architecture overview” on page 50 for a detailed description.

ESX/ESXi 4.x provides default SATPs that support non-specific active-active (VMW_SATP_DEFAULT_AA) and ALUA storage system (VMW_SATP_DEFAULT_ALUA). Each SATP accommodates special characteristics of a certain class of storage systems. It can perform the storage system-specific operations required to detect path state and activate an inactive path.

Note: Starting with XIV software Version 10.1, the XIV Storage System is a T10 ALUA-compliant storage system.

ESX/ESXi 4.x automatically selects the appropriate SATP plug-in for the IBM XIV Storage System based on the XIV Storage System software version. For versions before 10.1 and for ESX 4.0, the Storage Array Type is VMW_SATP_DEFAULT_AA. For XIV versions later than 10.1 and with ESX/ESXi 4.1, the Storage Array Type is VMW_SATP_DEFAULT_ALUA.

PSPs run with the VMware NMP, and are responsible for choosing a physical path for I/O requests. The VMware NMP assigns a default PSP to each logical device based on the SATP associated with the physical paths for that device.

VMware ESX/ESXi 4.x supports the following PSP types:

- ▶ Fixed (VMW_PSP_FIXED): Always use the preferred path to the disk if available. If the preferred path is not available, a random alternative path to the disk is chosen. When the preferred path is restored, an automatic failback to the preferred path occurs. This policy is not ALUA-aware, precluding exploitation of the distinction between active-optimized and active non-optimized paths, and so on, in the path selection policy. This can result in a phenomenon known as path thrashing when implemented with Active/Passive or asymmetric Active/Active arrays that are based on a traditional monolithic storage subsystem architecture.

Note: The potential for path thrashing is not applicable to the XIV Storage System.

- ▶ Most Recently Used (VMW_PSP_MRU): Selects the first working path, discovered at boot time, and continues this path (the most recently used) while the path remains available. Whenever a path failure occurs, an alternative path is chosen leveraging ALUA-awareness for compatible storage subsystems, meaning that whenever possible paths are chosen to incorporate the managing controller of processor for a given LUN. It is important to note that there is no automatic failback to the original path, and that manual intervention is necessary to restore the original state of access after a path failure/repair.
- ▶ Round-Robin (VMW_PSP_RR): The Round Robin policy employs a path selection algorithm that is ALUA-aware and implements load balancing by rotating the physical access to the LUN through all active-optimized paths by default (and also exploits active non-optimized paths, if configured to do so). The criteria for transitioning to the next available *active* path to a given LUN, and thus to optimize the distribution of workload across paths, relies upon either a path-centric I/O counter or a path-centric byte-wise counter exceeding a pre-set threshold (depending on configuration settings). Paths designated as standby, unavailable, or transitioning status will not be included in the rotation until they are re-activated or re-established.

Note: At the time of this writing, the Round Robin PSP is not supported for Logical Units that are part of a Microsoft Cluster Service (MSCS) virtual machine.

ESX has built-in rules defining relations between SATP and PSP for the storage system.

4.3.5 Configuring ESX/ESXi 4.x host for multipathing with XIV

With ESX/ESXi 4.x, VMWare supports a round-robin multipathing policy for production environments. The round-robin multipathing policy is always preferred over other choices when attaching to the IBM XIV Storage System.

Before proceeding with the multipathing configuration, complete the tasks described in the following sections:

- ▶ 4.3.1, “Installing HBA drivers” on page 52
- ▶ 4.3.2, “Identifying ESX host port WWN” on page 52
- ▶ “Considerations for the size and quantity of volumes” on page 53.

To add a datastore:

1. Start the VMware vSphere Client, and connect to your vCenter server.
2. Select the server that you plan to add a datastore to.
3. In the vSphere Client main window, click the **Configuration** tab for your host, and select **Storage**, as shown in Figure 4-17.

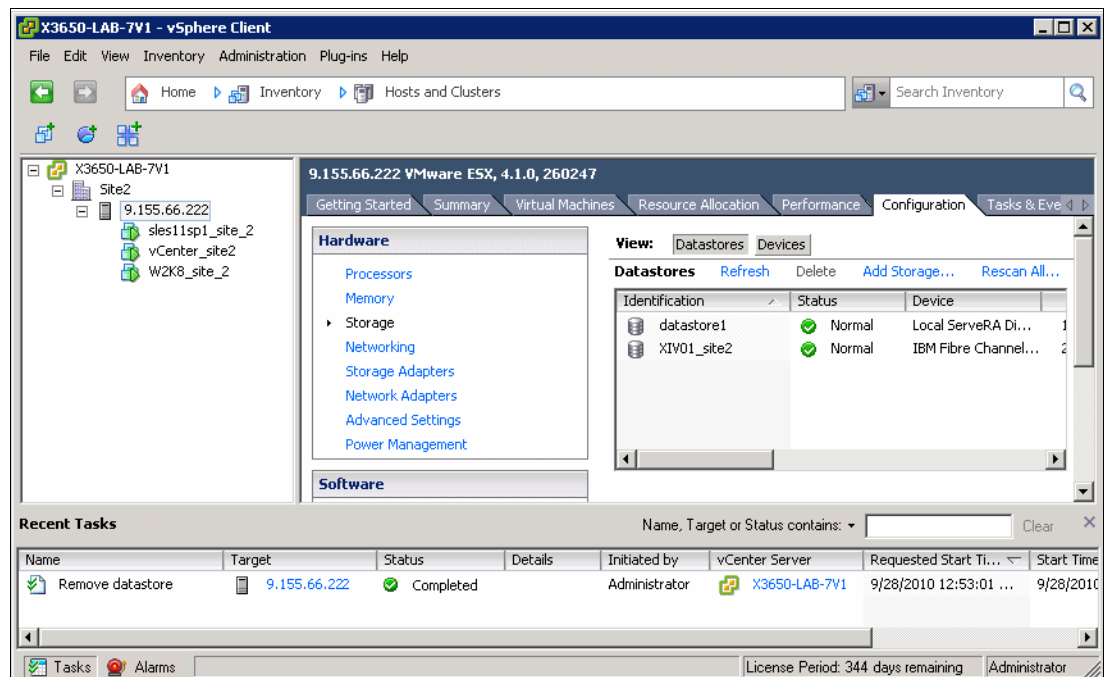


Figure 4-17 ESX/ESXi 4.x defined datastore

Here you can see datastore currently defined for the ESX host.

4. Click **Add Storage** to open the window shown in Figure 4-18.

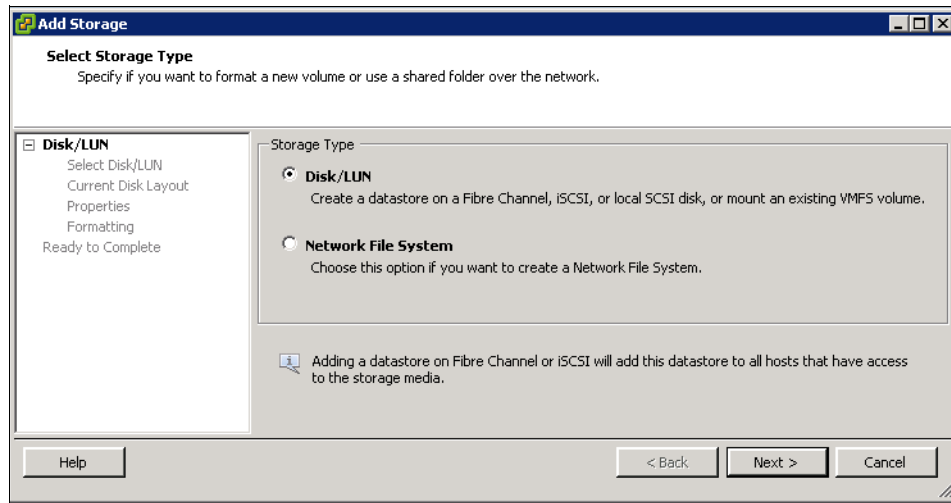


Figure 4-18 Add Storage dialog

5. In the Storage Type box, select **Disk/LUN**, and click **Next** to get to the window shown in Figure 4-19. You can see listed the Disks and LUNs that are available to use as a new datastore for the ESX Server.

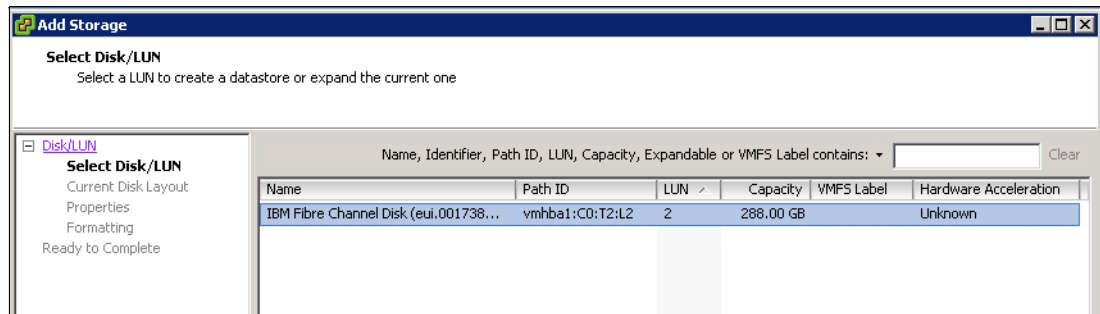


Figure 4-19 List of disks/LUNs for use as a datastore

6. Select the LUN that you want to use as a new datastore, and click **Next**. A new window similar to Figure 4-20 on page 59 is displayed.

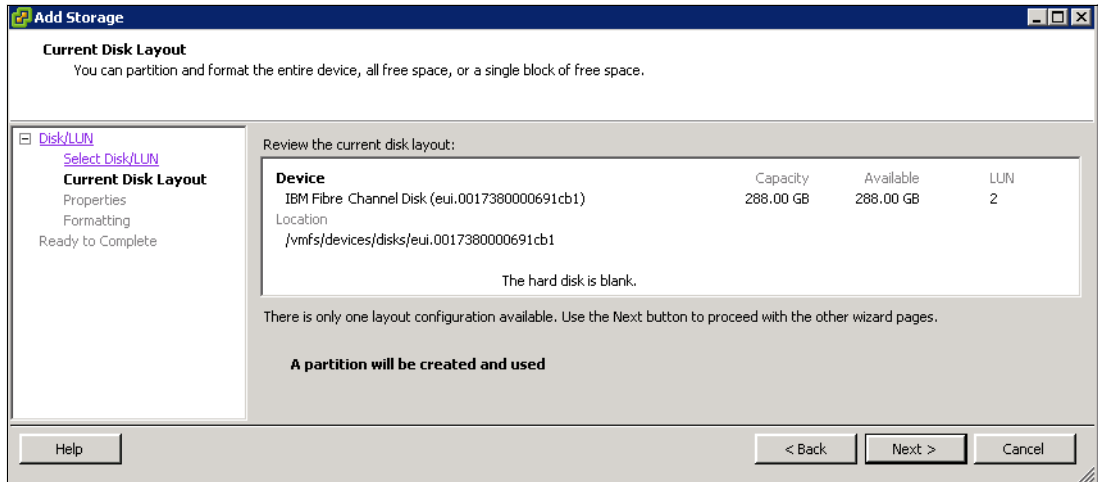


Figure 4-20 Partition parameters

Figure 4-20 shows the partition parameters that are used to create the partition. If you need to change the parameters, click **Back**. Otherwise, click **Next**. The window shown in Figure 4-21 displays.

7. Type a name for the new datastore, and click **Next**. In this example, the name is XIV_demo_store.

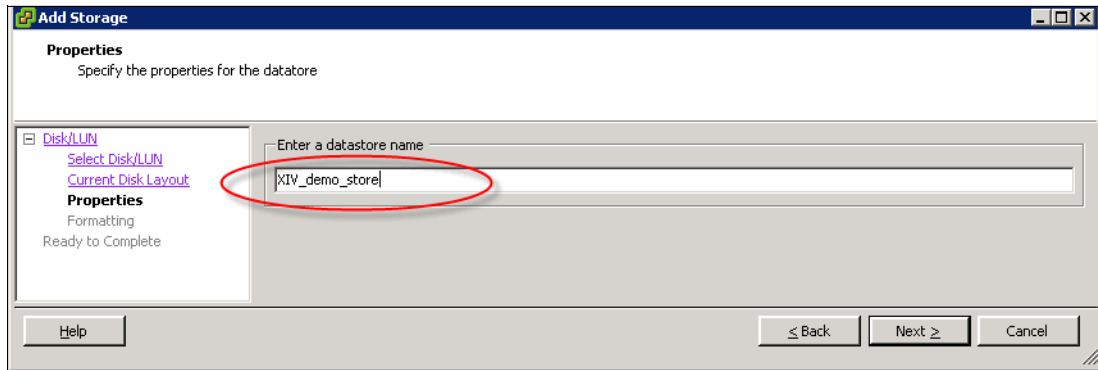


Figure 4-21 edatastore name

8. Enter the file system parameters for your new datastore, and click **Next** to continue. Figure 4-22 on page 60 example shows a 1-MB block size, but you can choose to use a larger size based on your requirements. See:

<http://kb.vmware.com/kb/1003565>

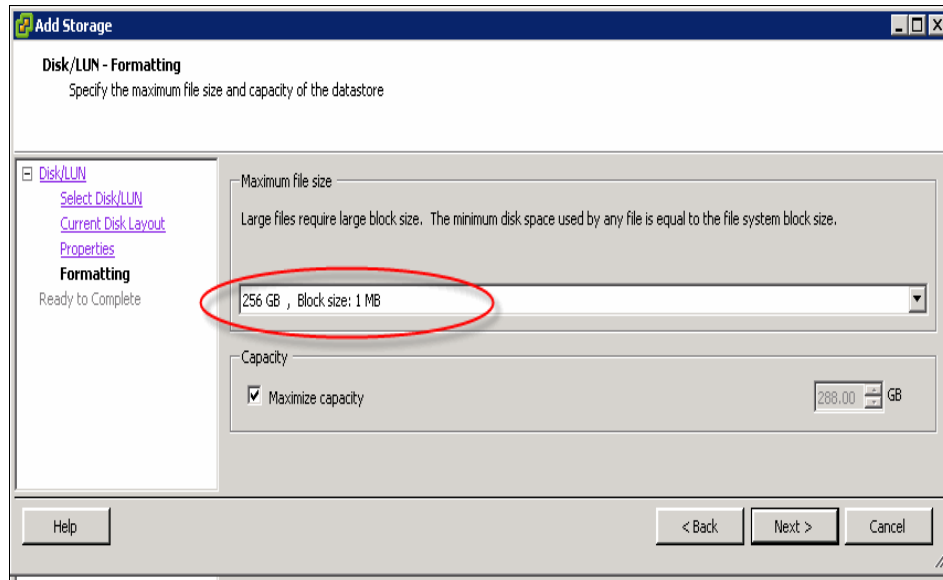


Figure 4-22 Selecting the file system parameters for ESX datastore

Tip: For more information about selecting the correct values for file system parameters for your specific environment, see your VMware ESX/ESXi 4.x documentation.

9. In the summary window shown in Figure 4-23, check the parameters that you entered. If everything is correct, click **Finish**.

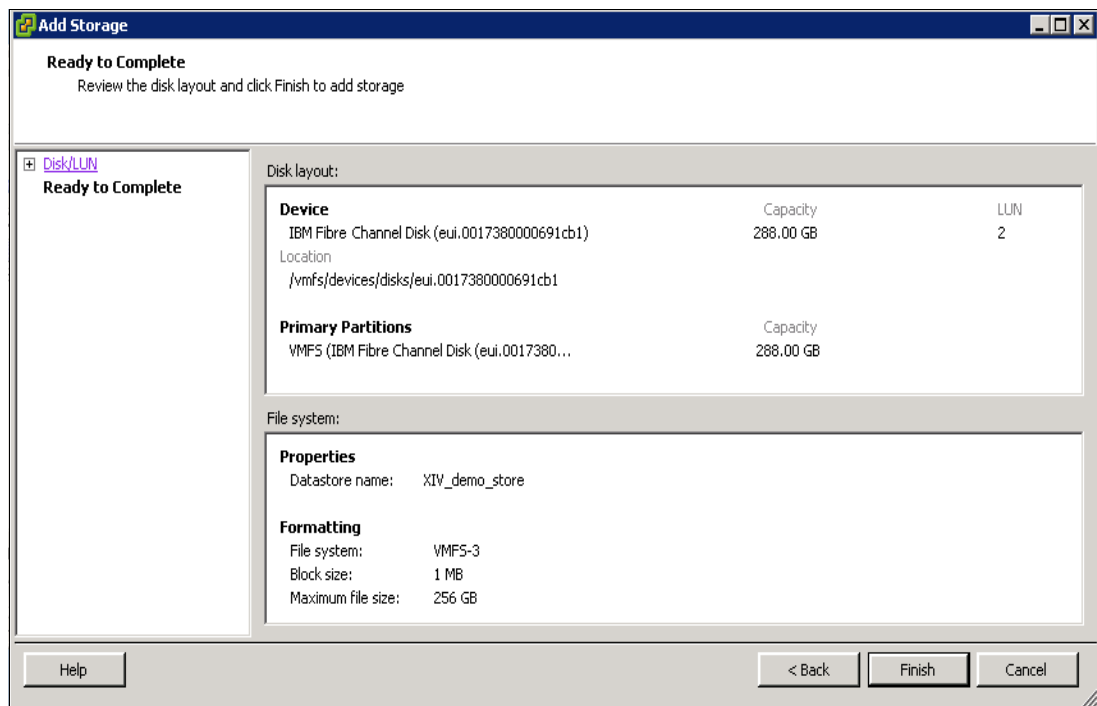


Figure 4-23 Summary of datastore selected parameters

10. In the vSphere Client main window, two new tasks are displayed in the recent task pane, as shown in Figure 4-24. They indicate the completion of the new datastore creation.

Recent Tasks							Name, Target or Status contains:	Clear	X
Name	Target	Status	Details	Initiated by	vCenter Server	Requested Start Time	Start Time		
Create VMFS datastore	9.155.66.222	Completed		Administrator	X3650-LAB-7V1	9/28/2010 1:29:42 PM	9/28/2010		
Compute disk partition ...	9.155.66.222	Completed		Administrator	X3650-LAB-7V1	9/28/2010 1:29:42 PM	9/28/2010		

Tasks Alarms License Period: 344 days remaining Administrator

Figure 4-24 Tasks related to datastore creation

Set up the round-robin policy for the new datastore by following these steps:

1. From the vSphere Client main window (Figure 4-25), you can see a list of all datastores, including the new one you created. Select the datastore you want to change the policy on, and click **Properties**.

View: **Datstores** Devices

Datstores Refresh Delete Add Storage... Rescan All...

Identification	Status	Device	Capacity	Free	Type	Last Upd
datastore1	Normal	Local ServeRA Di...	135.25 GB	43.08 GB	vmfs3	9/28/20
XIV_demo_store	Normal	IBM Fibre Channel...	287.75 GB	287.20 GB	vmfs3	9/28/20
XIV01_site2	Normal	IBM Fibre Channel...	287.75 GB	234.69 GB	vmfs3	9/28/20

Datstore Details Properties...

XIV_demo_store 287.75 GB Capacity

Location: /vmfs/volumes/4ca1d1b6-5...
Hardware Acceleration: Unknown

562.00 MB Used
287.20 GB Free

Path Selection	Properties	Extents
Most Recently Us...	Volume Label: XIV_demo_s...	IBM Fibre Channel Disk (eui... 288.00 GB
	Datastore Name: XIV_demo_s...	Total Formatted Capacity 287.75 GB

Paths
Total: 2
Broken: 0
Disabled: 0

Formatting
File System: VMFS 3.46
Block Size: 1 MB

Figure 4-25 sdatastore updated list

- In the Properties window shown in Figure 4-26, click **Manage Paths**. The Manage Paths window shown in Figure 4-27 is displayed.

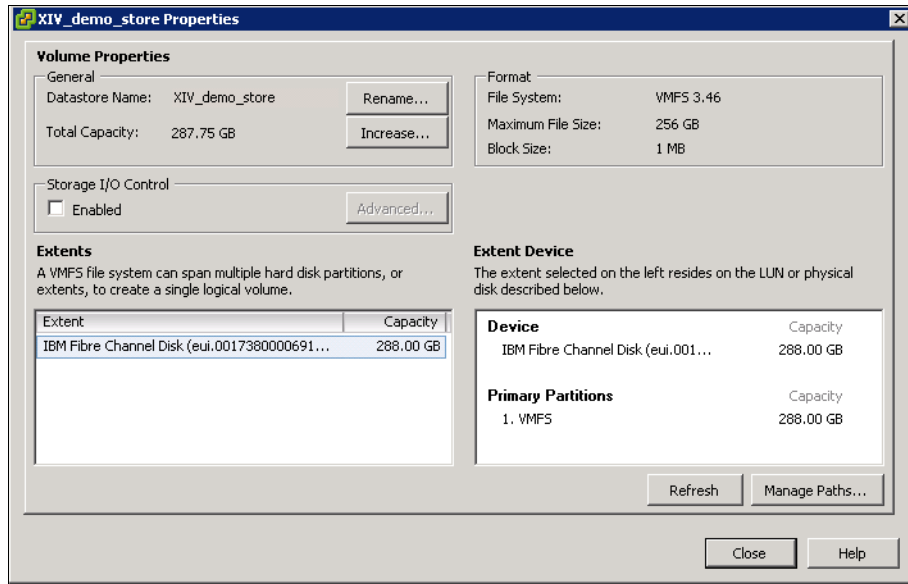


Figure 4-26 edatastore properties

- Select any of the vmhbas listed in Figure 4-27.

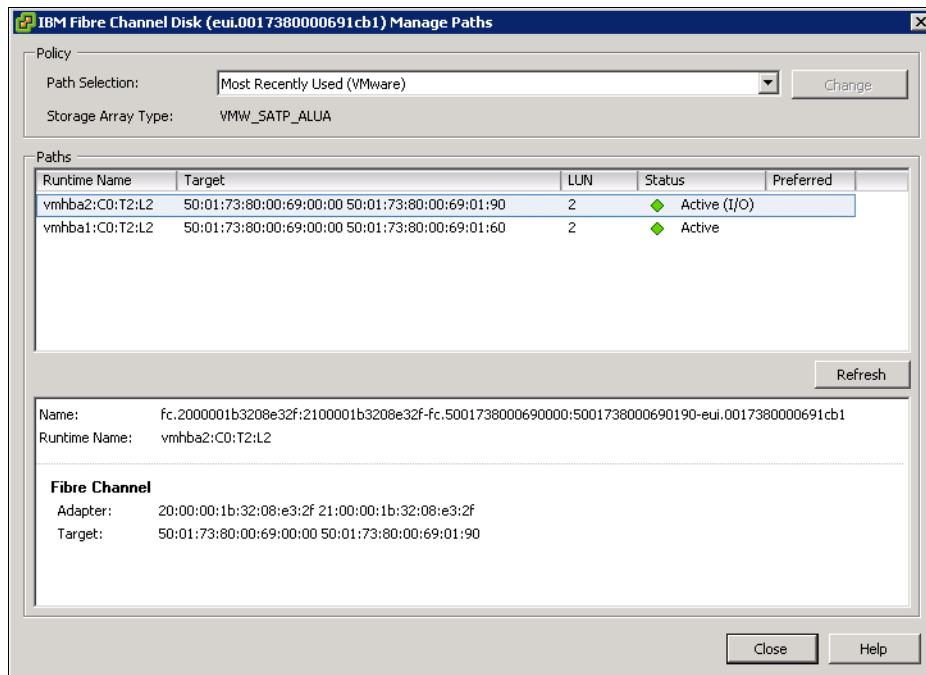


Figure 4-27 Manage Paths window

- Click the **Path selection** → **Round Robin (VMWare)**, as shown in Figure 4-28.

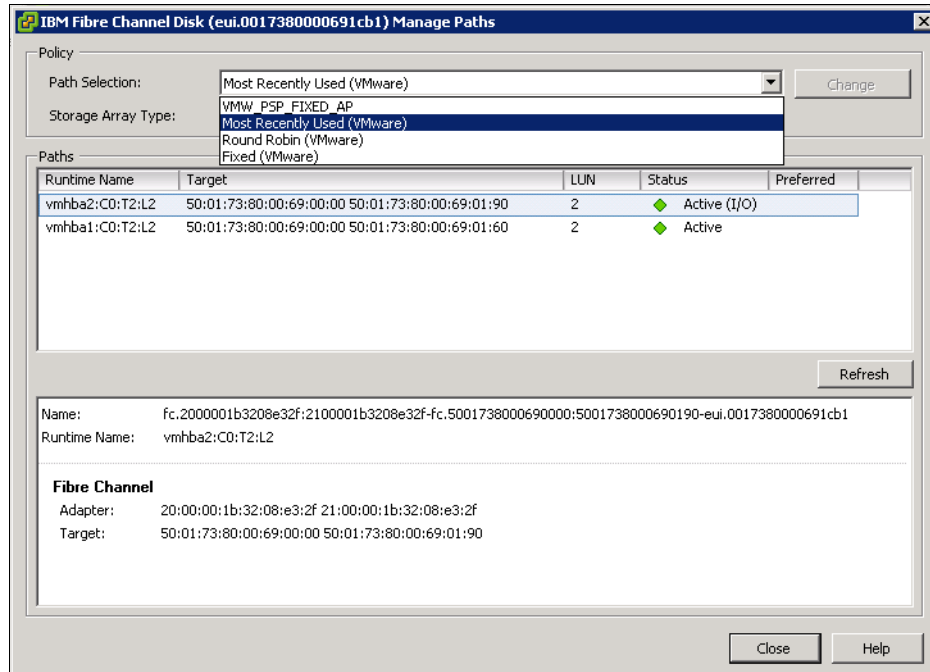


Figure 4-28 List of the path selection options

- Click **Change** to confirm your selection and return to the Manage Paths window, as shown in Figure 4-29.

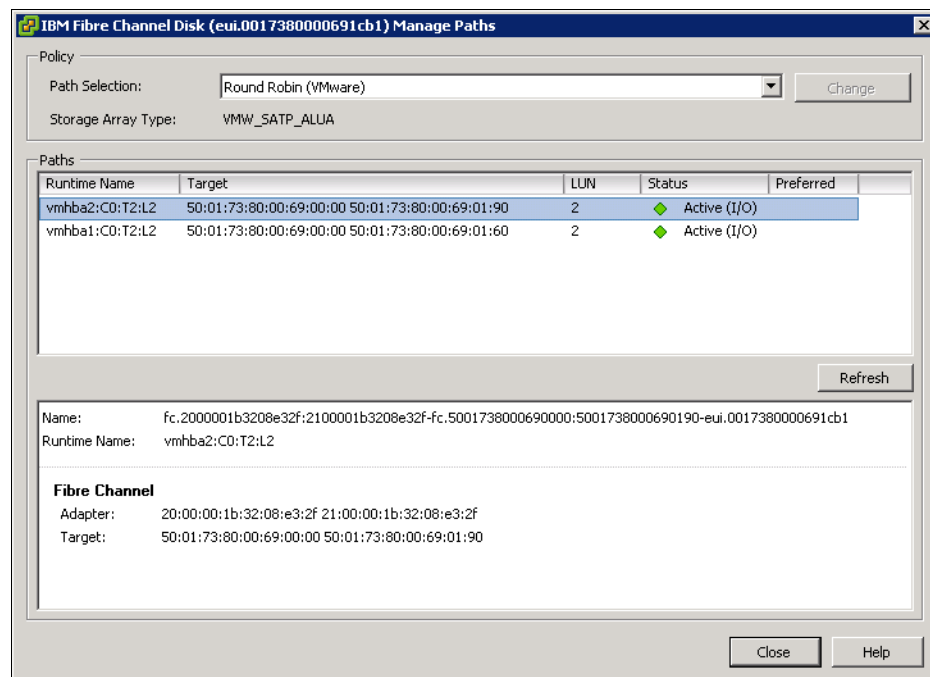


Figure 4-29 edatastore paths with selected round robin policy for multipathing

Apply the round-robin policy to any previously created datastores. Your ESX host is now connected to the XIV Storage System with the correct settings for multipathing.

4.3.6 Performance tuning tips for ESX/ESXi 4.x hosts with XIV

Review settings in ESX/ESXi 4.x to see whether they affect performance in your environment and with your applications.

Settings you might consider changing are:

- ▶ Using larger LUNs rather than LVM extents.
- ▶ Using a smaller number of large LUNs instead of many small LUNs.
- ▶ Increasing the queue size for outstanding I/O on HBA and VMWare kernel levels.
- ▶ Using all available paths for round-robin up to a maximum of 12 paths. For additional considerations on the number of paths, refer to 4.3.2, “Identifying ESX host port WWN” on page 52
- ▶ Decreasing the amount of I/O run by one path when using round-robin.
- ▶ If Windows 2003 guests are using LSI Logic drivers, see the following VMware knowledge base topic regarding block size:

<http://kb.vmware.com/kb/9645697>

Generally, use a maximum block size of 1 MB.

- ▶ You do not need to manually align your VMFS partitions.

Tip: Commands using `esxcli` need either the vSphere CLI installed on a management workstation or the Tech Support Mode enabled on the ESX server itself. Enabling the Tech Support Mode allows remote SSH shell access. If `esxcli` is run from a command prompt without any form of configuration file, each command normally uses the following syntax:

```
esxcli --server 9.155.113.135 --username root --password passw0rd <command>
```

If you run `esxcli` from a Tech Support Mode shell or on a host with UNIX utilities, you can use commands, such as `grep` and `egrep`. For more information, see the following knowledge base topics:

<http://kb.vmware.com/kb/1003677>

<http://kb.vmware.com/kb/1017910>

<http://kb.vmware.com/kb/2004746>

Queue size for outstanding I/O

In general, you do not need to change the HBA queue depth and the corresponding `Disk.SchedNumReqOutstanding` VMWare kernel parameter. When there is one virtual machine active on a volume, set only the maximum queue depth. If there are multiple virtual machines active on a volume, the value of `Disk.SchedNumReqOutstanding` value becomes relevant. The queue depth value is effectively equal to the lower of the queue depth of the adapter and the value of `Disk.SchedNumReqOutstanding`. Generally, set the `Disk.SchedNumReqOutstanding` parameter and the adapter queue depth to the same number. Consider the following suggestions:

- ▶ Set both the `queue_depth` and the `Disk.SchedNumReqOutstanding` VMWare kernel parameter to 128 on an ESX host that has exclusive access to its LUNs.
- ▶ Set both the `queue_depth` and the `Disk.SchedNumReqOutstanding` VMWare kernel parameter to 64 when a few ESX hosts share access to a common group of LUNs.

To change the queue depth:

1. Log on to the service console as root.
2. For Emulex HBAs, verify which Emulex HBA module is currently loaded, as shown in Example 4-4.

Example 4-4 Emulex HBA module identification

```
#vmkload_mod -l|grep lpfc
lpfc820          0x418028689000    0x72000    0x417fe9499f80    0xd000 33 Yes
```

For Qlogic HBAs, verify which Qlogic HBA module is currently loaded, as shown in Example 4-5.

Example 4-5 Qlogic HBA module identification

```
#vmkload_mod -l|grep qla
qla2xxx          2    1144
```

3. Set the new value for the `queue_depth` parameter, and check that new values are applied. For Emulex HBAs, see Example 4-6.

Example 4-6 Setting new value for queue_depth parameter on Emulex FC HBA

```
# esxcfg-module -s lpfc0_lun_queue_depth=64 lpfc820
# esxcfg-module -g lpfc820
lpfc820 enabled = 1 options = 'lpfc0_lun_queue_depth=64'
```

For Qlogic HBAs, see Example 4-7.

Example 4-7 Setting new value for queue_depth parameter on Qlogic FC HBA

```
# esxcfg-module -s ql2xmaxqdepth=64 qla2xxx
# esxcfg-module -g qla2xxx
qla2xxx enabled = 1 options = 'ql2xmaxqdepth=64'
```

You can also change the `queue_depth` parameters on your HBA using the tools or utilities provided by your HBA vendor.

To change the corresponding `Disk.SchedNumReqOutstanding` parameter in the VMWare kernel after changing the HBA queue depth:

1. Start the VMWare vSphere Client, and choose the server for which you plan to change the settings.
2. In the Software section, click the **Configuration** tab, and click **Advanced Settings** to display the Advanced Settings window.

3. Select **Disk** (circled in green in Figure 4-30), and set the new value for **Disk.SchedNumReqOutstanding** (circled in red on Figure 4-30). Click **OK** to save your changes.

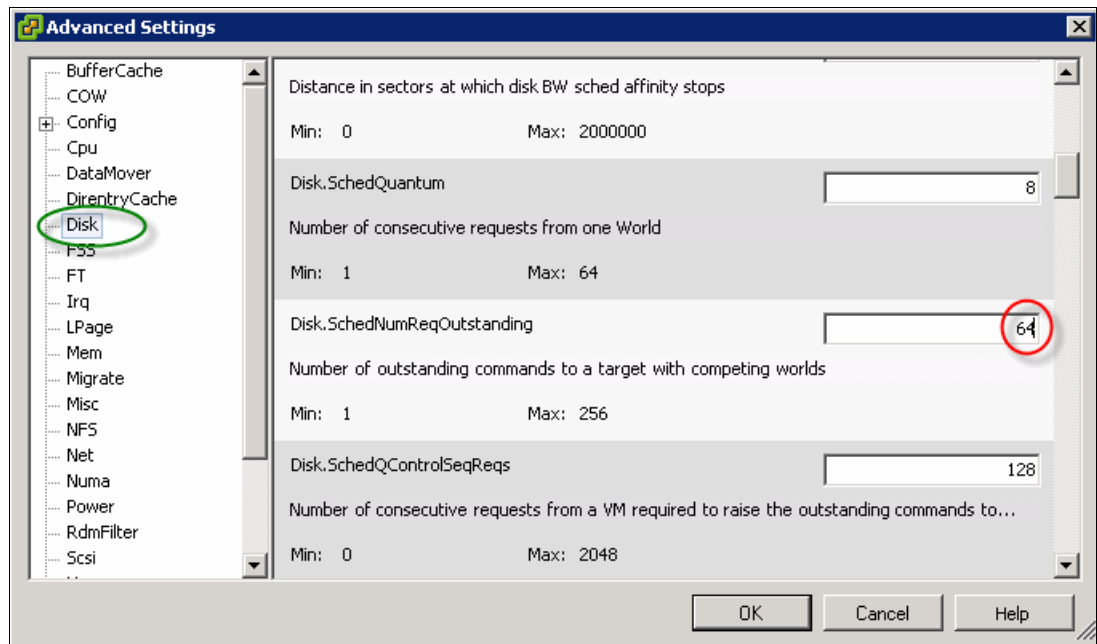


Figure 4-30 Changing `Disk.SchedNumReqOutstanding` parameter in VMware ESX/ESXi 4.x

Tuning multipathing settings for round-robin

Important: The default ESX VMware settings for round-robin are adequate for most workloads. Do not change them normally.

If you need to change the default settings, enable the non-optimal use for round-robin, and decrease the amount of I/O going over each path. This configuration can help the ESX host use more resources on the XIV Storage System.

If you determine that a change is required:

1. Start the VMware vSphere Client, and connect to the vCenter server.
2. From the vSphere Client, select your server.

- Click the **Configuration** tab and select **Storage** in the Hardware section, as shown in Figure 4-31, where you can view the device identifier for your datastore (circled in red).

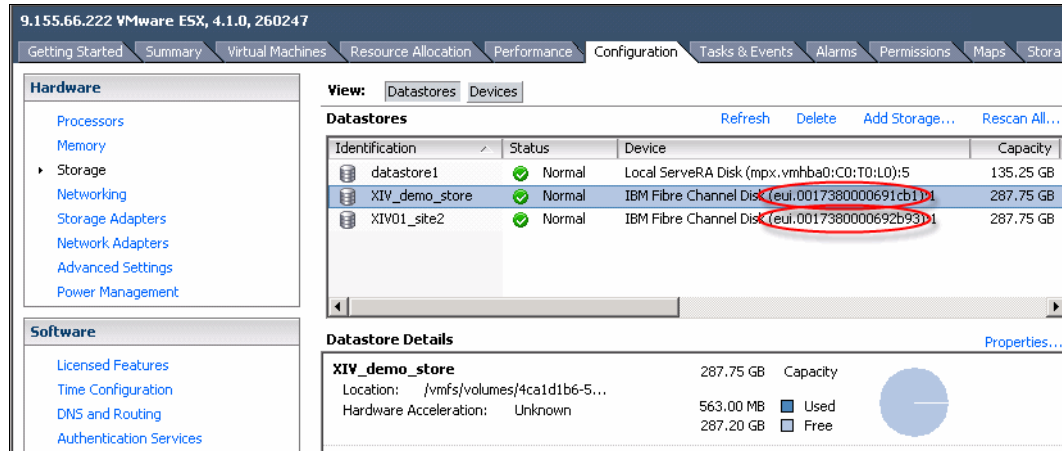


Figure 4-31 Identification of device identifier for your datastore

- Log on to the service console as root or access the esxcli. You can also get the device IDs using the esxcli, as shown in Example 4-8.

Example 4-8 Listing device IDs using esxcli

```
#esxcli nmp device list
eui.00173800278200ff
```

- Enable use of non-optimal paths for round-robin with the `esxcli` command, as shown in Example 4-9.

Example 4-9 Enabling use of non-optimal paths for round-robin on ESX/ESXi 4.x host

```
#esxcli nmp roundrobin setconfig --device eui.0017380000691cb1 --useANO=1
```

- Change the amount of I/O run over each path, as shown in Example 4-10. This example uses a value of 10 for a heavy workload. Leave the default (1000) for normal workloads.

Example 4-10 Changing the amount of I/O run over one path for round-robin algorithm

```
# esxcli nmp roundrobin setconfig --device eui.0017380000691cb1 --iops=10
--type "iops"
```

- Check that your settings are applied, as illustrated in Example 4-11.

Example 4-11 Checking the round-robin options on datastore

```
#esxcli nmp roundrobin getconfig --device eui.0017380000691cb1
Byte Limit: 10485760
Device: eui.0017380000691cb1
I/O Operation Limit: 10
Limit Type: Iops
Use Active Unoptimized Paths: true
```

If you need to apply the same settings to multiple datastores, you can also use scripts similar to the ones shown in Example 4-12 on page 68.

Example 4-12 Setting round-robin tweaks for all IBM XIV Storage System devices

```
# script to display round robin settings
for i in `ls /vmfs/devices/disks/ | grep eui.001738*|grep -v \:~ ; \
do echo "Current round robin settings for device" $i ; \
esxcli nmp roundrobin getconfig --device $i
done

# script to change round robin settings
for i in `ls /vmfs/devices/disks/ | grep eui.001738*|grep -v \:~ ; \
do echo "Update settings for device" $i ; \
esxcli nmp roundrobin setconfig --device $i --useANO=1;\
esxcli nmp roundrobin setconfig --device $i --iops=10 --type "iops";\
done
```

4.3.7 VMware vStorage API Array Integration (VAAI)

Starting with software version 10.2.4a the IBM XIV Storage System supports VAAI for ESX and ESXi 4.1. For more details, see Chapter 3, “VMware vStorage APIs Array Integration” on page 27.

4.4 VMware ESXi 5.0/5.1 and XIV

This section describes attaching ESXi 5.0/5.1 hosts to XIV through Fibre Channel.

4.4.1 ESXi 5.0/5.1 Fibre Channel configuration

The steps required to attach an XIV to a vSphere 5.0 server are similar to the steps in 4.3, “VMware ESX and ESXi 4.x and XIV” on page 52.

To attach an XIV to a vSphere 5.0 server:

1. Identify your ESX host ports, as shown in 4.3.2, “Identifying ESX host port WWN” on page 52.
2. Scan for new LUNs, as shown in “Considerations for the size and quantity of volumes” on page 53.
3. Create your datastore as per 4.3.5, “Configuring ESX/ESXi 4.x host for multipathing with XIV” on page 57. However when adding a datastore, there are three variations from the panels seen in ESXi 4.1.

4. You are prompted to create either a VMFS-5 or VMFS-3 file system, as shown in Figure 4-32. If you do not use VMFS-5, you cannot create a datastore larger than 2 TiB.

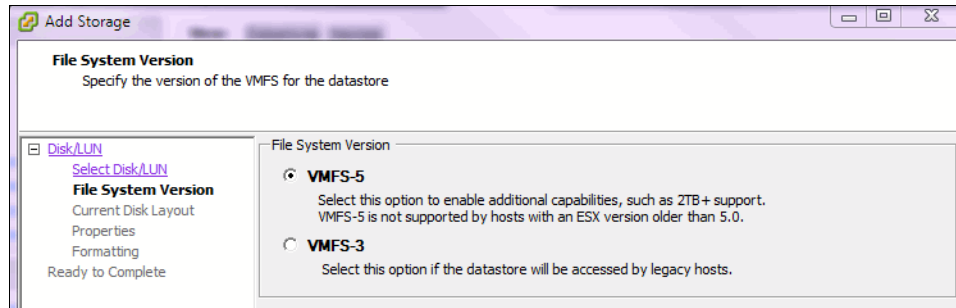


Figure 4-32 edatastore file system prompt in vSphere 5.0

5. If you use VMFS-5, you are not prompted to define a maximum block size. You are given the option to use a custom space setting, limiting the size of the datastore on the volume. You can expand the datastore at a later time to use the remaining space on the volume. However, you cannot use that space for a different datastore.

There is no need to manage the paths to the XIV because round robin must already be in use by default.

Considerations for the size and quantity of volumes

The following configuration maximums are documented for vSphere 5.0 and vSphere 5.1:

- ▶ The maximum number of LUNs per server is 256.
- ▶ The maximum number of paths per server is 1024.
- ▶ The maximum number of paths per LUN is 32.

These facts have some important design considerations. If each XIV volume can be accessed through 12 fabric paths (which is a large number of paths), the maximum number of volumes is 85. Dropping the paths to a more reasonable count of six increases the maximum LUN count to 170. For installations with large numbers of raw device mappings, these limits can become a major constraint.

More details are at:

<http://www.vmware.com/pdf/vsphere5/r50/vsphere-50-configuration-maximums.pdf>

4.4.2 Performance tuning tips for ESXi 5 hosts with XIV

Performance tips for ESXi 5.0/5.1 are similar to those in 4.3.6, “Performance tuning tips for ESX/ESXi 4.x hosts with XIV” on page 64. However the syntax of some commands changed, so they are documented here.

Queue size for outstanding I/O

In general, it is not necessary to change the HBA queue depth and the corresponding `Disk.SchedNumReqOutstanding` VMWare kernel parameter. If more than one virtual machine is active on a volume, you need to set only the maximum queue depth. If there are multiple virtual machines active on a volume, the value of `Disk.SchedNumReqOutstanding` becomes relevant. The queue depth value is effectively equal to the lower of the queue depth of the adapter and the value of `Disk.SchedNumReqOutstanding`. Normally, set both the `Disk.SchedNumReqOutstanding` parameter and the adapter queue depth to the same number.

Tip: Commands using `esxcli` require either the vSphere CLI installed on a management workstation or the Tech Support Mode enabled on the ESXi server. Enabling Tech Support Mode also allows remote SSH shell access. If `esxcli` is run from a command prompt without any form of configuration file, the command uses the following syntax:

```
esxcli --server 9.155.113.135 --username root --password passw0rd <command>
```

If `esxcli` is run from a Tech Support Mode shell or on a host with UNIX utilities, commands like `grep` and `egrep` can be used. For more information, see:

<http://kb.vmware.com/kb/1017910>

<http://kb.vmware.com/kb/2004746>

To set the queue size:

1. Issue the `esxcli system module list` command to determine which HBA type you have (Example 4-13). The output looks similar to Example 4-4 on page 65. However, in this example both HBA types are suggested, which is not usual.

Example 4-13 Using the module list command

```
# esxcli system module list | egrep "qla|lpfc"
Name                Is Loaded  Is Enabled
-----
qla2xxx              true       true
or
lpfc820              true       true
```

2. Set the queue depth for the relevant HBA type. In both Example 4-14 and Example 4-15, the queue depth is changed to 64. In Example 4-14, the queue depth is set for an Emulex HBA.

Example 4-14 Setting new value for queue_depth parameter on Emulex FC HBA

```
# esxcli system module parameters set -p lpfc0_lun_queue_depth=64 lpfc820
```

In Example 4-15 the queue depth is set for a Qlogic HBA.

Example 4-15 Setting new value for queue_depth parameter on Qlogic FC HBA

```
# esxcli system module parameters set -p ql2xmaxqdepth=64 -m qla2xxx
```

3. Reboot your ESXi server. After reboot, confirm that the new settings are applied using the command shown in Example 4-16. The example shows only one of a great many parameters. You need to change the syntax if you have an Emulex HBA.

Example 4-16 Checking the queue depth setting for a Qlogic HBA

```
# esxcli system module parameters list -m qla2xxx | grep qdepth
Name                Type  Value  Description
-----
ql2xmaxqdepth      int   64     Maximum queue depth
```

After changing the HBA queue depth, change the Disk.SchedNumReqOutstanding parameter in the VMWare kernel. To change the parameter:

1. Start the VMWare vSphere Client.
2. Select the server for which you plan to change the settings.
3. Click the **Configuration** tab under Software section, and click **Advanced Settings** to display the Advanced Settings window.
4. Select **Disk** (circled in green in Figure 4-33) and set the new value for Disk.SchedNumReqOutstanding (circled in red on Figure 4-33).

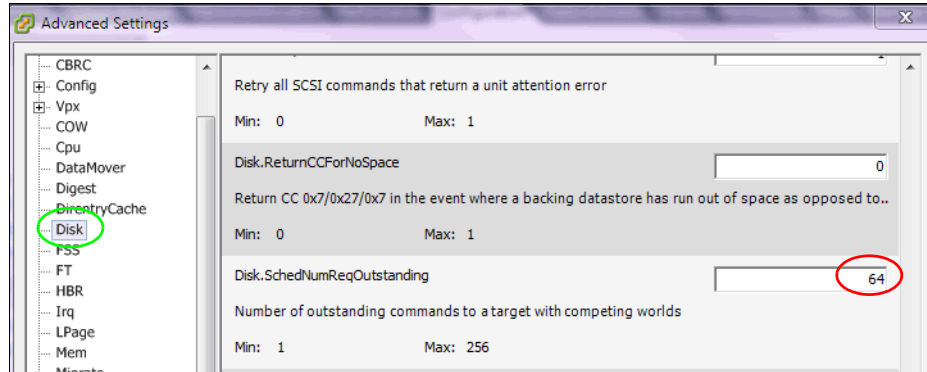


Figure 4-33 Changing Disk.SchedNumReqOutstanding parameter in VMWare ESXi 5

5. Click **OK** to save your changes.

Tuning multipathing settings for round-robin

Important: The default ESX VMware settings for round-robin are adequate for most workloads and must not normally be changed.

If you need to change the default settings, enable the non-optimal use for round-robin and decrease the amount of I/O going over each path. This configuration can help the ESX host use more resources on the XIV Storage System.

If you determine that a change is required, follow these instructions:

1. Start the VMware vSphere Client, and connect to the vCenter server.
2. From the vSphere Client, select your server, click the **Configuration** tab, and select **Storage** in the Hardware section, as shown in Figure 4-34.

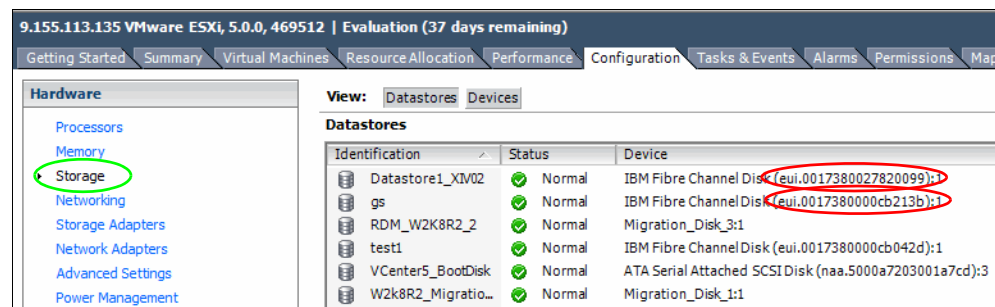


Figure 4-34 Identification of device identifier for your datastore

Here you can view the device identifier for your datastore (circled in red). You can also get this information using ESXCLI, as shown in Example 4-17.

Example 4-17 Listing storage devices

```
# esxcli storage nmp device list | grep "IBM Fibre Channel Disk (eui.001738"
  Device Display Name: IBM Fibre Channel Disk (eui.0017380000cb11a1)
  Device Display Name: IBM Fibre Channel Disk (eui.0017380027820099)
  Device Display Name: IBM Fibre Channel Disk (eui.00173800278203f4)
```

3. Change the amount of I/O run over each path, as shown in Example 4-18. This example uses a value of 10 for a heavy workload. Leave the default (1000) for normal workloads.

Example 4-18 Changing the amount of I/O run over one path for round-robin algorithm

```
# esxcli storage nmp psp roundrobin deviceconfig set --iops=10 --type "iops"
--device eui.0017380000cb11a1
```

4. Check that your settings are applied, as illustrated in Example 4-19.

Example 4-19 Checking the round-robin options on the datastore

```
#esxcli storage nmp device list --device eui.0017380000cb11a1
eui.0017380000cb11a1
  Device Display Name: IBM Fibre Channel Disk (eui.0017380000cb11a1)
  Storage Array Type: VMW_SATP_ALUA
  Storage Array Type Device Config: {implicit_support=on;explicit_support=off;
explicit_allow=on;alua_followover=on;{TPG_id=0,TPG_state=A0}}
  Path Selection Policy: VMW_PSP_RR
  Path Selection Policy Device Config:
{policy=iops,iops=10,bytes=10485760,useA
NO=0,lastPathIndex=1: NumIOsPending=0,numBytesPending=0}
  Path Selection Policy Device Custom Config:
  Working Paths: vmhba1:C0:T6:L1, vmhba1:C0:T5:L1, vmhba2:C0:T6:L1,
vmhba2:C0:T
5:L1
```

If you need to apply the same settings to multiple datastores, you can also use scripts similar to the ones shown in Example 4-20.

Example 4-20 Setting round-robin tweaks for all IBM XIV Storage System devices

```
# script to display round robin settings
for i in `ls /vmfs/devices/disks/ | grep eui.001738*|grep -v \:` ; \
do echo "*** Current settings for device" $i ; \
esxcli storage nmp device list --device $i
done

# script to change round robin settings
for i in `ls /vmfs/devices/disks/ | grep eui.001738*|grep -v \:` ; \
do echo "Update settings for device" $i ; \
esxcli storage nmp psp roundrobin deviceconfig set --device $i --iops=1000 --type
"iops";\
done
```

4.4.3 Creating datastores that are larger than 2 TiB

With VMFS-3, the largest possible datastore is 2 TiB. With VMFS-5 (introduced in vSphere 5.0), this limit is raised to 64 TiB. Combined with Atomic Test & Set (ATS), the VAAI primitive that current XIV software levels support, you can use much larger datastores. ATS locks only the blocks containing the relevant metadata when acquiring the on-disk locks necessary to perform metadata updates, rather than implementing SCSI2 reservations to serialize host access with a minimum scope of an entire LUN backing the datastore. This procedure improves performance and eliminates the risk of SCSI reservation conflicts.

Do not create a single giant datastore rather than multiple smaller ones for the following reasons:

- ▶ Each XIV volume is assigned a SCSI queue depth by ESXi. More volumes mean more SCSI queues, which means more commands can be issued at any one time.
- ▶ The maximum number of concurrent storage vMotions per datastore is still limited to 8.

Presenting an XIV volume larger than 64 TiB

If an XIV volume larger than 64 TiB is mapped to an ESXi 5.0/5.1 server, a datastore formatted with VMFS-5 uses only the first 64 TiB. In Figure 4-35, a 68.36 TiB XIV volume is presented to ESXi 5.0/5.1, but only the first 64 TiB is used.

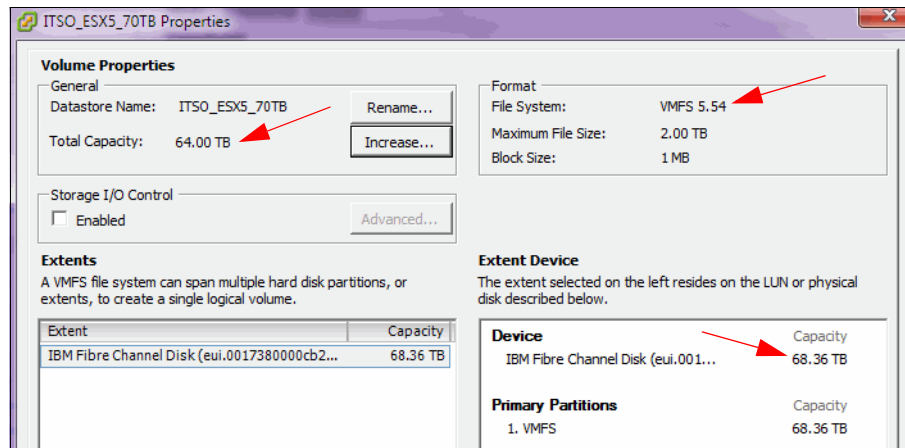


Figure 4-35 vSphere 5.0 volume larger than 64 TiB

If you want to create a datastore that approaches the maximum size, limit the maximum XIV volume size as follows:

2nd generation XIV 70368 GB (65536 GiB or 137438953472 Blocks)

XIV Gen3 70364 GB (65531 GiB or 137428467712 Blocks)

Example 4-21 shows the largest possible datastore, which is exactly 64 TiB in size. The `df` command was run on the ESX server using the tech support mode shell.

Example 4-21 Largest possible VMFS-5 datastore

```

~ # df
file system          Bytes          Used          Available Use% Mounted on
VMFS-5               44560285696 37578866688    6981419008 84% /vmfs/volumes/Boot
VMFS-5               70368744177664 1361051648 70367383126016 0% /vmfs/volumes/Giant

```




IBM Storage Management Console for VMware vCenter

This chapter describes the IBM Storage Management Console for VMware vCenter.

The IBM Storage Management Console for VMware vCenter enables VMware administrators to independently and centrally manage their storage resources on IBM storage systems.

5.1 The IBM Storage Management Console for VMware vCenter

The IBM Storage Management Console for VMware vCenter is a software plug-in that integrates into the VMware vCenter server platform. It enables VMware administrators to independently and centrally manage their storage resources on IBM storage systems. These resources include XIV, Storwize V7000, and SAN Volume Controller.

The plug-in runs as a Microsoft Windows Server service on the vCenter server. Any VMware vSphere Client that connects to the vCenter server detects the service on the server. The service then automatically enables the IBM storage management features on the vSphere Client.

Version 2.6.0 of the plug-in added support for both XIV Gen 3 (using Version 11) and VMware vCenter version 5.0. However, the remainder of this discussion will also address version 3.1.0 of the IBM Storage Management Console for VMware vCenter, which is the version that is installed as of the time of this writing.

5.1.1 Installation

Install the IBM Storage Management Console for VMware vCenter onto the Windows server that is running VMWare vCenter version 4.0, 4.1, 5.0, or 5.1. There are separate installation packages for x86 and x64. You can save time by downloading the correct package for your server architecture. During package installation, you are prompted to do the following steps:

1. Confirm which language you want to use.
2. Accept license agreements.
3. Select an installation directory location (a default location is offered).
4. When installation completes, a command-line configuration wizard starts, as shown in Example 5-1. Normally you can safely accept each prompt in the wizard. When prompted for a user name and password, you need a user ID that is able to log on to the VMware vSphere Client. If you do not either change or accept the SSL certificate, you get Certificate Warning windows when starting the Client. For more information about how to replace the SSL certificate, see the plug-in user guide.

Example 5-1 IBM Storage Management Console for VMWare vCenter Configuration wizard

```
Welcome to the IBM Storage Management Console for VMware vCenter setup wizard, version 2.6.0.
Use this wizard to configure the IBM Storage Management Console for VMware vCenter.
Press [ENTER] to proceed.
```

```
-----
The Wizard will now install the Management Console service and register the extension in the
vCenter server.
```

```
Do you want to continue? [default: yes ]:
```

```
-----
The IBM Storage Management Console for VMware vCenter requires a valid username for connecting
to the vCenter server.
```

```
This user should have permission to register the Plug-in in the Plug-ins Manager.
```

```
Please enter a username : Administrator
```

```
-----
Please enter the password for user Administrator :
```

```
-----
The IBM Storage Management Console for VMware vCenter web component requires a valid network
port number.
```

```
Please enter a port number for the web component [default: 8880 ]:
```

Please wait while configuring the service and registering the extension

The IBM Storage Management Console for VMware vCenter is now configured.
This product is using an SSL certificate which is not signed.
Please consult the User Guide for SSL certificate replacement instructions.
Press [ENTER] to exit.

5. After you configure the IBM Storage Management Console for VMware vCenter, restart the client if you were already logged on. A new IBM Storage icon plus a new IBM Storage tab with all their associated functions are added to the VMware vSphere Client.

You can access the IBM Storage icon from the Home view, as shown in Figure 5-1.

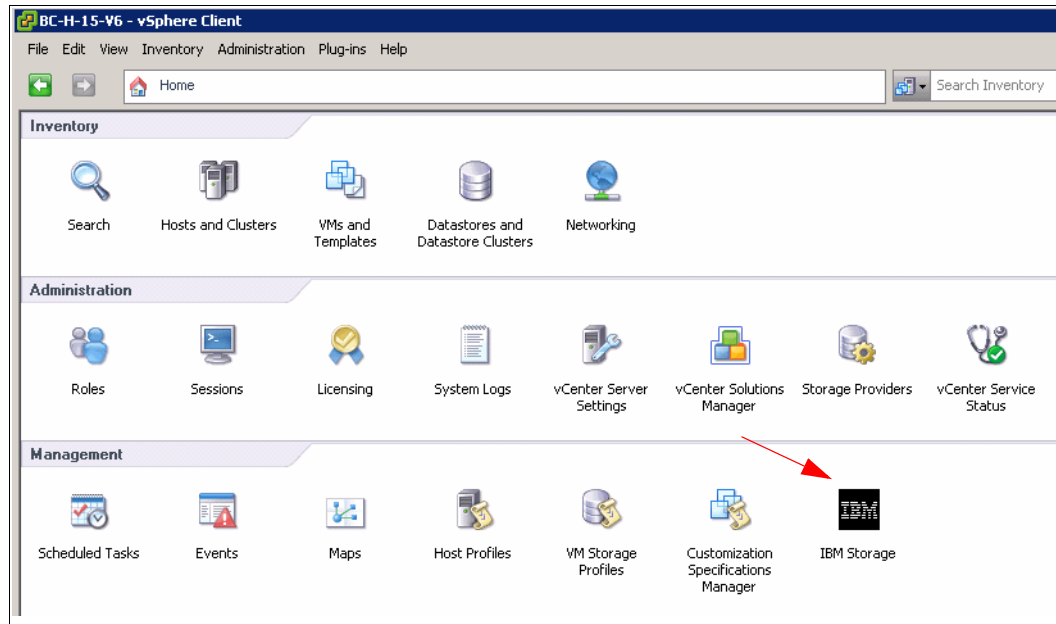


Figure 5-1 IBM Storage plug-in from Home menu

5.1.2 Customizing the plug-in

There are several options to customize the plug-in documented in the user guide. The relevant area in the registry is shown in Figure 5-2 on page 78.

Several parameters can be changed. To modify these registry parameters from the Windows task bar:

1. Click **Start** → **Run**. The Run dialog box is displayed.
2. Type **regedit** and then select **OK**. The Registry Editor is displayed.
3. Go to the following registry tree path:
HKEY_LOCAL_MACHINE → SYSTEM → CurrentControlSet → Services → IBMConsoleForvCenter → Parameters

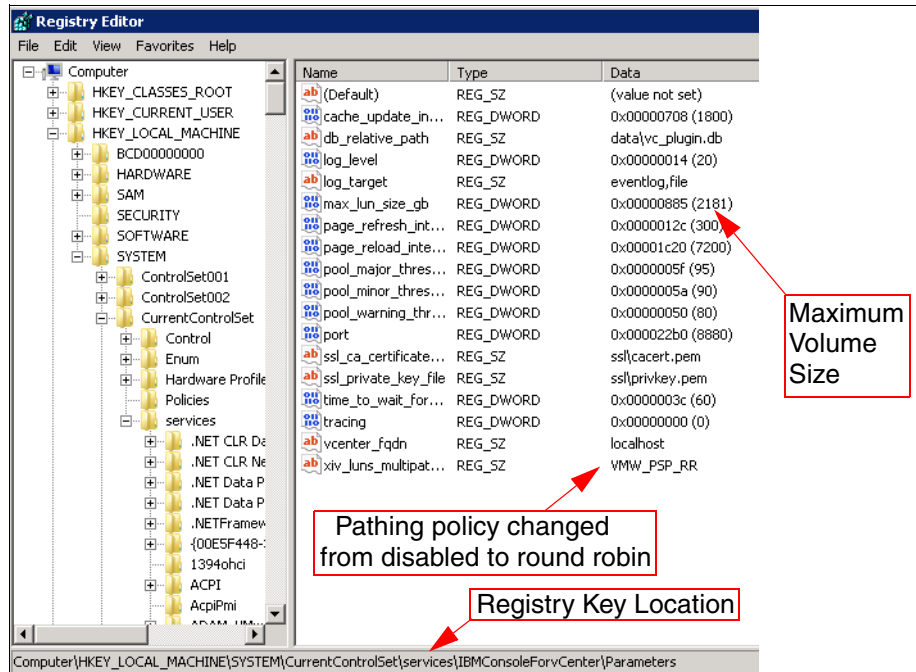


Figure 5-2 vCenter Server registry settings

After you make the required registry modifications:

1. Close the Registry Editor.
2. Close the vSphere Client application. Users connected remotely must also close their client applications.
3. Click **Start** → **Run** to open the Windows Services window. The **Run** dialog box is displayed.
4. Type `services.msc` and then select **OK**.
5. Stop and then start the following Windows service IBM Storage Management Console for VMware vCenter, as shown in Figure 5-3. You can then close the services console.

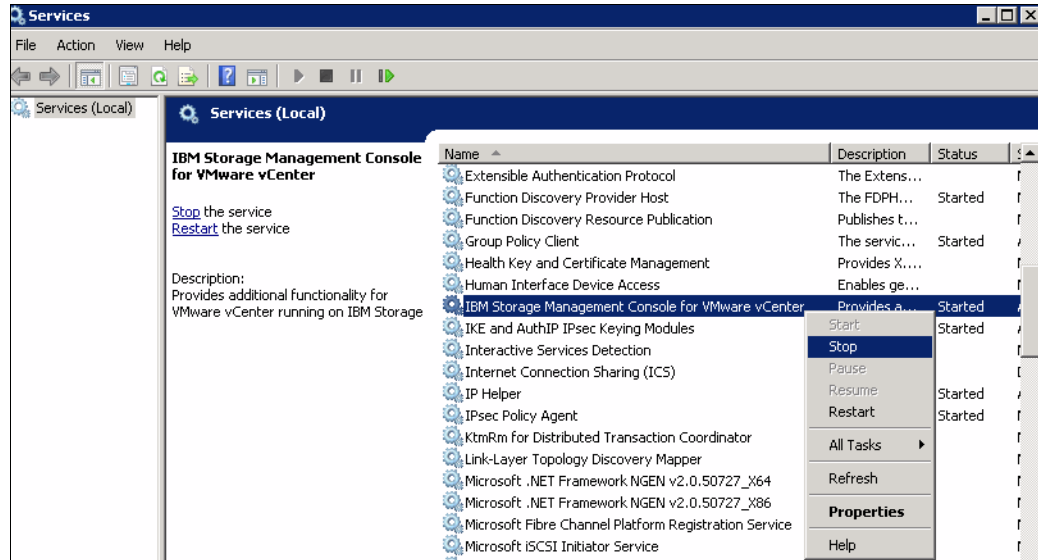


Figure 5-3 Stopping the vCenter plug-in

6. Start the vSphere Client application.

Two possible changes you might consider are in the following sections.

Maximum volume size

The maximum volume size is set to 2 TiB (2181 GB) because this is the largest volume size that VMFS-3 can work with. If you move to VMFS-5, you can use a larger volume size. To modify this parameter, select **max_lun_size_gb** and change the *Base* value from Hexadecimal to Decimal. Enter a new value in decimal. Figure 5-4 shows the default value.

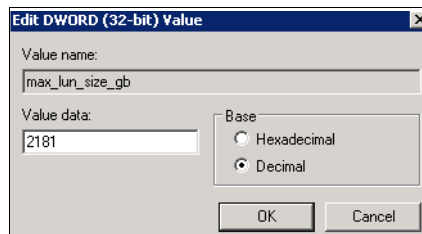


Figure 5-4 Change maximum volume size

For more information about maximum values, see 4.4.3, “Creating datastores that are larger than 2 TiB” on page 73. To change this value globally for the plug-in, ensure that you create 2181 GB or smaller volumes for VMFS-3.

Automatic multipath policy

Automatic multipath policy can be set to ensure all existing and new XIV volumes use round robin mode. The policy also checks and corrects the multipathing settings every 30 minutes to ensure that they continue to use round robin. Automatic multipath is not the default setting due to restrictions with Windows cluster virtual quorum devices. To enable this function, access the registry and then change the value from disabled to VMW_PSP_RR as shown in Figure 5-5.

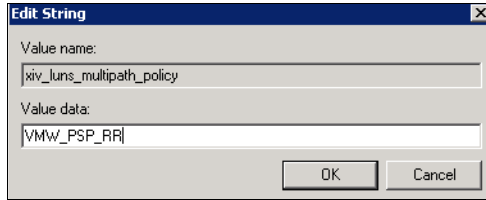


Figure 5-5 Changing multipathing policy

5.1.3 Adding IBM Storage to the plug-in

To add IBM storage to the plug-in:

1. From the Home view of the vSphere Client, double-click the **IBM Storage** icon.
2. The Storage Systems view opens showing all defined IBM Storage Systems. Select the option to **Add** a Storage System, as shown in Figure 5-6 on page 80.

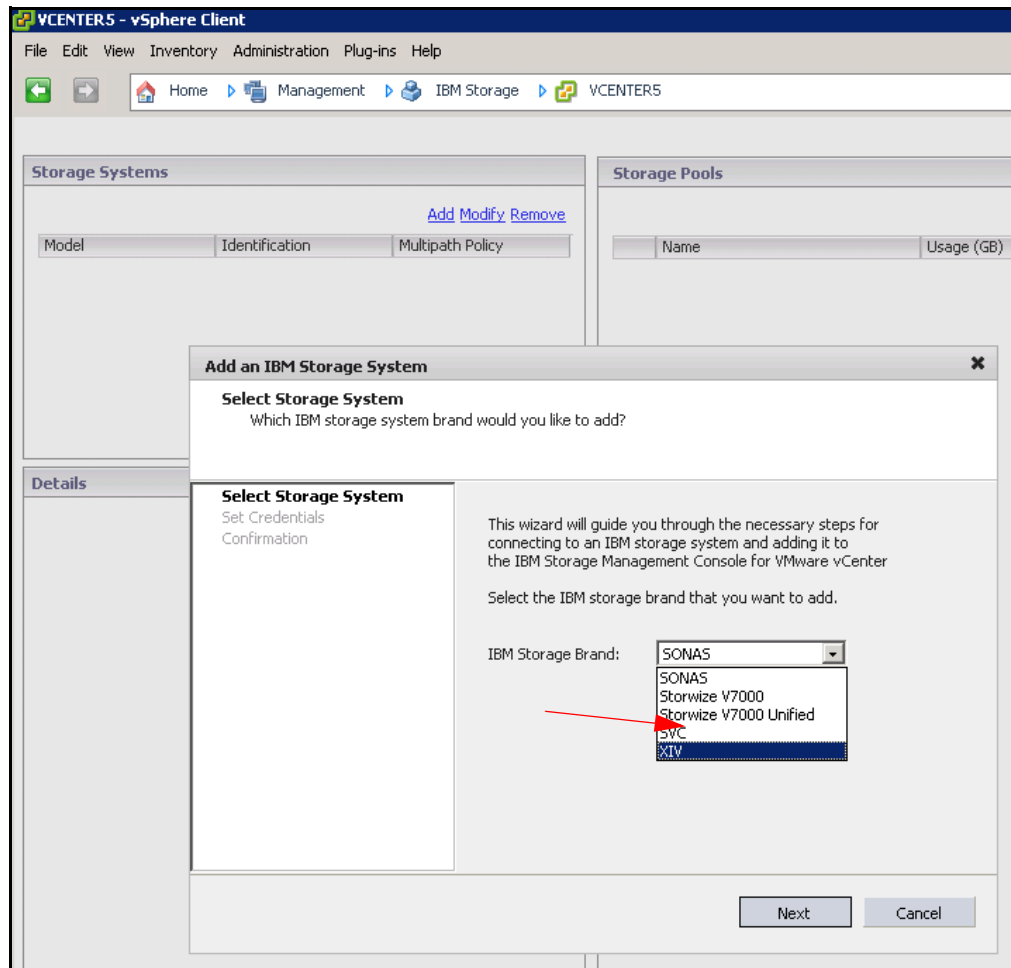


Figure 5-6 Selecting the Add option

3. A window prompts you for an IP address, user name, and password, as shown in Figure 5-7. Use an XIV management IP address, user ID, and password.

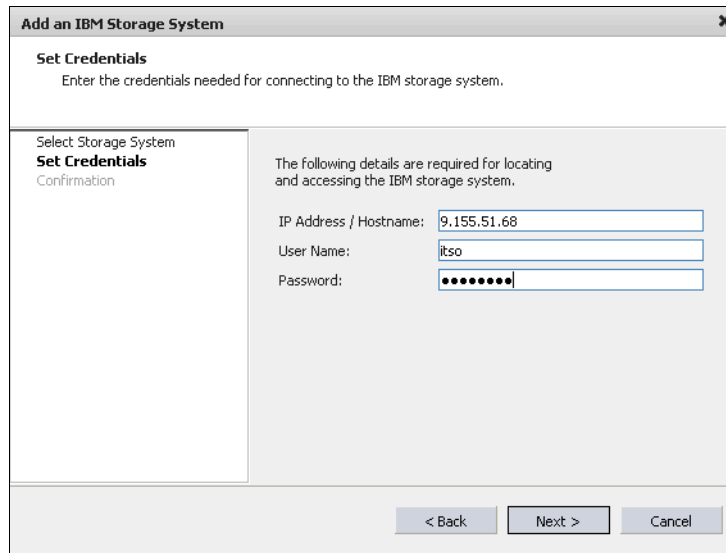


Figure 5-7 Adding an XIV to the plug-in

4. If the vSphere Client connects to the XIV successfully you get a list of storage pools on that XIV. Select the pools that the VMware administrator will allocate storage from, as shown in Figure 5-8. Additional pools can be added later if required.

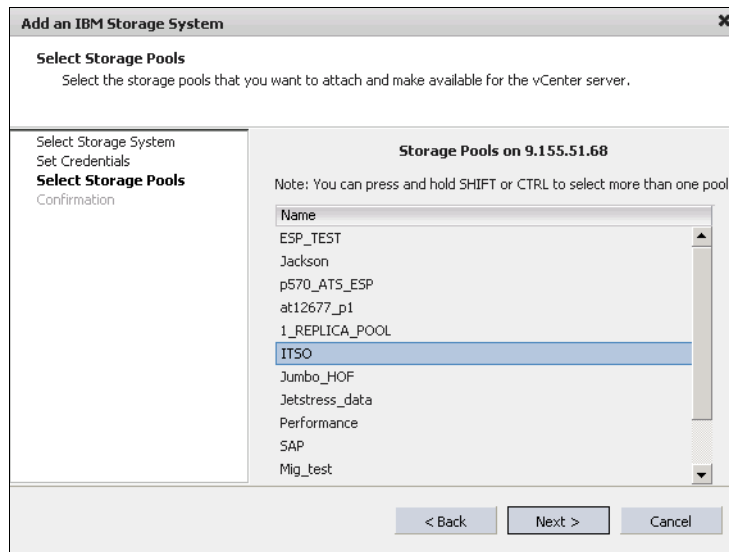


Figure 5-8 Selecting a pool from the plug-in

- The XIV is displayed in your list of Storage Systems. Although you must define only one management IP address (out of three), all three are discovered, as shown in Figure 5-9.

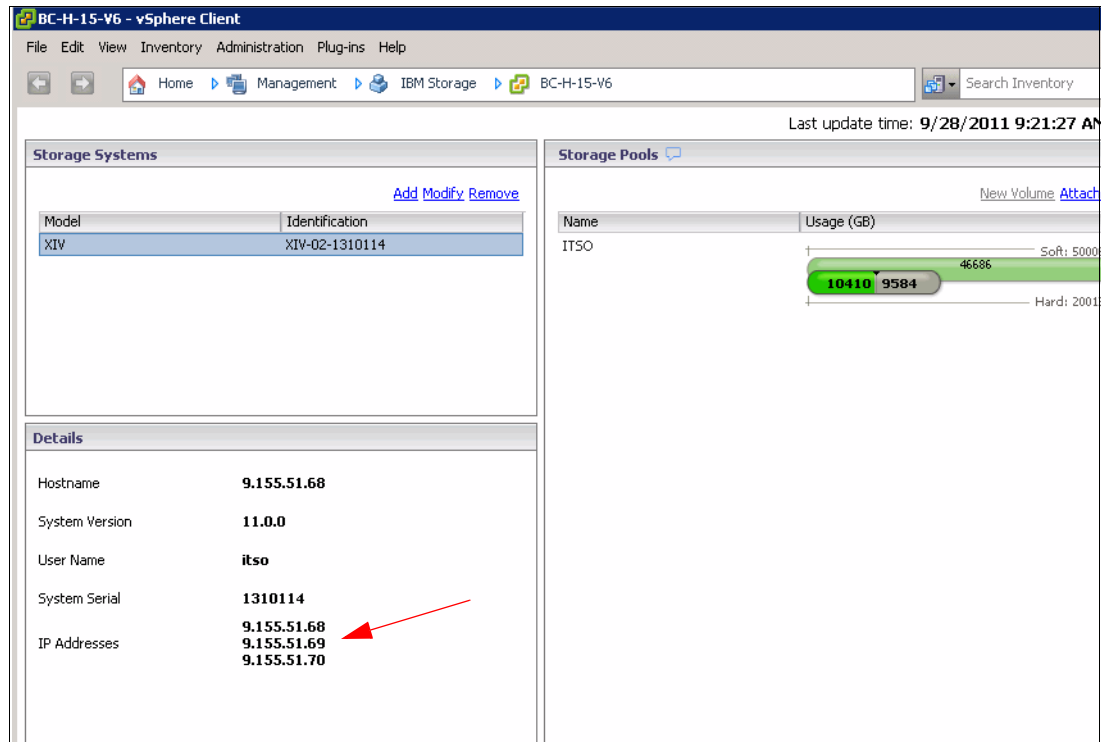


Figure 5-9 vSphere IBM Storage plug-in

- Select an XIV from the Storage Systems box and then select a pool from the Storage Pools box.
- Select the **New Volume** option to create a volume.

Tip: When creating a volume, use the same name for both the volume and the datastore. Using the same name ensures that the datastore and volume names are consistent.

- You are prompted to map the volume to either individual VMware servers or the whole cluster. Normally you select the entire cluster.

For the plug-in to work successfully, the SAN zoning to allow SAN communication between the VMWare cluster and the XIV must already be completed. On the XIV, the Cluster and hosts definitions (representing the VMWare cluster and its servers) must have also been created. This process cannot be done from the plug-in, and is not done automatically. If the zoning and host definitions are not done, volume creation fails and the requested volume is created and then deleted.

Tip: If you perform changes, such as renaming or resizing volumes, updates might take up to 60 seconds to display in the plug-in.

5.1.4 Checking and matching XIV Volumes

Use the IBM Storage tab to identify the properties of the volume. The IBM Storage tab allows you to perform many useful storage tasks. From this tab, shown in Figure 5-10 on page 83, you can perform these tasks:

- ▶ Extend a volume. This task allows you to grow an existing volume and then later resize the datastore using that volume.
- ▶ Rename a volume. Use this task to ensure that the XIV volume name and the datastore name are the same.
- ▶ Move a volume to a different pool on the XIV.
- ▶ Confirm which datastore is which XIV volume.
- ▶ Confirm the status of all of the volumes, snapshots, and mirrors. Mirrors cannot be confirmed if the user has read-only access.
- ▶ Confirm the size of the datastore in binary GiB and decimal GB.
- ▶ If the volume is not being used by a datastore, it can be unmapped and deleted. This process allows a VMware administrator safely return a volume back to the XIV.

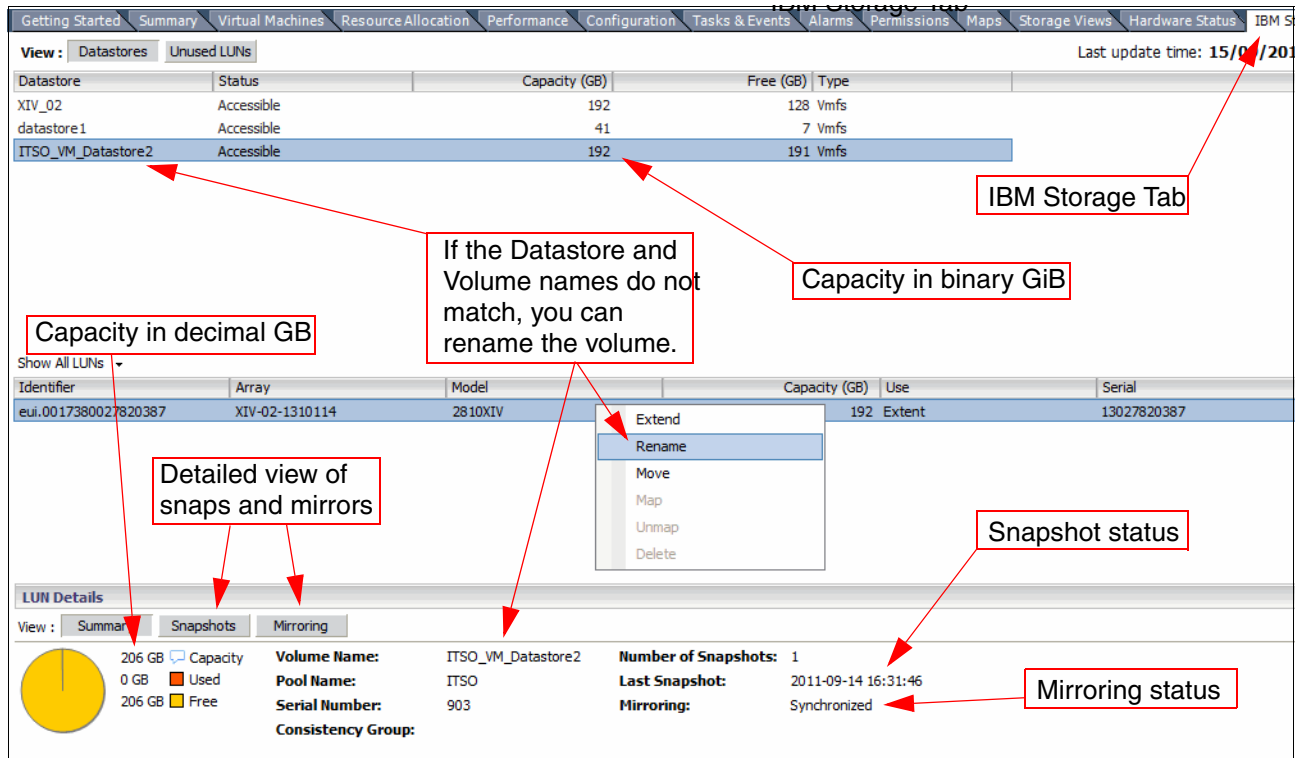


Figure 5-10 The IBM Storage tab added by the plug-in

5.1.5 Creating a datastore

Creating a datastore involves the same steps as 4.3.5, “Configuring ESX/ESXi 4.x host for multipathing with XIV” on page 57 for ESX/ESXi 4.x. ESXi 5.0/5.1 is addressed in 4.4.1, “ESXi 5.0/5.1 Fibre Channel configuration” on page 68.

Before creating the datastore:

1. Open the IBM Storage tab to confirm the LUN identifier.
2. Click the Unused LUNs tab to locate your newly created XIV volume, as shown in Figure 5-11 on page 84. Take note of the Identifier for each unused LUN (for example *eui.0017380027821838*). You need these identifiers to cross match when creating datastores.

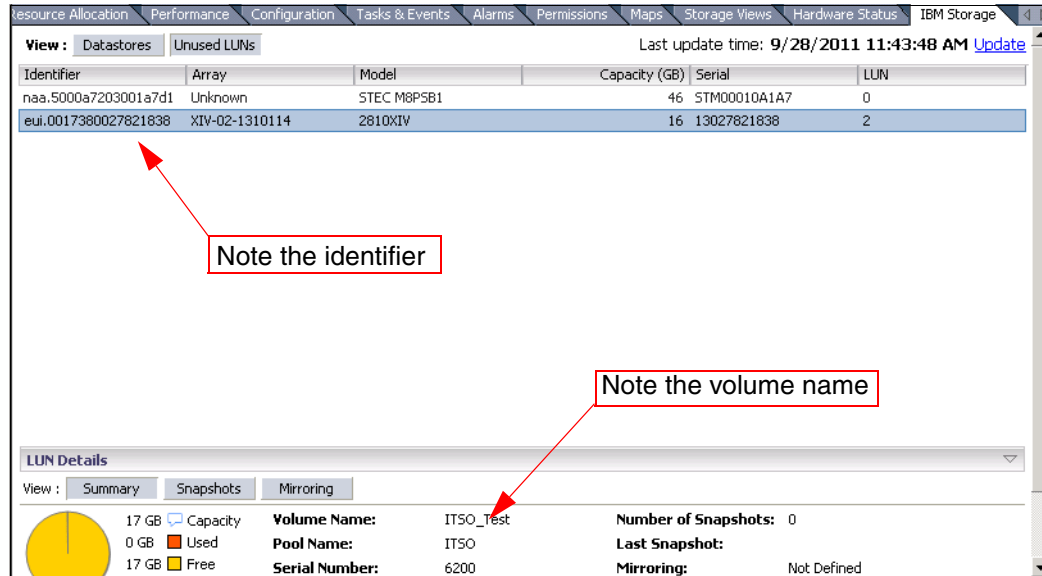


Figure 5-11 Unused LUNs in the IBM Storage tab

3. Select **Add Storage** from Configuration tab.
4. You are prompted to select a LUN, as shown in Figure 5-12. Use the identifier to ensure that you select the correct LUN. If the Identifier column is not displayed, right-click the heading area and add that column.

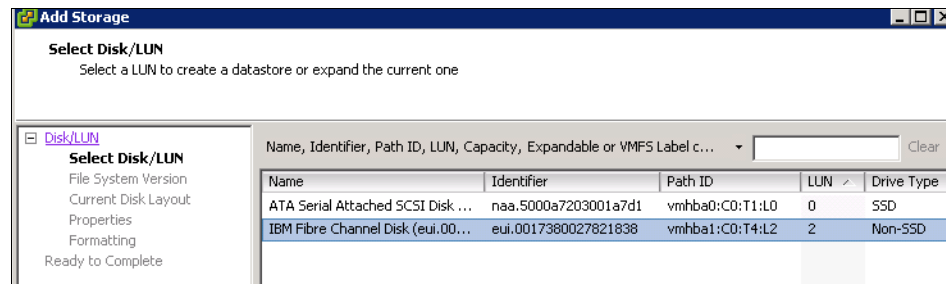


Figure 5-12 Locating the matching datastore

5. When you are prompted to enter a datastore name in the Properties tab, use the same name you used when creating the volume.

5.1.6 Using a read-only user

If your organizational structure does not allow the VMware administrator to make storage administration decisions, you can still use the plug-in with read-only access. To do so, create a user on the XIV that is in the Read Only category. When adding the XIV to the vCenter plug-in as shown in Figure 5-7 on page 81, use this restricted Read Only user.

The plug-in confirms that the permission level is Read Only, as shown in Figure 5-13

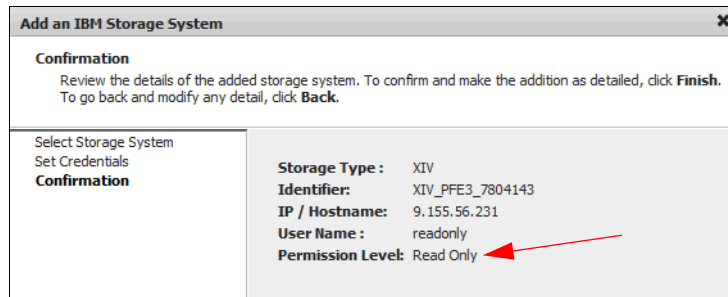


Figure 5-13 Read Only permission level

You are not prompted to select pools, as shown in Figure 5-8 on page 81 because the user has no authority to work with pools. However you still get to view the IBM Storage tab, as shown in Figure 5-11 on page 84. The advantage is that the VMware administrator can now be sure which hardware matches which datastore. This system allows you to identify the following data without any ability to change or configure the XIV:

- ▶ Exact XIV name
- ▶ XIV serial number
- ▶ XIV pool name
- ▶ XIV volume name
- ▶ XIV snapshot name

5.1.7 Locating the user guide and release notes

The IBM Storage Management Console for VMware vCenter includes a user guide and release notes that are available for download from:

[http://www.ibm.com/support/fixcentral/swg/selectFixes?parent=ibm/Storage_Disk&product=ibm/Storage_Disk/XIV+Storage+System+\(2810,+2812\)&release=All&platform=All&function](http://www.ibm.com/support/fixcentral/swg/selectFixes?parent=ibm/Storage_Disk&product=ibm/Storage_Disk/XIV+Storage+System+(2810,+2812)&release=All&platform=All&function)

5.1.8 Troubleshooting

This section discusses the two issues that were seen during testing for the book.

Plug-in disabled

If you do not see the IBM Storage icon, you might find the plug-in is disabled. The plug-in on a remote vSphere Client can be disabled if the client does not resolve the host name of the vCenter server. Click **Plug-ins** → **Manage Plug-ins**. If `ibm-vcplugin` is disabled, confirm if the issue is that the remote name cannot be resolved, as shown in Figure 5-14 on page 86.

As a simple test, ping the host name listed in the error. If the remote name is the problem, correct the issue with name resolution. One simple solution is to add an entry to the HOSTS file of the server on which you are trying to run the vSphere Client. This issue did not occur if

the vSphere Client was run locally on the vSphere vCenter server because the client was local to the server. After you ping the host name from the remote client, you can enable the plug-in, as shown in Figure 5-14.

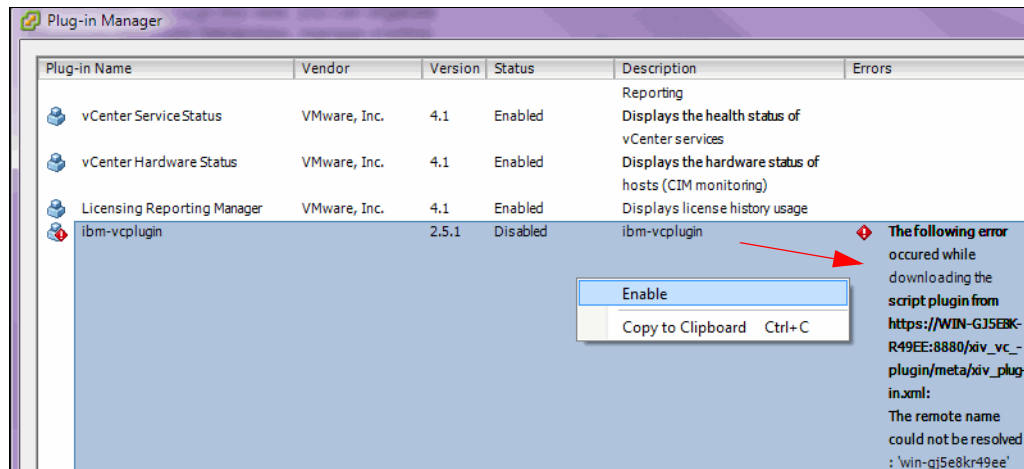


Figure 5-14 vCenter plug-in error

No detailed information is available

If you open the IBM Storage tab and highlight a datastore or a LUN, you might see the message No detailed information is available for this storage device. This message is shown in Figure 5-15. This error occurs because the LUN in question is being provided by a device that the IBM Storage plug-in cannot manage. It can also occur if the IBM Storage device in question is not added to the plug-in. If the device is not added to the plug-in, the plug-in does not have the logon credentials necessary to confirm device details.

To correct this error, add the relevant device using the process documented in 5.1.3, “Adding IBM Storage to the plug-in” on page 80. Figure 5-15 shows that the undefined system is an XIV as indicated in the Model column. The hint as to which XIV is given in the identifier column where we can see the identifier is: eui.00173800279502fb. This number derives from the WWNN. The WWNN, in this case, is 50:01:73:80:27:95:00:00 (note that 002795 is the unique portion of the WWNN).

You can also determine the XIV serial by using the identifier. The first part of the identifier is 001738, which is the IEEE Object ID for IBM. The next part is 002795, which is the serial number of the XIV in hexadecimal. If you convert that number from hex to decimal, you get the serial number of the XIV. In this example, it is 10133.

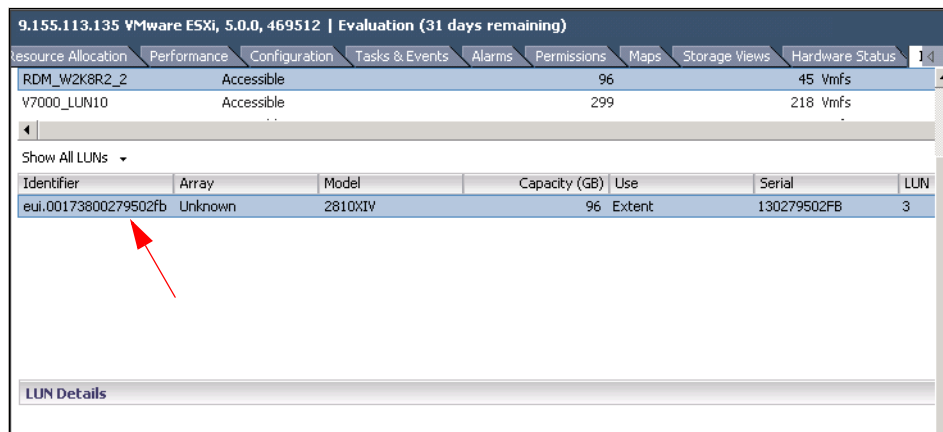


Figure 5-15 No detailed information is available



Data Protection in vSphere environments with XIV

This chapter describes how to implement data protection for VMware vSphere environments. The solution presented uses the vStorage APIs for Data Protection (VADP) along with the Tivoli Storage FlashCopy Manager for VMware to take advantage of the robust, versatile snapshot capabilities of the XIV Storage System.

6.1 Data Protection in vSphere environments with XIV

To provide motivation for the potentially crucial business value realized by the end-to-end integration of data protection technologies from IBM and VMware, first consider the typical business-driven design criteria and associated terminology inherent to data protection solutions.

6.1.1 Data Protection solution concepts and terminology

Data backup refers to the process of making one or more copies of data that can be used to restore a functionally viable version of the original data after a data loss or data corruption event. Simply capturing a data backup by no means guarantees that a successful data restore will be possible as defined by specific business needs. All data that is co-dependent must have point-in-time consistency across all layers of the stack, beginning with the application layer. Therefore, the concepts of point-in-time consistency and data interdependency are crucial when considering both the scope of the data backup and the mechanisms employed by integrated backup technologies, from the application layer through the storage subsystem layer.

There are two important objectives that must be carefully defined when devising a business-viable data protection scheme:

1. The *Recovery Point Objective (RPO)* represents the most recent point-in-time at which consistent inter-dependent data can be recovered. Specifying higher RPOs indicates that more data loss is acceptable.
2. The *Recovery Time Objective (RTO)* represents the amount of time elapsed between the data loss or corruption event and the full restoration of business functions. Obviously, the ability to achieve an RTO depends on much more than just the availability of a viable backup.

As a result of the necessity of preserving data consistency, there are two methodologies for creating backups:

1. During a *Cold Backup (offline backup)*, the applications are closed and locked to ensure that there are no further changes to the data occur until the backup is complete.
2. During a *Hot Backup (active backup)*, the application remains online and in use.

Because cold backups are obviously impractical for enterprise applications, another mechanism is necessary to capture point-in-time data. While the implementation and scope of operations fundamentally differ, in terms of basic functionality, both vSphere and XIV utilize a form of snapshot technology to preserve the state of consistent data at a point-in-time transparently and non-disruptively to applications. Snapshots initiated by VMware occur at the virtual machine level and comprise all state-specific information necessary to restore a given virtual machine, potentially including even the virtual machine's memory state. In contrast, snapshots initiated within the XIV Storage System maintain data consistency with a scope spanning one or more logical volumes, and include the following types:

- ▶ *Local Snapshots* are snapshots created within the storage pool where the original volume exists.
- ▶ *Consistency Group Snapshots* are snapshots created at the same point-in-time for all volumes in the consistency group, enabling the scope of consistent data recovery to span multiple volumes as required by business-critical applications and processes.

- ▶ *Mirrored Snapshots* are snapshots created at the same point-in-time for one or more volumes participating in a remote mirroring relationship, and at one or both sites (primary site and disaster recovery site).

For additional insight into the topic of snapshot technology and its role in the vSphere environment, refer to 2.1.5, “XIV and VMware thin provisioning” on page 18, and 3.1, “VAAI overview” on page 28.

6.1.2 vSphere Data Protection solution components

Large-scale VMware vSphere environments with hundreds or thousands of virtual machines are at high risk without a seamless backup and recovery implementation that complies with organizational RTO and RPO service level agreements. IBM offers a comprehensive set of products with robust features that enable customers to effectively protect and recover data in vSphere environments while simultaneously managing retention and controlling costs.

vStorage APIs for Data Protection overview

Fundamentally, *vStorage APIs for Data Protection* (VADP) serve the purpose of providing a standard framework to enable centralized, efficient, off-host, LAN-free backup of vSphere virtual machines using compatible storage products, like Tivoli Storage FlashCopy Manager for VMware leveraging the robust, versatile snapshot capabilities of the XIV Storage System.

Introduced in vSphere 4.0, VADP is the next generation of VMware’s data protection framework. Backup applications supporting VADP integration can backup vSphere virtual machines from a central backup server or virtual machine without requiring backup agents or incurring backup processing overhead on the ESX/ESXi hosts’s virtual machine guests. This is possible because the VADP framework exploits compatible backup software in centralizing the actual backup processing and data movement activity within the VADP-compatible storage subsystem(s).

Since VADP leverages backups incorporating the snapshot capabilities of VMFS, file-level backups can be performed transparently and non-disruptively at any time of the day without requiring extended backup windows or any interruption to applications and users associated with backup windows. The benefits of VMFS snapshot technology are augmented by leveraging VADP to cohesively integrate Tivoli Storage FlashCopy Manager and XIV snapshot technology, and include:

- ▶ The flexibility to perform to full, differential, and incremental image backups and restores of virtual machines.
- ▶ The ability to capture file-level backups of virtual machines running supported Windows and Linux operating systems.
- ▶ Ensuring data consistency by using Microsoft Volume Shadow Copy Services (VSS) for virtual machines running supported Microsoft Windows operating systems.

Sophisticated XIV snapshots

The IBM XIV Storage System performs snapshots with virtually no overhead using redirect-on-write technology instead of the traditional latency-causing copy-on-write snapshot processing used in other storage subsystems. Supplementing this efficiency, XIV incremental backup snapshots are optimally configured (just like primary volumes) helping ensure enterprise-class read and write performance while eliminating the need for processing and space overhead. Refer to the IBM Redbooks publication *IBM XIV Storage System: Copy Services and Migration*, SG24-7759, for detailed information and comprehensive implementation guidance about the topic of XIV snapshots.

VMware snapshots best practices

The purpose of snapshot technology in VMware is to preserve the state information in metadata and the data content of a virtual machine at a specific point in time. For example, the state includes the virtual machine's power state:

- ▶ Powered-on
- ▶ Powered-off
- ▶ Suspended

The data includes all of the files that represent the virtualized hardware components of the virtual machine, including (among others):

- ▶ Hard disks
- ▶ Memory
- ▶ Virtual network interface cards.

A virtual machine provides several operations for creating and managing both snapshots and entire snapshot chains, which consist of a linked chain of dependent snapshots. These operations allow administrators to create snapshots, revert the state information and data files of the virtual machine to those captured by a given snapshot in the chain, and remove snapshots. Because of the incremental, linked nature of VMware snapshots, extensive snapshot trees consisting of multiple branch points can exist without unnecessary duplication of data.

It is important to keep in mind that VMware Snapshots are not backups, although backup software that integrates with VADP exploits VMware snapshots in conjunction with XIV snapshot to preserve VMFS, VM, and file-level awareness for usage in backup and restore operations. The white paper entitled *IBM XIV Storage Systems - Snapshots Reinvented* is an excellent source of information about the benefits of XIV's powerful and extremely versatile snapshot technology, and is at the following web site:

ftp://public.dhe.ibm.com/storage/disk/xiv/ibm_xiv_snapshots_paper.pdf

The following policies are appropriate when implementing VMware snapshots on the XIV Storage System:

- ▶ Consistently monitor snapshot status using VMware snapshot alarm options (dependent on vSphere version):
 - Configure vCenter Snapshot alarms – Knowledge Base Article1018029:
http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1018029
- ▶ Limit the total number of snapshots in a chain to 2-3 to prevent performance degradation at the VMware level:
 - This is not specific to the storage subsystem.
- ▶ Delete all snapshots prior to Virtual Machine disk changes.
- ▶ If any uncertainty of snapshot state exists, confirm status using the CLI.
- ▶ As a general rule, retain VMware snapshots for no more than 24-72 hours:
 - VADP and Tivoli Storage FlashCopy Manager for VMware to take advantage of temporary VMware snapshots to create XIV snapshots for longer point-in-time data retention.

The ESX improvement history:

- ▶ ESX(i) 4.0 U2: Snapshot deletion takes up less space on disk

- ▶ ESXi 5.0: New functionality to monitor snapshots and provide warning if snapshots need consolidation

Consult these VMware knowledge base articles for an in depth description of VMware snapshot technology concepts and best practices:

VMware Snapshot Best Practices – Knowledge Base Article 1025279

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1025279

Understanding Virtual Machine Snapshots in ESX – Knowledge Base Article 1015180

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1015180

6.1.3 Tivoli Storage FlashCopy Manager for VMware

Traditional file-level backups are no longer sufficient when the amount of data to protect is large, and the window to perform backups is short. Scheduling backup operations outside of business hours is generally the solution, but many organizations require restore points created throughout the day too. As a result, administrators want a solution that can protect critical data in a way that minimizes downtime associated with a backup operation, and also increases the frequency with which backups can be created.

Tivoli data protection software leverages vSphere VADP to orchestrate virtual machine snapshots with XIV logical volume snapshots. IBM offers two Tivoli data protection software products that work cohesively with XIV and vSphere:

- ▶ *IBM Tivoli Storage FlashCopy Manager for VMware (FCM)* integrates hardware snapshot management and synchronizes the backup process with vSphere, reducing backup times and promoting backup consistency and recoverability.
- ▶ *IBM Tivoli Storage Manager for Virtual Environments (TSM for VE)* performs block-level incremental backups of VMware guests using VADP. The IBM TSM Data Protection for VMware Recovery Agent mounts snapshots to enable file-level and instant volume restores.

IBM Tivoli Storage FlashCopy Manager for VMware V3.2 is designed to deliver high levels of data protection for business-critical applications using integrated application snapshot backup and restore capabilities. These capabilities are achieved through the utilization of advanced XIV redirect-on-write snapshot technology to help create a high performance, low impact application data protection solution. It is easy to install, configure, and deploy, and deeply integrates with XIV's sophisticated high-performance snapshot technology.

Tivoli Storage FlashCopy Manager for VMware provides the ability to create and manage volume-level snapshots of VMware VMFS volumes, providing a backup of the virtual machines residing on the volume. The snapshots are created while virtual machines are running and by leveraging the VMware snapshot capability with no downtime. In addition, FCM integrates with XIV snapshots to promote simple, efficient recovery.

FCM orchestrates VMware-based snapshots with XIV snapshot of LUN(s) backing the associated VMFS datastore in vSphere, as follows:

- ▶ FCM initiates a VMware software snapshot of the VMs residing on the datastore through the vStorage APIs for Data Protection (VADP). VMware snapshots can trigger application quiescence for individual virtual machines, helping ensure application backup consistency.

- ▶ FCM determines the XIV LUNs that are associated with the datastore.
- ▶ FCM invokes an XIV snapshot, creating a persistent copy of the virtual disks and associated VMware snapshots.
- ▶ In order to adhere to performance-oriented best practices governing deletion of VMware snapshots, FCM creates persistent XIV snapshots for use as source for recovery operations, while VMware snapshots are deleted.
- ▶ Until specified for deletion, the XIV snapshot is retained for restore purposes.
- ▶ FCM optionally creates an additional copy of data on Tivoli Storage Manager server:
 - Enables individual file recovery using IBM TSM Data Protection for VMware Recovery Agent mount.

Tivoli Storage FlashCopy Manager for VMware limitations

It is important to become familiar with the known integration limitations prior to deploying Tivoli Storage FlashCopy Manager for VMware in specific vSphere environments. The following resources contain the most recent insights and restrictions that must be incorporated into a successful data protection plan structured upon storage snapshot technology:

- ▶ Tivoli Storage FlashCopy Manager for VMware Version 3.1:
<http://www-01.ibm.com/support/docview.wss?uid=swg21567738>
- ▶ Tivoli Storage FlashCopy Manager for VMware Version 3.2:
<http://www-01.ibm.com/support/docview.wss?uid=swg21612307>

Single-Pass backup operations

FCM helps protect entire vSphere environments by leveraging VADP to create snapshots of LUN-level VMFS datastores with a single backup server. These snapshots enable restore operations to be performed for individual virtual machine images, virtual volumes owned by a virtual machine, or individual files from a particular VM. The topic of restore operations is explored further in 6.1.4, “Comprehensive data protection for vSphere” on page 93.

With FCM and XIV, vSphere data backups can be scheduled or performed ad hoc with the following attributes and benefits:

- ▶ Application-transparent backups occur off-host (proxy based) harnessing nearly instantaneous redirect-on-write XIV snapshots.
- ▶ XIV snapshot technology supports a combined total of up to 8,000 volumes and persistent snapshots concurrently:
 - Persistent snapshots are retained locally on the XIV system.
 - Nearly instantaneous restores are possible from snapshots
- ▶ Backup operations are performed with VMFS datastore granularity, with scheduled or ad-hoc execution.
- ▶ Tivoli VM backup software covers all major operating systems, so there is no requirement to deploy OS-specific agents to each VM.
- ▶ Optionally, Tivoli Storage Manager for Virtual Environments can backup the point-in-time data to a storage pool located within the appropriate lower-tiered storage repository for general backup, archival, and mandated data retention purposes. TSM backups have the following attributes, among others:
 - Performs multiple, parallel VM-granular backups simultaneously to minimize backup windows.

- Transfers data outboard of application servers to minimize impact to application.
- Preserves copies on the TSM server to provide long-term retention and disaster recovery.
- Supports block-level, progressive incremental backups with VMware Changed Block Tracking.
- Facilitates application-consistent backups of Microsoft SQL Server and Exchange hosted on VMware.

Figure 6-1 illustrates the FlashCopy Manager and Tivoli Storage Manager for Virtual Environment components and their interactions.

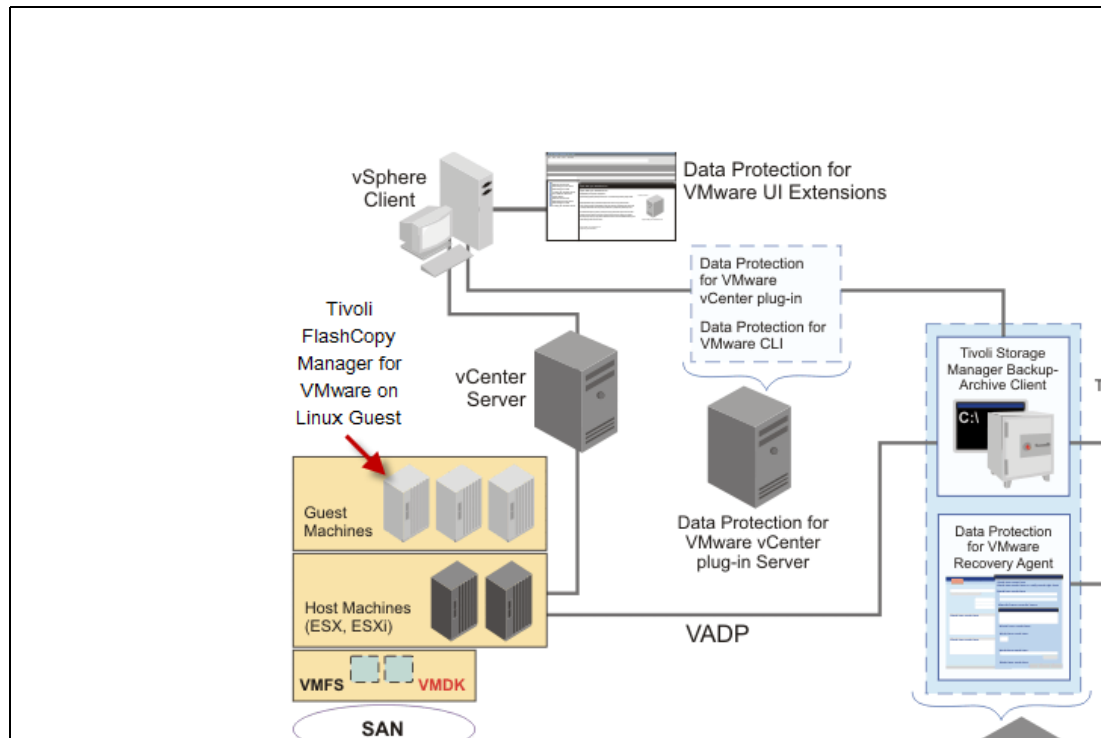


Figure 6-1 FCM and TSM Components - Comprehensive Data Protection for vSphere Environments

6.1.4 Comprehensive data protection for vSphere

Customers can choose multiple XIV snapshot recovery options. Individual VMs can be restored from an XIV hardware snapshot of VMFS datastores or from an offloaded TSM backup. The recovery can be performed to either the original or an alternate VMFS datastore, under the original or a new name. The Tivoli vCenter plug-in presents a unified view of all available FCM-based and TSM-based backups, including both XIV snapshots and TSM storage pool backups.

For pinpoint recovery, FCM provides the flexibility to selectively restore one or more individual virtual disks, without the need to restore an entire virtual machine. Individual files can be restored by attaching one or more virtual disks from a datastore snapshot to a guest VM.

The TSM for Virtual Environments recovery agent accesses the storage pool data directly in TSM, offering similar robust recovery options. TSM VE recovery agent applies TSM storage pool access rules, enabling authorized recovery initiation by help desk personnel, backup administrators, or application owners.

With XIV storage, recovery operations have the following properties and benefits:

- ▶ Flexible and granular recovery options allow administrators to mount the required scope of point-in-time data to guest VMs:
 - Individual files or folders
 - Individual VMDK or groups of VMDKs (also referred to as virtual disks)
 - Full virtual machine (or machines), even if different from the original VM.
- ▶ Reliable, high performance, nearly instantaneous restore capability using XIV snapshots of VMFS datastores:
 - Efficient XIV snapshot technology gives administrators the flexibility to create frequent restore points as determined by business requirements to minimize data-loss exposure.
- ▶ XIV snapshots created by FCM can be progressively and incrementally preserved on external, long-term retention media using TSM for VE, enhancing the viability of creating frequent restore points and the likelihood of exceeding RPO and RTO objectives during hardware outages.

FlashCopy Manager Data Protection for VMware vCenter plug-in

The IBM Tivoli Storage FlashCopy Manager Data Protection for VMware vCenter plug-in provides scheduled datastore-level XIV hardware snapshots. XIV storage offers datastore-level snapshots in conjunction with Tivoli Storage FlashCopy Manager. The web-based FCM and TSM plug-in expands the familiar vSphere client to include a wide range of management processes, from incremental backups of individual VMs for long-term TSM storage pool retention to leveraging space-efficient XIV snapshots for high-speed full datastore backups. Snapshot retention policies can be defined separately for datastore snapshots managed by FCM and for VM backups sent to TSM, and incorporate automatic reuse of local snapshot storage as older snapshot versions expire.

The FCM and TSM plug-in expands the vSphere client to drive quick and easy monitoring of in-progress backup and restore activities and historic reporting of past backup and recovery operations.

The rich TSM monitoring and reporting functionality features summary views with detailed drill-down analysis, important backup and restore statistics, and managed-capacity reporting to facilitate effective data protection strategies and promote timely, accurate restores.

Command Line interface

The IBM Tivoli Data Protection for VMware command line interface (CLI) offers a powerful common front end for Tivoli Storage FlashCopy Manager for VMware and Tivoli Storage Manager for Virtual Environments. The CLI can correlate backups created by FCM and TSM, combining multiple backup runs into one logical backup. It also offers a simple backup scheduler to configure recurring backup tasks. The CLI features robust custom scripting and specialized external scheduling capabilities.

In summary, by integrating with the vStorage APIs for Data Protection (VADP) and leveraging advanced XIV storage snapshot technology, Tivoli Storage FlashCopy Manager (FCM) provides powerful data protection with excellent performance and simplified management for VMware environments. FCM-captured XIV snapshots can be retained on the XIV system for nearly instantaneous, reliable restores. Customers can also exploit the powerful data protection and data reduction capabilities available through TSM-FCM integration with Tivoli Storage Manager for the Virtual Environment (TSM for VE). The wide range of integrated functionality with VMware includes content-aware backup, changed block tracking (CBT), data deduplication, progressive incremental backup, LAN-free backup, hierarchical storage management, and centrally managed policy-based administration.

FCM for VMware Basic installation steps

The following fundamental steps are required to install FCM for VMware leveraging XIV snapshot capabilities:

1. Install XIV CLI.
2. Install FCM for VMware 3.2.
3. Configure FCM for VMware.

Reference the publication at the following web site for detailed guidance, including the installation and configuration of FCM on a Linux server and the FCM for VMware plug-in on the vCenter server:

http://pic.dhe.ibm.com/infocenter/tsminfo/v6r4/index.jsp?topic=%2Fcom.ibm.itsm.fcm.doc%2Ft_protect_fcm.html

Additional practical guidance spanning the installation and configuration of all necessary solution components are in the white paper located at:

[http://www-03.ibm.com/support/techdocs/atsmastr.nsf/5cb5ed706d254a8186256c71006d2e0a/e80fb1aadbe9b3b0852579ca00554356/\\$FILE/SAP%20with%20IBM%20FlashCopy%20Manager%20for%20VMware%20for%20XIV%20and%20Stowize%20V7000.pdf](http://www-03.ibm.com/support/techdocs/atsmastr.nsf/5cb5ed706d254a8186256c71006d2e0a/e80fb1aadbe9b3b0852579ca00554356/$FILE/SAP%20with%20IBM%20FlashCopy%20Manager%20for%20VMware%20for%20XIV%20and%20Stowize%20V7000.pdf)

Check the latest pre-installation checklist before installing FlashCopy Manager for VMware on the Linux server. The following link provides detailed information about the hardware and software requirements: Version 3.2 FlashCopy Manager for VMware:

<http://www-01.ibm.com/support/docview.wss?uid=swg21612309>

FCM for VMware implementation and usage with XIV

As a functional demonstration of FCM's role in providing data protection services for vSphere environments containing datastores residing on XIV storage, a step-by-step examination of these fundamental administrator data protection activities follows:

1. Defining a backup task.
2. Initiating the restore of a virtual machine.
3. Attaching a restored VMDK to a virtual machine.

Accessing FCM for VMware in the vSphere client and defining a backup task

To begin using FCM for VMware:

1. Navigate to the vSphere client Home panel, and click **Tivoli Data Protection for VMware** under the Solutions and Applications category, as shown in Figure 6-2 on page 96.

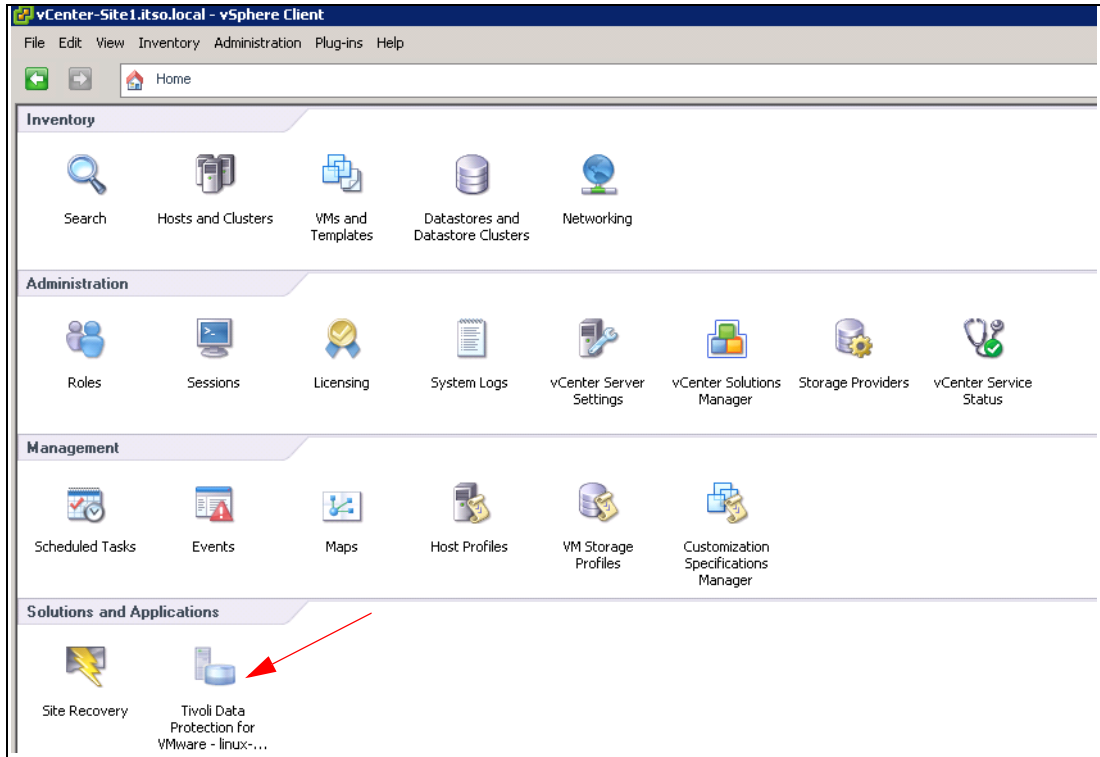


Figure 6-2 Tivoli Storage FlashCopy Manager for VMware: vSphere Client Solutions

2. In the tree view on the left side of the FCM GUI, a drop-down menu allows the hierarchical structure representing the vSphere infrastructure to be organized using either Hosts and Clusters or Data Stores viewpoints. Since the XIV snapshots invoked by FCM encompass the entire logical volume(s) backing the datastore(s), this example will reference the selection of the datastore-centric view. Select **Data Stores** from the drop-down menu, as illustrated in Figure 6-3.

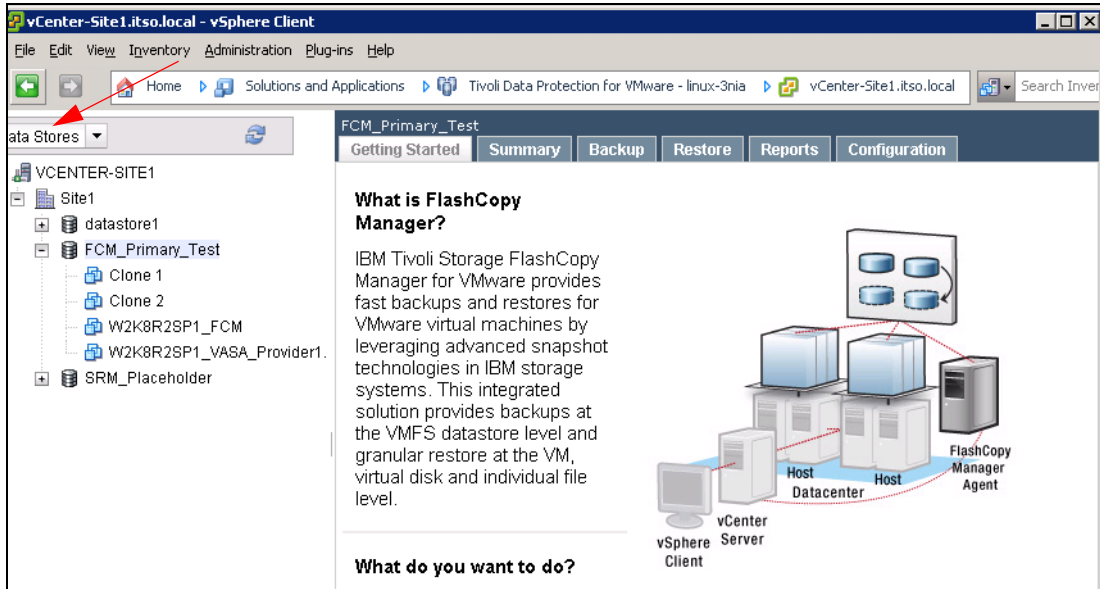


Figure 6-3 Tivoli Storage FlashCopy Manager for VMware: Getting Started

- Click the **Configuration** tab shown in Figure 6-4 to validate that the desired versions of FlashCopy Manager and the FCM Plug-in are implemented. The Plug-in Domain, consists of the set of VMware data centers managed by the plug-in through VADP, can also be verified.

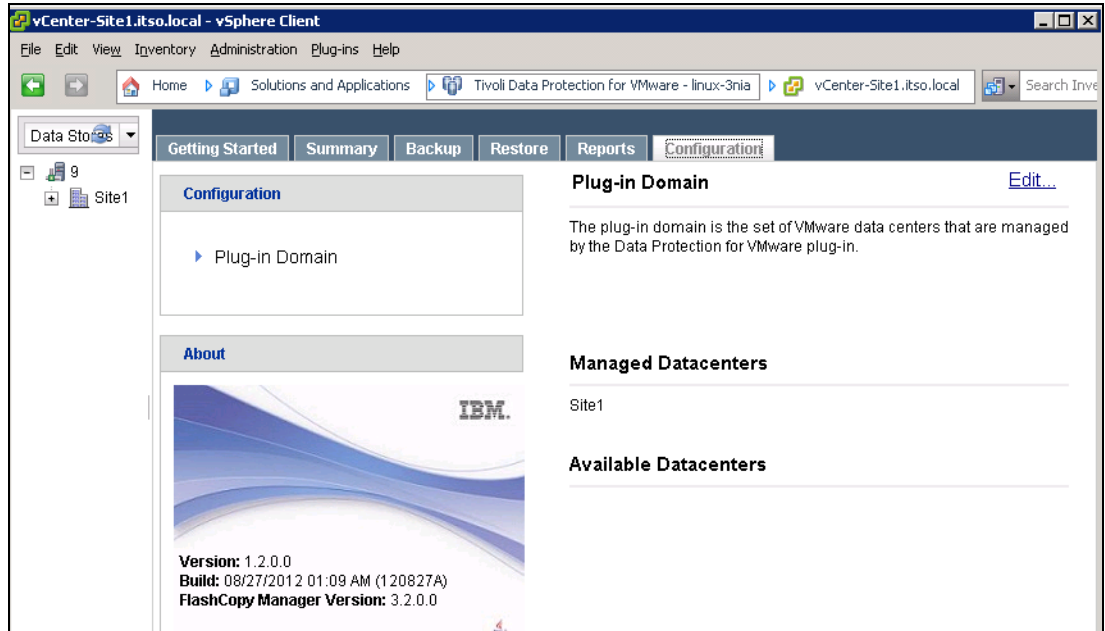


Figure 6-4 Tivoli Storage FlashCopy Manager for VMware: Plug-in Domain

- To begin the process of creating a backup task, select a datastore from the left tree view, and click either the **Getting Started** or **Summary** tabs and then click **Define a backup task**, as shown in Figure 6-5. It is not necessary to attempt to select multiple datastores, hosts, and so on, at this stage because multiple selections can be made as required within the backup task configuration wizard that follows.

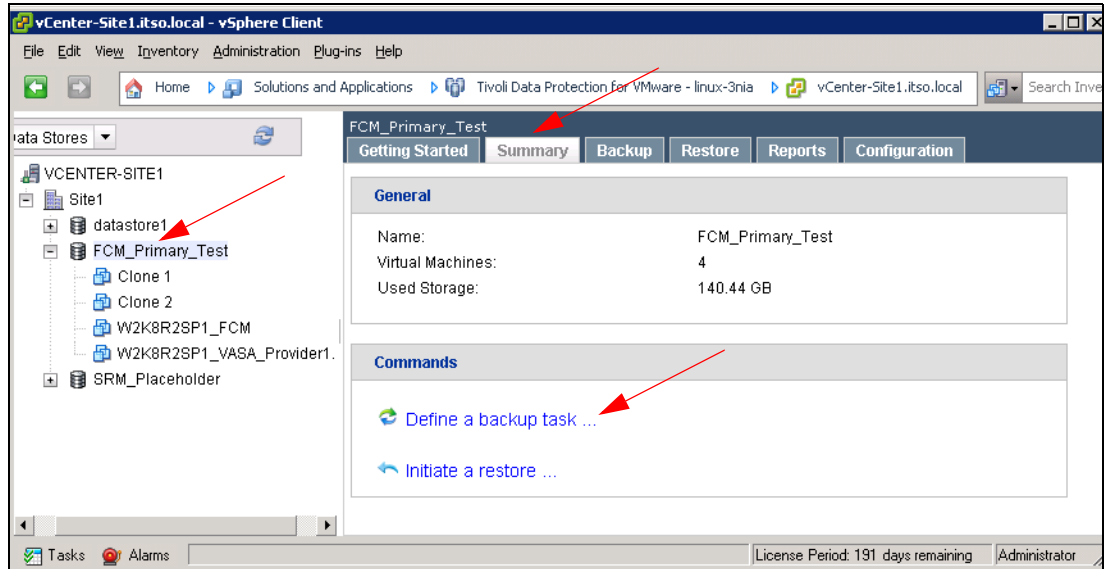


Figure 6-5 Tivoli Storage FlashCopy Manager for VMware: Datastore Summary

5. As shown in Figure 6-6, click **Next** to begin the process of configuring a backup task using the wizard.

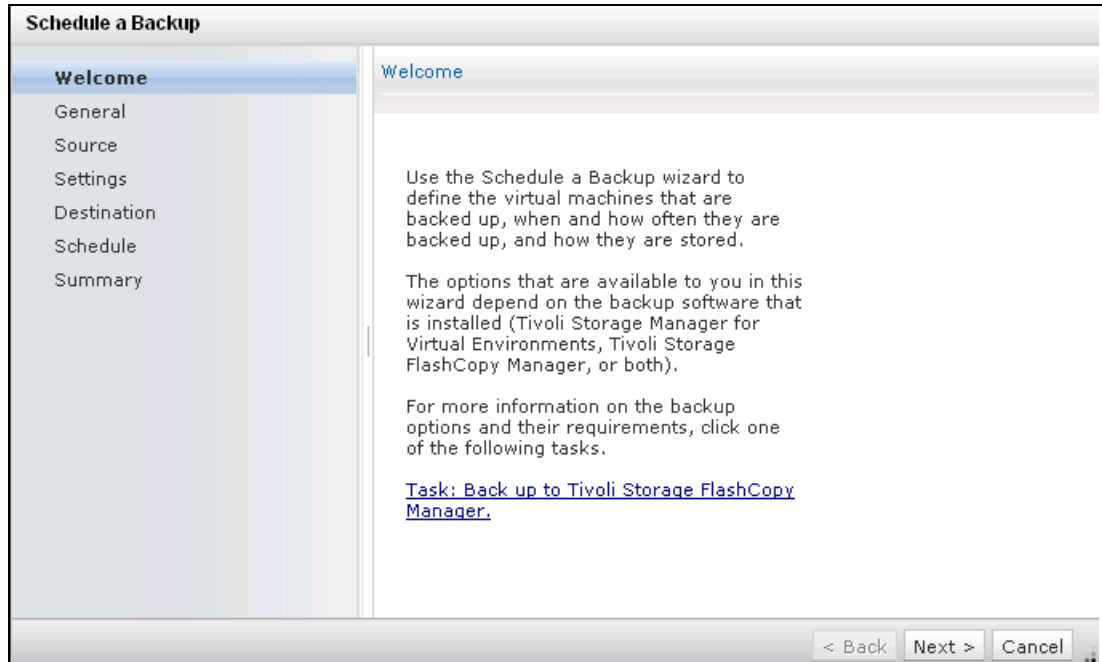


Figure 6-6 FCM Backup Process: Welcome

6. The General panel in the Properties window is used to enter a name and description for the backup process, as shown in Figure 6-7. Click **OK**.

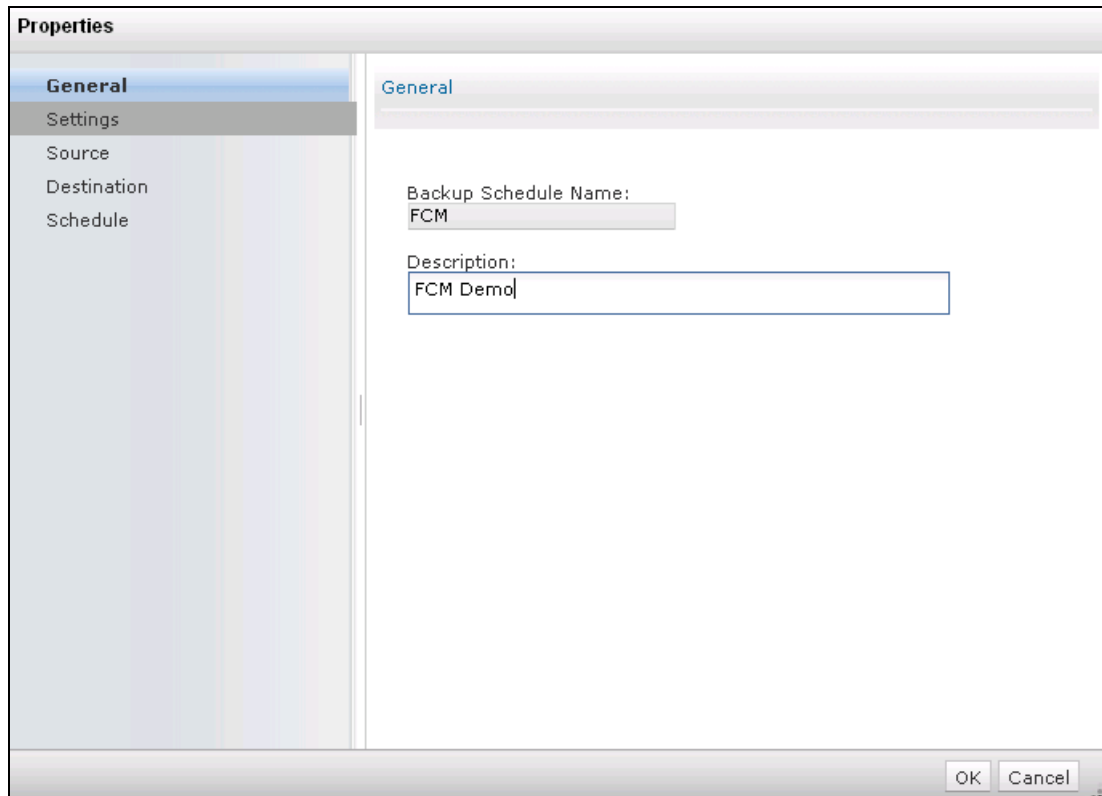


Figure 6-7 FCM Backup Process: Backup Name

7. Within the Settings panel illustrated in Figure 6-8, administrators can choose the behavior of the backup process with regard to virtual machine settings.

Review the white paper at the following site for a detailed description of these options:

[http://www-03.ibm.com/support/techdocs/atsmastr.nsf/5cb5ed706d254a8186256c71006d2e0a/e80fb1aadbe9b3b0852579ca00554356/\\$FILE/SAP%20with%20IBM%20FlashCopy%20Manager%20for%20VMware%20for%20XIV%20and%20Stowize%20V7000.pdf](http://www-03.ibm.com/support/techdocs/atsmastr.nsf/5cb5ed706d254a8186256c71006d2e0a/e80fb1aadbe9b3b0852579ca00554356/$FILE/SAP%20with%20IBM%20FlashCopy%20Manager%20for%20VMware%20for%20XIV%20and%20Stowize%20V7000.pdf)

Select the desired behavior, and click **OK** to proceed.

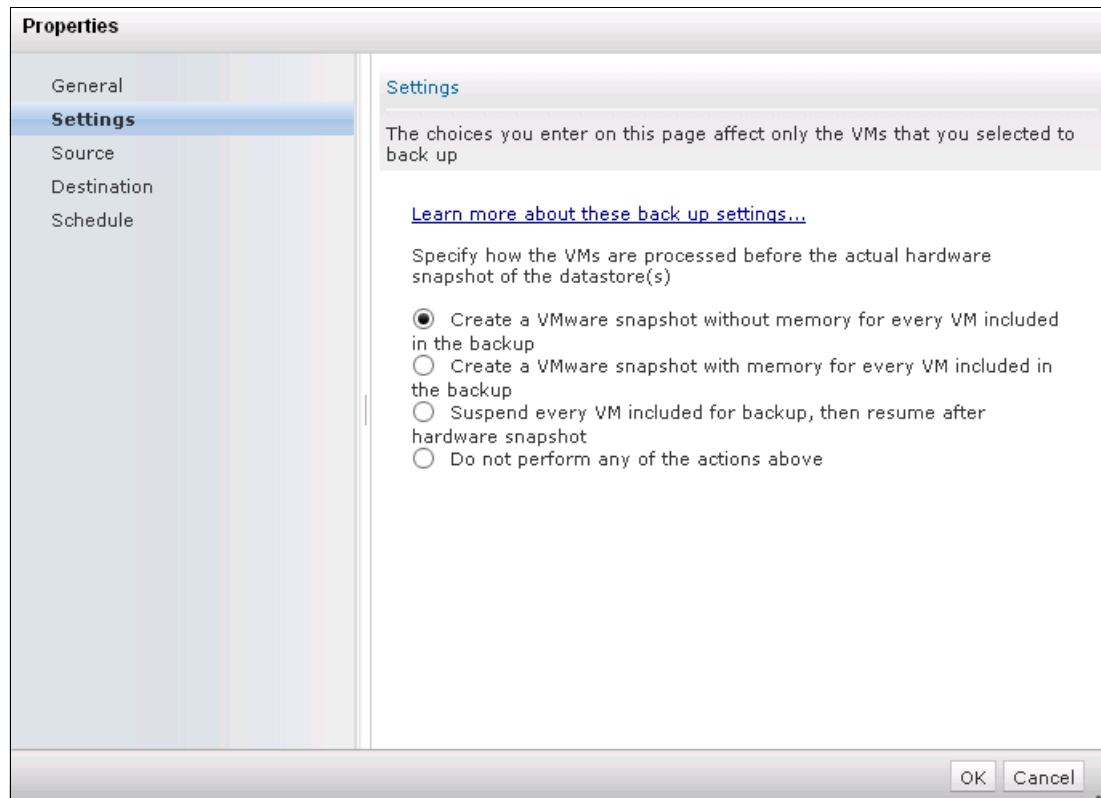


Figure 6-8 FCM Backup Process: VMware Snapshot Settings, Etc.

8. The Source panel facilitates the selection of one or more virtual machines and datastores, and provides flexibility in specifying the scope of the backup process to ensure that all dependent elements are included.

If a given virtual machine to be backed-up spans multiple datastores, all datastores are included. Figure 6-9 on page 100 illustrates the selection of a single virtual machine, and necessarily, the datastore associated with it.

After the desired selections are made, click **OK** to proceed.

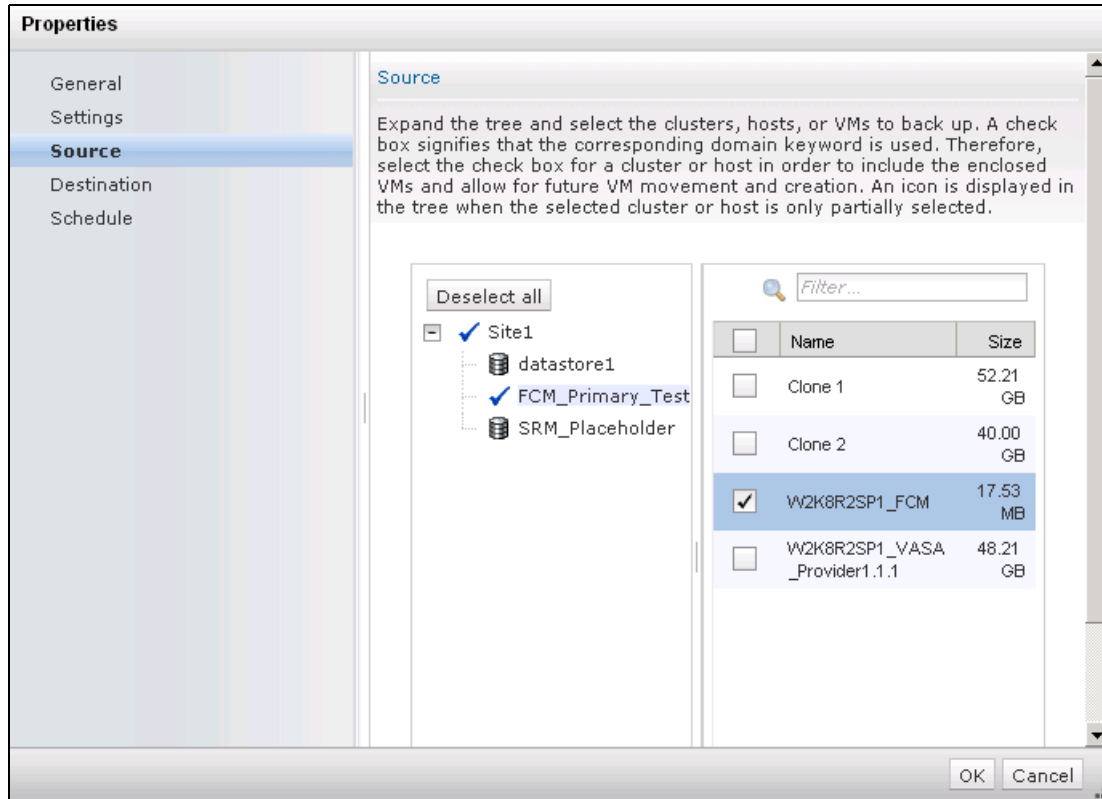


Figure 6-9 FCM Backup Process: Source Datastore(s) and Virtual Machine(s)

9. Ensure that XIV is selected from the drop-down list in the Destination panel, as demonstrated in Figure 6-10, and click **OK** to proceed. This defines the mechanism whereby VMware snapshots are synchronized with XIV snapshots through FCM orchestration.

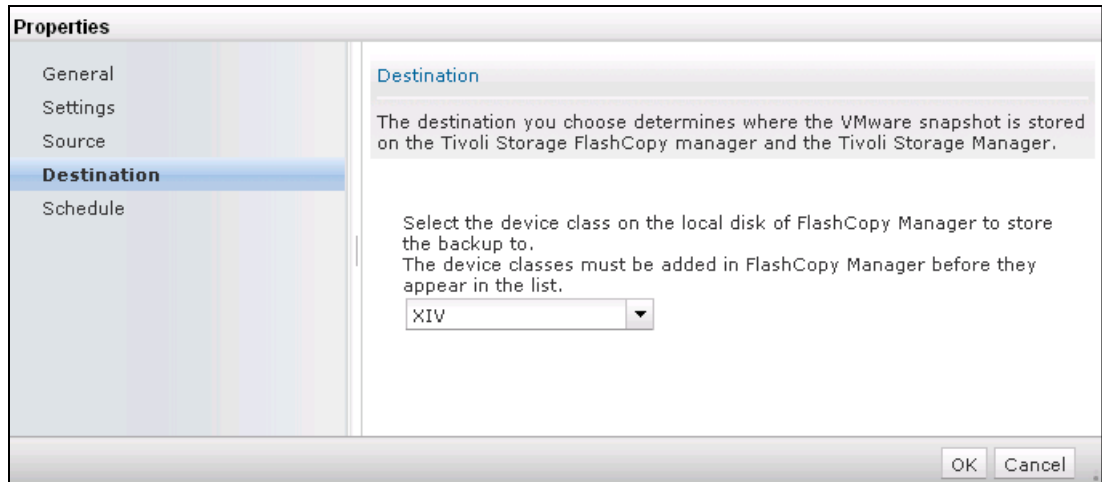


Figure 6-10 FCM Backup Process: FCM Storage Snapshot Destination

10. The Schedule panel displayed in Figure 6-11 on page 101 might contain the option to either run the backup process immediately or to create a schedule. Although the latter is the more common and practical approach to provide data protection for production environments, for purposes of this example and for performing a test, select **Run the backup now**, and click **OK**.

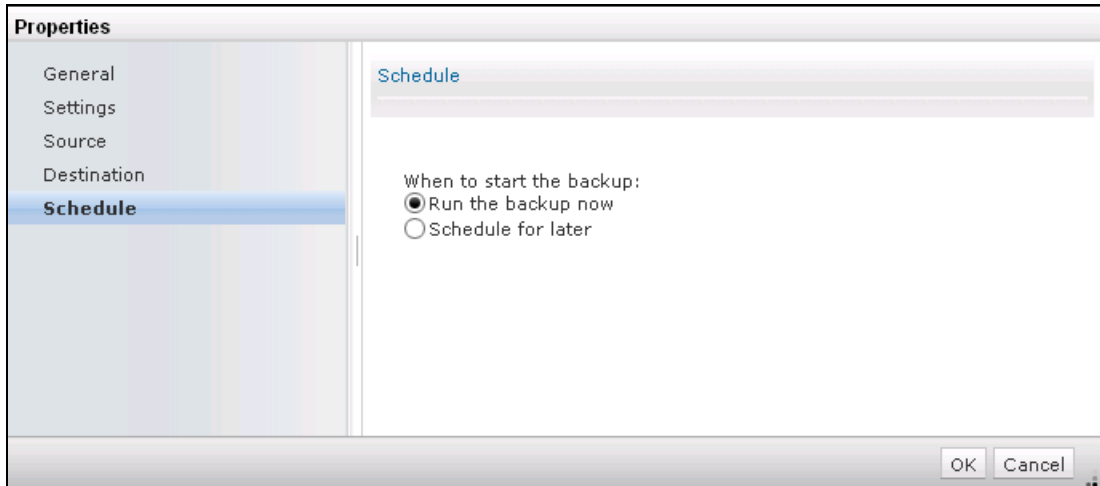


Figure 6-11 FCM Backup Process: Backup Schedule

11. As shown in Figure 6-12, click the **Backup** tab to view the status of the backup and verify successful completion status.

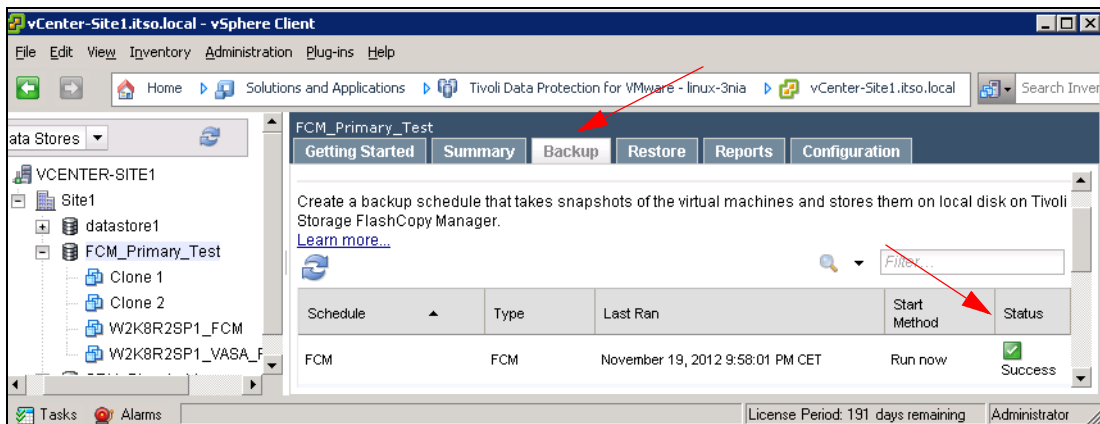


Figure 6-12 FCM Backup Process: Backup Job Status

12. Navigate to the Reports tab, and click **Backup Status** from the View options to display options for generating reports conveying virtual machine backup status and the capability of organizing virtual machines based on various attributes, including age of most recent backup, back-up status other than success, and so on.

Figure 6-13 on page 102 illustrates that the virtual machine specified in the backup job has been successfully backed-up, and provides the backup date and the duration of the job.

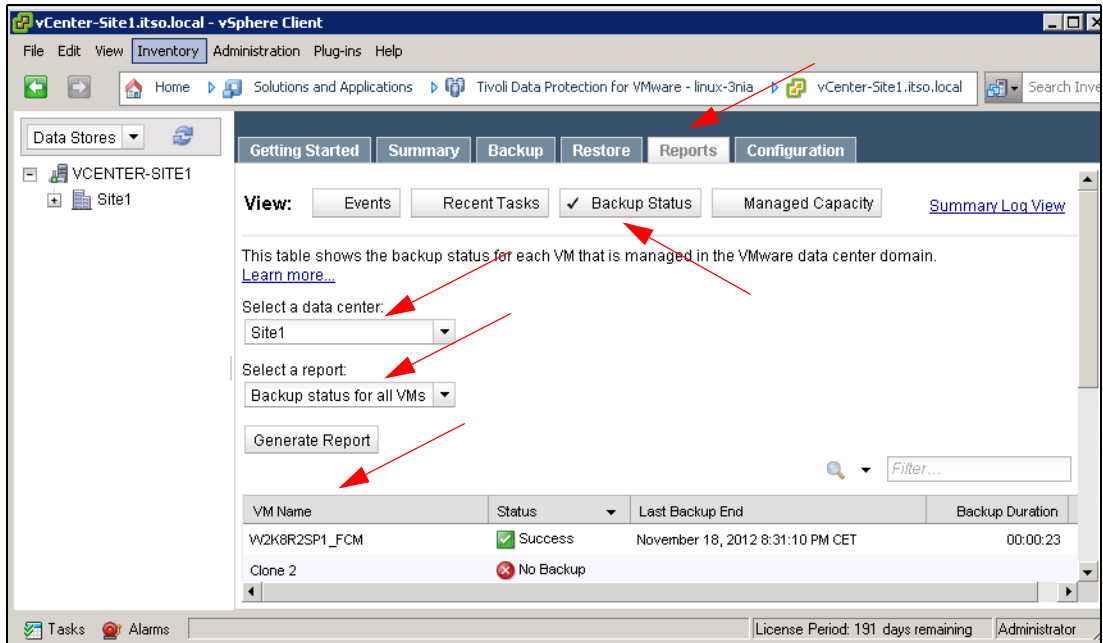


Figure 6-13 FCM Backup Process: Results Tab Backup Status by Virtual Machine

13. Open the XIV GUI and navigate to the Volumes by Pools interface to view the persistent XIV snapshots that exist for the logical volumes backing the datastores on which the virtual machines in scope for backup reside. Refer to Figure 6-14.

Note that the XIV snapshots are created, managed, and maintained by FCM, and require no administrator management labor.

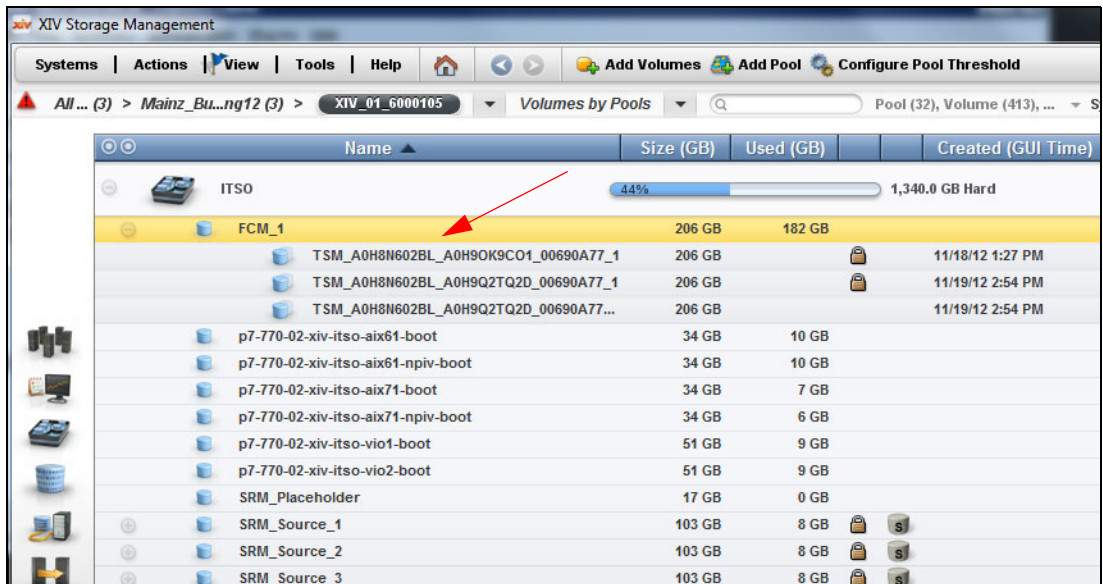


Figure 6-14 FCM Backup Process: XIV Snapshots Created and Managed by FCM

Note: It is important to plan backup operations according to the virtual machines' locations in the context of specific LUNs to minimize the number of snapshots required and thus save space since XIV snapshots are performed at the LUN level and all virtual machine disks that reside on a LUN will be part of that snapshot. However, virtual machines that were not selected in the FCM GUI will not be aware of the backup and as a result they cannot be restored using FCM with a guaranteed state of consistency.

FCM Virtual Machine restore

FCM restore points represent point-in-time VMware snapshots of virtual machine files backed by persistent XIV snapshots. They can be used to rollback a virtual disk to an earlier state, perhaps before a data corruption event occurred. The level of data consistency, for example representing application-consistent, file system-consistent, or crash-consistent classification, is determined by the type of VMware snapshot that was implemented during the creation of the restore point. The following example explores the steps necessary to initiate an FCM restore process:

1. Click the Restore tab, select the virtual machine in scope for recovery, select an available Restore Point with a time stamp prior to the disrupting event, and click the **Restore** hypertext, as illustrated in Figure 6-15.

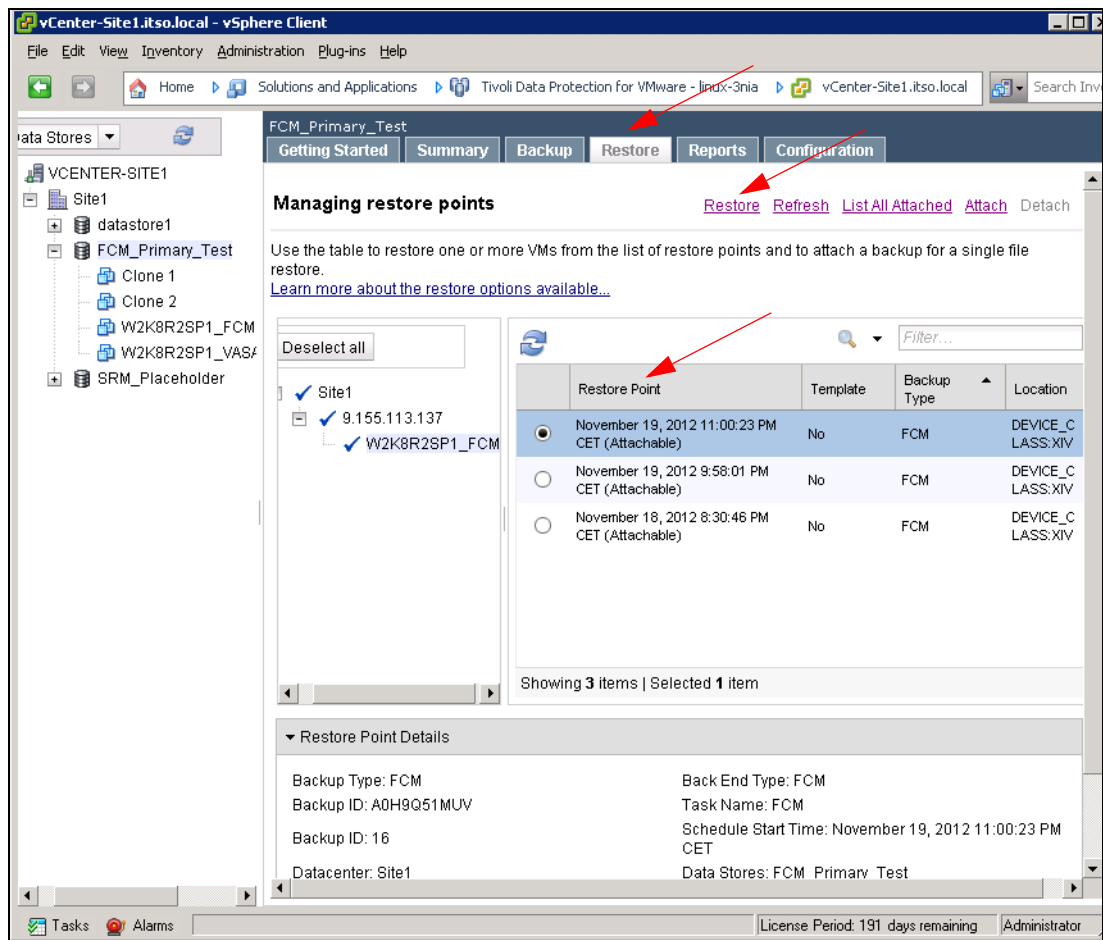


Figure 6-15 FCM Restore: Managing restore points

2. In the resulting wizard, review the Welcome panel shown in Figure 6-16, and click **Next**.

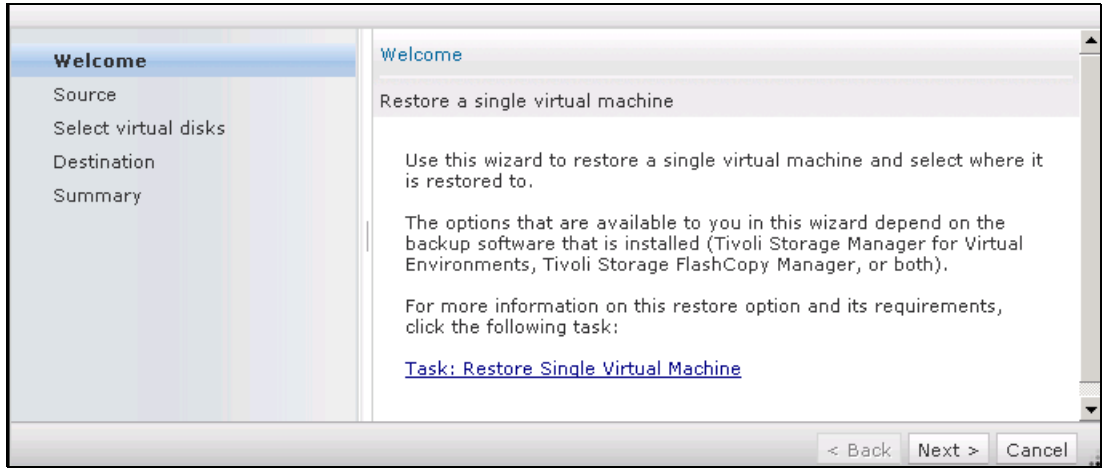


Figure 6-16 FCM Restore Process: Welcome

3. Within the Source panel, administrators can choose the scope of the restore process to be the virtual machine itself, or alternatively a subset of the virtual disks associated with the virtual machine, if applicable. Make the desired selection, as shown in Figure 6-17, and click **Next**.

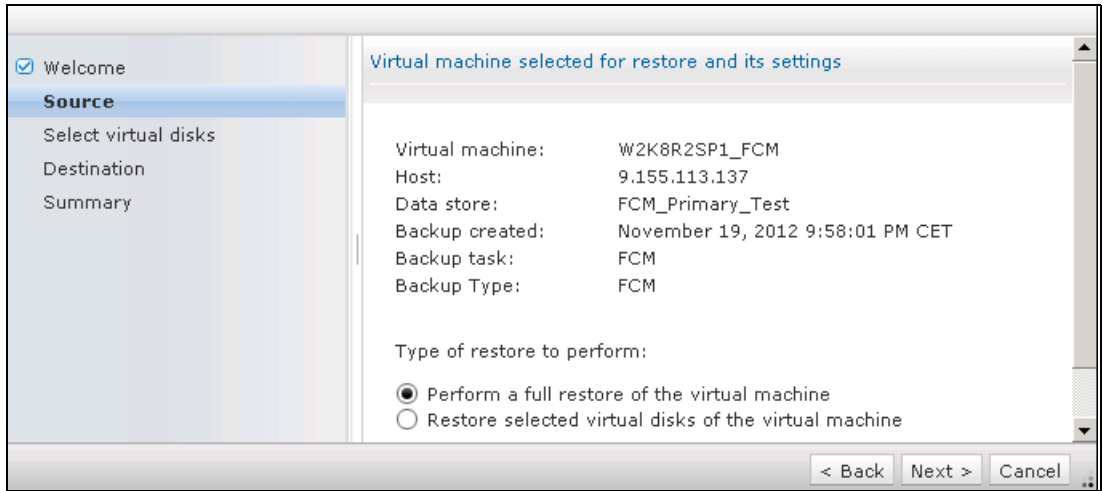


Figure 6-17 FCM Restore Process: Restore Type

- As shown in Figure 6-18, the Destination dialog grants the flexibility to restore virtual machines with associated virtual disks contained within a single datastore to be mapped to an alternate location. To follow this example, select **Restore to original location**, and click **Next**.

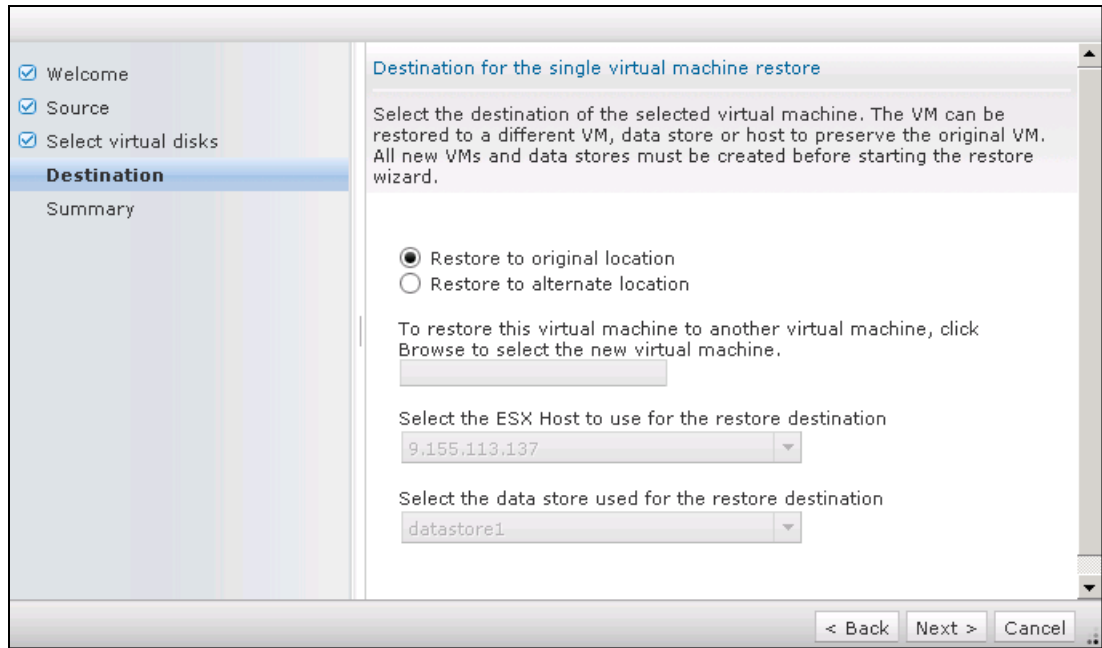


Figure 6-18 FCM Restore Process: Restore Destination

- Review the Summary panel shown in Figure 6-19 to validate that the desired restore source and destination are specified and then click **Finish**.

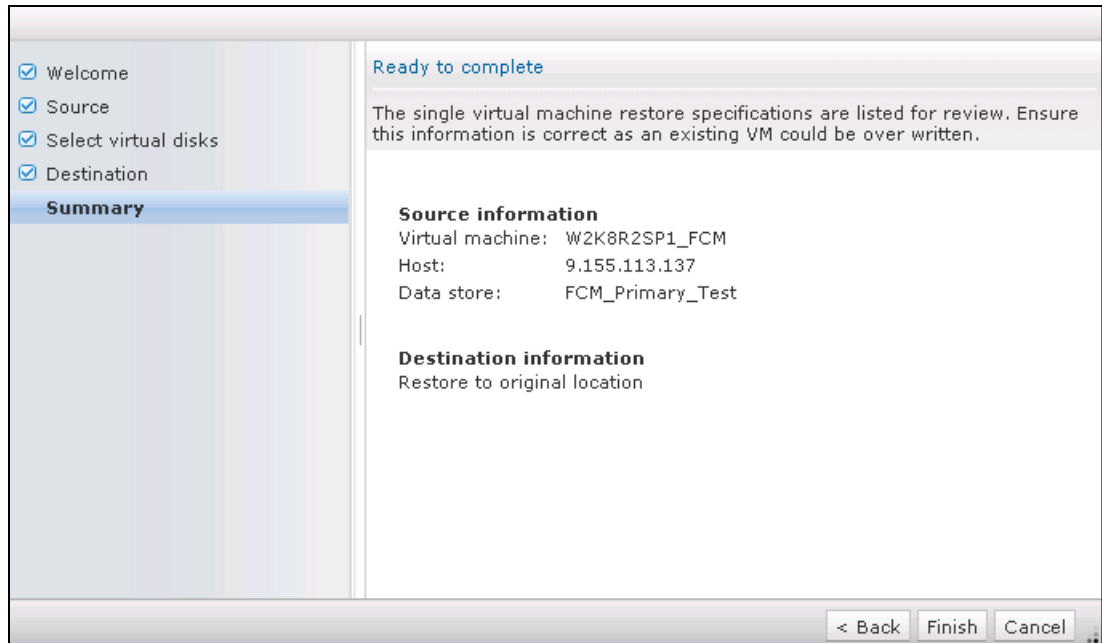


Figure 6-19 FCM Restore Process: Review and Complete

- Figure 6-20 on page 106 illustrates the warning message that is displayed after choosing to restore a virtual machine to its original location. Confirm the over-write by clicking **OK**.

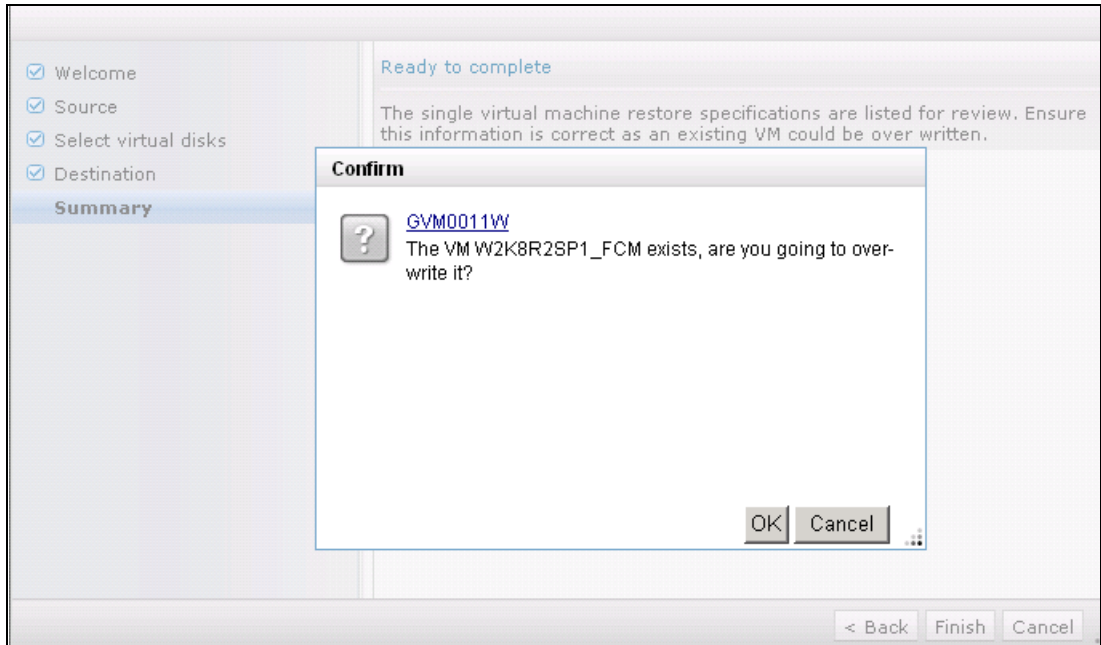


Figure 6-20 FCM Restore Process: Confirm Over-Write

7. As shown in Figure 6-21, click **OK** to monitor the status of the restore task.

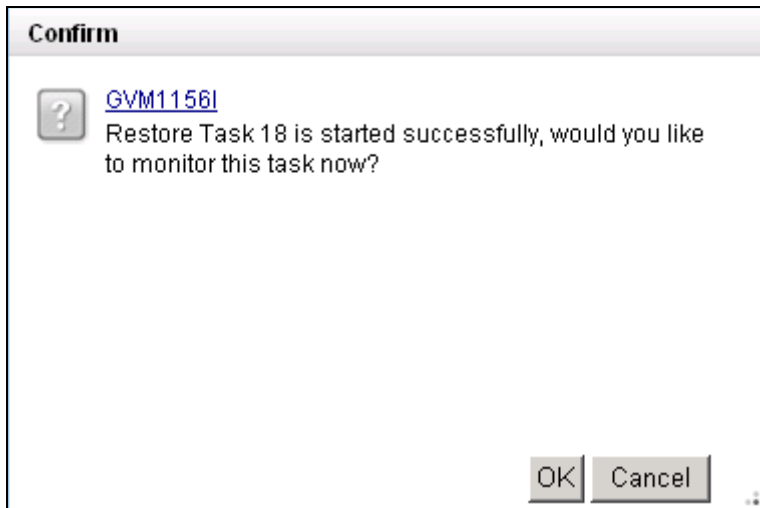


Figure 6-21 FCM Restore Process: Restore Initiated

8. The Recent Tasks view under the Reports tab now displays to provide monitoring of the restore task's completion progress through periodic automatic window refreshes. Note the progress bar and completion percentage shown in Figure 6-22 on page 107.

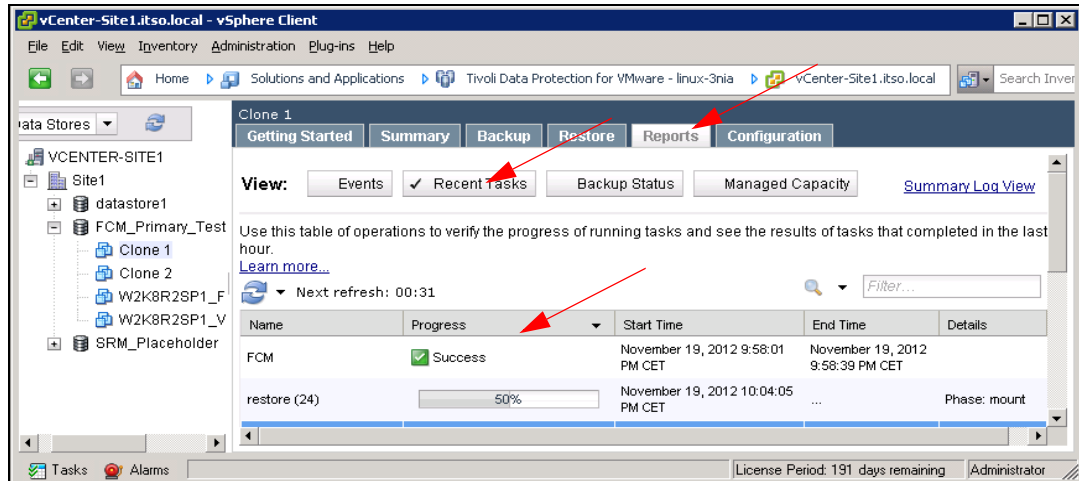


Figure 6-22 FCM Restore Process: Monitoring Restore Process

FCM VMDK attach

The attachment process facilitates the restoration of individual files contained within an available restore point by automating the process of mapping a single VMDK file to a running virtual machine, thereby enabling the associated guest operation system to access the necessary files. To initiate the attachment process:

1. Under the Restore tab, select the source virtual machine, defined as the virtual machine originally mapped to the VMDK file in scope to be restored at the time the backup was performed. Select an available restore point, and click the **Attach** hypertext shown in the upper-left corner of the panel displayed in Figure 6-23.

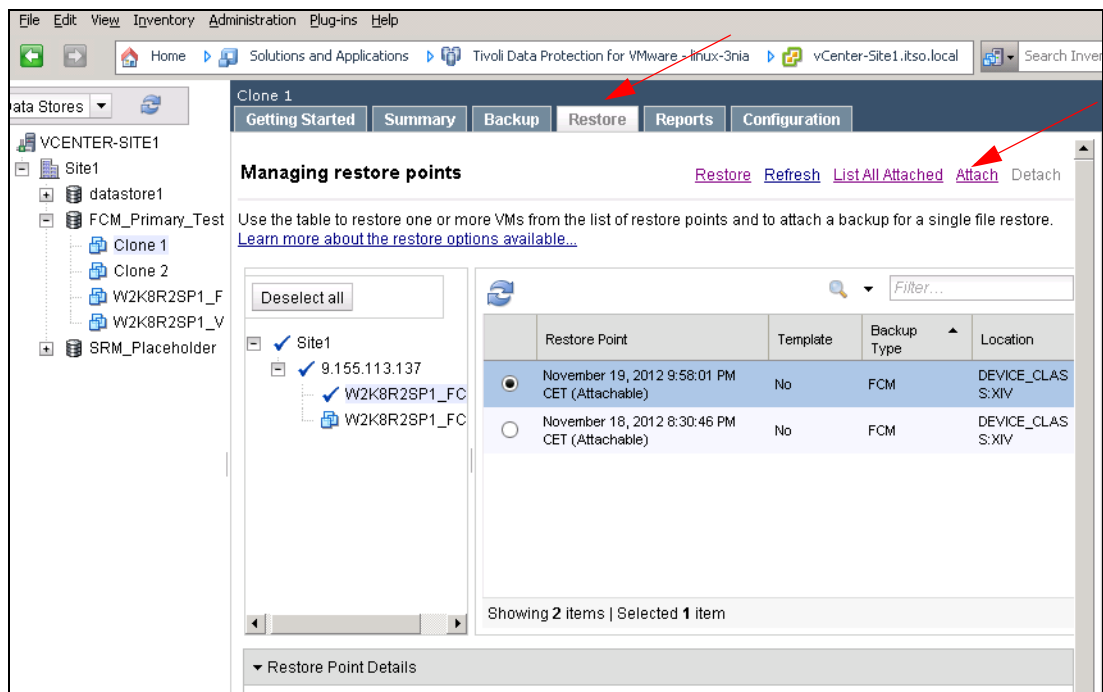


Figure 6-23 FCM Attach Process: Initiate Attachment

2. The resulting Attach Virtual Disk wizard illustrated in Figure 6-24 on page 108 allows administrators to specify a single VMDK file and the target virtual machine for the

attachment process. Click **Browse** to specify the target, and bear in mind that the target must be a running virtual machine.

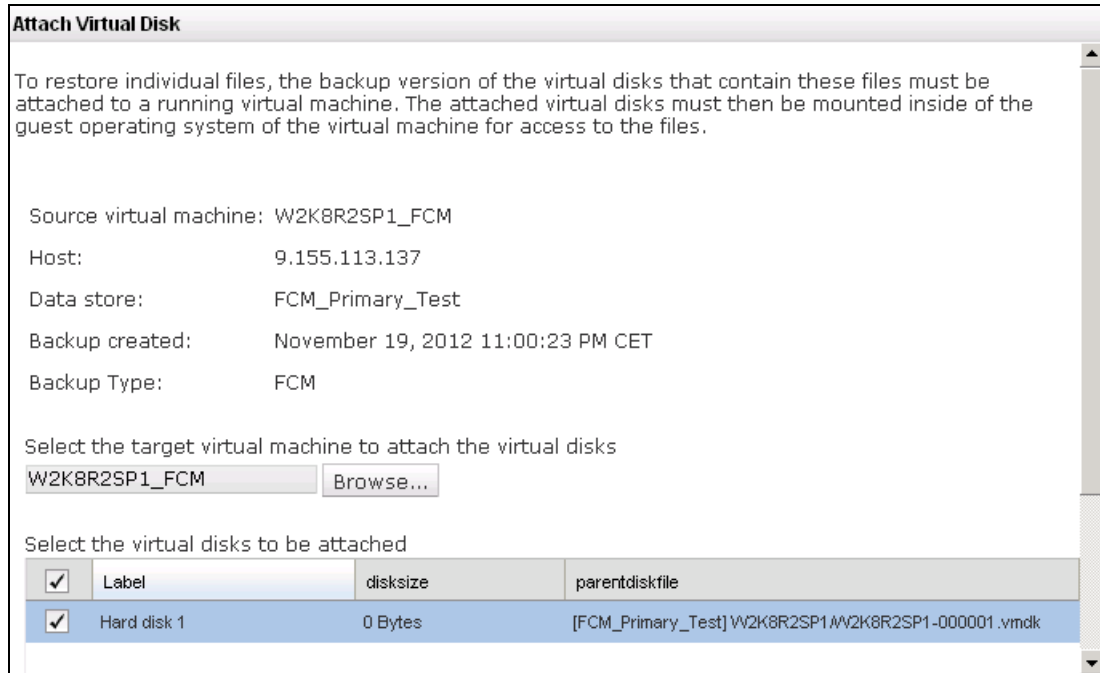


Figure 6-24 FCM Attach Process: Attach Virtual Disk to Target Virtual Machine

3. Select the desired target virtual machine from the Browse Virtual Machines pop-up window shown in Figure 6-25, and click **OK**. There is no requirement for the target virtual machine to be the same as the virtual machine owning the VMDK when the restore point was created.

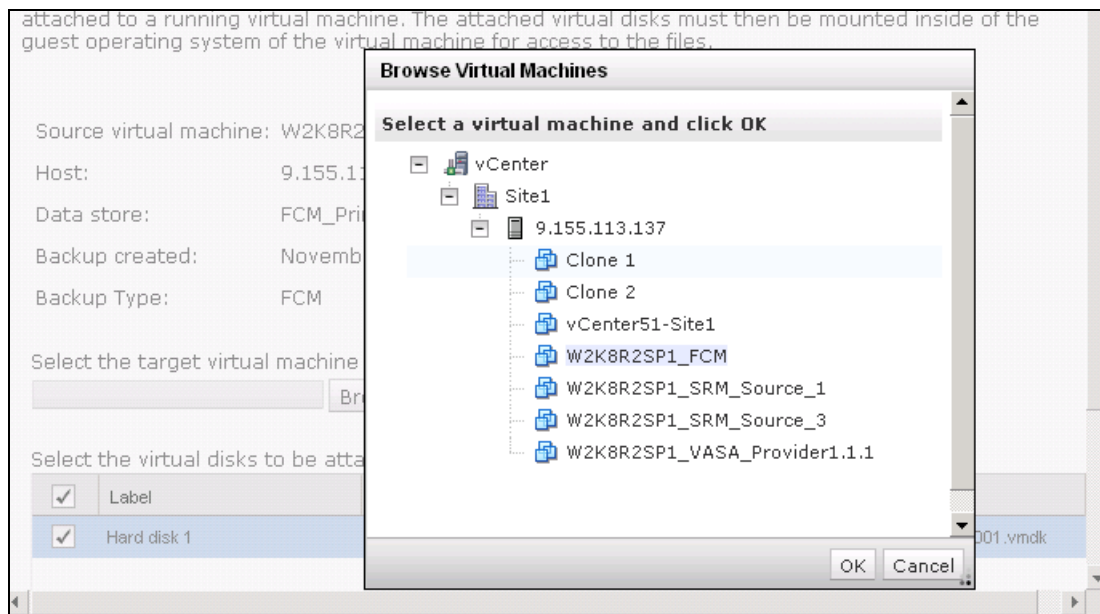


Figure 6-25 FCM Attach Process: Browse Virtual Machines

4. Navigate to the Reports tab, and click **Recent Tasks** in the Views options to monitor the progress of the attachment process, as shown in Figure 6-26 on page 109.

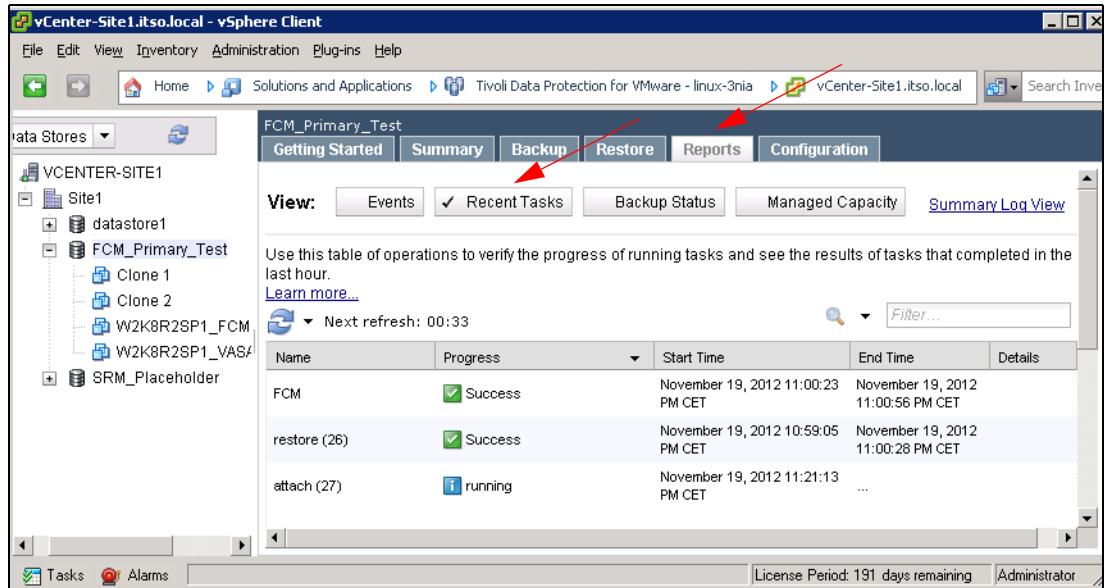


Figure 6-26 FCM Attach Process: Monitoring Attachment Process

5. Click **OK** on the informational alert designating the successful attachment of the VMDK file to the target virtual machine, as illustrated in Figure 6-27.

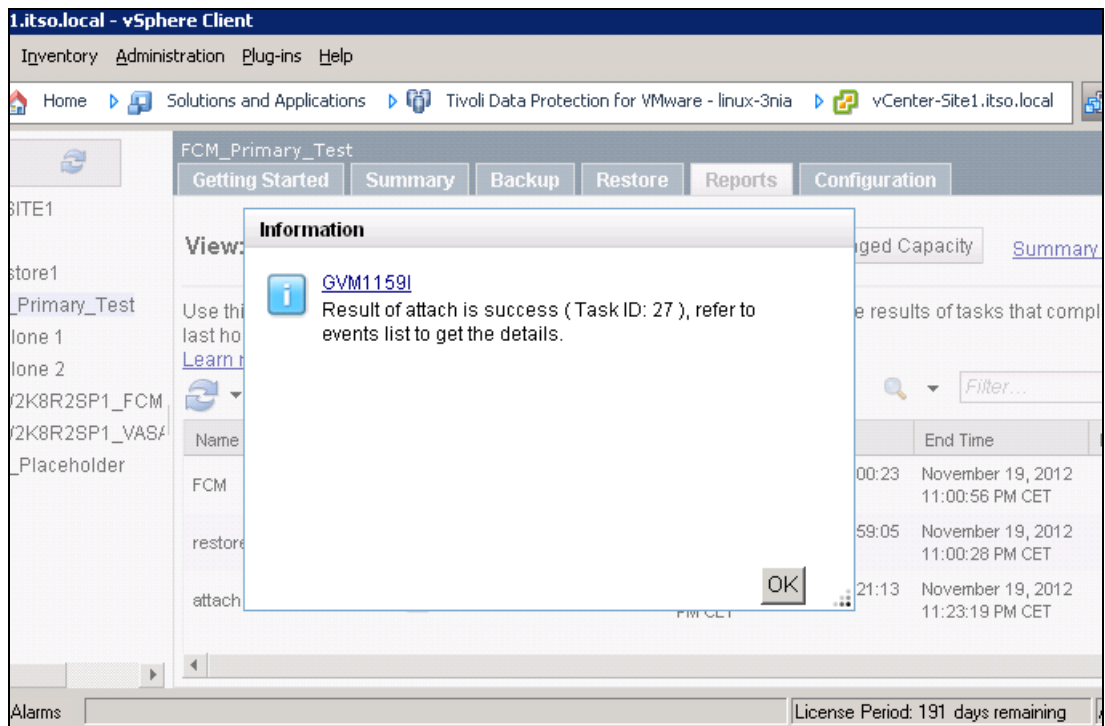


Figure 6-27 FCM Attach Process: Task Completed Successfully

- As illustrated in Figure 6-28, click **Events** in the Views options for a summary of the backup, restore, and attachment tasks in addition to the associated statuses and completion times.

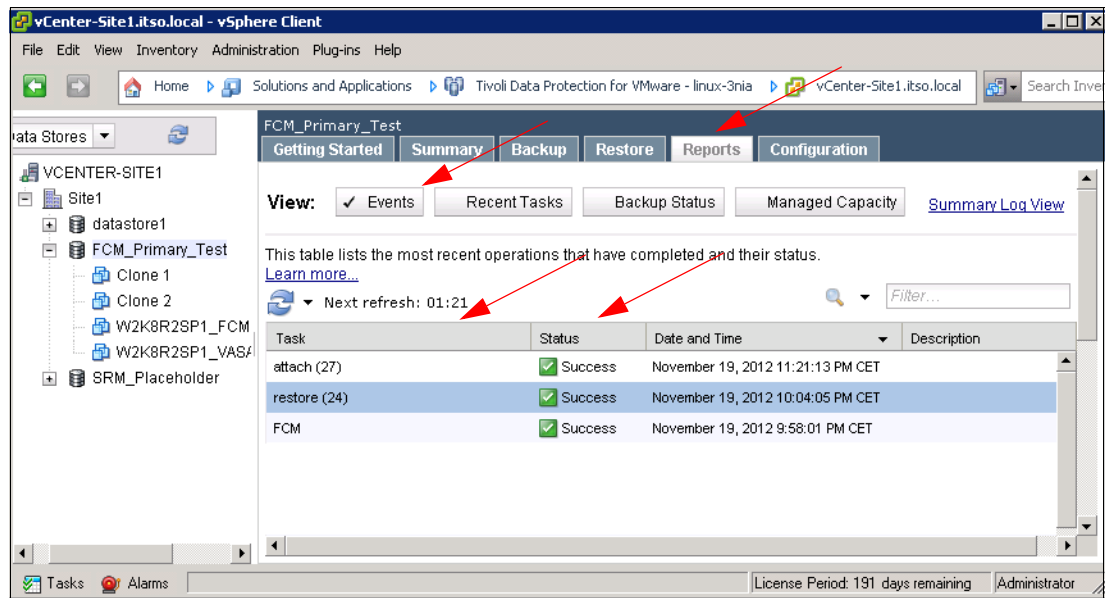


Figure 6-28 FCM Reports: Status of Backup, Restore, and Attach Events



XIV Storage System and VMware Site Recovery Manager

This chapter describes the combined solution of the IBM XIV Storage System and VMware vCenter Site Recovery Manager for business continuity and disaster recovery.

We describe how to deploy and operate the solution.

For the convenience of those willing to implement the solution, the last section of the chapter contains useful information and references about how to initially set up VMware vCenter Site Recovery Manager.

7.1 XIV Storage System and VMware Site Recovery Manager

Planning for all contingencies necessary to meet essential business-driven objectives during events ranging from site-wide disasters to equipment outages constitutes one of the greatest challenges faced by IT professionals, and this challenge is amplified when applied to complex, densely-provisioned virtual infrastructures. Fundamentally, these complexities and challenges stem from the fact that they necessitate a seamless, end-to-end scope of integration that coordinates the underlying mechanics of multiple, dependent components spanning all layers of the infrastructures. Clearly, the scope of potential complexity involved in realizing the rapid, reliable transition of data center services between geographically dispersed sites means that even relatively trivial errors in planning or execution can be devastating to the business in the absence of a comprehensive disaster recovery framework featuring powerful orchestration with turn key execution.

With the combined solution of the IBM XIV Storage System and VMware vCenter Site Recovery Manager, the activities required for meeting business continuity objectives before, during, and after a critical event resulting in site-wide interruption of service are vastly simplified by exploiting features, such as:

- ▶ Drag-and-drop storage replication configuration
- ▶ Automatic discovery of replicated volumes
- ▶ Policy-driven power up procedures at the Disaster Recovery (DR) site
- ▶ Flexible failover and failback
- ▶ High performance incremental snapshots for testing DR without disruption to existing replication

In addition to streamlining classic disaster recovery operations, the combined IBM XIV Storage System and VMware Site Recovery Manager solution can provide:

- ▶ Testing or data mining on existing replicated production data
- ▶ Backup at the DR site
- ▶ Planned data center migrations:
 - Disaster avoidance
 - Site maintenance

In summary, clients implementing Site Recovery Manager with XIV's advanced replication technology benefit from a comprehensive service-oriented framework jointly spanning disaster recovery and avoidance methodologies while realizing predictable, reliable, and efficient delivery of business continuity services and objectives.

The remainder of this section summarizes the solution components underpinning the disaster recovery and avoidance capabilities inherent to the joint IBM and VMware solutions. It assumes the reader has a fundamental level of proficiency with XIV, VMware, and general disaster recovery concepts and terminology.

Topics including solution design principles, planning considerations, and best practices recommendations are explored as well, in addition to a conceptual examination of several common usage cases. Reference materials covering the details of installation and configuration will also follow.

7.1.1 Overview of XIV Remote Mirroring

Fundamentally, the XIV Remote Mirroring function maintains a real-time copy of consistent data by creating a persistent relationship between two or more XIV storage systems physically connected to the replication fabric using Fibre Channel (FC) or iSCSI links.

As an introduction to the rich features inherent to XIV remote mirroring technology, consider that the traditional core functions typical of remote mirroring solutions are augmented by the following unique capabilities with associated advantages:

- ▶ Both synchronous and asynchronous mirroring are supported on a single XIV system.
- ▶ XIV mirroring is supported for consistency groups and individual volumes and mirrored volumes might be dynamically moved into and out of mirrored consistency groups.
- ▶ XIV mirroring is data aware. Only actual data is replicated.
- ▶ Synchronous mirroring automatically resynchronizes couplings when a connection recovers after a network failure.
- ▶ Both FC and iSCSI protocols are supported, and both can be used to connect between the same XIV systems.
- ▶ XIV mirroring provides an option to automatically create subordinate volumes.
- ▶ XIV allows user specification of initialization and resynchronization speed.

Furthermore, as an overview of the management capabilities facilitated through deep XIV and VMware Site Recovery Manager (XSRM) 5.x integration, consider the comprehensive features available when performing the following DR common tasks:

- ▶ Create the configuration within the vSphere client:
 - Add and remove XIV systems from the SRM configuration.
 - Create recovery plans for XIV-based data stores.
 - Create and manage SRM protection groups for XIV-based data stores.
 - Enable, disable, and view the connectivity status.
 - Review mirroring status between volumes and consistency groups.
- ▶ Leverage SRM 5.x to orchestrate end-to-end workflow of failover and failback operations by harnessing XIV's robust remote mirroring capabilities:
 - Fail the operation over to the recovery site by reversing mirroring, designating the Recovery LUNs as primary for updates, and mapping them to new Primary hosts. Subsequent to restoration of service at the original protected site, enable the re-protect capability to reinstate comprehensive protection of the data center running at the recovery site.
 - Failback by reverse mirroring from the recovery site XIV systems to the protected site XIV systems. Invoke re-protection again to restore the normal steady-state operation of SRM with the protected and recovery sites fulfilling their standard roles.
- ▶ Test the disaster recovery plan:
 - Create and use XIV snapshots of target mirror LUNs in failover testing without interrupting the replication.
 - Create backup LUN snapshot replication points before they become mirroring targets and are overwritten (applies to both failover and failback scenarios).
 - Perform cleanup (delete snapshots).
- ▶ Monitor and manage the XIV systems:
 - Query XIV array details and connectivity health status
 - Detect and display paired XIV arrays (mirrored relationships)

7.1.2 VMware Site Recovery Manager overview

VMware vCenter Site Recovery Manager (SRM) represents a purpose-built business continuity solution that empowers administrators to implement robust, customized, end-to-end disaster recovery, disaster avoidance, and data center migration strategies.

SRM includes features for seamless automation, simplified management, and functionality addressing both testing and execution of the planned or un-planned relocation of data center services from one site to another.

VMware vCenter Site Recovery Manager uses the IBM XIV Gen3 remote mirroring capabilities to create a copy of the data at a secondary location. Software for replicating data on the XIV Storage System is included with every system. Also included is the IBM XIV Storage Replication Adapter (SRA), which allows VMware Site Recovery Manager to suspend, snapshot, re-enable, and reverse replication on the XIV Gen3 system.

VMware vCenter SRM 5.x Disaster Recovery enhancements

VMware vCenter Site Recovery Manager 5.x enhances the ability to build, manage, and execute reliable disaster recovery plans spanning bi-direction relocation of data center services, and thus provides unprecedented levels of protection. The sophistication of SRM disaster recover strategies have expanded through the addition of the following capabilities:

- ▶ **Automated re-protection:** The re-protection capability allows SRM to take advantage of the advanced XIV remote mirroring function to wrap the roles of the primary and DR sites following a failover, thereby continuing data center protection while meeting RPOs and RTOs regardless of which site is currently operating in the role of the primary. In effect, the environment running at the recovery site exploits the original primary data center to establish replication and protection of the environment back to the original protected site through a single click.
- ▶ **Automated failback:** After the re-protection process ensures that XIV remote mirroring is reversed and data synchronization subsequently established at the original primary site, the automated failback capability can restore the roles of the two data centers to their original states, simultaneously restoring the state of the entire vSphere environment and maintaining full site-wide data protection operations. To accomplish this, failback will run the same workflow that was used to migrate the environment to the protected site, thus ensuring that the systems included in the recovery plan are returned to their original environment.
- ▶ **Enhanced dependency definition:** This feature organizes the failover workflow to enforce the order in which virtual machines are restarted at the DR site by expanding the number of priority groups that are available to vSphere administrators and permitting SRM to recognize virtual machine dependencies within a priority group.

Common disaster recovery planning terms and definitions

This section offers a brief review of the universal concepts inherent to disaster recovery and avoidance planning presented using terms and definitions specific to the joint SRM and XIV solution:

- ▶ *Site pairing:* Site pairing establishes the connection between two sites and ensures authentication. After this is done, the two sites can exchange information. This requires administrative privileges at both the sites.
- ▶ *Bi-directional operation:* You can use a single set of paired Site Recovery Manager sites to protect in both directions. In this scenario, each site can simultaneously be a protected and a recovery site for different VMs and XIV volumes, also known as an Active/Active protection scheme. Bi-directional does not mean a single volume or VM is replicated in

both directions at the same time; instead, bidirectional refers to the ability of the Site Recovery Manager to fulfill the role of a protected site, a recovery site, or both simultaneously.

- ▶ *Protected and recovery sites*: In a typical installation, the protected site hosts the mission-critical services while the recovery site is an alternative facility where these services can be migrated. Here again, Site A can be a protected site for some VMs and a recovery site for other VMs at Site B.
- ▶ *Protection groups*: A container for VMs and templates that use the same replicated data store group. Protection Groups consist of pointers to the replicated vSphere data stores containing collections of VMs that get failed over from the protected site to the recovery site during actual disaster recovery or testing operations. Conceptually, SRM protection groups specify the relationship of protected virtual machines in the same way that XIV consistency groups specify recovery relationships among logical volumes.
- ▶ *Recovery plan*: A recovery plan specifies how the virtual machines in a specified set of protection groups are recovered.
- ▶ *Storage Replication Adapter (SRA)*: The XIV software required for Site Recovery Manager to issue replication commands to the XIV array. This software is included with the XIV Gen3 array and available for download on the VMware website.
- ▶ *Recovery point objective (RPO)*: The recovery point objective (RPO) indicates how current the target replica needs to be relative to the source. The RPO reflects the maximal amount of data (within the specified RPO time frame) that is acceptable to lose upon failure or unavailability of the main peer.

XIV Gen3 reports a mirror state as RPO OK or RPO Lagging. This status is determined by:

- The RPO parameter associated with the mirror.
- The time stamp of the master's current *last_replicated snapshot*.
- The current system time. Note that this can be affected by time zone differences between the data centers.

An XIV Gen3 asynchronous mirror state is refreshed based on a system-defined schedule.

- ▶ *Recovery time objective (RTO)*: Unlike RPO, which defines how much data is lost, RTO defines how much downtime is acceptable for a particular application or infrastructure. There are several options in the Site Recovery Manager for assigning recovery priority and scheduling. This defines how VMs recover after a failover is initiated.
- ▶ *Synchronous replication*: Remote Mirroring can be a synchronous copy solution where write operations are completed on both copies (local and remote sites) before they are considered to be complete. This type of remote mirroring is normally used for short distances to minimize the effect of I/O delays inherent to the distance to the remote site. Synchronous replication ensures that the data volumes on both the source and target XIV Gen3 storage systems are exact mirrors. Typically, the host application might notice that the writes take longer to process due to the distance delays (latency) in sending the update to the secondary and receiving the response.
- ▶ *Asynchronous replication*: Remote Mirroring can also be an asynchronous solution where consistent sets of data are copied to the remote location at specified intervals and host I/O operations are complete after writing to the primary. This is typically used for long distances between sites. In asynchronous replication, the host update is acknowledged immediately and replication of the updates is sent later to the remote system. In this case, the host avoids the write latency inherent to synchronous mirror designs. Although asynchronous replication has performance benefits, it is a non-zero RPO, which means

some data loss is acceptable when a failover is initiated. Long distances require asynchronous replication in most cases.

- ▶ *Consistency group*: A consistency group consists of volumes that share the same atomic point in time (RPO). When a recovery is issued at the remote site, all volumes will recovery at the exact same point in time. In Site Recovery Manager installations, consistency groups must only be used if actual dependencies across XIV volumes exist (such as data stores with multiple extents).

7.1.3 Minimum XIV and SRM solution prerequisites

In a typical SRM installation, the protected site provides business-critical data center services, and the recovery site provides an alternative facility to which these services can be migrated. The protected site can be any site where virtual infrastructure supports a critical business need. The recovery site can be located thousands of miles away or in the same site. In the typical case, the recovery site is located in a facility that is unlikely to be affected by any environmental, infrastructure, or other disturbances that affect the protected site.

To build an SRM solution featuring the robust XIV Storage System remote mirroring technology, the vSphere and XIV Storage System configurations deployed at both sites must meet the following minimum hardware and software requirements:

The vSphere requirements are:

- ▶ Each site must include at least one vSphere data center with the following components.
 - The vCenter server, SRM server and SRM Client plug-in must be configured at both sites.
 - ESX hosts must exist at both the protected site and the recovery site
 - At least one virtual machine must be located on a replicated data store at the protected site.
- ▶ A Site Recovery Manager license must be installed with a sufficient number of per-virtual machine licenses to cover the systems protected at each site.

Databases must be installed and configured at each site to support vCenter Server and SRM. Supported databases and related interoperability information is located at:

http://partnerweb.vmware.com/comp_guide2/sim/interop_matrix.php?

- ▶ Server hardware resources capable of supporting the same virtual machines and associated workloads as the protected site must be provisioned at the DR site.
- ▶ A comprehensive list of SRM installation prerequisites including the specific minimum hardware requirements for the SRM server deployed on either a physical or virtual machine can be found at:

http://www.vmware.com/pdf/srm_admin_5_0.pdf

The SAN and Networking requirements are:

- ▶ Transmission Control Protocol (TCP) connectivity must be available between the Site Recovery Manager servers and the vCenter servers.
- ▶ The protected and recovery sites must be connected by a reliable Fibre Channel and TCP/IP network with sufficient bandwidth available to meet RPOs and RTOs in the case of asynchronous remote mirroring, and both application workload requirements and RTOs in the case of synchronous remote mirroring.
- ▶ The recovery site must have access to the same public and private networks as the protected site, though not necessarily the same range of network addresses.

- ▶ At both the sites, applicable networking and domain name server (DNS) resources must be set up, configured, and tested before installing SRM.
- ▶ Network hardware resources capable of supporting the same virtual machines and associated workloads as the protected site must be provisioned at the DR site.

XIV Storage System requirements are:

- ▶ The logical volumes backing the data stores must reside on the XIV Storage Systems at both the protected site and the recovery site.
- ▶ The IBM XIV Storage Replication Adapter software must be downloaded from the VMware.com web site and subsequently installed and configured.
- ▶ Both XIV systems must be physically and logically configured for remote mirroring.
Refer to the IBM Redbooks publication *IBM XIV Storage System: Copy Services and Migration*, SG24-7759, for detailed instructions.
- ▶ XIV and SAN hardware resources capable of supporting the same virtual machines and associated workloads as the protected site must be provisioned at the DR site.
- ▶ Sufficient hard capacity must be available on the XIV at each site as dictated by the specific implementation and must include a relatively small amount of capacity dedicated to *placeholder data stores*, which are used by SRM to store virtual machine placeholder files, each roughly 1 KB in size.

The conceptual diagram presented in Figure 7-1 offers a comprehensive view of the topology of the SRM environment harnessing XIV Remote Mirroring with all essential components and their relationships.

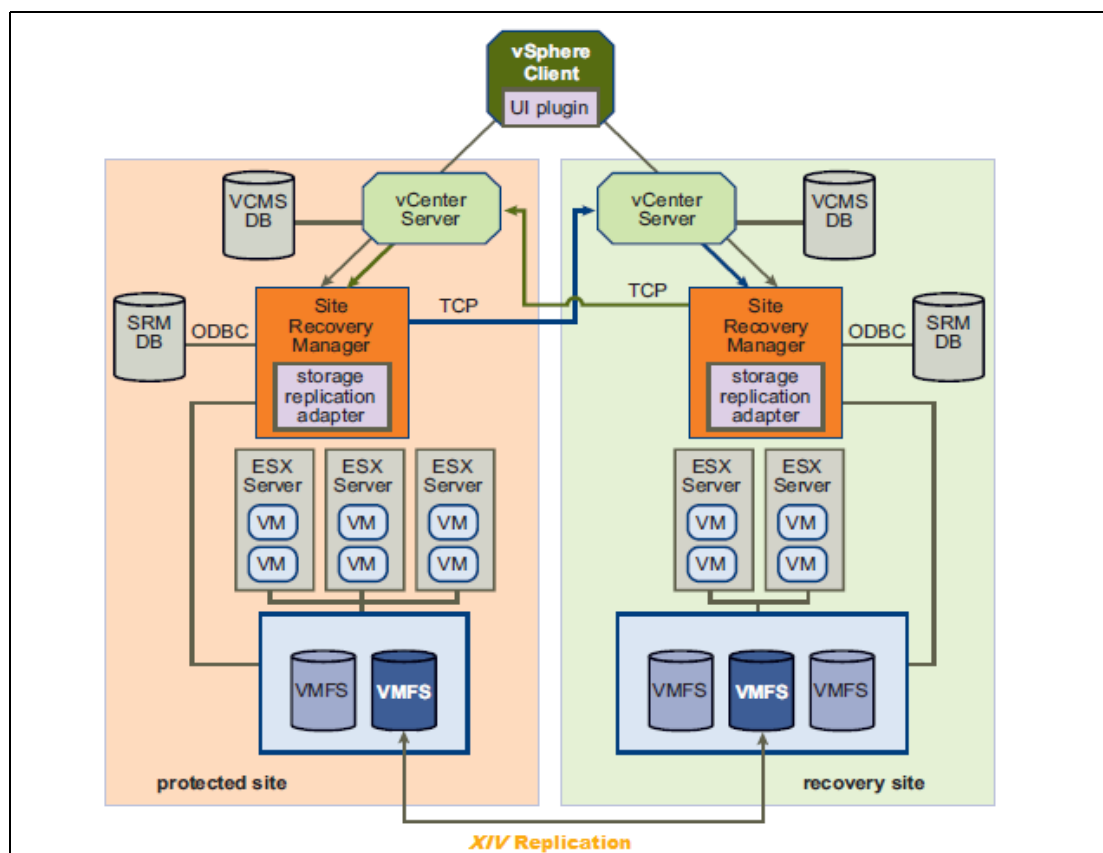


Figure 7-1 XIV and VMware SRM Environment with minimum required components

7.1.4 XIV and SRM integration with the XIV Storage Replication Adapter

As depicted in Figure 7-1 on page 117, the IBM XIV Storage Replication Adapter (SRA) functions as an interface between the XIV Storage Systems and SRM by translating standardized commands generated by SRM into XIV-specific commands. For example, SRA enables SRM commands encompassing execution of the vSphere-centric tasks and workflows, including querying replicated data stores and promoting replicated data stores, to proceed in concert with XIV-specific functions including remote mirroring and snapshot management:

- ▶ Discovering LUNS and their associations in the context of the remote mirroring relationship(s) spanning XIV systems at both the primary and disaster recovery sites.
- ▶ Executing test failover, which is used to test the implementation of the planned disaster recovery workflows by invoking XIV snapshots to create copies of the data stores without impacting operations at either the primary or disaster recovery environments.
- ▶ Automating the failover of an XIV system at the primary SRM site to an XIV system at a recovery (secondary) SRM site.

Immediately upon a failover, the ESX/ESXi servers at the secondary SRM site start using the replicated data stores on the mirrored volumes of the secondary XIV system.

- ▶ Invoking re-protect for either the entirety of the VMware environment or a subset.
- ▶ Executing a failback following a previously completed failover and re-protect.

In summary, the IBM XIV SRA extends SRM capabilities and allows it to seamlessly employ XIV replication and mirroring as part of the SRM comprehensive Disaster Recovery Planning (DRP) solution.

At the time of writing this book, the IBM XIV Storage Replication Adapter supports the following versions of VMware SRM server:

- ▶ 1.0
- ▶ 1.0 U1
- ▶ 4.X
- ▶ 5.X

For instructions detailing the installation of SRA and deploying SRM, see 7.2, “Quick install guide for VMware Site Recovery Manager” on page 160.

7.1.5 Site Recovery Manager operations

VMware vCenter Site Recovery Manager (SRM) is tightly integrated with VMware vSphere. vSphere administrators use the product to initiate automated failover from a primary (protected) site to a secondary (recovery) site. Starting with VMware v5.0, Site Recovery Manager also automates failback to the primary site.

Figure 7-2 on page 119 illustrates a high-level view of the XIV remote mirroring capability orchestrated using key vCenter integration components to provide comprehensive vSphere data center protection.

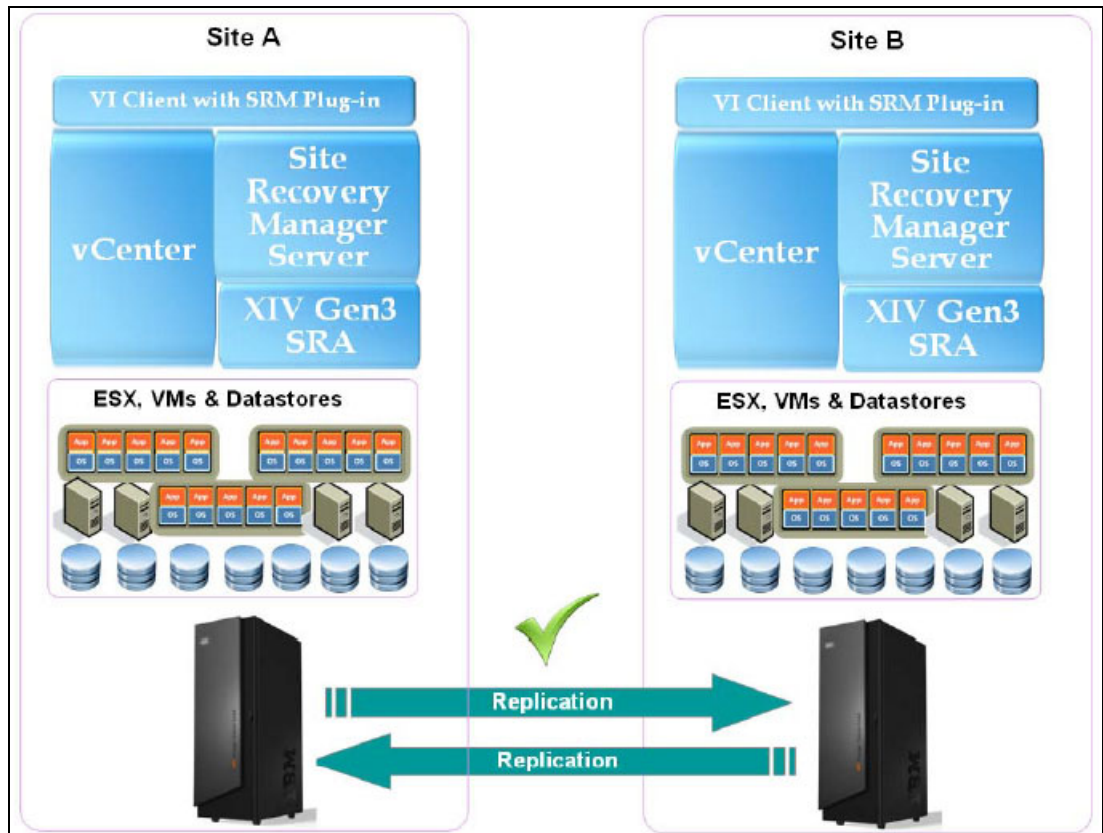


Figure 7-2 XIV Remote Mirroring and integration components for vSphere data center protection

Continuous protection

In normal production, the virtual machines (VMs) run on ESX hosts and storage devices in the primary data center. Additional ESX servers and storage devices are on stand by in the backup data center. As a brief overview, the following common configuration steps are necessary to implement disaster recovery plans:

- ▶ Define protection groups of associated virtual machines that must be recovered together, for example Infrastructure (Active Directory or DNS), Mission Critical, Business Critical, and so on.
- ▶ Define actions prior to a test or failover as the recovery site, such as cloning or suspending low-priority virtual machines to free recovery resources at the recovery site.
- ▶ Define the allocation of resources and any networking changes required by the virtual machines.
- ▶ Build call-outs that cause test or failover process to pause and present instructions to the administrator, or specify scripts in the recovery process.
- ▶ Identify finite values of time or specific numbers of heartbeats to wait for virtual machines to respond after their power-on process is complete.

SRM does not automatically trigger the failover of a recovery plan; instead, human intervention is required to evaluate and declare a disaster. Following the initiation of storage remote mirroring failover processes, SRM executes the earlier steps and automatically manages the mapping of compute resources, network resources, and recovery of VMFS, associated data stores and ultimately virtual machine and associated applications to the recovery site.

During failover

XIV remote mirroring functionality maintains a copy of the data with the necessary scope of consistency and currency to meet the predefined RPO on the XIV system at the disaster recovery location.

In the event that a failover process is triggered, all VMs shut down at the primary site if still possible/required. They are restarted on the ESX hosts at the backup data center, accessing the data on the backup storage system.

SRM servers coordinate the operations of the replicated XIV systems and vCenter servers at two sites with the following sequence of steps:

- ▶ Shuts down the protected virtual machines if there is still connectivity between the sites and they are online:
 - When virtual machines at the protected site are shut down, virtual machines at the recovery site are started-up. The virtual machines subsequently access the data previously replicated from the protected site to assume responsibility for providing the same services.
- ▶ Commands XIV remote replication to synchronize any final data changes between sites.
- ▶ Suspends data replication on the XIV Storage System and the XIV system provides read/write replica devices mapped to vSphere at the recovery site.
- ▶ Rescans the ESX servers at the recovery site to find devices and mounts the data stores
- ▶ Registers the replicated virtual machines.

With the introduction of *Enhanced Dependency Definition* in SRM 5.X, transfer of services from one site to the other is controlled by a recovery plan that specifies the order in which virtual machines are shut down and started up, and the allocation of host and network resources that might be accessed.

- ▶ Completes power-up of replicated protected virtual machines in accordance with the recovery plan.

VMware SRM can automatically perform all these steps and failover complete virtual environments in just one click. This process saves time, eliminates user errors, and provides a detailed documentation of the disaster recovery plan.

After failover

Following a failover operation and the subsequent restoration of the original protected site, site protection can be reinstated by enacting the reversal of site mirroring roles in the Site Recovery Manager 5.x using the Reprotect option.

The Reprotect option helps to communicate with the IBM XIV SRA to reverse the direction of replication. Protection groups now will be replicated from the recovery site to the original primary site, as depicted in Figure 7-3 on page 121.

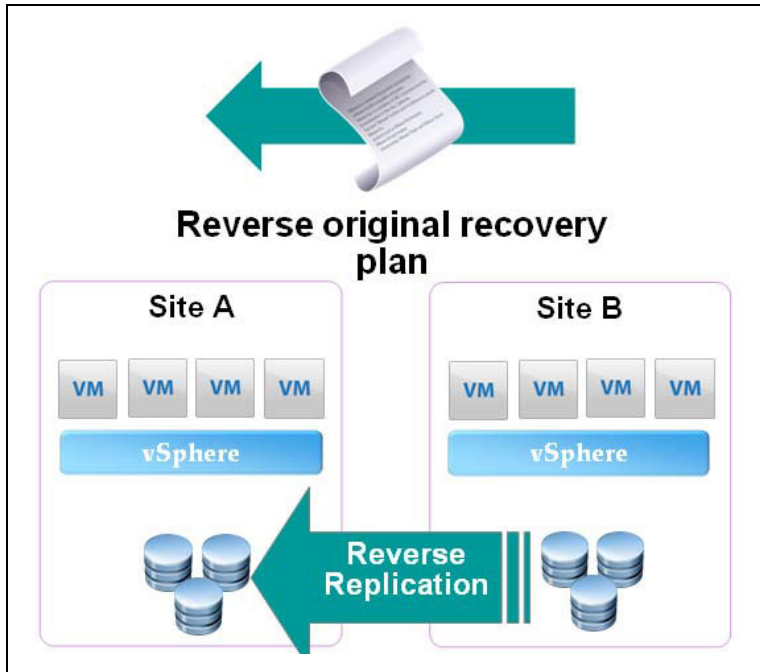


Figure 7-3 SRM Failback

During the reverse, XIV switches the roles of the volumes and replication is reversed. When this is done, only the changes between the volumes will be replicated. An entire resynchronization of the volume is not necessary unless the original volume was completely lost in a disaster event.

After the data is synchronized to the original site, you can failover and then select the **Reprotect** option again to restore the original operational configuration, as shown in Figure 7-4.

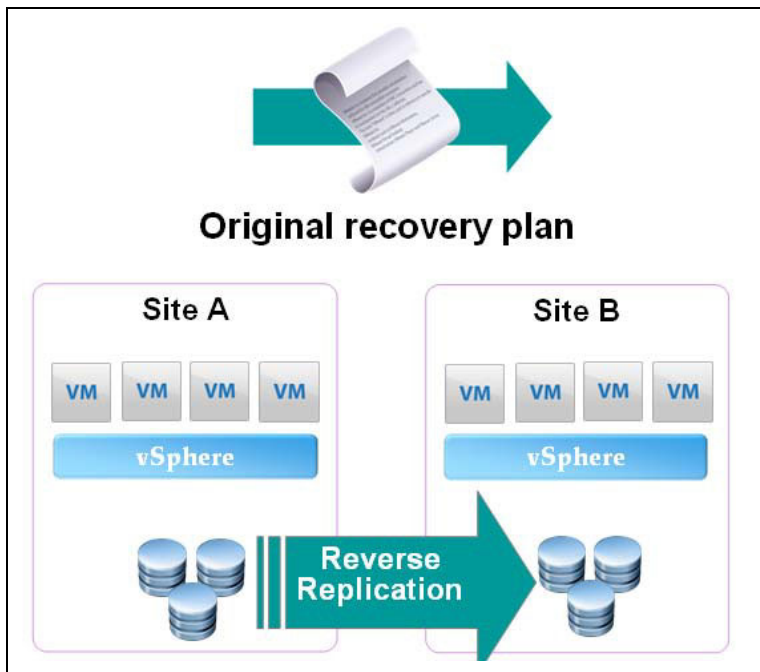


Figure 7-4 SRM Final Re-Protect

Testing failover with the XIV Storage System

SRM can also perform a test of the failover plan by creating an additional copy of the data at the backup site and starting the virtual machines from this copy without actually connecting them to any network. This feature allows vSphere administrators to test recovery plans without interfering with the production environment. In addition, disaster recovery scripts can be tested using XIV high performance, point-in-time, incremental snapshots at the remote site that occur transparently to ongoing data replication processes. These tests provide a deep understanding and functional assessment of the viability of achieving business-driven RPOs and RTOs, and provide an invaluable opportunity to pre-tune performance characteristics of hardware resources and address gaps in business continuance strategies prior to an actual disaster. Figure 7-5 illustrates the fundamental sequence of steps invoked by SRM's testing capability and demonstrates the operational synergy of VMware and XIV snapshot technology necessary for highly efficient and transparent disaster recovery plan testing.

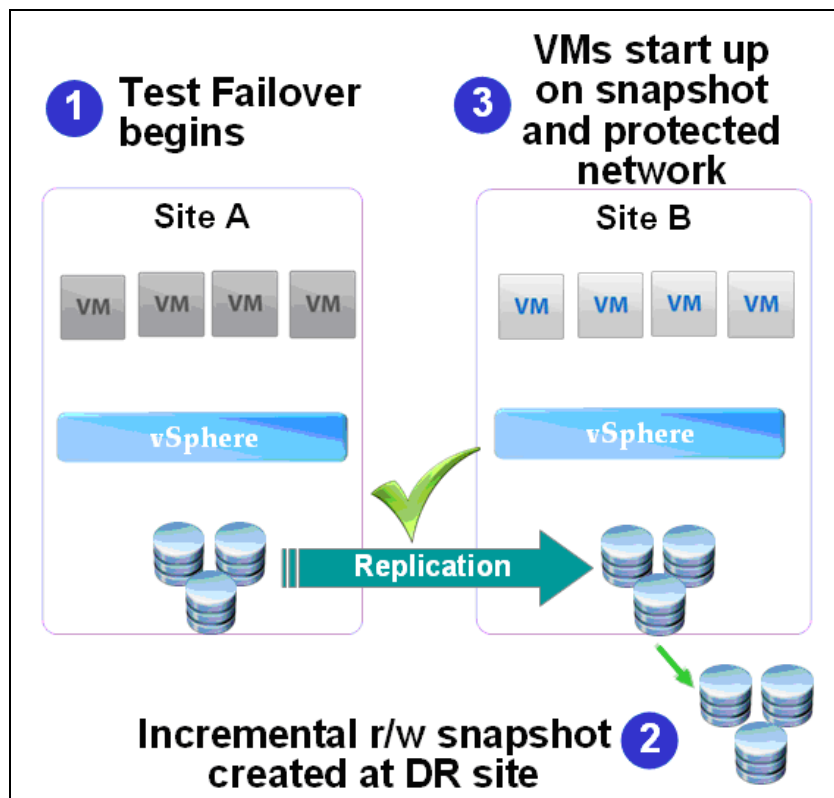


Figure 7-5 SRM Test Failover Plan Leveraging XIV Snapshots

For purposes of testing the DR plan, Site Recovery Manager performs the following tasks:

- ▶ Creates a test environment, including network infrastructure.
- ▶ Rescans the ESX servers at the recovery site to find iSCSI and FC devices, and mounts the XIV snapshot.
- ▶ Registers the replicated virtual machines.
- ▶ Suspends nonessential virtual machines (if specified) at the recovery site to free up resources.
- ▶ Completes power-up of replicated, protected virtual machines in accordance with the recovery plan.
- ▶ Automatically deletes temporary files and resets storage configuration in preparation for a failover or for the next scheduled Site Recovery Manager test.

- ▶ Provides a report of the test results.

High-performance snapshots on the XIV Gen3 system open additional usage cases during test failovers. These can include:

- ▶ Backup at the remote site
- ▶ Data mining
- ▶ Data analytics
- ▶ Test and development

SRM Disaster Recovery plan configuration with XIV

This section offers a step-by-step guide illustrating the flexible, comprehensive SRM disaster recovery features as they are implemented in the SRM utility within vCenter client. To this end, the functional attributes of creating active/passive disaster recovery plan and subsequent invocation of the SRM testing, recovery, and reprotect capabilities will be presented.

Detailed installation guidance including all necessary components necessary to deploy SRM in a vSphere environment provisioned on XIV storage can be found in the 7.2, “Quick install guide for VMware Site Recovery Manager” on page 160. For an overview of the physical and logical configuration steps necessary to deploy XIV systems in a dual-site disaster recovery topology that includes guidance addressing the configuration of connectivity, performance tuning settings, and provisioning mirrored logical volumes and associated consistency groups, consult the following white paper, entitled “VMware vCenter Site Recovery Manager version 5.x guidelines for IBM XIV Gen3 Storage System,” located at:

[https://www.ibm.com/partnerworld/wps/servlet/download/DownloadServlet?id=VZPYFkT7gvZiPCA\\$cnt&attachmentName=VMware_vCenter_Site_Recovery_Manager_version_guidelines_IBM_XIV_Storage.pdf&token=MTM1MjM4MzY5MDg5NA==&locale=en_ALL_ZZ](https://www.ibm.com/partnerworld/wps/servlet/download/DownloadServlet?id=VZPYFkT7gvZiPCA$cnt&attachmentName=VMware_vCenter_Site_Recovery_Manager_version_guidelines_IBM_XIV_Storage.pdf&token=MTM1MjM4MzY5MDg5NA==&locale=en_ALL_ZZ)

SRM setup best practices

In addition to performing installation and setup with the proper order of operations, it is also important to follow the best practices when initially creating the Site Recovery Manager configuration.

VMware vCenter Site Recovery Manager version 5.x guidelines for IBM XIV Gen3 Storage System are:

- ▶ Specify a non-replicated data store for swap files.
This not only avoids unwanted bandwidth consumption, but also improves the recovery time as vCenter does not have to remove all the swap files from the VMs during recovery.
- ▶ Install VMware tools (strongly recommended) on all the VMs participating in a protection group.
Many recovery operations depend on proper VMware tools installation.
- ▶ Configure the VM dependencies across priority groups instead of setting it per VM.
This assures that VMs are started in parallel. The XIV Gen3 Storage System is optimized for parallel workloads so that this greatly improves performance.

Refer to the *VMware vCenter Site Recovery Manager 5.0 Performance and Best Practices* white paper at the following site for a comprehensive guide to SRM 5.0 implementation best practices:

<http://www.vmware.com/files/pdf/techpaper/srm5-perf.pdf>

SRM 5.1 caveats and limitations:

<https://www.vmware.com/support/srm/srm-releasenotes-5-1-0.html#upgrading>

The remainder of this discussion assumes that these prerequisite steps were completed, resulting in the state of the protected XIV volumes appearing similar to those in the SRM CG Demo consistency group depicted in Figure 7-6.

Name	RPO	Status	Remote Volume
Mirrored Volumes			
geocluster_test		Synchronized	geocluster_test XIV_PFE_03_7804143
izik1_1		Consistent	izik1 XIV_PFE_03_7804143
ESXi5U1_DS1	00:00:30	RPO Lagging	ESXi5U1_DS1 XIV_02_1310114
SAP_DATA_MIG	00:00:30	RPO OK	SAP_DATA_MIG XIV_02_1310114
TA_LS_PROD_Volume_2	00:02:30	RPO OK	TA_LS_DR_Volume_2 XIV_02_1310114
TA_LS_PROD_Volume_1	00:02:30	RPO OK	TA_LS_DR_Volume_1 XIV_02_1310114
TA_LS_PROD_Volume_3	00:02:30	RPO OK	TA_LS_DR_Volume_3 XIV_02_1310114
SRM_CG_Demo			
SRM_Source_3		Synchronized	SRM_Target_3 XIV_02_1310114
SRM_Source_2		Synchronized	SRM_Target_2 XIV_02_1310114
SRM_Source_1		Synchronized	SRM_Target_1 XIV_02_1310114
izikcg	00:00:30	RPO OK	izikcg XIV_PFE_03_7804143
TA_LS_PROD_CG	00:02:30	Inactive	TA_LS_DR_CG XIV_02_1310114

Figure 7-6 XIV Mirrored Volumes and Consistency Groups

SRM implementation and usage

Implementing SRM effectively requires the following sequence of steps in addition to the optional but recommended step of testing the recovery plan(s):

1. Connect the SRM instances at both sites by specifying the IP addresses and authentication credentials of the vCenter Servers.
2. Setup "Inventory Mappings" between sites:
 - a. *Resource Mappings* specify the resources that will be recovered for protected virtual machines.
 - b. *Folder Mappings* correlate the folders between the sites.
 - c. *Network Mappings* link the vSphere data center networks of the two sites.
3. Assign *Placeholder Data stores* at each site:

These data stores serve as repositories for small virtual machine files that function as placeholders, which can be registered and activated when SRM restores operations at the recovery site.
4. Specify alarms and permissions as required.
5. Add and configure *Array Managers*.
6. Create *Protection Groups*.
7. Create a *Recovery Plan*.
8. Test and Cleanup *Recovery Plan(s)*.
9. Invoke *Recovery*.
10. Invoke *Reprotect* (SRM 5.X Only).

The remainder of the discussion focuses primarily on the final five steps because these are specific to the implementation of SRM disaster recovery planning and testing with XIV. The processes outlined in some case make references to SRM capabilities only available in SRM 5.X.

Connect the sites

To configure SRM server for the protected and recovery sites, perform these steps:

1. Run the vCenter Client, and connect to the vCenter server.
2. In the vCenter Client main window, click **Home**, as shown in Figure 7-7.

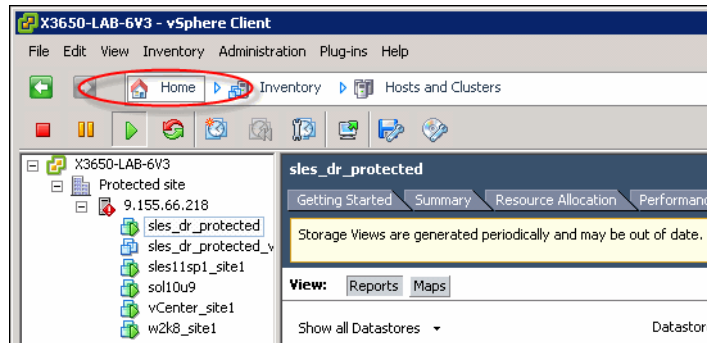


Figure 7-7 Selecting the main vCenter Client window

3. Click **Site Recovery** at the bottom of the main vSphere Client window, as shown in Figure 7-8.

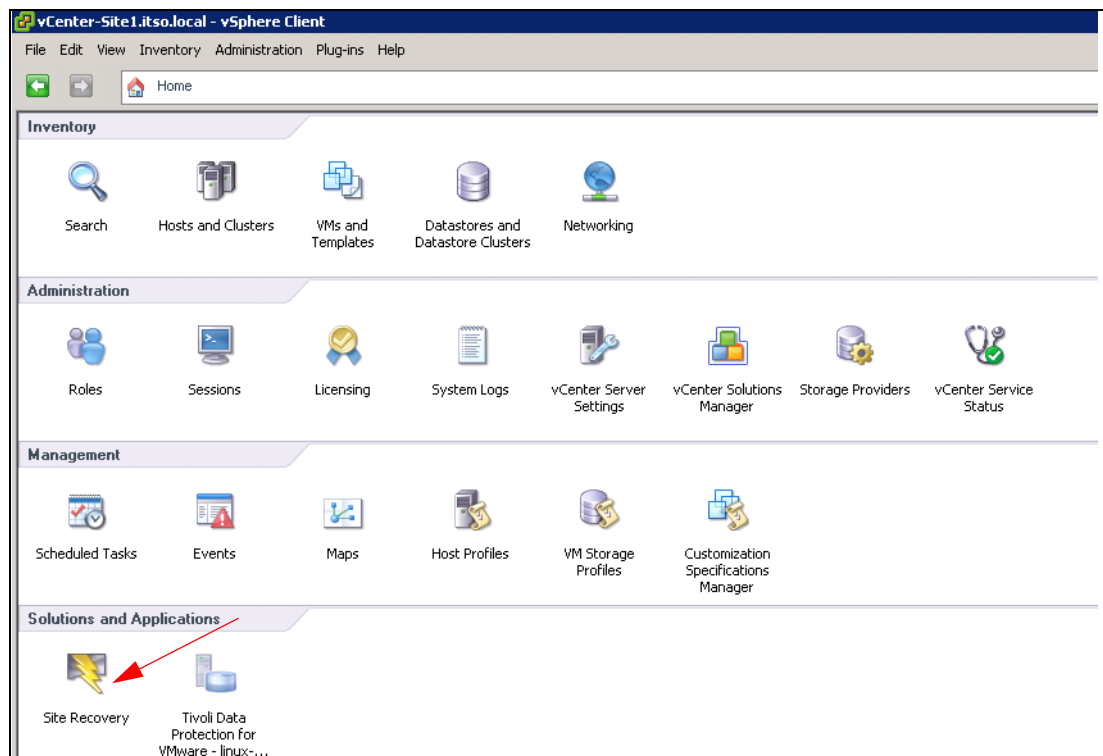


Figure 7-8 Solutions and Applications: Site Recovery Manager

- The SRM Getting Started tab will appear with a list of seven steps. Click the **Configure Connections** hypertext under step one, as shown in Figure 7-9.

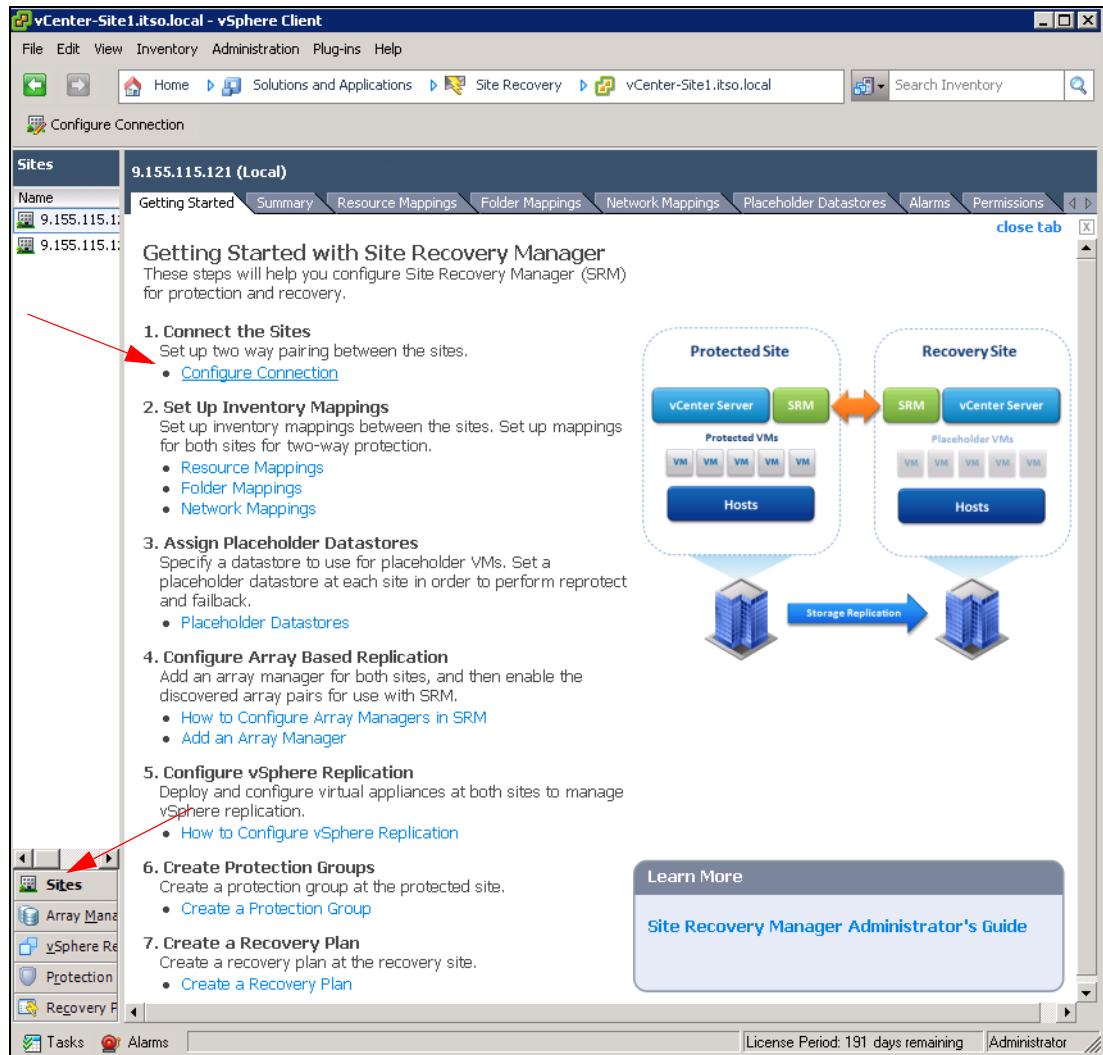


Figure 7-9 Getting Started with Site Recovery Manager: Configure Connection

5. In the Configure Connection window that appears, enter the Remote Site Information by specifying the IP address and communication port of the vCenter server at the recovery site, as shown in Figure 7-10. Proceed by clicking **Next**.

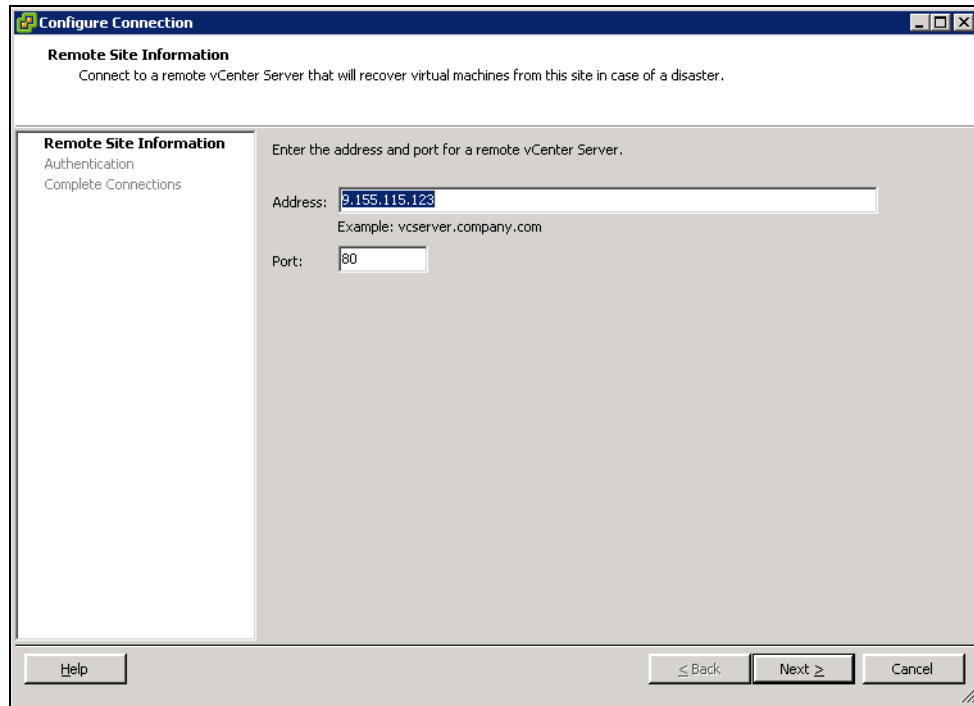


Figure 7-10 Configure SRM Connections: Recovery Site Information

6. A remote vCenter server certificate error might display, as shown in Figure 7-11. Ignore the error message, and click **OK**.

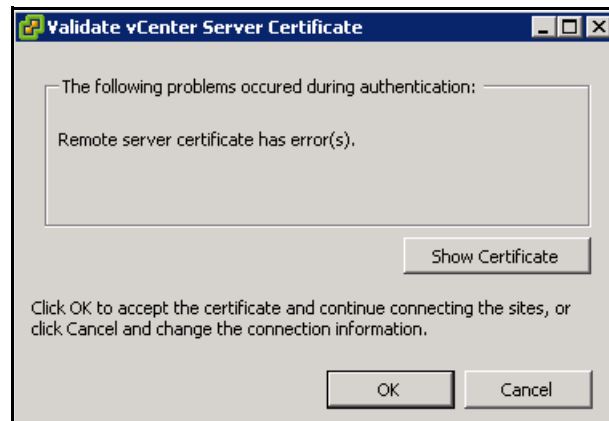


Figure 7-11 vCenter Server Certificate Error Warning

7. Type the administrator login credentials for the remote vCenter Server in the authentication step within the Configure Connection wizard, as illustrated in Figure 7-12. Proceed by clicking **Next**.

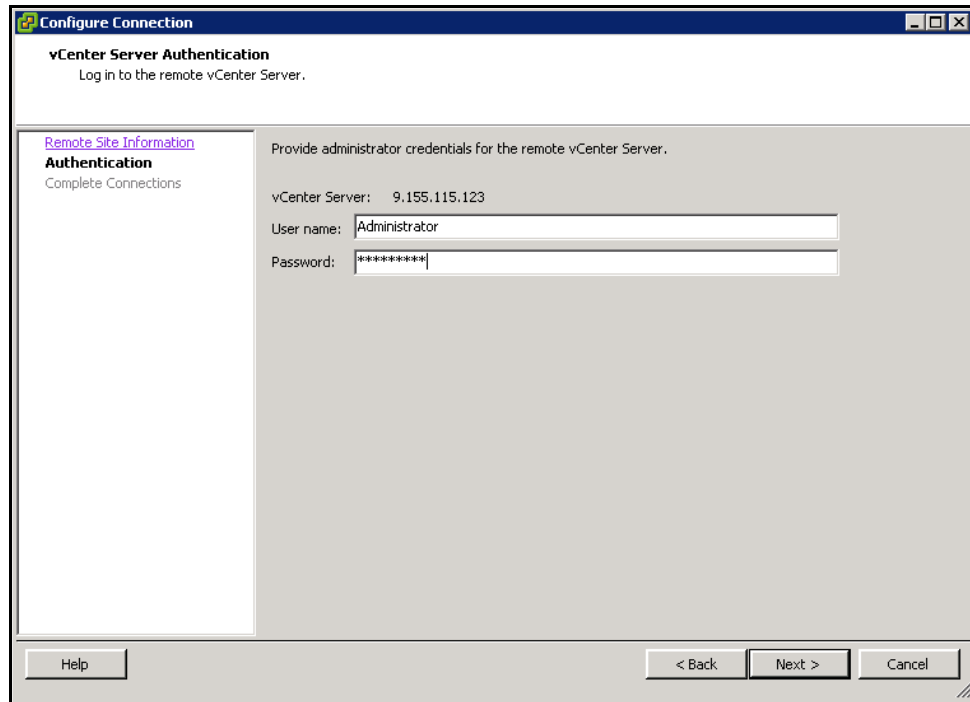


Figure 7-12 Configure SRM Connections: Authentication

8. A configuration summary for the SRM server connection is displayed, as shown in Figure 7-13. Verify the data, and click **Finish**.

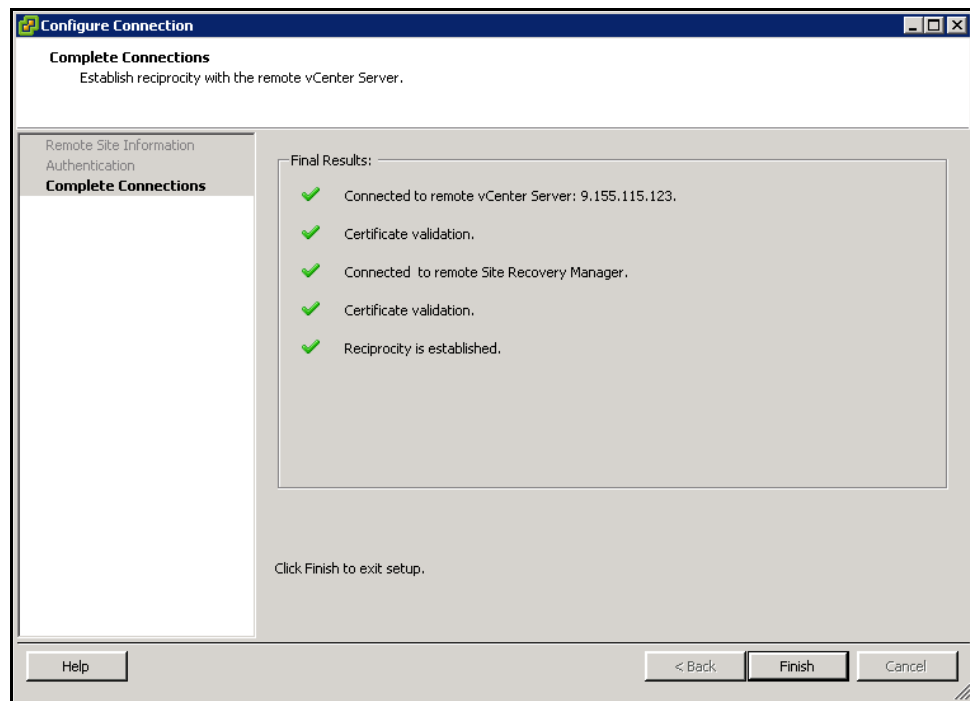


Figure 7-13 Configure SRM Connections: Complete Connections

9. Select the remote site under the left-tree view and then highlight the **Summary** tab to confirm the connectivity information for the remote vCenter server, as shown in Figure 7-14.

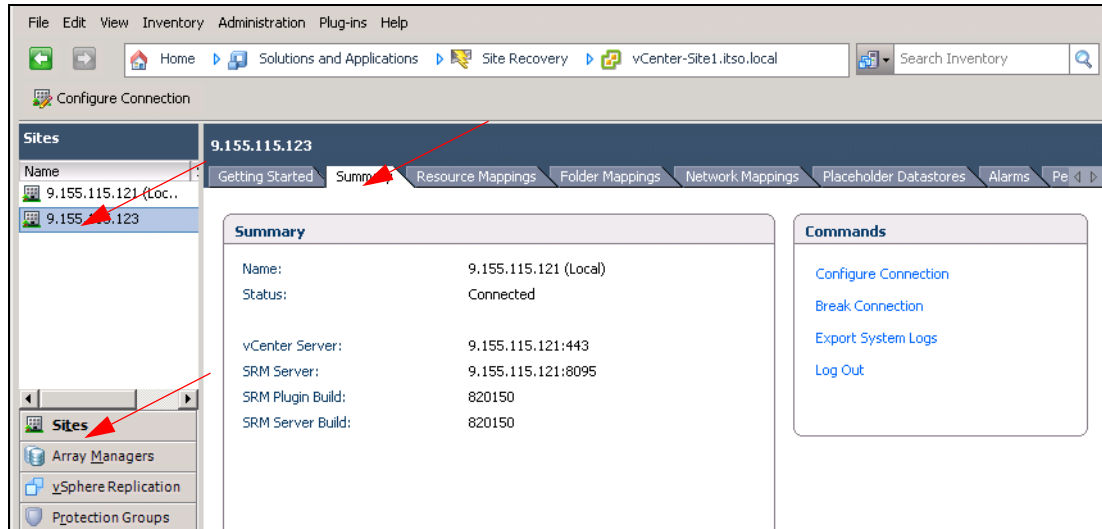


Figure 7-14 Configure SRM Connections: Remote Site Summary

Setup inventory mappings

Inventory mappings specify locations and networks for the vCenter server to use when placeholder virtual machines are initially created at the recovery site, effectively defining how SRM maps resources at the protected site to resources at the recovery site. This phase of the SRM implementation details the process of making inventory mappings for both the protected and recovery sites:

1. Select the protected site under the left-tree view and then highlight the **Resource Mappings** tab to access the SRM GUI panel used to specify resource mappings. Next, select the desired host or existing resource pool under the Protected Site Resources column and then click the **Configure Mapping** hypertext, as shown in Figure 7-15.

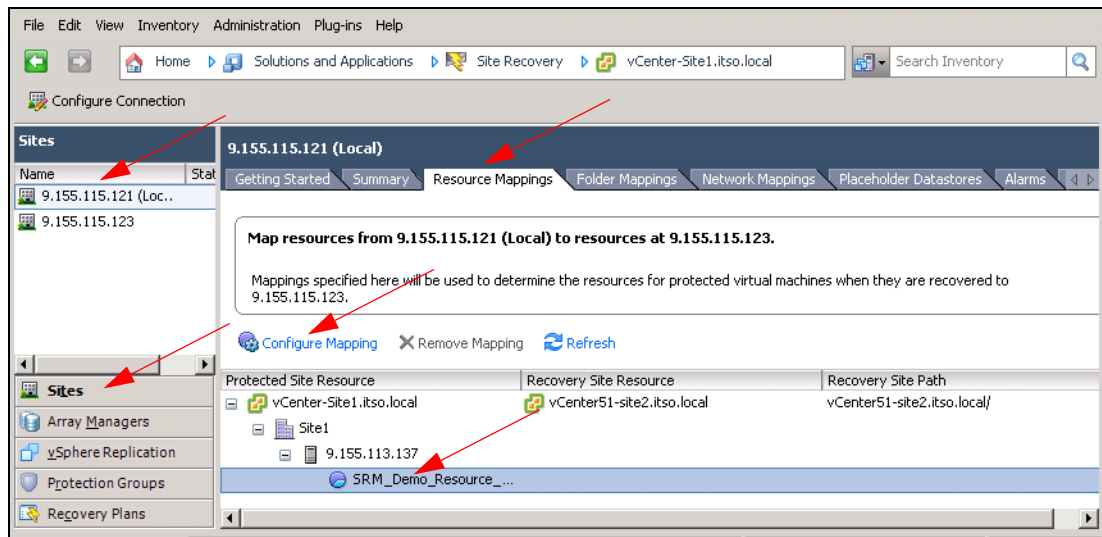


Figure 7-15 SRM Resource Mappings: No Mapping

2. If a new resource pool is desired, click **New Resource Pool** in the Mapping window that is displayed, as shown in Figure 7-16.

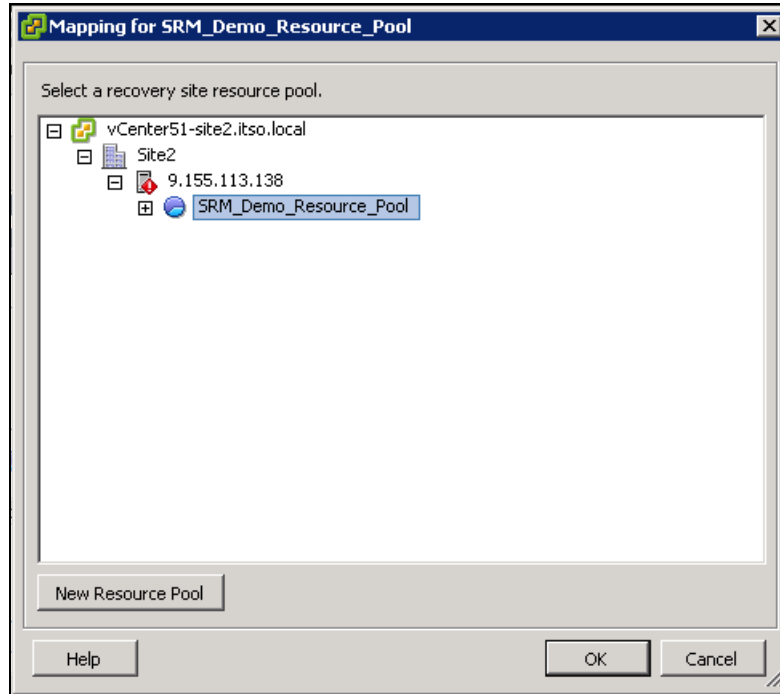


Figure 7-16 SRM Resource Mappings: New Resource Pool

3. The resulting window, shown in Figure 7-17, informs the administrator that resource pools settings can only be changed from the default specifications within the vSphere client GUI, but only the name can be changed within the SRM GUI. Click **OK** to proceed, edit the name of the pool if desired, and click **OK** again in the wizard.

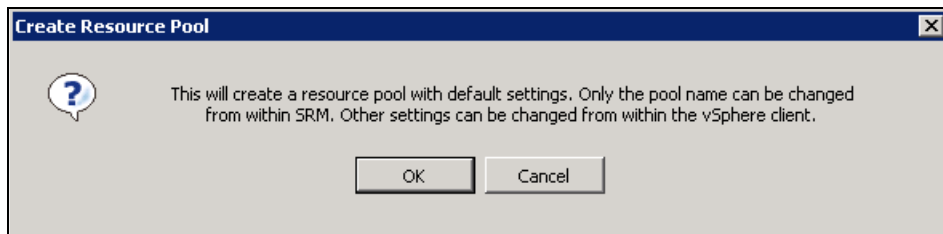


Figure 7-17 SRM Resource Mappings: Resource Pool Notification

- View the Resource Mappings tab to confirm the resource mappings specified, as illustrated in Figure 7-18.

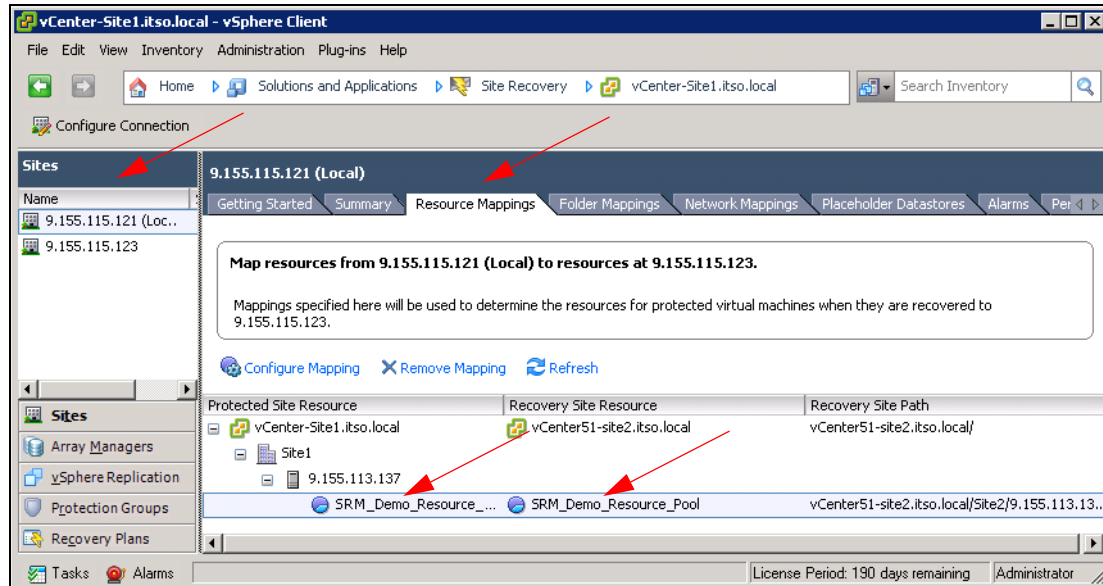


Figure 7-18 SRM Resource Mappings: Completed Mappings

- Select the protected site under the left-tree view and then highlight the **Folder Mappings** tab to access the SRM GUI panel used to specify folder mappings. Next, select the desired sites or existing folder names under the Protected Site Resources column, and then click the **Configure Mapping** hypertext, as shown in Figure 7-19.

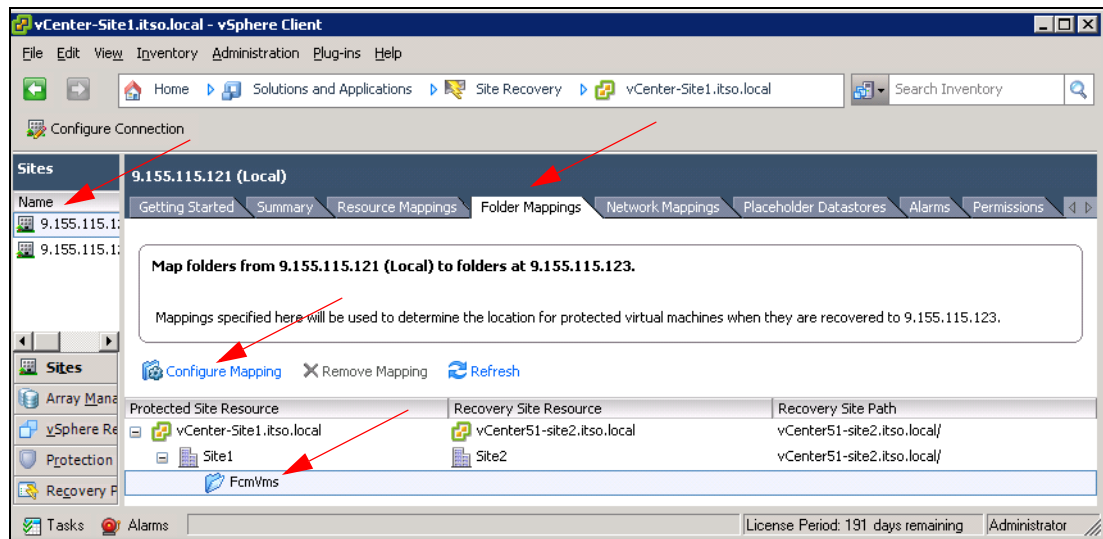


Figure 7-19 SRM Folder Mappings: No Mappings

- The window illustrated in Figure 7-20 allows the administrator to select an existing folder at the recovery site or to create a new one by clicking **New Folder**, renaming the folder as desired and then clicking **OK**.

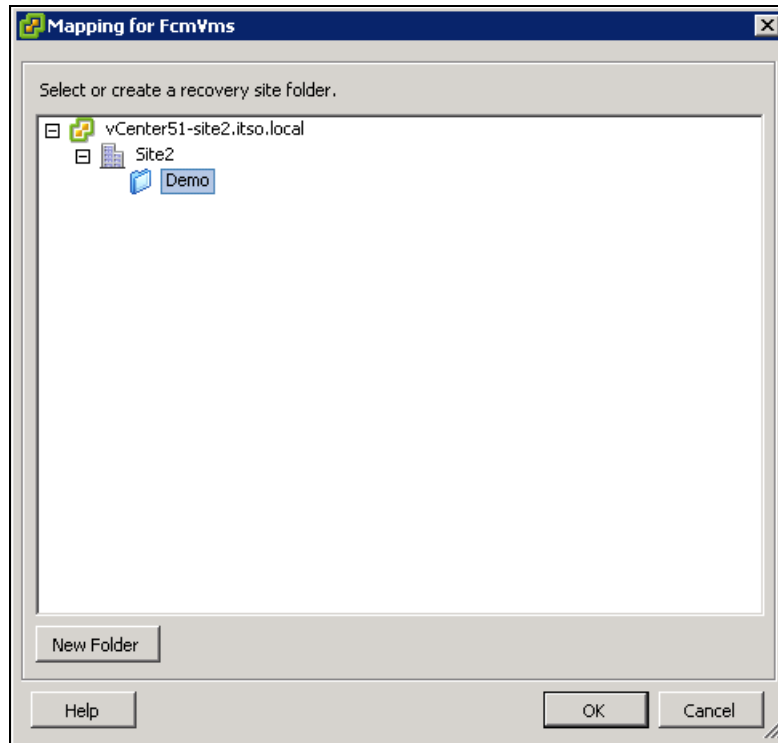


Figure 7-20 SRM Folder Mappings: Create Folder

- View the Folder Mappings tab to confirm the folder mappings specified, as illustrated in Figure 7-21.

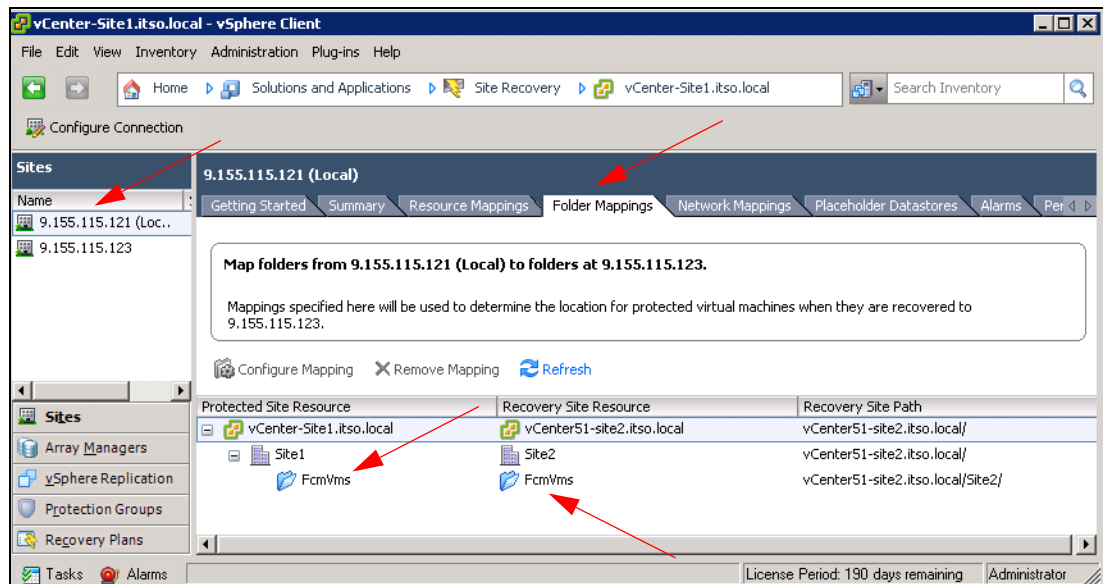


Figure 7-21 SRM Folder Mappings: Completed Mapping

8. Select the protected site under the left-tree view and then highlight the **Network Mappings** tab to access the SRM GUI panel used to specify network mappings. Next, select the existing network name for the protected site under the Protected Site Resources column and then click the **Configure Mapping** hypertext. Note that an existing network must be defined in the vCenter Server at the recovery site before the mapping can be defined in the SRM GUI. After the desired networks are mapped using a process similar to the previous steps describing mapping resource pools and folders, validate the completed mapping, as shown in Figure 7-22.

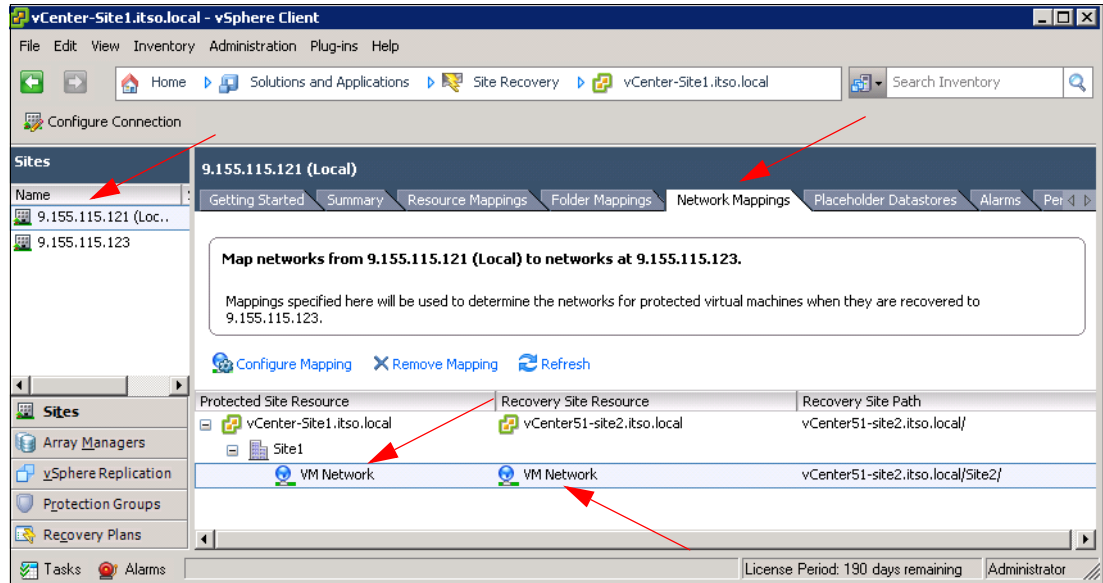


Figure 7-22 SRM Network Mappings: Completed Mapping

9. Repeat the previous eight steps, but this time select the recovery site from the left-tree view in each step. Note that the protected and recovery site designators in the column labels change to indicate that these mappings apply to a failback process.

Configuring placeholder data stores

Placeholder data stores must be assigned at both sites to function as repositories for a subset of metadata files associated with the protected virtual machines that are used to register them in vCenter inventory following an actual failover, failback, or planned migration. Placeholder data stores can be relatively trivial in size, so for most applications backing them with an XIV logical volume with the minimum allocation of 17 GB is probably sufficient. These logical volumes and placeholder data stores must be created and configured in the vSphere client prior to completing the following steps necessary to define them in the SRM configuration.

1. Select the protected site under the left-tree view and then highlight the **Placeholder Datastores** tab to access the SRM GUI panel used to assign placeholder data stores. Next, click the **Configure Placeholder Datastore** hypertext, as shown in Figure 7-23.

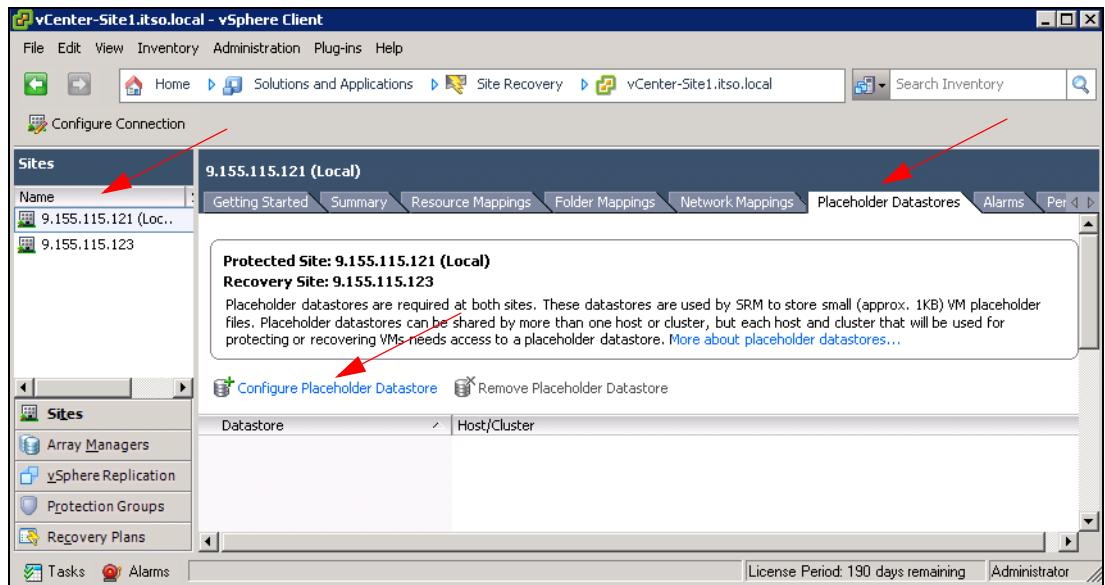


Figure 7-23 SRM Placeholder data stores: No data stores Assigned

2. The Configure Placeholder Datastore window that launches allows the administrator to select a previously defined data store to function in the role of an SRM placeholder for protected virtual machines. As shown in Figure 7-24, select the appropriate data store (named SRM_Placeholder in the example), and click **OK**.

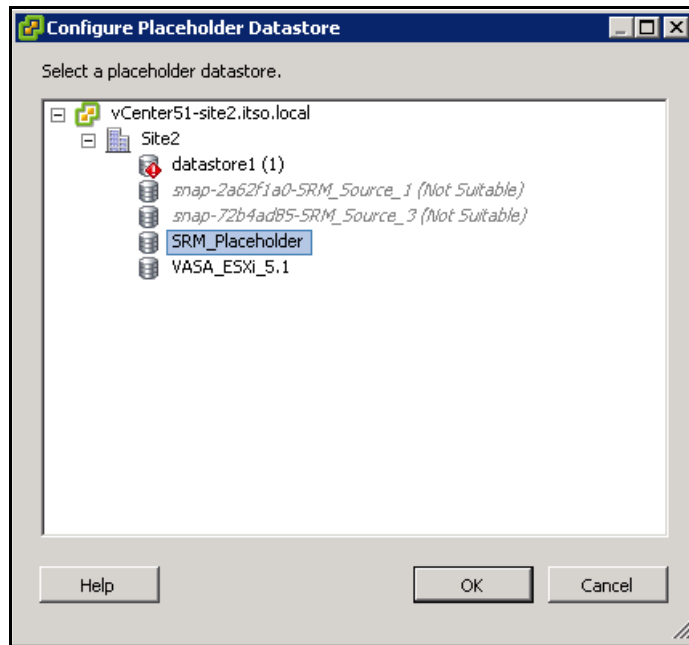


Figure 7-24 SRM Placeholder data stores: Select Previously-Configured Datastore

3. Review the Placeholder Datastores tab to confirm the data stores assigned to function as placeholders, as illustrated in Figure 7-22 on page 133.

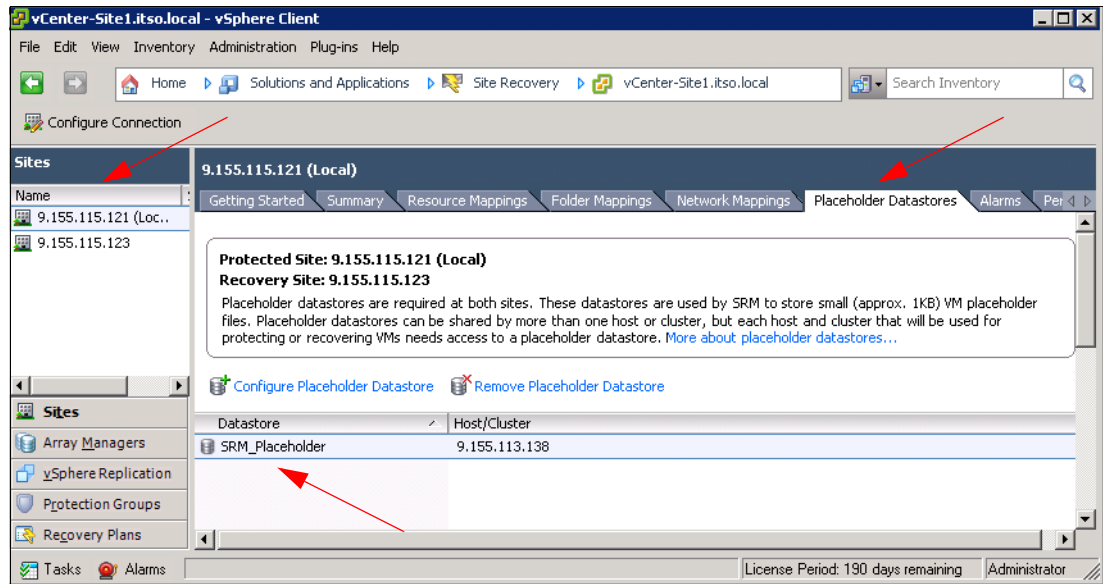


Figure 7-25 SRM Placeholder Data stores: Completion Status

4. Repeat the previous three steps, but this time select the recovery site from the left-tree view in each step. Note that the protected and recovery site designators in the column labels change to indicate that these mappings apply to a failback process.

Adding and configuring Array Managers

In order for SRA to invoke the XIV remote mirroring processes that will underpin the data center site protection workflows in SRM, it is necessary to designate the storage subsystems at each site that will be managed by SRA. To configure the SRA within the SRM GUI:

1. Navigate to the SRM Getting Started tab, select the protected site from the left-tree view, and click the **Configure Connections** hypertext under step four, as shown in Figure 7-26.

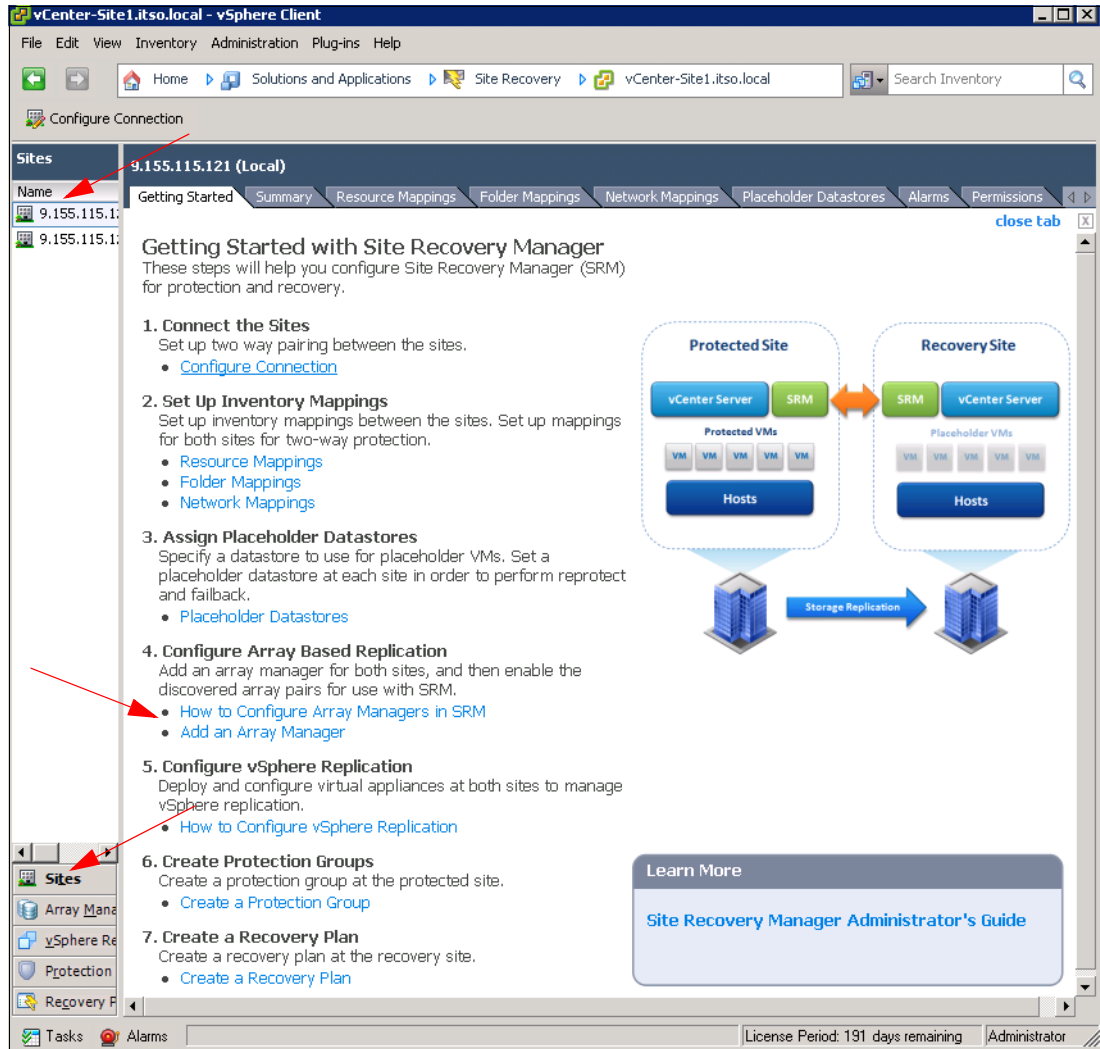


Figure 7-26 Getting Started with Site Recovery Manager: Add Array Managers

2. As shown in Figure 7-27, specify a display name for the XIV system, and ensure the SRA Type is set to IBM XIV SRA. Click **Next** to proceed.

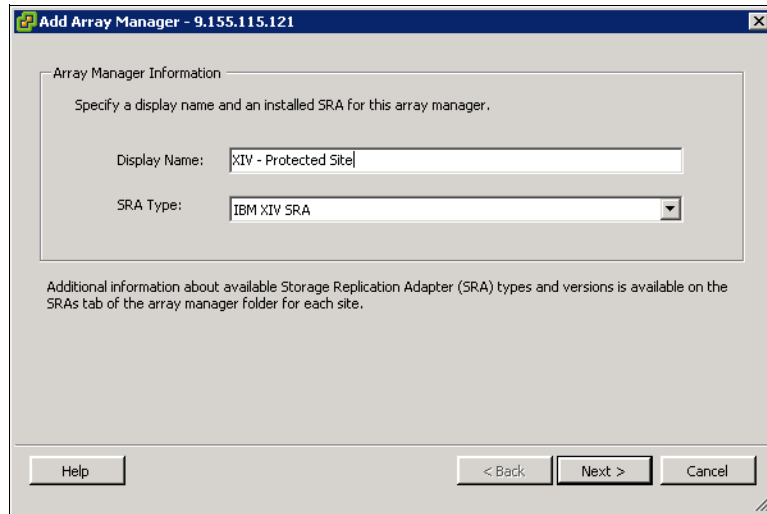


Figure 7-27 Add Array Managers: Set Display Name and SRA Type

3. Populate the fields designating the three IP addresses and the administrator login credentials associated with the XIV systems residing at the protected site, as shown in Figure 7-28. Click **Next** to proceed.

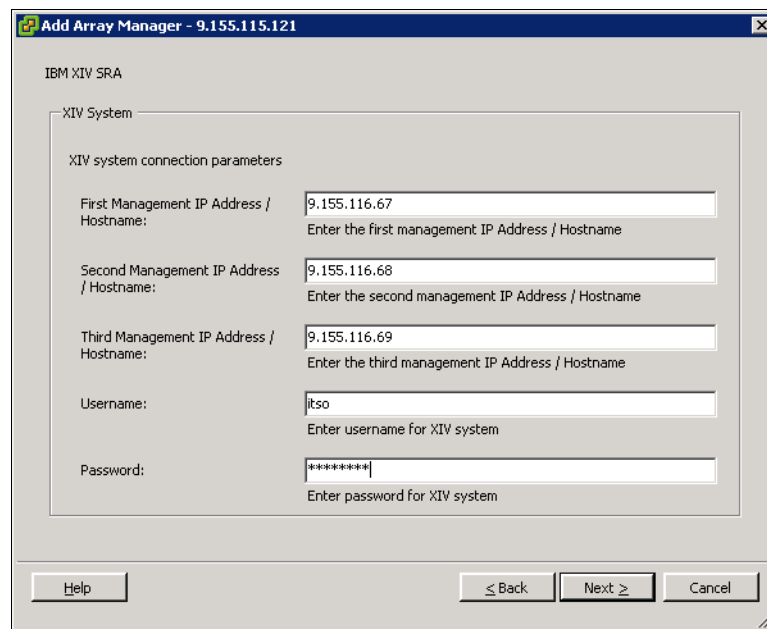


Figure 7-28 Array Managers: Define Connection Parameters using XIV SRA

- Validate that the array manager for the protected site was successfully configured, as demonstrated in Figure 7-29, and click **Finish**.

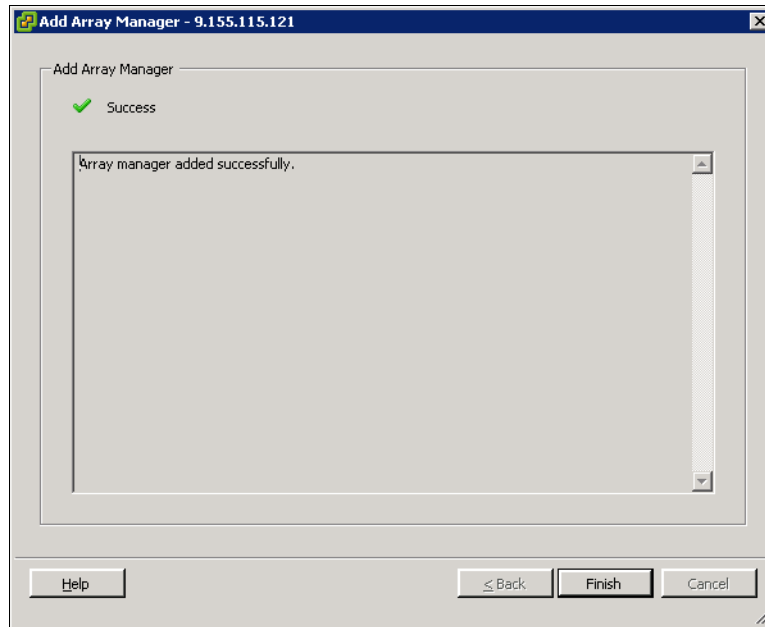


Figure 7-29 Array Managers: Add Array Manager Status Successful

- Repeat the previous four steps, but this time select the recovery site from the left-tree view in the initial step, and ensure the parameters associated with the recovery site XIV system are input into the wizard.
- Under the Array Pairs tab, enable the Discovered Array Pair by clicking the **Enable** hypertext in the Actions column within the row containing the desired local and remote arrays, as shown in Figure 7-30.

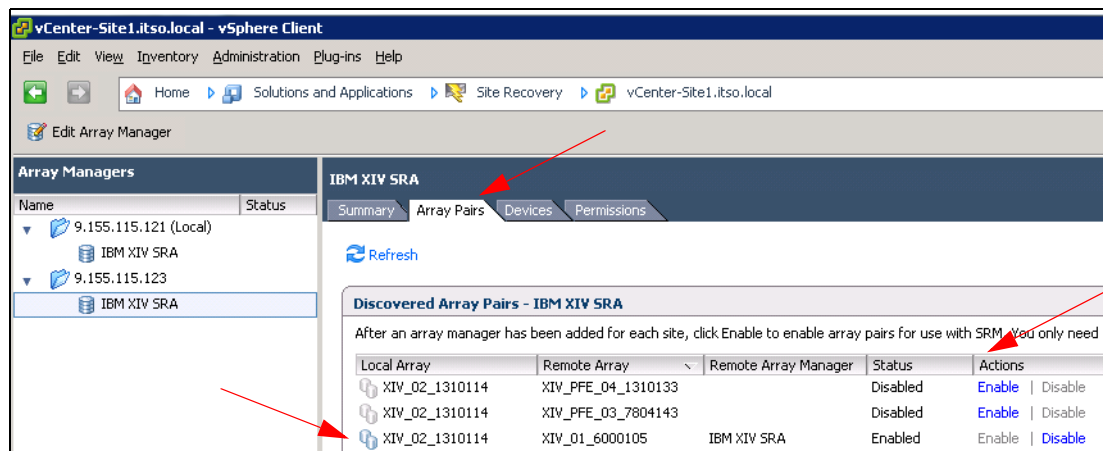


Figure 7-30 Array Managers: Enable Array Pairs for Both Local and Remote Sites

- Navigate to the Devices tab to review the previously defined mirrored logical volumes in the Local Device column and their counterparts in the Remote Device column. Validate that the associated consistency groups and data stores are also defined, as displayed in Figure 7-31.

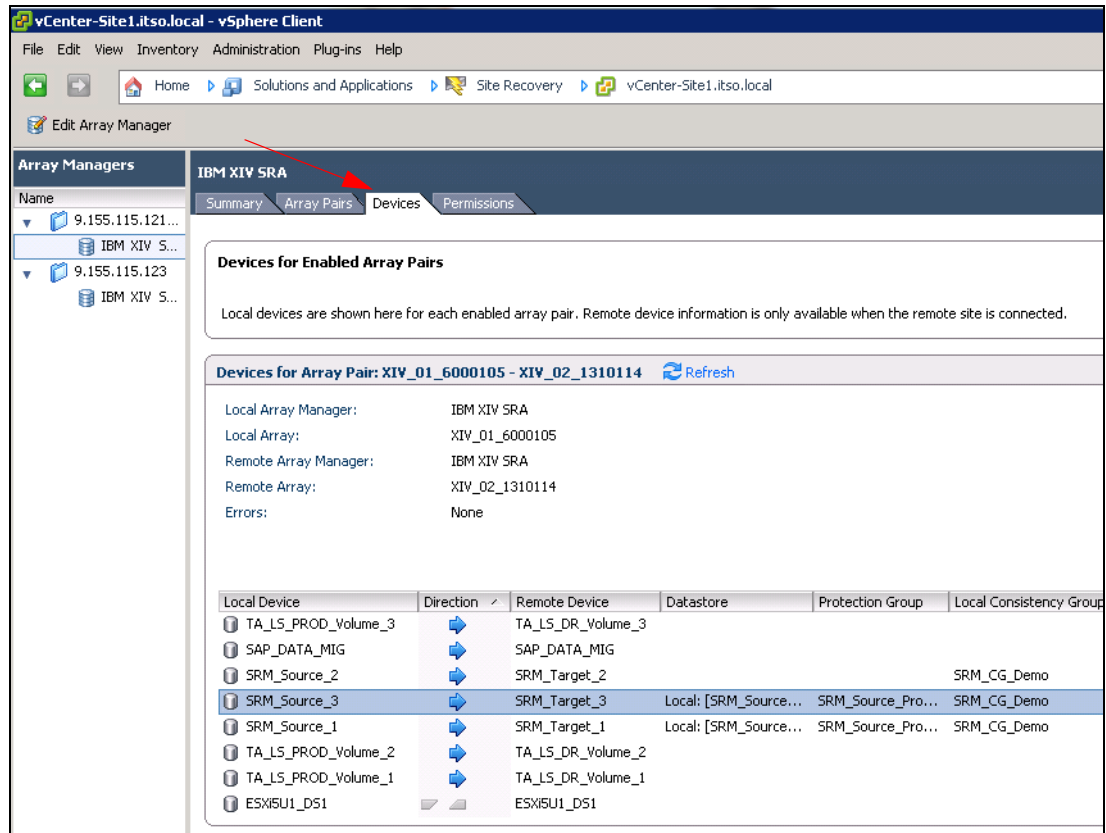


Figure 7-31 Array Managers: XIV SRA Devices

Creating Protection Groups

Protection Groups consist of pointers to the replicated vSphere data stores containing collections of VMs that will be failed over from the protected site to the recovery site during SRM disaster recovery, planned migration, or testing operations. In a way, protection groups are the VMware equivalent of storage consistency groups in that they are logical groupings defined at the data store level instead of the logical volume level. The process of creating protection groups within SRM consists of the following steps, and is a prerequisite to running SRM workflows:

1. Access the SRM Getting Started with Protection Groups panel by clicking **Protection Groups** in the lower-left menu of the parent SRM GUI. Click the **Create a Protection Group** hypertext under step three. Alternatively, click **Create Protection Group** in the upper left of the window, as shown in Figure 7-32.

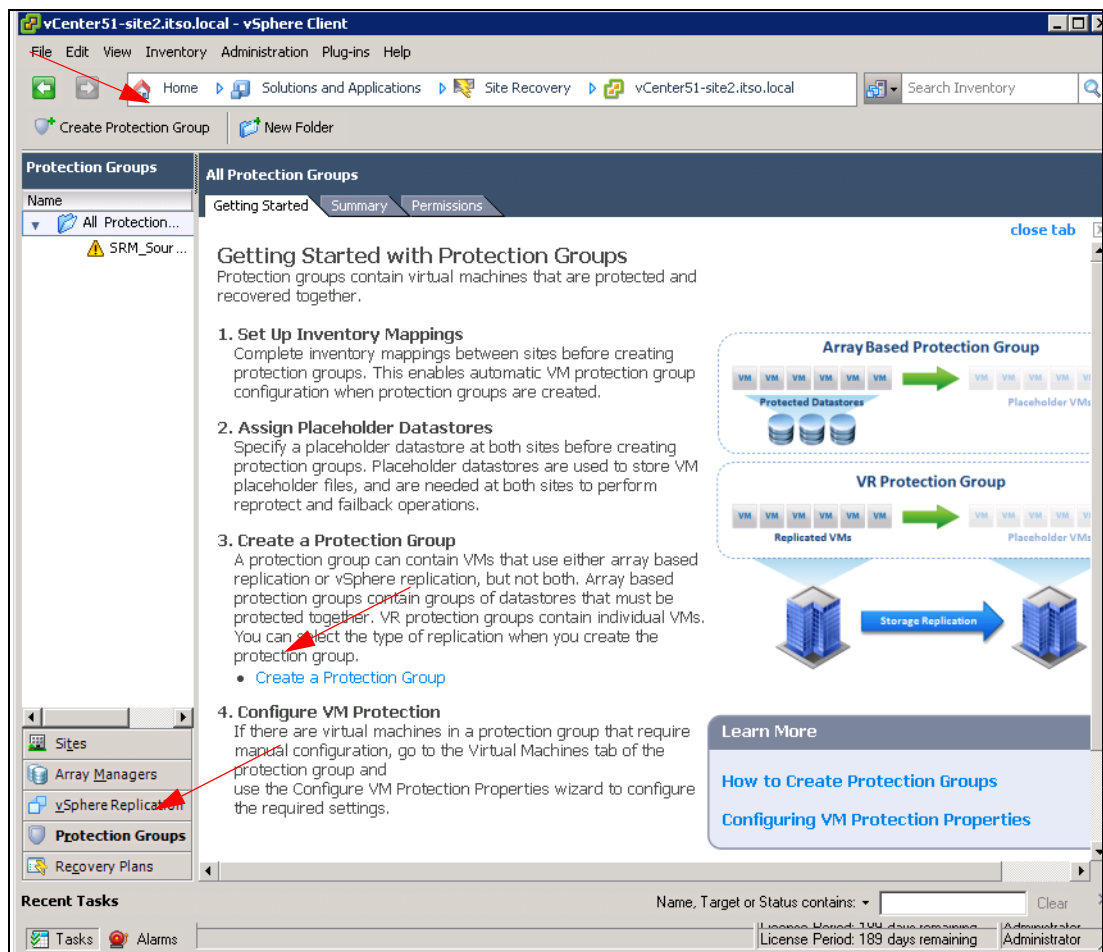


Figure 7-32 Protection Groups: Create New Protection Group

- In the resulting Create Protection Group wizard that is displayed in Figure 7-33, the Protected Site panel contains tasks consisting of selecting the protected site, defining the Protection Group Type to be Array based replication (SAN), and specifying the previously-configured SRA pair under Array Pair. When these tasks are complete, click **Next** to proceed.

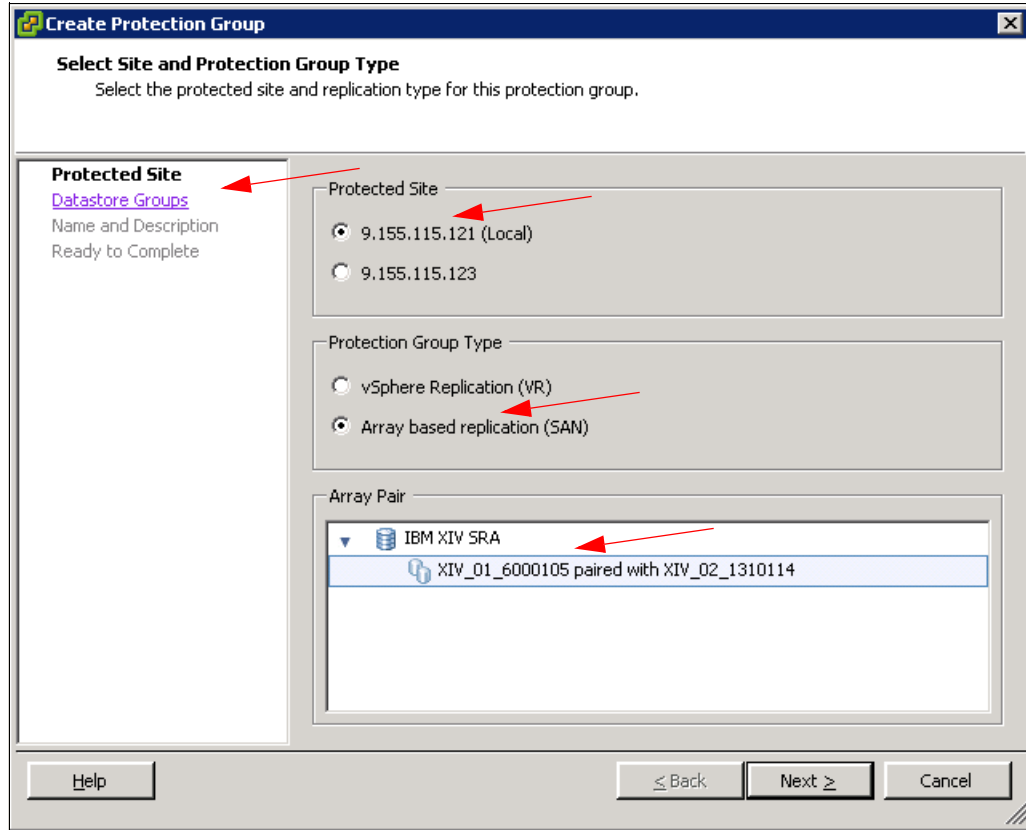


Figure 7-33 Protection Groups: Select Site and Protection Group Type

- The Datastore Groups panel of the create/edit protection groups wizard allows the administrator to specify a data store group in terms of its constituent data stores and associated virtual machines, as illustrated in Figure 7-34 on page 142. **Click** the check box next to the desired data store group and then **click** next to proceed.

Note: If the XIV logical volumes backing the data stores are defined as part of a consistency group, SRM will force them to be included within the data store group.

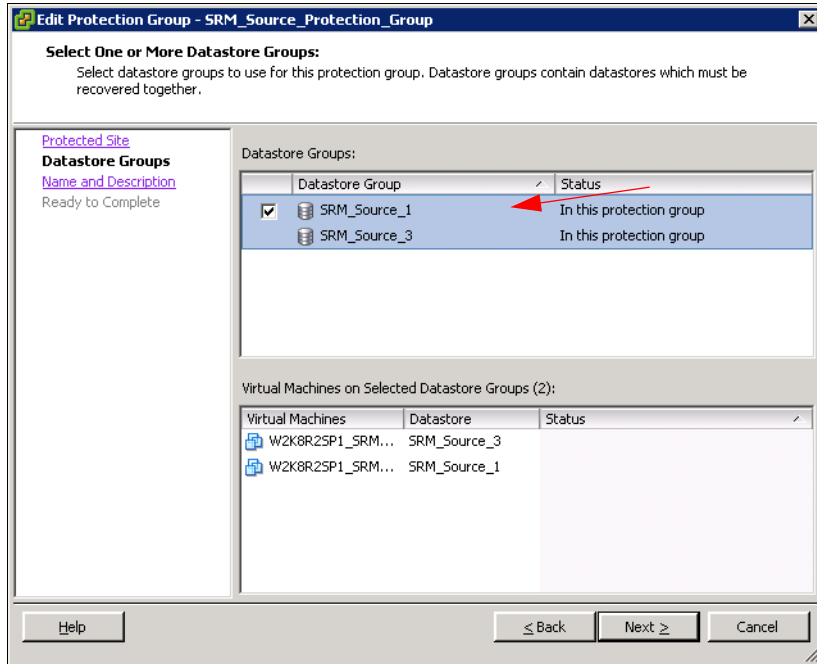


Figure 7-34 Protection Groups: Select Data stores

- As shown in Figure 7-35, type in the desired name and description for the data store group, and click **Next** to proceed.

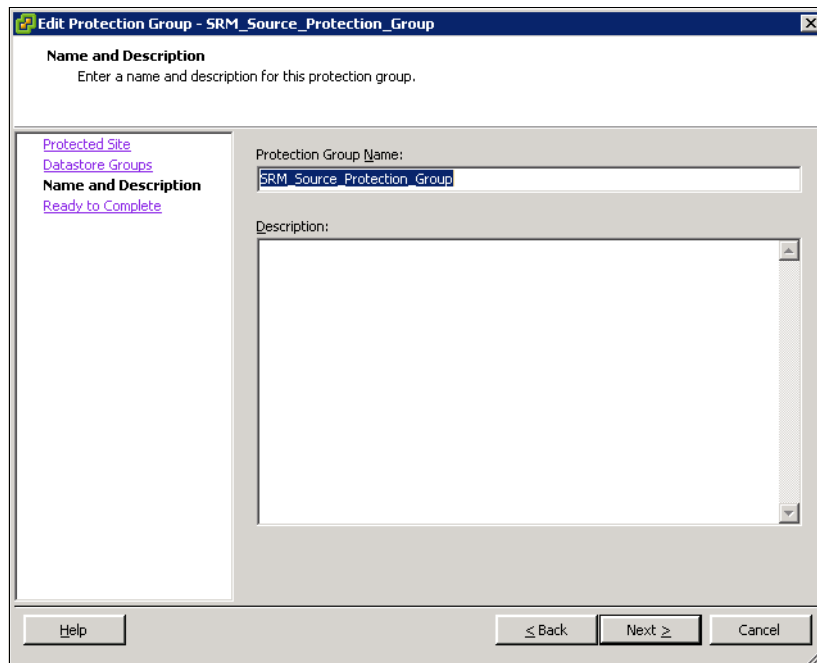


Figure 7-35 Protection Groups: Name Protection Group

- On the Ready to Complete panel, click **Finish**, as shown in Figure 7-36.

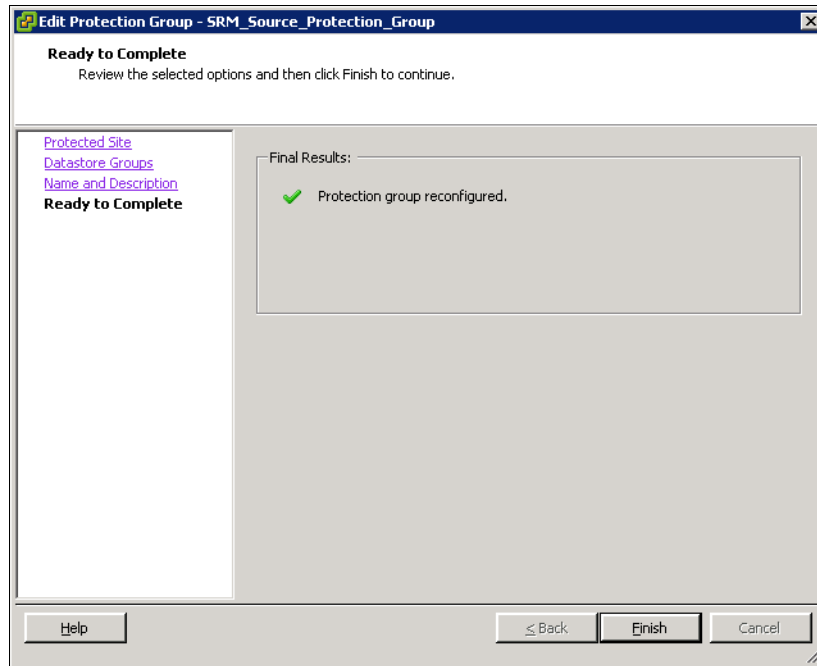


Figure 7-36 Protection Groups: Configuration Successful

- Select the newly configured protection group from the left-tree view of the Protection Groups panel, highlight the **Virtual Machines** tab, and validate that all virtual machines in scope for protection are present and accurately described, as shown in Figure 7-37.

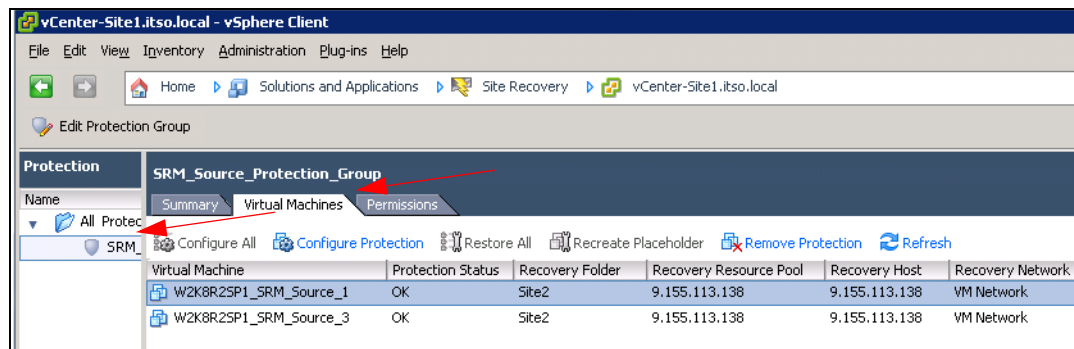


Figure 7-37 Protection Groups: Virtual Machines and Associated Resources

Note: Define protection groups consisting of associated virtual machines that must be recovered together, for example Infrastructure (Active Directory or DNS), Mission Critical, Business Critical, and so on.

Creating recovery plans

Recovery plans govern how virtual machines in one or more protection groups are restored at the recovery site. The steps that follow demonstrate how to customize the plan to meet specific needs.

1. Access the SRM Getting Started with Recovery Plans panel by clicking **Recovery Plans** in the lower-left menu of the parent SRM GUI. Next, click the **Create Recovery Plan** hypertext under step one, or alternatively click **Create Recovery Plan** in the upper left of the window, as shown in Figure 7-38.

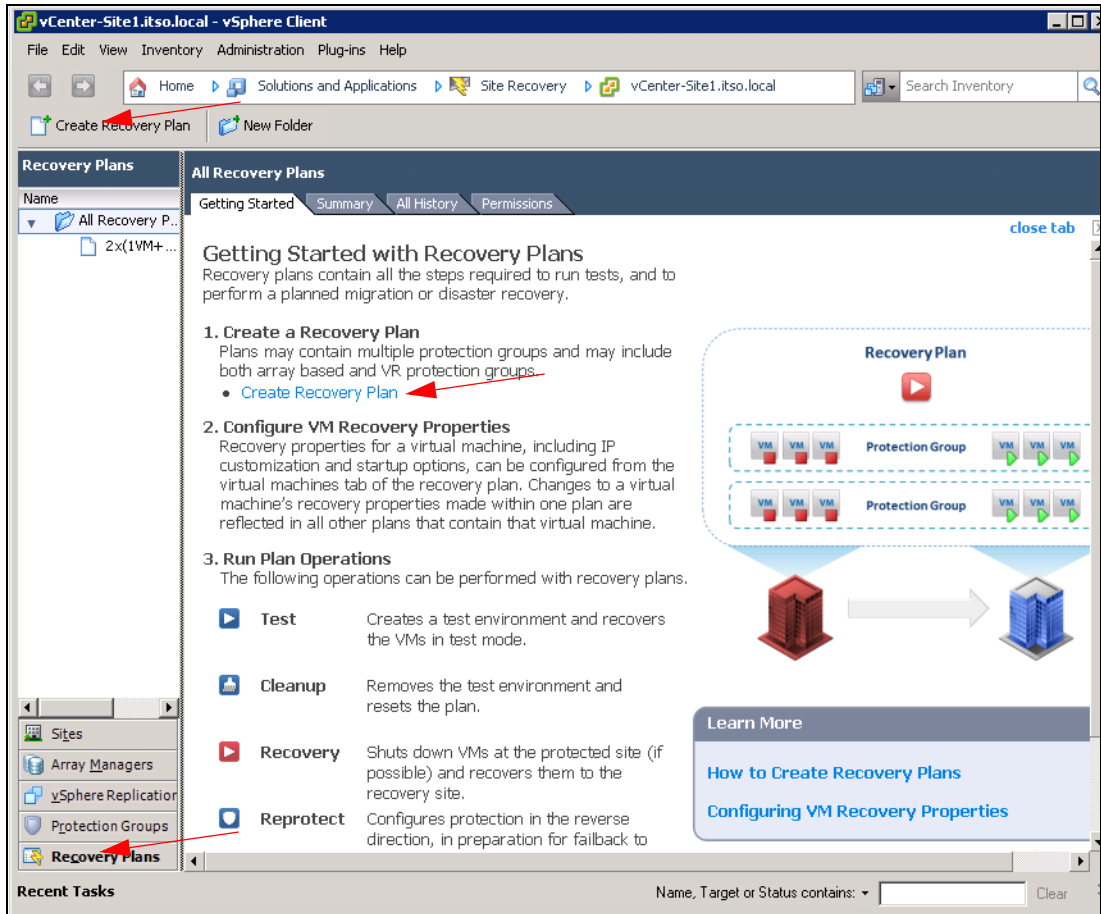


Figure 7-38 Recovery Plans: Create Recovery Plan

2. In the resulting Create Recovery Plan wizard shown in Figure 7-39 on page 145, specify which site will act as the recovery site, and click **Next**.

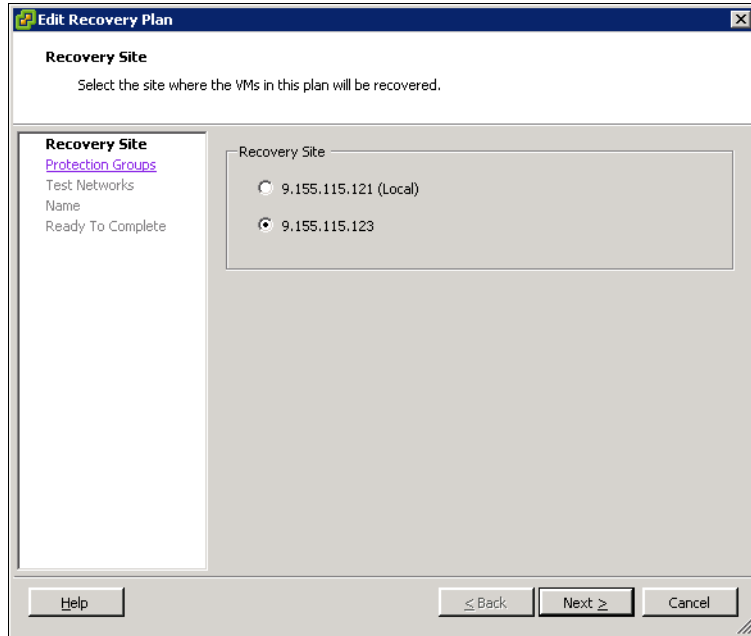


Figure 7-39 Recovery Plans: Select Recovery Site

- Place a check mark next to the desired protection groups that were previously defined in the Protection Groups panel of the Create Recovery Plan wizard, as illustrated in Figure 7-40. Click **Next** to proceed.

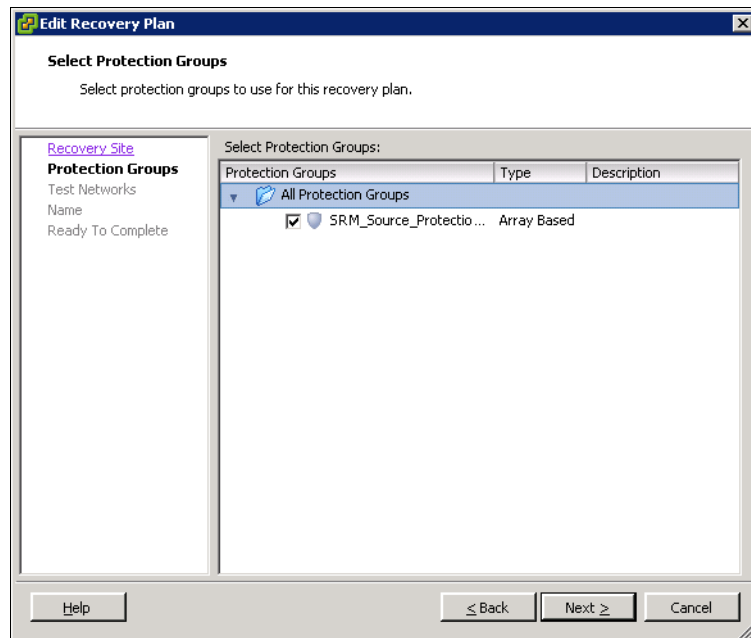


Figure 7-40 Recovery Plans: Select Protection Group(s)

- The Test Networks panel of the wizard, shown in Figure 7-41, allows the administrator to specify which network to use when testing the recovery plan. Specifying “Auto” grants SRM the capability to automatically create an isolated network environment when performing tests as a precaution against impacting the production network. After specifying SRM’s networking behavior during recovery plan tests, click **Next** to proceed.

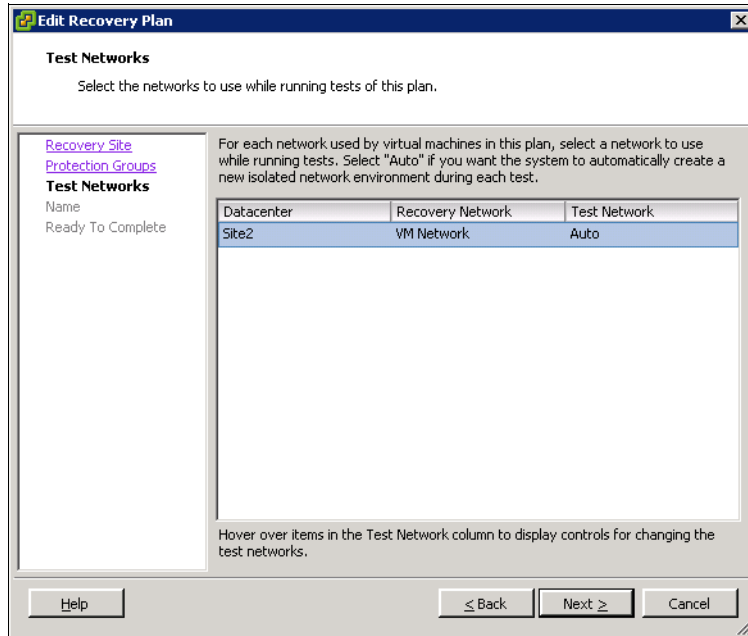


Figure 7-41 Recovery Plans: Specify Test Networks

- As shown in Figure 7-42, fill in the Name and Description panel of the Create Recovery Plan wizard with information that will uniquely identify the recovery plan. Click **Next** to proceed.

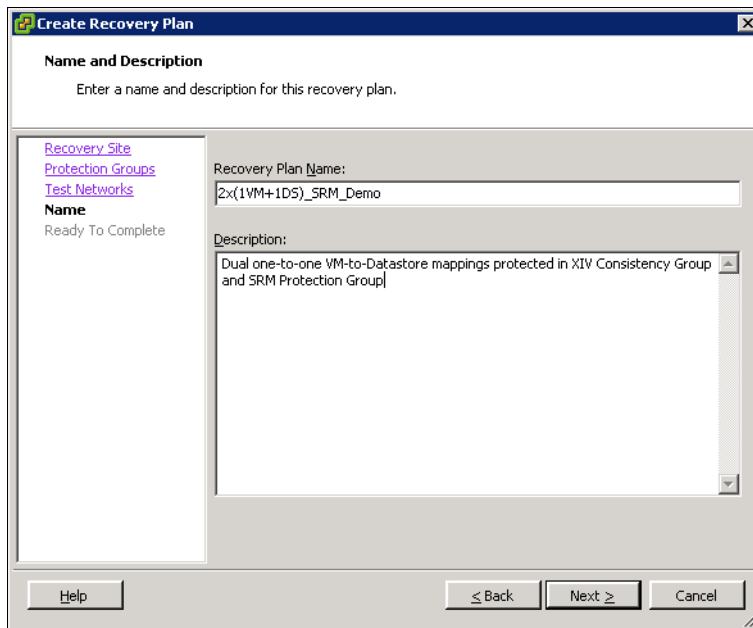


Figure 7-42 Recovery Plans: Name the Plan

- Review the recovery plan parameters summarized on the Ready to Complete panel of the Create Recovery Plan wizard that is shown in Figure 7-43 and then click **Finish** if all settings appear to be correct. Click **Back** if settings need modification.

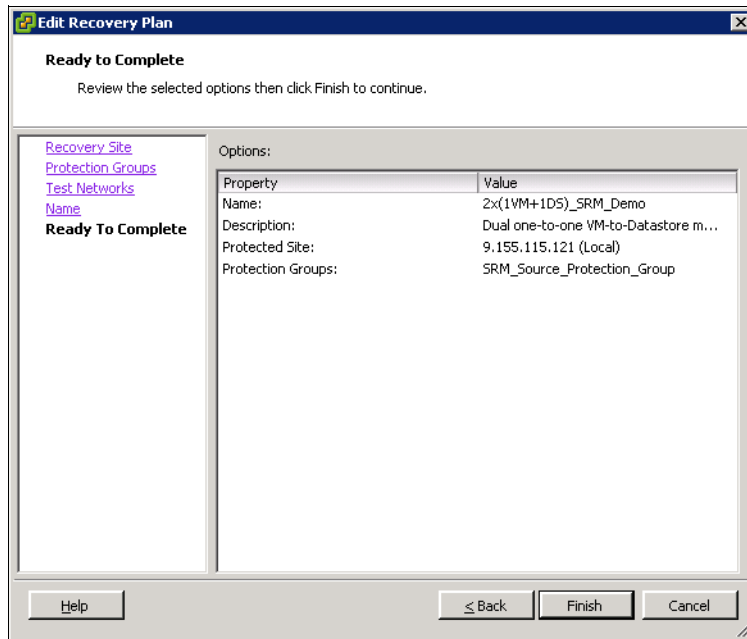


Figure 7-43 Recovery Plans: Validate and Complete

- The name of the new recovery plan now appears in the left-tree view of the main SRM recovery plan GUI panel, as does a series of tabs representing various elements and attributes of the recovery plan. After the name of the recovery plan is selected, click the **Summary** tab shown in Figure 7-44 to review the basic attributes of the plan that were specified in the previous steps.

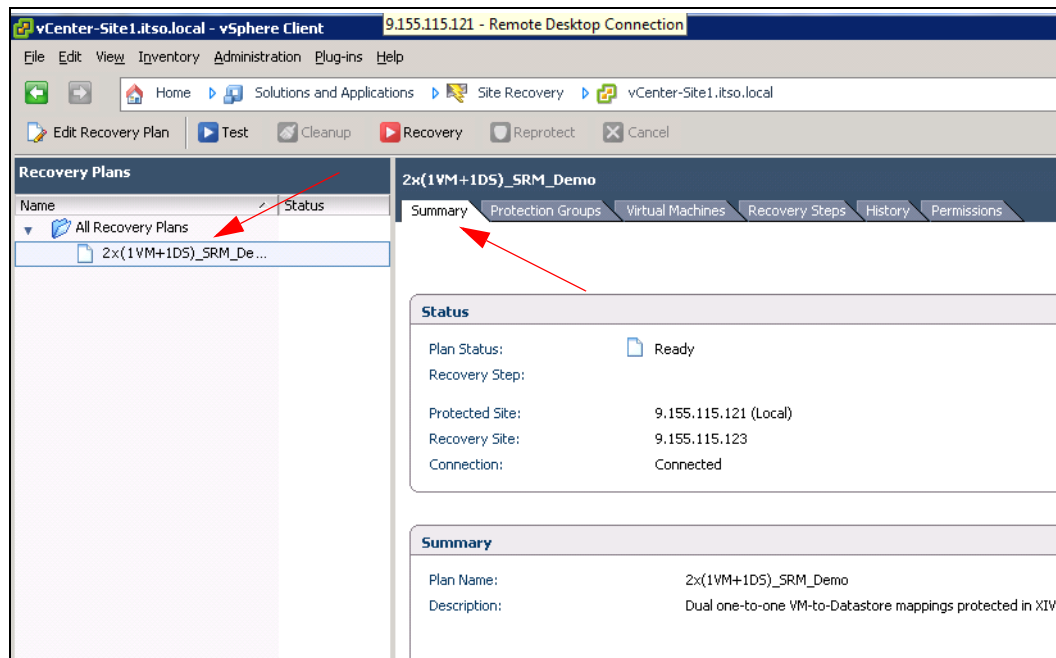


Figure 7-44 Recovery Plans: Review Summary

- Clicking the name of the newly created recovery plan in the left-tree view of the SRM Recovery Plans panel and highlighting the **Protection Groups** tab provides a view of the protection group(s) in scope for the recovery plan, as illustrated in Figure 7-45.

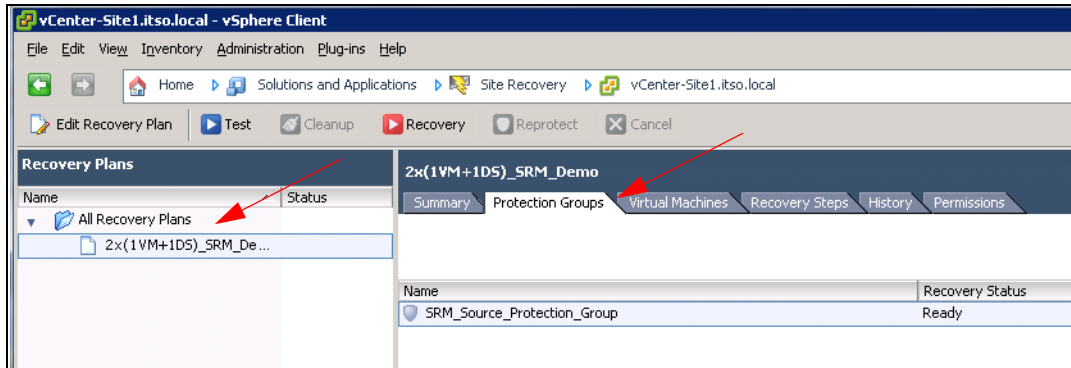


Figure 7-45 Recovery Plans: Review Protection Group(s)

- By highlighting the **Virtual Machines** tab, administrators can review the recovery plan attributes associated with the virtual machines and modify any necessary parameters by using the Recovery Properties menus that are available upon clicking the **Configure Recovery** hypertext. Note the attributes and their values prior to configuring these properties, as shown in Figure 7-46. Review the SRM documentation for additional guidance on the topic of configuring the recovery properties of virtual machines.

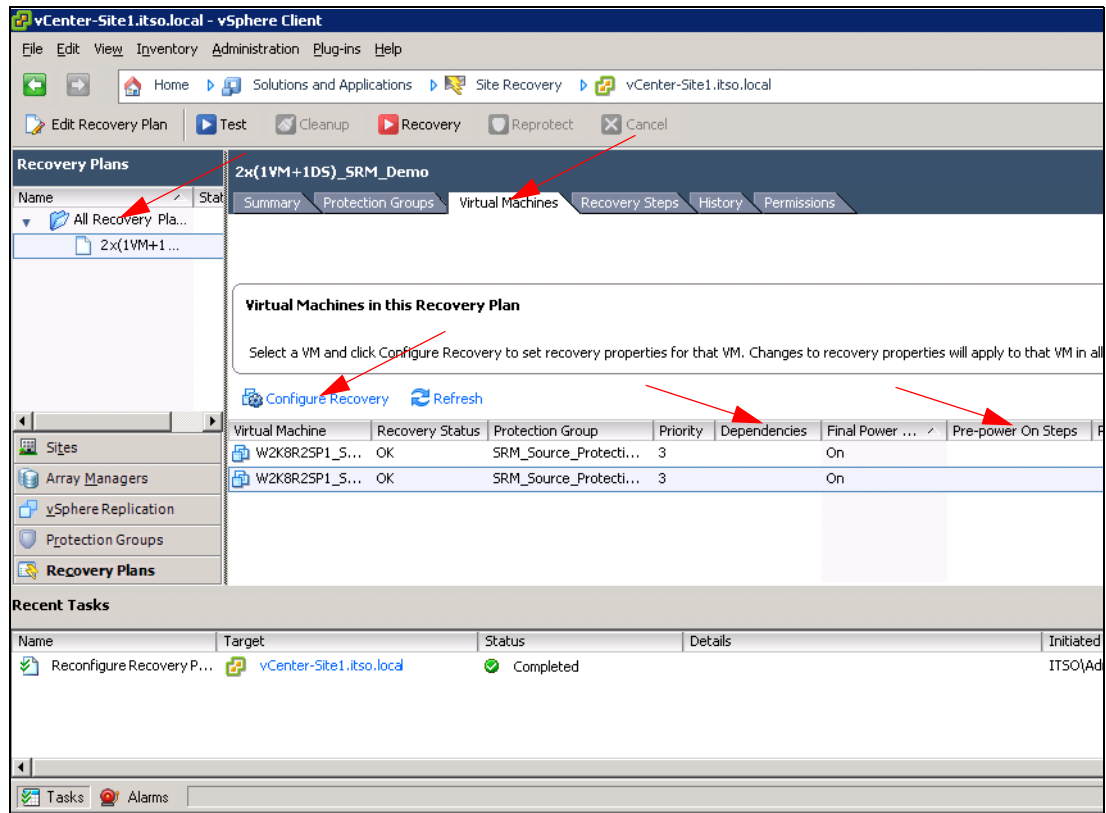


Figure 7-46 Recovery Plans: Configure Virtual Machines' Recovery Properties

Note: Identifying finite values of time or specific numbers of heartbeats to wait for virtual machines to respond after their power-on process is complete is recommended.

10. After the virtual machines' recovery properties are set as appropriate for the intended recovery plan, review the values that are subsequently displayed under the Virtual Machines tab illustrated in Figure 7-47.

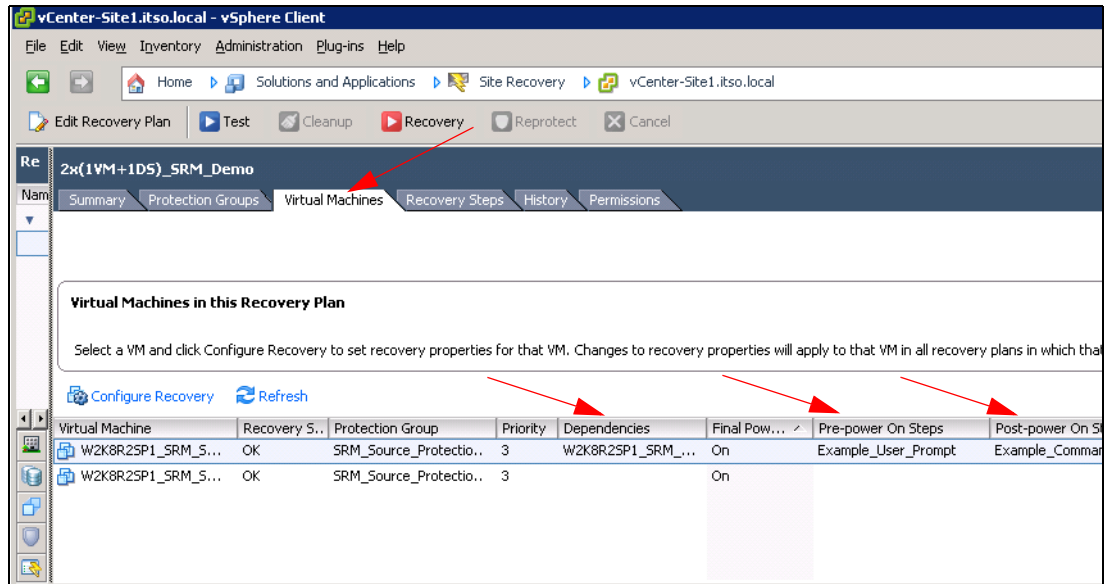


Figure 7-47 Recovery Plans: Review Virtual Machines' Recovery Properties

11. By default, every recovery plan includes a sequence of steps, also known as a workflow, which consists of pre-set values and directives to control how virtual machines in a protection group are recovered at the recovery site. These workflows can be uniquely defined for each of the activity categories consists of Test Steps, Cleanup Steps, Recovery Steps, and Reprotect Steps. A default workflow sequence representing the Test Steps is illustrated in Figure 7-48 within the highlighted Recovery Steps tab. For detailed information about developing recovery plans by specifying customized recovery steps, consult the SRM documentation.

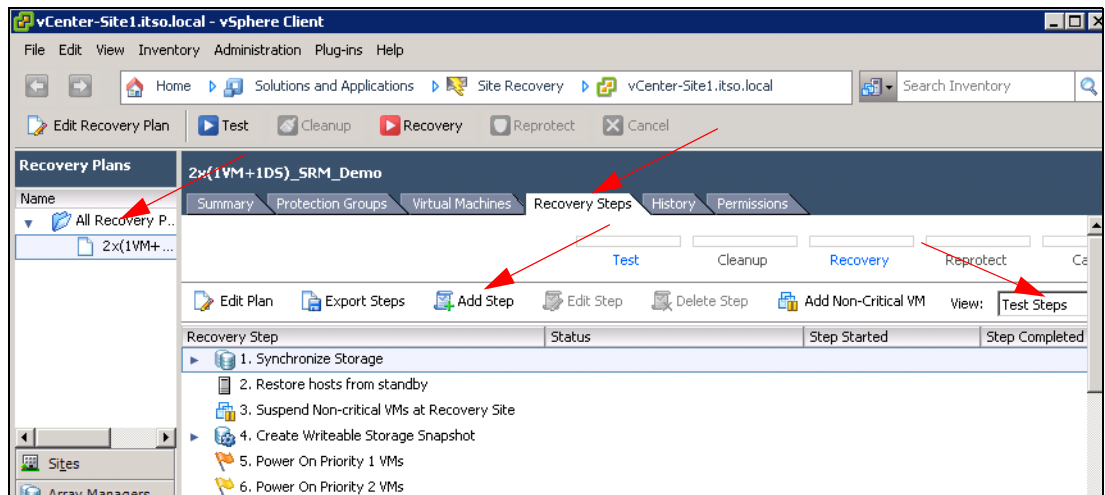


Figure 7-48 Recovery Plans: Review Recovery Steps

Note: Consider these basic guidelines when defining recovery plan workflows:

- ▶ Configure the VM dependencies across priority groups instead of setting VM dependencies individually per VM.

This assures that VMs are started in parallel. The XIV Storage System is optimized for parallel workloads, so this greatly improves performance.

- ▶ Define action prior to a test or failover as the recovery site, such as cloning down or suspending low-priority virtual machines to free recovery resources at the recovery site.
- ▶ Define the allocation of resources and any networking changes required by the virtual machines.
- ▶ Build call-outs that cause test or failover process to pause and present instructions to the administrator, or specify scripts in the recovery process.

Testing recovery plans

SRM's recovery plan testing capability represents an invaluable asset in the development of robust recovery strategies because it empowers administrators to eliminate uncertainty and dramatically reduce risk of failed recovery by identifying and addressing any issues pro-actively without interrupting operation of the production environment. The two primary phases of this process include the test itself and the subsequent removal of temporary elements created during the test and restore the recovery plan to its initial state using the cleanup process. Both of these procedures are illustrated in the next series of steps:

1. After selecting the recovery plan to be tested from the left-tree view of the SRM Recovery Plans GUI panel, click **Test**, as shown in Figure 7-49.

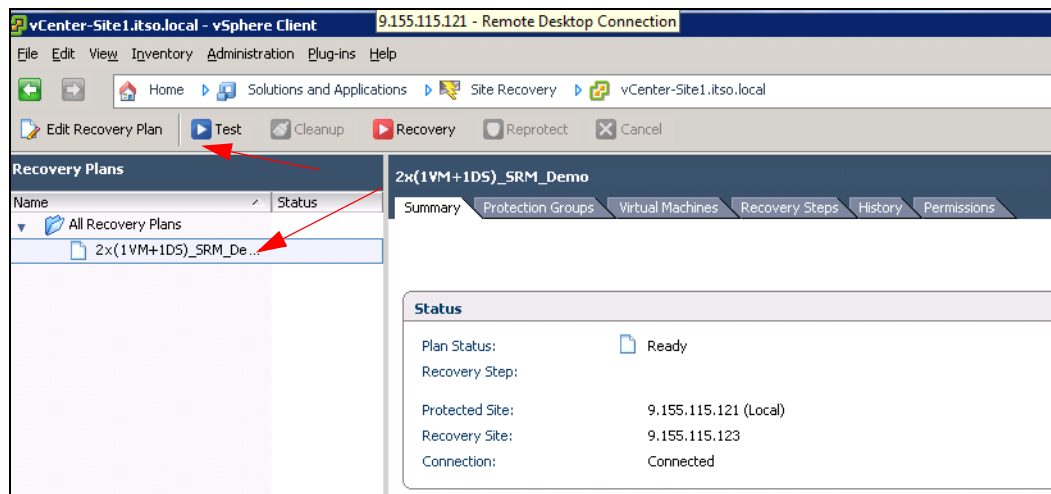


Figure 7-49 Testing Recovery Plans: Initiate Test

- As illustrated in Figure 7-50, the resulting Test Confirmation window contains an optional check box that can be activated to ensure that the recovery site has the most recent copy of the protected virtual machines. It is important to note that the synchronization can take additional time as a result of specifying this option. Click **Next** to proceed.

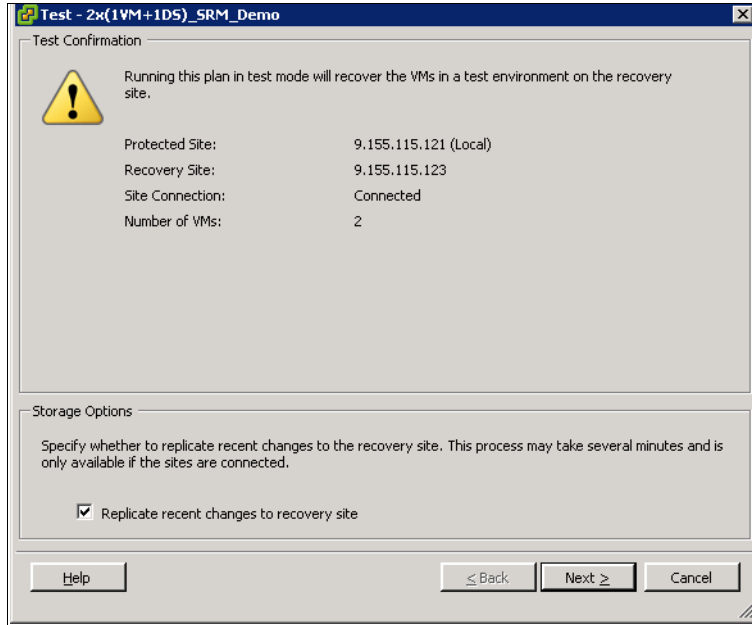


Figure 7-50 Testing Recovery Plans: Confirm Test and Specify Storage Options

- After testing specifications, similar to those illustrated in Figure 7-51, are validated, click **Start** to execute the test of the recovery plan.

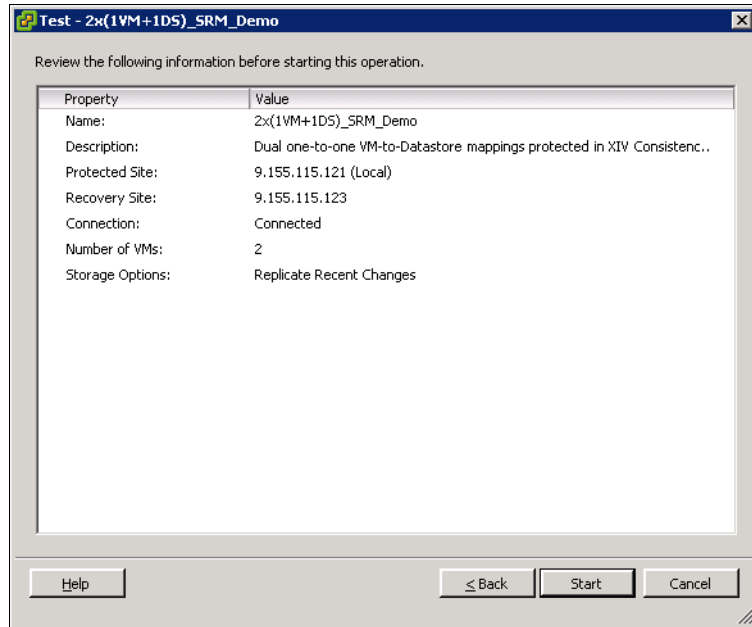


Figure 7-51 Testing Recovery Plans: Review Testing Parameters

Cleanup

The cleanup process is necessary to restore the recovery plan to its initial state after a test is performed:

1. Following test completion, administrators receive a notification of the test completion status, and the option to click the **Cleanup** button within the toolbar of the SRM Recovery Plans GUI panel that is illustrated in Figure 7-52 becomes available.

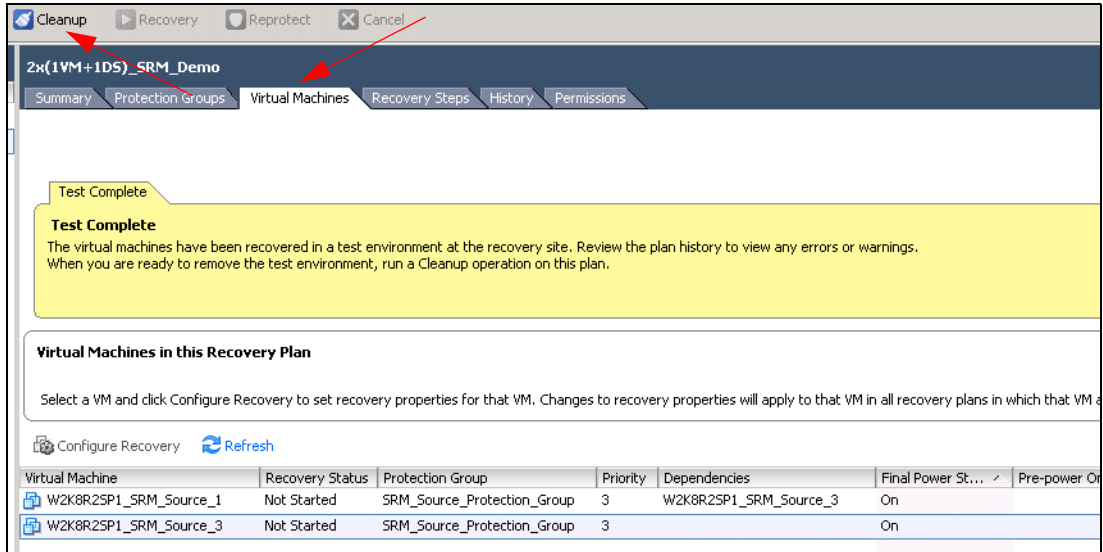


Figure 7-52 SRM Test Environment Cleanup Initialization

2. As illustrated in Figure 7-53, review the important information regarding cleanup operations contained in the initial panel of the Cleanup wizard that was invoked in the previous step, noting that the Force Cleanup check box is initially unavailable, and click **Next** to proceed.

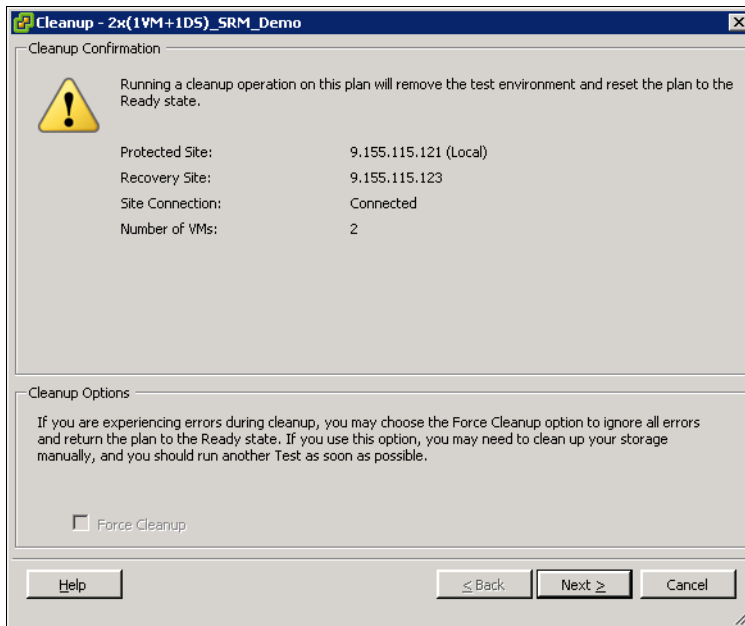


Figure 7-53 SRM Test Environment: Cleanup Confirmation

3. Validate the properties and associated values shown in the wizard panel similar to the example in Figure 7-54, and click **Start**.

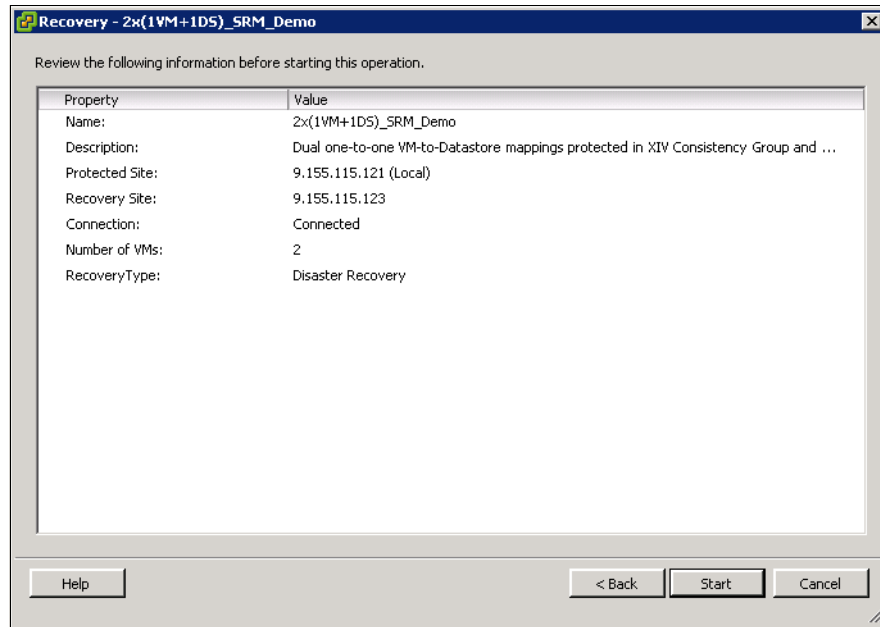


Figure 7-54 SRM Test Environment: Cleanup Review

Note: If the cleanup process fails to successfully complete, revisit the Cleanup wizard to specify the Force Cleanup option and return the recovery plan to the ready state. The caveat of this method consists of the potential requirement for administrator intervention to restore the storage to its initial state.

Recovery

While the workflows outlined in SRM recovery plans are fully automated, vSphere administrators are responsible for initiating a site recovery using the following process:

1. Access the SRM Recovery Plans panel by clicking **Recovery Plans** in the lower-left menu of the parent SRM GUI and then selecting the recovery plan to be invoked from the left-tree view. Click **Recovery** in the toolbar to initiate recovery, as shown in Figure 7-55.

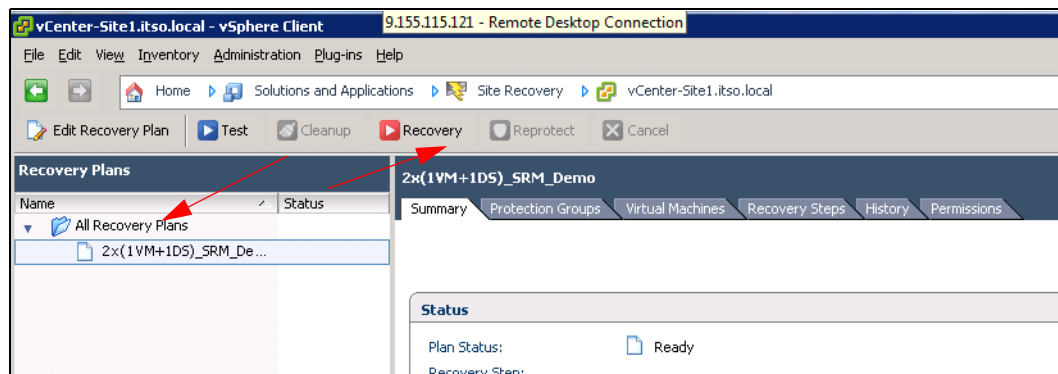


Figure 7-55 SRM Recovery Process: Initiate Recovery

2. As illustrated in Figure 7-56 on page 154, the administrator must confirm that he or she is aware that the scope of impact of initiating a recovery process spans infrastructure at both

the protected and the recovery sites. In addition, the administrator must specify whether the Recovery Type comprises a Planned Migration or Disaster Recovery. For purposes of this demonstration, the Disaster Recovery option is specified. After the appropriate selections are made, click **Next** to proceed.

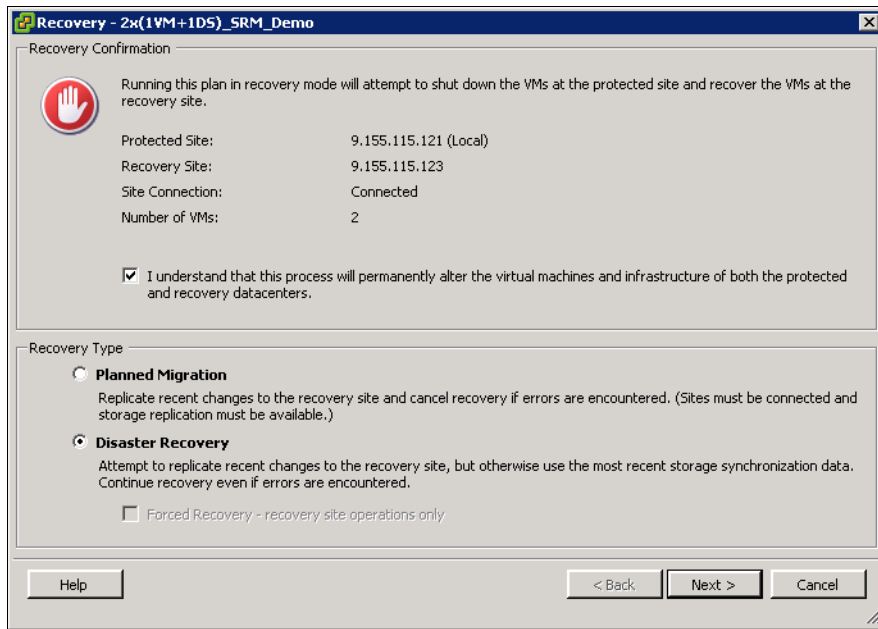


Figure 7-56 SRM Recovery Process: Confirm Recovery and Select Recovery Type

- Review the recovery plan displayed in the recovery wizard panel to ensure that the intended recovery plan will be invoked in the site recovery process, which is similar to the example in Figure 7-57. Click **Start** to begin site recovery.

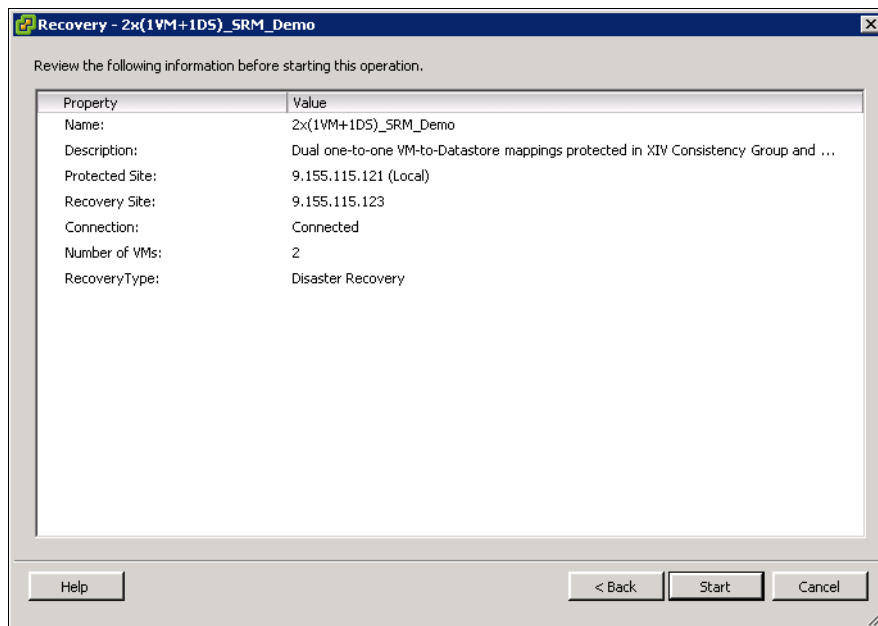


Figure 7-57 SRM Recovery Process: Review Recovery Plan

- When the recovery process successfully completes, click the **History** tab to view the status message illustrated in Figure 7-58 on page 155, which informs the administrator

that the virtual machines within the protection group were relocated to the recovery site, and so on. The history tab also enables the administrator to review the status of all recovery tasks comprising the recovery process workflow.

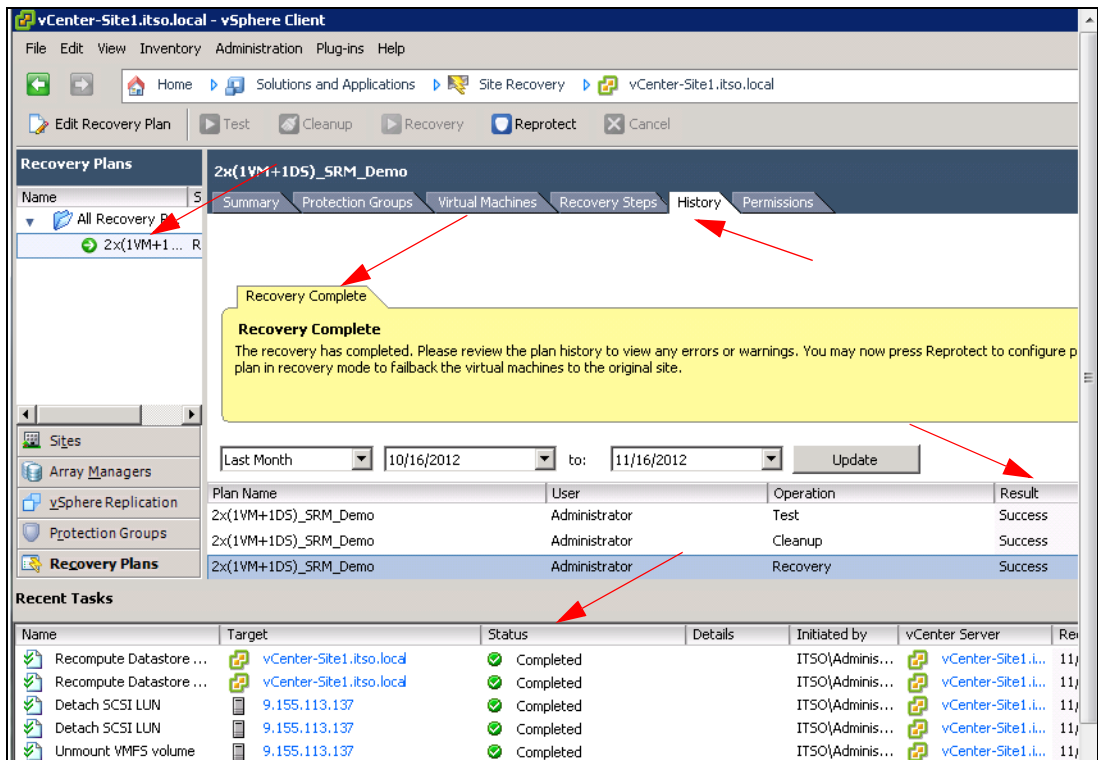


Figure 7-58 SRM Recovery Process: Recovery Complete

5. Figure 7-59 illustrates the status of the recovery steps outlined in the recovery process workflow that can be viewed by clicking the **Recovery Steps** tab following recovery process completion.

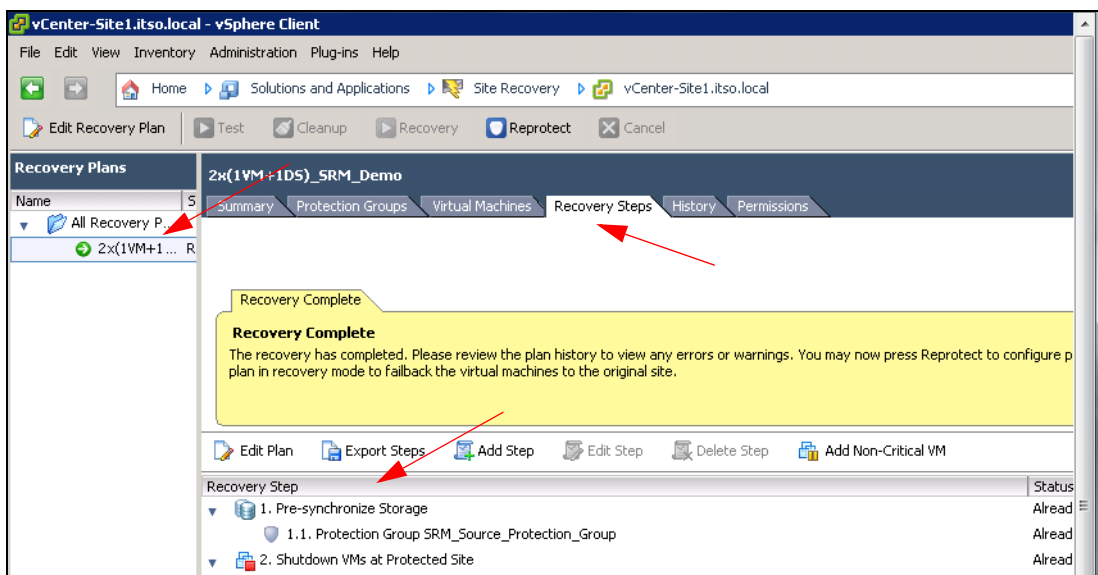


Figure 7-59 SRM Recovery Process: Status of Recovery Steps

- Navigate to the recovery site's vSphere client Hosts and Clusters menus to view the migrated virtual machines, as illustrated in Figure 7-60.

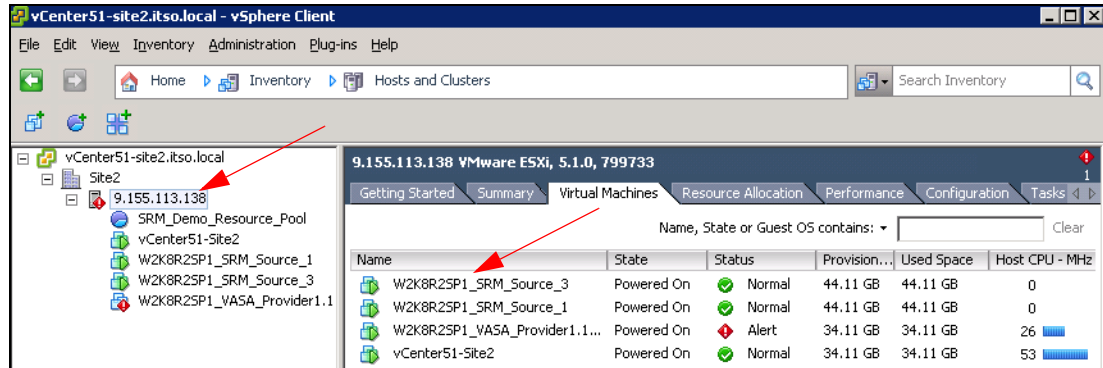


Figure 7-60 SRM Recovery Process: Virtual Machines at Protected Site Restarted at Recovery Site

Reprotect

The reprotect process can only be invoked following the successful completion of a recovery process, and effectively automates the process of reversing the roles of the protected and recovery sites to facilitate the continuance of site-level protection for the vSphere production environment that is running at the recovery site following a planned or unplanned migration. Activating the reprotect process can be completed by employing the upcoming sequence of steps:

- Access the SRM Recovery Plans panel by clicking **Recovery Plans** in the lower-left menu of the parent SRM GUI and then selecting the recovery plan to be initiated in the reprotection process from the left-tree view. Click **Reprotect** in the toolbar to begin the process of site reprotection, as shown in Figure 7-61.

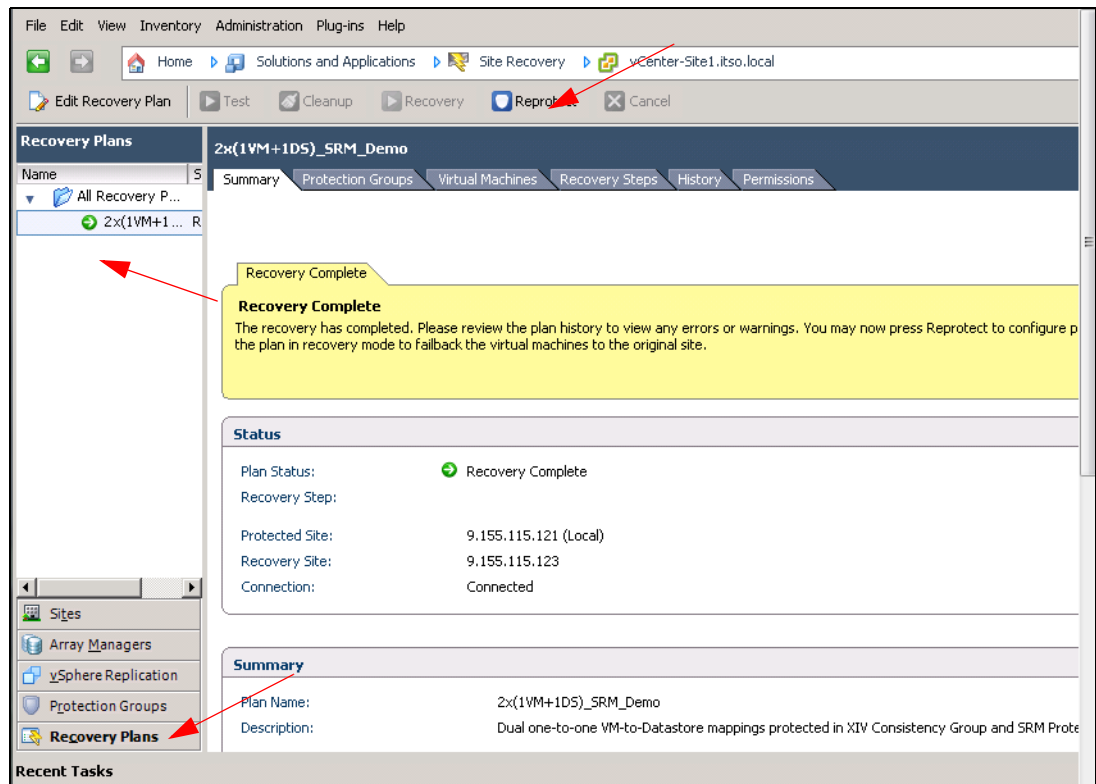


Figure 7-61 SRM Reprotect Process: Initiate Reprotection

- As illustrated in Figure 7-62, review the important information regarding reprotect operations contained in the initial Reprotect panel of the wizard that was invoked in the previous step, noting that the Force Cleanup check box is initially unavailable, and click **Next** to proceed.

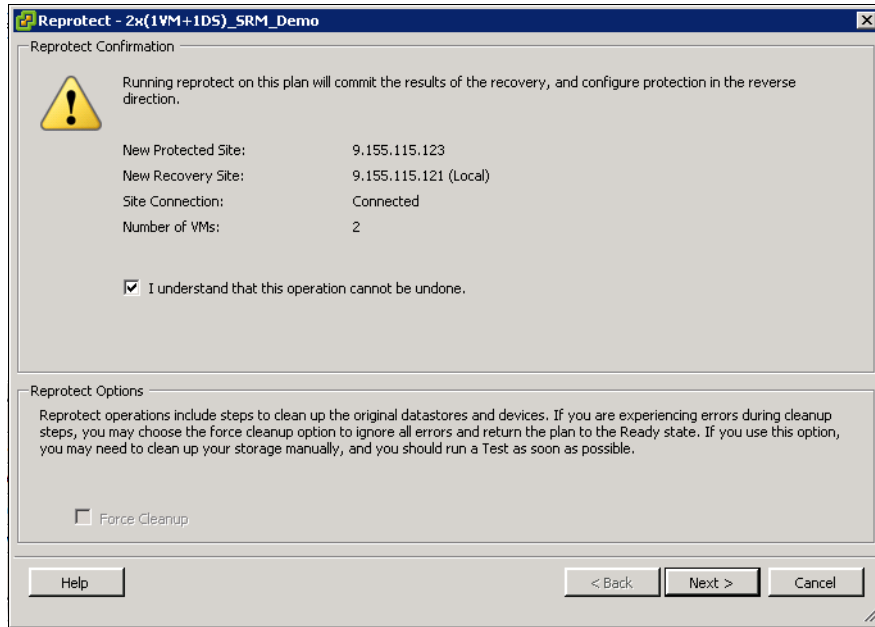


Figure 7-62 SRM Reprotect Process: Confirm Reprotection

- Review the recovery plan displayed in the Reprotect wizard panel to ensure that the intended recovery plan is invoked in the site reprotection process, which is similar to the example in Figure 7-63. Click **Start** to begin site reprotection.

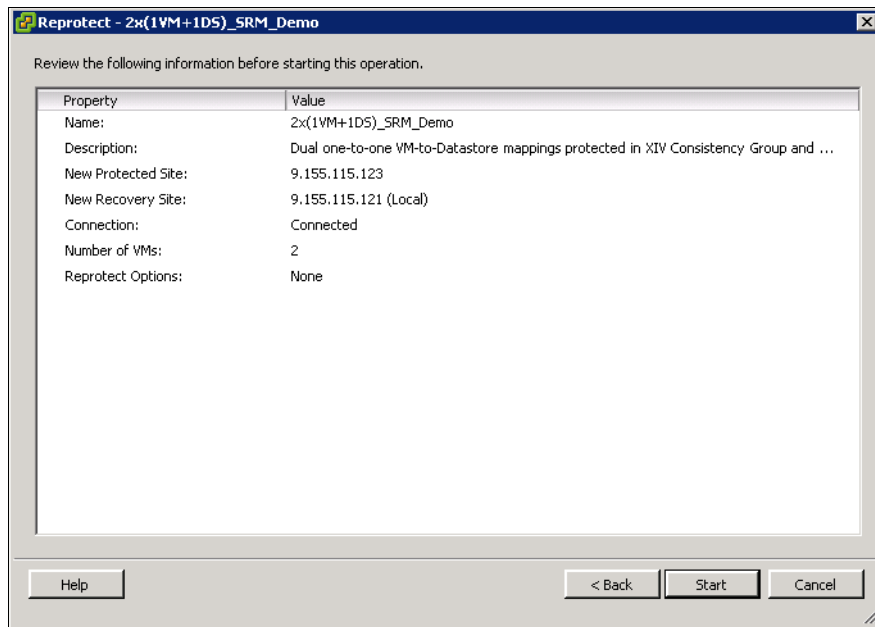


Figure 7-63 SRM Reprotect Process: Review Recovery Plan

- When the reProtection process successfully completes, click the **Summary** tab to view the status message illustrated in Figure 7-64 informing the administrator that the virtual machines within the protection group were reprotected at the new recovery site, which was the protected site prior to failover through the SRM recovery process.

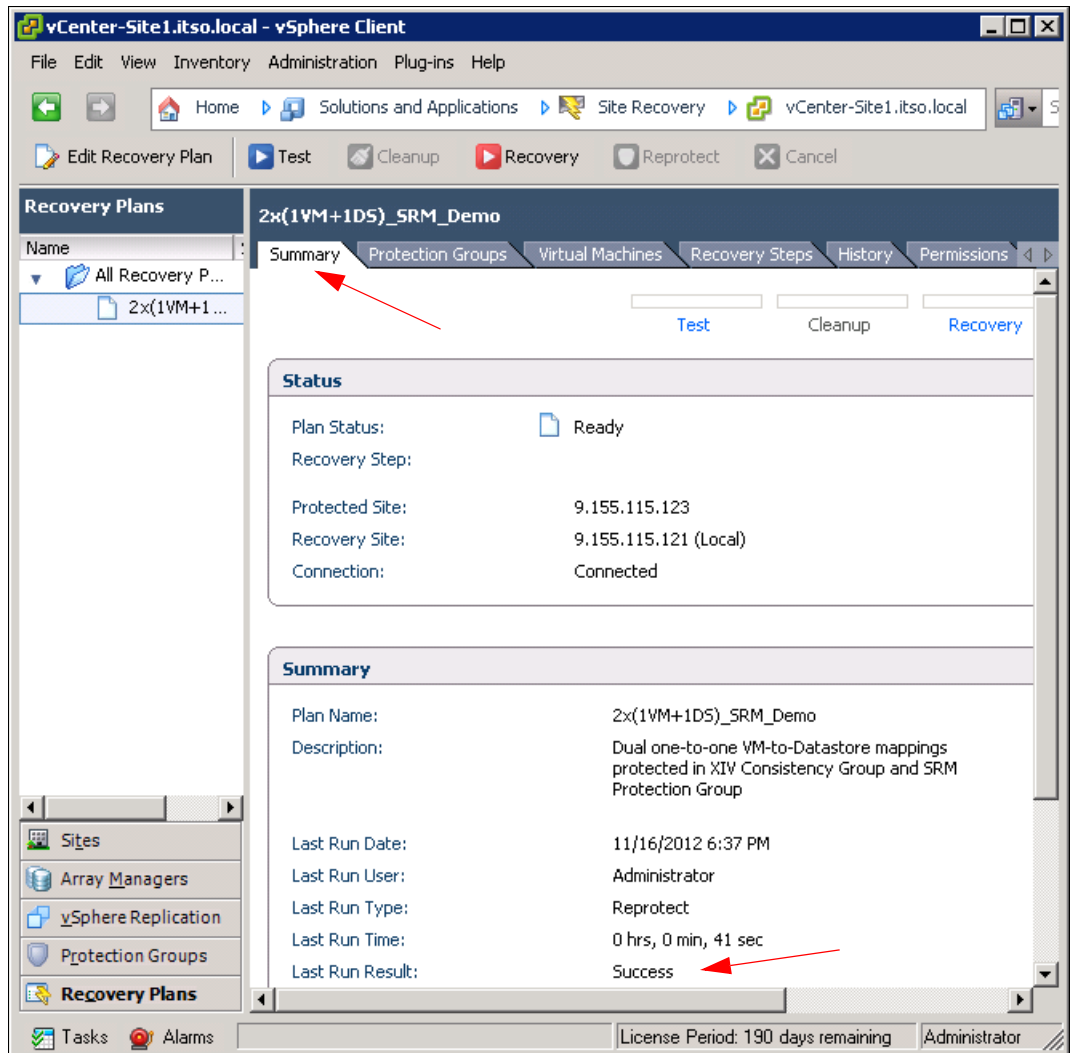


Figure 7-64 SRM Reprotect Process: Reprotection Complete

Figure 7-65 and Figure 7-66 illustrate how the XIV remote mirroring configuration changes and reflects the reversal of roles of the master and subordinate logical volumes backing the data stores in the protection group. In Figure 7-65, the SRM CG Demo consistency group residing on the XIV at the original protected site now contains only subordinate volumes, while Figure 7-66 demonstrates that the original volumes residing in the XIV at the original recovery site have concurrently been defined as remote mirroring master volumes.

Name	RPO	Status	Remote Volume
geocluster_test		Synchronized	XIV_PFE_03_7804143
izik1_1		Consistent	XIV_PFE_03_7804143
ESXi5U1_DS1	00:00:30	RPO Lagging	XIV_02_1310114
SAP_DATA_MIG	00:00:30	RPO OK	XIV_02_1310114
TA_LS_PROD_Volume_2	00:02:30	RPO OK	XIV_02_1310114
TA_LS_PROD_Volume_1	00:02:30	RPO OK	XIV_02_1310114
TA_LS_PROD_Volume_3	00:02:30	RPO OK	XIV_02_1310114
SRM_CG_Demo		Consistent	XIV_02_1310114
SRM_Source_3		Consistent	XIV_02_1310114
SRM_Source_2		Consistent	XIV_02_1310114
SRM_Source_1		Consistent	XIV_02_1310114
izikcg	00:00:30	RPO OK	XIV_PFE_03_7804143
TA_LS_PROD_CG	00:02:30	Inactive	XIV_02_1310114

Figure 7-65 SRM Reprotect Process: XIV Mirrored LUNs Role Reversal - Protected Site Masters Becomes Recovery Site Slaves

Name	RPO	Status	Remote Volume
SRM_CG_Demo		Synchronized	XIV_01_6000105
SRM_Target_3		Synchronized	XIV_01_6000105
SRM_Target_1		Synchronized	XIV_01_6000105
SRM_Target_2		Synchronized	XIV_01_6000105
TA_LS_DR_CG	0...	Inactive	XIV_01_6000105

Figure 7-66 SRM Reprotect Process: XIV Mirrored LUNs Role Reversal - Recovery Site Slaves Becomes Protected Site Masters

7.2 Quick install guide for VMware Site Recovery Manager

This section addresses VMware Site Recovery Manager-specific installation considerations, including information related to XIV configurations. The goal is only to give you enough information to quickly install, configure, and experiment with Site Recovery Manager (SRM). It is not meant as a guide on how to deploy SRM in a real production environment

For more information about the concepts, installation, configuration, and usage of VMware Site Recovery Manager, see the VMware product site at:

http://www.vmware.com/support/pubs/srm_pubs.html

At the time of this writing, versions 1.0, 1.0 U1, 4.X, 5.0, and 5.0.1 of Storage Replication Agent for VMware SRM server are supported with XIV Storage Systems.

Tip: The *VMware vCenter Site Recovery Manager version 5.x Guidelines for IBM XIV Gen3 Storage System* provides a valuable supplemental resource addressing topics including concepts, planning, testing, and best practices necessary to fully implement vSphere SRM integration with XIV remote replication technology:

[https://www.ibm.com/partnerworld/wps/servlet/download/DownloadServlet?id=VZPYFkT7gvZiPCA\\$cnt&attachmentName=VMware_vCenter_Site_Recovery_Manager_version_guide_lines_IBM_XIV_Storage.pdf&token=MTM1MTU5NTQwNzA2NA==&locale=en_ALL_ZZ](https://www.ibm.com/partnerworld/wps/servlet/download/DownloadServlet?id=VZPYFkT7gvZiPCA$cnt&attachmentName=VMware_vCenter_Site_Recovery_Manager_version_guide_lines_IBM_XIV_Storage.pdf&token=MTM1MTU5NTQwNzA2NA==&locale=en_ALL_ZZ)

The SRM server needs to have its own database for storing recovery plans, inventory information, and similar data. SRM supports the following databases:

- ▶ IBM DB2®
- ▶ Microsoft SQL
- ▶ Oracle

The SRM server has a set of requirements for the database implementation. Some of these requirements are general and do not depend on the type of database used, but others are not. For more information about specific database requirements, see the VMware SRM documentation.

The SRM server database can be on the same server as vCenter, on the SRM server host, or on a separate host. The location depends on the architecture of your IT landscape and on the database that is used.

Information about compatibility for SRM server versions are at the following locations:

- ▶ Version 5.X:
<https://www.vmware.com/support/srm/srm-compat-matrix-5-0.html>
- ▶ Version 4.X:
http://www.vmware.com/pdf/srm_compat_matrix_4_x.pdf
- ▶ Version 1.0 update 1:
http://www.vmware.com/pdf/srm_101_compat_matrix.pdf
- ▶ Version 1.0:
http://www.vmware.com/pdf/srm_10_compat_matrix.pdf

7.2.1 Installing and configuring the database environment

This section illustrates the step-by-step installation and configuration of the database environment for the VMware vCenter and SRM server. The example uses Microsoft SQL Server 2005 Express and Microsoft SQL Server Management Studio Express as the database environment for the SRM server. Microsoft SQL Express database is installed on the same host server as vCenter.

The Microsoft SQL Express database is available at no additional cost for testing and development purposes. It is available for download from the Microsoft website at:

<http://www.microsoft.com/downloads/en/details.aspx?FamilyID=3181842A-4090-4431-ACD-9A1C832E65A6&displaylang=en>

The graphical user interface for the database can be downloaded for at no additional cost from the Microsoft website at:

<http://www.microsoft.com/downloads/details.aspx?FamilyId=C243A5AE-4BD1-4E3D-94B8-5A0F62BF7796&DisplayLang=en>

Further Information: For specific requirements and details about installing and configuring the database application, see the database vendor and VMware documentation for SRM.

Installing the Microsoft SQL Express database

After you download the Microsoft SQL Express software, start the installation process:

1. Double-click **SQLEXPR.EXE** in Windows Explorer, as shown in Figure 7-67. The installation wizard starts.

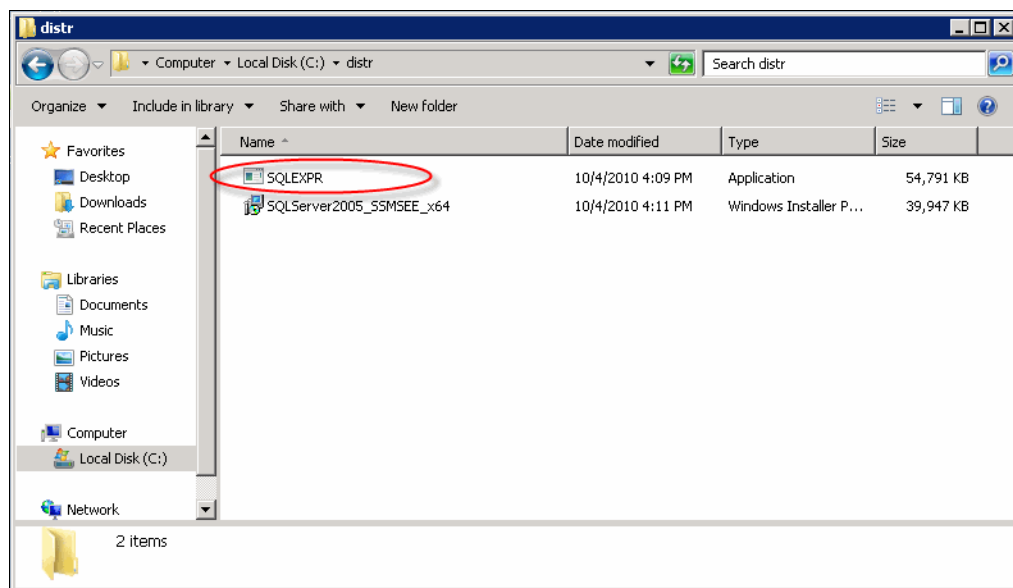


Figure 7-67 Starting Microsoft SQL Express installation

2. Proceed through the prompts until you reach the Feature Selection window shown in Figure 7-68 on page 162. Select **Connectivity Components** for installation. Click **Next**.

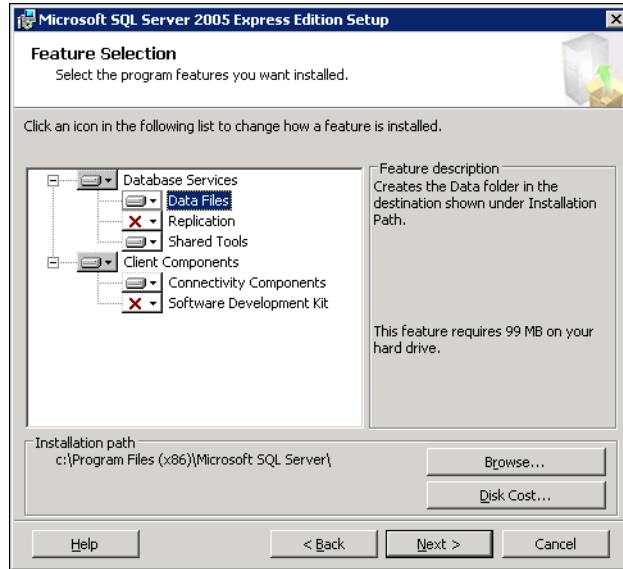


Figure 7-68 List of components for installation

3. The Instance Name window is displayed, as shown in Figure 7-69. Select **Named instance** and type `SQLExpress`, and click **Next**. This name is also used for SRM server installation.

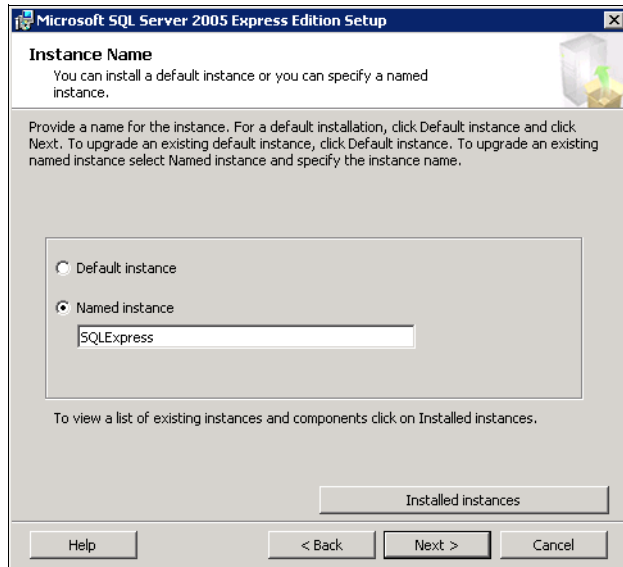


Figure 7-69 Instance naming

4. The Authentication Mode window is displayed, as shown in Figure 7-70 on page 163. Select **Windows Authentication Mode** for a simple environment. Depending on your environment and needs, you might need to choose another option. Click **Next**. The Configuration Options window is displayed as shown in Figure 7-71 on page 163.

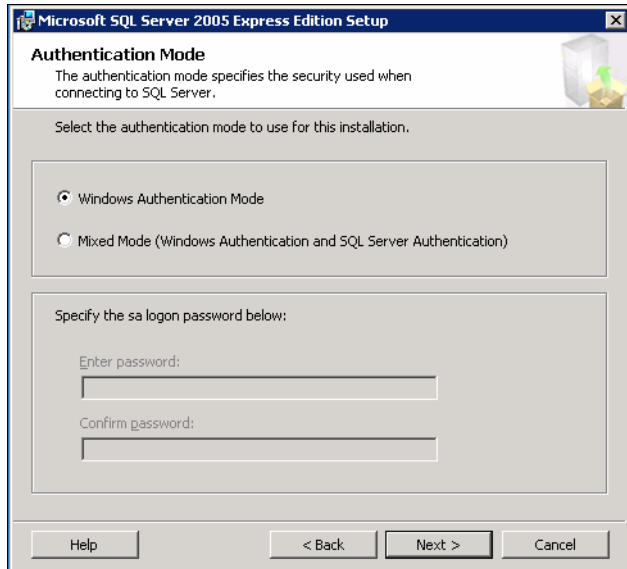


Figure 7-70 Selecting the type of authentication

5. Select **Enable User Instances** and click **Next**. The Error and Usage Report Settings window is displayed as shown in Figure 7-72 on page 164.

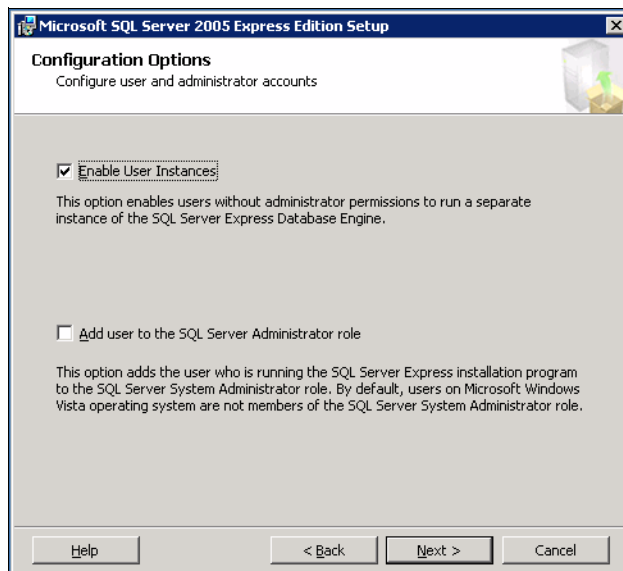


Figure 7-71 Selecting configuration options

6. Select whether you want to report errors to Microsoft Corporation, and click **Next**.

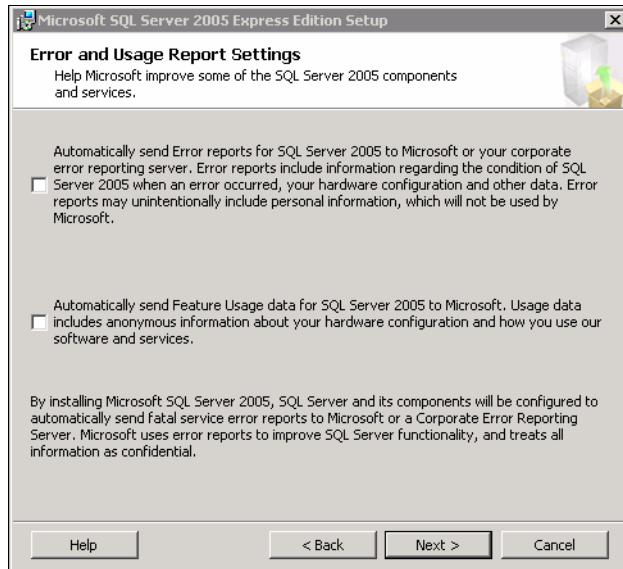


Figure 7-72 Configuration on error reporting

7. The Ready to Install dialog window is displayed, as shown in Figure 7-73. Start the MS SQL Express 2005 installation process by clicking **Install**. If you decide to change previous settings, you can go backwards by clicking **Back**.

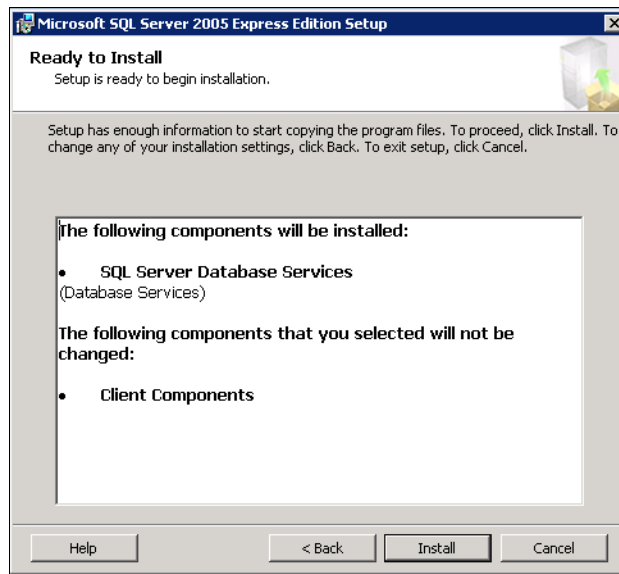


Figure 7-73 Ready to install

8. After the installation process is complete, the dialog window shown in Figure 7-74 on page 165 is displayed. Click **Next** to complete the installation procedure.

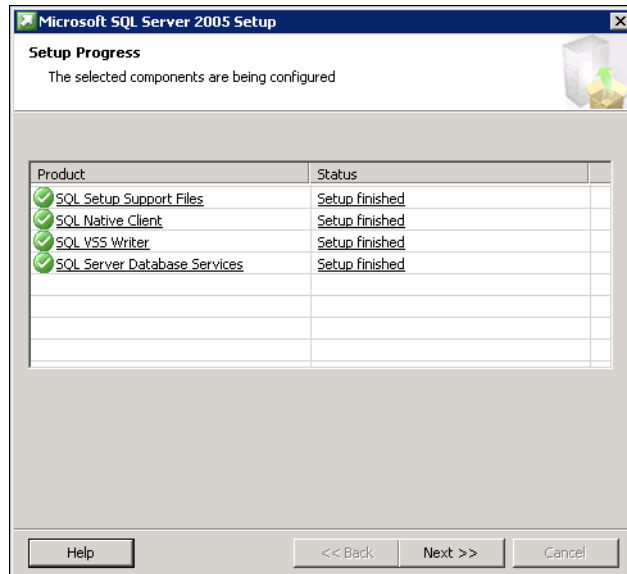


Figure 7-74 Install finished

- The final dialog window displays the results of the installation process, as shown in Figure 7-75. Click **Finish** to complete the process.

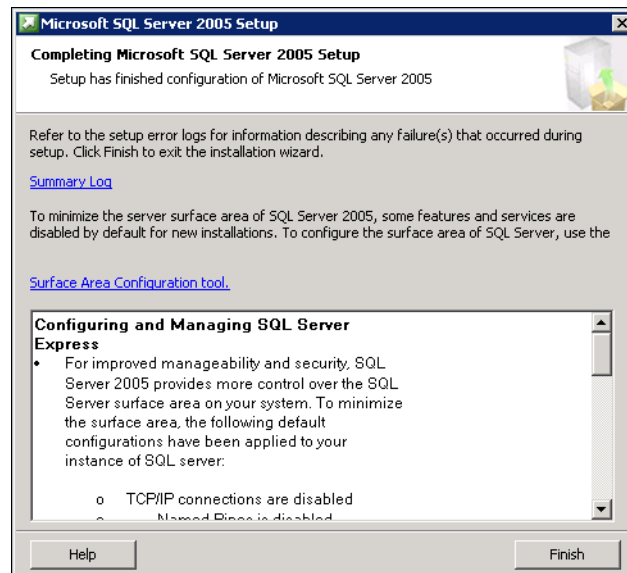


Figure 7-75 Install completion

Installing SQL Server Management Studio Express

To install the visual tools to configure the database environment:

- Download the SQL Server Management Studio Express installation files from the Microsoft web site.
- Start the installation process by double-clicking **SQLServer2005_SSMSEE_x64.msi** in Windows Explorer, as shown in Figure 7-76 on page 166.

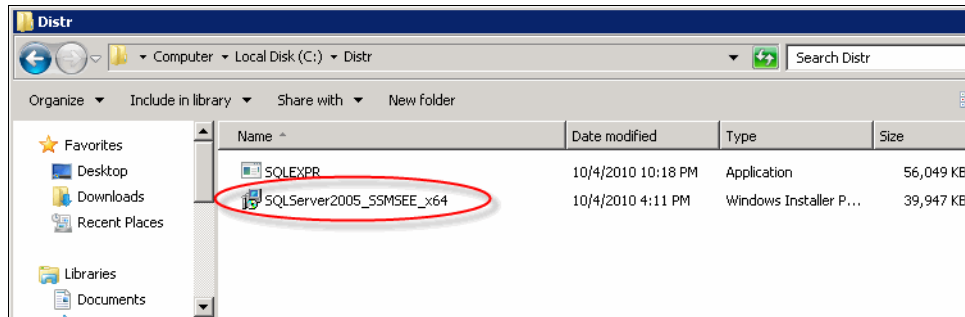


Figure 7-76 Starting installation for Microsoft SQL Server Management Studio Express

After clicking the file, the installation wizard starts. Proceed with the required steps to complete the installation.

The Microsoft SQL Server Management Studio Express software must be installed at all locations involved in your continuity and disaster recovery solution.

3. Create additional local users on your host by clicking **Start** and selecting **Administrative Tools** → **Computer Management**, as shown in Figure 7-77.

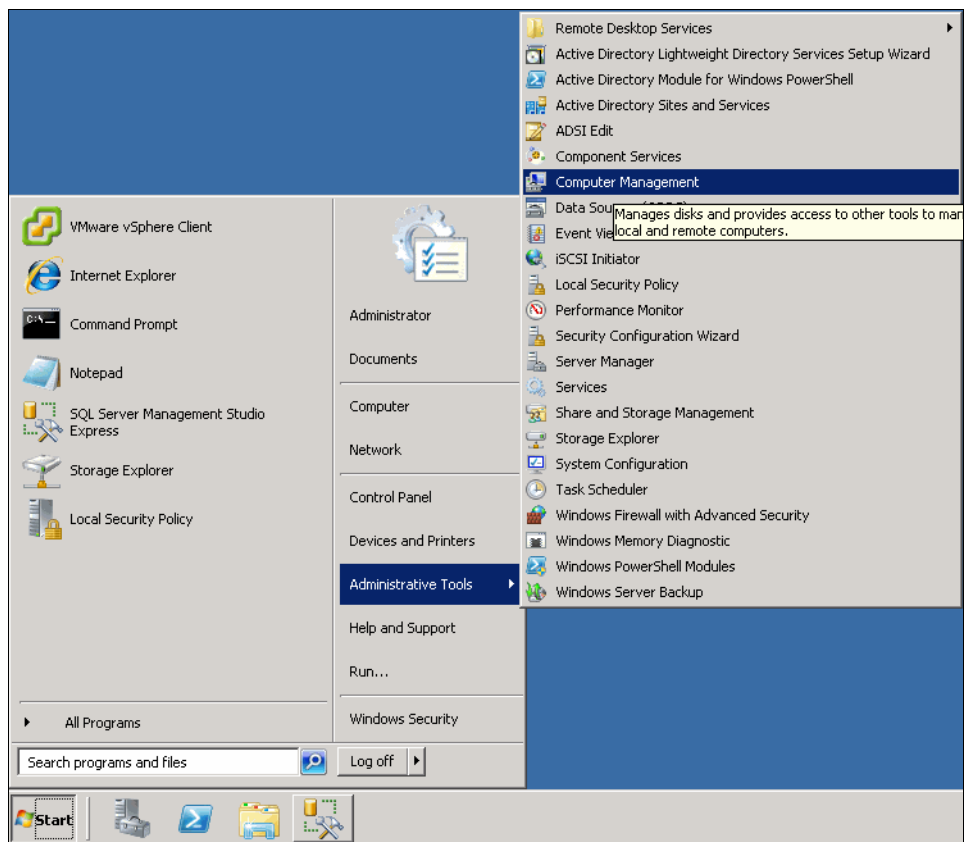


Figure 7-77 Running Computer Management

4. Navigate to the subfolder **Computer Management (Local)\System Tools\Local Users and Groups** and then right-click **Users**. Click **New User**.

5. The New User window is displayed, as shown in Figure 7-78. Enter details for the new user, and click **Create**. You need to add two users: One for the vCenter database and one for the SRM database.

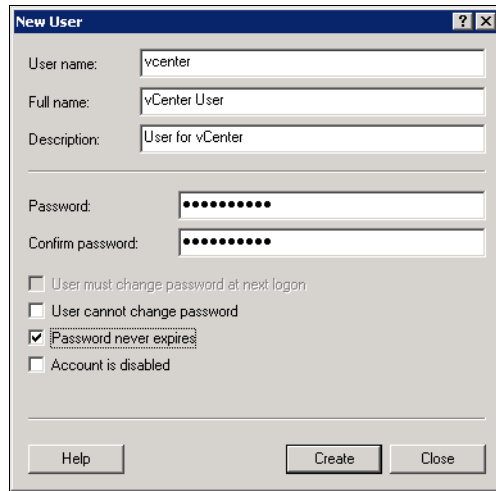


Figure 7-78 Adding a user

6. Configure one vCenter database and one SRM database for each site. The examples provide instructions for the vCenter database. Repeat the process for the SRM server database and the vCenter database at each site. Start Microsoft SQL Server Management Studio Express by clicking **Start** → **All programs** → **Microsoft SQL Server 2005** and then click **SQL Server Management Studio Express** (Figure 7-79).

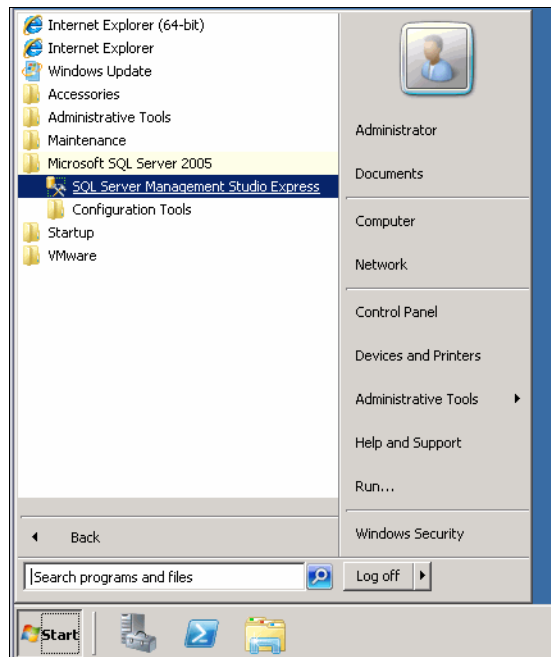


Figure 7-79 Starting MS SQL Server Management Studio Express

- The login window shown in Figure 7-80 is displayed. Do not change any of the values in this window. Click **Connect**.



Figure 7-80 Login window for MS SQL Server Management Studio

- After successful login, the MS SQL Server Management Suite Express main window is displayed (Figure 7-81). In this window, use configuration tasks to create databases and logins. To create databases, right-click **Databases**, and select **New database**.

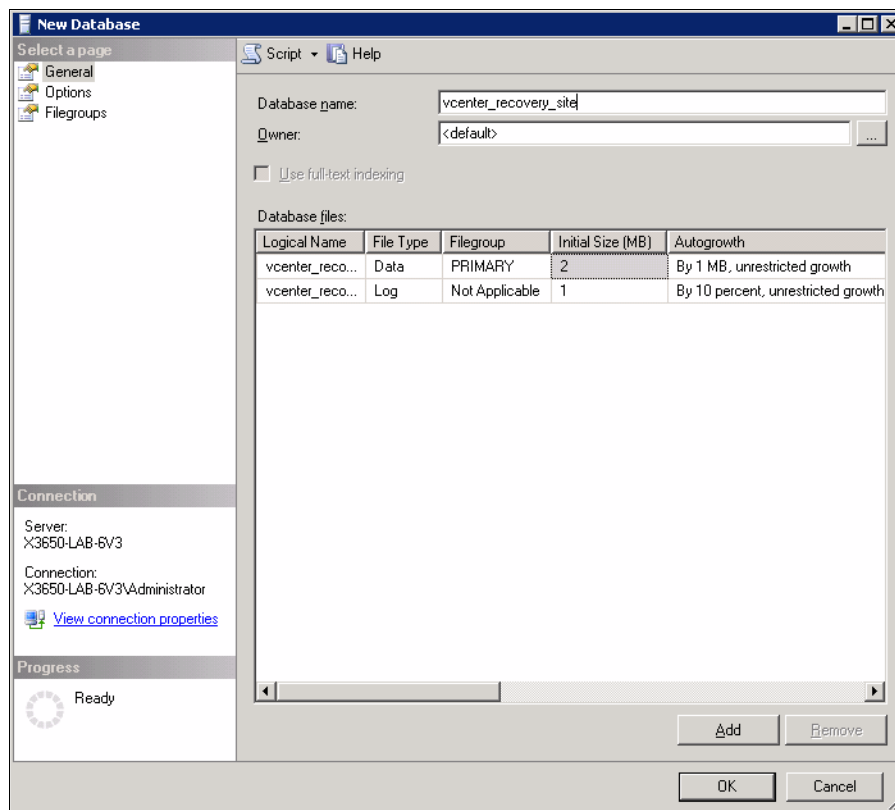


Figure 7-81 Add database window

- Enter the information for the database name, owner, and database files. In this example, only the database name is set, leaving all others parameters at their default values. Click **OK** to create your database.

10. Check to see whether the new database was created using the Object Explorer. Expand **Databases** → **System Databases**, and verify that there is a database with the name you entered. In Figure 7-82, the names of the created databases are circled in red. After creating the required databases, you must create a login for them.

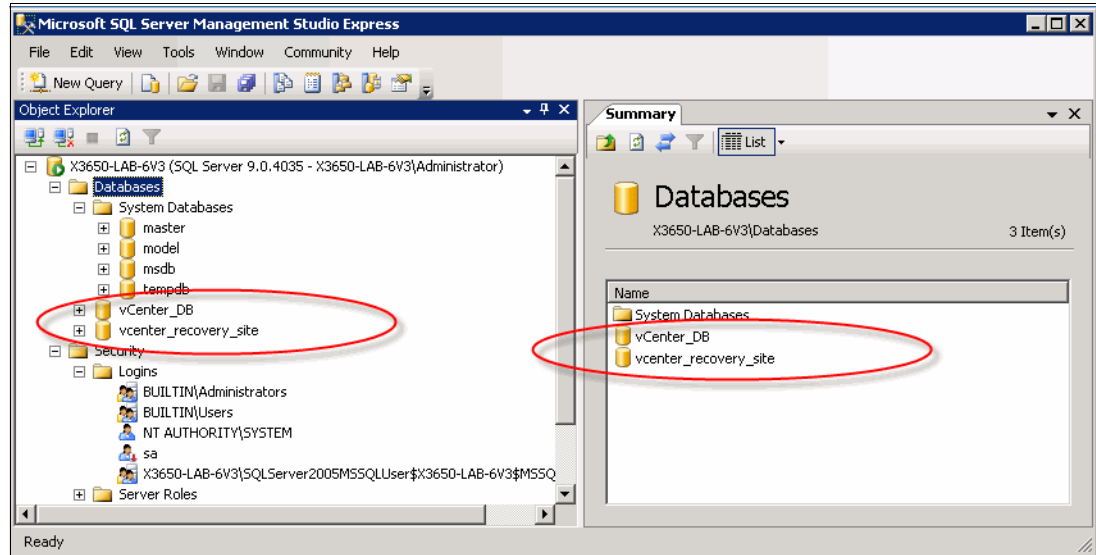


Figure 7-82 Verifying that the database is created

11. Right-click the subfolder logins, and select new login in the window (Figure 7-83 on page 170). Enter the user name, type of authentication, default database, and default code page. Click **OK**. Repeat this action for the vCenter and SRM servers databases.

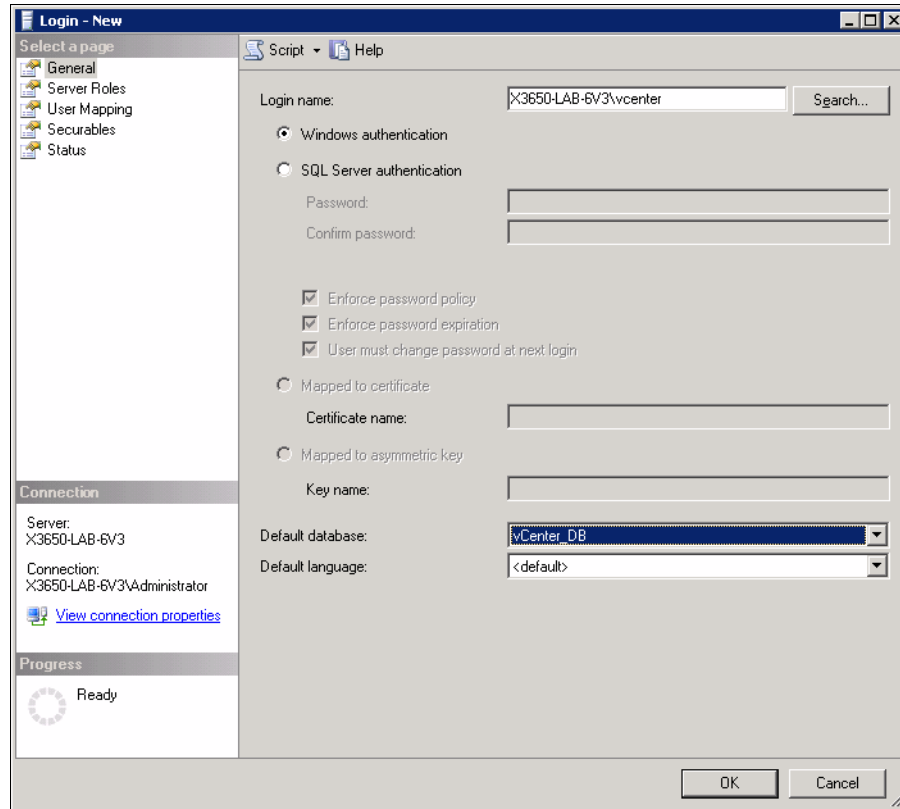


Figure 7-83 Defining database logins

12. Now grant rights to the database objects for these logins. Right-click **Logins** on the vcenter user login, and select **Properties**.

13. A new window opens, as shown in Figure 7-84 on page 171. Select **User Mappings**, and select **vCenter** database in the upper-right pane. In the lower-right pane, select the **db_owner** and **public** roles. Finally, click **OK**, and repeat those steps for the srmuser.

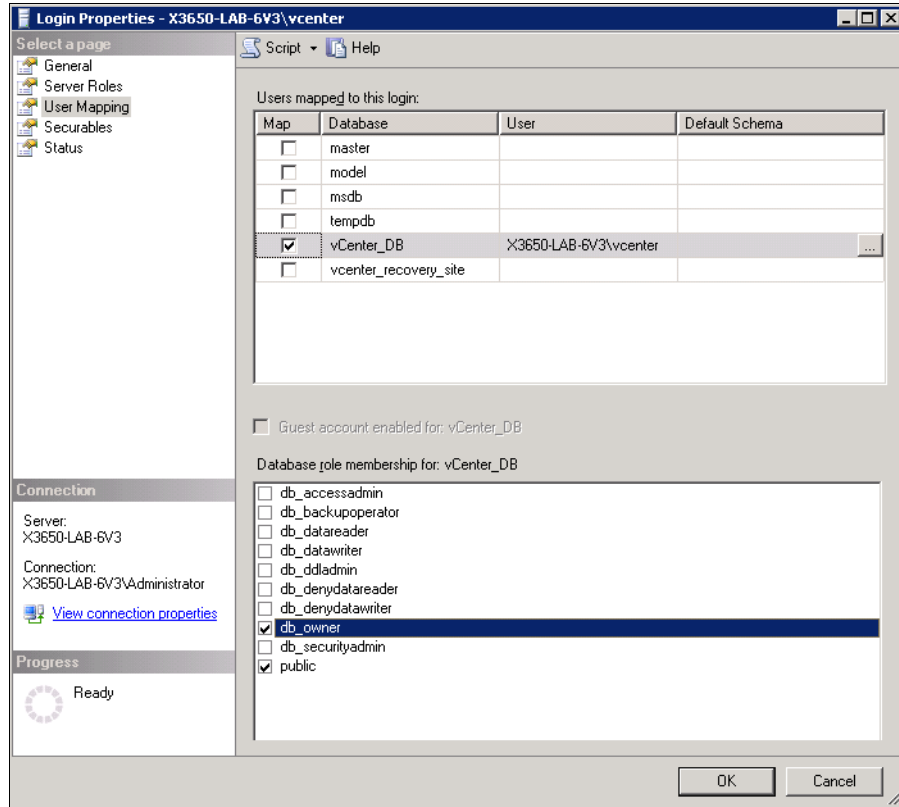


Figure 7-84 Granting the rights on a database for the login created

You are ready to configure the ODBC data sources for the vCenter and SRMDB databases on the server you plan to install them on.

14. To configure ODBC data stores, click **Start** in the Windows desktop task bar, and select **Administrative Tools** → **Data Source (ODBC)**. The ODBC Data Source Administrator window is now open, as shown in Figure 7-85.
15. Click the **System DSN** tab, and click **Add**.

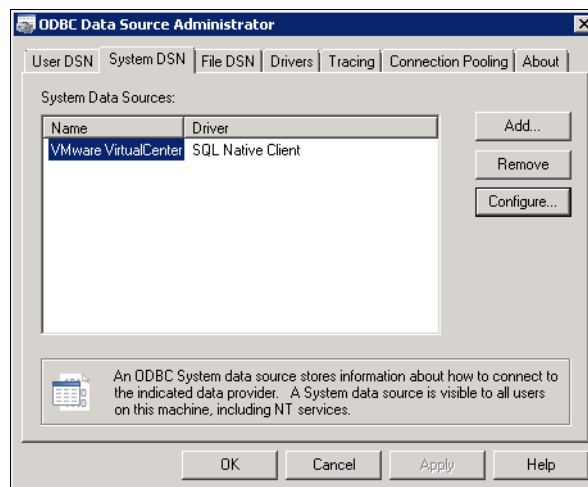


Figure 7-85 Selecting system DSN

16. In the Create New Data Source window, Figure 7-86, select **SQL Native Client**, and click **Finish**.

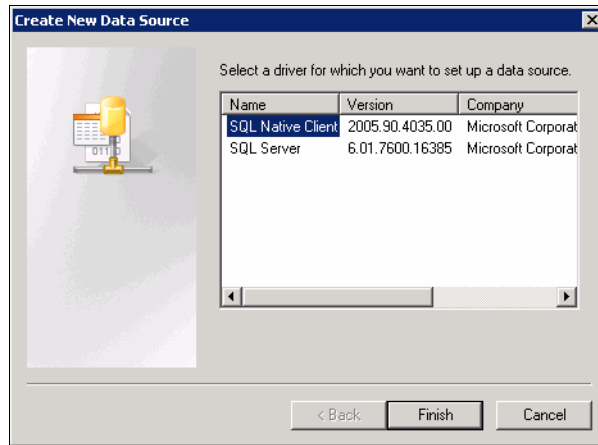


Figure 7-86 Selecting SQL driver

17. The window shown in Figure 7-87 opens. Enter information for your data source, such as the name, description, and server, for the vcenter database. In the example, the parameters are set as follows:

- Name parameter to vcenter
- Description parameter to database for vmware vcenter
- Server parameter to SQLEXPRESS.

Click **Next**.

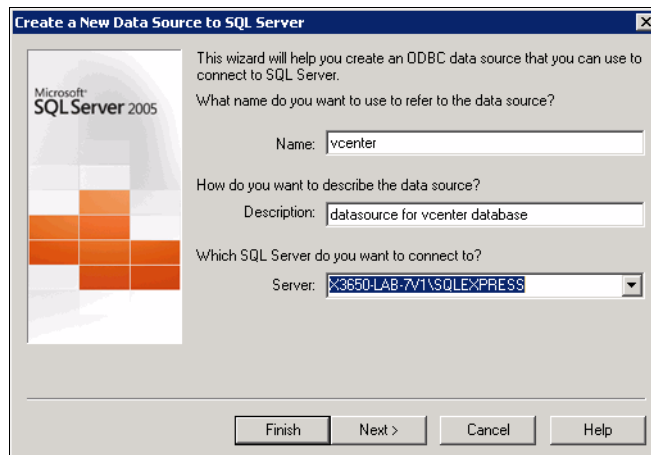


Figure 7-87 Defining data source name and server

18. The window shown in Figure 7-88 on page 173 opens. Select **With Integrated Windows Authentication** and **Connect to SQL Server to obtain default settings to the additional configuration options**, and click **Next**.

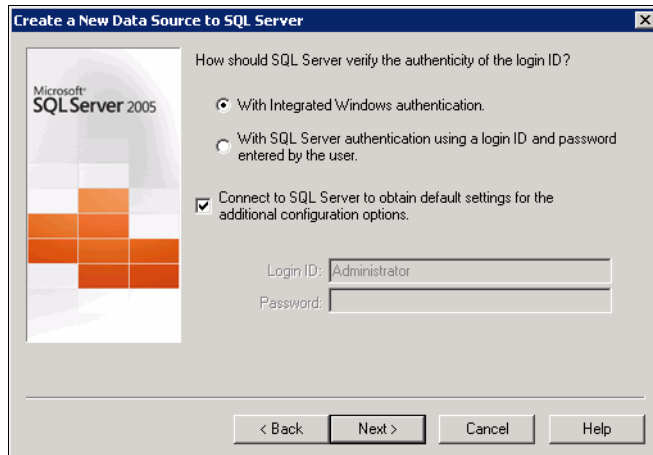


Figure 7-88 Selecting authorization type

19. Select **Change default database, vCenter_DB**, and the two check boxes as shown in Figure 7-89. Click **Next**. The window shown in Figure 7-90 is displayed.

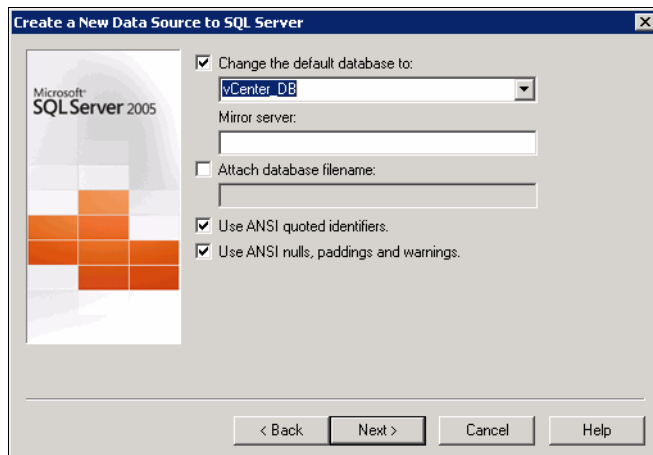


Figure 7-89 Selecting default database for data source

20. Select **Perform translation for the character data** and then click **Finish**.

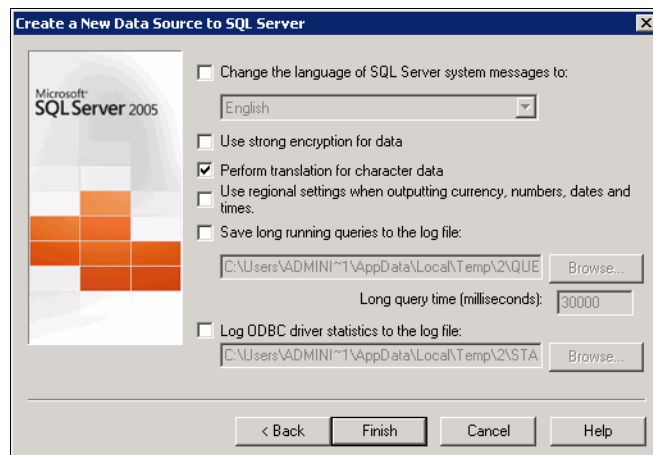


Figure 7-90 SQL server database locale-related settings

21. In the window shown in Figure 7-91, inspect the information for your data source configuration, and click **Test Data Source**.

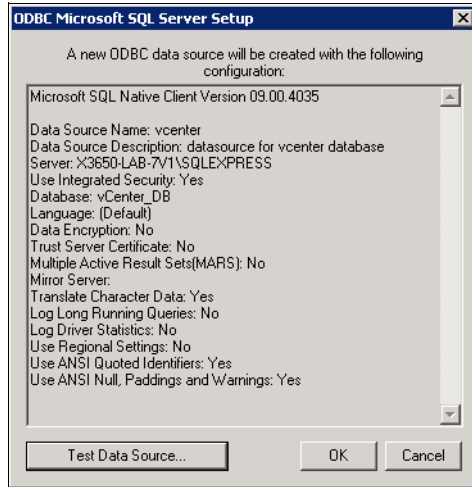


Figure 7-91 Testing data source and completing setup

22. The window shown in Figure 7-92 indicates that the test completed successfully. Click **OK** to return to the previous window and then click **Finish**.

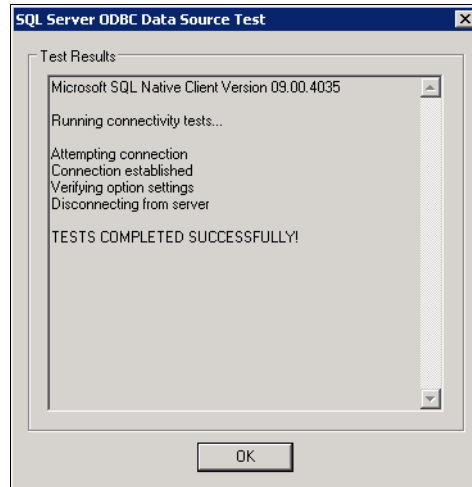


Figure 7-92 Results of data source test

23. You are returned to the window shown in Figure 7-93. You can see the list of Data Sources defined system-wide. Check the presence of the **vcenter** data source.

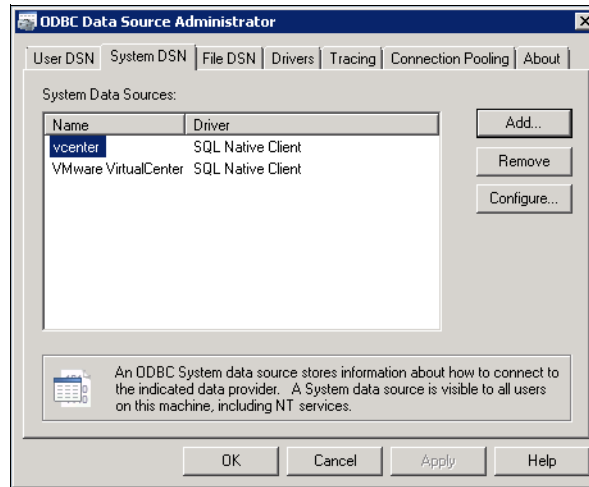


Figure 7-93 Defined system data sources

Install and configure databases on all sites that you plan to include into your business continuity and disaster recovery solution.

Now you are ready to proceed with the installation of vCenter server, vCenter client, SRM server, and SRA agent.

7.2.2 Installing the vCenter server

This section illustrates the step-by-step installation of vCenter server under Microsoft Windows Server 2008 R2 Enterprise.

Further Information: For detailed information about vCenter server installation and configuration, see the VMware documentation. This section includes only common, basic information for a simple installation used to demonstrate SRM server capabilities with the IBM XIV Storage System.

Perform the following steps to install the vCenter Server:

1. Locate the vCenter server installation file (either on the installation CD or a copy you downloaded from the Internet).
2. Follow the installation wizard guidelines until you reach the step where you are asked to enter information about database options.

3. Choose the database for vCenter server. Select the **Using existing supported database** option, and enter vcenter as the **Data Source Name** as seen in Figure 7-94. The name of the DSN must be the same as the ODBC system DSN that was defined earlier. Click **Next**.

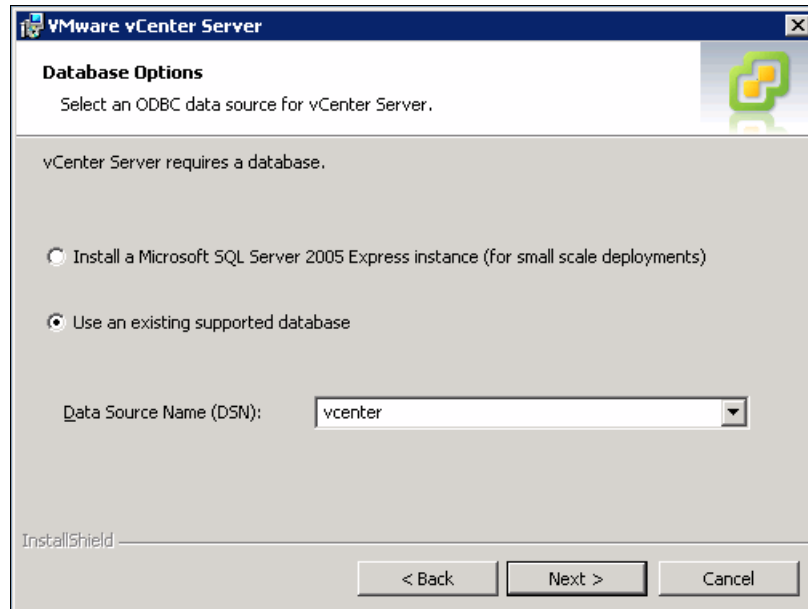


Figure 7-94 Choosing database for vCenter server

4. In the window shown in Figure 7-95, type the password for the system account, and click **Next**.

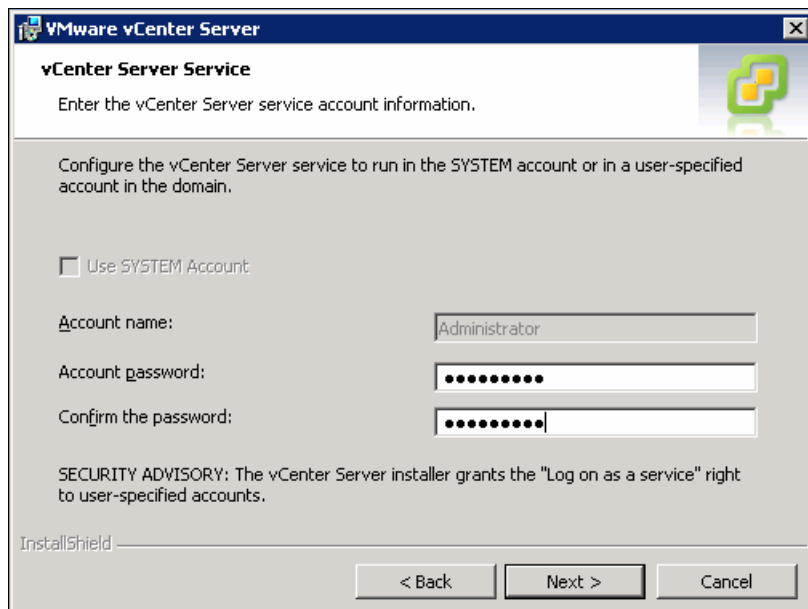


Figure 7-95 Requesting password for the system account

- In Figure 7-96, choose a Linked Mode for the installed server. For a first time installation, select **Create a standalone VMware vCenter server instance**. Click **Next**.

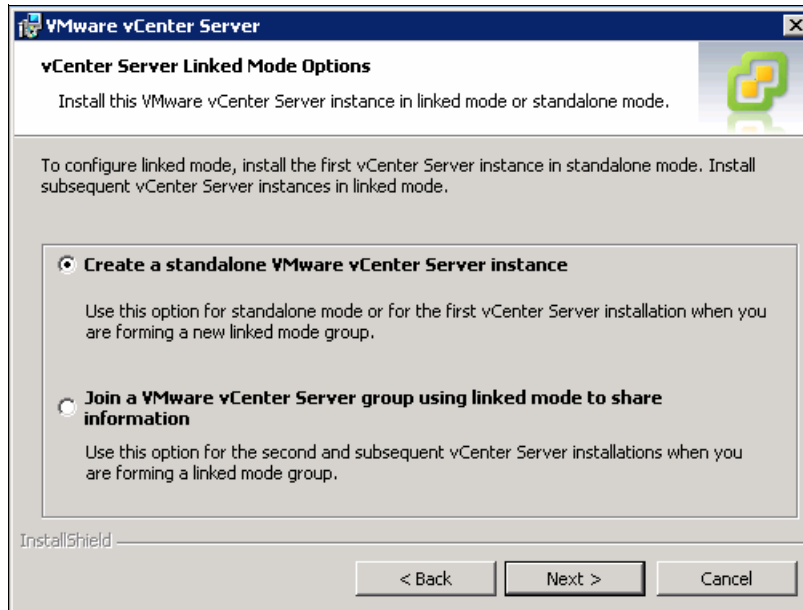


Figure 7-96 Selecting Linked Mode options for the vCenter server

- In the window shown in Figure 7-97, you can change default settings for ports used for communications by the vCenter server. For most implementations, keep the default settings. Click **Next**.

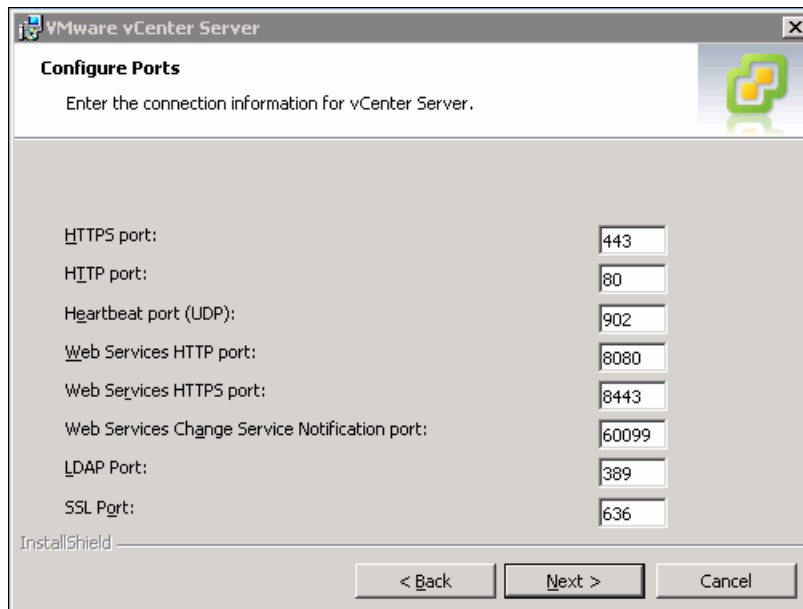


Figure 7-97 Configuring ports for the vCenter server

7. In the window shown in Figure 7-98, select the required memory size for the JVM used by vCenter Web Services according to your environment. Click **Next**.

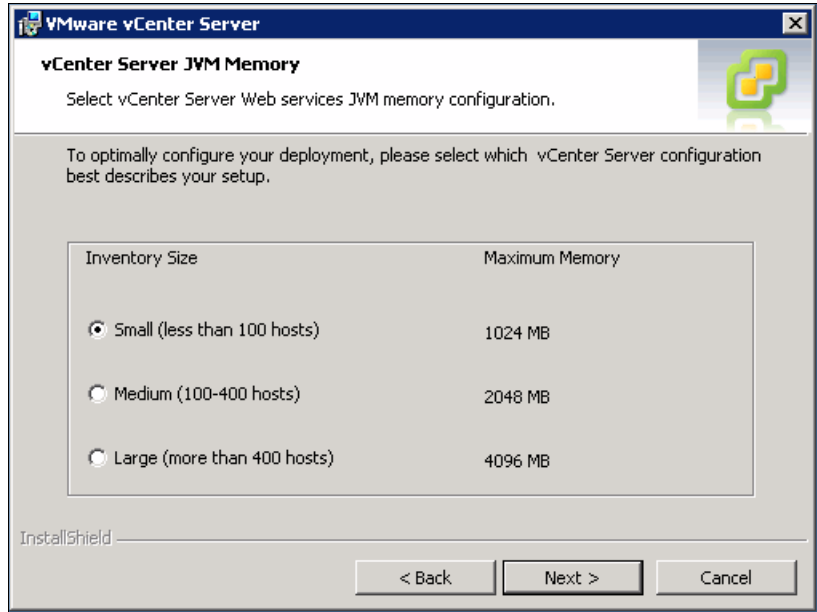


Figure 7-98 Setting inventory size

8. The window shown in Figure 7-99 indicates that the system is now ready to install vCenter. Click **Install**.

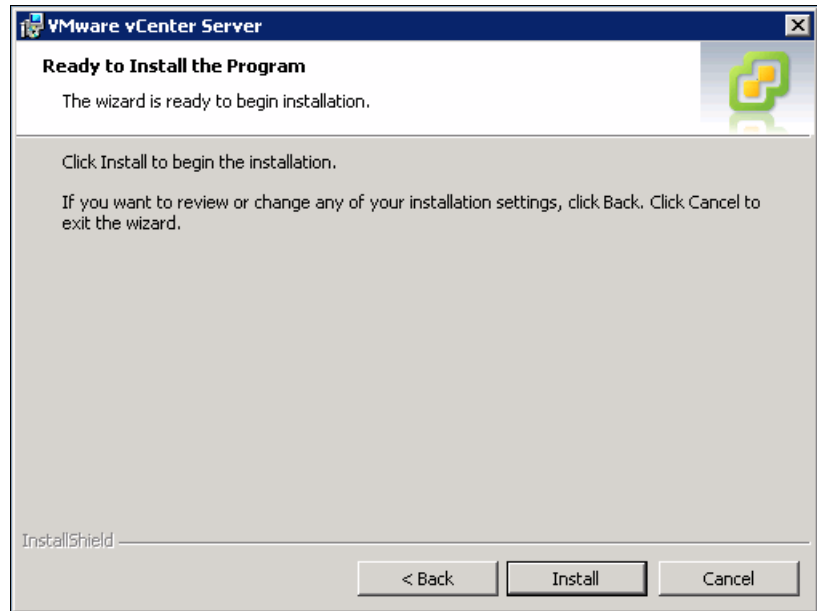


Figure 7-99 Ready to Install the Program window

9. After the installation completes, the window shown in Figure 7-100 is displayed. Click **Finish**.

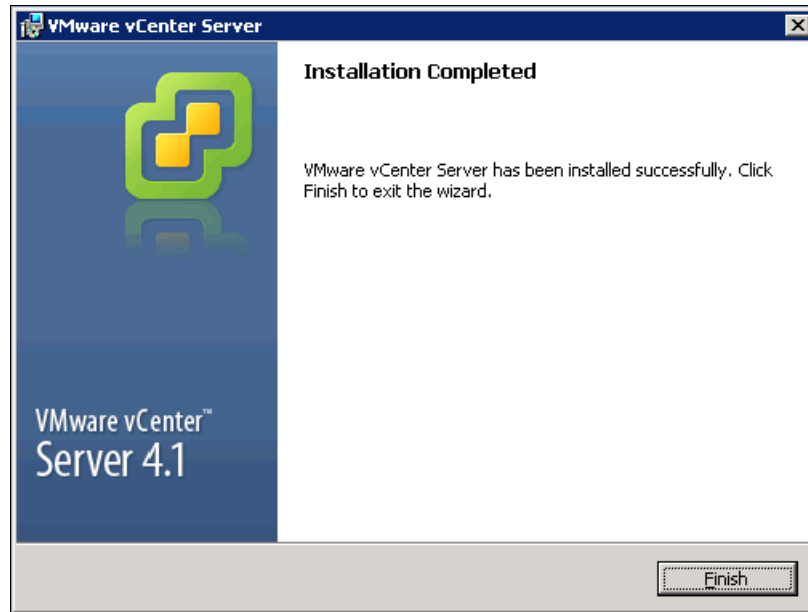


Figure 7-100 The vCenter installation is completed

You need to install vCenter server on all sites that you plan to include as part of your business continuity and disaster recovery solution.

7.2.3 Installing and configuring vCenter client

This section illustrates the installation of the vSphere Client under Microsoft Windows Server 2008 R2 Enterprise.

Note: For detailed information about vSphere Client, and complete installation and configuration instructions, see the VMware documentation. This chapter includes only basic information about installing the vSphere Client and using it to manage the SRM server.

Locate the vCenter server installation file (either on the installation CD or a copy you downloaded from the Internet). Running the installation file displays the vSphere Client installation wizard welcome dialog. Follow the installation wizard instructions to complete the installation. You need to install vSphere Client on all sites that you plan to include in your business continuity and disaster recovery solution.

After you finish installing SQL Server 2005 Express, vCenter server, and vSphere Client, place existing ESX servers under control of the newly installed vCenter server. To perform this task, follow these instructions:

1. Start the vSphere Client.
2. In the login window shown in Figure 7-101, type the IP address or system name of your vCenter server, and a user name and password. Click **Login**.

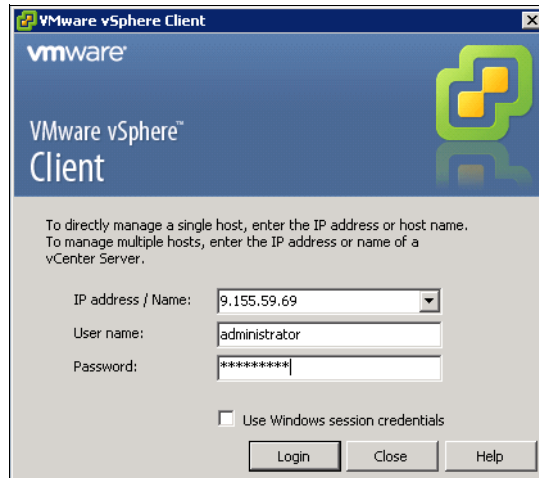


Figure 7-101 vSphere Client login window

3. Add the new data center under control of the newly installed vCenter server. In the main vSphere Client window, right-click the server name, and select **New Datacenter**, as shown in Figure 7-102.

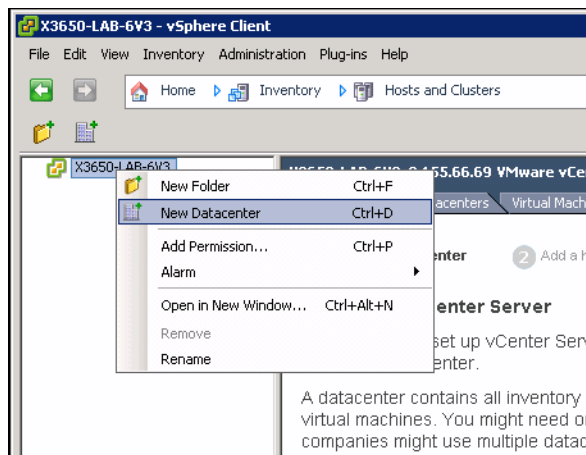


Figure 7-102 Defining the data center

4. Enter a new name for the data center, as shown in Figure 7-103.

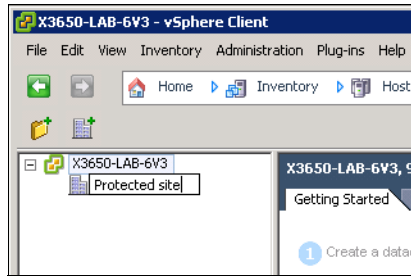


Figure 7-103 Specifying the name of the data center

5. The Add Host wizard is started. Enter the name or IP address of the ESX host, user name for the administrative account on this ESX server, and the account password (Figure 7-104). Click **Next**.

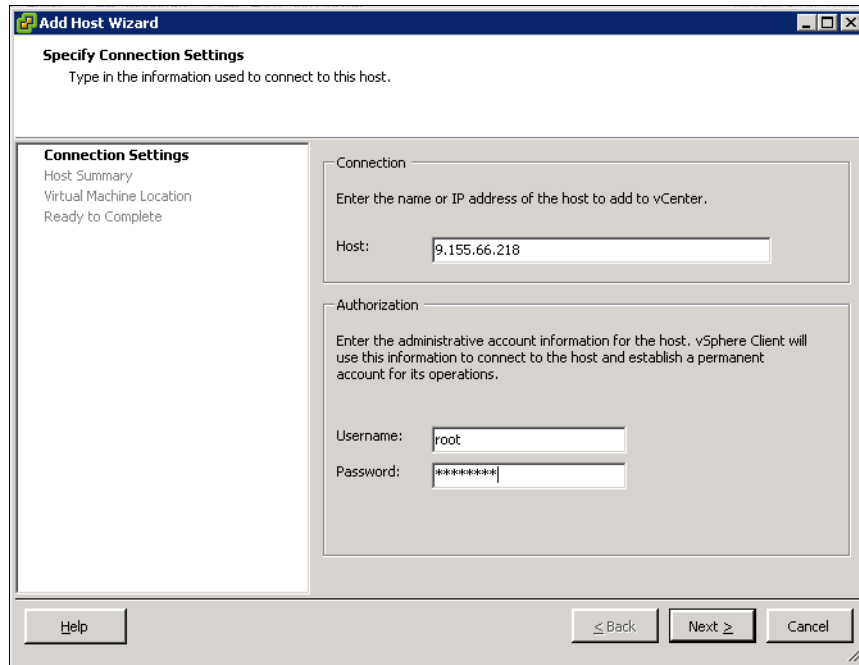


Figure 7-104 Specifying host name, user name, and password

6. Verify the authenticity of the specified host, as shown in Figure 7-105. If it is correct, click **Yes** to continue to the next step.

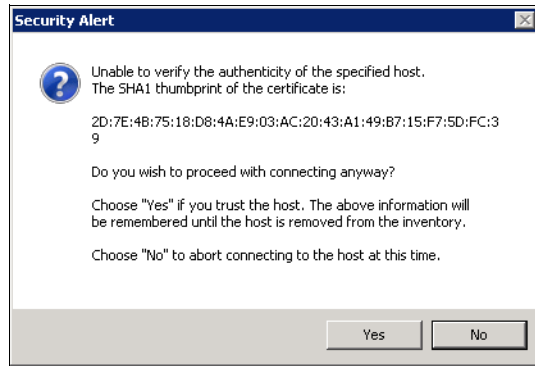


Figure 7-105 Verifying the authenticity of the specified host

7. Verify the settings discovered for the specified ESX host, as shown in Figure 7-106. Check the information, and, if all is correct, click **Next**.

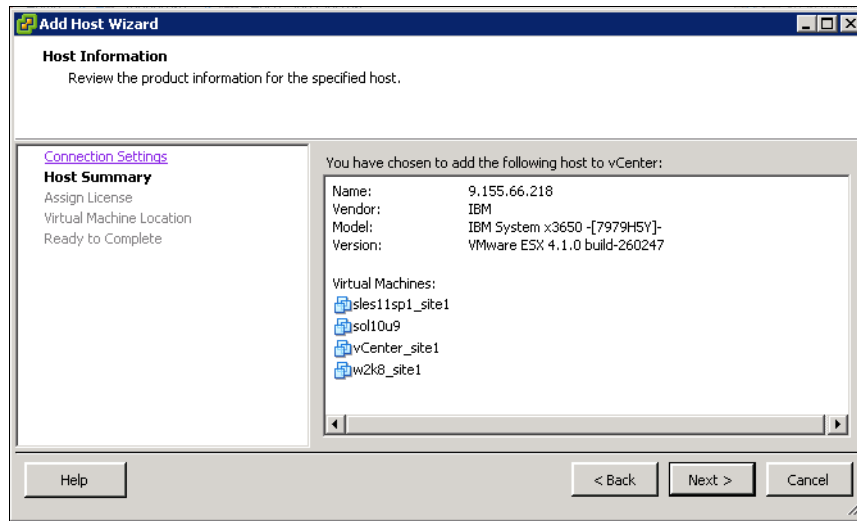


Figure 7-106 Configuration summary of the discovered ESX host

8. Select ESX host in evaluation mode or enter a valid license key for the ESX server, as shown in Figure 7-107. Click **Next**.

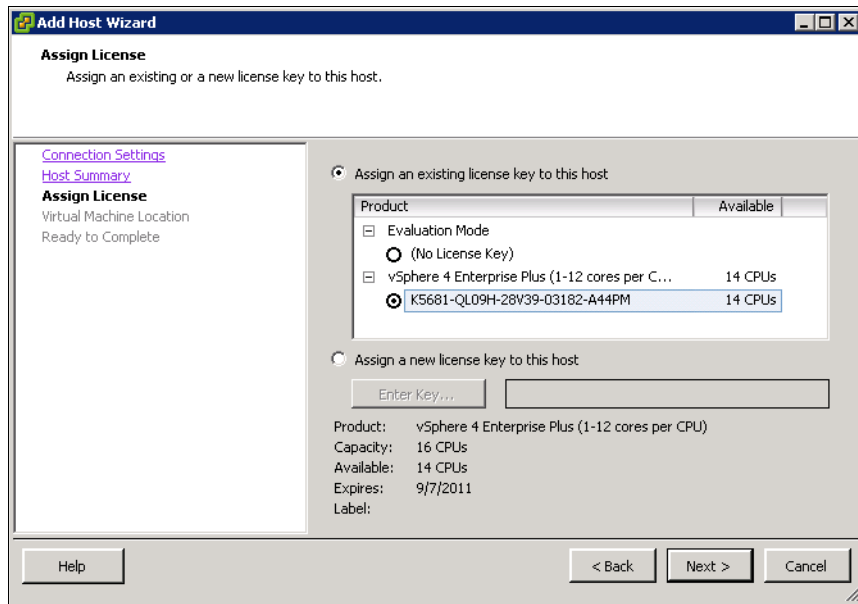


Figure 7-107 Assigning license to the host

9. Select a location for the newly added ESX server, as shown in Figure 7-108, and click **Next**.

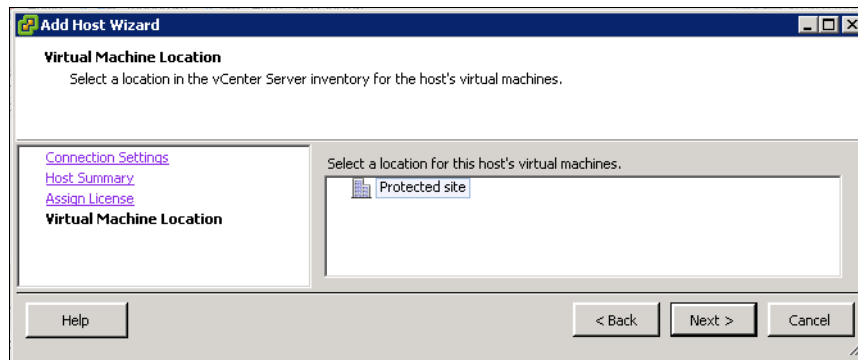


Figure 7-108 Selecting the location in the vCenter inventory for the virtual machines of the host

10. The window shown in Figure 7-109 summarizes your settings. Check the settings and, if they are correct, click **Finish**.

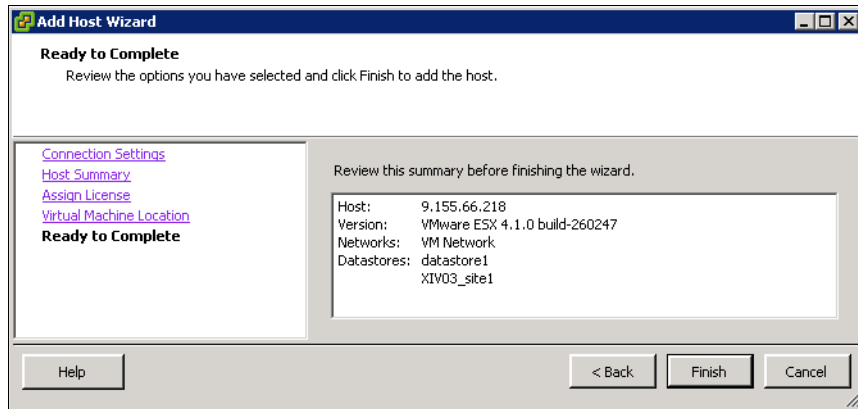


Figure 7-109 Review summary

You return to the vSphere Client main window as shown in Figure 7-110.

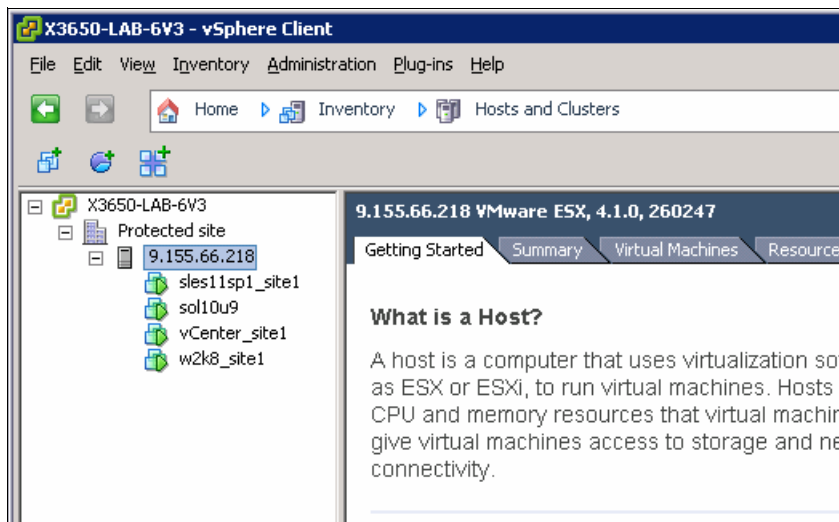


Figure 7-110 Presenting inventory information about ESX server in the vCenter database

Repeat all these steps for all the vCenter servers you want to include into your business continuity and disaster recovery solution.

7.2.4 Installing the SRM server

This section describes the basic installation tasks for the VMware SRM server version 4.X or 5.X under Microsoft Windows Server 2008 R2 Enterprise. To install VMware SRM server:

1. Locate the vCenter server installation file, either on the installation CD or a copy you downloaded from the Internet.
2. Run the installation file.
3. The Welcome window for the vCenter Site Recovery Manager wizard is displayed, as shown in Figure 7-111 on page 185. Click **Next**, and follow the installation wizard guidelines.

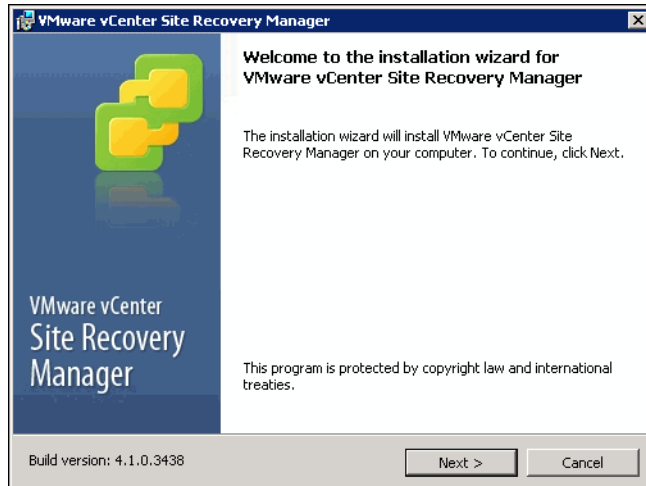


Figure 7-111 SRM Installation wizard welcome message

4. Provide the vCenter server IP address, vCenter server port, vCenter administrator user name, and password for the administrator account (Figure 7-112). Click **Next**.

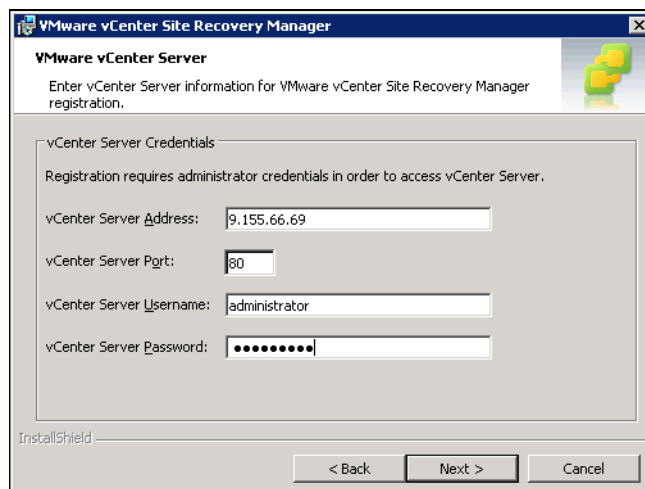


Figure 7-112 SRM settings on paired vCenter server

5. You might get a security warning similar to Figure 7-113. Check the vCenter server IP address and, if it is correct, click **OK**.

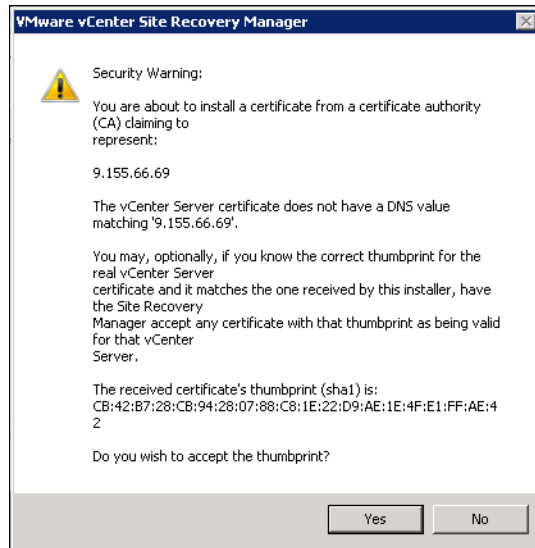


Figure 7-113 Certificate acceptance window

6. Select **Automatically generate certificate**, as shown in Figure 7-114, and click **Next**.

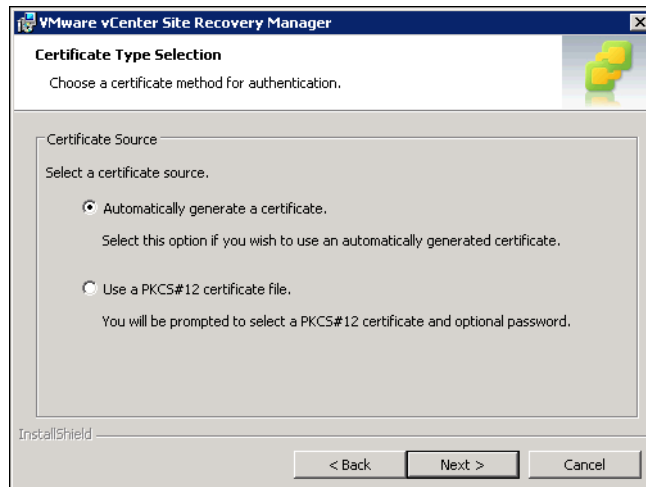


Figure 7-114 Selecting certificate type

Tip: If your vCenter servers are using NON-default (that is, self signed) certificates, select **Use a PKCS#12 certificate file**. For more information, see the VMware vCenter Site Recovery Management Administration Guide at:

http://www.vmware.com/pdf/srm_admin_4_1.pdf

7. Enter details, such as organization name and organization unit, that are used as parameters for certificate generation (Figure 7-115). When complete, click **Next**.

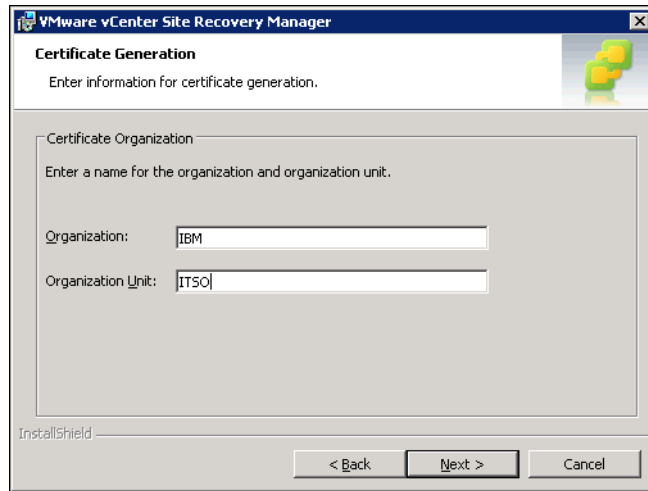


Figure 7-115 Setting up certificate generation parameters

8. The window shown in Figure 7-116 asks for general parameters pertaining to your SRM installation. Provide the location name, administrator email, additional email, local host IP address or name, and the ports to be used for connectivity. When complete, click **Next**.

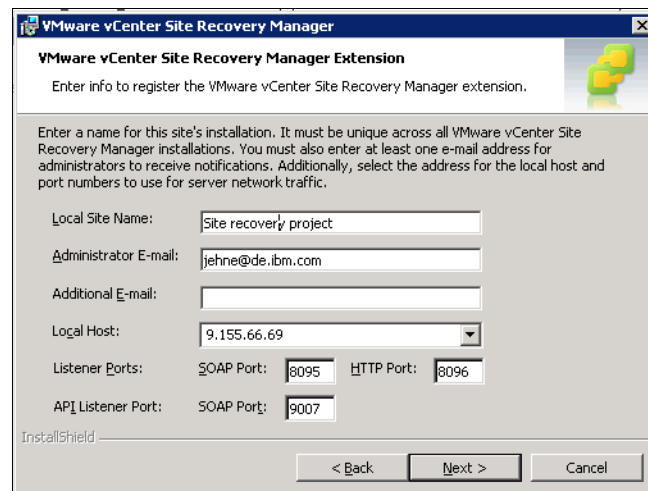


Figure 7-116 General SRM server settings for the installation location

9. Enter parameters related to the database that was previously installed, as shown in Figure 7-117. These parameters are types of the database, ODBC System data source, user name and password, and connection parameters. Click **Next**.

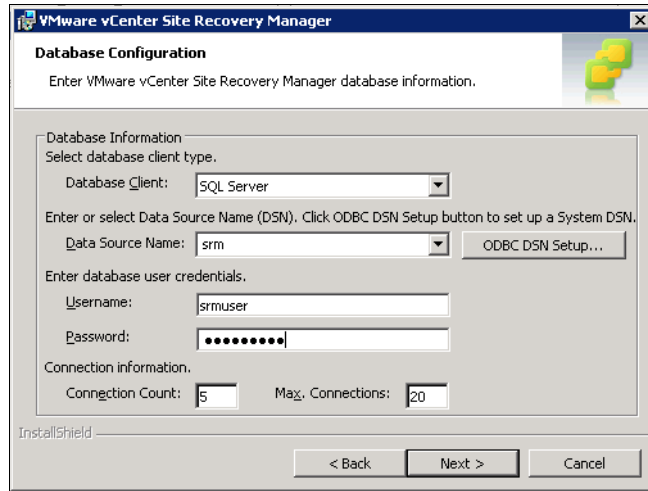


Figure 7-117 Specifying database parameters for the SRM server

10. The next window informs you that the installation wizard is ready to proceed, as shown in Figure 7-118. Click **Install** to start the installation process.

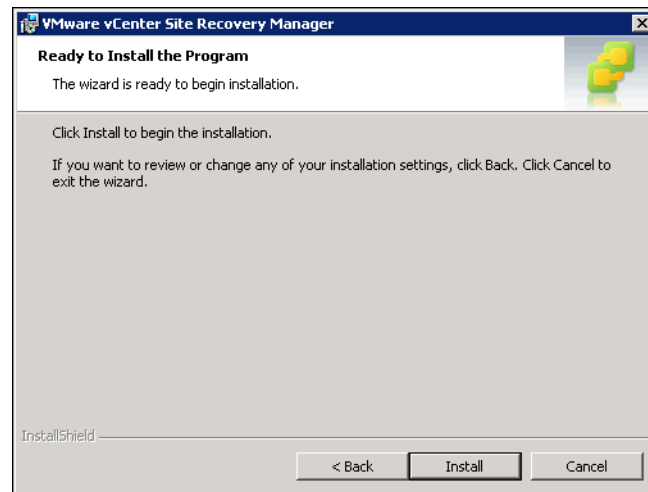


Figure 7-118 Ready to Install the Program window

You need to install SRM server on each protected and recovery site that you plan to include into your business continuity and disaster recovery solution.

7.3 Installing the vCenter Site Recovery Manager plug-in

Now that you installed the SRM server, install the SRM plug-in on the system that is hosting your vSphere Client:

1. Run the vSphere Client.
2. Connect to the vCenter server on the site where you are planning to install the SRM plug-in.

- In the vSphere Client console, click **Plug-ins** → **Manage Plug-ins**, as shown in Figure 7-119. The Plug-in Manager window opens.

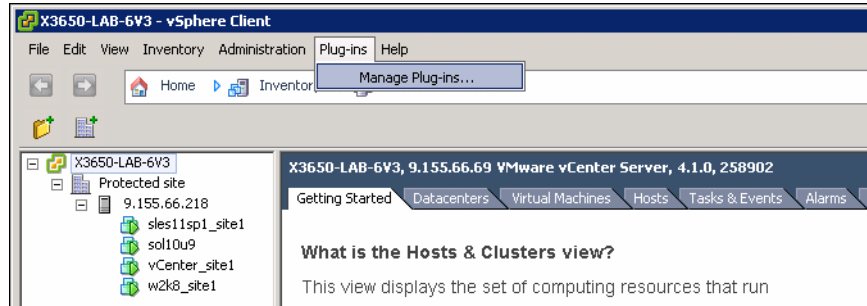


Figure 7-119 Selecting the Manage Plug-ins option

- Under the category Available plug-ins, right-click **vCenter Site Recovery Manager Plug-in**, and select **Download and Install**, as shown in Figure 7-120.

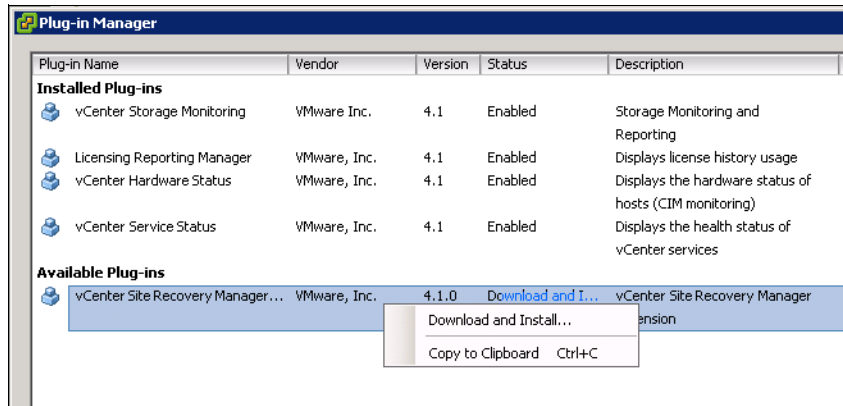


Figure 7-120 Downloading and installing SRM plug-in

- The vCenter Site Recovery Manager Plug-in wizard is started. Follow the wizard guidelines to complete the installation.

You need to install SRM plug-in on each protected and recovery site that you plan to include in your business continuity and disaster recovery solution.

7.3.1 Installing XIV Storage Replication Adapter for VMware SRM

This section provides the steps for installing the XIV SRA for VMware SRM server versions 4.X and 5.X under Microsoft Windows Server 2008 R2 Enterprise. Download and install XIV SRA for VMware on each SRM server in your business continuity and disaster recovery solution:

- Locate the XIV Storage Replication Adapter installation file:
 - For SRM 4.X, the SRM and SRA installation files are located at: <https://my.vmware.com/web/vmware/details/srm412/ZCVwYnRocHBidGR0cA==>
 - For SRM 5.0 and 5.0.1, the SRM and SRA installation files are located at: <https://my.vmware.com/web/vmware/details/srm501/dHdwYnd1KiVi dHdAZQ==>

- For SRM 5.1.0, the SRM and SRA installation files are located at:
https://my.vmware.com/web/vmware/details?downloadGroup=SRM510_GA&productId=291&rPid=2941
2. Run the installation file.
 3. The vCenter Site Recovery Adapter installation wizard is displayed, as shown in Figure 7-121. Click **Next**.

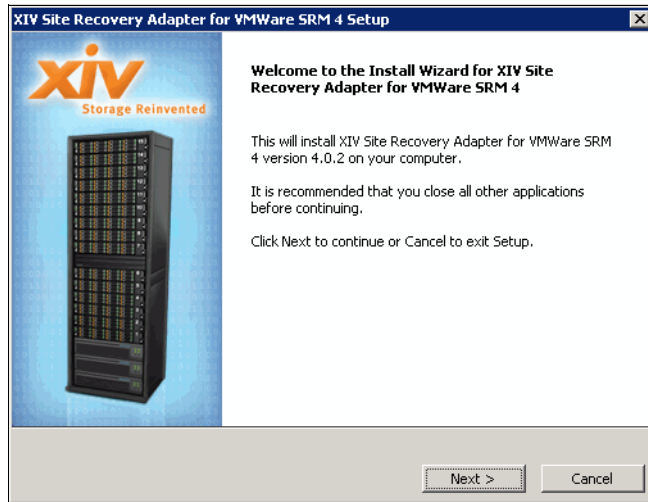


Figure 7-121 Welcome to SRA installation wizard window

4. Follow the wizard guidelines to complete the installation.

Tip: For step-by-step instructions containing detailed installation guidance, error message definitions and resolutions, and basic configuration practices for the IBM XIV Site Recovery Adapter for VMware version 2.1.0, refer to the guide at:

http://pic.dhe.ibm.com/infocenter/strhosts/ic/topic/com.ibm.help.strghosts.doc/PDFs/XIV_Adapter_for_VMware_VC_SRM_2.1.0_UG.pdf

7.3.2 Configuring the IBM XIV System Storage for VMware SRM

Make sure that all virtual machines that you plan to protect are on IBM XIV Storage System volumes. If there are any virtual systems that are not on IBM XIV Storage System, move them using these steps:

1. Create volumes on XIV.
2. Add the data store to the ESX server.
3. Migrate or clone that virtual machine to relocate it to XIV volumes.

For more information about connecting ESX hosts to the IBM XIV Storage, see Chapter 4, “Attaching VMware ESX to XIV” on page 41.

Create a storage pool on the IBM XIV Storage System at the recovery site. The new storage pool contains the replicas of the ESX host data stores that are associated with virtual machines that you plan to protect.

Remember: Configure a snapshot size of at least 20 percent of the total size of the recovery volumes in the pool. For testing failover operations that can last several days, increase the snapshot size to half the size of the recovery volumes in the pool. For longer-term or I/O intensive tests, the snapshot size might have to be the same size as the recovery volumes in the pool.

For information about IBM XIV Storage System LUN mirroring, see the IBM Redbooks publication *IBM XIV Storage System: Copy Services and Migration*, SG24-7759.

At least one virtual machine for the protected site must be stored on the replicated volume before you can start configuring SRM server and SRA adapter. In addition, avoid replicating swap and paging files.

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

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only:

- ▶ *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659
- ▶ *IBM XIV Storage System: Copy Services and Migration*, SG24-7759
- ▶ *XIV Storage System: Host Attachment and Interoperability*, SG24-7904
- ▶ *XIV Storage System SSD Caching Implementation*, REDP-4842
- ▶ *Using the IBM XIV Storage System with OpenStack Cloud Environments*, REDP-4971

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

- ▶ *IBM XIV Storage System Planning Guide*, GC27-3913
- ▶ *IBM XIV Storage System Pre-Installation Network Planning Guide for Customer Configuration*, GC52-1328-01
- ▶ *IBM XIV Storage System: Product Overview*, GC27-3912
- ▶ *IBM XIV Remote Support Proxy Installation and User's Guide*, GA32-0795
- ▶ *IBM XIV Storage System User Manual*, GC27-3914
- ▶ *IBM XIV Storage System Management Tools version 4.0 User Guide* SC27-4230-00
- ▶ *IBM XIV Storage System XCLI Utility User Manual*, GC27-3915

Online resources

These websites are also relevant as further information sources:

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

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services



XIV Storage System in a VMware Environment



XIV and VMware Implementation Best Practices

Data Protection with FlashCopy Manager

XIV and VMware Site Recovery Manager

The IBM XIV Storage System is an excellent choice for VMware storage requirements. XIV achieves consistent high performance by balancing the workload across physical resources. This paper includes information about the following topics:

- ▶ XIV Storage System and VMware Integration Concepts
- ▶ XIV Storage System and VMware Implementation Best Practices
- ▶ XIV Storage System integration harnessing VMware APIs including:
 - vStorage APIs for Array Integration (VAAI)
 - vSphere API for Storage Awareness (VASA)
 - vStorage API for Data Protection (VADP) interfacing with Tivoli Storage FlashCopy Manager (FCM) and Tivoli Storage Manager for Virtualization Environments (TSM for VE)
- ▶ Connection for ESX version 3.5, ESX/ESXi version 4.x, and ESXi 5.0/5.1
- ▶ The IBM vCenter plug-in
- ▶ The XIV Storage Replication Adaptor (SRA)
- ▶ VMware Site Recovery Manager (SRM)

This IBM Redpaper is intended for those who want to plan, design, and configure an XIV based storage solution in a VMware environment.

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