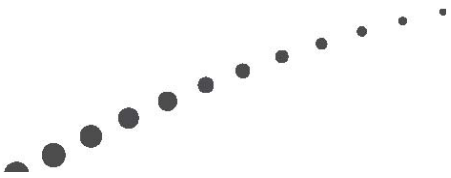




***Citrix XenServer Design:
Configuring iSCSI Multipathing
Support for XenServer***

www.citrix.com





Contents

About.....	3
Visual Legend	4
Additional Terminology	5
Chapter 1: Introduction to iSCSI Multipathing	6
Overview	6
Creating Redundancy for Storage Traffic.....	6
Multipathing iSCSI Storage Traffic	8
Understanding Multipathing.....	9
Chapter 2: Configuring Software Initiator iSCSI Multipathing	12
Overview	12
Before Configuring iSCSI Multipathing	12
Multipath Handler Support.....	13
Correct Physical and Subnet Configuration	13
Configuring iSCSI Software Initiator Multipathing	14
Chapter 3: Creating the Storage Repository.....	17
Overview	17
Creating the Storage Repository after Enabling Multipathing	17
Revision History	21

About

This document helps you understand how to configure software initiator iSCSI multipathing for XenServer. It includes the following topics:





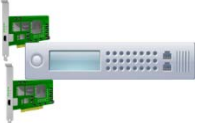
- An overview of multipathing in general and iSCSI multipathing in particular
- Instructions for configuring software initiator iSCSI multipathing
- After configuring multipathing, instructions for selecting Target IQNs when creating a storage repository using XenCenter

Due to scope, this guide provides some, but not much, device-specific information. For device-specific configuration, Citrix suggests reviewing the storage vendor's documentation, the storage vendor's hardware compatibility list, and contacting the vendor's technical support if necessary.

Note: XenServer supports multipathing for Fibre Channel and iSCSI SANs. For information about Fibre Channel multipathing, see the *XenServer Administrator's Guide* and the Citrix Knowledge Center.

Visual Legend

This guide relies heavily on diagrams to explain key concepts. These diagrams use the following icons:

Icon	Meaning
	<p>Host. A XenServer host is the physical server on which the XenServer hypervisor is running.</p>
	<p>NIC. The physical network card (NIC) in your host.</p>
	<p>Physical Switch. The device on a physical network that connects network segments together.</p>
	<p>Storage Array. This icon represents a generic storage array with a LUN configured in it.</p>
	<p>Storage Controller. This icon represents a storage controller on a storage array.</p>



Additional Terminology

These terms appear in the sections that follow:

Management Interface. The primary management interface is a NIC assigned an IP address that XenServer uses for its management network, including, but not limited to, traffic between hosts, between a host and Workload Balancing and for live migration. In previous versions of XenServer, the management interface was referred to as the *primary management interface*.

VM traffic. VM traffic refers to network traffic that originates or terminates from a virtual machine. This is sometimes referred to as guest traffic or VM/guest traffic.

Chapter 1: Introduction to iSCSI Multipathing

Overview

This chapter provides an introduction to configuring iSCSI multipathing support. It includes the following topics:

- An overview of the importance of redundancy for software initiator iSCSI storage traffic
- An explanation of multipathing and how it works

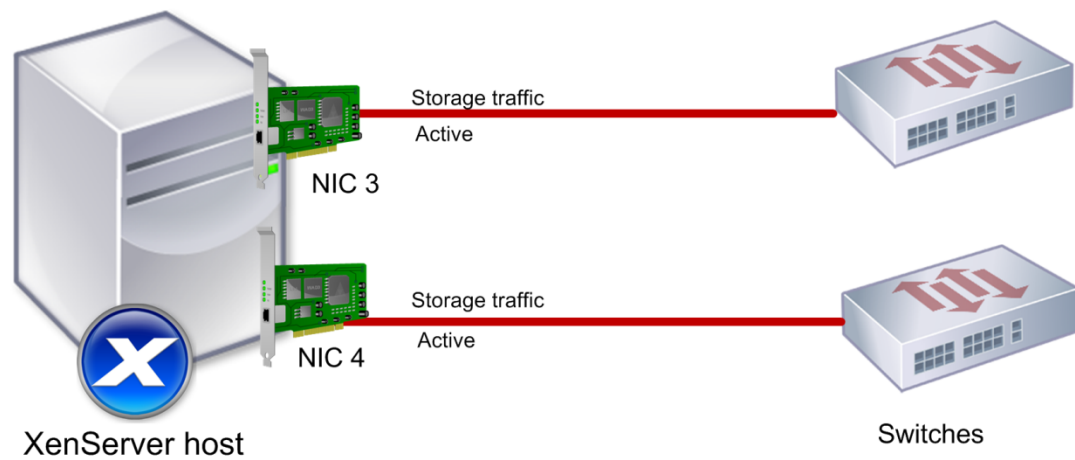
For some audiences, the information in this chapter may be unnecessary. However, this information is given to provide a common baseline of information and avoid pitfalls that result from common misconceptions about iSCSI software initiator multipathing.

Creating Redundancy for Storage Traffic

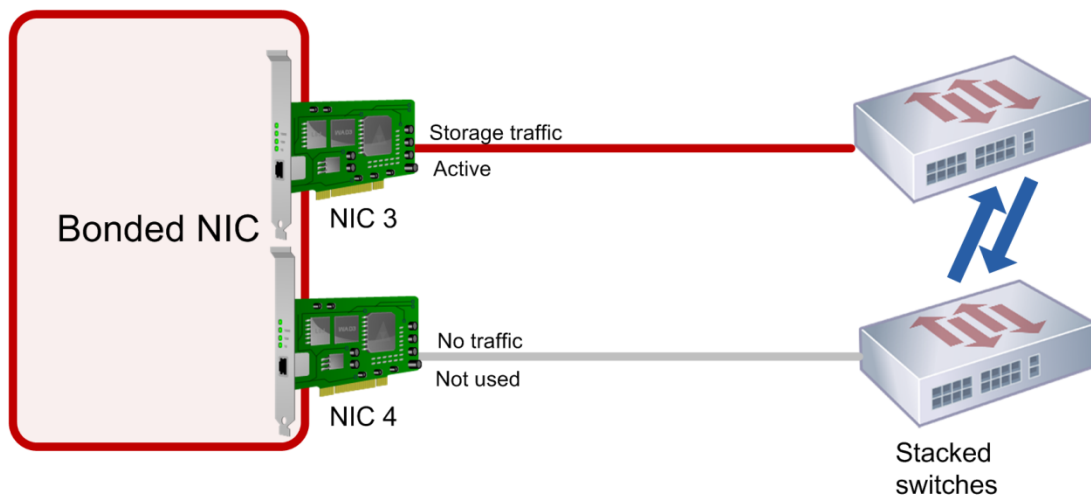
Configuring multipathing helps provide redundancy for network storage traffic in case of partial network or device failure. The term *multipathing* refers to routing storage traffic to a storage device over multiple paths for redundancy (failover) and increased throughput.

While NIC bonding can also provide redundancy for storage traffic, Citrix recommends configuring multipathing instead of NIC bonding whenever possible. When you configure multipathing, XenServer can send traffic down both paths: multipathing is an active-active configuration. (By default, multipathing uses round-robin mode load balancing, so both routes will have active traffic on them during normal operation, which results in increased throughput.) The illustration that follows provides a visual guide to the differences.

Multipathing



NIC Bonding



This illustration shows how, for storage traffic, both paths are active with multipathing whereas only one path is active with NIC bonding in active-active mode.

Note: LACP NIC bonding can load balance storage traffic as described in *Designing XenServer Network Configurations* for XenServer 6.1. However, multipathing is generally recommended over LACP bonding.

Citrix strongly recommends that you do not mix NIC bonding and iSCSI multipathing. There is no benefit from layering multipathing and NIC bonding on the same connection. After you enable



multipathing, you not only have better performance but you also have the failover that bonding would have provided.

Multipathing is incompatible with the following technologies, so you should consider using NIC bonding when:

- You have an NFS storage device.
- Your storage device does not support iSCSI connections over multiple IPs (for example, Dell EqualLogic or HP LeftHand SAN).
- You have limited number of NICs and need to route iSCSI traffic and file traffic (such as CIFS) over the same NIC.

For more information about choosing NIC bonding or multipathing, see *Designing XenServer Network Configurations*.

Tip: To determine what XenServer hosts have multipathing enabled on them, see **Multipathing** on the **General** tab in XenCenter.

Multipathing iSCSI Storage Traffic

For iSCSI software initiator storage traffic, it is important to consider redundancy. The iSCSI protocol was not designed for failure, resending packets, or data loss. Consequently, resiliency is very important. If there is a hardware issue with an iSCSI storage component, the throughput will be poor.

When working with iSCSI, it is important to design the infrastructure to be solid enough to support I/O. Ideally, iSCSI storage traffic should be sent over a separate physical network to prevent issues at the Ethernet level. If there are issues at the Ethernet level, storage transactions will be a problem.

Likewise, it is recommended that you use high quality switches that are on your storage vendor's hardware compatibility list and have been tested with your storage array.

When configuring iSCSI software initiator multipathing, there are two major points to consider:

- The type of multipath handler your array supports.
- Whether your array returns just one iSCSI Qualified Name (IQN) or multiple IQNs.

If you enable multipathing, you must do so on all hosts in the pool.



Understanding Multipathing

In order to understand the issues potentially associated from misconfiguring multipathing, it is important to know the purpose of multipathing, the function of the DM-MP handler in Linux, and how hosts establish links with storage repositories.

As previously stated, multipathing is a method of providing redundant *access* to storage devices if one or more components between the XenServer host and the storage array fail. (Multipathing protects against connectivity failures and not storage device failures.)

Multipathing creates multiple connections between the XenServer host and the storage controller; these connections are known as *paths*. When organizations configure multipathing, they are configuring multiple paths to a storage device (LUN) on a storage subsystem.

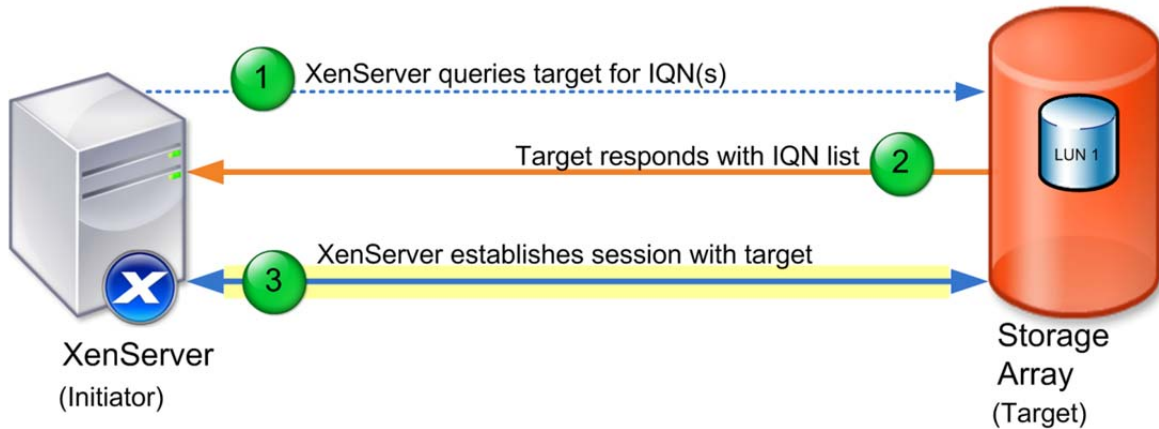
XenServer uses the DM-MP multipath handler in its multipathing implementation. The primary purpose of the DM-MP handler is that it creates a storage device for each LUN instead of creating a storage device for each *path*. That is, DM-MP reconciles multiple paths to a LUN so that Linux only creates one storage device in spite of seeing multiple paths.

Without DM-MP, Linux would create a storage device for each path to a LUN. This means in an environment with two paths to a LUN, Linux would create two storage devices. This would make it difficult to specify a storage device or find the path to that device.

However, for Linux to establish multiple active links, or *sessions*, to a LUN, Linux must use DM-MP so that it can treat multiple paths as representing only one LUN yet still be able to recognize both paths.

Establishing Multiple Active Links to a LUN

For a XenServer host to establish a link with the storage repository, it must, using iSCSI terminology, create a target and initiator connection. XenServer, the *initiator*, does so by querying the storage device, the *target*, and waiting for the target to reply saying it is available. After XenServer receives a list of target IQNs, XenServer, in its role as initiator, logs into the target. The target now has a link for sending traffic.



This illustration shows the process for creating a session. (1) XenServer, in its role as initiator, queries the storage target for a list of IQNs; (2) the storage device (target) responds with the list; and (3) after receiving the IQN list, XenServer establishes a session with the target.

This link (the session) remains up and only needs to be established if there is a reboot. After you configure both paths to the storage array (the multipath) and two paths are created, XenServer can create a session for each link, as shown in the following illustration.

Multipathing Enabled



This illustration shows how when multipathing is enabled XenServer creates two sessions with the storage target.

The multipath handler uses target IQNs to determine if the storage devices discovered on the target are different LUNs or different paths. The handler makes this determination by querying the storage target. The target replies with the IQN, which includes the LUN serial number. (Ideally, regardless of the number of paths connecting to a LUN, the serial number in the IQN is always the same.)

The multipath handler checks IQNs for matching serial numbers to determine how many paths are associated with each LUN. When the serial numbers in the IQNs match, the handler assumes that the IQNs are associated with the same LUN and therefore must represent different paths to that LUN.



When you create the storage repository and multipathing is enabled (specifically, when you create the Physical Block Device (PBD)), XenServer includes a *multihome* parameter that resets XenServer to expect a multihomed device. (The term multihome refers to computer or, in this case, a storage device that has multiple IP addresses connected to a network.)

If XenServer is not aware the storage device is multihomed (because multipathing was not enabled before the PBD/storage repository was created), XenServer can only create one session (or path) to the array.

For iSCSI arrays, it is better to configure multipathing first; however, if you created the storage repository first, you can put the host into maintenance mode and then configure multipathing. (For Fibre Channel, Citrix strongly recommends configuring multipathing and enabling the multipathing check box in XenCenter before creating the storage repository.)

With all types of SANs, it is best to plan and design your storage and networking configuration before implementation, determine you want multipathing, and configure it before putting the pool into production. Configuring multipathing after the pool is live results in a service interruption: configuring multipathing affects all VMs connected to the storage repository.

Chapter 2: Configuring Software Initiator iSCSI Multipathing

Overview

This chapter provides high-level steps for configuring iSCSI software initiator multipathing in XenCenter. It includes:

- Steps for configuring multipathing for iSCSI software initiator storage devices, including how to check if the target ports are operating in portal mode
- How to enable MPP RDAC Handler Support for LSI Arrays

Before Configuring iSCSI Multipathing

When configuring iSCSI software initiator multipathing, you must ensure that you use:

- The type of multipath handler your array requires
- The correct physical and subnet configuration

Both of these topics are described in more depth in the sections that follow. Review this information before setting up iSCSI multipathing and incorporate it into your design.



Multipath Handler Support

XenServer supports two different multipath handlers: Device Mapper Multipathing (DM-MP) and Multipathing Proxy Redundant Disk Array Controller (MPP RDAC). It also indirectly supports DMP RDAC.

By default XenServer uses Linux native multipathing (DM-MP), the generic Linux multipathing solution, as its multipath handler. However, XenServer supplements this handler with additional features so that XenServer can recognize vendor-specific features of storage devices. Typical examples of DM-MP arrays with portal mode include many NetApp arrays, HP StorageWorks Modular Smart Array (MSA) arrays, and HP StorageWorks Enterprise Virtual Array (EVA) arrays.

However, for LSI arrays, XenServer also supports LSI Multi-Path Proxy Driver (MPP) for the Redundant Disk Array Controller (RDAC). MPP RDAC works differently than the DMP handler and requires steps be performed in a specific sequence. As of XenServer 5.6 Feature Pack 1, XenServer now includes a CLI command for enabling MPP RDAC support that performs this sequence of steps for you in the correct order. Typical examples of MPP RDAC arrays include Dell MD3000-series arrays, IBM DS4000-series arrays, and others mentioned in [CTX122824](#).

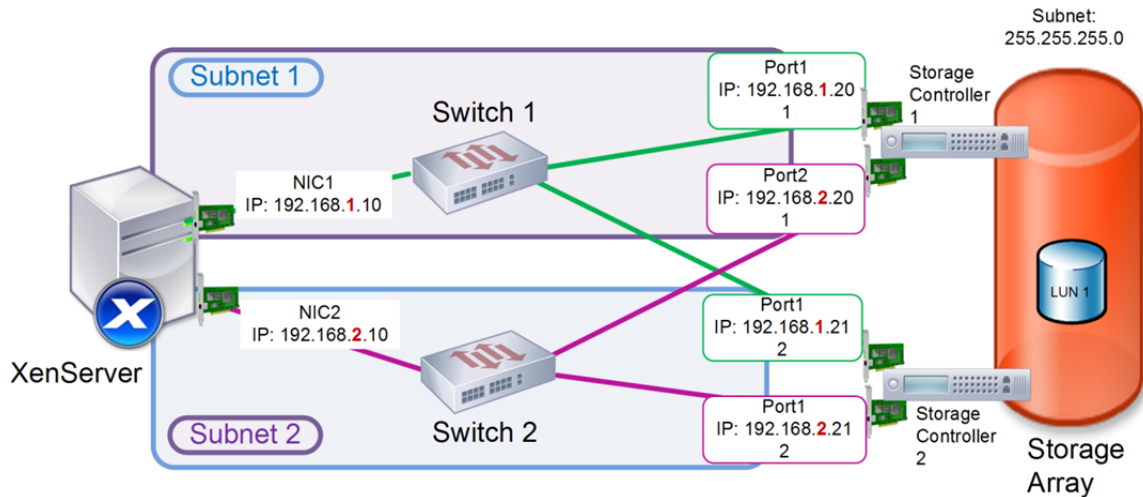
Citrix recommends users consult their array-vendor documentation or best-practice guide to determine precisely which multipath handler to select. If in doubt, LSI array data paths typically work best with the MPP RDAC drivers.

Correct Physical and Subnet Configuration

Setting up your storage solution correctly, including switch configuration, is critical for multipathing success.

The best-practice recommendation for failover is to use two switches; however, even if you do not configure two switches, you must logically separate the paths (that is, put each NIC on the host on separate subnets). Because TCP/IP acts as the transport protocol for iSCSI storage traffic, correct IP configuration is essential for iSCSI multipathing to work.

To add IP addresses to NICs, you must configure a *secondary interface* for storage. For more information, see the *XenCenter Help*.



This illustration shows how both NICs on the host in a multipathed iSCSI configuration must be on different subnets. In this illustration, NIC 1 on the host along with Switch 1 and NIC 1 on both storage controllers are on a different subnet than NIC2, Switch 2, and NIC 2 on the storage controllers.

In addition, one NIC on each storage controller must be on the same subnet as each NIC. For example, in the illustration that follows, XenServer NIC 1 is on the same subnet as NIC1 on Storage Controller 1 and NIC 1 on Storage Controller 2.

After performing the physical configuration, you enable and disable storage multipathing *support* in XenCenter in the **Multipathing** tab on the host's **Properties** dialog. It is easier to do this before you create the storage repository. The overall process for enabling multipathing requires a series of tasks, which is shown in the procedure that follows.

Important: Do not route iSCSI storage traffic through the XenServer host's management interface.

Configuring iSCSI Software Initiator Multipathing

Important: Citrix recommends either (a) enabling multipathing in XenCenter before you connect the pool to the storage device or (b) if you already created the storage repository, putting the host into Maintenance Mode before you enable multipathing.

If you enable multipathing while connected to a storage repository, XenServer may not configure multipathing successfully. If you already created the storage repository and want to configure multipathing, put all hosts in the pool into Maintenance Mode before configuring multipathing and then configure multipathing on all hosts in the pool. This ensures that any running virtual machines that have virtual disks in the affected storage repository are migrated before the changes are made.

The following procedure assumes you are using XenCenter.

To configure multipathing for iSCSI software initiator

1. Create the redundant physical paths (that is, set up the cables, switches, and subnets before configuring any storage settings in XenServer, including creating your storage repository.
 - a) Make sure that each NIC on the host is on a different subnet as shown in the diagram.
 - b) On each controller on the storage array, put one of the NICs on one of those subnets. (For example, make sure that on Controller A, NIC 1 is on Subnet 1 and NIC 2 is on Subnet 2. Likewise, on Controller B, make sure that NIC 1 is on Subnet 1 and NIC 2 is on Subnet 2.)
2. Follow any vendor multipathing configurations specific to your storage device and create any LUNs you will require.

Important: Make sure the iSCSI target and all servers in the pool must not have the same IQN set. It is imperative that every iSCSI target and initiator have a unique IQN. If a non-unique IQN identifier is used, data corruption can occur and/or access to the target may be denied.

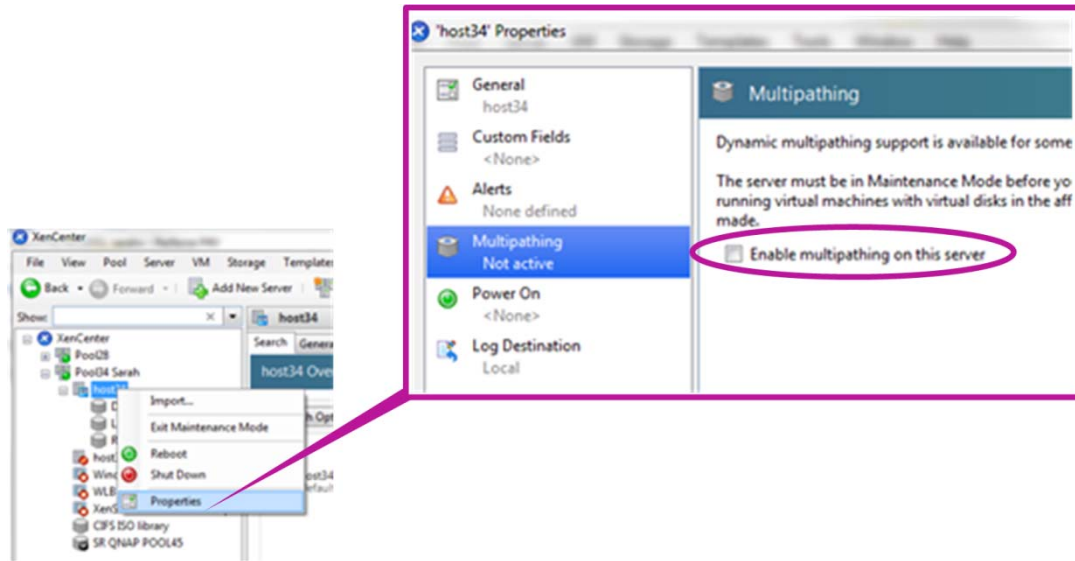
3. Verify that the iSCSI target ports are operating in portal mode:
 1. In XenCenter, start the **New Storage Repository** wizard (**Storage** menu > **New Storage Repository**).
 - i. Click through the options until you reach the **Enter a name and path for the new iSCSI storage** page, click **Discover IQNs**. XenServer queries the storage array for a list of IQNs.
 - ii. Check the **Target IQN** list box on the **Location** page of the Storage Repository Wizard.

If the iSCSI target ports are operating in portal mode, all target IPs should show up in the on the **Location** page of the Storage Repository Wizard.

4. Do one of the following to enable support for your multipath handler:
 - For MPP RDAC, enable multipathing in XenCenter (that is, select the **Enable multipathing on this server** check box in the **Multipathing** tab on the host's **Properties** dialog) and see "Enabling MPP RDAC Handler Support for LSI Arrays" on page 16.

Note: Citrix recommends users consult the array vendor's documentation or best-practice guide to determine precisely which multipath handler to select. If in doubt, LSI array data paths will typically work best with the MPP RDAC drivers.

- For DMP, enable multipathing in XenCenter (that is, select the **Enable multipathing on this server** check box in the **Multipathing** tab on the host's **Properties** dialog). Enabling multipathing is shown in the following illustration:



*This screenshot show the XenCenter **Enable multipathing on this server** check box.*

Enabling MPP RDAC Handler Support for LSI Arrays

To enable support for the MPP RDAC handler perform the following procedure on each host in the pool.

To enable the MPP RDAC handler

1. Open a console on the host, and run the following command:

```
# /opt/xensource/libexec/mpp-rdac --enable
```
2. Reboot the host.

Chapter 3: Creating the Storage Repository

Overview

This chapter explains how to create a storage repository after enabling iSCSI multipathing, including:

- How to select the most appropriate options in XenCenter based on the number of Target IQNs your array returns
- How to interpret the IQN as XenCenter displays it
- How to use the wildcard masking option in XenCenter

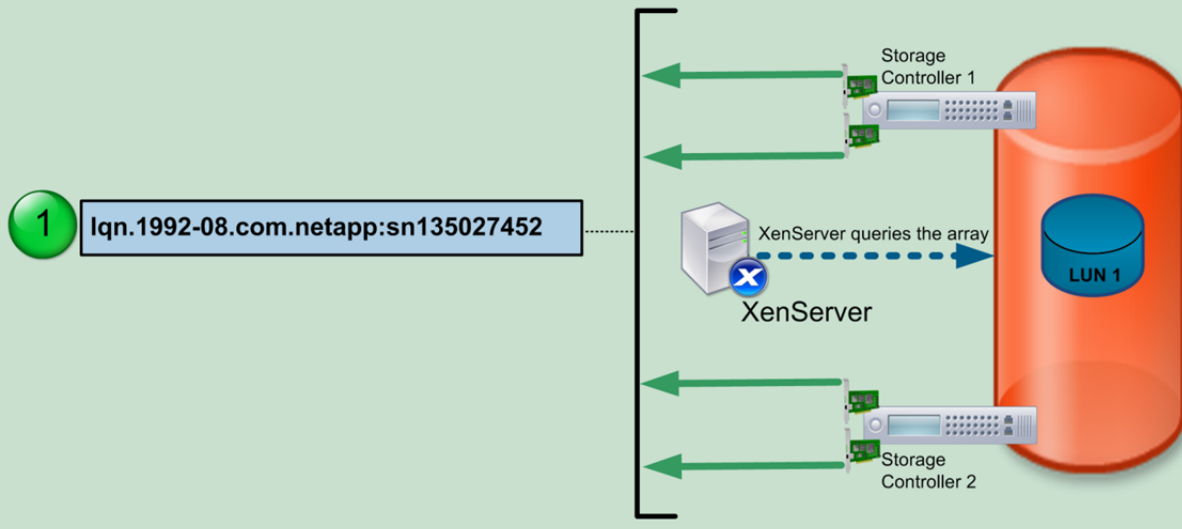
Creating the Storage Repository after Enabling Multipathing

Ideally, you should create the storage repository after configuring multipathing on all hosts in the pool. However, when you create the storage repository for hosts with multiple paths, you must be sure to select the correct option for the number of Target IQNs your array returns.

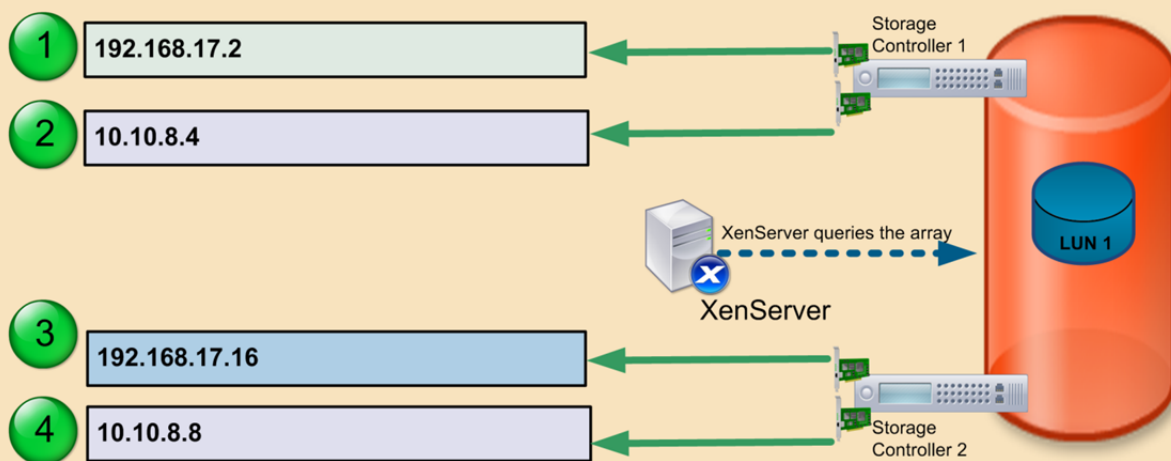
When you create a storage repository, XenServer requires that you provide it with the target IQN, which is the address used to identify the iSCSI storage device. However, when you query the storage device for its IQN, it may return one or more IQNs, depending on the specific device.

After multipathing is enabled, both paths should see the same LUN and the target should only return one IQN when XenServer queries it. However, even though multipathing is enabled, some storage vendors treat each path as a different LUN and return multiple IQNs. The illustration that follows compares two arrays, one that returns one IQN and one that returns multiple IQNs.

Target returns one IQN



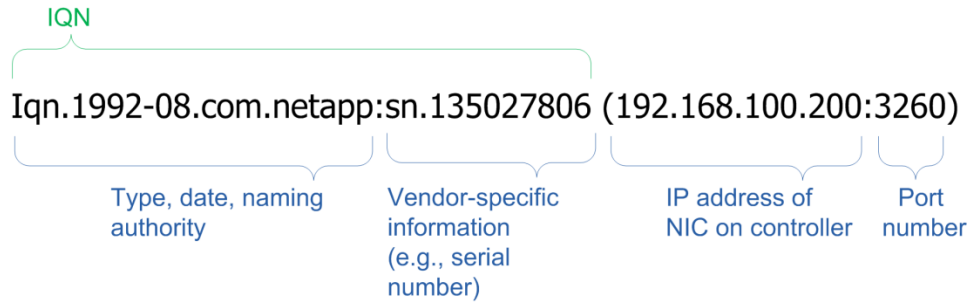
Target returns multiple IQNs



This diagram shows how some iSCSI storage devices (targets) return only one IQN whereas other iSCSI storage devices return multiple IQNs when XenServer queries the target storage device for the IQN.

Looking at the IQNs in the Storage Repository wizard, may be slightly confusing since XenCenter displays additional information. Unlike with the `xe sr-probe` command, the IQN displayed in XenCenter also includes the IP address of the NIC on the controller and the port number the host and array use for communication.

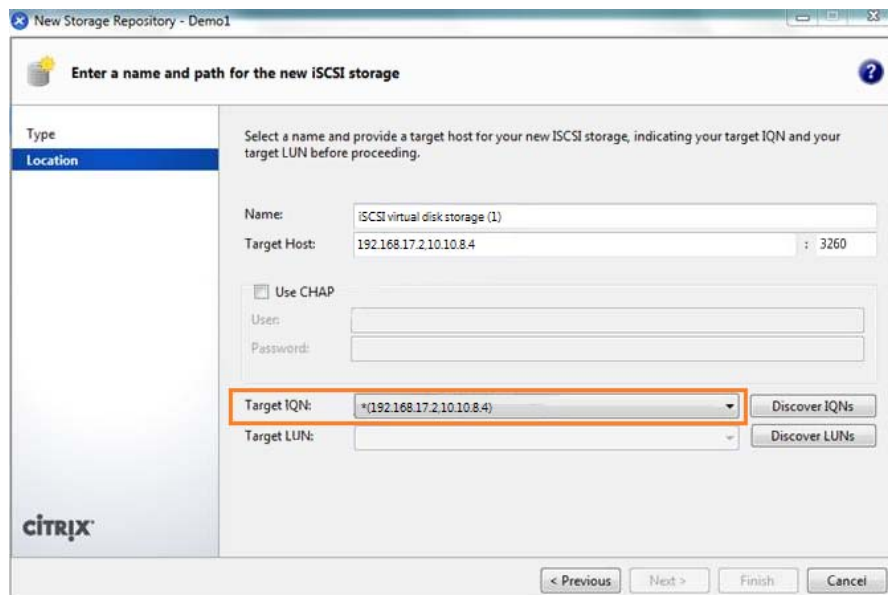
To determine if target query (that is, clicking **Discover IQN**) returned one or more IQNs, note the following illustration of a sample IQN value XenCenter might display:



This illustration shows a sample IQN number, as it might be displayed in XenCenter. The first part of the IQN is vendor specific and does not change. It includes the date the storage vendor applied for the IQN number and the storage vendor name. The second part of the IQN number includes vendor specific information. This is the part of the IQN to examine when determining whether or not to select the Wild Masking () option in the XenCenter Storage Repository wizard.*

Some storage devices, such as DataCore and StarWind arrays, have a multi-IQN feature and return multiple IQNs. For these arrays, you must specify the IP addresses of all targets, in the **Target Host** box on the **Location** page of the New Storage Repository wizard, and separate them with a comma. After you enter the IP addresses, if you press **Discover IQNs**, XenCenter will return multiple IQNs with different names.

For storage devices that return multiple IQNs, when you are creating the storage, you must select the **Wildcard Masking (*)** option in XenCenter, which is denoted by an asterisk in front of the Target IQN in the **Target IQN** list box. The wildcard masking option appears as follows:



This screen capture shows the XenCenter Storage Repository wizard's Wildcard Masking () option, which is used for storage that require multiple IQN support. After entering all the IP addresses of the targets separated by commas, the Storage Repository wizard returns multiple IQNs with different IQN numbers.*



To create a storage repository after enabling multipathing

1. In XenCenter, create the new storage repository for the pool using the New Storage Repository wizard.
2. Click through the options until you reach the **Enter a name and path for the new iSCSI storage** page. Fill out the options in this page until you reach **Discover IQNs**.
3. Click **Discover IQNs**. XenServer queries the storage array for a list of IQNs. Do one of the following:
 - If only one identical IQN value is returned, select the IQN on the storage device (target) that corresponds with the LUN you want to configure.
 - If you are using a storage array that returns multiple distinct IQNs, as described previously in this chapter, select the * option from the **Target IQN** list box.
4. Finish creating the storage repository and exit the wizard.



Revision History

Revision	Comments	Date
1.0	Initial Release	April 28, 2011
1.1	<p>Changed our recommendations about MPP RDAC and DMP drivers, especially for LSI arrays.</p> <p>Corrected step #4 in the “To configure multipathing for iSCSI software initiator” procedure to instruct people to <i>select</i> the Enable multipathing on this server check box. Previously, it said to leave this check box cleared.</p> <p>Changed information about the arrays that return multiple IQNs, and noted the need to enter multiple IP addresses for such arrays.</p>	July 15, 2011
1.2	Redrew multi-IQN illustration so that it aligns with example in text. Removed statement about not using the MPP RDAC handler with LSI arrays. Added a third reason to using bonding instead of multipathing. Minor editorial changes.	July 25, 2011
1.3	Clarified names of handlers, clarified guidance for when to use DMP and MPP RDAC. Minor editorial changes.	October 18, 2011
2.0	Updated terminology for XenServer 6.1. Additional minor changes.	September 2012



About Citrix

Citrix Systems, Inc. (NASDAQ:CTXS) is the leading provider of virtualization, networking and software as a service technologies for more than 230,000 organizations worldwide. Its Citrix Delivery Center, Citrix Cloud Center (C3) and Citrix Online Services product families radically simplify computing for millions of users, delivering applications as an on-demand service to any user, in any location on any device. Citrix customers include the world's largest Internet companies, 99 percent of Fortune Global 500 enterprises, and hundreds of thousands of small businesses and prosumers worldwide. Citrix partners with over 10,000 companies worldwide in more than 100 countries. Founded in 1989, annual revenue in 2010 was \$1.87 billion.

©2011 Citrix Systems, Inc. All rights reserved. Citrix®, Access Gateway™, Branch Repeater™, Citrix Repeater™, HDX™, XenServer™, XenCenter™, XenApp™, XenDesktop™ and Citrix Delivery Center™ are trademarks of Citrix Systems, Inc. and/or one or more of its subsidiaries, and may be registered in the United States Patent and Trademark Office and in other countries. All other trademarks and registered trademarks are property of their respective owners.