Technical Report

MetroCluster Upgrade Planning Guide

Jim Lanson, NetApp

April 2010 | TR-3517

ABSTRACT

This technical report is design to assist in planning the various forms of upgrades relative to NetApp® FAS MetroCluster. It includes planning considerations, checklists, and technical tips surrounding the completion of a successful upgrade. This document is intended to be a supplement to the standard active-active configuration guide. It is not a step-by-step guide to actually performing the upgrade.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 PURPOSE

The purpose of this document is to present a guide for planning various types of upgrades of a FAS MetroCluster configuration. It does not presently include V-Series configurations of MetroCluster. This document is intended to be a supplement to the standard Data ONTAP® Upgrade Guide and the Active-Active or High-Availability Configuration Guides. It is not a step-by-step guide to actually performing the upgrade.

The type of upgrade under consideration and a number of other factors need to be considered in the planning process. These upgrades and the considerations necessary to planning a successful upgrade are discussed in the following sections.

1.2 INTENDED AUDIENCE

This technical report is designed for storage administrators and architects who are already familiar with Data ONTAP administration, active-active or HA configurations, and MetroCluster and are considering deployments for production environments. This guide should be of equal interest to both NetApp and partner system engineers along with professional services personnel.

1.3 DOCUMENTATION

Active-Active or High-Availability Configuration Guide (NOW™ [NetApp on the Web] site)

Data ONTAP Upgrade Guide (NOW site)

MetroCluster Compatibility Matrix (NOW site)

Brocade Switch Configuration Guide (NOW site)

TR-3548: MetroCluster Best Practices for Design and Implementation (Technical Library on www.netapp.com)

2 TYPES OF UPGRADES

There are many variations of upgrades involving a NetApp FAS MetroCluster configuration. There are also several keys to planning a successful upgrade which include one or more of the following:

- Obtaining the necessary hardware components
- Obtaining the necessary software licenses
- Scheduling downtime (if necessary)
- Verification of hardware and software compatibility

For each of the upgrades described in the following sections a checklist is provided in order make sure all needs are addressed.

2.1 ACTIVE-ACTIVE TO STRETCH METROCLUSTER

This upgrade includes the conversion of a standard active-active or HA pair to a stretch MetroCluster configuration. It is assumed that the same controllers and disk shelves will be used.

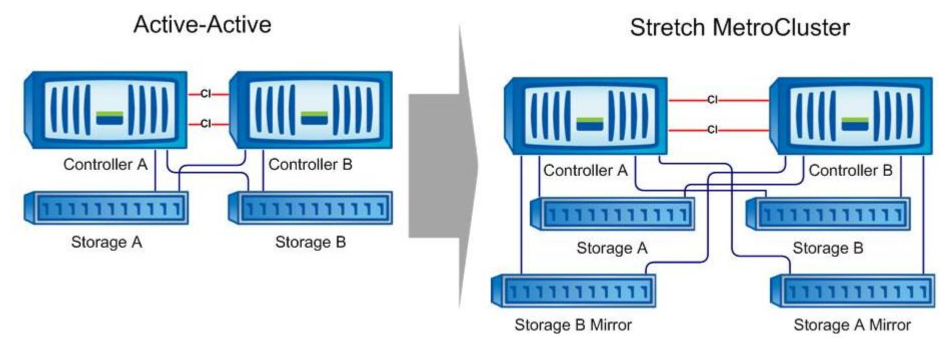


Figure 1) Active-active to stretch MetroCluster.

HARDWARE

- Additional shelves for mirroring

Figure 1 shows the addition of disk shelves to both sides assuming that this is truly an active-active configuration. If the original pair was not running applications on both sides, in other words, was running active-passive, then shelves for mirroring would only be required on the passive controller in order to protect the active controller. In a stretch MetroCluster configuration, both SATA and FC drive shelves are supported with the same cautions as exist on a regular active-active configuration.

- Additional cables for disk shelves

Two cables will be required for each additional shelf, including the last on each loop if using multipath HA. Keep in mind that MPHA does add resiliency to a stretch MetroCluster configuration.

- Disk and shelf firmware compatibility?

Inventory the current disk shelves to verify that firmware levels are compatible with the MetroCluster Compatibility Matrix located on the NetApp NOW site. If they are not, then the firmware will have to be upgraded as part of the process.

- Additional chassis necessary (if upgrading 31xx)?

The architecture of the FAS3100 series accommodates two controllers per physical chassis. Since part of the value of MetroCluster is the physical separation of the controllers, a separate chassis is needed to accommodate one of the controllers.

- FC/VI card required (if upgrading 31xx)?

If the current active-active configuration is one of the 3100 series of systems with two controllers per chassis, then the internal cluster interconnect can no longer be used. A separate FC/VI card must be installed in each controller as part of the upgrade process. This card, when installed, disables the internal backplane cluster interconnect. This card is described later in this document.

- Proper FC cabling

Make sure you have the proper length, number, and type of Fibre Channel cables for interconnecting the disk paths and the two cluster interconnect ports between controllers. Also remember the connector types, especially if the connections between MetroCluster nodes will traverse through patch panels of any kind.

- Copper to fiber adapters (non-31xx) for cluster interconnect

In the case of a stretch MetroCluster configuration using a non-3100 series platform, the existing cluster interconnect on the NVRAM card may be used. In order to accommodate the increased distances, the copper interface needs to be converted to fiber so that FC cabling may be used.

- Maximum spindle count = platform maximum

The total maximum number of disk spindles for a FAS stretch MetroCluster configuration is equal to whatever the limit is for that platform. Refer to the System Configuration Guide on the NOW site for the platform being used.

SOFTWARE

- CF-Remote and SyncMirror licenses

Assuming this was an active-active or HA configuration, the Cluster license should already be installed on both systems. In that case all that has to be ordered and installed are the SyncMirror_local and Cluster_remote licenses. Remember without all three it is not a MetroCluster configuration.

- Will the Data ONTAP version change? Check compatibility.

If a Data ONTAP operating system upgrade is planned, make sure it is compatible based on the MetroCluster Compatibility Matrix located on the NOW site.

2.2 ACTIVE-ACTIVE TO FABRIC METROCLUSTER

This upgrade includes the conversion of a standard active-active or HA pair to a stretch MetroCluster configuration. It is assumed that the same controllers will be used.

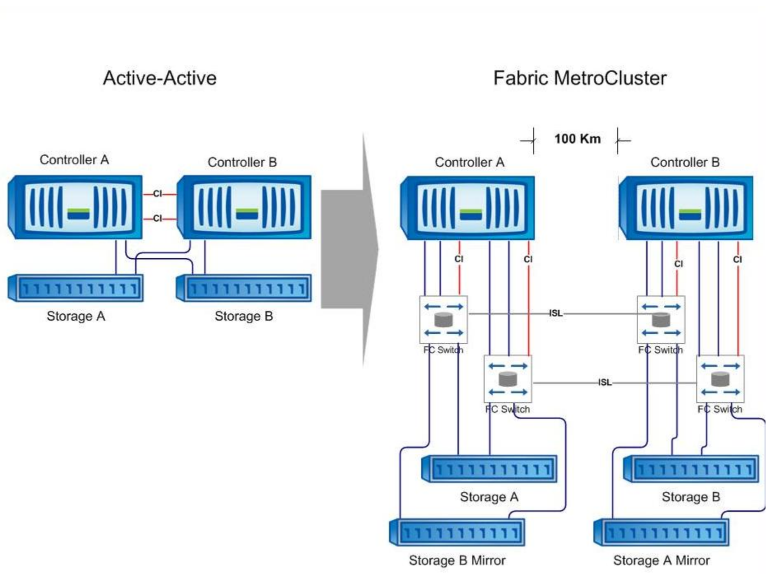


Figure 2) Active-active to fabric MetroCluster.

Hardware

- Additional shelves for mirroring

Figure 2 shows the addition of disk shelves to both sides assuming that this is truly an active-active configuration. If the original pair was not running applications on both sides, in other words, was running active-passive, then shelves for mirroring would only be required on the passive controller in order to protect the active controller. In a fabric MetroCluster configuration, only FC drive shelves are supported.

NOTE: Remember that SATA disks are not supported on a fabric MetroCluster configuration so additional FC shelves might be necessary in order to migrate the data from the SATA disks.

- Fibre Channel switches

Four dedicated Fibre Channel (FC) switches will be required for the fabric MetroCluster configuration. Two will be located at each site. The switches must be supplied by NetApp. There will be redundant FC fabrics with one switch at each location as shown in Figure 2. Verify that the switches chosen are compatible with the controllers and Data ONTAP version (see MetroCluster Compatibility Matrix). Also make sure that the firmware installed on the switches is compatible with the MetroCluster matrix. The specific switch needed will depend on several factors described later in this document.

- Fibre Channel HBAs

Depending on the exact configuration you might need to add FC HBAs. All of the currently shipping platforms have on-board FC ports. However, if these controllers are also connecting to a front-end SAN, then more ports might be necessary. Consult the System Configuration Guide on the NOW site for specific models and support. Keep in mind that many on-board ports are capable of only 2Gbps speeds, so a 4GBPS HBA might be necessary if that speed is required.

- Fibre Channel switch interswitch links

The fabric MetroCluster configuration in Figure 2 shows two interswitch links (ISLs). These FC connections carry the cluster interconnect and disk traffic between nodes. If the traffic isolation feature is being used, there will be a total of four ISLs requiring 4 fiber connections (traffic isolation separates the disk traffic from the cluster interconnect traffic). The distance between the two locations will determine the type of cabling, the Small Form-Factor pluggables (SFPs) needed for the ISLs, and possibly the switch model itself. Refer to the distance considerations section of this document for further information.

- Customer supplied wave division multiplexors

Since the cost of fiber cabling is expensive, customers will often deploy multiplexing devices in order to use the fiber for multiple purposes. NetApp MetroCluster supports whatever device is used as long as it is listed on Brocade's Compatibility Guide for the model switch being used. This guide may be found at www.brocade.com.

- Additional cables for controllers and disk shelves

In a standard active-active configuration, there can be many shelves connected in a physical disk loop. In a fabric MetroCluster configuration there is a limit of two physical disk shelves per loop so accommodation might have to be made for additional loops and additional Fibre Channel switch ports (1 loop [2 shelves] per port).

Since the controller FC ports, FC/VI cards, and disk shelves all have physical connection to both of the FC fabrics, additional cables will be required. Use Table 1 to calculate.

Table 1) Cable requirements.

Cable	Qty Local	Qty Remote
Controller FC/VI to local switches[1]	2	2
Controller FC ports (on-board or HBA) to local switches[1]	4	4
Interswitch link cables[2]	2 4 (with traffic isolation)	2 4 (with traffic isolation)
Disk shelves[1]	2	2

[1] Length is dependent on distance to switches or previous shelf.

[2] Length is dependent on distance to other node, patch panel, or WDM device if used.

- Disk and shelf firmware compatibility?

Inventory the current disk shelves to verify that firmware levels are compatible with the MetroCluster Compatibility Matrix located on the NetApp NOW site. If they are not, then the firmware will have to be upgraded as part of the process.

- Additional chassis necessary (if upgrading 31xx)?

The architecture of the FAS3100 series accommodates two controllers per physical chassis. Since part of the value of MetroCluster is the physical separation of the controllers, a separate chassis is needed to accommodate one of the controllers.

- FC/VI card required

If the current active-active configuration is one of the 3100 series of systems with two controllers per chassis, then the internal cluster interconnect can no longer be used. A separate FC/VI card must be installed in each controller as part of the upgrade process. This card, when installed, disables the internal backplane cluster interconnect. This card is described later in this document.

- Maximum spindle count

The total maximum number of disk spindles for a FAS fabric MetroCluster configuration is equal to the lesser of whatever the limit is for that platform or 672.

Software

- CF-Remote and SyncMirror licenses

Assuming this was an active-active or HA configuration, the Cluster license should already be installed on both systems. In that case all that has to be ordered and installed is the SyncMirror_local and Cluster_remote licenses. Remember without all three it is not a MetroCluster configuration.

- Will the Data ONTAP version change? Check compatibility

If a Data ONTAP operating system upgrade is planned, make sure it is compatible based on the MetroCluster Compatibility Matrix.

- If older FAS series (3020, 3050), then plan for possible conversion from hardware disk ownership to software disk ownership

If the current active-active or HA pair is based on the older 3020 or 3050 then the disk ownership model will have to be changed in order to upgrade to a fabric MetroCluster configuration. Make sure that these controllers will be compatible with the Data ONTAP version and switches being used.

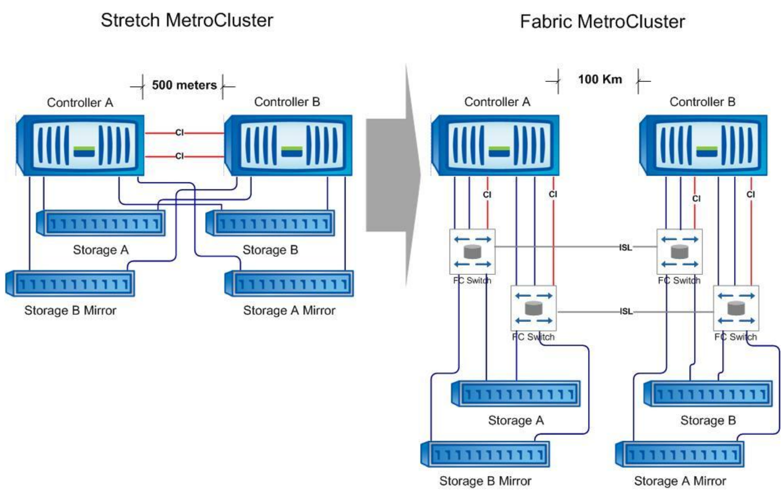
- Licenses for switches

It is extremely important that the proper licenses for the switches be ordered and obtained prior to the start of the upgrade. The following are the minimum required licenses needing to be installed on each switch:

- Full Fabric license (without this there will be no E-ports for ISLs)
- Ports on demand license (for however many ports are needed)
- Web license (Web-based switch tool)
- Zoning license (required for disk spindle configurations of 512 or greater)

Optionally, if the distance between the nodes is greater than 10km, an Extended Fabric license is also required. This enables setting of the proper buffer credits for the ISLs.

2.3 STRETCH TO FABRIC METROCLUSTER



HARDWARE

- Additional shelves for mirroring

Remember that SATA disks are not supported on a fabric MetroCluster configuration so additional FC shelves might be necessary in order to migrate the data from the SATA disks.

- Fibre Channel switches

Four dedicated Fibre Channel switches will be required for the fabric MetroCluster configuration. Two will be located at each site. The switches must be supplied by NetApp. There will be redundant FC fabrics with one switch for each fabric at each location, as shown in Figure 3. Verify that the switches chosen are compatible with the controllers and Data ONTAP version (see MetroCluster Compatibility Matrix). Also make sure that the firmware installed on the switches is compatible with the MetroCluster matrix. The specific switch needed will depend on several factors described later in this document.

- Fibre Channel switch interswitch links

The fabric MetroCluster configuration in Figure 2 shows two interswitch links (ISLs). These FC connections carry the cluster interconnect and disk traffic between nodes. If the traffic isolation feature is being used there will be a total of four ISLs requiring 4 fiber connections (traffic isolation separates the disk traffic from the cluster interconnect traffic). The distance between the two locations will determine the type of cabling, the SFPs needed for the ISLs, and possibly the switch model itself. Refer to the distance considerations section of this document for further information.

- Customer supplied wave division multiplexors

Since the cost of fiber cabling is expensive, customers will often deploy multiplexing devices in order to use the fiber for multiple purposes. NetApp MetroCluster supports whatever device is used as long as it is listed on Brocade's Compatibility Guide for the model switch being used. This guide may be found at www.brocade.com.

- Additional cables for controllers and disk shelves

In a standard active-active configuration, there can be many shelves connected in a physical disk loop. In a fabric MetroCluster configuration there is a limit of two physical disk shelves per loop so accommodation might have to be made for additional loops and additional Fibre Channel switch ports (1 loop [2 shelves] per port).

Since the controller FC ports, FC/VI cards, and disk shelves all have physical connection to both of the FC fabrics, additional cables will be required. Use Table 1 to calculate.

- Disk and shelf firmware compatibility

Inventory the current disk shelves to verify that firmware levels are compatible with the MetroCluster Compatibility Matrix located on the NetApp NOW site. If they are not, then the firmware will have to be upgraded as part of the process.

- FC/VI card required

If the current active-active configuration is one of the 3100 series of systems with two controllers per chassis, then the internal cluster interconnect can no longer be used. A separate FC/VI card must be installed in each controller as part of the upgrade process. This card, when installed, disables the internal backplane cluster interconnect. This card is described later in this document.

- Maximum spindle count

The total maximum number of disk spindles for a FAS fabric MetroCluster configuration is equal to the lesser of whatever the limit is for that platform or 672.

<u>Software</u>

- Will the Data ONTAP version change? Check compatibility

If a Data ONTAP operating system upgrade is planned, make sure it is compatible based on the MetroCluster Compatibility Matrix.

- If older FAS series (3020, 3050), then plan for possible conversion from hardware disk ownership to software disk ownership

If the current stretch MetroCluster configuration is based on the older 3020 or 3050 then the disk ownership model will have to be changed in order to upgrade to a fabric MetroCluster configuration. Make sure that these controllers will be compatible with the Data ONTAP version and switches being used.

- Licenses for switches

It is extremely important that the proper licenses for the switches be ordered and obtained prior to the start of the upgrade. The following are the minimum required licenses needing to be installed on each switch:

- Full Fabric license (without this there will be no E-ports for ISLs)
- Ports on demand license (for however many ports are needed)
- Web license (Web-based switch tool)
- Zoning license (required for disk spindle configurations of 512 or greater)

Optionally, if the distance between the nodes is greater than 10km an Extended Fabric license is also required. This enables setting of the proper buffer credits for the ISLs.

2.4 CONTROLLER-ONLY UPGRADES

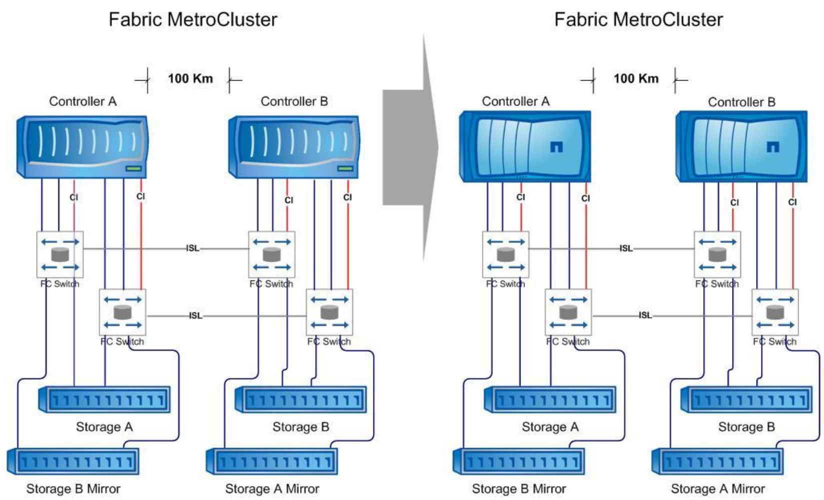


Figure 4) Controller-only upgrade.

HARDWARE

- Disk and shelf firmware compatibility

Inventory the current disk shelves to verify that firmware levels are compatible with the MetroCluster Compatibility Matrix located on the NetApp NOW site. If they are not then the firmware will have to be upgraded as part of the process.

- Additional chassis necessary (if upgrading to 31xx)?

The architecture of the FAS3100 series accommodates two controllers per physical chassis. Since part of the value of MetroCluster is the physical separation of the controllers, a separate chassis is needed to accommodate one of the controllers.

- FC/VI card required

Depending on the specific model controllers being replaced and the version of Data ONTAP to be used, a new FC/VI card might be necessary for each of the controllers. For example, the FC/VI card used on a 9xx or 3020/3050 is a 2Gbps card. The newer controllers and Data ONTAP operating system releases like 8.x do not support this card. A 4Gbps card will have to be ordered and installed as part of the upgrade process. These cards are described later in this document.

- Maximum spindle count

Depending on the specific model controllers being replaced, the type of MetroCluster configuration (stretch or fabric) and the version of Data ONTAP to be used, the maximum number of FC disks might change. Refer to the System Configuration Guide or the MetroCluster Compatibility Matrix on the NOW site for the platform being used.

- Will the Data ONTAP version change? Check compatibility

If a Data ONTAP operating system upgrade is planned, make sure it is compatible based on the MetroCluster Compatibility Matrix.

- If older FAS series (3020, 3050), then plan for possible conversion from hardware disk ownership to software disk ownership

If the current active-active or HA pair is based on the older 3020 or 3050 then the disk ownership model will have to be changed in order to upgrade to a fabric MetroCluster configuration. Make sure that these controllers will be compatible with the Data ONTAP version and switches being used.

3 DISTANCE CONSIDERATIONS

In order to determine the upgrade components necessary, the estimated distance between nodes needs to be determined. For example, with controllers running at 2Gbps, if the length of cable between the nodes is 500 meters or less, then a stretch MetroCluster configuration may be used with OM3 (or better) cabling. If 4Gbps speeds are used, the maximum distance drops from 500m to 270m. If the distance is greater than either of these maximums, then a fabric MetroCluster configuration must be chosen, making Fibre Channel switches necessary.

4 STORAGE REQUIREMENTS

In analyzing the additional storage requirements necessary, it is important to remember that ATA shelves are not supported on a fabric MetroCluster configuration. Typically in a MetroCluster configuration each aggregate of disks on a storage controller has a mirror on its partner controller. This includes the aggregate containing the root volume. Not all aggregates on a node need to be mirrored (other than the aggregate containing the root volume). However, any data contained on an aggregate not mirrored on its partner will obviously not be available in a failover situation. Figure 5 shows a typical Figure 5) active-active configuration with each node containing two shelves.

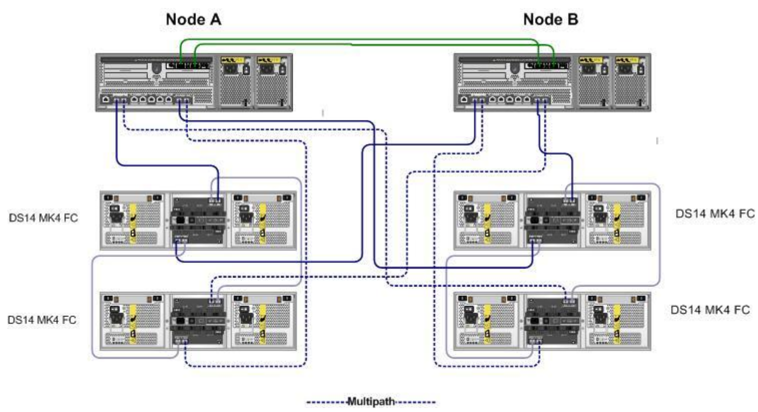


Figure 5) Standard active-active configuration.

If this configuration were to be converted to a stretch MetroCluster configuration, then four (4) additional shelves would be necessary, two for each controller. Figure 6 below shows the converted stretch MetroCluster configuration. Notice each controller now has two additional shelves on the opposite side. These contain the mirrored aggregates of their partner.

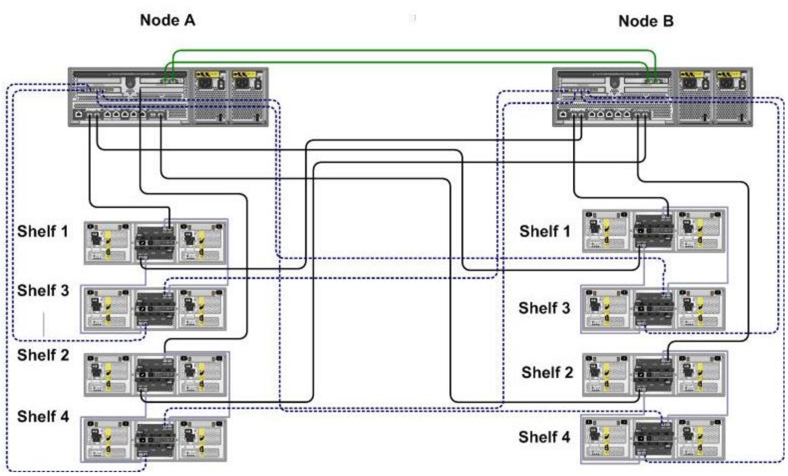


Figure 6) Stretch MetroCluster configuration.

In a fabric MetroCluster configuration the additional storage requirements remain the same, except that, as stated before, ATA drive shelves are not supported.

Additionally, for the FAS3020- and FAS3050-based active-active configuration, if software disk ownership was being used, it should be reverted back to the hardware ownership model prior to upgrading to a fabric MetroCluster configuration. The FAS3040, FAS3070, and FAS6000 series fabric MetroCluster configurations use software disk ownership. While use of the software disk ownership model provides more flexibility in terms of cabling, it requires more attention to detail in assigning disks/shelves to the proper pools for mirroring. Additional information may be found in TR-3548, “Best Practices for MetroCluster Design and Implementation.”

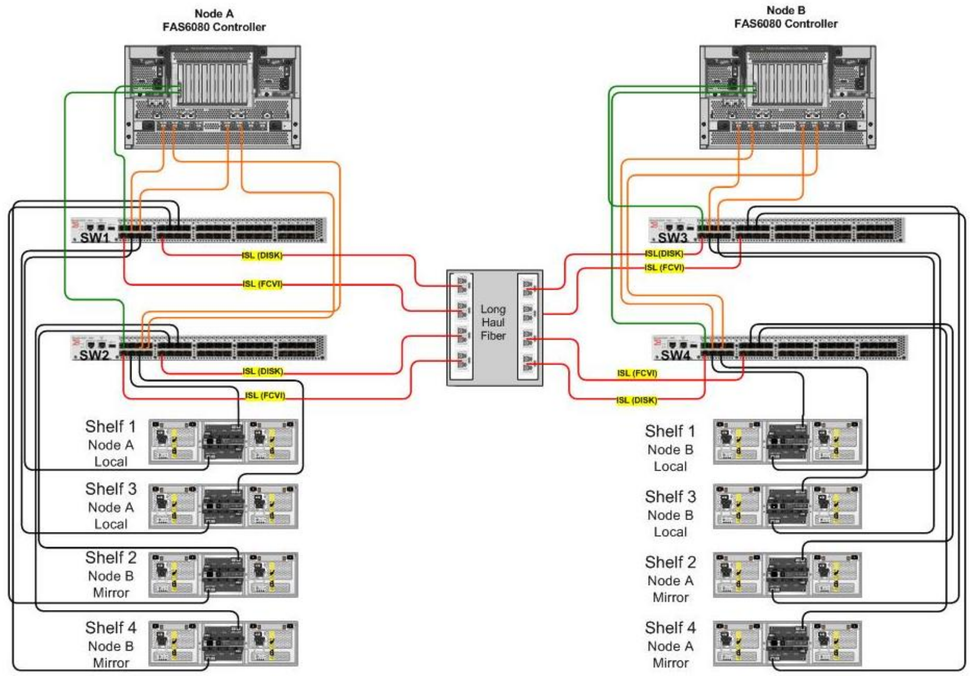


Figure 7) FAS6000 Series fabric MetroCluster configuration.

5 CLUSTER INTERCONNECT

Connectivity of the NVRAM5 or NVRAM6 cards is as follows:

For a stretch MetroCluster configuration (other than the 31xx series), the cluster interconnect embedded on the NVRAM card will continue to be used. A 4xIB MPO optical cluster interconnect cable is required in the stretch configuration.

For a fabric MetroCluster configuration or a 31xx stretch MetroCluster configuration, the cluster interconnect embedded on the NVRAM card may no longer be used. A separate FC/VI card will have to be installed. Consult the System Configuration Guide on the NOW site for the correct card and slot placement of both the NVRAM and FC/VI cards.

6 LICENSING

For either MetroCluster configuration, the following licenses are required on both nodes of the MetroCluster configuration:

- Cluster
- Cluster remote
- SyncMirror® local

Also remember that if upgrading to a fabric MetroCluster configuration, there are FC switch licenses that need to be installed. The following are the minimum required licenses needing to be installed on each switch:

- Full Fabric license (without this there will be no E-ports for ISLs)
- Ports on demand license (for however many ports are needed)
- Web license (Web-based switch tool)
- Zoning license (required for disk spindle configurations of 512 or greater)

Optionally, if the distance between the nodes is greater than 10km an Extended Fabric license is also required. This enables setting of the proper buffer credits for the ISLs.

7 CABLE CONSIDERATIONS

7.1 CABLING OVERVIEW

Stretched MetroCluster can support a maximum distance of either 500-meter with OM3 cable at 2Gbps or 270m at 4Gbps. Fabric MetroCluster through the use of Fibre Channel switches further extends this distance to approximately 100Km. This extended distance capability gives customers much greater flexibility in the physical location of the clustered controllers while maintaining the high-availability benefits of clustered failover.

The maximum permissible distance is a function of a combination of factors including speed, number of connections, cable type, and interconnection type.

To deploy clustered failover with extended distance, customers must adhere to the following requirements:

7.2 GENERAL NOTES

For cable length greater than 30 meters, customers must purchase cables directly from approved cable vendors.

In order to achieve the maximum distance between stretch MetroCluster controllers, the 4xIB MPO optical cluster interconnect cable must be a direct point-to-point connection with no intermediate device between them (such as a patch panel).

The maximum distance between stretch MetroCluster controllers will be reduced when the cluster interconnect is routed through an intermediate passive device (such as a patch panel) due to degradation in optical signal quality. The actual distance reduction will vary. The only practical way for establishing the maximum supported distance is to test the desired configuration in the actual customer environment. Cable runs should be tested end to end by a qualified installer to make sure they meet desired specifications. Refer to NetApp TR-3552, “Optical Network Installation Guide,” for additional information.

Do not mix cable types. The signal loss due to splicing will probably result in a total attenuation greater than specifications allow.

The minimum number of extended distance cables required is four (two for cluster interconnect, one for each remote loop from each clustered storage).

7.3 CLUSTER INTERCONNECT

FAS controllers have a copper interface for the cluster interconnect. Customers have the option of ordering a 2m, 5m, or 10m copper cable (with the 5m copper cable being the default). The integrated NVRAM/cluster adapter supports the 4xInfiniBand standard. For MetroCluster, fiber interconnect is necessary.

Customers who want to run either stretch or fabric MetroCluster will need to run fiber optic cable between the clustered controllers. Copper (Cu)-to-fiber converters are necessary in addition to the fiber cables. For each clustered pair, four X1949A-R6 converters are required. The part number is X1949A-R6 (description: Cu-to-fiber converter).

Figure 8 shows a picture of the Cu-to-fiber converter.

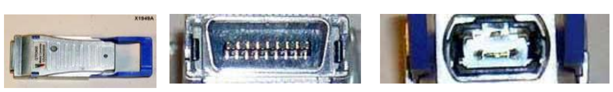


Figure 8) Cu-to-fiber converter.

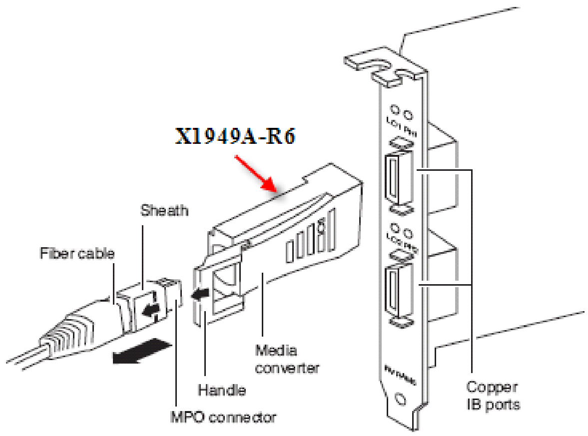


Figure 9) Cu-to-fiber converter installation.

Figure 9 shows installation of the converter.

For a stretch MetroCluster cluster interconnect, the fiber cables to be used can be MPO to MPO. The more likely scenario is that interconnection through patch panels will be required.

Depending on the type of connector (LC, SC, ST) that is needed, customers would need to buy four optical cables (of the same type) per cluster pair. Figure 11 shows the different types of connectors.

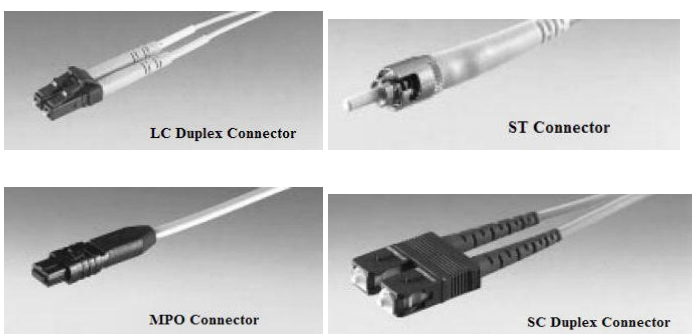


Figure 10) Connector types.

Customers now have the choice of ordering 30m 1xIB cables for select connectors from NetApp directly. The connectors for which cables are available and their part numbers are:

```
X1951A-R6      : for SC termination
X1952A-R6      : for ST termination
X1953A-R6      : for LC termination
```

Customers needing 1xIB cables of other lengths and for other connector types would need to order them directly from an authorized vendor and submit a PVR for NetApp Engineering validation. The part numbers for each type of connector and the vendor contact information are:

```
Part number:
AFL # FTF-PDL-P20Z-002-R-XX     : for LC termination
AFL # FTF-PSF-P20Z-002-R-XX     : for SC termination
AFL # FTF-PST-P20Z-002-R-XX     : for ST termination
```

Note: `XX` can be replaced with the exact length in meters of cables required. Example: FTF-PFC-P20Z-002-R-25 will provide a 25m 1xIB cable with an FC termination. A customer would need four of these per cluster (two per storage head).

Vendor info:

Company: Fujikura America
Contact: Cary Sakamoto
Phone: 408-988-7424
Fax: 408-727-3415
E-mail: Cary@Fujikura.com

7.4 DISK INTERCONNECT

The same distance considerations apply to the disk shelf cables. The length of B-loop cables connecting appliances to storage should match the length of the cluster interconnect.

The minimum number of disk cables required is two (one for each remote loop on each clustered appliance). Each additional loop requires at least one additional long cable. The number of disk cables doubles for a stretch MetroCluster configuration connected through patch panels or in a fabric MetroCluster configuration. (Some of the long cables connect to the "A" side of the shelf.)

Multipathing is supported on the stretch MetroCluster configuration, further maximizing resiliency. These cables will have to be considered also. Fabric MetroCluster already has redundant disk paths, so multipathing is not necessary.

Speed considerations. With the variety of disk shelves installed including the DS14-MK4 shelves and other 4Gbps devices, it is important to consider the possible mixed environment (2Gbps and 4Gbps) of an upgraded MetroCluster configuration. The same guidelines apply as in a standard active-active configuration. Also, remember that in the case of the current FAS6000 series, the on-board FC ports are only capable of 2Gbps operation. For those customers requiring a full 4Gbps solution, separate 4Gbps FC host adapters must be used.

7.5 CABLING DISTANCE CALCULATIONS AND OPTICAL LINK DESIGN

This section provides a mechanism to determine if the desired optical cable length is within the supported maximum specification. This is provided as a general guidance to determine the distances. The best way to really optimize a network design is by testing and characterizing the actual network environment (as there are many variables like types of cables, panels, and so on).

Example: A customer needs to run an optical cable and transceivers rated at 850nm and wants to operate at 2Gbps. How do we determine the maximum distance that can be allowed for cabling between nodes?

Table 2) Optical cable parameters chart.

Cable Type	Fibre Type	Mode	Wavelength (nm)	Maximum Distance (m)	Attenuation (db/km)	Splice Loss (db)	Connector Pair Loss (dB)
1 Gbps							
OM2	50/125 um	Multi	850	500	3	3	.75
OM3	50/125 um	Multi	850	860	3	3	.75
OM3+	50/125 um	Multi	850	1,100	3	3	.75
OS1*	9/125 um	Single	1,310	2,000	.4	.3	.75
2 Gbps							
OM2	50/125 um	Multi	850	300	3	3	.75
OM3	50/125 um	Multi	850	500	3	3	.75
OM3+	50/125 um	Multi	850	750	3	3	.75
OS1*	9/125 um	Single	1,310	2,000	.4	.3	.75
4Gbps							
OM2	50/125 um	Multi	850	150	3	3	.75
OM3	50/125 um	Multi	850	270	3	3	.75
OM3+	50/125 um	Multi	850	500	3	3	.75
OS1*	9/125 um	Single	1,310	500	.4	.3	.75
8Gbps							
OM2	50/125 um	Multi	850	50	3	3	.75
OM3	50/125 um	Multi	850	150	3	3	.75
OM3+	50/125 um	Multi	850	221	3	3	.75
OS1*	9/125 um	Single	1,310	500	.4	.3	.75
IB 1X 250 MB/sec							
OM2	50/125 um	Multi	850	300	3.5	3	.75
OM3	50/125 um	Multi	850	500	3.5	3	.75

Table 1 summarizes "typical" data related to optical cabling for data communications available in documents published by various standards organizations. The distances shown are calculated for worst-case fiber and transceiver characteristics. The focus is on data that is relevant to fiber deployments supported on NetApp systems.

To determine whether the desired optical cable length is within the supported maximum specification:

Determine the needed transfer rate based on the type of shelf being used.

Find out what fiber (http://now.netapp.com/NOW/knowledge/docs/san/guides/CFO_cable/cabledistance.shtml - 4#4) type is installed in the customer network for the system.

Determine the type of optical interfaces (transceivers) provided with (or recommended for) the equipment to be connected.

Determine the number of connectors and splices in the path between nodes.

The maximum optical link length is the shorter of the following two distance limitations:

- Distance limitation due to optical power budget (Figure 8)
- Distance limitation due to the bandwidth limitation of the type of fiber used (expressed as Max Distance in Table 1)

DISTANCE LIMITATION DUE TO AVAILABLE OPTICAL POWER BUDGET

To design a fiber optic link, one needs to analyze the so-called optical link loss budget against the available optical power budget. Figure 12 illustrates the required optical calculations for designing a fiber link.

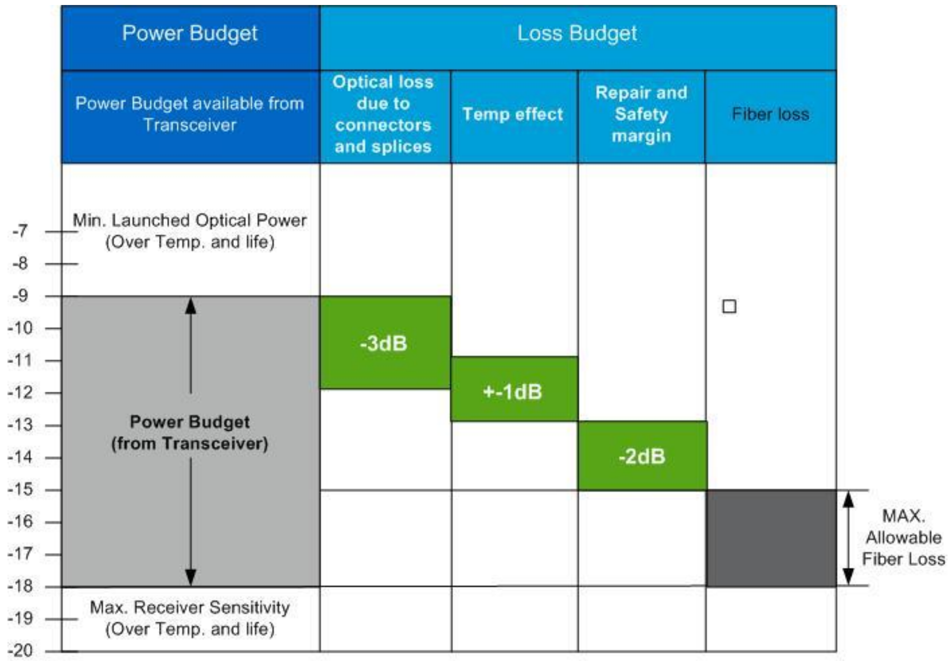


Figure 11) Illustration of key contributors to optical link loss budget.

The link example considered here consists of 2Gbps (2xFC) SFP transceivers operating at 850nm, 300 meters of OM2 fiber, and two patch panels, as depicted in Figure 10.

The illustration in Figure 12 starts on the left side with the optical power budget available from the transmit and receive ports of the transceivers specified for the link. The difference between the minimum launch power and the worst-case sensitivity of the optical receiver constitutes the minimum available optical power budget. Worst-case optical launch power and receiver sensitivity are key parameters that are guaranteed by transceiver vendors over operating temperature range and over transceiver lifetime. In the example illustrated in Figure 12, a 2Gbps SFP transceiver operating at 850nm has a minimum launch power of –9dBm and a maximum receiver sensitivity of –18dBm, resulting in an available power budget of -9 – (-18) = 9dB. This is the range shaded in blue diagonal lines in Figure 12. This is also specified in Table 2.

The optical loss budget is shown by the solid-red blocks in Figure 12. One factor in the loss budget is the presence of connectors and/or splices. The optical path in our examples involves 3dB loss due to the presence of four optical connector pairs (source, destination, two patch panels) and no splices. Since each mated connector pair results in about 0.75dB loss (refer to Table 1), the total loss due to connectors is 4 x 0.75dB = 3dB.

Next, ±1dB is added to the optical link loss budget as a consideration for temperature effects on the fiber itself and the mated connectors.

The next factor in the optical link loss budget is the repair and safety margin for the addition of future patch panels or splices, in this case specified as -2dB.

The maximum fiber length can now be determined by calculating the maximum allowable fiber loss, which is obtained by subtracting the total power loss (sum of all solid-red blocks) from the available power budget provided by the transceiver (block in blue diagonal lines). In our example, the maximum allowable fiber loss is (9dB – 6dB = 3dB), represented by the dotted green block in Figure 12. The maximum fiber length is then calculated by dividing up the allowable fiber loss by the fiber attenuation. As an example, the optical attenuation in OM2 fiber is 3dB/km as shown in Table 1. The maximum fiber link distance is then:

L_{max} = (Max Allowable fiber loss) / (Fiber attenuation) = (3dB) / (3dB/km) = 1km.

So, the power-limited maximum link distance for our example is 1,000m. Now, one needs to check the distance limitation due to the type of fiber used.

The chart in Figure 12 can be easily applied to other link scenarios to determine the power-limited maximum link distance by using the relevant values for the available power budget from the transceivers and the loss due to connectors and splices. The appendixes at the end of this guide also provide a practical worksheet to do such calculations.

Distance Limitation Due to Fiber Bandwidth

Contrary to popular belief, optical fiber does not provide unlimited bandwidth. Every type of fiber has a characteristic bandwidth•distance limitation, which basically implies that the allowable fiber link length decreases as the data rate of the optical signal increases. This is true for single-mode fiber (SMF), but much more so for multimode fiber (MMF), where effects such as modal noise conspire to limit the maximum bandwidth•distance (having unit of MHz•km) attainable. Fiber bandwidth•distance limitation is also a function of the type and wavelength of the optical source used in the transmitter. The maximum distance due to fiber bandwidth•distance limitation is shown in Table 1 and is specified as 500m (see previous notes on 500m) for our example of OM2 type multimode fiber at 2Gbps and at 850nm.

So, the maximum guaranteed distance for our example is limited by the fiber bandwidth to 500m.

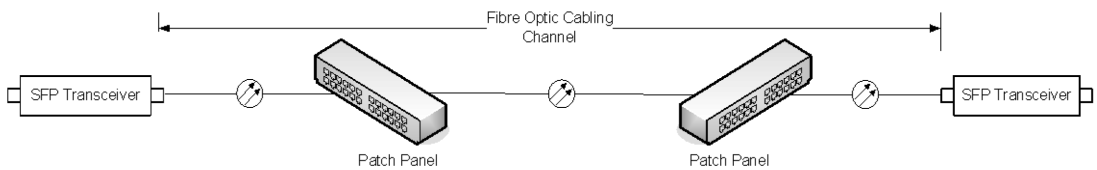


Figure 12) Patch panel interconnections.

It must be noted that the typical parameters specified in Table 1 are conservative, and specific installations might operate beyond the maximum distance specified in the table. While the analysis above provides a useful tool to design and assess an optical link based on nominal values, actual testing and characterization of the fiber link and its components by a qualified fiber technician are the most accurate ways to determine the real limitation of an optical link. This naturally implies that by physically measuring a channel's characteristics and finding that it is within the limits of the specification for the fiber in use, a longer distance than the specification's maximum could be achieved.

NetApp does not support distances beyond the stated maximums. Careful reevaluation of the optical link must be conducted when upgrading the link to higher data rates such as migrating from 2 to 4 Gbps.

7.6 FIBRE CHANNEL SFPS

There are four types of SFPs associated with fabric MetroCluster configurations. They are:

- **Short-wavelength laser (SWL)**. Short-wavelength laser transceivers based on 850nm lasers are designed to transmit over short distances. This is the most common type of media and is the default on the Brocade 200E. These lasers are also used in DS14-type shelf controllers and on-board FC ports on FAS6000 and FAS3000.
- **Long-wavelength laser (LWL)**. Long-wavelength laser transceivers may be based on 1310nm lasers. They are used for long-distance native FC links. Generally, these media types are used with SMF cable and only used for the interswitch link port on fabric MetroCluster.
- **Extended long-wavelength laser (ELWL)**. Extended long-wavelength laser transceivers may be based on 1550nm lasers. They are used to run native Fibre Channel connections over even greater distance than LWL media can support. Generally these media types use SMF cable. They are only used for the interswitch link port on fabric MetroCluster and require a "distance option license" for the fabric switch hardware.
- **Dense wavelength-division multiplexing (DWDM)**, DWDM works by combining and transmitting multiple signals simultaneously at different wavelengths on the same fiber. In effect, one fiber is transformed into multiple virtual fibers.

The type required is a function of the distance and the interconnect technology used. Table 3 summarizes the supported types and specifications.

Table 3) Supported types and functions.

SFP Type (Wavelength)	Max. Distance	Speed Gbps	NetApp Part Number	Used On	Outside Vendor Part/Part Number
SWL (850nm)	500m[1]	4	X6539-R5	32XX,38XX,200e	Finisar/FTLF8524P2BNV
LWL (1310nm)	10km	2	X1554B	32XX,38XX,200e	Finisar/FTRJ1319P1BTL-B1
ELWL (1310nm)	52km	2	X1696-R5	32XX,38XX,200e	Finisar/FTRJ1419P1BTL-B1
LWL (1310nm)	10km	4	X1670A	200e, 5000	Finisar/FTLF1424P2BCD
ELWL (1550nm)	30Km[2]	4	X1504-R6	200e, 5000	Finisar/FTLF1424P2BCV
LWL (1310nm)	4km	4	X1868-R6	300, 5100	Brocade XNA-000142
LWL (1310nm)	10km	4	X1880-R6	300, 5100	Brocade XNA-000144
LWL (1310nm)	30km	4	X1881-R6	300, 5100	Brocade XNA-000146
LWL(1310nm)	10km	8	X1869-R6	300, 5100	Brocade XNA-000153

1 See Table 1.

2 Using 9/125 single-mode cable.

8 APPENDIX

8.1 OPTICAL LINK LENGTH CALCULATION WORKSHEET

This worksheet is an aid for planning. Look for the actual spreadsheet on the MetroCluster product page.

OPTICAL FIBER LINK BUDGET CALCULA

Project:			
Transceiver Manufacturer:	Finisar		
Transceiver Manufacturer P/N:	FTLF1319P1xTL		
Type of Fiber:	MM		
Transmitter Wavelength:	1360		nm
Bit Rate:	2		Gbps
A) Minimum Transmitter Output Power:	-9.5		dBm
B) Worst Case (or Maximum) Receiver Sensitivity:	-15		dBm
C) System Gain: (A - B) or manufacturer specified optical budget		**5.5**	dB
D) Losses Due to Connectors: (per manufacturer's specifications) (loss / connector = **0.75**) x (number of connectors =)	2		dB
E) Losses Due to Installation Splices: (loss / splice = **0.3**) x (number of splices =)	0		dB
F) Safety Margin and Margin for future Repair Splices*:	1		dB
G) Margin for future WDM Upgrade (addition of optical Mux/Demux, Splitters, etc.):			dB
H) Maximum Allowable Optical Fiber Link Loss: (C-D-E-F-G)		**2.5**	dB
I) Fiber Attenuation (from Table 1 or fiber manufacturer):	3.5		dB/km
J) Maximum Allowable power-limited Fiber Span Length (1000*H/I)	**714**		m
K) Maximum Allowable Bandwidth-Limited Fiber Span Length (from Table below; Applies to MM Fiber)	500		m
L) Maximum Allowable Fiber Span Length (Minimum of J and K)		**500**	m

If (L) is less than the fiber span required in your application, installation tests by a qualified technician are recommended to assess the possibility of extending the fiber span length based on measurement results and/or actual tests with MetroCluster e

Note: The presence of significant optical reflections will make link performance poorer than calculated values.

* : Safety margin includes power penalty due to dispersion penalty caused by signal propagation in fiber.

9 REVISION HISTORY

Date	Name	Description
9/19/2006	Jim Lanson	Creation
11/20/06	Optellent Inc.	Update of optical link design sections
12/01/06	Optellent Inc.	Final revision of optical link design sections
7/15/07	Jim Lanson	Updated for 4Gbps support and new platforms
4/19/10	Jim Lanson	Updated for Data ONTAP 7.3.2 and new platforms

© Copyright 2010 NetApp, Inc. All rights reserved. No portions of this document may be reproduced without prior written consent of NetApp, Inc. Specifications are subject to change without notice. NetApp, the NetApp logo, Go further, faster, Data ONTAP, NOW, and SyncMirror are trademarks or registered trademarks of NetApp, Inc. in the United States and/or other countries. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.

NetApp provides no representations or warranties regarding the accuracy, reliability or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

© 2010 NetApp. All rights reserved.no portions of this document may be reproduced without prior written consent of NetApp, Inc. sp Specifications are subject to change without notice. NetApp, the NetApp logo, Go further, faster, and DataONTAP, NOW and SyncMirror are trademarks or registered trademarks of NetApp, Inc. in the United States and/or other countries. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such. TR-3517