

Dell™ PowerVault™ Systems

Performance Considerations for Tape Drives and Libraries

**Information in this document is subject to change without notice.
© 2005 Dell Inc. All rights reserved.**

Reproduction in any manner whatsoever without the written permission of Dell Inc. is strictly forbidden.

Trademarks used in this text: *Dell*, the *DELL* logo, and *PowerVault* are trademarks of Dell Inc.; *EMC* and *PowerPath* are registered trademarks of EMC Corporation.

Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and names or their products. Dell Inc. disclaims any proprietary interest in trademarks and trade names other than its own.

June 2005

Contents

Introduction	5
General Host Backup Considerations	5
Tape Drive and Data Considerations	5
Hard Drive and RAID Array Configuration	6
General Performance Considerations When Using Multiple Drives in Tape Libraries	12
SCSI Configurations.	12
SAN Configurations.	13

Introduction

With recent improvements in tape drive transfer rates, many host-side factors, such as RAID (Redundant Array of Inexpensive [or Independent] Disks) configuration and hard-drive specifications, must be considered when determining whether the host server and tape drive can process data at the same rate. General configurations and attributes that may limit throughput from the host server to the tape drive are discussed in "General Host Backup Considerations."

As multiple drives are placed into tape libraries, greater host bandwidths are needed to keep pace with the potential throughput of multiple tape drives. Potential fibre limitations for multidrive units, as well as recommended cabling configurations, are discussed in "General Performance Considerations When Using Multiple Drives in Tape Libraries."

General Host Backup Considerations

The considerations in this section apply to both SCSI and storage area network (SAN) tape backup configurations.

Tape Drive and Data Considerations

The following issues should be considered when evaluating performance:

Overhead from SCSI commands. Command overhead on the SCSI bus restricts *all* SCSI devices in achieving theoretical maximum transfer speeds. Tape backup software does not account for this overhead; instead, the software only measures the rate at which data is written to the tape. For example, the drive may be processing 80 MB/sec of data, but only writing 77 MB/sec of data. The latter rate is what the backup software will report.

Tape block sizes. 64 Kb block sizes are optimal for most tape drives. However, some backup applications allow the user to change block size, even though a larger size will not enhance performance. Using block sizes less than 64 Kb can actually hinder performance. See your backup software User's Guide for information on adjusting the block size of your tape device.

Backup software buffer size. For optimal backup performance, backup software buffers should be as large as possible. Some applications allow users to change the buffer size, which can help maintain a steady stream of data to and from the drive and significantly increase transfer rates, especially of small files. The larger the buffer, the more data it can hold and the less time the disk spends seeking the data; however, this can affect memory and CPU performance. See your tape backup application User's Guide for specific details.

Drivers and firmware. Always ensure that the SCSI or fibre controller and tape drive have the latest drivers and firmware installed. Visit support.dell.com to download the latest drivers and firmware for your Dell PowerVault tape product.

Attach tape drives and hard drives on separate controllers (internal or external) on separate host bus adapters (HBAs). This depends somewhat on the performance capabilities of your controller, but best practice is to keep the tape drive HBA separate from the hard drive HBA to ensure maximum throughput. Most onboard dual-mode SCSI/RAID controllers share one processor, which must share bandwidth between the RAID array and tape drive. Thus, one controller is handling reads and writes between the hard disks and the tape drive, as well as calculating and writing any necessary parity information to the hard drives. See "Hardware RAID Configuration Considerations" for specific information on RAID arrays and parity bytes.

Dirty drive heads or old media. A dirty tape drive head or old media can cause high error rates and a corresponding reduction in read/write speeds. Each time a drive attempts to rewrite or reread a block on a tape, performance is degraded. Once a certain threshold of read/write errors is reached, the drive will usually request cleaning. It is important to clean the drive heads on regular intervals, or when requested.

The chance of encountering a bad block increases as media ages or is excessively used. The typical lifecycle of a piece of LTO media is approximately 75 full tape read/writes.

Speed matching. Newer LTO drives will match the speed of the data being provided to the drive, down to approximately one-half of the maximum uncompressed data transfer speed. If data is provided to the drive at less than the lower speed matching limit, the drive must stop, wait for the buffer to fill, rewind, and then attempt to write the buffer (this is known as "back hitching").

For example, the Dell PowerVault 110T (LTO2 and LTO3) tape drive matches speed down to 30 MB/sec while writing to LTO-3 media. If the host server can only provide data at 20 MB/sec, the drive will "back hitch" while waiting for its buffers to fill. In this situation, the effective throughput will be something less than 20 MB/sec (probably closer to 15 MB/sec).

Confirming Performance of Your Tape Drive

Certain tape drive manufacturers have a performance diagnostic mode built into the drive that can be used to confirm throughput. The PowerVault 110T LTO-2 and LTO-3 (firmware 53XX or later) offer a diagnostic mode "F," which performs a quick read/write performance test on the drive and media. If the performance rate is not within 6 percent of the maximum specified drive speed, the test fails with an error message. No error message is displayed if the test passes. Consult your tape drive User's Manual for specific details on diagnostic mode "F."

 **NOTICE:** Diagnostic mode "F" requires media that can be safely overwritten as part of the diagnostic test. Do not use media containing critical data. Any data residing on the media used in the diagnostic test will be lost.

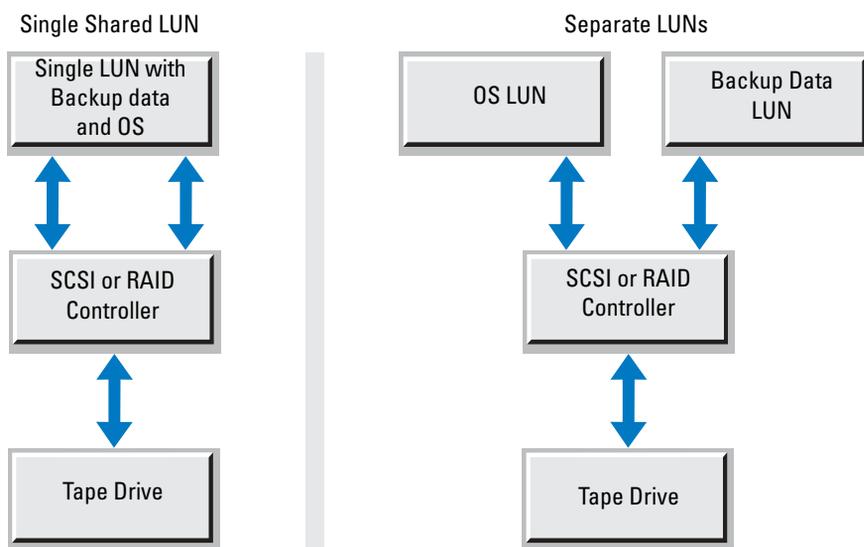
Hard Drive and RAID Array Configuration

Several hard drive and disk array (both internal and external) attributes can affect backup or restore performance. These attributes, as well as recommended configurations that help achieve maximum backup and restore speeds, are discussed in the following subsections. If the tape drive's sustainable throughput exceeds that of the disk array, then the tape drive's peak performance will not be realized.

General Hard Drive Configuration Considerations

Data/operating system (OS) on different LUNs. Backing up data on a logical unit number (LUN) separate from the OS LUN ensures that the hard drive is not splitting access and overhead between OS operations and backup operations. This can be accomplished by having one hard drive or disk array contain the OS and a physically separate hard drive or disk array contain the data to be backed up.

Figure 1-1. Single-Channel vs. Two-Channel Bandwidth



Hard Drive Performance

By design, tape drives write data sequentially and require a constant data feed to keep the drive operating sequentially (avoiding back hitching). Conversely, hard drives are random access devices. Therefore, hard drives can sometimes struggle to provide sequential data to tape drives if that data is spread out over the drive platter. This forces the drive to continuously seek small blocks of data.

Additionally, other hard drive attributes can further affect the throughput of data to the tape drive.

Spindle speed. Typically measured in RPMs (revolutions per minute), the hard drive's spindle speed determines how many times per minute the drive platter assembly can perform a full revolution. This has a direct effect on both random access times and sequential transfer rates. The higher the spindle speed, the faster the drive can access data.

Random access time or seek time. Usually measured in milliseconds, seek time is the length of time a drive's heads take to find a piece of data on the disk. The seek time of a hard disk measures the amount of time required for the read/write heads to move between tracks on the surface of the platters. Because hard disks are random access devices, data can be stored on virtually any sector of the disk. The longer it takes to access that data, the slower the overall throughput of the drive. This attribute is very significant when a hard drive contains many small files. The smaller the files, the more "seeks" the drive must make to read or write the file to disk; therefore, disks tend to read or write very slowly when many small files are being transferred.

Sequential/sustained transfer rates (STR). STR measures how fast a drive actually reads data from and writes data to its platters. If the data being backed up is one large contiguous file, the sustained throughput will be close to the drive's maximum STR. However, in real-world applications, data becomes more scattered about the platter as data is deleted and written. Defragmenting a hard drive can help the drive reach its maximum STR.

Buffer (cache). The buffer is the amount of memory on the drive that holds the most recently written or stored data. The bigger the buffer, the more data it can hold, resulting in less time seeking data on the disk.

Hardware RAID Configuration Considerations

General overview of RAID

This section presents an overview of typical RAID configurations and how they affect backup and restore rates. A RAID array is a set of hard disks that act as a single storage system or LUN. Data can be potentially transferred through the channel of each hard drive at once, allowing for total throughput to be a multiple of the total number of drives in the array, minus overhead and any redundancy as described in the following sections.

In the case of a RAID configuration, the speed of the interface becomes important because the drives share the bandwidth of the interface. For example, a single Ultra160 drive may only sustain 40 MB/sec. Thus, a five-disk RAID 0 array consisting of the same drive type should be able to read/write at 200 MB/sec. However, the Ultra160 interface will limit the array to a maximum of 160 MB/sec.

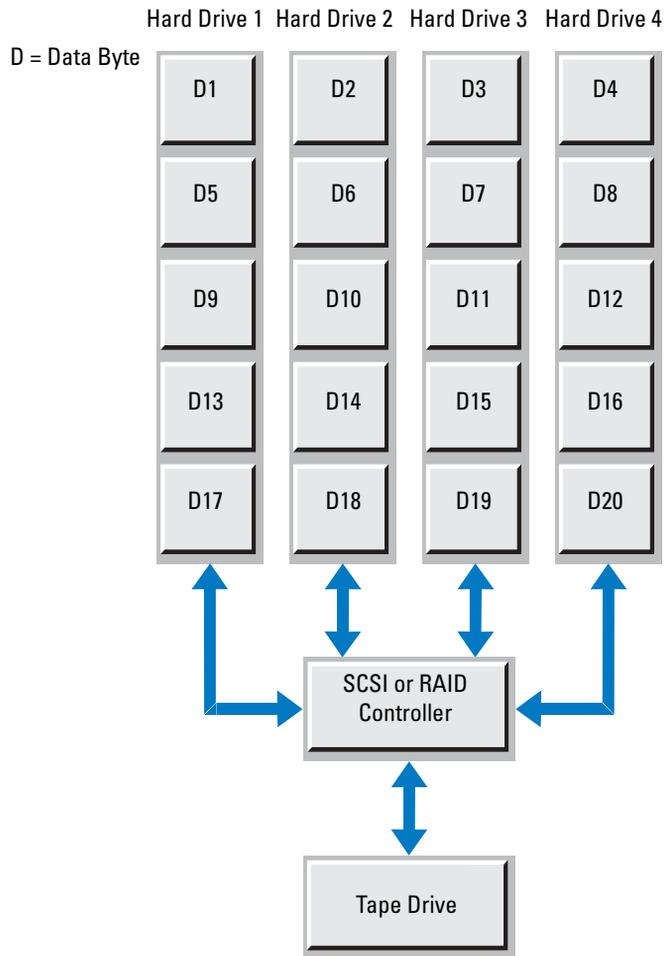
External disk arrays, particularly in SANs, may offer significant levels of cache memory to improve I/O performance. This cache will greatly improve performance when writing to the array and may store frequently accessed data to improve read performance. With respect to its impact on tape performance, the cache will mask most RAID limitations when restoring data to the array or backing up data from the array. However, backup operations from external arrays with cache may still feel the impact of RAID configuration limitations because the data still needs to be read from the disks.

RAID 0

Commonly known as striping, RAID 0 allows two or more disks to be joined to create one virtual drive in the fashion of a single LUN. It is referred to as striping because data is written across all of the disks in the array, not just to one disk at a time. Thus, the throughput is spread across n channels (n being the number of hard drives in the array) instead of a single channel for a single hard disk. This results in excellent read/write performance, but no fault tolerance.

Figure 1-2 shows four hard drives in a RAID 0 configuration. Data is striped across all four hard drives, resulting in four channels for reading and writing to the array.

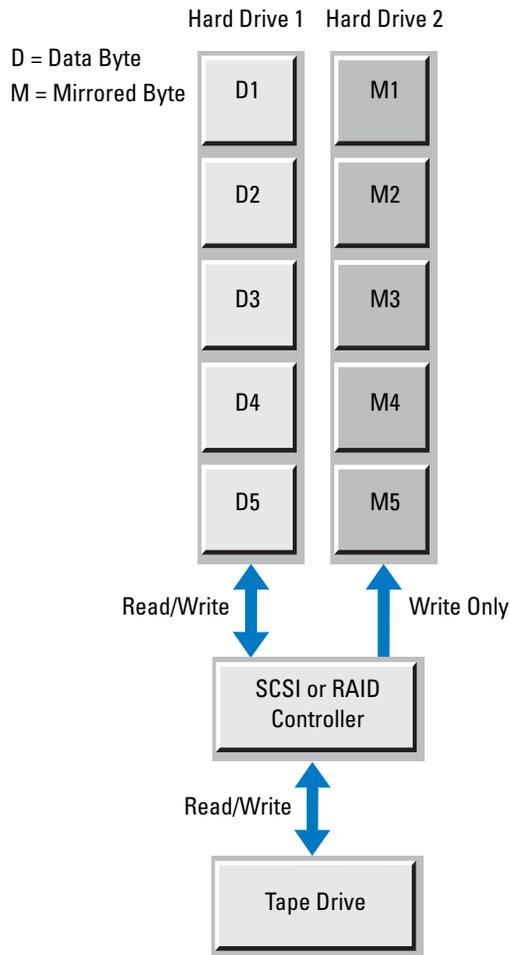
Figure 1-2. Example RAID 0 Configuration



RAID 1

Also known as mirroring, RAID 1 means data is written twice across two disks simultaneously. If one drive fails, the system switches to the other drive without losing data. During tape drive backups, the read rate from the RAID 1 array is approximately the same as a single drive because it is reading from the primary drive. However, restore performance from the tape drive writing to the RAID 1 array can be slower due to error checking/correction (ECC) included in writing to the primary and mirrored disks. Much of this inefficiency is due to the fact that the mirroring is often performed on the CPU or RAID controller. Thus, newer RAID controllers tend to be faster due to newer and more capable processors.

Figure 1-3. Example RAID 1 Configuration

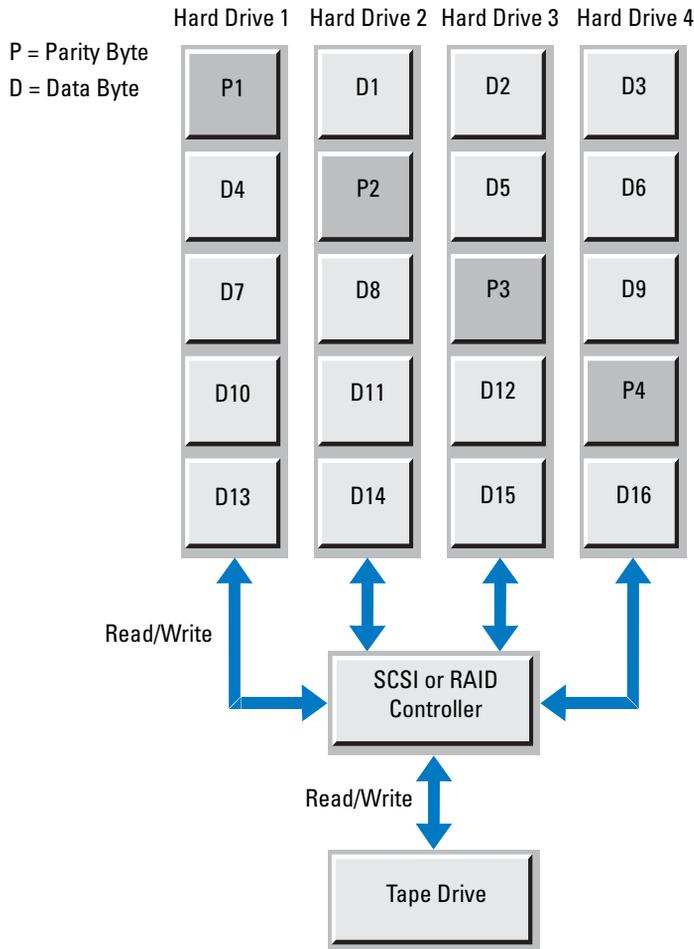


RAID 5

With a RAID 5 array, data is striped across the disk array at the byte level and error correction data, or parity data, is also striped across the disk array. RAID 5 arrays tend to have very good random read performance; this read performance generally improves as the number of disks in the array increases. With the larger disk arrays, read performance can actually outperform RAID 0 arrays because the data is distributed over an additional drive. In addition, parity information is not required during normal reads.

Restores from tape to a RAID 5 array tend to be nominal because it involves additional overhead for calculating and writing the parity information.

Figure 1-4. Example RAID 5 Configuration



In conclusion, RAID 0 tends to be the best overall configuration for read and write performance, but does not allow for redundancy. RAID 1 is the worst performer overall, as all data written to the array is mirrored and reads come from a single disk. RAID 5 tends to be a good read performer but average write performer; however, RAID 5 improves if more disks are added to the array. If the RAID is within an enclosure that offers significant levels of cache memory, then performance limitations during restore operations may be abated. Backup operations will still be subject to limitations of the RAID configuration. In addition, the characteristics of the array still depend heavily on the specific hard drive characteristics listed in "Hard Drive Performance."

General Performance Considerations When Using Multiple Drives in Tape Libraries

When multiple tape drives are utilized simultaneously to perform data backups (such as in tape libraries), additional aspects of the hardware configuration must be considered. By employing simple performance-minded methods in setting up hardware and cabling configurations, additional throughput bottlenecks can be limited.

SCSI Configurations

The latest high-performance tape drives offered in tape libraries support the Ultra160 specification of the SCSI interface standard. Therefore, to achieve maximum performance, backup servers utilizing SCSI must have an HBA installed that supports data speeds of Ultra160 or higher. A SCSI HBA that meets this requirement will allow each tape drive to communicate with the host at a rate of 160 MB/sec on the SCSI bus. The higher data rate of the SCSI bus compared to tape drive speeds allows multiple devices to be connected to the same bus without sacrificing device performance. But only to a point.

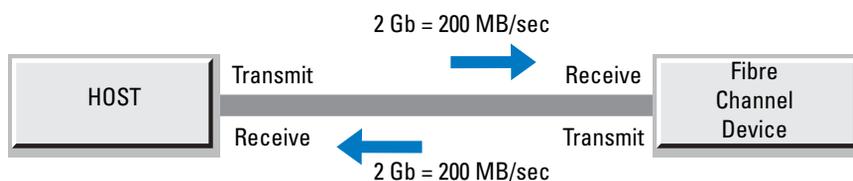
The 160 MB/sec data rate of an Ultra160 bus is the maximum possible data throughput rate to all devices connected to the bus. Therefore, a single tape drive will not consume the full bandwidth of the bus because it can read or write data to tape at up to 80 MB/sec (native). Multiple tape drives, however, can combine to consume the full 160 MB/sec offered by the bus if each is operating at its maximum native performance. Each additional drive connected to the same bus after this point will reduce the average performance of each drive.

Therefore, to achieve maximum performance from a tape library, it is recommended to connect no more than two tape drives to each SCSI bus. See "Recommended Cabling Configurations" for specific details and illustrations. A SCSI HBA supporting at least Ultra160 should be used, but upgrading to an Ultra320 HBA will not lead to an additional improvement in performance if the tape drive's specification is Ultra160.

SAN Configurations

Fibre Channel (FC) offers many advantages over SCSI. First, it overcomes the distance limitations of SCSI (12 meters for LVD SCSI versus 300 meters for a short-wave 2-Gb FC link) and allows for the transmission of data at higher speeds. As a serial network protocol rather than a bus-based architecture like SCSI, FC has also become the protocol of choice for the implementation of SANs, allowing for the consolidation of data storage resources. In addition, each FC connection is made up of a transmit link and a receive link, allowing for full-duplex operation. This means that data can be transmitted in two directions simultaneously. Therefore, during a backup operation across a single FC connection, data can be read from a source and written to tape without taking turns in communication, effectively doubling the bandwidth of a connection. See Figure 1-5.

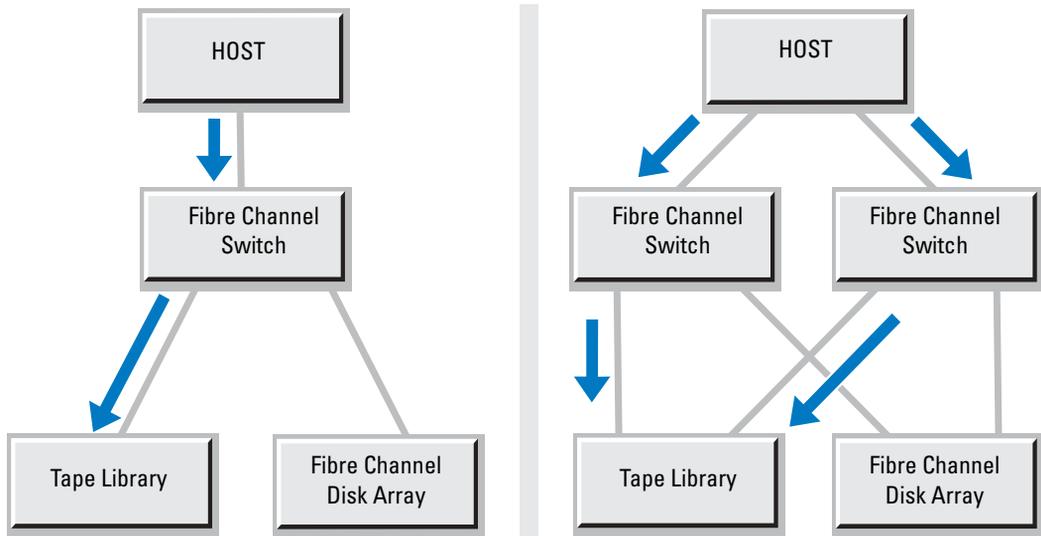
Figure 1-5. Fibre Channel Link Diagram



When setting up tape libraries in a SAN, performance can still be affected by various factors. These factors include FC link speeds, data flow between the source and tape library, and performance limitations of external storage arrays. With an understanding of the overall setup and management of the solution, many of these factors can be avoided.

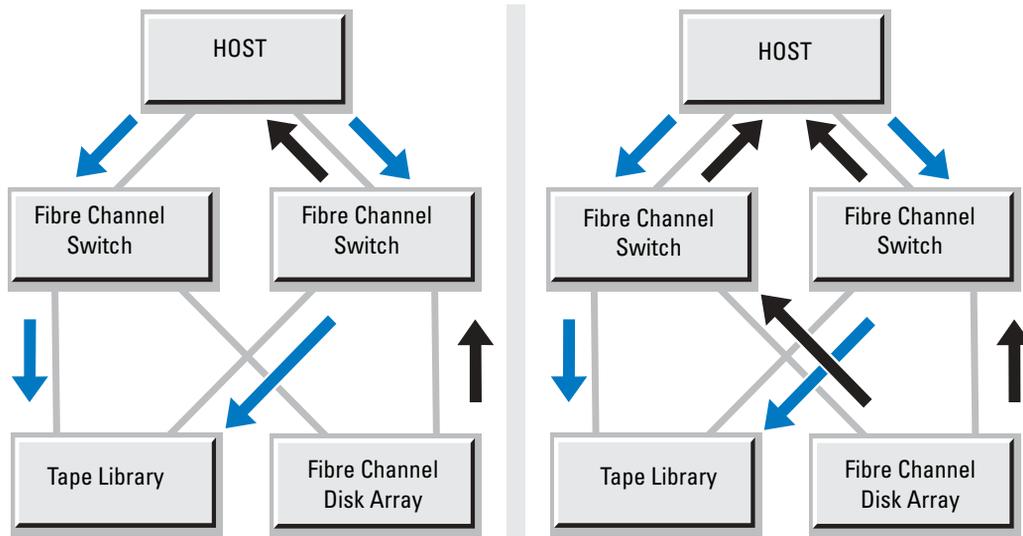
Even with the high data bandwidth offered by the FC protocol in SANs, proper considerations must be made for tape drives in order to avoid a situation in which the FC link may limit performance. The data rate of a 2-gigabit (Gb) FC link is 200 MB/sec (that is, 200 MB/sec on the transmit link and 200 MB/sec on the receive link). Therefore, attempting to operate multiple tape drives across the same link can potentially exceed the full bandwidth of a link. If the host is operating with a legacy 1-Gb adapter, backing up data to two drives may be sufficient to reveal significant performance limitations.

Therefore, when using three or more tape drives simultaneously on a 2-Gb link, you may need to distribute the backups across a number of connections, rather than relying on a single link. This is where understanding the SAN solution's topology is beneficial. Following the data path during a backup operation as it is read from the source and then written out to tape will help administrators recognize any potential bottlenecks. If any bottlenecks are identified, measures may be taken depending on the configuration. For example, if the backup solution requires multiple drives to be in operation at once, splitting the tape hardware across separate fabrics may improve performance by splitting the connections. See Figure 1-6.

Figure 1-6. Single vs. Split Data Flow to Tape Library

At the same time, if a bottleneck exists in the data being read from an external FC disk array, the use of I/O management software such as EMC® PowerPath® with an additional fabric connection will automatically load-balance the data across multiple paths and increase availability through path failover. See Figure 1-7. The left side of the figure represents a SAN disk array in which all of the data is forced through a single link, creating a bottleneck that slows data flow to the tape hardware. The right side shows how load balancing doubles the I/O bandwidth coming out of the array by allowing the data to transmit across two links.

Figure 1-7. Bottlenecked Data Flow vs. Load-Balanced Disk Array



Finally, FC disk arrays on the SAN can also experience the same performance limiters described in "Hard Drive and RAID Array Configuration." Therefore, improving the performance characteristics of the disk arrays will also have a direct effect on backup speed across the SAN.

SAN Configurations Utilizing the Library Fibre Channel Bridge

Certain tape libraries may be connected to a SAN by way of a Fibre Channel bridge module. The module acts as a bridge between the SCSI and FC protocols and provides additional management, security, and operational features unavailable in most native FC libraries. For details on these features, see the Fibre Channel bridge User's Guide for your tape library.

In some tape library configurations, the Fibre Channel bridge module may act as a bottleneck and decrease performance of tape drives. This is because the processing capability in the Fibre Channel bridge module required to bridge the SCSI and FC communication cannot meet the aggregate data throughput offered by certain multidrive configurations. Despite this, most data backup solutions will not experience the Fibre Channel bridge module as the primary limiting factor in tape performance. Dedicated backup servers will frequently encounter a situation where the limitations at the host will be compounded by the exertion of feeding data to multiple tape drives. This results in average drive performance below the level where the Fibre Channel bridge module becomes a factor.

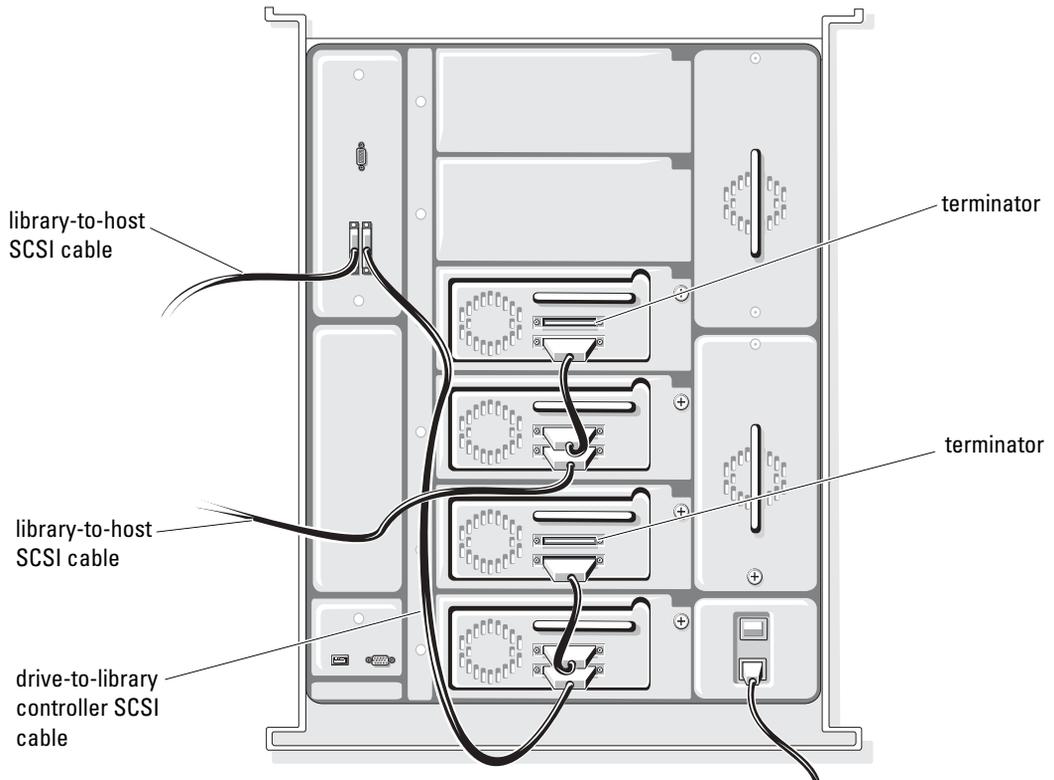
Recommended Cabling Configurations

SCSI Cabling to the Host

Tape Library With up to Six Tape Drives

When the tape library is SCSI-attached to a host, ensure that no more than two drives are on a single bus. Additional SCSI controllers are required for libraries with five or six tape drives to ensure that no SCSI bus becomes a barrier to maximizing throughput.

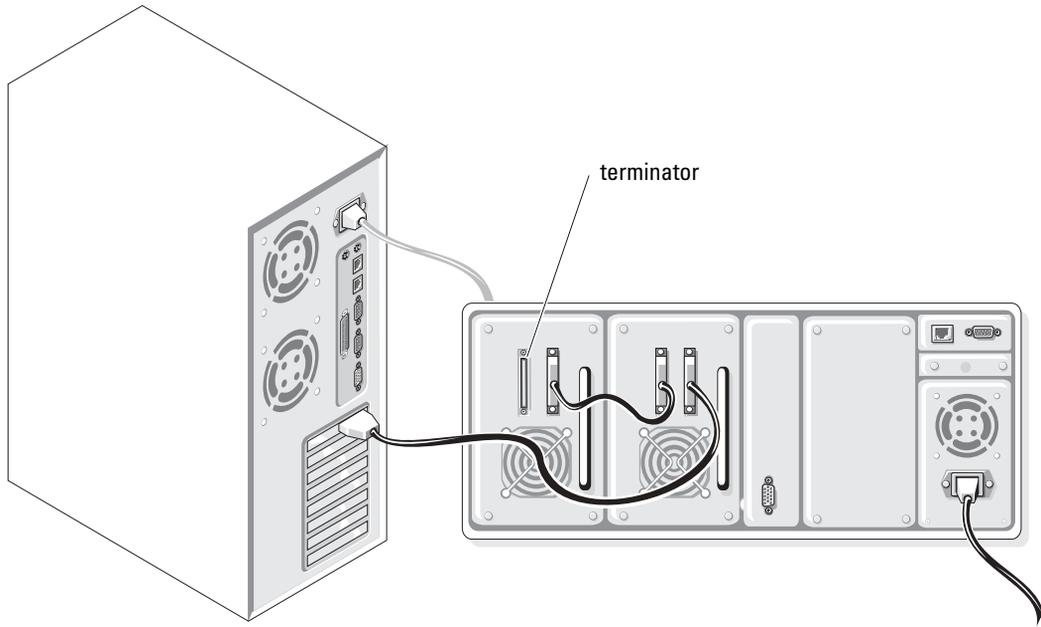
Figure 1-8. SCSI Cabling for Library With up to Six Tape Drives



Tape Library With up to Two Tape Drives

The drives in a fully configured two-drive tape library can be cabled to a host on the same SCSI bus without encountering significant limitations to backup performance. The backup rates for two drives on a single SCSI bus will match the backup rates for two drives on separate buses. However, customers who enable the verify feature in backup applications may wish to improve the verify performance by splitting two drives onto two SCSI buses. By doing so, verify performance may improve by up to 25 percent.

Figure 1-9. SCSI Cabling for Library With up to Two Tape Drives



Cabling Drives to the Fibre Channel Bridge

Figures 1-10 through 1-17 illustrate how a tape library with up to six drives should be configured with a Fibre Channel bridge module in order to optimize tape performance over FC.

Figure 1-10. Fibre Channel Bridge Cabling With One Tape Drive

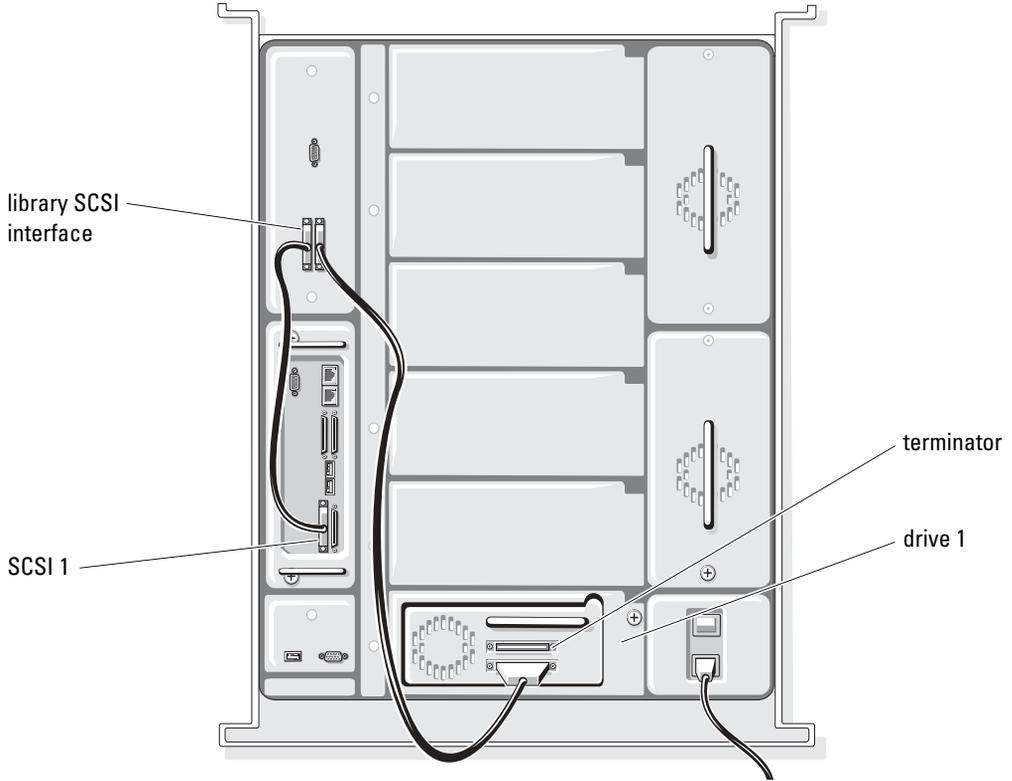


Figure 1-11. Fibre Channel Bridge Cabling With Two Tape Drives

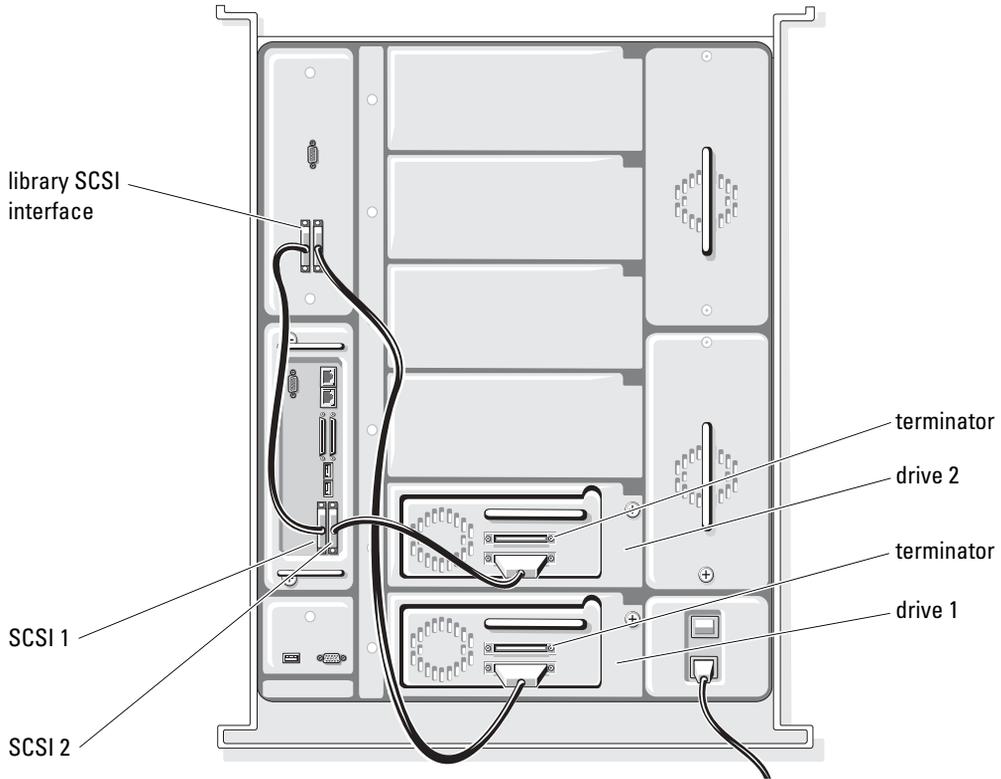


Figure 1-12. Fibre Channel Bridge Cabling With Three Tape Drives

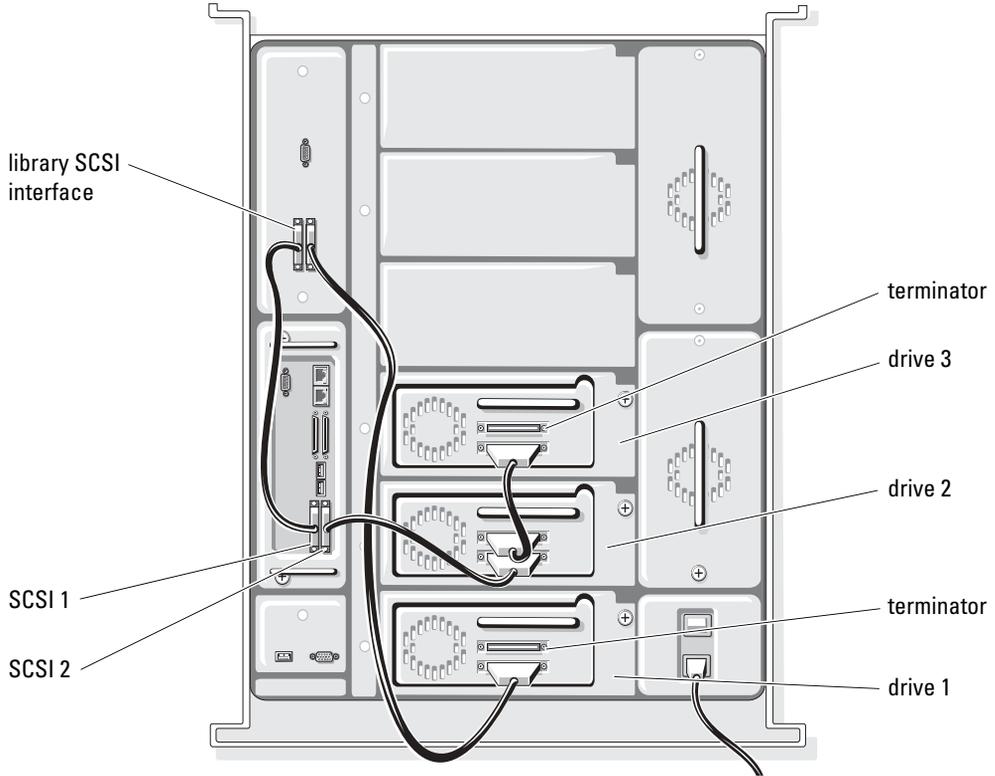


Figure 1-13. Fibre Channel Bridge Cabling With Four Tape Drives

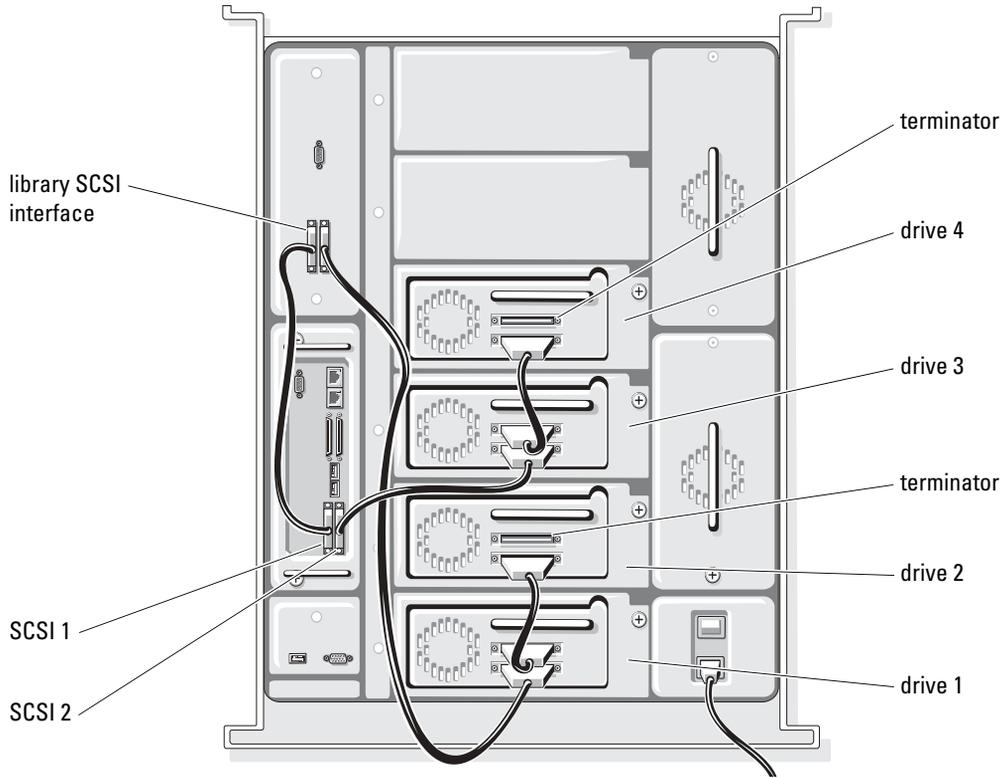


Figure 1-14. Channel Zoning Settings for Tape Library With One to Four Drives

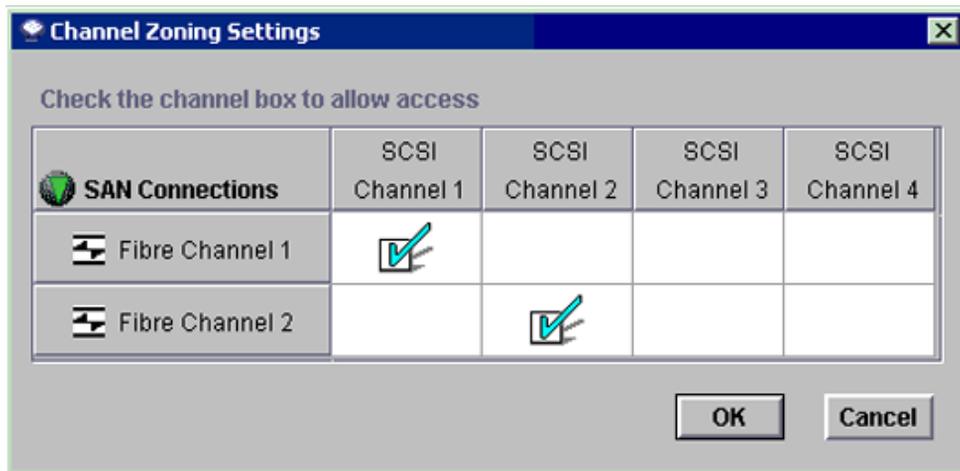


Figure 1-15. Fibre Channel Bridge Cabling With Five Tape Drives

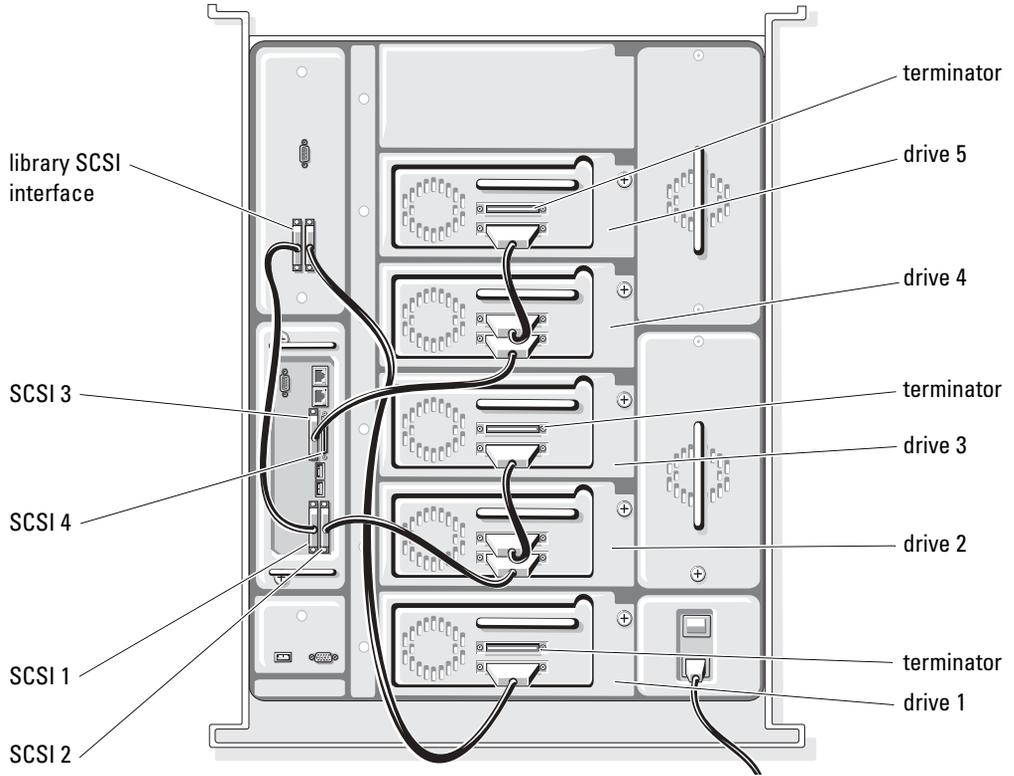


Figure 1-16. Fibre Channel Bridge Cabling With Six Tape Drives

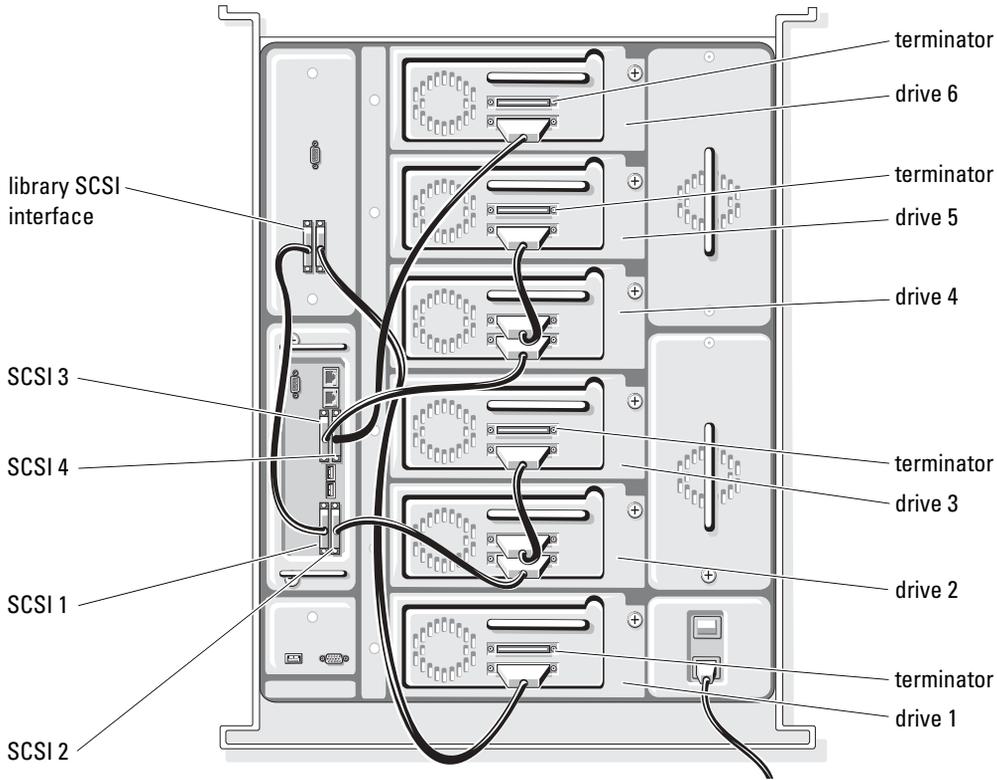


Figure 1-17. Channel Zoning Settings for Tape Library With Five or Six Drives

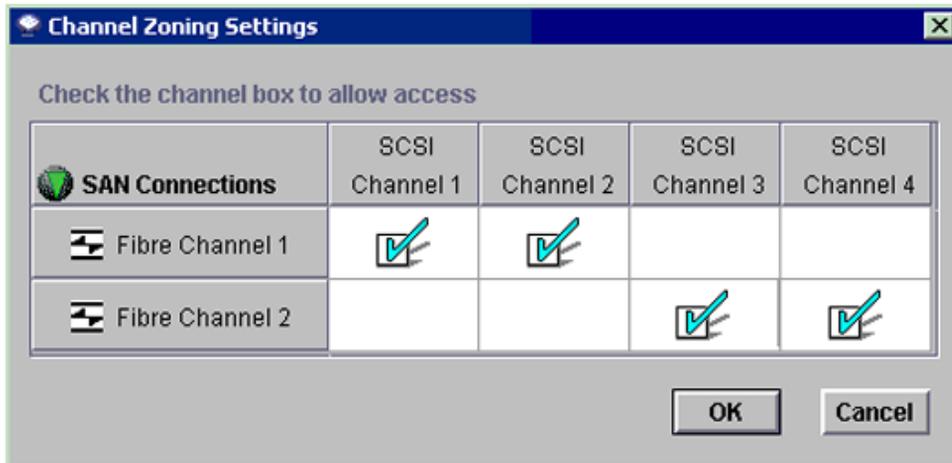


Figure 1-18 and Figure 1-19 illustrate how a tape library with up to two tape drives should be configured with a Fibre Channel bridge module in order to optimize tape performance over FC.

Figure 1-18. Tape Library With One Drive

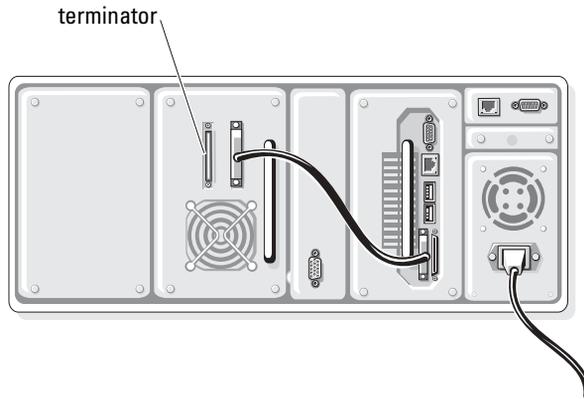
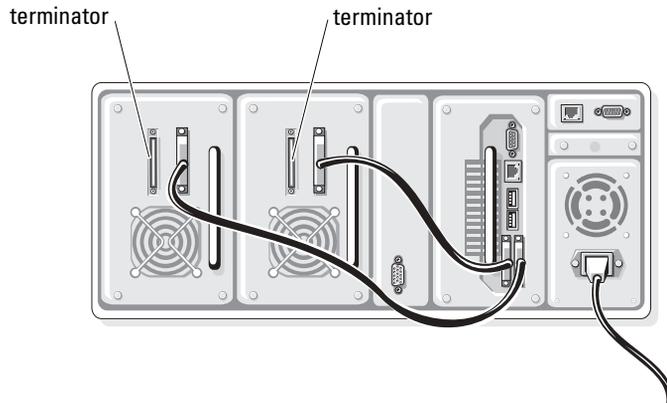


Figure 1-19. Tape Library With Two Drives



Dell™ PowerVault™ 系统

磁带驱动器和磁带存储库 性能注意事项

本文件中的信息如有更改，恕不另行通知。

© 2005 Dell Inc. 版权所有，翻印必究。

未经 Dell Inc. 书面许可，严禁以任何方式进行复制。

本文中使用的商标：*Dell*、*DELL* 徽标和 *PowerVault* 是 Dell Inc. 的商标；*EMC* 和 *PowerPath* 是 EMC Corporation 的注册商标。

本文件中提及的其它商标和产品名称是指拥有相应商标和名称的公司或其制造的产品。Dell Inc. 对本公司的商标和产品名称之外的其它商标和名称不拥有任何专有权。

2005 年 6 月

目录

简介	29
主机备份一般注意事项	29
磁带驱动器和数据注意事项	29
硬盘驱动器和 RAID 阵列配置	30
在磁带存储库中使用多个驱动器时的一般性能注意事项	36
SCSI 配置	36
SAN 配置	36

简介

随着近来磁带驱动器传输速率的提高，在确定主机服务器和磁带驱动器能否以相同速率处理数据时，必须考虑许多主机方面的因素，如 RAID（廉价 [或独立] 磁盘冗余阵列）配置和硬盘驱动器的规格。在“主机备份一般注意事项”中，将讨论可能会限制从主机服务器到磁带驱动器吞吐量的一般配置和属性。

由于磁带存储库中放入了多个驱动器，因此，为了满足多个磁带驱动器的潜在吞吐量要求，需要增加主机带宽。“在磁带存储库中使用多个驱动器时的一般性能注意事项”中将讨论有关多驱动器单元的潜在光纤限制以及建议的布线配置。

主机备份一般注意事项

本节中的注意事项同时适用于 SCSI 和存储区域网络 (SAN) 磁带备份配置。

磁带驱动器和数据注意事项

在评估性能时应考虑以下问题：

SCSI 命令所带来的额外开销。SCSI 总线上的命令所带来的额外开销限制了所有 SCSI 设备，使之无法获得理论上的最大传输速率。磁带备份软件并未说明这一额外开销，而只是测量数据写入磁带的速率。例如，驱动器可能以 80 MB/ 秒的速度处理数据，但写入数据时的速度仅为 77 MB/ 秒。后一个速率就是备份软件将报告的速度。

磁带区块大小。对于大多数磁带驱动器，64 KB 的区块大小是最佳的。然而，某些备份应用程序允许用户更改区块大小，即使较大的区块并不能提高性能。如果区块大小低于 64 KB，则的确会妨碍性能。有关调整磁带设备的区块大小的信息，请参阅备份软件的用户指南。

备份软件缓冲区大小。为达到最佳的备份性能，备份软件缓冲区应尽可能大。某些应用程序允许用户更改缓冲区大小，这样可能有助于在将数据传出和传入驱动器时保持稳定的数据流，从而显著提高传输速率，特别是对于小文件。缓冲区越大，存储的数据就越多，磁盘寻找数据所需的时间就越少；但这样可能会影响内存和 CPU 性能。有关详细信息，请参阅磁带备份应用程序的用户指南。

驱动程序和固件。始终确保 SCSI 或光纤控制器及磁带驱动器安装了最新的驱动程序和固件。请访问 support.dell.com，为您的 Dell PowerVault 磁带产品下载最新的驱动程序和固件。

在各个主机总线适配器 (HBA) 上的各个控制器（内部或外部）上连接磁带驱动器和硬盘驱动器。这在一定程度上取决于控制器的性能，但最佳的方法是保持磁带驱动器 HBA 与硬盘驱动器 HBA 分离，以确保获得最大吞吐量。大多数板载双模式 SCSI/RAID 控制器共享一个处理器，这就意味着必须在 RAID 阵列与磁带驱动器之间共享带宽。因此，一个控制器既要处理硬盘与磁带驱动器之间的读和写操作，又要计算和向硬盘驱动器写入任何需要的奇偶校验信息。有关 RAID 阵列和奇偶校验字节的具体信息，请参阅“硬件 RAID 配置注意事项”。

有污垢的驱动器磁头或旧的介质。有污垢的磁带驱动器磁头或旧的介质可能会导致很高的错误率，并相应地降低读 / 写速度。每次驱动器尝试在磁带上重写或重读区块时，性能都会下降。一旦到达了读 / 写错误的某个阈值，驱动器通常会要求清洗磁头。定期或在请求时清洗驱动器磁头是十分必要的。

当介质老化或过度使用时，区块损坏的机率就会增加。一片 LTO 介质的典型寿命大约为 75 次完全磁带读 / 写。

速率匹配。更高级的 LTO 驱动器将与向驱动器提供的数据的速度相匹配，最低速率为最大未压缩数据传输速度的一半左右。如果向驱动器提供数据的速度低于速度匹配的下限，驱动器必须停止传输，等待缓冲区填充，重绕磁带，然后尝试写入缓冲区（此过程称为“后系留”）。

例如，Dell PowerVault 110T（LTO2 和 LTO3）磁带驱动器在向 LTO-3 介质写入数据时，将速率匹配为最低 30 MB / 秒。如果主机服务器仅以 20 MB / 秒的速率提供数据，在等待缓冲区填充时，驱动器将出现“后系留”现象。在这种情况下，有效的吞吐量将低于 20 MB / 秒（可能接近 15 MB / 秒）。

确定磁带驱动器的性能

某些磁带驱动器制造商在驱动器中内置了性能诊断模式，用于确定吞吐量。PowerVault 110T LTO-2 和 LTO-3（固件 53XX 或更高版本）提供了一个诊断模式“F”，可以在驱动器和介质上执行快速读 / 写性能测试。如果性能速率超出了指定的最大驱动器速度的百分之六，则测试失败，并显示一条错误信息。如果测试通过了，则不会显示错误信息。有关诊断模式“F”的详细信息，请参阅磁带驱动器的用户指南。

 **注意：** 作为诊断测试的一部分，诊断模式“F”要求可以安全地覆盖介质。不要使用包含重要数据的介质。用于诊断测试的介质上包含的任何数据都将丢失。

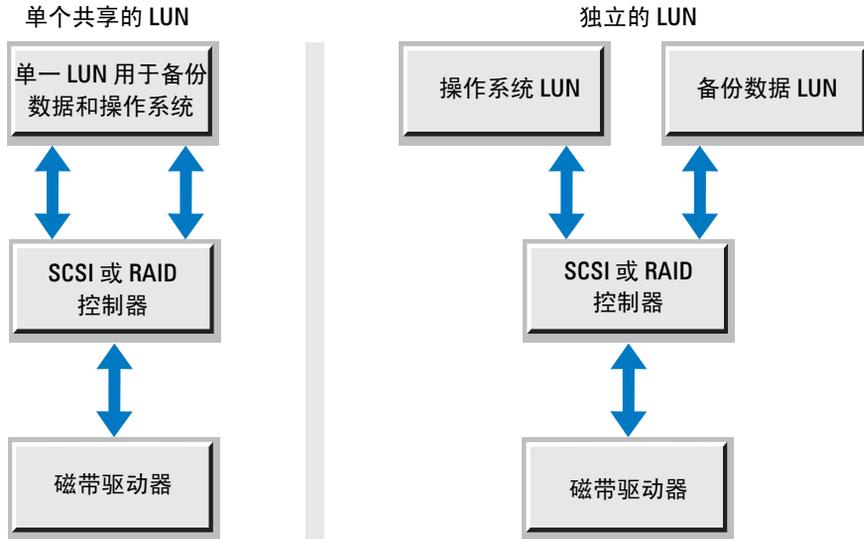
硬盘驱动器和 RAID 阵列配置

许多硬盘驱动器和磁盘阵列（包括内部和外部）属性可能会影响备份或恢复性能。以下各小节将分别介绍这些属性以及建议采用的可帮助获得最大备份和恢复速度的配置。如果磁带驱动器的可承受吞吐量超过了磁盘阵列的吞吐量，则磁带驱动器将无法达到最佳性能。

硬盘驱动器配置一般注意事项

不同 LUN 上的数据 / 操作系统 (OS)。在与操作系统 LUN 不同的逻辑设备号 (LUN) 上备份数据，可确保硬盘驱动器不会在操作系统操作与备份操作之间拆分访问量和系统开销。为此，可通过使用一个硬盘驱动器或磁盘阵列包含操作系统，而使用一个物理上独立的硬盘驱动器或磁盘阵列来包含要备份的数据来实现这一要求。

图 1-1. 单信道与双信道带宽



硬盘驱动器性能

根据设计，磁带驱动器应按顺序写入数据，并要求稳定的数据传输，以确保驱动器按顺序操作（避免出现后系留现象）。相反，硬盘驱动器是随机访问设备。因此，如果数据广泛分布在驱动器盘片上，则硬盘驱动器有时需要尽量向磁带驱动器按顺序提供数据。这样，就迫使驱动器持续查找小的数据块。

另外，其它硬盘驱动器属性可能进一步影响传输到磁带驱动器的数据吞吐量。

转速。通常以 RPM（转 / 分钟）为单位，硬盘驱动器的转速决定了驱动器盘片每分钟可以执行完整转动的次数。这不仅直接影响随机访问时间，也直接影响顺序传输速率。转速越高，驱动器访问数据的速度就越快。

随机访问时间或查找时间。通常以毫秒为单位，查找时间是驱动器磁头在磁盘上查找数据所用的时间。硬盘的查找时间衡量的是读 / 写磁头在盘片表面的磁道间移动所需的时间。因为硬盘是随机访问设备，故而数据可以存储在磁盘上的几乎任何扇区。硬盘访问数据所需的时间越长，驱动器的整体吞吐量就越慢。当硬盘驱动器包含许多小文件时，这一属性就显得十分重要。文件越小，驱动器就必须执行更多的“查找”，以便向磁盘读或写文件；因此，当要传输许多小文件时，磁盘的读或写操作往往会非常慢。

顺序 / 持续传输速率 (STR)。STR 衡量驱动器实际从盘片读取数据及向盘片写入数据的速度。如果要备份的数据是一个大的存储空间相邻的文件，则持续吞吐量将接近驱动器的最大 STR。然而，在实际应用中，由于数据的删除和写入，使得数据在盘片上的分布变得非常分散。对驱动器进行碎片管理有助于驱动器获得其最大 STR。

缓冲区（高速缓存）。缓冲区是驱动器上保存最近写入的数据或存储的数据的内存量。缓冲区越大，所能保存的数据就越多，从而减少了在磁盘上查找数据的时间。

硬件 RAID 配置注意事项

RAID 概述

本节概述了典型的 RAID 配置以及它们如何影响备份和恢复速率。RAID 阵列是一组可作为单一存储系统或 LUN 的硬盘。由于可以一次通过每个硬盘驱动器的信道传输数据，这样，总的吞吐量可以达到阵列中所有驱动器的吞吐量的若干倍，但要减去系统开销和任何冗余，如以下各节中所述。

就 RAID 配置来说，接口的速率很关键，因为驱动器共享接口的带宽。例如，一个 Ultra160 驱动器只能承受 40 MB / 秒的速度。这样，由相同驱动器类型组成的 5 个磁盘所形成的 RAID 0 阵列能够以 200 MB / 秒的速率进行读 / 写。但 Ultra160 接口将阵列的速率限制在最大 160 MB / 秒。

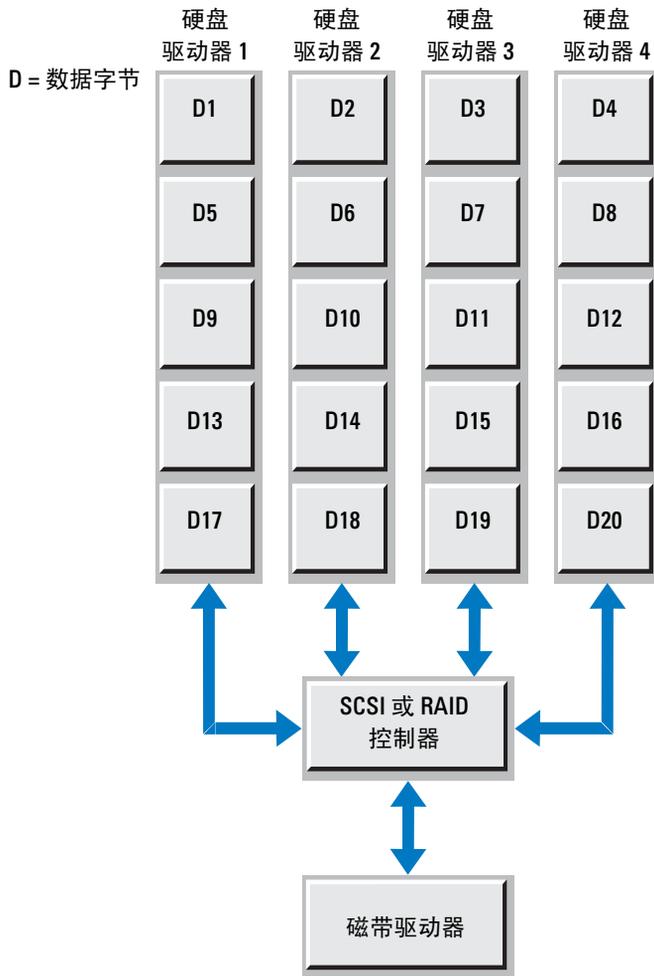
外部磁盘阵列（特别是在 SAN 中）可以提供更高的高速缓存容量，以提高 I/O 性能。这种高速缓存可以大大提高向阵列写入数据的性能，且可以存储经常访问的数据以提高读性能。至于它对磁带性能的影响，在向阵列恢复数据或从阵列备份数据时，高速缓存可以屏蔽大多数 RAID 限制。但是，使用高速缓存从外部阵列进行备份操作时，还是会感受到 RAID 配置限制所带来的影响，因为仍需从磁盘读取数据。

RAID 0

RAID 0 通常称为分拆，允许将两个或更多磁盘组合起来，以单一 LUN 方式生成一个虚拟驱动器。之所以称为分拆，是因为在同一时间数据将写入阵列中的所有磁盘，而不仅仅是一个磁盘。因此，吞吐量分布在 n 个信道上（ n 是阵列中硬盘驱动器的数量），而不是单一硬盘的单一信道上。这就提供了优异的读 / 写性能，但不具备容错性能。

图 1-2 显示了以 RAID 0 配置四个硬盘驱动器。数据将分拆到所有四个硬盘驱动器，这样，四个信道同时对阵列进行读和写操作。

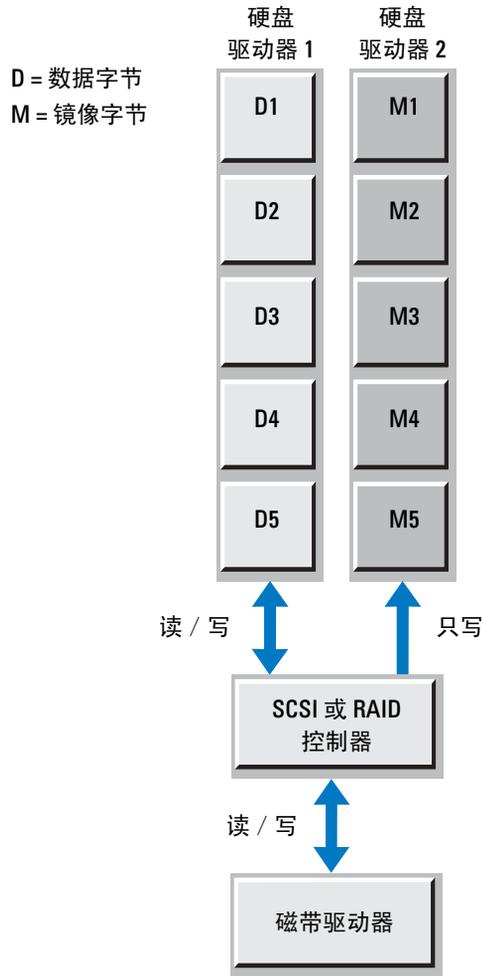
图 1-2. 示例 RAID 0 配置



RAID 1

RAID 1 也称作镜像，它意味着将数据同时写入两个磁盘中。如果其中一个驱动器出现故障，系统将切换到另一个驱动器，数据将不会丢失。在磁带驱动器备份期间，从 RAID 1 阵列读取数据的速率大约与从单一驱动器读取数据的速率相同，因为数据是从主驱动器读取的。然而，从磁带驱动器写入 RAID 1 阵列的恢复性能可能会较慢，这是由于在向主磁盘和镜像磁盘写入数据时，要进行错误检查 / 纠正 (ECC)。这种低效率很大程度上是由于镜像操作常常是在 CPU 或 RAID 控制器上执行而造成的。因此，更高级的 RAID 控制器由于使用了更新和功能更强大的处理器，其速度将更快。

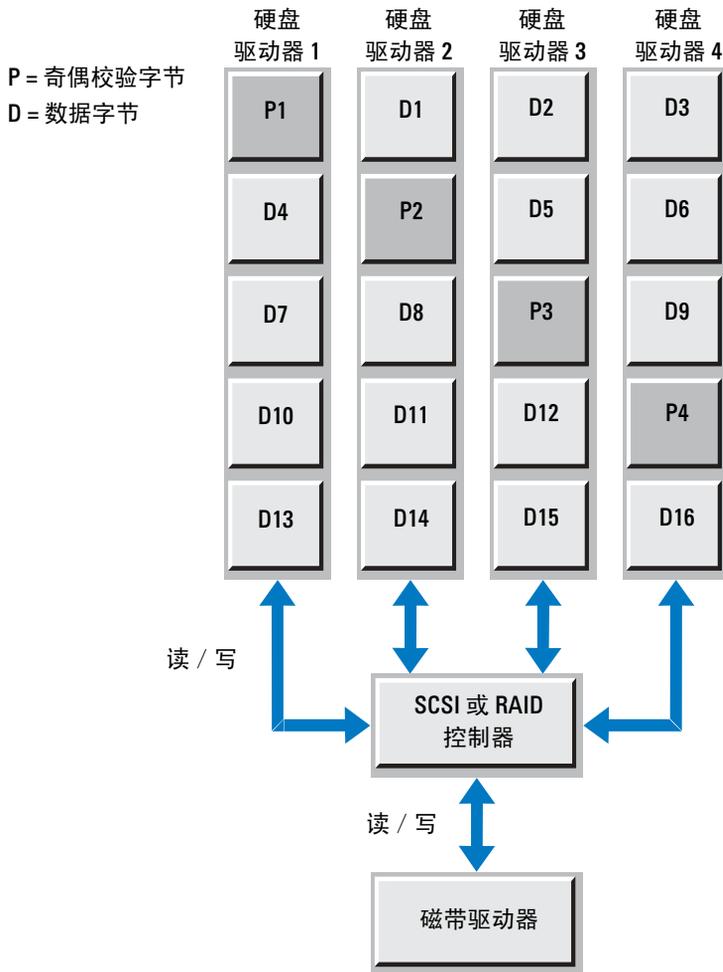
图 1-3. 示例 RAID 1 配置

**RAID 5**

使用 RAID 5 阵列，数据以字节级方式分拆到磁盘阵列中，且纠错数据或奇偶校验数据也分拆到磁盘阵列中。RAID 5 阵列具有非常好的随机读取性能，随着阵列中磁盘的数量增加，读性能通常会更好。使用较大的磁盘阵列，读性能实际上高于 RAID 0 阵列，因为数据分布在一个附加的驱动器上。另外，在执行正常的读操作时，不需要奇偶校验信息。

从磁带恢复到 RAID 5 阵列时，这种读优势往往是微不足道的，因为它会引发用于计算和写入奇偶校验信息的额外开销。

图 1-4. 示例 RAID 5 配置



总之，RAID 0 可以获得读和写性能的最佳整体配置，但没有考虑冗余性能。RAID 1 总体上效果最差，因为需要对所有写入阵列的数据进行镜像，且读取数据时只涉及一个磁盘。RAID 5 的读性能很好，但写性能一般；然而，如果阵列中添加更多的磁盘，RAID 5 性能将得以改进。如果 RAID 位于一个可提供更高高速缓存能力的硬盘柜中，则恢复操作期间的性能限制可能会减少。但备份操作仍会受到 RAID 配置限制的影响。另外，阵列的特性依然很大程度上依赖于“硬盘驱动器性能”中列出的具体硬盘驱动器的特性。

在磁带存储库中使用多个驱动器时的一般性能注意事项

同时使用多个磁带驱动器执行数据备份时（如在磁带存储库中），就必须考虑硬盘配置的其他方面。通过使用简单的注重性能的方式设置硬盘和布线配置，可以限制发生额外吞吐量瓶颈问题。

SCSI 配置

磁带存储库中提供的最新的高性能磁带驱动器支持 SCSI 接口标准的 Ultra160 规格。因此，要想获得最佳性能，采用 SCSI 的备份服务器必须安装 HBA，以支持 Ultra160 或更高的数据速度。满足这一要求的 SCSI HBA 允许每个磁带驱动器在 SCSI 总线上以 160 MB / 秒的速度与主机通信。SCSI 总线的速率高于磁带驱动器速率，这就允许多个驱动器连接到同一总线上，而不会牺牲设备的性能。但这仅仅是一个方面。

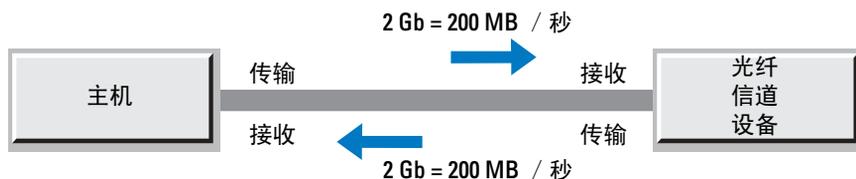
Ultra160 总线 160 MB / 秒的数据速率是对于连接到总线的所有设备所能提供的最大数据吞吐量。因此，单一磁带驱动器不会占用总线的全部带宽，因为它从磁带读数据或向磁带写入数据的速率最高为 80 MB / 秒（本机）。然而，对于多个磁带驱动器而言，如果每个驱动器都以其最佳本机性能运行，则联合起来可占用总线提供的全部 160 MB / 秒带宽。在这一临界点后，连接到该同一总线的每个附加驱动器都会降低每个驱动器的总体性能。

因此，要想从磁带存储库获得最佳性能，建议每个 SCSI 总线所连接的磁带驱动器数量不超过两个。有关详细信息和说明，请参阅“建议的布线配置”。如果磁带驱动器的规格是 Ultra160，应使用至少支持 Ultra160 的 SCSI HBA，但升级到 Ultra320 HBA 将不会使性能有所改善。

SAN 配置

光纤信道 (FC) 与 SCSI 相比提供了更多的优势。首先，它不仅克服了 SCSI 的距离限制（与 300 米的短波 2 Gb FC 链路相比，LVD SCSI 的距离为 12 米），还能够以更高速度传输数据。由于 FC 是一个串行网络协议，而不是基于总线的体系结构（如 SCSI），因此它还可以成为用于实施 SAN 的理想协议，并允许合并数据存储资源。另外，每个 FC 连接都是由一个传输链路和一个接收链路组成的，允许全双工操作。这意味着数据可以在两个方向同时进行传输。因此，在单一 FC 连接之间执行备份操作时，可以从源读取数据，同时向磁带中写入数据，而不必轮流通信，从而有效地将连接的带宽增加了一倍。请参阅图 1-5。

图 1-5. 光纤信道链路图表

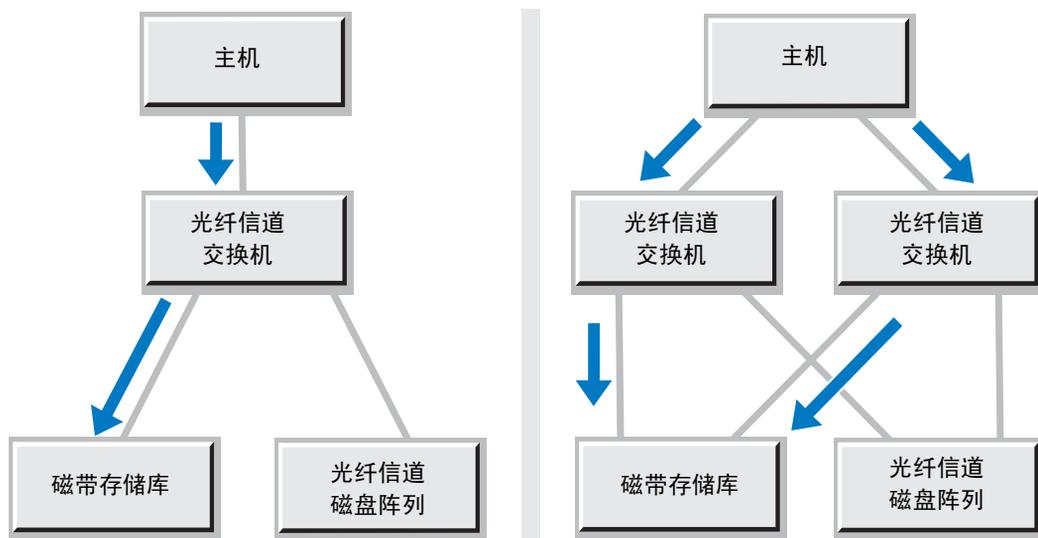


在 SAN 中设置磁带存储库时，性能仍然会受到各种因素的影响。包括 FC 链路速度，源与磁带存储库之间的数据流，以及外部存储阵列的性能限制。随着对解决方案的总体设置和管理的逐步了解，可以避免其中的许多因素。

即使 SAN 中的 FC 协议提供了很高的数据带宽，也必须适当地考虑磁带驱动器的各个方面，以避免发生 FC 链路可能限制性能的情况。2 千兆位 (Gb) FC 链路的数据速率为 200 MB / 秒（即传输链路上数据速率为 200 MB / 秒，而接收链路上的数据速率同样为 200 MB / 秒）。因此，在同一链路上尝试运行多个磁带驱动器可能会超过该链路的全部带宽。如果主机使用传统 1 Gb 适配器运行，则在向两个驱动器备份数据时，可能会很明显地暴露出性能限制。

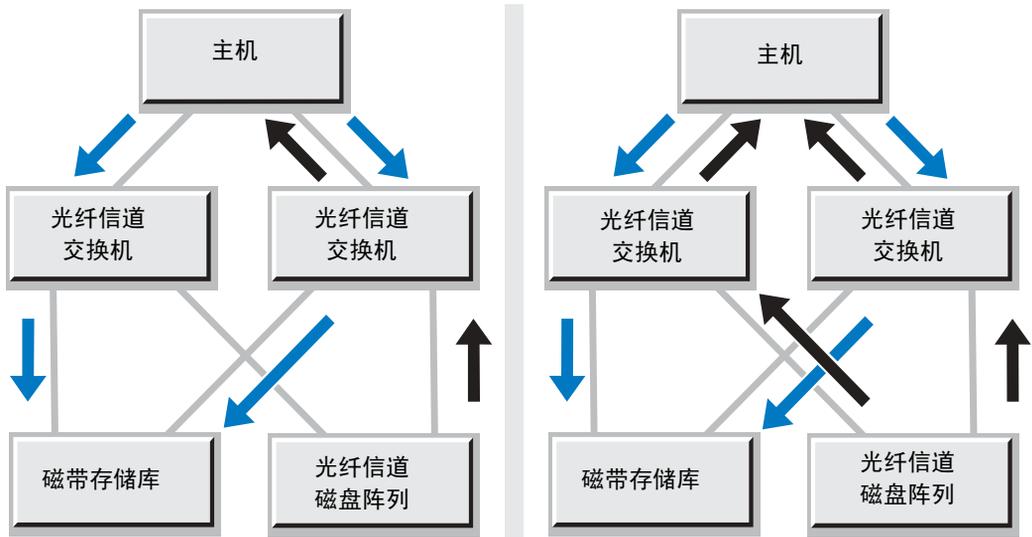
因此，在一条 2 Gb 链路上同时使用三个或更多磁带驱动器时，可能需要将备份工作分配到多个连接上，而不是依赖于单一链路。这样，您就了解了 SAN 解决方案拓扑结构的优势所在。了解在备份过程中从源读数据和向磁带写入数据的数据通道，可帮助管理员认识到任何潜在的瓶颈问题。如果确定了任何瓶颈，则可以根据配置采取适当的措施。例如，如果备份解决方案一次需要多个驱动器进行操作，则通过拆分连接将磁带硬件分配到若干单独的光纤通道上，以改善性能。请参阅图 1-6。

图 1-6. 单一数据流与拆分数据流（到磁带存储库）



同时，如果瓶颈存在于要从外部 FC 磁盘阵列读取的数据中，则使用 I/O 管理软件（如 EMC® PowerPath®）以及一个额外光纤连接可以自动地实现负载均衡，从而将数据均衡分配到多个通道中，并通过通道故障转移提高可用性。请参阅图 1-7。图的左侧表示一个 SAN 磁盘阵列，在这个阵列中所有数据被迫通过单一链路，导致产生瓶颈，降低了数据流到磁带硬件的速度。右侧显示了负载均衡如何通过让数据使用两条链路进行传输，从而使来自阵列的 I/O 带宽翻倍。

图 1-7. 瓶颈数据流与负载均衡磁盘阵列



最后，SAN 上的 FC 磁盘阵列也可能会遇到在“硬盘驱动器和 RAID 阵列配置”中介绍的相同的性能限制因素。因此，通过改进磁盘阵列的性能特性，可直接影响到在 SAN 上进行备份的速度。

使用存储库光纤信道桥接器进行 SAN 配置

某些磁带存储库可以利用光纤信道桥接器模块连接到 SAN。该模块充当 SCSI 与 FC 协议之间的桥梁，可以提供大多数本机 FC 存储库所不具备的额外管理、安全和操作功能。有关这些功能的详细信息，请参阅磁带存储库的光纤信道桥接器用户指南。

在某些磁带存储库配置中，光纤信道桥接器模块可能会成为瓶颈，而降低磁带驱动器的性能。这是因为要在 SCSI 与 FC 通信之间起到桥接作用，光纤信道桥接器模块中的处理功能需要足够强大，但此时处理功能无法满足处理特定多驱动器配置所提供的集合数据吞吐量的需要。尽管如此，在大多数数据备份解决方案中，光纤信道桥接器模块不会成为限制磁带性能的主要因素。专用备份服务器经常会遇到以下这种情况，即由于要向多个磁带驱动器提供数据，主机上的多种限制会结合起来发挥作用。这会导致驱动器的总体性能低于一定级别，而使光纤信道桥接器模块变为限制因素。

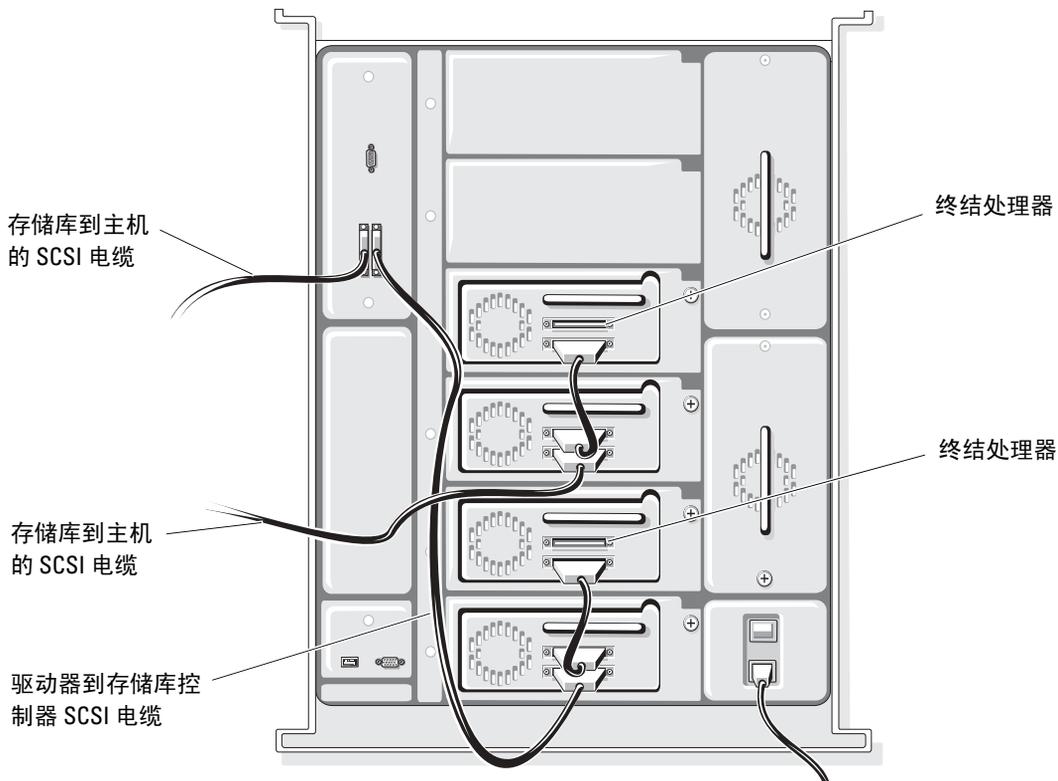
建议的布线配置

SCSI与主机之间的布线

配有最多六个磁带驱动器的磁带存储库

当磁带存储库通过 SCSI 连接到主机时，应确保单一总线上不超过两个驱动器。对于配有五至六个磁带驱动器的存储库，需添加其它 SCSI 控制器，以确保 SCSI 总线不会成为障碍，从而实现最大吞吐量。

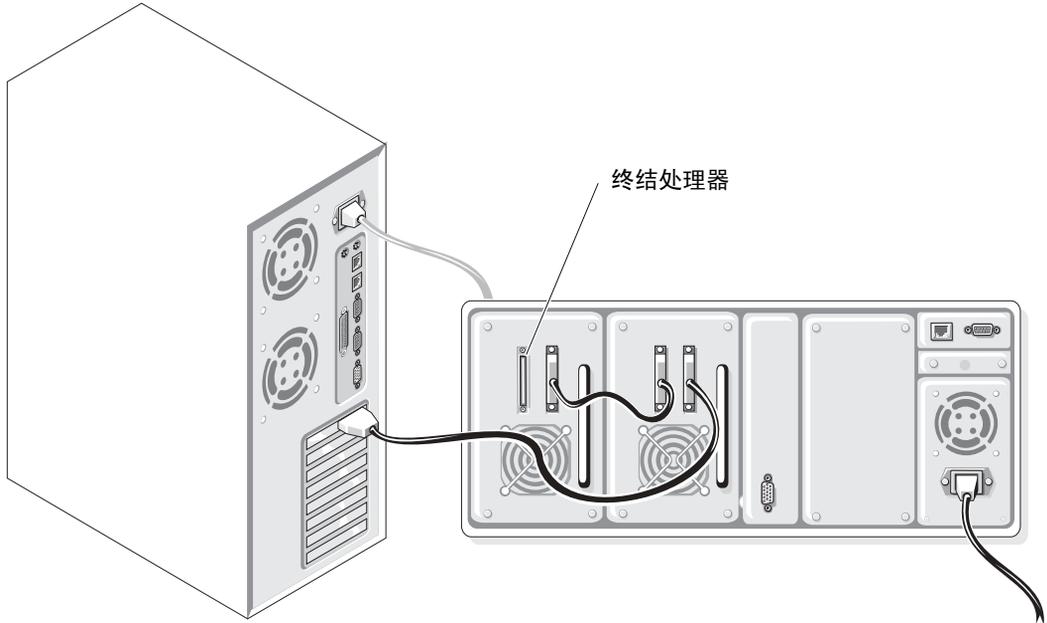
图 1-8. 配有最多六个磁带驱动器的存储库的 SCSI 布线



配有最多两个磁带驱动器的磁带存储库

位于完全配置的双驱动器磁带存储库中的驱动器可以连接到位于同一 SCSI 总线上的主机，而不会对备份性能产生严重限制。位于单一 SCSI 总线上的两个驱动器的备份速率与两个驱动器分别位于单独总线上的备份速率持平。然而，在备份应用程序中启动验证功能的用户可能希望通过将两个驱动器拆分到两个 SCSI 总线上来改进该验证性能。这样一来，验证性能可以提供多达百分之二十五。

图 1-9. 配有最多两个磁带驱动器的存储库的 SCSI 布线



将驱动器连接至光纤信道桥接器

图 1-10 至图 1-17 图示说明了如何使用光纤信道桥接器模块配置一个配有最多六个驱动器的磁带存储库，以便通过 FC 优化磁带性能。

图 1-10. 采用一个磁带驱动器时的光纤信道桥接器布线

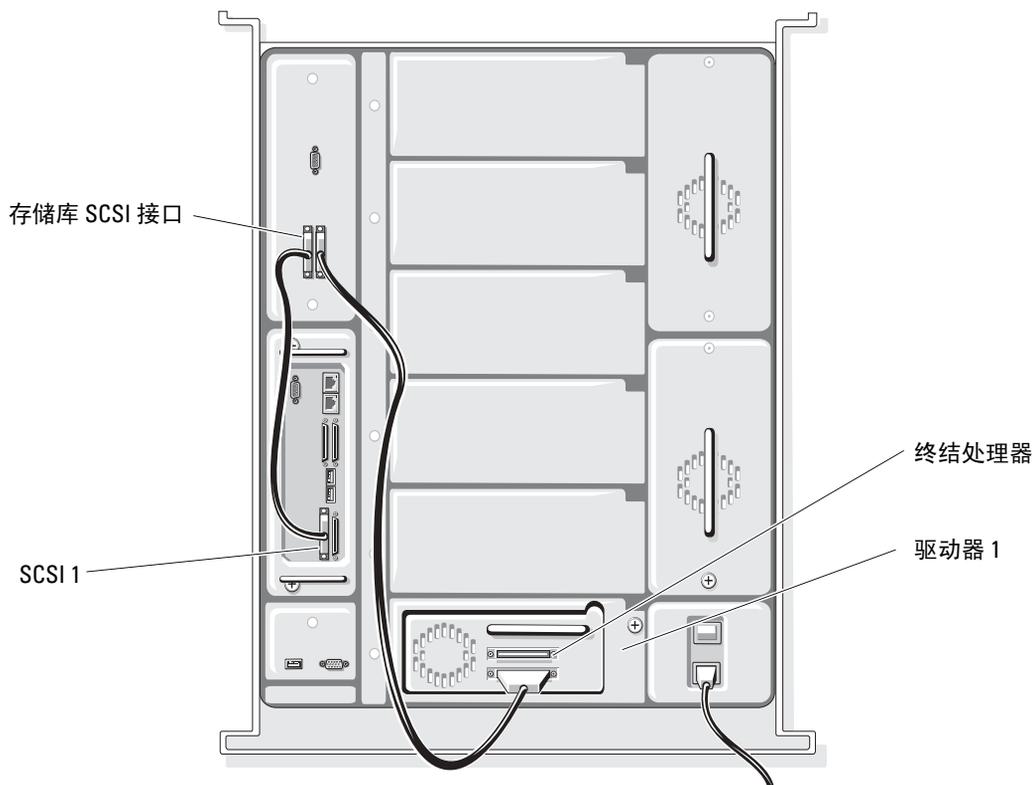


图 1-11. 采用两个磁带驱动器时的光纤信道桥接器布线

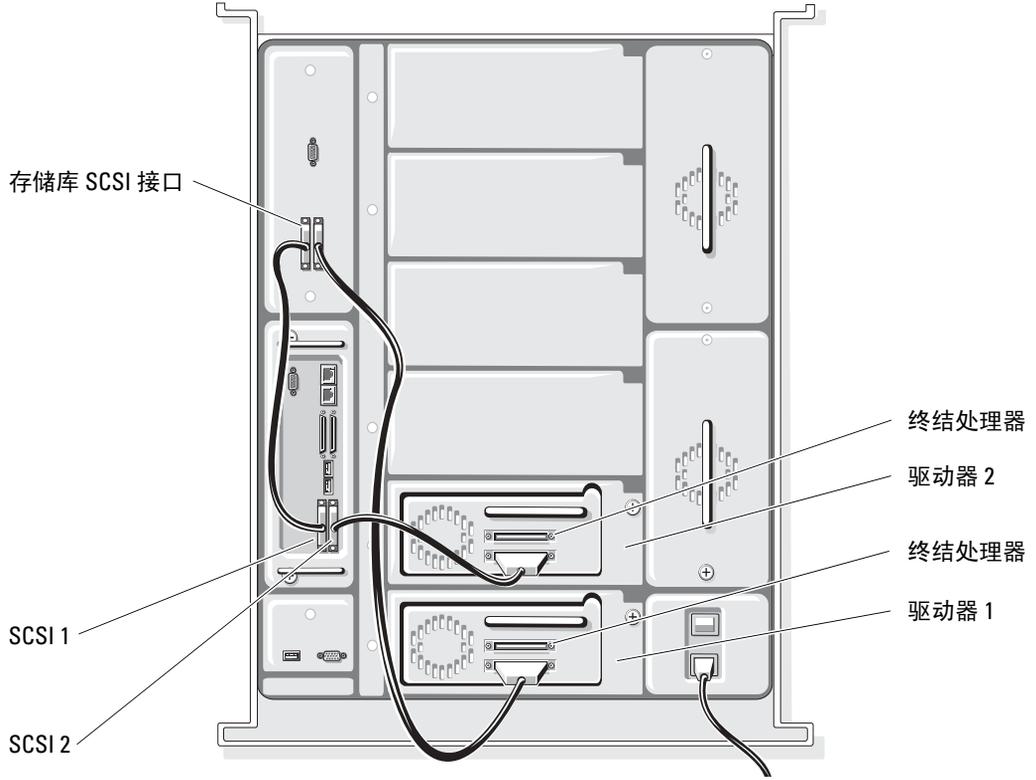


图 1-12. 采用三个磁带驱动器时的光纤信道桥接器布线

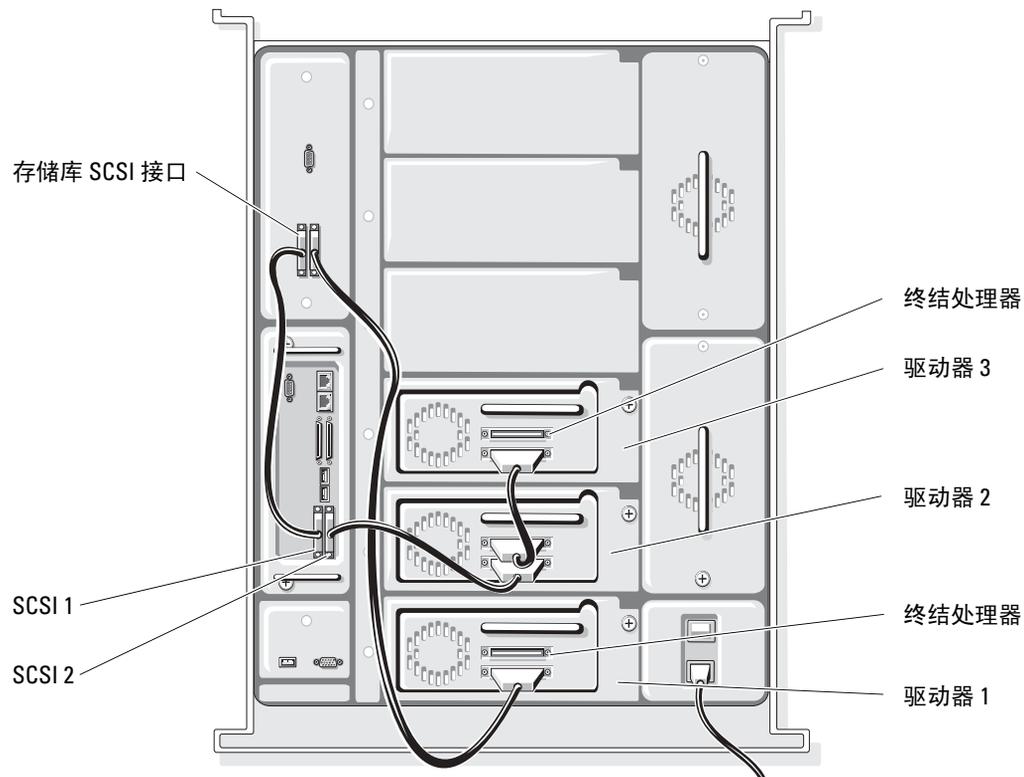


图 1-13. 采用四个磁带驱动器时的光纤信道桥接器布线

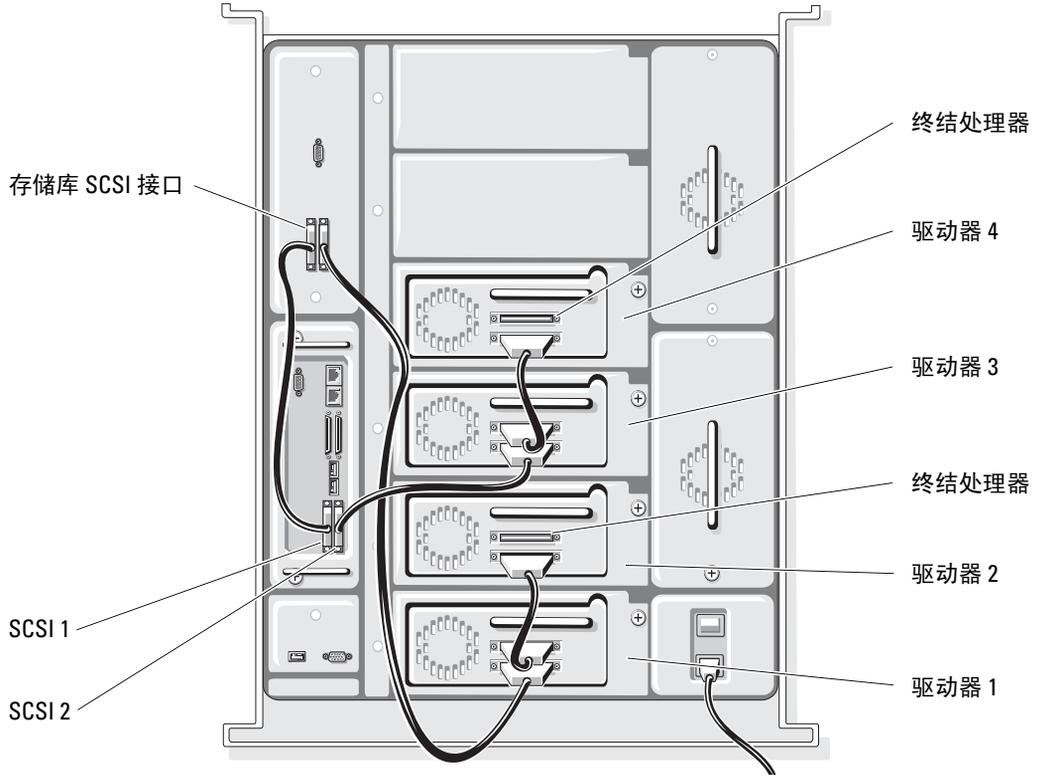


图 1-14. 配有一到四个驱动器的磁带存储库的信道分区设置

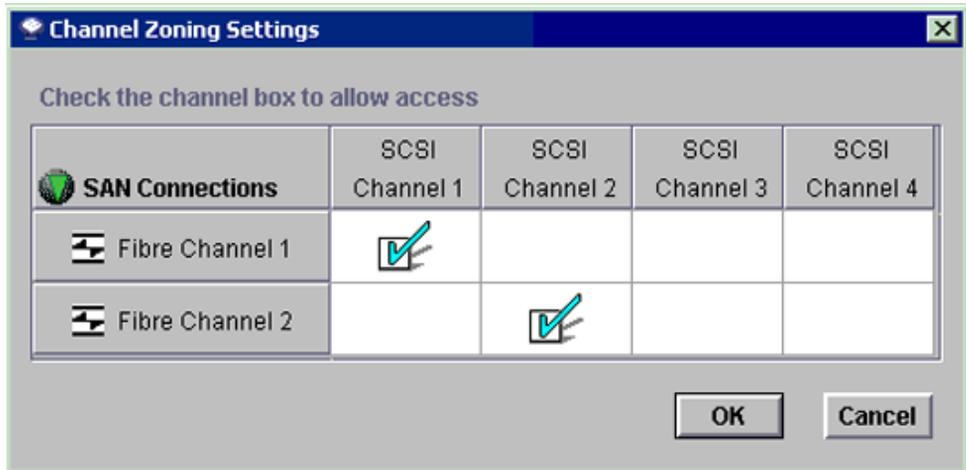


图 1-15. 采用五个磁带驱动器时的光纤信道桥接器布线

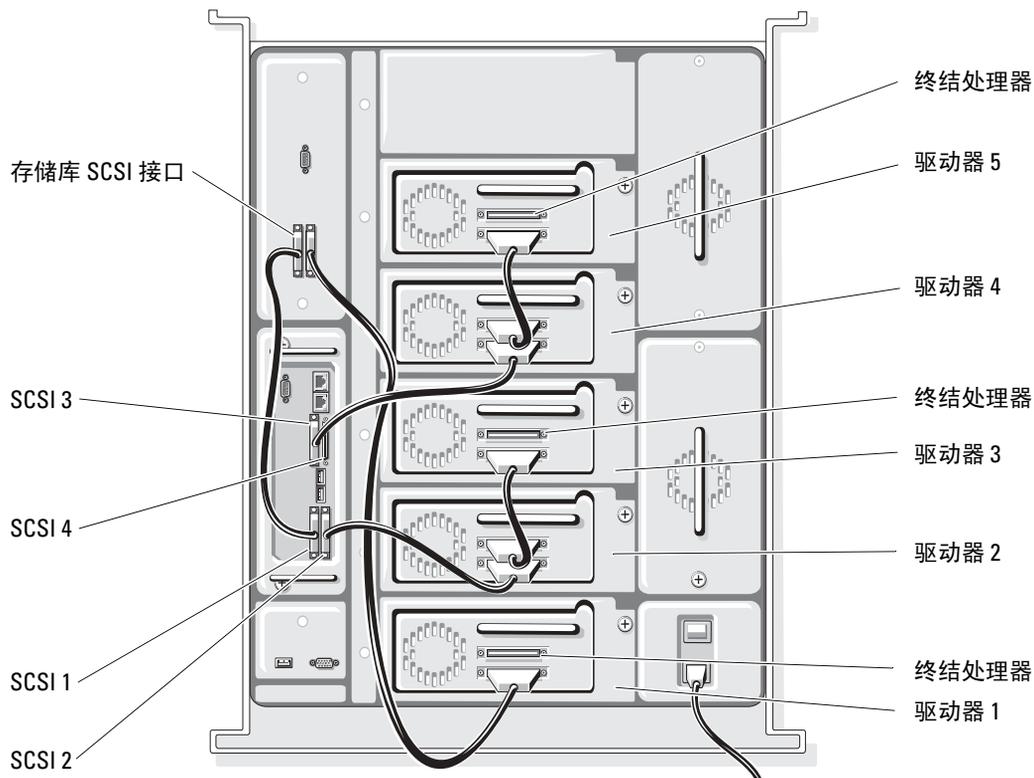


图 1-16. 采用六个磁带驱动器时的光纤信道桥接器布线

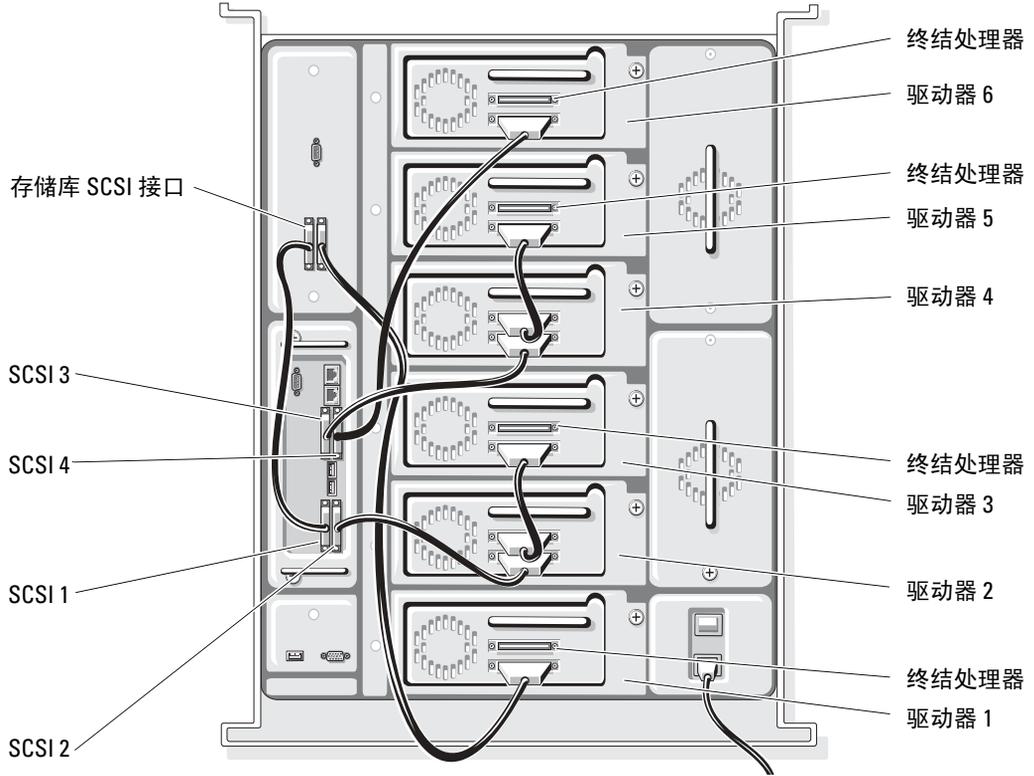


图 1-17. 配有五或六个驱动器的磁带存储库的信道分区设置

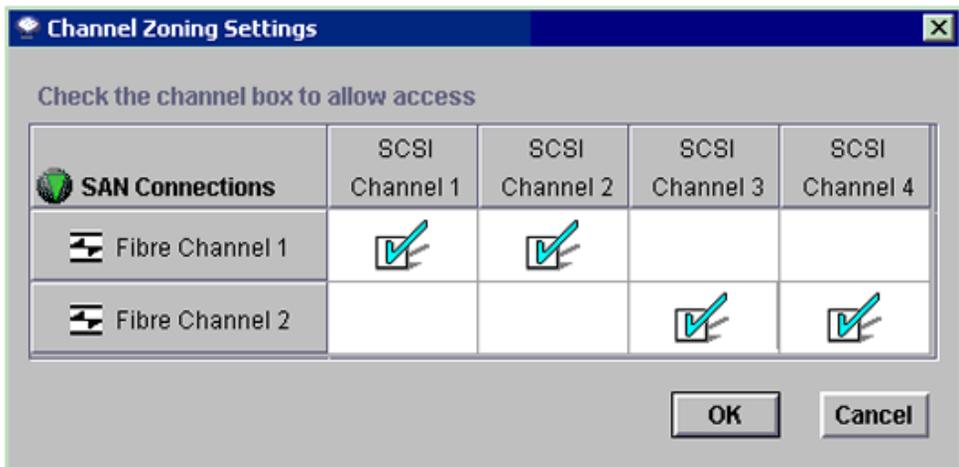


图 1-18 和图 1-19 图示说明了如何使用光纤信道桥接器模块配置一个配有最多两个磁带驱动器的磁带存储库，以便通过 FC 优化磁带性能。

图 1-18. 配有一个驱动器的磁带存储库

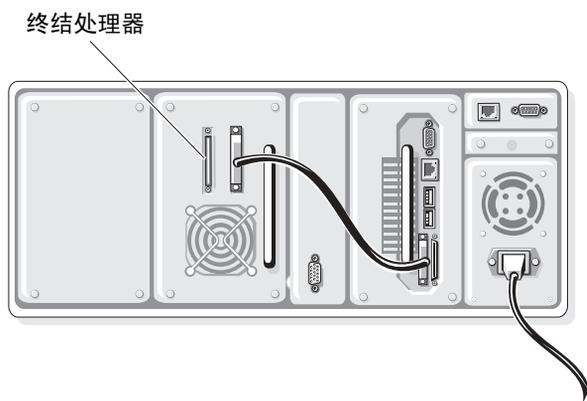
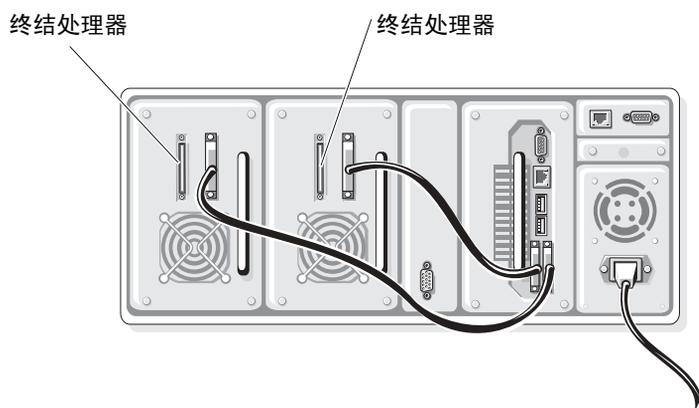


图 1-19. 配有两个驱动器的磁带存储库



Systemes Dell™ PowerVault™

**Remarques sur
les performances
des lecteurs et des
bibliothèques de bandes**

Les informations contenues dans ce document peuvent être modifiées sans préavis.

© 2005 Dell Inc. Tous droits réservés.

La reproduction de ce document de quelque manière que ce soit sans l'autorisation écrite de Dell Inc. est strictement interdite.

Marques utilisées dans ce document : *Dell*, le logo *DELL* et *PowerVault* sont des marques de Dell Inc. ;
EMC et *PowerPath* sont des marques déposées d'EMC Corporation.

Tous les autres noms de marques et marques commerciales utilisés dans ce document se rapportent aux sociétés propriétaires des marques et des noms de ces produits. Dell Inc. décline tout intérêt dans l'utilisation des marques déposées et des noms de marques ne lui appartenant pas.

Juin 2005

Sommaire

Présentation	53
Remarques générales sur la sauvegarde du système hôte	53
Remarques sur les données et les lecteurs de bande	53
Configuration des disques durs et des matrices RAID	55
Remarques générales sur les performances lors de l'utilisation de bibliothèques de bandes à plusieurs lecteurs.	61
Configurations SCSI.	61
Configurations SAN.	62

Présentation

Récemment, les vitesses de transfert des lecteurs de bande ont évolué de manière significative. Pour cette raison, un grand nombre de facteurs liés au système hôte, tels que la configuration RAID et les spécifications des disques durs, doivent être pris en compte pour savoir si le serveur et le lecteur de bande peuvent traiter les données à débit équivalent. Voir la section “Remarques générales sur la sauvegarde du système hôte” pour plus d’informations sur les configurations et les attributs provoquant une réduction du débit entre le serveur hôte et le lecteur de bande.

Étant donné que les bibliothèques de bandes regroupent plusieurs lecteurs, il est nécessaire d’utiliser des bandes passantes hôtes plus importantes de manière à pouvoir utiliser le débit potentiel de tous les lecteurs. Voir la section “Remarques générales sur les performances lors de l’utilisation de bibliothèques de bandes à plusieurs lecteurs” pour plus d’informations sur les configurations de câblage recommandées et les restrictions potentielles liées à l’utilisation de fibres optiques sur les unités à plusieurs lecteurs.

Remarques générales sur la sauvegarde du système hôte

Les remarques figurant dans cette section s’appliquent aux configurations de sauvegarde sur bande SCSI et SAN.

Remarques sur les données et les lecteurs de bande

Lors de l’évaluation des performances, les points suivants doivent être pris en compte :

Surcharge causée par les commandes SCSI. La surcharge de commandes sur le bus SCSI a un impact sur les vitesses de transfert théoriques de *tous* les périphériques SCSI. Le logiciel de sauvegarde sur bande ne prend pas cette surcharge en compte. Il mesure uniquement la vitesse à laquelle les données sont écrites sur la bande. À titre d’exemple, il est possible que le lecteur traite un débit de 80 Mo/s mais écrive les données à un débit de 77 Mo/s. Le logiciel de sauvegarde ne prend en compte que le dernier débit.

Taille des blocs de bande. La taille optimale des blocs pour la plupart des lecteurs de bande est de 64 Ko. Bien que l’utilisation de blocs plus importants n’améliore en rien les performances, certaines applications de sauvegarde permettent à l’utilisateur de modifier cette taille. En revanche, l’utilisation de blocs de moins de 64 Ko risque de diminuer les performances. Reportez-vous au guide d’utilisation du logiciel de sauvegarde pour plus d’informations sur l’ajustement de la taille des blocs de l’unité de bande.

Taille de tampon du logiciel de sauvegarde. Plus la taille des tampons du logiciel de sauvegarde est importante, plus il est possible d'optimiser les performances de sauvegarde. Certaines applications permettent à l'utilisateur de modifier cette taille, ce qui peut aider à assurer un flux constant des données en entrée et en sortie sur le lecteur et augmenter de façon significative les vitesses de transfert, particulièrement pour des fichiers de petite taille. Plus la taille du tampon est importante, plus ce dernier peut contenir de données. Le disque passe ainsi moins de temps à rechercher ces dernières. Cependant, une taille de tampon trop importante risque de diminuer les performances de la mémoire et du processeur. Reportez-vous au guide d'utilisation de l'application de sauvegarde pour plus de détails.

Pilotes et micro-code. Vérifiez toujours que le contrôleur SCSI ou Fibre Channel et le lecteur de bande utilisent la dernière version des pilotes et du micro-code. Les pilotes et le micro-code les plus récents des lecteurs de bande Dell PowerVault peuvent être téléchargés à partir du site support.dell.com.

Reliez les lecteurs de bande et les disques durs à des contrôleurs distincts (internes ou externes) se trouvant sur des cartes HBA différentes. Il est préférable, pour permettre un débit optimal, que la carte HBA du lecteur de bande et celle du disque dur soient différentes (cela dépend toutefois des performances du contrôleur). La plupart des cartes double fonction SCSI/RAID intégrées n'ont qu'un seul processeur qui doit partager la bande passante entre la matrice RAID et le lecteur de bande. Un contrôleur peut ainsi traiter les opérations de lecture et d'écriture entre les disques durs et le lecteur de bande. Il peut également calculer et écrire toute information de parité requise sur les disques durs. Voir la section "Remarques sur la configuration du RAID matériel" pour plus de détails sur les matrices RAID et les octets de parité.

Têtes de lecture sales ou supports obsolètes. Des têtes de lecture sales ou un support obsolète peuvent être à l'origine d'un fort taux d'erreurs et d'une réduction des vitesses de lecture et d'écriture. Les performances diminuent chaque fois qu'un lecteur tente de réécrire ou de relire un bloc. Lorsqu'un seuil d'erreurs de lecture et d'écriture donné est atteint, le lecteur doit être nettoyé. Il est important d'effectuer ce nettoyage à des intervalles réguliers ou lorsque le système vous y invite.

La probabilité de trouver un bloc endommagé augmente avec l'âge du support ou si celui-ci est utilisé de façon excessive. L'espérance de vie d'un support LTO correspond en général à 75 opérations complètes de lecture/écriture.

Correspondance des débits. Les lecteurs LTO récents se règlent sur la vitesse des données entrantes. La limite inférieure correspond à environ la moitié de la vitesse maximum de transfert des données non compressées. Si la vitesse utilisée pour le transfert des données entrantes est inférieure à ce seuil minimal, le lecteur doit s'arrêter, attendre que le tampon soit plein, rembobiner et essayer d'écrire le contenu du tampon (cette opération est appelée "repositionnement").

Par exemple, pour le lecteur de bande Dell PowerVault 110T (LTO2 et LTO3), le débit minimal est de 30 Mo/s lors d'une opération d'écriture sur un support LTO-3. Si le serveur hôte peut uniquement transmettre des données à un débit de 20 Mo/s, le lecteur se "repositionne" en attendant que les tampons soient pleins. Dans ce cas, le débit effectif sera légèrement inférieur à 20 Mo/s (environ 15 Mo/s).

Vérification des performances du lecteur de bande

Certains fabricants de lecteurs de bande intègrent à leurs appareils un mode de diagnostic des performances pouvant être utilisé pour vérifier le débit. Le PowerVault 110T LTO-2 et LTO-3 (version 53XX ou suivante du micro-code) comprend le mode de diagnostic “F” qui permet d'effectuer un test rapide des performances de lecture/écriture sur le lecteur et le support. Si le débit constaté diffère de 6 % par rapport à la vitesse maximum spécifiée pour le lecteur, le test échoue et un message d'erreur s'affiche. Si le test réussit, aucun message d'erreur ne s'affiche. Consultez le guide d'utilisation du lecteur de bande pour plus de détails sur le mode de diagnostic “F”.



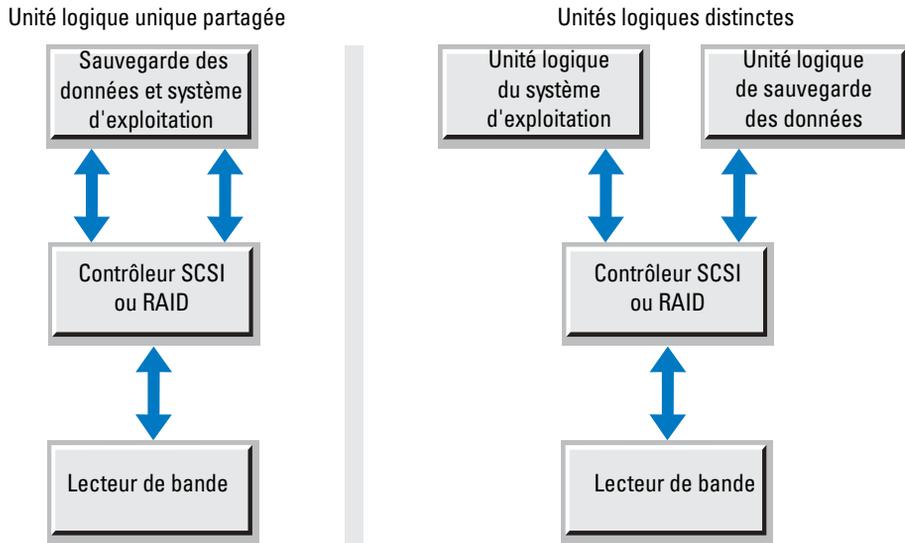
AVIS : pour le test du mode de diagnostic “F”, vous devez utiliser un support dont les données peuvent être écrasées. N'utilisez pas de support contenant des données importantes. Toutes les données sont supprimées au cours du test.

Configuration des disques durs et des matrices RAID

Plusieurs attributs des disques durs et des matrices de disques (internes et externes) peuvent affecter les performances de sauvegarde ou de restauration. Les sous-sections ci-après présentent ces attributs et les configurations recommandées pour optimiser les vitesses de sauvegarde et de restauration. Si le débit moyen du lecteur de bande est supérieur à celui de la matrice de disques, les performances du lecteur ne seront pas optimales.

Remarques générales sur la configuration des disques durs

Système d'exploitation et données sur des unités logiques distinctes. Il est recommandé de sauvegarder les données sur une unité logique différente de celle contenant le système d'exploitation. Cela permet d'éviter que le disque dur doive traiter à la fois les opérations liées au système d'exploitation et celles liées à la sauvegarde. Pour ce faire, il est possible de dédier un disque dur ou une matrice de disques au système d'exploitation et un(e) autre aux données à sauvegarder.

Figure 1-1. Bande passante simple canal/double canal**Performances des disques durs**

Les lecteurs de bande écrivent les données de façon séquentielle. La continuité de ce mode de fonctionnement doit être assurée par un flux de données constant évitant un repositionnement. À l'inverse, les disques durs sont des périphériques à accès aléatoire. Ils peuvent donc rencontrer des difficultés pour envoyer des données de façon séquentielle aux lecteurs de bande si ces données sont réparties sur toute la surface du disque, car cela les oblige à rechercher continuellement de petits blocs de données.

De plus, d'autres attributs des disques durs peuvent affecter la vitesse à laquelle les données sont transmises au lecteur de bande.

Vitesse de rotation du disque. La vitesse de rotation est mesurée en rpm (tours par minute). Elle détermine le nombre de révolutions complètes effectuées par le disque dur en une minute. Ce paramètre affecte directement les temps de l'accès aléatoire aux données et la vitesse des transferts séquentiels. Plus la vitesse est élevée, plus le disque accède rapidement aux données.

Temps d'accès aléatoire. Le temps d'accès est mesuré en millisecondes. Il correspond au temps que met la tête de lecture à trouver un fragment de données sur le disque. Sur un disque dur, il mesure le temps nécessaire aux têtes de lecture/écriture pour se déplacer entre les pistes. Les données peuvent être stockées sur n'importe quel secteur du disque. Plus l'accès aux données est long, plus le débit global du disque est lent. Cet attribut est très important lorsque le disque dur contient un grand nombre de fichiers de petite taille, car la vitesse de lecture et d'écriture est ralentie par le nombre de recherches que le lecteur doit effectuer.

Vitesses de transfert séquentielles/moyennes (STR). Ce paramètre correspond à la vitesse à laquelle le lecteur lit et écrit les données sur le disque. Si les données à sauvegarder se trouvent dans un seul fichier contigu, la vitesse moyenne est proche de la vitesse maximale de transfert du disque. Cependant, en général, les opérations successives d'écriture et de suppression entraînent une fragmentation des données, qui sont réparties en divers endroits du disque. La défragmentation du disque dur permet d'optimiser la vitesse de transfert.

Tampon (cache). Le tampon est la partie de la mémoire qui contient les données les plus récemment écrites ou enregistrées sur le disque. Plus il est important, plus il peut contenir de données, réduisant ainsi la durée des recherches sur le disque.

Remarques sur la configuration du RAID matériel

Présentation générale de la configuration RAID

Cette section présente les configurations RAID habituelles et leur impact sur les vitesses de sauvegarde et de restauration. Une matrice RAID est un ensemble de disques durs agissant comme un système de stockage ou comme une unité logique unique. Les données peuvent être transférées simultanément par le canal de chaque disque dur, ce qui permet d'obtenir un débit total correspondant à un multiple du nombre total de lecteurs de la matrice (moins la surcharge et les fonctions de redondance - voir les sections ci-après).

Dans le cas d'une configuration RAID, la vitesse de l'interface est importante car les lecteurs partagent la même bande passante. Par exemple, si un lecteur Ultra160 offre un débit de 40 Mo/s, une matrice RAID 0 composée de 5 disques identiques doit pouvoir prendre en charge un débit de 200 Mo/s. Cependant, l'interface du lecteur Ultra160 limite le débit de la matrice à un maximum de 160 Mo/s.

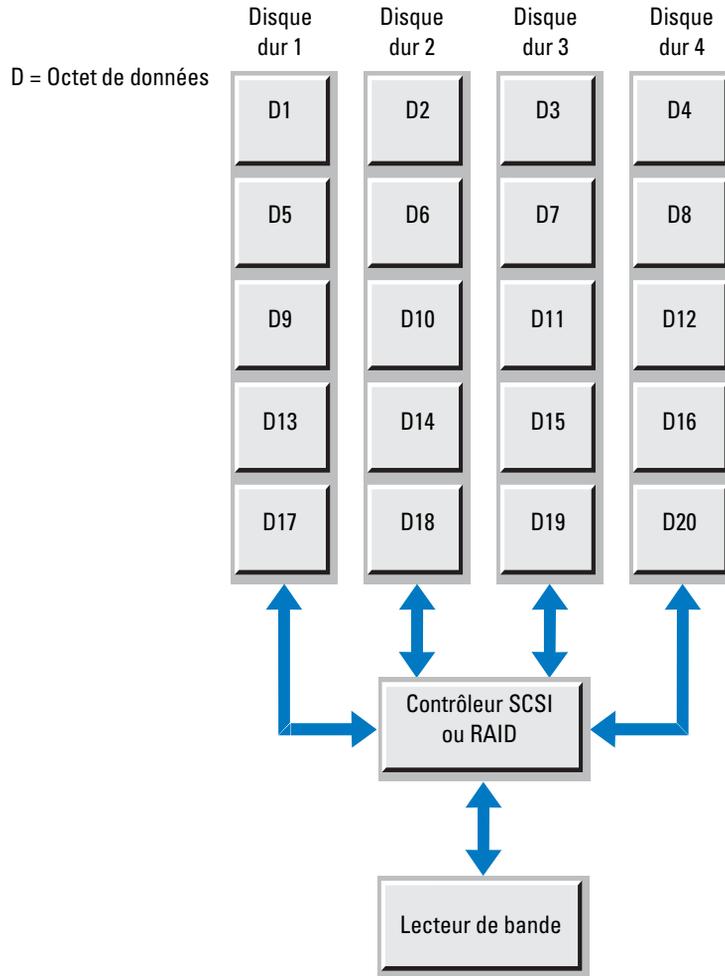
Les matrices de disques externes, particulièrement dans les SAN, peuvent offrir des tailles de mémoire cache importantes optimisant les performances des E-S, notamment lors de l'écriture de données sur la matrice. En outre, le stockage des données fréquemment utilisées améliore les performances de lecture. Dans le cas des lecteurs de bande, la mémoire cache compense la plupart des restrictions imposées par la configuration RAID lors de la restauration de données sur la matrice ou de la sauvegarde de données à partir de celle-ci. Cependant, les données devant être lues à partir des disques, il est possible que les sauvegardes effectuées à partir de matrices externes dotées d'une mémoire cache subissent l'impact de ces restrictions.

RAID 0

Le RAID 0, ou "striping" (étalement des données), permet d'associer deux ou plusieurs disques pour créer un lecteur virtuel unique semblable à unité logique. Les données sont écrites simultanément sur tous les disques de la matrice au lieu d'un seul à la fois. Le débit est ainsi réparti sur n canaux (n correspondant au nombre de disques durs de la matrice), ce qui permet d'obtenir d'excellentes performances de lecture/écriture. L'inconvénient de cette configuration est qu'elle n'offre pas de tolérance aux pannes.

La figure 1-2 montre quatre disques durs configurés en RAID 0. Les données sont réparties sur les quatre disques durs, ce qui permet de disposer de quatre canaux de lecture et d'écriture sur la matrice.

Figure 1-2. Exemple de configuration RAID 0

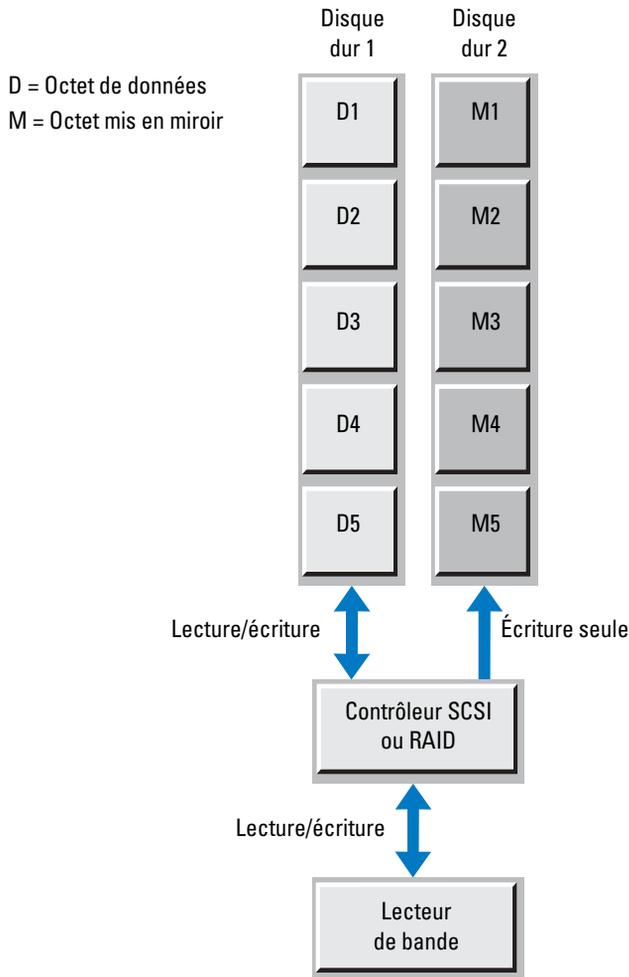


RAID 1

La configuration RAID 1 consiste en une mise en miroir des données, c'est-à-dire en leur écriture simultanée sur deux disques. Si un disque tombe en panne, le système bascule sur l'autre et aucune donnée n'est perdue. Lors de sauvegardes sur bande, la vitesse de lecture d'une matrice RAID 1 est approximativement la même que pour un seul disque, car les données sont lues sur le lecteur principal.

Cependant, la vérification ECC (vérification et correction d'erreurs) qui est effectuée lors de l'écriture sur les disques principal et miroir peut réduire les performances observées lors de la restauration de données à partir du lecteur de bande vers la matrice. Ce ralentissement est dû au fait que la mise en miroir s'effectue souvent via le processeur ou le contrôleur RAID. Les capacités supérieures des nouveaux processeurs permettent généralement aux contrôleurs RAID récents d'être plus rapides.

Figure 1-3. Exemple de configuration RAID 1



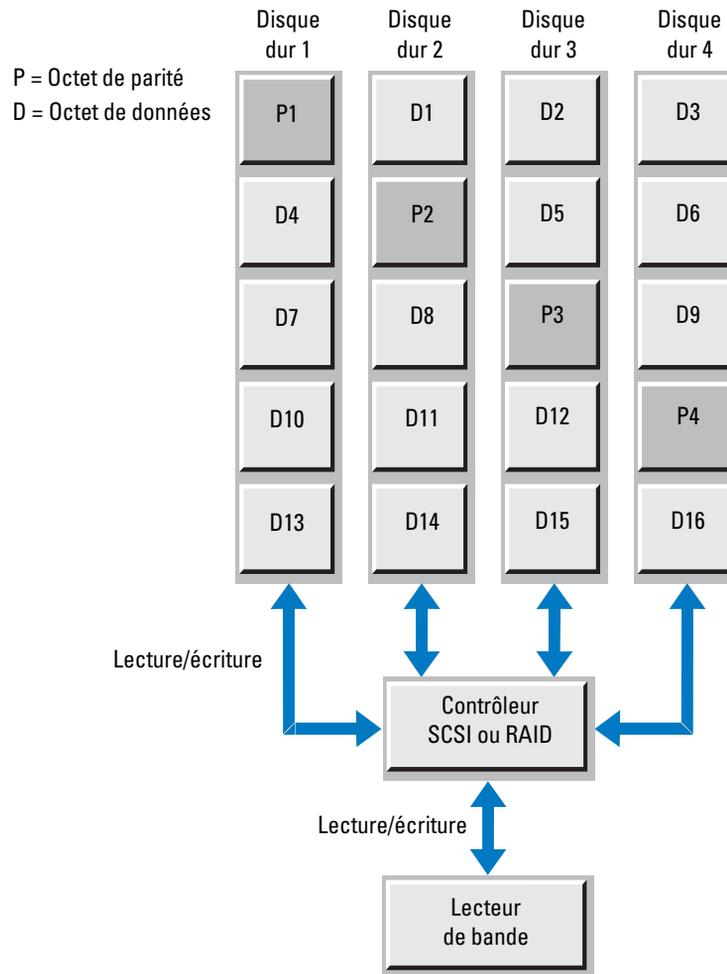
RAID 5

Dans une configuration RAID 5, les données sont réparties sur la matrice de disques au niveau des octets, ainsi que les données de parité (ECC). Les performances de lecture aléatoire des matrices RAID 5 sont généralement très bonnes.

Plus la matrice contient de disques, plus les performances s'améliorent. Les matrices comprenant un grand nombre de disques peuvent afficher des performances de lecture supérieures à celles des matrices RAID 0, car les données sont réparties sur un lecteur supplémentaire. De plus, les opérations de lecture normales ne requièrent aucune information de parité.

Les restaurations de données à partir d'une bande vers une matrice RAID 5 sont rares car elles impliquent le calcul et l'écriture des informations de parité, ce qui entraîne une surcharge supplémentaire.

Figure 1-4. Exemple de configuration RAID 5



En résumé, la configuration RAID 0 offre généralement les meilleures performances en lecture/écriture, mais elle a pour défaut de n'offrir aucune fonction de redondance. Le niveau RAID 1 est le moins performant, car toutes les données écrites sur la matrice sont mises en miroir tandis que les lectures sont effectuées à partir d'un seul disque. Le niveau RAID 5 offre généralement de bonnes performances de lecture. Les performances d'écriture sont médiocres mais peuvent être améliorées par l'ajout de disques dans la matrice. Si la matrice RAID est intégrée à une baie offrant une mémoire cache importante, il est possible de réduire les baisses de performances subies lors des restaurations. En revanche, les sauvegardes sont toujours soumises aux restrictions liées à la configuration RAID. En outre, les caractéristiques de la matrice sont toujours intimement liées à celles des disques durs (voir la section "Performances des disques durs").

Remarques générales sur les performances lors de l'utilisation de bibliothèques de bandes à plusieurs lecteurs

Lorsque plusieurs lecteurs sont utilisés simultanément pour les sauvegardes (comme dans le cas des bibliothèques de bandes), divers aspects supplémentaires de la configuration matérielle doivent être pris en compte. Il est possible de réduire les ralentissements du débit en utilisant des méthodes simples, axées sur les performances, lors de la configuration du matériel et du câblage.

Configurations SCSI

Les lecteurs de bande haute performance les plus récents intégrés aux bibliothèques de bandes prennent en charge la spécification Ultra160 de la norme d'interface SCSI. Pour des performances optimales, les serveurs de sauvegarde utilisant la norme SCSI doivent être équipés d'une carte HBA prenant en charge les débits Ultra160 ou supérieurs. Une carte HBA SCSI répondant à ce critère permet à chaque lecteur de bande de communiquer avec l'hôte à une vitesse de 160 Mo/s sur le bus SCSI. Le débit du bus SCSI étant supérieur à celui des lecteurs de bande, ce bus peut prendre en charge plusieurs périphériques sans pour autant diminuer leurs performances, du moins dans une certaine mesure.

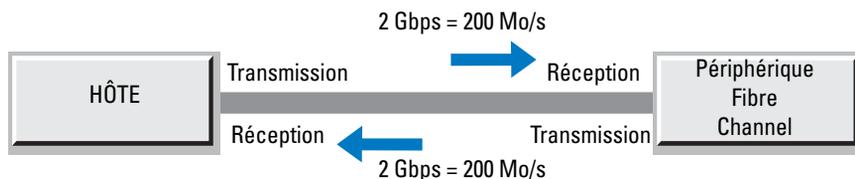
La vitesse maximale d'un bus Ultra160 pour le transfert de données vers les périphériques connectés est de 160 Mo/s. Un seul lecteur de bande ne peut pas utiliser la totalité de la bande passante du bus, car son débit natif pour l'écriture et la lecture des données sur la bande est de 80 Mo/s. Cependant, plusieurs lecteurs utilisés ensemble peuvent consommer la totalité de la bande passante du bus s'ils fonctionnent chacun au maximum de leurs capacités natives. Chaque lecteur supplémentaire connecté au même bus ultérieurement réduira les performances moyennes de tous les lecteurs.

Par conséquent, pour optimiser les performances d'une bibliothèque de bandes, il est recommandé de ne pas connecter plus de deux lecteurs à chaque bus SCSI. Voir la section "Configurations de câblage recommandées" pour plus de détails. Utilisez une carte HBA SCSI prenant en charge au minimum la norme Ultra160. La mise à niveau vers une carte HBA Ultra320 n'améliorera pas les performances si les spécifications du lecteur de bande correspondent à la norme Ultra160.

Configurations SAN

La technologie Fibre Channel (FC) présente de nombreux avantages par rapport à l'interface SCSI. Elle est plus performante en termes de distance (12 mètres pour les périphériques SCSI à différentiel basse tension contre 300 mètres pour une liaison FC à onde courte de 2 Gbps) et offre des vitesses de transfert des données plus élevées. Contrairement à l'interface SCSI, la technologie Fibre Channel ne fait pas appel à une architecture basée sur les bus. Elle utilise un protocole réseau série qui fait désormais figure de norme dans les réseaux SAN et permet la consolidation des ressources de données de stockage. De plus, chaque connexion FC est composée d'un lien de transmission et d'un lien de réception, ce qui permet l'utilisation du mode duplex intégral. Les données peuvent être transmises simultanément dans les deux directions. Lors d'une sauvegarde effectuée sur une seule connexion FC, les données peuvent être lues à partir d'une source et écrites sur la bande sans que les communications doivent être alternées. La bande passante de la connexion est donc doublée. Voir la figure 1-5.

Figure 1-5. Diagramme de liaison Fibre Channel



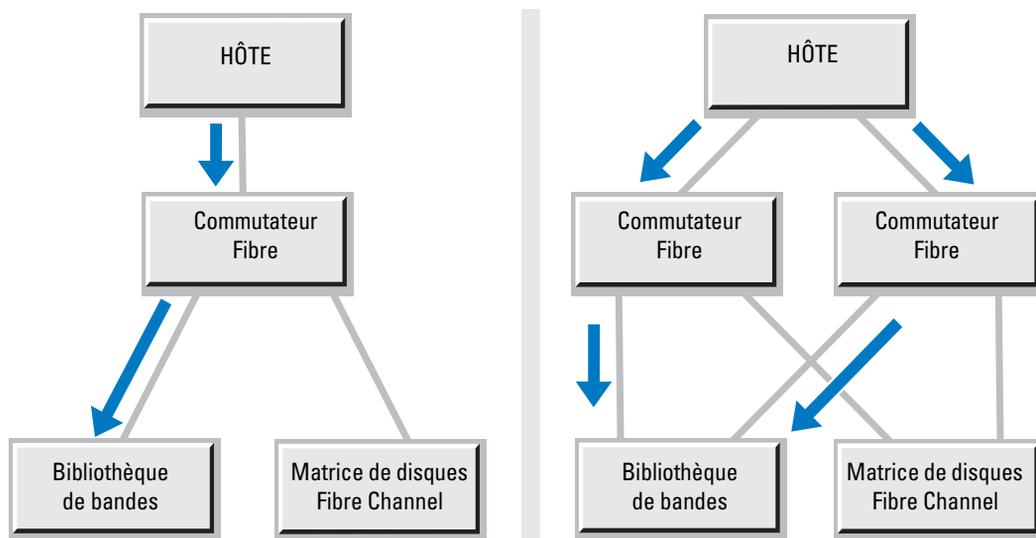
Les performances des bibliothèques de bande intégrées à un SAN sont influencées par divers facteurs : vitesses du lien FC, flux des données entre la source et la bibliothèque de bandes, restrictions liées aux baies de stockage externes en termes de performances. Ces inconvénients peuvent être évités grâce à une bonne compréhension de la configuration et de la gestion générale de la solution.

Bien que l'utilisation du protocole FC dans un SAN permette de bénéficier de débits supérieurs, certains points concernant les lecteurs de bande doivent être pris en compte afin d'éviter que les liaisons FC n'affectent les performances. Un débit de 2 Gbps sur une liaison FC correspond à 200 Mo/s (c'est-à-dire 200 Mo/s en envoi et 200 Mo/s en réception). L'utilisation de plusieurs lecteurs de bande sur un même lien risque d'aboutir à une saturation de la bande passante potentielle de ce dernier. Si l'hôte fonctionne avec un adaptateur de 1 Gbps de génération antérieure, la sauvegarde des données sur deux lecteurs peut suffire à entraîner une réduction importante des performances.

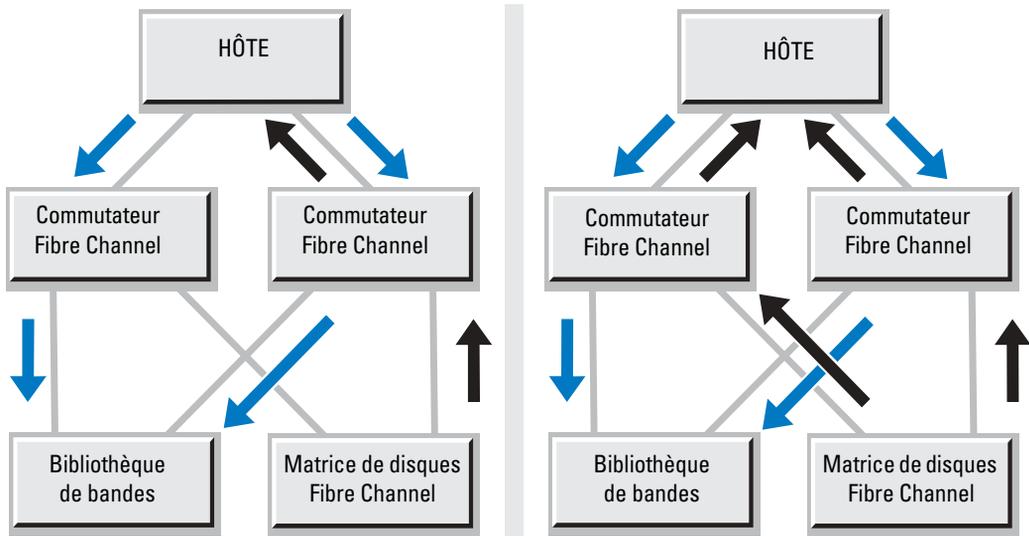
Par conséquent, si vous utilisez trois lecteurs (ou plus) simultanément sur une liaison à 2 Gbps, vous devrez peut-être répartir les sauvegardes sur plusieurs connexions plutôt que d'utiliser un seul lien. C'est pourquoi il est important de bien comprendre les notions de topologie d'une solution SAN. Le fait de suivre le chemin parcouru par les données lors des sauvegardes, depuis leur lecture sur le périphérique source jusqu'à leur écriture sur bande, permet aux administrateurs d'identifier tout risque de saturation potentiel et de prendre les mesures appropriées.

Par exemple, si la solution de sauvegarde nécessite l'utilisation simultanée de plusieurs lecteurs, le fait de répartir les périphériques de sauvegarde sur différentes structures peut permettre de diviser les connexions, et par conséquent d'améliorer les performances. Voir la figure 1-6.

Figure 1-6. Utilisation d'un flux de données unique ou divisé vers une bibliothèque de bandes



Si un ralentissement du débit survient lors de la lecture des données à partir d'une matrice de disques FC externe, vous pouvez utiliser un logiciel de gestion des E-S tel qu'EMC® PowerPath® avec une connexion vers une structure supplémentaire pour équilibrer automatiquement la charge sur plusieurs liens et tirer parti des fonctions de basculement pour améliorer la disponibilité des données. La partie gauche de la figure 1-7 représente une matrice de disques SAN dont les données sont envoyées via un seul lien, ce qui entraîne une saturation et un ralentissement du débit vers le matériel de sauvegarde sur bande. La partie droite montre comment l'équilibrage de charge permet de transférer les données sur deux liens et de doubler la bande passante des E-S.

Figure 1-7. Ralentissement du débit des données/Matrice d'équilibrage de charge

Les matrices de disques FC du SAN peuvent également être soumises aux restrictions de performances décrites à la section “Configuration des disques durs et des matrices RAID”. L'amélioration des performances des matrices de disques a donc un effet direct sur la vitesse de sauvegarde sur le SAN.

Configurations de SAN utilisant un pont Fibre Channel vers la bibliothèque

Certaines bibliothèques de bandes sont connectées à un SAN via un module de pont Fibre Channel. Ce dernier relie les protocoles SCSI et FC et offre des fonctions supplémentaires de gestion, de sécurité et d'exploitation qui ne sont pas disponibles dans la plupart des bibliothèques FC natives. Pour plus de détails sur ces fonctionnalités, consultez le guide d'utilisation du pont Fibre Channel de votre bibliothèque de bandes.

Dans certaines configurations de bibliothèques, ce module peut provoquer un ralentissement du débit et réduire les performances des lecteurs de bande. En effet, la capacité de traitement du module, nécessaire pour les communications SCSI-FC, est inférieure au débit agrégé de certaines configurations à plusieurs lecteurs. Cependant, pour la plupart des solutions de sauvegarde, ce module ne constitue pas un facteur majeur de réduction des performances. Il est fréquent que sur les serveurs dédiés à la sauvegarde, les restrictions au niveau de l'hôte soient compensées par le fait que les données sont envoyées vers plusieurs lecteurs de bande. Les performances moyennes du lecteur sont alors suffisamment basses pour que le module de pont Fibre Channel ne pose pas de problème de débit.

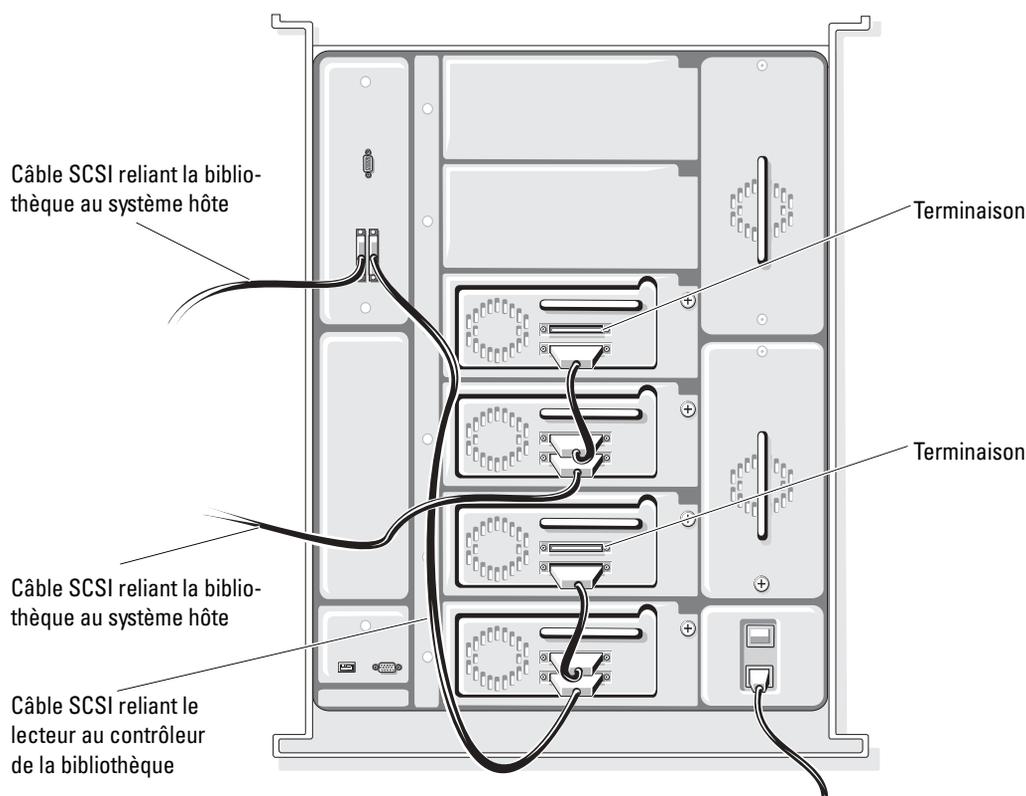
Configurations de câblage recommandées

Câblage SCSI vers le système hôte

Bibliothèque comportant un maximum de six lecteurs de bande

Lorsque la bibliothèque de bandes est connectée à l'hôte via l'interface SCSI, chaque bus doit être connecté à un maximum de deux lecteurs. Des contrôleurs SCSI supplémentaires sont nécessaires pour les bibliothèques de cinq ou six lecteurs de bande, et ce afin d'éviter que la saturation des bus SCSI n'entraîne une dégradation des débits de transfert.

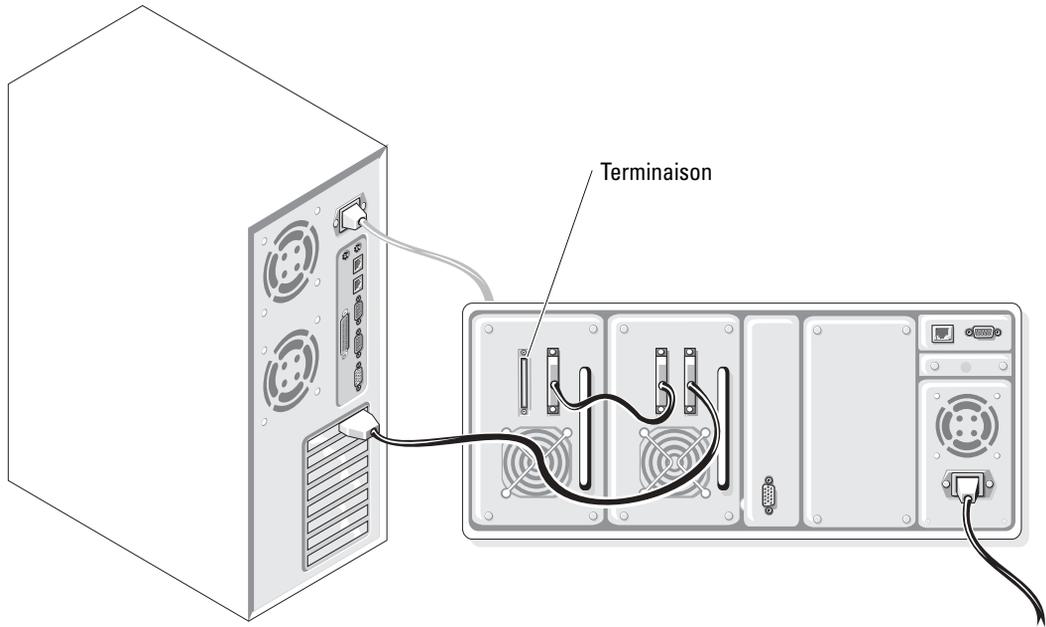
Figure 1-8. Câblage SCSI pour les bibliothèques comportant un maximum de six lecteurs de bande



Bibliothèque comportant un maximum de deux lecteurs de bande

Dans une bibliothèque de bande contenant deux lecteurs, ces derniers peuvent être reliés à un système hôte sur le même bus SCSI sans risque de baisse des performances de sauvegarde. La vitesse de sauvegarde est alors identique à celle de deux lecteurs connectés à des bus différents. Cependant, si vous activez la fonction de vérification des applications de sauvegarde, vous voudrez peut-être améliorer les performances de vérification en répartissant les deux lecteurs sur deux bus SCSI. Cette opération augmente les performances de vérification de 25 %.

Figure 1-9. Câblage SCSI pour une bibliothèque comportant un maximum de deux lecteurs de bande



Câblage de lecteurs vers le pont Fibre Channel

Les figures 1-10 à 1-17 indiquent comment configurer une bibliothèque comportant un maximum de six lecteurs avec un module de pont Fibre Channel pour optimiser les performances FC.

Figure 1-10. Câblage reliant un pont Fibre Channel à un lecteur de bande

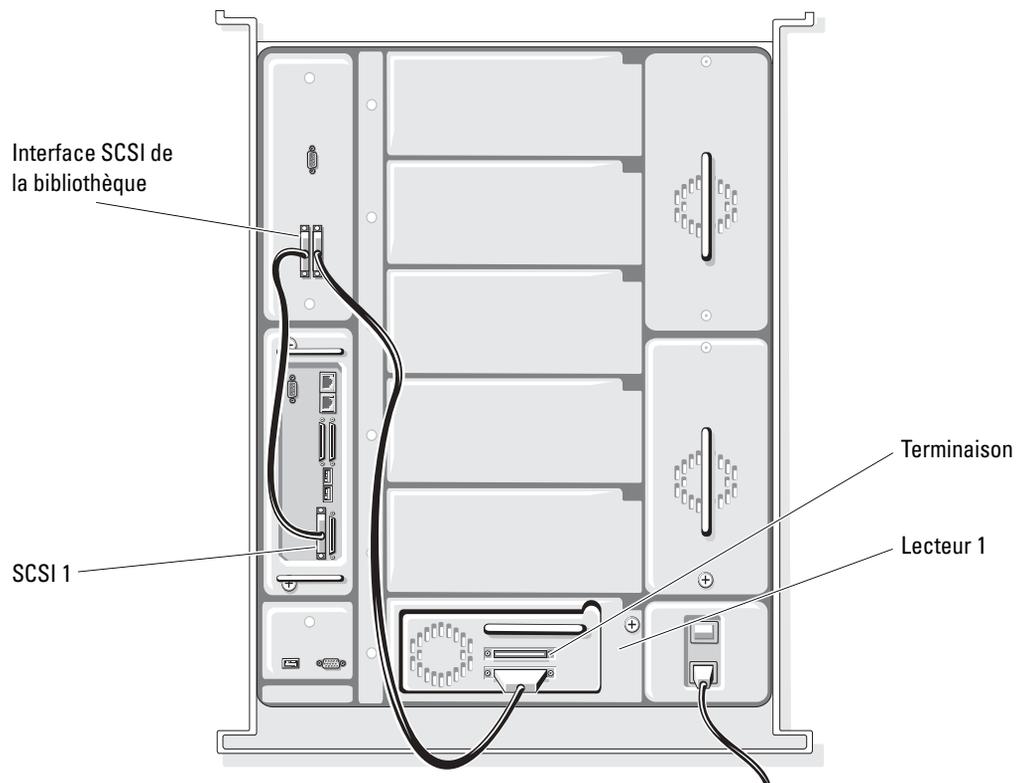


Figure 1-11. Câblage reliant un pont Fibre Channel à deux lecteurs de bande

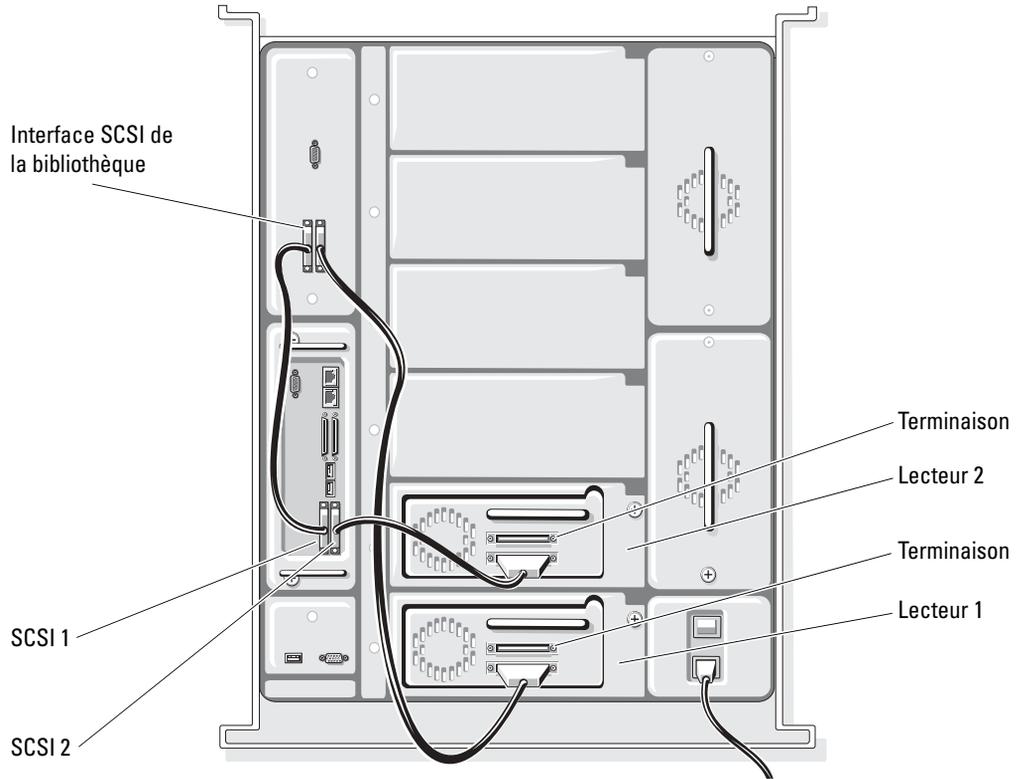


Figure 1-12. Câblage reliant un pont Fibre Channel à trois lecteurs de bande

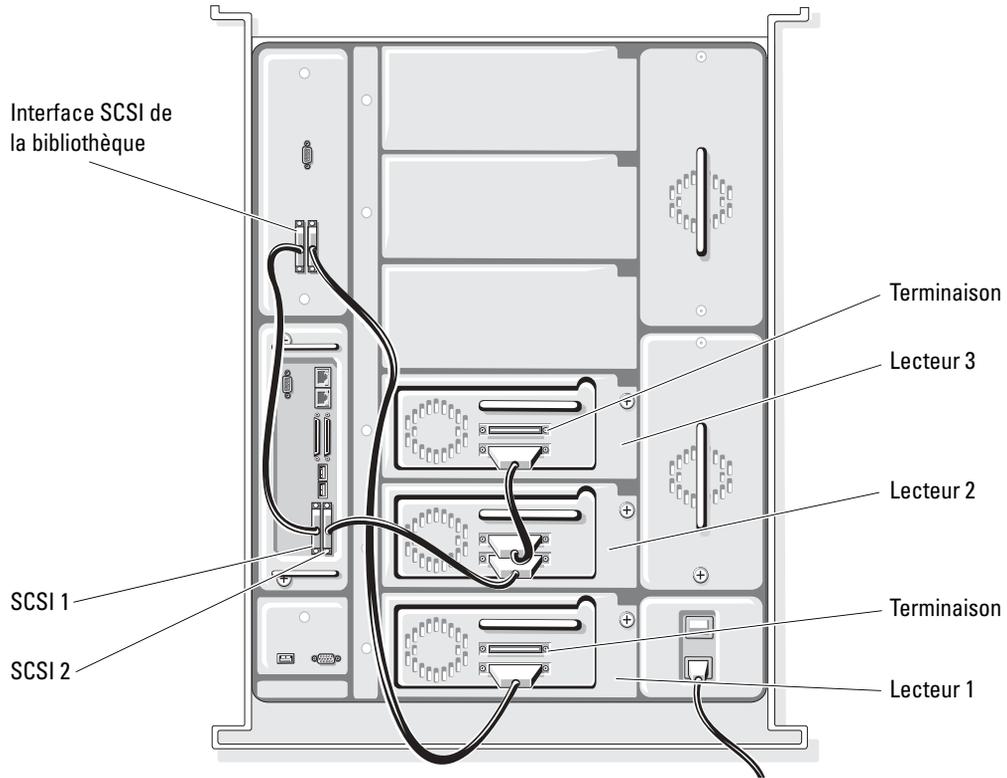


Figure 1-13. Câblage reliant un pont Fibre Channel à quatre lecteurs de bande

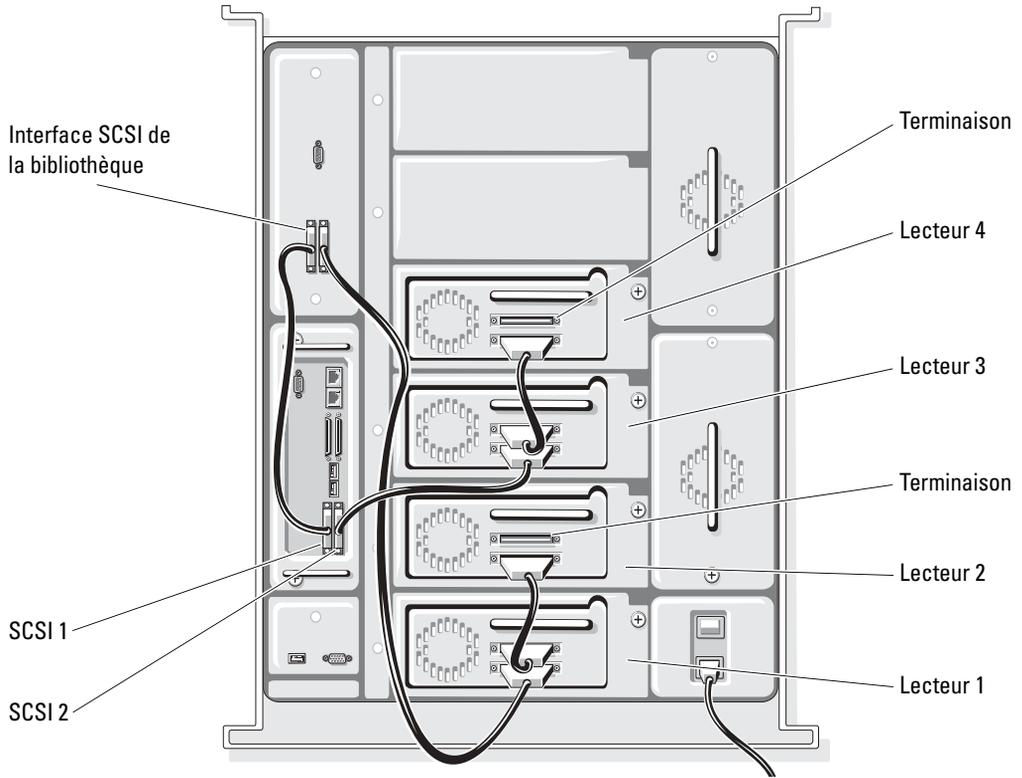


Figure 1-14. Paramètres de division en zones du canal pour une bibliothèque de bandes de un à quatre lecteurs

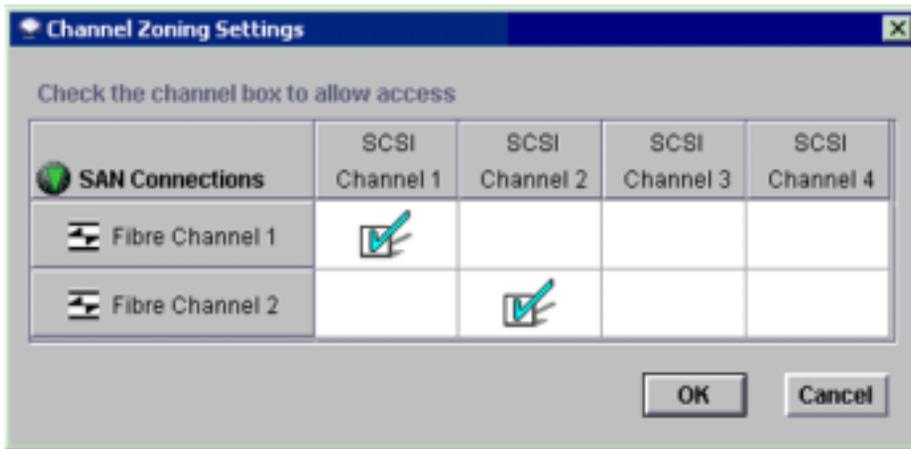


Figure 1-15. Câblage reliant un pont Fibre Channel à cinq lecteurs de bande

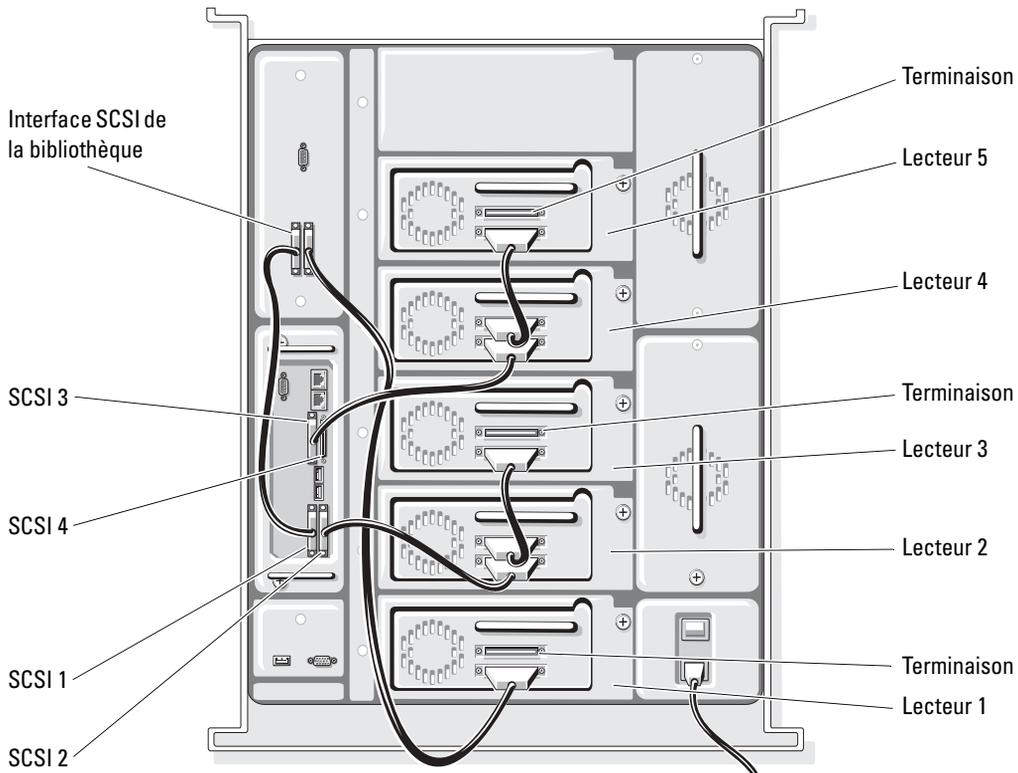


Figure 1-16. Câblage reliant un pont Fibre Channel à six lecteurs de bande

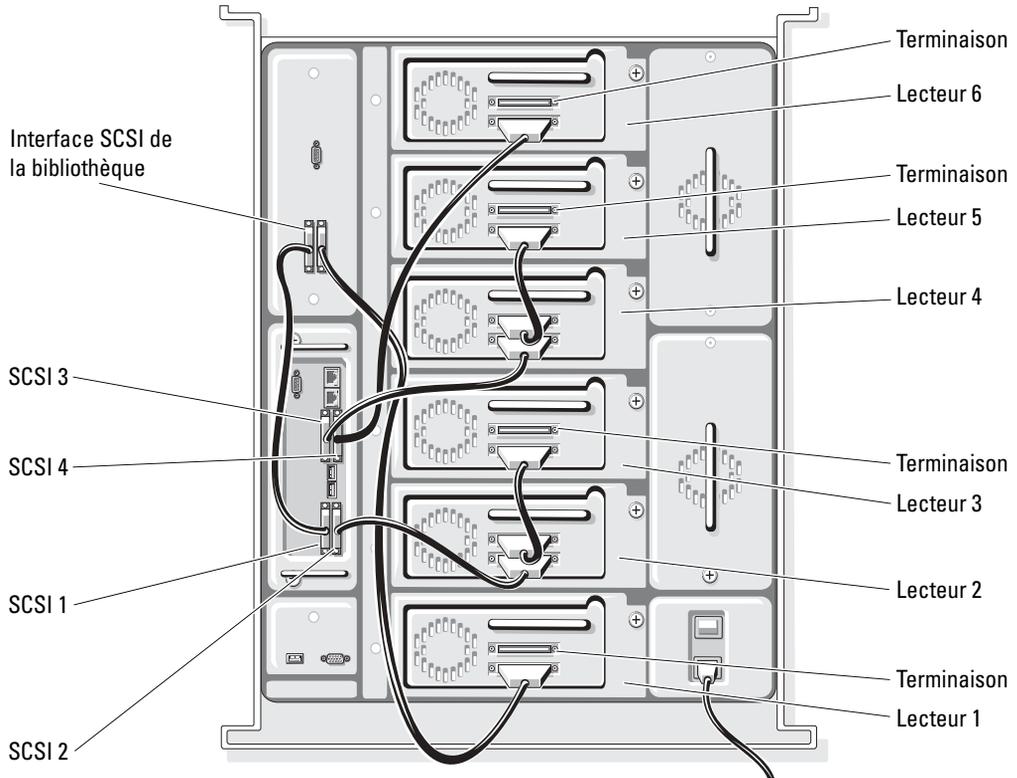
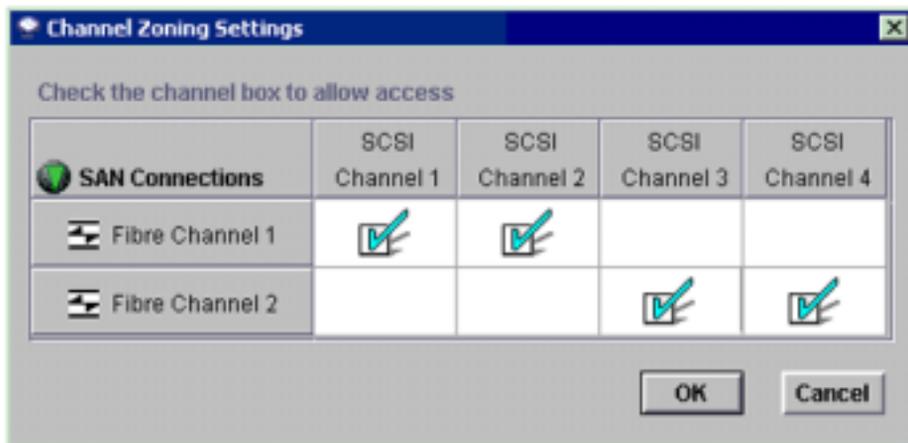


Figure 1-17. Paramètres de division en zones du canal pour une bibliothèque de bandes de cinq ou six lecteurs



Les figures 1-18 et 1-19 indiquent comment configurer une bibliothèque comportant un maximum de deux lecteurs de bande avec un module de pont Fibre Channel pour optimiser les performances FC.

Figure 1-18. Bibliothèque de bandes comportant un lecteur

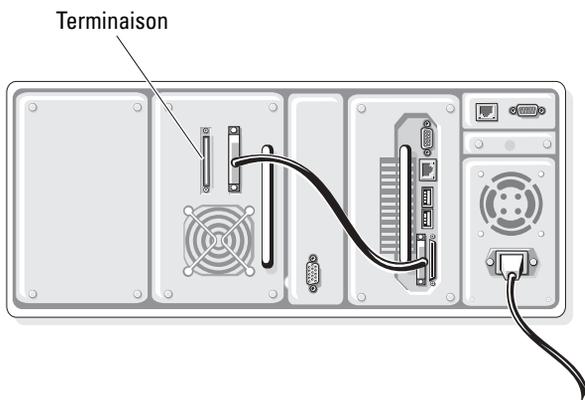
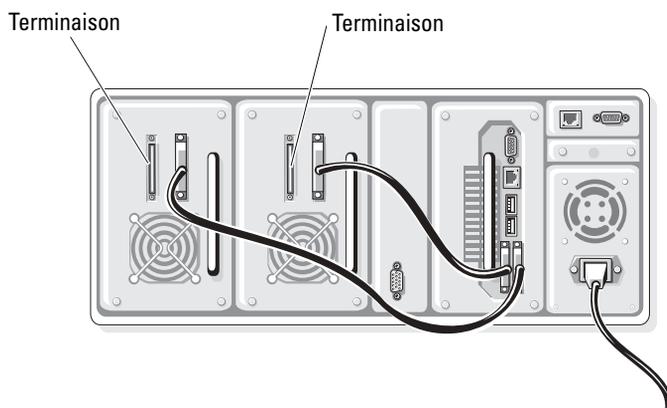


Figure 1-19. Bibliothèque de bandes comportant deux lecteurs



Dell™ PowerVault™-Systeme

**Hinweise zur
Leistungsoptimierung
bei Bandlaufwerken
und Bandbibliotheken**

Irrtümer und technische Änderungen vorbehalten.
© 2005 Dell Inc. Alle Rechte vorbehalten.

Die Reproduktion dieses Dokuments in jeglicher Form ohne schriftliche Genehmigung von Dell Inc. ist streng verboten.

Marken in diesem Text: *Dell*, das *DELL* Logo und *PowerVault* sind Marken von Dell Inc.; *EMC* und *PowerPath* sind eingetragene Marken von EMC Corporation.

Alle anderen in dieser Dokumentation genannten Marken und Handelsbezeichnungen sind Eigentum der jeweiligen Hersteller und Firmen. Dell Inc. erhebt keinen Anspruch auf Marken und Handelsbezeichnungen mit Ausnahme der eigenen.

Juni 2005

Inhalt

Einführung	79
Allgemeine Datensicherungshinweise für Hostserver	79
Hinweise zu Bandlaufwerken und Daten.	79
Konfiguration von Festplatten und RAID-Arrays	81
Allgemeine leistungsbezogene Hinweise zum Einsatz mehrerer Laufwerke in Bandbibliotheken	87
SCSI-Konfigurationen.	87
SAN-Konfigurationen.	88

Einführung

Angesichts der jüngsten Verbesserungen bei den Übertragungsraten für Bandlaufwerke müssen viele Faktoren auf der Hostseite – wie RAID-Konfigurationen und Festplattenspezifikationen – berücksichtigt werden, um zu ermitteln, ob Hostserver und Bandlaufwerk Daten mit der selben Übertragungsrates verarbeiten können. Im Abschnitt „Allgemeine Datensicherungshinweise für Hostserver“ werden allgemeine Konfigurationen und Attribute erörtert, die den Durchsatz vom Hostserver zum Bandlaufwerk beeinträchtigen können.

Um dem potenziellen Durchsatz mehrerer Bandlaufwerke, wie sie in Bandbibliotheken enthalten sind, gerecht zu werden, sind auf der Hostseite größere Bandbreiten erforderlich. Der Abschnitt „Allgemeine Hinweise zur Performance beim Einsatz mehrerer Laufwerke in Bandbibliotheken“ erörtert potenzielle Einschränkungen durch faseroptische Kabel bei Einheiten mit mehreren Laufwerken und gibt Empfehlungen für die Kabelkonfiguration.

Allgemeine Datensicherungshinweise für Hostserver

Die Hinweise in diesem Abschnitt gelten sowohl für SCSI- als auch für SAN (Storage Area Network = Speichernetzwerk)-Bandsicherungskonfigurationen

Hinweise zu Bandlaufwerken und Daten

Bei der Leistungsbeurteilung sollten folgende Punkte berücksichtigt werden:

Verarbeitungszeit für SCSI-Befehle. Der Zeitaufwand für die Befehlsverarbeitung am SCSI-Bus führt bei *allen* SCSI-Geräten zu einer Abweichung von der theoretisch erreichbaren maximalen Datenübertragungsrates. Dieser Aufwand wird nicht durch die Bandsicherungssoftware verursacht. Vielmehr misst die Software lediglich die Geschwindigkeit, mit der die Daten aufs Band geschrieben werden. So ist es beispielsweise möglich, dass das Laufwerk 80 MB Daten pro Sekunde verarbeitet, aber nur 77 MB pro Sekunde auf Band schreibt. Der letztere Wert wird von der Datensicherungssoftware angezeigt.

Blockgrößen für Bandlaufwerke. Für die meisten Bandlaufwerke ist eine Blockgröße von 64 Kb optimal. Bei Datensicherungsanwendungen kann die Blockgröße durch den Benutzer geändert werden. Das Erhöhen der Blockgröße führt allerdings nicht zu einer Leistungsverbesserung. Die Verwendung von Blockgrößen von unter 64 Kb kann sogar zu Leistungseinbußen führen. Informationen zur Anpassung der Blockgröße für Ihr Bandlaufwerk finden Sie in der Benutzeranleitung Ihrer Datensicherungssoftware.

Puffergröße der Datensicherungssoftware. Um bei der Datensicherung eine optimale Leistung zu erzielen, sollten die Puffer der Datensicherungssoftware so groß wie möglich sein. Bei einigen Anwendungen kann die Puffergröße durch den Benutzer geändert werden. Dies kann dazu beitragen, einen kontinuierlichen Datenfluss vom und zum Laufwerk zu gewährleisten und die Übertragungsraten – vor allem bei kleinen Dateien – beträchtlich zu erhöhen. Je größer der Puffer ist, desto mehr Daten kann er aufnehmen. Dies verringert den Zeitaufwand für das Suchen der Daten auf dem Laufwerk, kann jedoch andererseits die Speicher- und Prozessorleistung beeinträchtigen. Einzelheiten finden Sie in der Benutzeranleitung der Datensicherungssoftware für Ihr Bandlaufwerk.

Treiber und Firmware. Achten Sie darauf, dass für den SCSI- oder Fibre-Controller und das Bandlaufwerk immer die aktuellsten Treiber- und Firmwareversionen installiert sind. Unter support.dell.com können Sie die jeweils aktuellen Treiber- und Firmwareversionen für Ihr Dell PowerVault-Bandlaufwerk herunterladen.

Schließen Sie Bandlaufwerke und Festplatten an getrennten (internen oder externen) Controllern an, die sich an getrennten Host-Bus-Adaptern (HBAs) befinden. Auch wenn dies bis zu einem gewissen Grad von der Leistungskapazität Ihres Controllers abhängt, wird allgemein empfohlen, den HBA für das Bandlaufwerk vom HBA für das Festplattenlaufwerk zu trennen, um maximalen Durchsatz zu gewährleisten. Die meisten für den Dualbetrieb ausgelegten Onboard-SCSI/RAID-Controller nutzen gemeinsam einen Prozessor, der seine Bandbreite auf RAID-Array und Bandlaufwerk aufteilen muss. Dies bedeutet, dass ein einziger Controller die Lese- und Schreibvorgänge zwischen den Festplatten und dem Bandlaufwerk abwickelt und zugleich alle erforderlichen Paritätsinformationen berechnet und auf die Festplatten schreibt. Spezifische Informationen zu RAID-Arrays und Paritätsbytes finden Sie unter „Hinweise zur Hardware-RAID-Konfiguration“.

Verschmutzte Laufwerkköpfe oder alte Medien. Ein verschmutzter Laufwerkkopf oder ein altes Speichermedium kann hohe Fehlerraten und einen entsprechenden Rückgang der Lese-/Schreibgeschwindigkeiten verursachen. Immer, wenn ein Laufwerk versucht, einen Block erneut auf Band zu schreiben bzw. vom Band zu lesen, führt dies zu einer Beeinträchtigung der Leistung. Sobald eine gewisse Anzahl an Lese-/Schreibfehlern erreicht ist, fordert das Laufwerk in der Regel eine Reinigung an. Es ist sehr wichtig, die Laufwerkköpfe in regelmäßigen Abständen bzw. auf Anforderung zu reinigen.

Je älter oder abgenutzter das verwendete Medium ist, desto höher ist die Wahrscheinlichkeit, dass Datenblöcke beschädigt werden. Ein LTO-Bandmedium ist in der Regel für einen Produktzyklus von ca. 75 kompletten Lese-/Schreibvorgängen ausgelegt.

Geschwindigkeitsanpassung. Neuere LTO-Laufwerke passen sich an die Geschwindigkeit an, mit der die Daten an das Laufwerk übertragen werden. Der Anpassungsbereich reicht bis zur Hälfte der maximalen Datenübertragungsgeschwindigkeit ohne Komprimierung. Wenn die Daten noch langsamer an das Laufwerk übertragen werden, muss es anhalten, warten, dass sich der Puffer wieder auffüllt, zurückschleusen und versuchen, den Pufferinhalt zu schreiben (das so genannte „Back Hitching“).

So passt sich etwa das Bandlaufwerk Dell PowerVault 110T (LTO2 und LTO3) beim Beschreiben von LTO-3-Medien an Geschwindigkeiten bis hinunter zu 30 MB/s an. Wenn der Hostserver die Daten nur mit 20 MB/s bereitstellen kann, kommt es zu „Back Hitching“, da das Bandlaufwerk wartet, dass die Datenpuffer wieder gefüllt werden. In dieser Situation liegt der effektive Durchsatz unter 20 MB/s (eher bei 15 MB/s).

Bestätigung der Leistung Ihres Bandlaufwerks

Einige Hersteller von Bandlaufwerken bieten für ihre Produkte einen integrierten Leistungsdiagnosemodus an, der zur Bestätigung des Datendurchsatzes verwendet werden kann. Die PowerVault-Bandlaufwerke 110T LTO-2 und LTO-3 (ab Firmwareversion 53XX) bieten einen Diagnosemodus „F“ an, bei dem ein schneller Lese-/Schreibtest auf Laufwerk und Medium ausgeführt wird. Wenn die gemessene Leistung nicht bei mindestens 6 Prozent der für das Laufwerk spezifizierten Maximalgeschwindigkeit liegt, schlägt der Test fehl, und es wird eine Fehlermeldung angezeigt. Wenn der Test erfolgreich verläuft, wird keine Meldung angezeigt. Einzelheiten zum Diagnosemodus „F“ finden Sie im Benutzerhandbuch Ihres Bandlaufwerks.



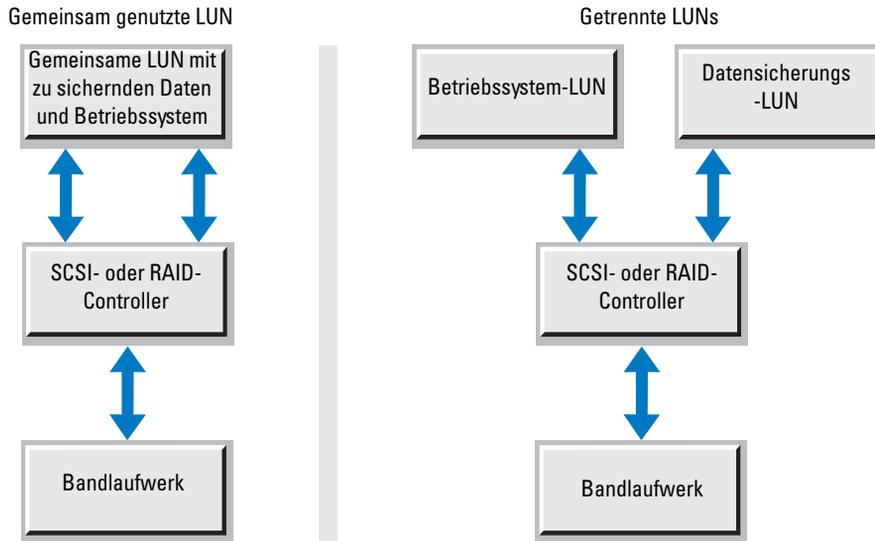
HINWEIS: Für den Diagnosemodus „F“ müssen Medien verwendet werden, die im Rahmen des Diagnosetests ohne Gefahr überschrieben werden können. Verwenden Sie hierzu keinesfalls Medien, die wichtige Daten enthalten. Alle Daten auf dem für den Diagnosetest verwendeten Medium gehen verloren.

Konfiguration von Festplatten und RAID-Arrays

Die Übertragungsgeschwindigkeit bei der Sicherung oder Wiederherstellung von Daten kann durch eine Reihe (interner und externer) Array-Attribute beeinträchtigt werden. In den nachstehenden Teilabschnitten werden diese Attribute erörtert und Konfigurationsempfehlungen gegeben, die die maximale Geschwindigkeit bei der Datensicherung und -übertragung gewährleisten. Wenn der Durchsatz des Bandlaufwerks dauerhaft über dem des Disk-Arrays liegt, kann die Maximalleistung des Bandlaufwerks nicht erreicht werden.

Allgemeine Hinweise zur Festplattenkonfiguration

Daten/Betriebssystem bei unterschiedlichen LUNs. Damit das Festplattenlaufwerk die Zugriffs- und Bearbeitungszeiten nicht zwischen Betriebssystemoperationen und Datensicherungsoperationen aufteilen muss, kann für die Datensicherung eine andere logische Einheitennummer (LUN = Logical Unit Number) als für das Betriebssystem verwendet werden. Hierzu ist es erforderlich, dass sich das Betriebssystem und die zu sichernden Daten auf physisch getrennten Festplatten bzw. Disk-Arrays befinden.

Abbildung 1-1. Single-Channel- vs. Two-Channel-Bandbreite

Festplattenleistung

Bandlaufwerke schreiben Daten konstruktionsbedingt sequentiell. Sie sind auf konstante Datenzufuhr angewiesen, um den sequentiellen Laufwerkbetrieb aufrechtzuerhalten (d.h. Back Hitching zu vermeiden). Festplatten sind dagegen Geräte mit wahlfreiem Datenzugriff (Random Access). Aus diesem Grund tun sich Festplatten zuweilen schwer damit, Daten für Bandlaufwerke in sequentieller Folge bereitzustellen, sofern diese Daten stark fragmentiert und über die gesamte Festplatte verteilt sind. Das Laufwerk muss in diesem Fall ständig winzige Datenblöcke zusammensuchen.

Darüber hinaus können auch weitere Merkmale der Festplatte den Durchsatz der Datenübertragung zum Bandlaufwerk beeinträchtigen.

Spindelgeschwindigkeit. Die Spindelgeschwindigkeit des Festplattenlaufwerks wird in der Regel in U/min (Umdrehungen pro Minute) gemessen. Dieser Wert bezieht sich auf die Anzahl der vollständigen Umdrehungen, die die Platte innerhalb einer Minute vollführt. Er wirkt sich direkt auf die Zugriffszeiten für die wahlfreie und die sequentielle Datenübertragungsrate aus. Je höher die Spindelgeschwindigkeit, desto schneller kann das Laufwerk auf die Daten zugreifen.

Direktzugriffszeit bzw. Suchzeit. Die Suchzeit wird in der Regel in Millisekunden gemessen. Bei ihr handelt es sich um die Zeit, die ein Laufwerkkopf benötigt, um ein Datenfragment auf der Festplatte zu finden. Die Suchzeit einer Festplatte ist der Zeitraum, den ihre Lese-/Schreibköpfe benötigen, um sich zwischen den Spuren auf der Plattenoberfläche hin- und herzubewegen. Da Festplattenlaufwerke Geräte mit wahlfreiem Zugriff sind, können die Daten auf praktisch jedem beliebigen Festplattensektor abgelegt werden. Je mehr Zeit der Datenzugriff in Anspruch nimmt, desto langsamer ist der Durchsatz des Festplattenlaufwerks insgesamt.

Dieses Merkmal ist von entscheidender Bedeutung, wenn sich auf einem Festplattenlaufwerk viele kleine Dateien befinden. Je kleiner die Dateien sind, desto mehr „Suchvorgänge“ muss das Laufwerk durchführen, um sie auf die Platte zu schreiben oder von ihr zu lesen. Deshalb ist das Lesen und Schreiben von Festplatten in der Regel sehr langsam, wenn viele kleine Dateien übertragen werden müssen.

Sequentielle/anhaltende Übertragungsraten (STR = Sequential/sustained transfer rates)

Die STR misst die tatsächliche Geschwindigkeit, mit der ein Laufwerk Daten auf die Festplatte schreibt und von ihr liest. Wenn sich die gesicherten Daten in einer großen, zusammenhängenden Datei befinden, nähert sich der anhaltende Durchsatz der maximalen STR des Laufwerks. Im Alltag werden die Daten jedoch durch Löschen und Wiederbeschreiben zunehmend fragmentiert und über die gesamte Platte verteilt. Das Defragmentieren der Festplatte kann dazu beitragen, dass ein Laufwerk seine maximale STR erreicht.

Puffer (Cache). Als Puffer wird der Speicherbereich des Laufwerks bezeichnet, der die zuletzt geschriebenen bzw. gespeicherten Daten enthält. Je größer der Puffer ist, desto mehr Daten kann er enthalten. Dies verkürzt die Zeit, die zum Suchen der Daten auf der Festplatte benötigt wird.

Hinweise zur Hardware-RAID-Konfiguration

RAID – Allgemeine Übersicht

Dieser Abschnitt gibt Überblick über typische RAID-Konfigurationen und ihre Auswirkungen auf die Übertragungsraten bei der Datensicherung und –wiederherstellung. Ein RAID-Array ist ein Verbund mehrerer Festplatten, die als einheitliches Speichersystem bzw. LUN fungieren. Die Datenübertragung erfolgt potenziell simultan über die Kanäle aller Festplatten. Damit erreicht der Gesamtdurchsatz ein Mehrfaches der Gesamtanzahl der Festplatten im Array – abzüglich der Verarbeitungszeit und der in den nachfolgenden Abschnitten beschriebenen Redundanzen.

Bei einer RAID-Konfiguration ist die Schnittstellengeschwindigkeit entscheidend, da die Bandbreite der Schnittstelle von allen Laufwerken gemeinsam genutzt wird. Ein Beispiel: Angenommen, die Übertragungsraten eines einzigen Ultra160-Laufwerks beträgt nur 40 MB/s. Ein RAID-0-Array aus fünf Festplatten dieses Laufwerktyps sollte entsprechend eine Schreib-/Leseleistung von 200 MB/s erreichen. Die Ultra160-Schnittstelle begrenzt die Maximalgeschwindigkeit des Arrays jedoch auf 160 MB/s.

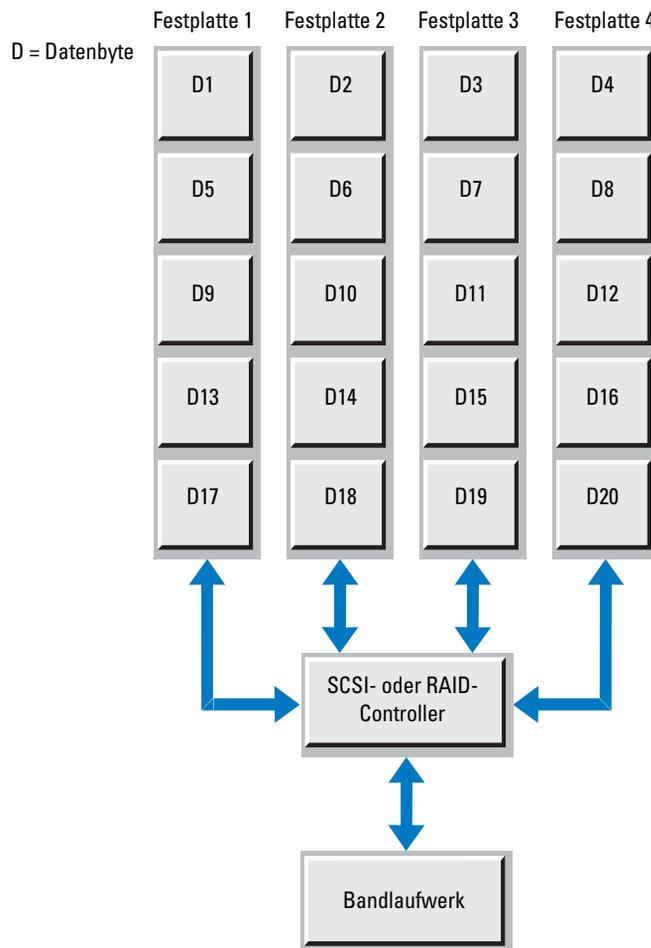
Externe Disk-Arrays, besonders in SANs, bieten unter Umständen umfassenden Cachespeicher, um die E/A-Leistung zu verbessern. Dies erhöht die Schreibgeschwindigkeit des Arrays erheblich und ermöglicht das Speichern häufig abgerufener Daten zwecks Steigerung der Leseleistung. In Bezug auf die Bandlaufwerkleistung kompensiert der Cachespeicher die meisten RAID-bedingten Beschränkungen beim Wiederherstellen von Daten auf dem Array oder beim Sichern der Daten vom Array. Bei Datensicherungen von externen Arrays sind jedoch unter Umständen trotz Cachespeicher Beschränkungen aufgrund der RAID-Konfiguration spürbar, da die Daten von den Festplatten gelesen werden müssen.

RAID 0

RAID 0 ist auch unter der Bezeichnung „Striping“ bekannt. Dabei werden zwei oder mehr Festplatten zu einem virtuellen Laufwerk mit einer einzigen logischen Gerätenummer (LUN) zusammengefasst). Dabei wird zu einem Zeitpunkt nicht jeweils nur auf eine Festplatte geschrieben, sondern die Daten werden auf alle Festplatten im Array verteilt. Damit wird der Datendurchsatz auf n Kanäle verteilt (n = die Anzahl der Festplattenlaufwerke im Array), anstatt auf einen einzigen Kanal bzw. eine einzige Festplatte beschränkt zu bleiben. Dies führt zu einer herausragenden Lese-/Schreibgeschwindigkeit, bietet jedoch keinerlei Fehlertoleranz.

Abbildung 1-2 zeigt vier Festplattenlaufwerke in einer RAID-0-Konfiguration. Die Daten werden auf alle vier Laufwerke verteilt, so dass das Array über vier Lese-/Schreibkanäle verfügt.

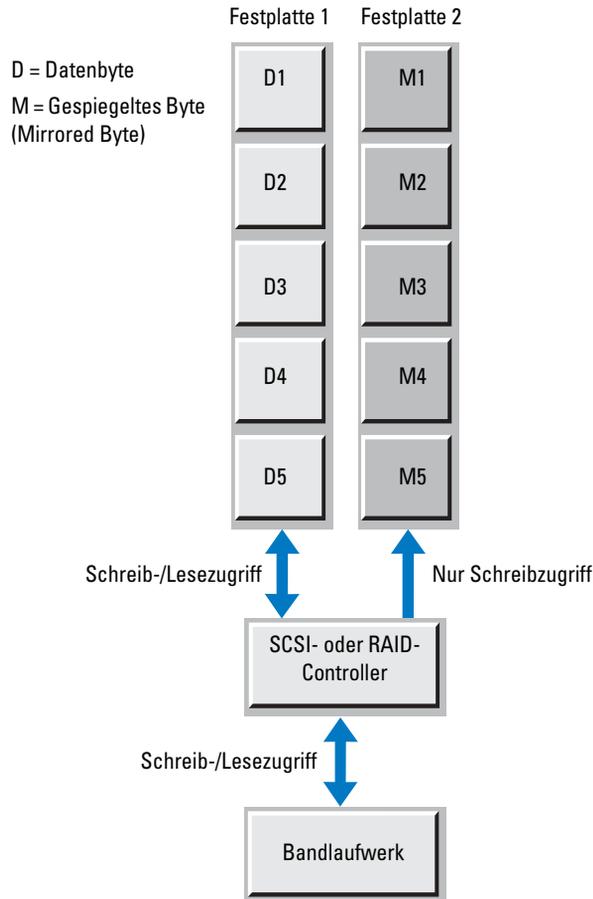
Abbildung 1-2. Beispiel für eine RAID-0-Konfiguration



RAID 1

RAID 1 wird auch als „Mirroring“ bezeichnet. Bei dieser Konfiguration werden dieselben Daten simultan auf zwei Laufwerke geschrieben. Beim Ausfall eines Laufwerks wechselt das System zum Zweitlaufwerk, ohne dass Daten verloren gehen. Bei Datensicherungen mit Bandlaufwerken entspricht die Lesegeschwindigkeit von einem RAID-1-Array etwa der einer einzelnen Festplatte, da die Daten vom Primärlaufwerk gelesen werden. Die Schreibgeschwindigkeit bei der Datenwiederherstellung vom Bandlaufwerk ist bei einem RAID-1-Array jedoch unter Umständen langsamer. Dies ist durch die Fehlerüberprüfung und –korrektur (ECC) beim Schreiben auf das primäre und das gespiegelte Laufwerk bedingt. Der Effizienzverlust ist zu einem Großteil darauf zurückzuführen, dass die Spiegelung häufig über den Prozessor oder den RAID-Controller erfolgt. Neuere RAID-Controller sind daher aufgrund der neueren und leistungsfähigeren Prozessoren in der Regel schneller.

Abbildung 1-3. Beispiel für eine RAID1-Konfiguration

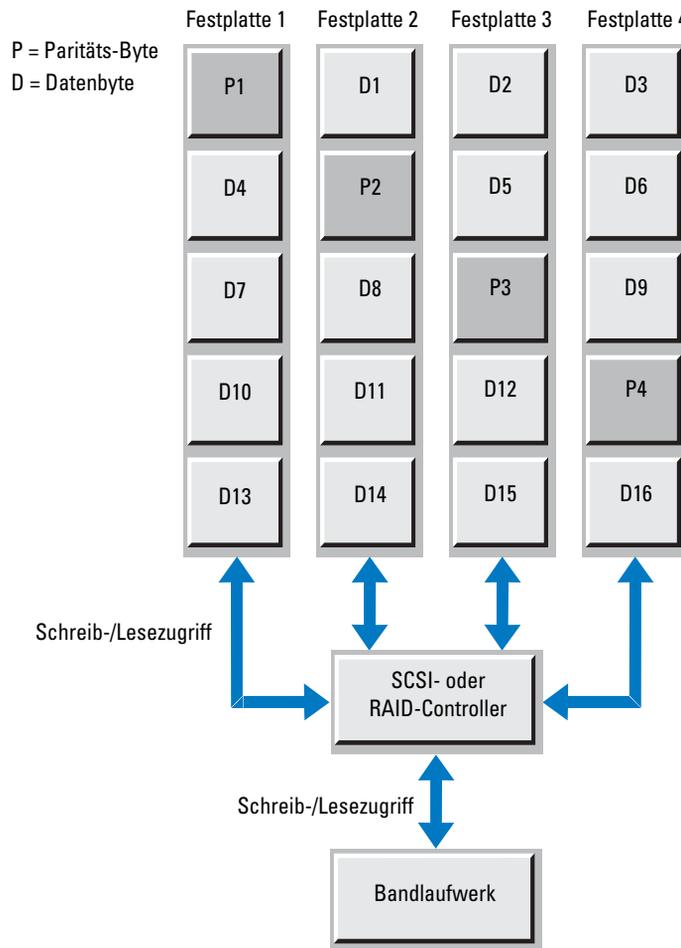


RAID 5

Bei einem RAID-5-Array werden Daten im Striping-Verfahren auf Byte-Ebene auf alle Festplatten im Array verteilt. Dies gilt auch für die Fehlerkorrektur- oder Paritätsdaten. RAID-5-Arrays weisen in der Regel eine sehr hohe Geschwindigkeit beim direkten Lesezugriff auf, die mit der Anzahl der Festplatten im Array zunimmt. Bei größeren Arrays kann die Lesegeschwindigkeit die von RAID-0-Arrays übertreffen, da für die Datenverteilung ein weiteres Festplattenlaufwerk zur Verfügung steht. Darüber hinaus werden die Paritätsdaten bei normalen Lesevorgängen nicht benötigt.

Datenwiederherstellungen vom Bandlaufwerk zu einem RAID-5-Array erfolgen aufgrund des zusätzlichen Aufwands für das Berechnen und Schreiben der Paritätsdaten in der Regel zur Nenngeschwindigkeit.

Abbildung 1-4. Beispiel für eine RAID5-Konfiguration



Insgesamt ist RAID 0 tendenziell die beste Konfiguration für maximale Lese- und Schreibgeschwindigkeit, bietet jedoch keine Datenredundanz. RAID 1 weist generell die niedrigsten Übertragungsraten auf, weil alle auf das Array geschriebenen Daten gespiegelt und die Daten von einer einzigen Festplatte gelesen werden. RAID 5 erreicht gute Werte bei der Lesegeschwindigkeit, aber nur durchschnittliche Werte bei der Schreibgeschwindigkeit. Die Leistung einer RAID-5-Konfiguration steigt mit der Anzahl der Festplatten im Array. Wenn das RAID-Array in einem Gehäuse mit umfassender Cachespeicherung untergebracht ist, werden die Leistungseinschränkungen während der Datenwiederherstellung unter Umständen abgemildert. Die Datensicherung unterliegt auch in diesem Fall den Einschränkungen der jeweiligen RAID-Konfiguration. Die Merkmale eines Arrays hängen darüber hinaus zu einem Großteil von den Merkmalen der darin verwendeten Festplatten ab (siehe Aufstellung unter „Festplattenleistung“).

Allgemeine leistungsbezogene Hinweise zum Einsatz mehrerer Laufwerke in Bandbibliotheken

Beim simultanen Einsatz mehrerer Bandlaufwerke zur Datensicherung – wie in Bandbibliotheken üblich – sind weitere Aspekte der Hardwarekonfiguration zu berücksichtigen. Durch einfache Maßnahmen zur Leistungsoptimierung bei Hardwareeinrichtung und Kabelkonfiguration kann zusätzlichen Engpässe beim Datendurchsatz vorgebeugt werden.

SCSI-Konfigurationen

Die neuesten Hochleistungs-Bandlaufwerke für Bandbibliotheken unterstützen die Ultra160-Spezifikation des SCSI-Standards. Um maximale Leistung zu erzielen, müssen Datensicherungs-Server, die SCSI verwenden, daher über einen HBA verfügen, der Datengeschwindigkeiten nach dem Ultra160-Standard oder höher unterstützt. Ein SCSI-HBA, der diese Anforderungen erfüllt, ermöglicht jedem Bandlaufwerk, mit einer Übertragungsrate von 160 MB/s über den SCSI-Bus mit dem Host zu kommunizieren. Da der SCSI-Bus im Vergleich zum Bandlaufwerk eine höhere Datenrate erreicht, können mehrere Laufwerke an denselben Bus angeschlossen werden, ohne dass dies zu Leistungseinbußen führt. Dies gilt jedoch nur bis zu einem gewissen Punkt.

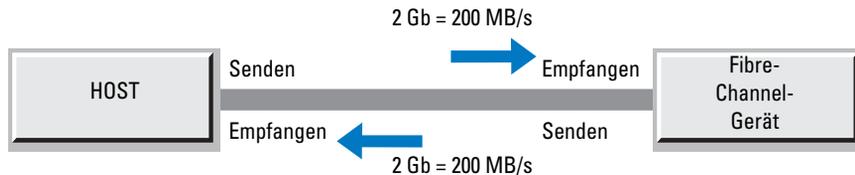
Die Datenrate des Ultra160-Busses – 160 MB/s – ist zugleich die maximale Gesamt-Durchsatzrate aller an den Bus angeschlossenen Geräte. Ein einziges Bandlaufwerk, das die Daten mit bis zu 80 MB/s (nativ) auf Band schreibt bzw. vom Band liest, kann nicht die gesamte Bus-Bandbreite ausnutzen. Durch Kombination mehrerer Bandlaufwerke, die mit maximaler nativer Geschwindigkeit betrieben werden, können jedoch die gesamten vom Bus angebotenen 160 MB/s ausgeschöpft werden. Ist dieser Punkt erreicht, so bewirkt jedes darüber hinaus zusätzlich angeschlossene Laufwerk eine Beeinträchtigung der durchschnittlichen Leistung pro Laufwerk.

Um sicherzustellen, dass eine Bandbibliothek mit maximaler Leistung arbeitet, sollten deshalb nicht mehr als zwei Bandlaufwerke an einen SCSI-Bus angeschlossen werden. Spezifische Hinweise und Illustrationen finden Sie im Abschnitt „Empfehlungen für die Kabelkonfiguration“. Der verwendete SCSI-HBA sollte mindestens den Standard Ultra160 unterstützen. Beachten Sie, dass der Einsatz eines Ultra320-HBAs nicht zu einer Leistungssteigerung führt, wenn das Bandlaufwerk für Ultra 160 spezifiziert ist.

SAN-Konfigurationen

Fibre-Channel (FC) bietet gegenüber SCSI zahlreiche Vorteile. Zum einen ist es nicht an die für SCSI geltenden Entfernungsbegrenzungen gebunden (LVD-SCSI: 12 m; 2-Gb-FC-Link: 300 m) und ermöglicht höhere Datenübertragungsgeschwindigkeiten. Zum anderen ist FC keine busbasierte Architektur wie SCSI, sondern ein serielles Netzwerkprotokoll, das auch für die Implementierung von Speichernetzwerken (SANs) bevorzugt wird. Dies ermöglicht die Konsolidierung von Datenspeicherressourcen. Darüber hinaus besteht jede FC-Verbindung aus einer Sende- und einer Empfangsverbindung. Dies ermöglicht den vollwertigen Duplexbetrieb, also die simultane Datenübertragung in beide Richtungen. Bei der Datensicherung über eine einzige FC-Verbindung können also Daten vom Quelllaufwerk ohne Unterbrechung durch Kommunikation in die Gegenrichtung gelesen und auf Band geschrieben werden. Dadurch wird die Bandbreite der Verbindung faktisch verdoppelt. Siehe Abbildung 1-5.

Abbildung 1-5. Fibre-Channel-Verbindungsdiagramm



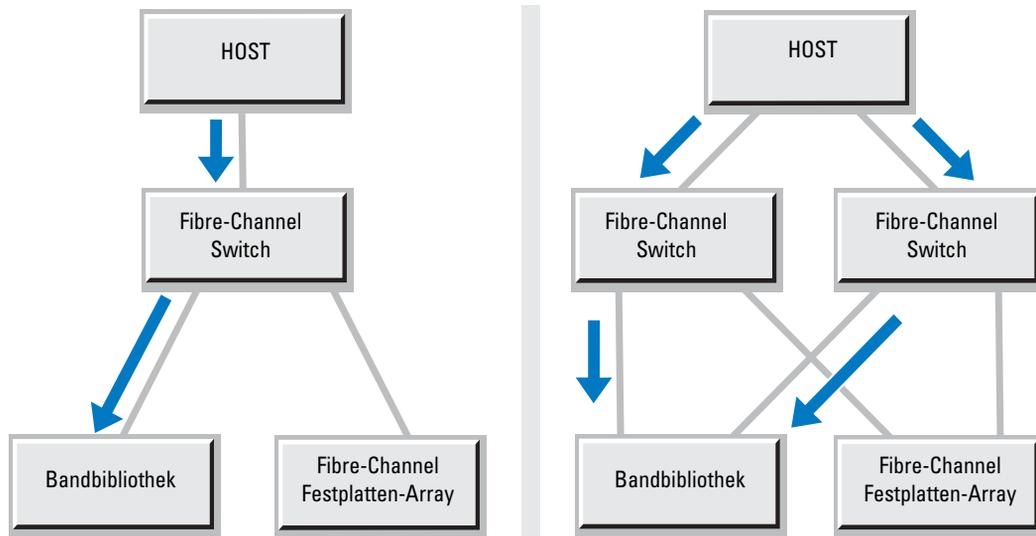
Bei der Einrichtung von Bandbibliotheken in einem Speichernetzwerk (SAN) kann die Leistung durch eine Reihe weiterer Faktoren beeinflusst werden. Hierzu zählen die Geschwindigkeiten der FC-Verbindungen, der Datenfluss zwischen Quelllaufwerk und Bandbibliothek sowie die Leistungsbeschränkungen externer Speicher-Arrays. Mit entsprechendem Verständnis der Gesamteinrichtung und -verwaltung der Lösung lassen sich viele die Leistung beeinträchtigende Faktoren vermeiden.

Ungeachtet der hohen Datenbandbreite, die das FC-Protokoll in SANs bereit stellt, müssen für Bandlaufwerke entsprechende Vorkehrungen getroffen werden, um eine Situation zu vermeiden, in der die Leistung durch die FC-Verbindung beeinträchtigt wird. Die Datenrate einer 2-Gigabit-FC-Verbindung beträgt 200 MB/s (d.h. je 200 MB/s auf der Sende- und der Empfangsverbindung). Beim Versuch, mehrere Bandlaufwerke über dieselbe Verbindung zu betreiben, kann deshalb die gesamte Bandbreite der Verbindung potenziell überschritten werden. Wenn der Host mit einem älteren 1-GB-Adapter betrieben wird, kann bereits die Datensicherung auf zwei Laufwerken zu erheblichen Leistungseinbußen führen.

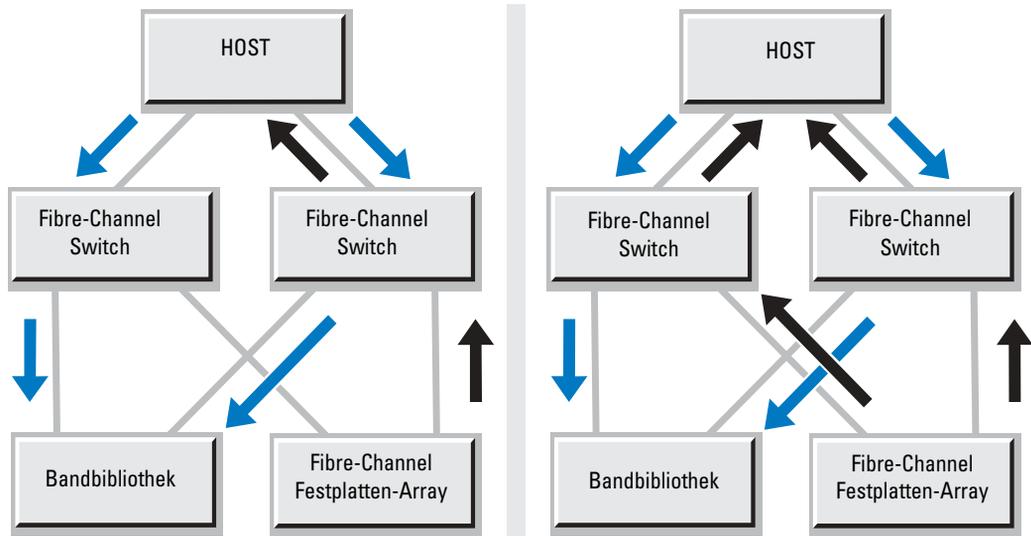
Wenn drei oder mehr Bandlaufwerke an einer 2-Gb-Verbindung betrieben werden, kann es deshalb erforderlich sein, die Datensicherung auf mehrere Verbindungen zu verteilen anstatt alles über eine einzige Verbindung laufen zu lassen. Hier ist ein angemessenes Verständnis der Topologie der SAN-Lösung hilfreich. Administratoren können potenzielle Engpässe erkennen, indem sie den Datenpfad während einer Datensicherung nachverfolgen – vom Lesen des Quelllaufwerks bis zum Schreiben auf das Band. Falls Engpässe identifiziert werden, können je nach Konfiguration entsprechende Maßnahmen ergriffen werden.

Wenn die Datensicherungslösung beispielsweise den gleichzeitigen Betrieb mehrerer Laufwerke erfordert, lässt sich die Leistung unter Umständen verbessern, indem die Bandlaufwerke auf unterschiedliche Strukturen verteilt und die Verbindungen geteilt werden. Siehe Abbildung 1-6.

Abbildung 1-6. Einheitlicher und aufgeteilter Datenfluss zur Bandbibliothek



Wenn beim Lesen der Daten von einem externen FC-Festplatten-Array ein Engpass auftritt, kann zudem durch den Einsatz einer E/A-Management-Software wie EMC® PowerPath® und die Verbindung zu einer weiteren Struktur ein automatischer Lastausgleich realisiert werden. Die Daten werden auf mehrere Pfade verteilt, und die Verfügbarkeit wird durch Pfad-Failover erhöht. Siehe Abbildung 1-7. Die linke Seite der Abbildung stellt ein SAN-Disk-Array dar, in dem alle Daten eine einzige Verbindung durchlaufen müssen. So entsteht ein Engpass, der den Datenfluss zum Bandlaufwerk verlangsamt. Die rechte Seite zeigt, wie die E/A-Bandbreite bei der Übertragung vom Array durch Lastausgleich verdoppelt wird, da die Daten über zwei Verbindungen übertragen werden.

Abbildung 1-7. Datenübertragungs-Engpass vs. Festplatten-Array mit Lastausgleich

Schließlich können FC-Festplatten-Arrays im SAN auch den im Abschnitt „Konfiguration von Festplatten und RAID-Arrays“ beschriebenen Leistungsbeeinträchtigungen unterliegen. Die Optimierung der Leistungsmerkmale der Festplatten-Arrays hat daher auch direkte Auswirkungen auf die Geschwindigkeit der Datensicherung über das SAN.

SAN-Konfigurationen unter Verwendung der Fibre-Channel-Brücke der Bandbibliothek

Bestimmte Bandbibliotheken lassen sich über ein Fibre-Channel-Brückenmodul an ein SAN anschließen. Das Modul dient als Überbrückung zwischen SCSI- und FC-Protokoll und bietet Zusatzfunktionen für Verwaltung, Sicherheit und Bedienung, die in den meisten nativen FC-Bibliotheken nicht verfügbar sind. Ausführliche Informationen zu diesen Funktionen finden Sie in der Benutzerdokumentation zur Fibre-Channel-Brücke für Ihre Bandbibliothek.

Bei einigen Bandbibliothekskonfigurationen kann die Fibre-Channel-Brücke Engpässe verursachen und die Leistung der Bandlaufwerke beeinträchtigen. Dies liegt daran, dass die auf die Überbrückung der SCSI/FC-Kommunikation ausgelegte Verarbeitungskapazität der Fibre-Channel-Brücke den gesammelten Datendurchsatz nicht bewältigt, den einige Mehrfachlaufwerk-Konfigurationen erzeugen. Gleichwohl wird das Fibre-Channel-Brückenmodul bei den meisten Datensicherungslösungen nicht als primärer Hemmfaktor für die Bandleistung in Erscheinung treten. Bei dedizierten Datensicherungsservern kommt es häufig dazu, dass die Einschränkungen am Host durch den Aufwand der Datenübermittlung an mehrere Bandlaufwerke verstärkt werden. Hierdurch sinkt die durchschnittliche Laufwerkleistung unter die Grenze, ab der das Fibre-Channel-Brückenmodul als Faktor in Erscheinung treten würde.

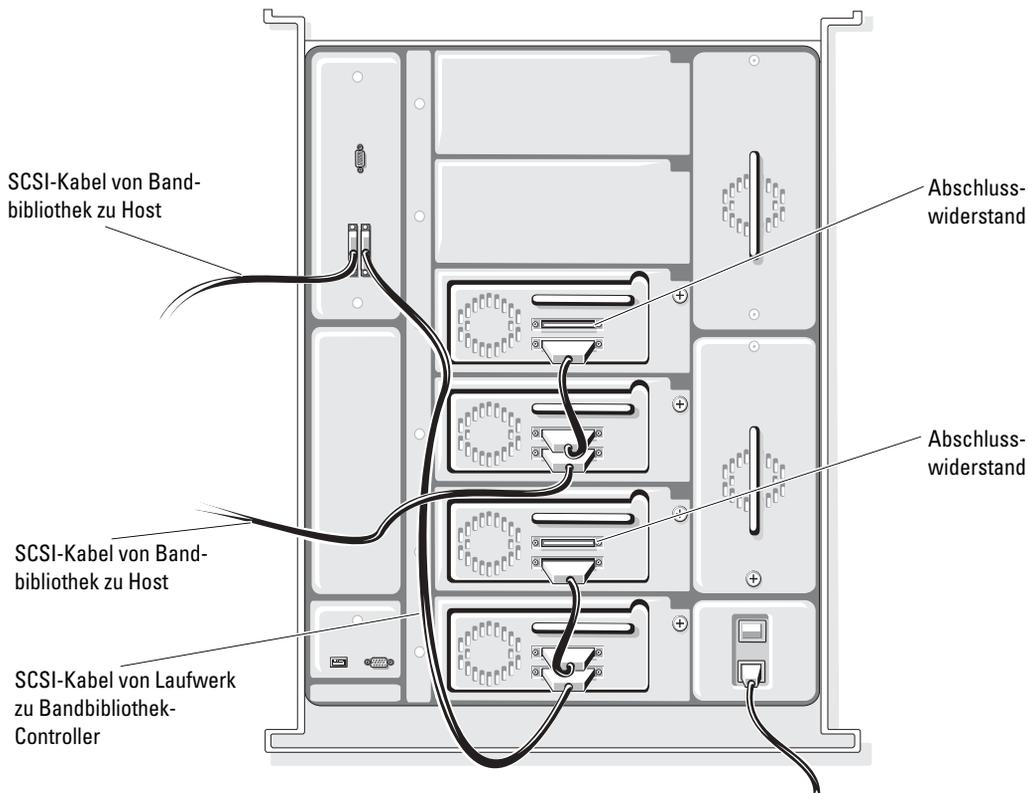
Empfehlungen für die Kabelkonfiguration

SCSI-Kabelverbindung zum Host

Bandbibliothek mit bis zu sechs Bandlaufwerken

Wenn die Bandbibliothek über SCSI an einen Host angeschlossen ist, sollten maximal je zwei Laufwerke an einem Bus betrieben werden. Für Bandbibliotheken mit fünf oder sechs Bandlaufwerken sind zusätzliche SCSI-Controller erforderlich, um sicherzustellen, dass der maximale Datendurchsatz nicht durch den SCSI-Bus eingeschränkt wird.

Abbildung 1-8. SCSI-Verkabelung für Bandbibliothek mit bis zu sechs Bandlaufwerken



Bandbibliothek mit bis zu zwei Bandlaufwerken

In einer vollständig konfigurierten Bandbibliothek mit zwei Laufwerken können diese an einen am selben SCSI-Bus betriebenen Host angeschlossen werden, ohne dass es zu signifikanten Leistungsbeeinträchtigungen bei der Datensicherung kommt. Die Datensicherungsraten für zwei Laufwerke an einem SCSI-Bus entsprechen den Datensicherungsraten zweier an getrennten Bussen betriebener Laufwerke. Kunden, die die Überprüfungsfunktion ihrer

Kabelkonfigurationen für Fibre-Channel-Brücke mit Bandlaufwerken

Die Abbildungen 1-10 bis 1-17 zeigen, wie eine Bandbibliothek mit bis zu sechs Laufwerken bei Einsatz eines Fibre-Channel-Brückenmoduls konfiguriert werden sollte, um die Bandleistung über FC zu optimieren.

Abbildung 1-10. Kabelkonfiguration für Fibre-Channel-Brücke mit einem Bandlaufwerk

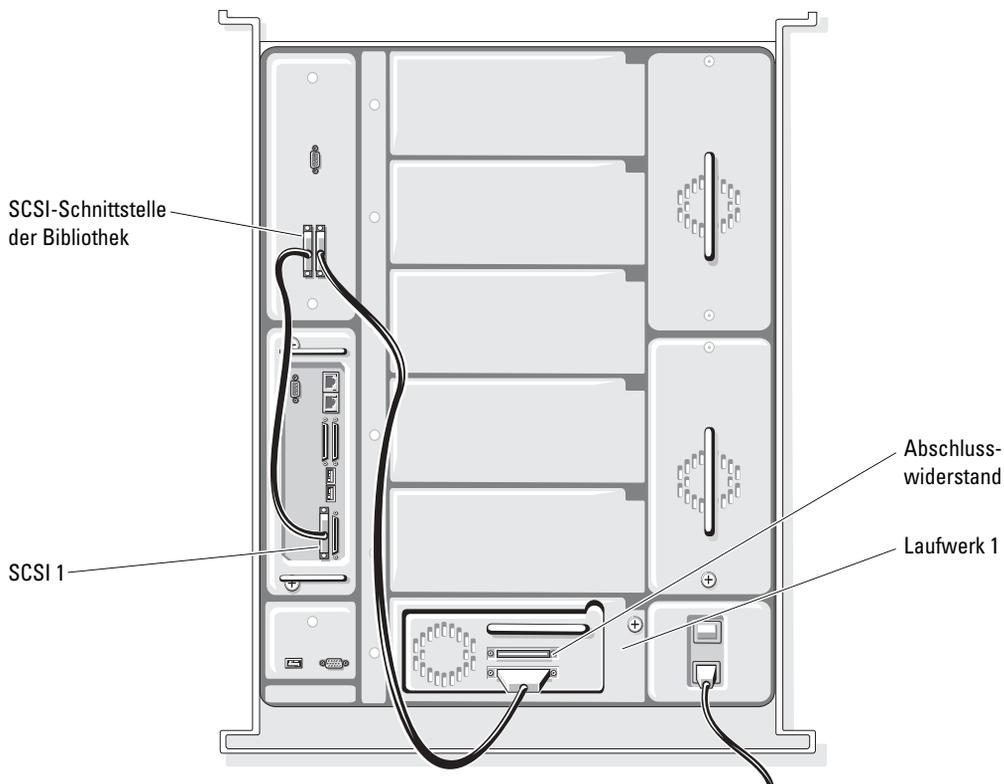


Abbildung 1-11. Kabelkonfiguration für Fibre-Channel-Bridge mit zwei Bandlaufwerken

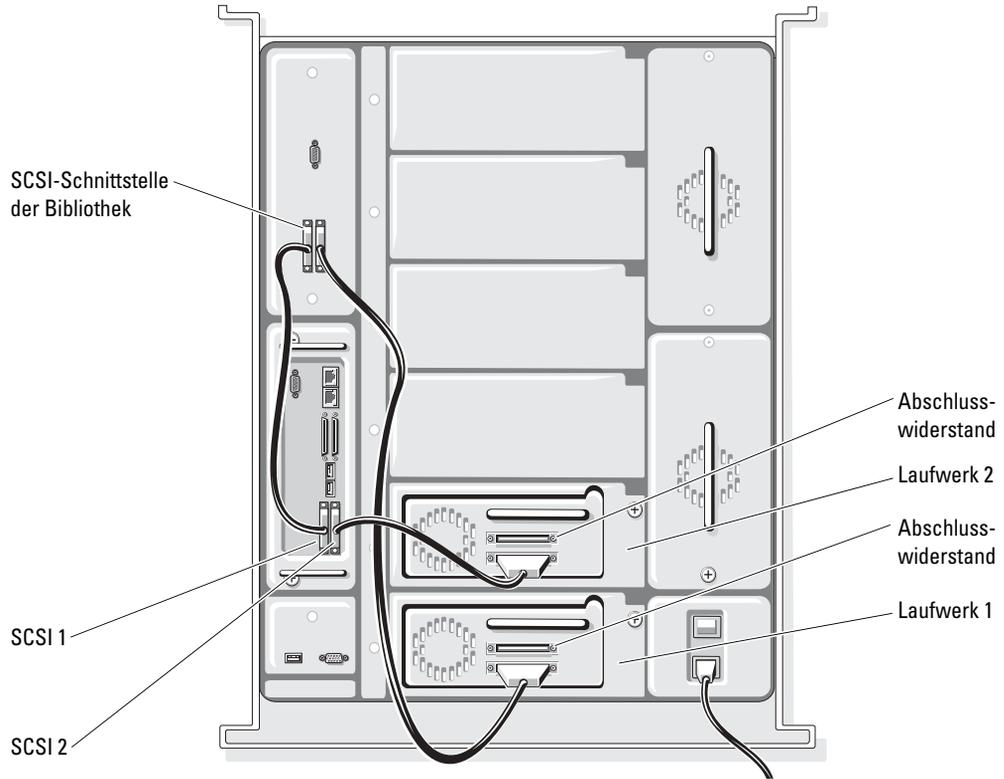


Abbildung 1-12. Kabelkonfiguration für Fibre-Channel-Bridge mit drei Bandlaufwerken

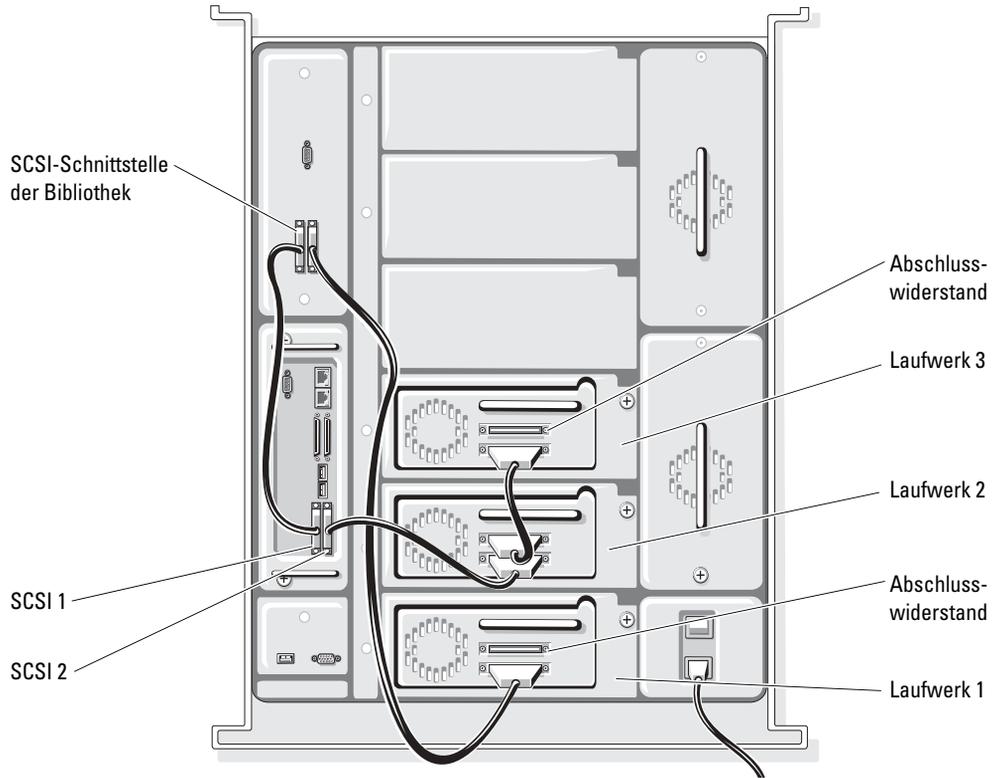


Abbildung 1-13. Kabelkonfiguration für Fibre-Channel-Bridge mit vier Bandlaufwerken

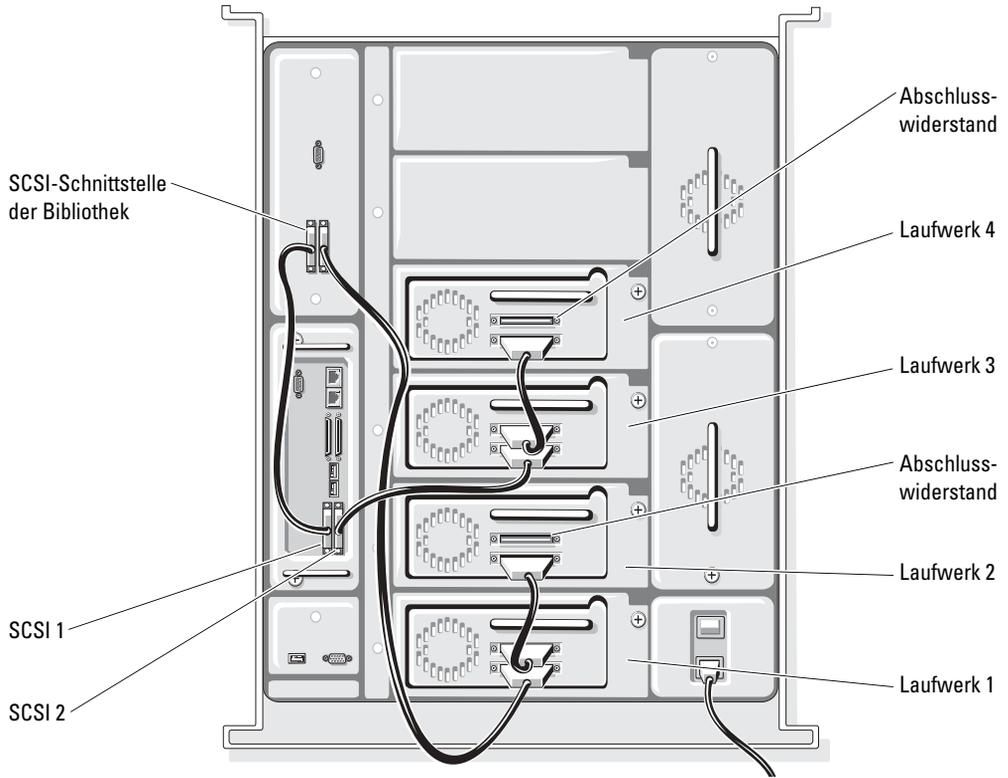


Abbildung 1-14. Kanalzoneinstellungen für Bandbibliothek mit einem bis vier Laufwerken

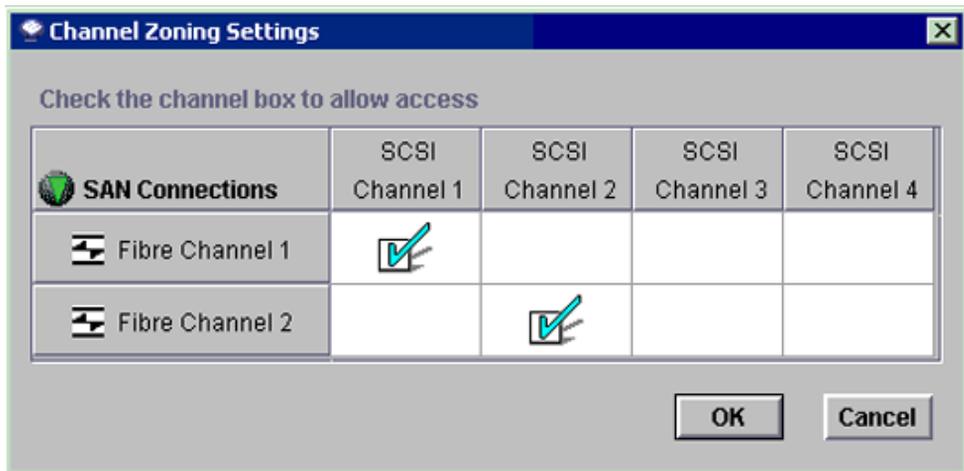


Abbildung 1-15. Kabelkonfiguration für Fibre-Channel-Bridge mit fünf Bandlaufwerken

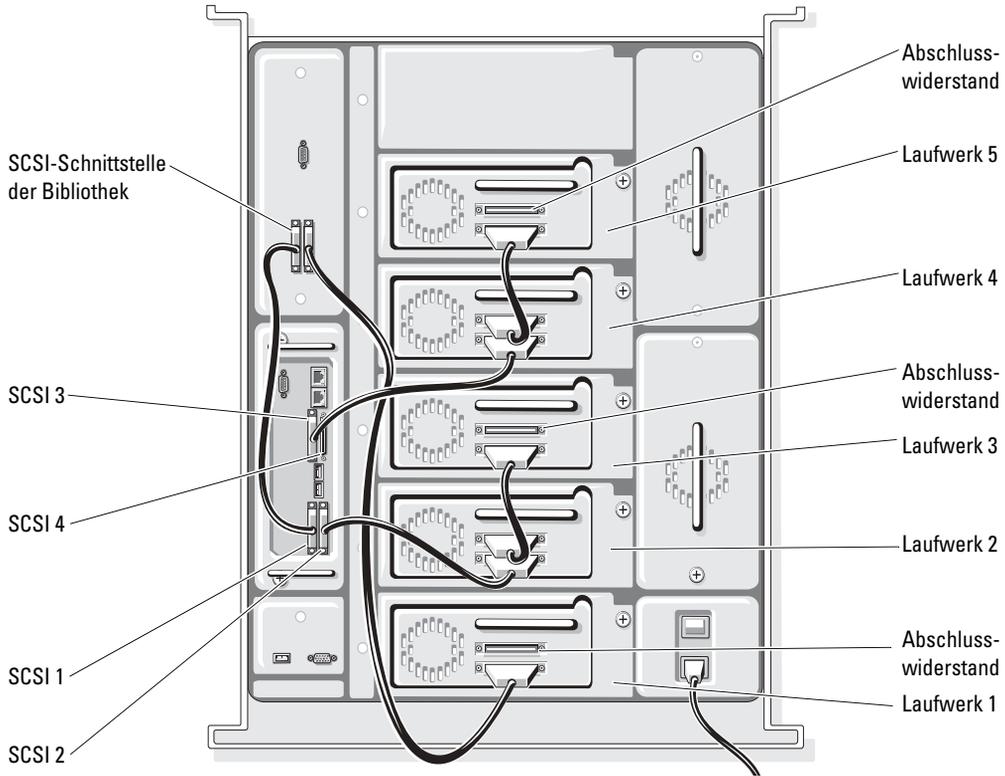


Abbildung 1-16. Kabelkonfiguration für Fibre-Channel-Bridge mit sechs Bandlaufwerken

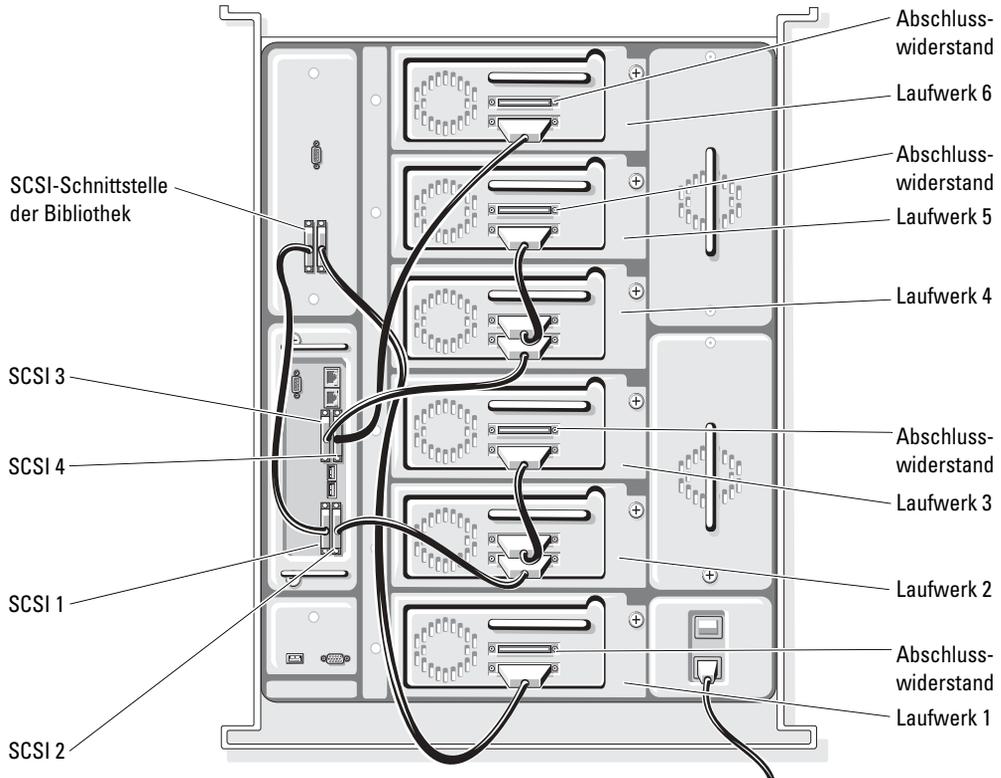
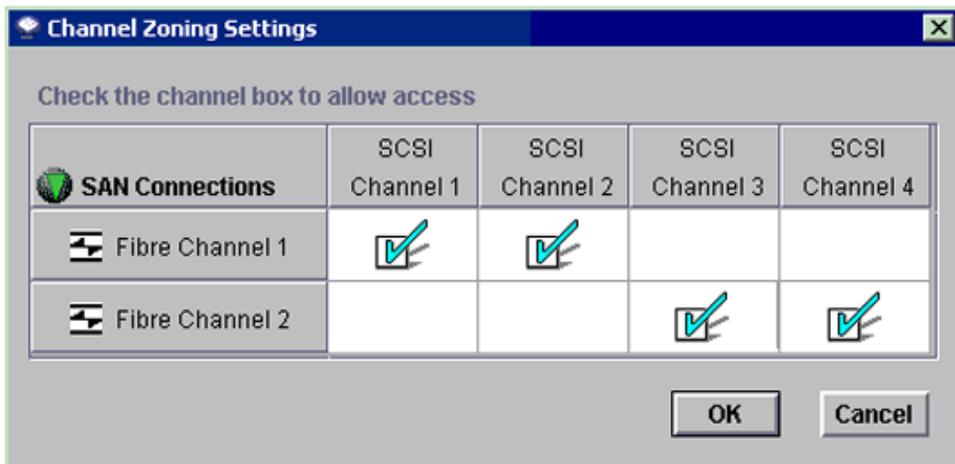


Abbildung 1-17. Kanalzoneneinstellungen für Bandbibliothek mit fünf bis sechs Laufwerken



Die Abbildungen 1-18 bis 1-19 zeigen, wie eine Bandbibliothek mit bis zu zwei Laufwerken bei Einsatz eines Fibre-Channel-Brückenmoduls konfiguriert werden sollte, um die Bandleistung über FC zu optimieren.

Abbildung 1-18. Bandbibliothek mit einem Laufwerk

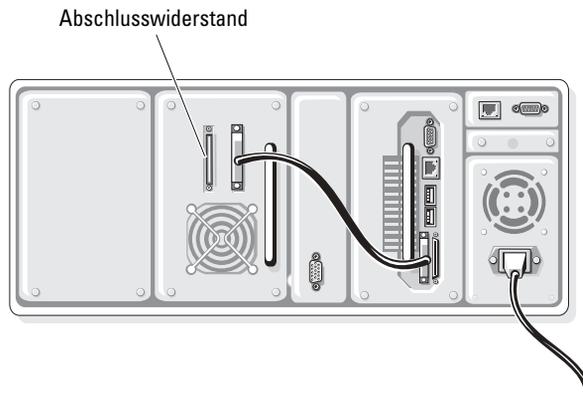
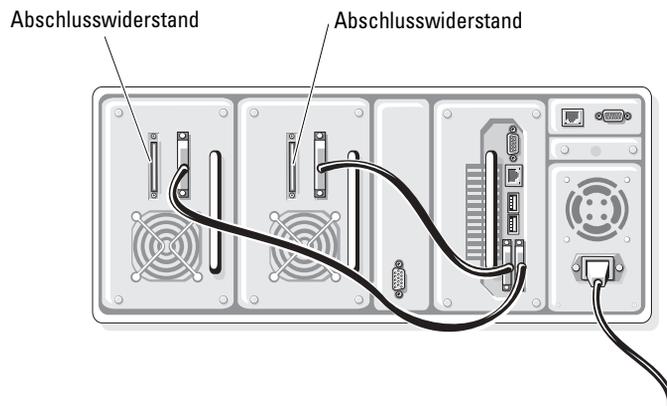


Abbildung 1-19. Bandbibliothek mit zwei Laufwerken



Dell™ PowerVault™ システム

テープドライブとライブラリの
パフォーマンスに関する考慮事項

本書の内容は予告なく変更されることがあります。
© 2005 すべての著作権は Dell Inc. にあります。

Dell Inc. の書面による許可のない複製は、いかなる形態においても厳重に禁じられています。

本書で使用されている商標について：Dell、DELL ロゴ、および PowerVault は Dell Inc. の商標です。
EMC および PowerPath は EMC Corporation の登録商標です。

本書では、必要に応じて上記以外の商標や会社名が使用されている場合がありますが、これらの商標や会社名は、一切 Dell Inc. に所属するものではありません。

2005 年 6 月

目次

はじめに	105
ホストのバックアップに関する一般的な考慮事項	105
テープドライブとデータに関する考慮事項	105
ハードドライブと RAID アレイの構成	107
テープライブラリで複数のドライブを使用する際のパフォーマンスに関する一般的な考慮事項	113
SCSI 構成	113
SAN 構成	114

はじめに

最近ではテープドライブの転送速度が向上したため、ホストサーバーとテープドライブが同じ速度でデータを処理できるかどうかを判断する際に、RAID (Redundant Array of Inexpensive [または Independent] Disks) 構成やハードドライブの仕様などのホスト側の多くの要因を考慮する必要があります。ホストサーバーからテープドライブへのスループットを制限する可能性のある一般的な設定と属性については、「ホストのバックアップに関する一般的な考慮事項」で説明します。

テープライブラリに複数のドライブが設置されているため、複数のテープドライブの潜在的なスループットに対応するには、ホストの帯域幅を増す必要があります。マルチドライブユニットの潜在的なファイバー制限、および推奨するケーブル設定については、「テープライブラリでマルチドライブを使用する際のパフォーマンスに関する一般的な考慮事項」で説明します。

ホストのバックアップに関する一般的な考慮事項

本項に挙げる考慮事項は、SCSI とストレージエリアネットワーク (SAN) の両方のテープバックアップ構成に適用されます。

テープドライブとデータに関する考慮事項

パフォーマンスの評価では、次の問題を検討する必要があります。

SCSI コマンドからのオーバーヘッド。 SCSI バスに対するコマンドのオーバーヘッドは、どの SCSI デバイスに対しても理論上の最高転送速度の達成を妨げる要因になります。このオーバーヘッドの原因はテープバックアップソフトウェアではありません。ソフトウェアは、データがテープに書き込まれる速度を測定するだけです。たとえば、ドライブがデータを 80 MB/秒 で処理していても、データ書き込みの速度は 77 MB/秒 でしかないかもしれません。バックアップソフトウェアが報告する速度は後者です。

テープのブロックサイズ。 ほとんどのテープドライブにとって最適なのは、64 Kb のブロックサイズです。ただし、サイズが大きくなってもパフォーマンスは向上しないものの、一部のバックアップアプリケーションではユーザーがブロックサイズを変更することができます。使用するブロックサイズが 64 Kb に満たないと、実際にパフォーマンスの低下を招くことがあります。テープデバイスのブロックサイズを調節する手順については、バックアップソフトウェアの『ユーザーズガイド』を参照してください。

バックアップソフトウェアのバッファサイズ。 バックアップのパフォーマンスを最適にするには、バックアップソフトウェアのバッファは可能な限り大きくすべきです。一部のアプリケーションではユーザーがバッファサイズを変更できるため、ドライブとの間でやり取りされるデータの流れを一定に保つのに役立ち、小さなファイルの場合は特に転送速度が大幅に向上します。バッファが大きいほど多くのデータを保持でき、ディスクがデータの検索に費やす時間が短縮できますが、その反面、メモリと CPU のパフォーマンスに影響を与える場合があります。詳細については、テープバックアップアプリケーションの『ユーザーズガイド』を参照してください。

ドライバおよびファームウェア。 SCSI またはファイバーコントローラとテープドライブに最新のドライバとファームウェアがインストールされていることを常に確認してください。support.dell.com から該当する Dell PowerVault テープ製品用の最新のドライバとファームウェアをダウンロードします。

テープドライブとハードドライブを別々のホストバスアダプタ (HBA) の別々のコントローラ (内蔵または外付け) に取り付けます。 コントローラの性能によっても多少左右されますが、テープドライブの HBA とハードドライブの HBA を別々にして最大のスループットを確保しておくことが最善です。オンボードのデュアルモード SCSI/RAID コントローラは、ほとんどが 1 つのプロセッサを共有するため、これらのコントローラは RAID アレイとテープドライブの間に帯域幅を分け合う必要があります。そのため、1 つのコントローラがハードディスクとテープドライブの間の読み取り / 書き込みを処理するだけでなく、必要なパリティ情報を計算してハードドライブに書き込んでいます。RAID アレイとパリティバイトの詳細については、「ハードウェア RAID 構成に関する考慮事項」を参照してください。

ドライブヘッドの汚れ、または古いメディア。 テープドライブヘッドが汚れていたり、メディアが古かったりすると、エラー率が高くなり、その結果、読み取り / 書き込みの速度が低下する場合があります。ドライブがテープ上のブロックの書き換えや再読み取りを試みるたびに、パフォーマンスが低下します。読み取り / 書き込みのエラーが一定のしきい値に達すると、ドライブは通常、クリーニングを要求します。ドライブヘッドのクリーニングは、要求時か、または定期的に行うことが重要です。

メディアが古くなったり使いすぎたりすると、不良ブロックが発生する可能性が高くなります。LTO メディアの標準的な寿命は、テープ全体の読み取り / 書き込みで約 75 回です。

スピードマッチング。 最近の LTO ドライブは、ドライブに送り込まれるデータの速度の低下が、圧縮されていないデータの最大転送速度の約 2 分の 1 までなら、速度を一致させます (スピードマッチング)。データがドライブに送り込まれる速度がスピードマッチングの下限を下回ると、ドライブは停止して、バッファが満たされるのを待ち、巻き戻し、バッファへの書き込みを試みる必要があります (この動作は「バックヒッチング」と呼ばれています)。

たとえば、Dell PowerVault 110T (LTO2 および LTO3) テープドライブが LTO-3 メディアに書き込みを行う際には、スピードマッチングの下限は 30 MB/ 秒です。ホストサーバーがデータを送り込む速度が 20 MB/ 秒なら、ドライブはバッファが満たされるのを待つ間、「バックヒッチ」します。この場合、有効なスループットは 20 MB/ 秒未満の値となります (おそらく、15 MB/ 秒に近い値)。

テープドライブのパフォーマンスの確認

テープドライブの製造元が、スループットの確認に使用できるパフォーマンス診断モードをドライブに組み込んでいる場合もあります。PowerVault 110T LTO-2 および LTO-3 (ファームウェア 53XX またはそれ以降) には、診断モード「F」が提供されており、ドライブとメディアに対して読み取り / 書き込みのパフォーマンステストをすばやく行うことができます。パフォーマンステストの結果がドライブの指定最高速度から 6% 以上外れた場合、テストは失敗してエラーメッセージが表示されます。テストが正常に終了した場合、エラーメッセージは表示されません。診断モード「F」の詳細については、テープドライブの『ユーザズマニュアル』を参照してください。



注意：診断モード「F」を実行するには、診断テストの一部として、安全に上書きできるメディアが必要です。重要なデータが保存されているメディアは使用しないでください。診断テストに使用するメディアに保存されているデータは、すべて失われます。

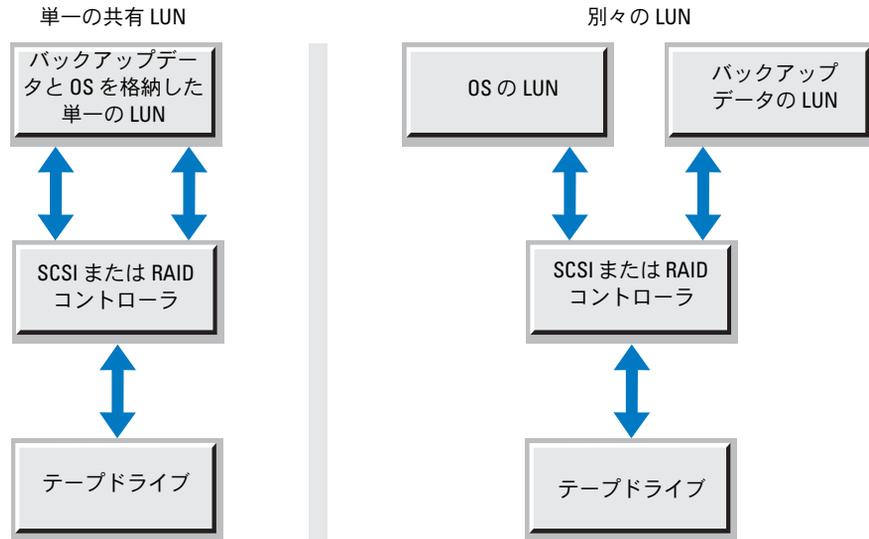
ハードドライブと RAID アレイの構成

ハードドライブとディスクアレイ (内蔵および外付け) のいくつかの属性は、バックアップまたは復元のパフォーマンスに影響する場合があります。それらの属性、およびバックアップと復元の速度を最適にするための推奨される構成については、以下の各項で説明します。テープドライブの持続可能なスループットがディスクアレイの持続可能なスループットを上回る場合、テープドライブのピークパフォーマンスは実現しません。

ハードドライブの構成に関する一般的な考慮事項

異なる LUN 上のデータとオペレーティングシステム (OS)。 論理ユニット番号 (LUN) 上のデータを OS の LUN とは別にバックアップすれば、ハードドライブが OS の動作とバックアップの動作の間でアクセスとオーバーヘッドを分割せずに済みます。そのためには、1 つのハードドライブまたはディスクアレイに OS を入れ、バックアップするデータは物理的に別のハードドライブまたはディスクアレイに入れる必要があります。

図 1-1 シングルチャネルの帯域幅対 2 チャネルの帯域幅



ハードドライブのパフォーマンス

テープドライブはデータの書き込みを連続して行うように設計されており、バックヒッチングが発生しないようにドライブの動作を連続的に行うには、データの送り込みを一定速度にする必要があります。他方、ハードドライブはランダムアクセスデバイスです。したがって、データがドライブ上で分散している場合、ハードドライブは時々、テープドライブに連続データを送り込むために苦勞することがあります。そのため、ドライブはデータの小さなブロックを探し続けなければなりません。

テープドライブに送り込まれるデータのスループットに大きな影響を与えるハードドライブの属性が、ほかにもあります。

回転速度。 通常 RPM（1 分間あたりの回転数）で表示されるハードドライブの回転速度は、ドライブのプラッターアセンブリが 1 分間に丸々何回転できるかを測定したものです。この回転速度は、ランダムアクセスの回数と連続的な転送速度の両方に直接影響します。回転速度が早いほど、ドライブはデータに速くアクセスできます。

ランダムアクセスタイムまたはシークタイム。 通常ミリ秒で表示されるシークタイムは、ドライブのヘッドがディスク上のデータを検出するのに要する時間です。ハードディスクのシークタイムは、読み取り / 書き込みヘッドがプラッターの表面上のトラック間を移動するのに要する時間を測定したものです。ハードディスクはランダムアクセスデバイスであるため、データは事実上、ディスクのどのセクターにでも格納できます。そのデータへのアクセスに要する時間が長いほど、ドライブの全体的なスループットは遅くなります。ハードドライブに小さなファイルが多数格納されている場合、この属性は非常に重要です。

ファイルが小さいほど、ドライブはファイルをディスクとの間で読み取り / 書き込みするための「シーク」が多くなります。したがって、多くの小さなファイルを転送するには、ディスクの読み取り / 書き込みが非常に遅くなります。

順次的 / 持続的転送速度 (STR)。STR は、ドライブが実際にそのプラッターからデータを読み込んだり、プラッターにデータを書き込んだりする速度を測定したものです。バックアップされるデータが 1 つの大きな連続したファイルである場合、持続的なスループットはドライブの最大 STR に近くなります。ただし、実環境においてはデータの削除と書き込みが繰り返されるため、データはプラッター内のあちこちに散在しています。ハードドライブのデフラグを行うと、速度をそのドライブの最大 STR に近づけることができます。

バッファ (キャッシュ)。バッファとは、直前に書き込みや保存が行われたデータを保持しておく、ドライブ上のメモリ容量のことです。バッファが大きいほど、保持できるデータ量は多くなり、ディスク上のデータをシークする時間が短縮されます。

ハードウェア RAID 構成に関する考慮事項

RAID の概要

本項では、通常の RAID 構成の概要を示し、RAID 構成がバックアップと復元の速度に与える影響について説明します。RAID アレイとは、単一のストレージシステムまたは LUN として動作する一組のハードディスクのことです。理論上、各ハードドライブのチャンネルを通じて一度にデータを転送できるため、全体のスループットは、アレイ内のドライブの総数の倍数から、以下の各項で説明するオーバーヘッドと冗長性を差し引いたものになります。

RAID 構成の場合は、インタフェースの速度が重要になります。各ドライブはインタフェースの帯域幅を共有しているからです。たとえば、シングルの Ultra160 ドライブ 1 台では、最高速度は 40 MB/秒です。したがって、同一のドライブタイプで構成されるディスク 5 台の RAID 0 アレイは、200 MB/秒で読み取り / 書き込みができるはずですが、ただし、Ultra160 のインタフェースはアレイの最大速度を 160 MB/秒に限定します。

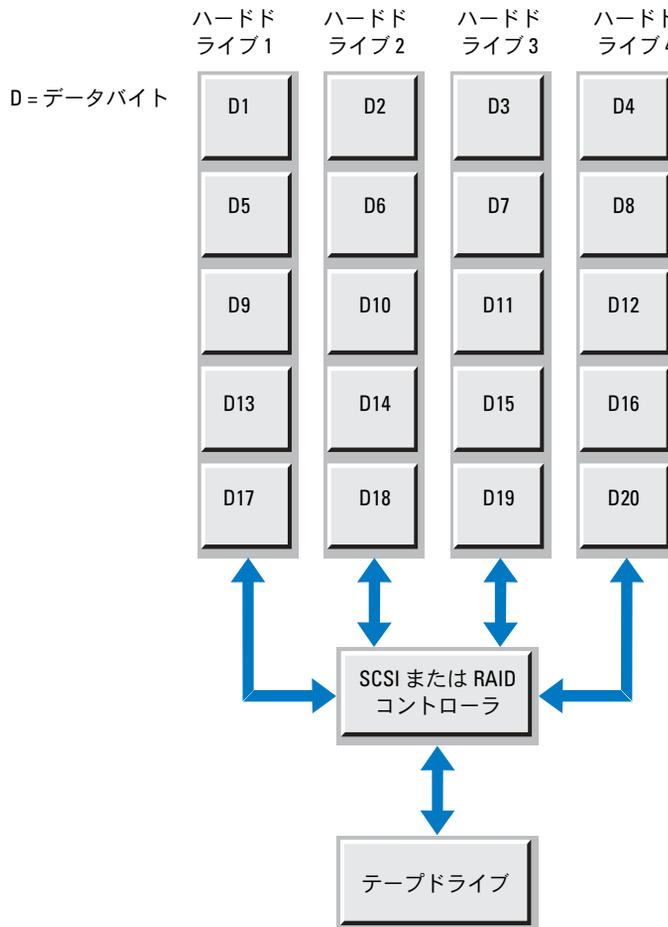
外付けディスクアレイは、SAN の場合は特に、キャッシュメモリが豊富に搭載され、I/O パフォーマンスが大幅に向上している場合があります。このキャッシュにより、アレイに書き込む動作のパフォーマンスが大幅に向上し、頻繁にアクセスするデータを格納することで読み取りのパフォーマンスも向上します。テープのパフォーマンスに対する影響に関しては、キャッシュにより、アレイにデータを復元したり、アレイからデータをバックアップしたりする際の RAID の制限のほとんどが隠されます。ただし、キャッシュを搭載した外付けアレイからのバックアップ動作には、依然として RAID 構成の制限による影響が見られる場合があります。依然としてディスクから読み込まなければならないデータがあるためです。

RAID 0

RAID 0 では、2 台以上のディスクを結合して、1 つの LUN のように 1 つの仮想ドライブを作成する（一般に「ストライピング」と呼ばれています）ことができます。ストライピングと呼ばれるのは、データが一度に 1 台のディスクにではなく、アレイ内のすべてのディスクにまたがって書き込まれるためです。したがってスループットは、1 台のハードディスクに対して 1 つのチャンネルではなく、 n のチャンネル（ n = アレイ内のハードドライブの数）に分散されます。そのため、読み取り / 書き込みのパフォーマンスは向上しますが、フォールトトレランスは実現しません。

図 1-2 は、RAID 0 構成の 4 台のハードドライブを示したものです。データは 4 台のハードドライブすべてにストライピング（分散）され、アレイとの間の読み取り / 書き込みのチャンネルが 4 つになっています。

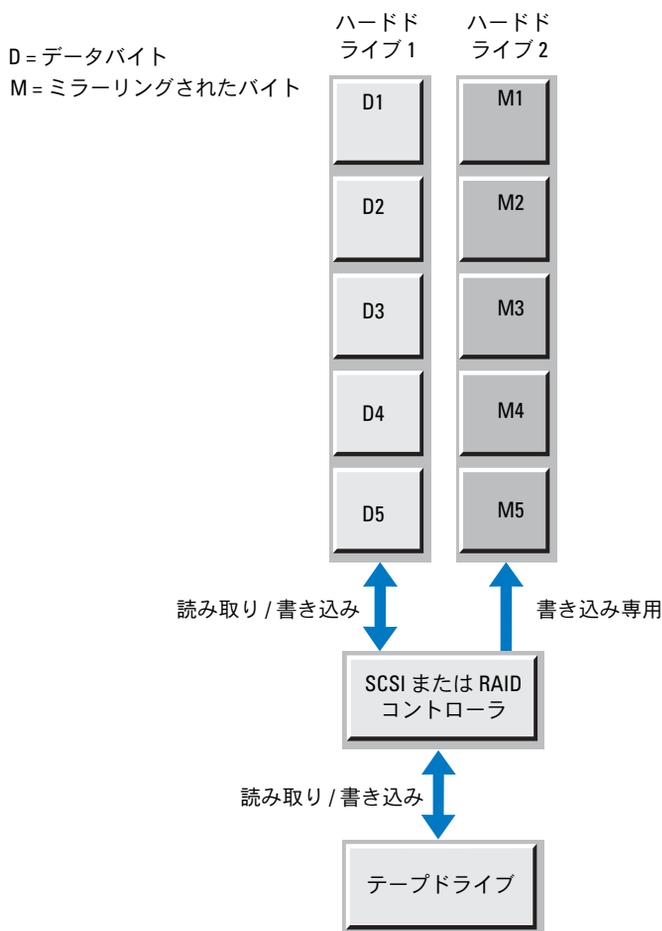
図 1-2 RAID 0 構成の例



RAID 1

RAID 1 では、データが 2 つのディスクにまたがって 2 度同時に書き込まれます（「ミラーリング」とも呼ばれます）。1 台のドライブに障害が発生した場合、システムはデータを失うことなく、もう一方のドライブへの切り替えを行います。テープドライブのバックアップ中は、プライマリドライブから読み込むため、RAID 1 のアレイからの読み取り速度はシングルドライブの場合とほぼ同じです。ただし、RAID 1 のアレイに書き込むテープドライブからの復元パフォーマンスは、プライマリディスクおよびミラーディスクへの書き込み中に行われるエラーチェックおよび訂正（ECC）のために低下する場合があります。このパフォーマンス低下の主な原因は、ミラーリングが多くの場合 CPU または RAID コントローラ上で行われることにあります。したがって、新しい RAID コントローラは、プロセッサが新しく性能も優れているため、速い場合が多いと言えます。

図 1-3 RAID 1 構成の例

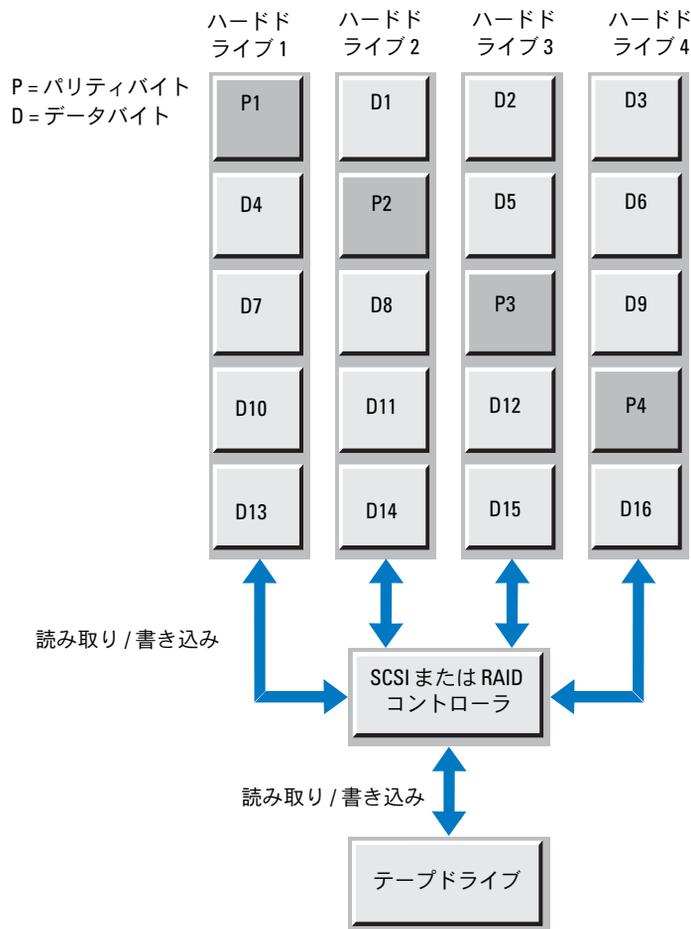


RAID 5

RAID 5 アレイでは、データはバイトレベルでディスクアレイにまたがってストライピングされ、エラー訂正データまたはパリティデータも、ディスクアレイにまたがってストライピングされます。RAID 5 アレイは、ランダム読み取りに非常に優れたパフォーマンスを発揮します。この読み取りパフォーマンスは一般に、アレイ内のディスク数が増えるにつれて向上します。大きなディスクアレイの場合、読み取りパフォーマンスは実際に、RAID 0 のアレイのパフォーマンスを上回ることがあります。データが追加ドライブに分散されるからです。また、通常読み取り中にはパリティ情報は不要です。

テープから RAID 5 アレイへの復元は非常に少ない場合が多くなります。パリティ情報を計算して書き込むための付加的なオーバーヘッドが発生するためです。

図 1-4 RAID 5 構成の例



結論として、RAID 0 は読み取り / 書き込みのパフォーマンスにとっては総合的に最適な構成である場合が多いと言えますが、冗長性は提供しません。RAID 1 では、アレイに書き込まれるすべてのデータがミラーリングされ、読み取りもシングルディスクから行われるため、総合的にパフォーマンスは最低です。RAID 5 は読み取りのパフォーマンスには優れているものの、書き込みのパフォーマンスは標準的です。ただし、アレイにディスクを追加するとパフォーマンスが向上します。RAID がキャッシュメモリが豊富に搭載されたエンクロージャ内にある場合、復元動作中のパフォーマンスの制約は軽減されます。バックアップ動作はそれでも RAID 構成の制約を受けます。また、アレイの特性は、「ハードドライブのパフォーマンス」に示されているハードドライブの具体的な特性によって左右されるところが依然として大きいと言えます。

テープライブラリで複数のドライブを使用する際のパフォーマンスに関する一般的な考慮事項

テープライブラリなどのデータのバックアップに複数のテープドライブを同時に使用する場合は、ハードウェア構成に関して考慮すべき側面がほかにもあります。ハードウェアとケーブル接続をセットアップする際に、パフォーマンスを考慮した単純な方法を用いることで、付加的なスループットのボトルネックを限定することが可能です。

SCSI 構成

テープライブラリに提供されている最新の高性能テープドライブは、SCSI インタフェース標準の Ultra160 仕様をサポートします。したがって、SCSI を利用するバックアップサーバーが最大のパフォーマンスを得るには、Ultra160 以上のデータ速度をサポートする HBA がインストールされている必要があります。この要件を満たす SCSI HBA がインストールされていると、各テープドライブは SCSI バス上で 160 MB/ 秒の速度でホストと通信できます。テープドライブの速度と比較して SCSI バスのデータ速度が速いほど、デバイスのパフォーマンスを犠牲にせずに、複数のデバイスを同じバスに接続できます。ただし、これには限度があります。

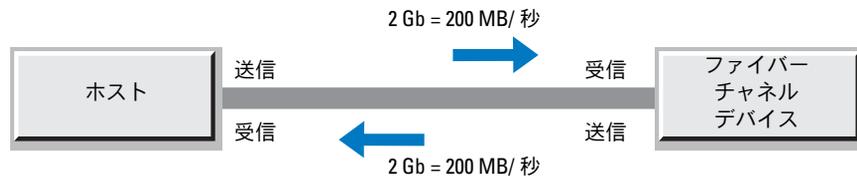
Ultra160 バスのデータ速度である 160 MB/ 秒は、バスに接続されるすべてのデバイスにとって可能な最大のデータスループット速度です。したがって、シングルテープドライブではバスの帯域幅すべてを使い切ることができません。なぜなら、テープとの間で読み取り / 書き込みができる速度は 80 MB/ 秒（ネイティブ）までに限られるからです。ただし、複数のテープドライブのそれぞれが最大のネイティブパフォーマンスで動作している場合は、テープドライブを組み合わせ、バスが提供している 160 MB/ 秒の速度をフルに消費することができます。これ以降に同じバスに追加のドライブを接続すると、台数が増えるごとに各ドライブの平均パフォーマンスが低下します。

したがって、テープライブラリで最大のパフォーマンスを得るには、各 SCSI バスに接続するテープドライブは 2 台までとすることをお勧めします。詳細および説明図については、「推奨するケーブル設定」を参照してください。Ultra160 以上をサポートする SCSI HBA を使用する必要がありますが、Ultra320 HBA にアップグレードしても、テープドライブの仕様が Ultra160 であれば、パフォーマンスはそれ以上向上しません。

SAN 構成

ファイバーチャネル (FC) は多くの点で SCSI よりも優れています。まず、SCSI の距離制限が克服されており (LVD SCSI では 12 m なのに対して、短波の 2 Gb FC リンクでは 300 m)、データの転送速度も高速です。SCSI のようなバスベースのアーキテクチャではなく、シリアルネットワークプロトコルとして、FC は SAN の実装にとって最適なプロトコルともなっており、データストレージリソースの統合を可能にしています。また、FC の各接続は送信リンクと受信リンクで構成され、全二重動作が可能になっています。このため、データを双方向に同時送信できます。したがって、シングル FC 接続によるバックアップ動作中に、通信を交代せずにソースとテープとの間でデータの読み取り / 書き込みを行うことができ、接続の帯域幅は事実上倍加します。図 1-5 を参照してください。

図 1-5 ファイバーチャネルリンクの図

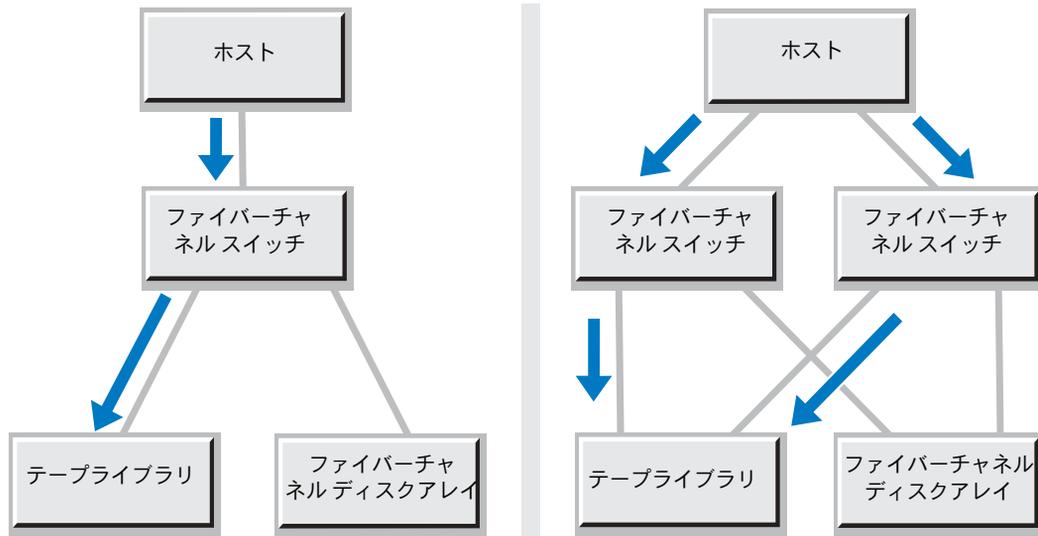


テープライブラリを SAN で構成しても、パフォーマンスに影響を与えるさまざまな要因があります。たとえば、FC リンクの色度、ソースとテープライブラリとの間のデータフロー、外付けストレージレイのパフォーマンスの限界などです。ソリューションの全体的なセットアップと管理について理解していれば、これらの要因の多くは回避することができます。

SAN では FC プロトコルによって広いデータ帯域幅が提供されているものの、FC リンクによってパフォーマンスが制限されるような状況避けるには、テープドライブに対して適切な配慮が必要です。2 ギガビット (Gb) による FC リンクの色度転送速度は、200 MB/秒です (すなわち、送信リンクと受信リンクでそれぞれ 200 MB/秒)。したがって、複数のテープドライブを同一リンクで操作しようとする、リンクの総帯域幅を超える場合があります。ホストがレガシーの 1 Gb アダプタで動作している場合、2 台のドライブにデータをバックアップするだけでパフォーマンスが大幅に限定される可能性があります。

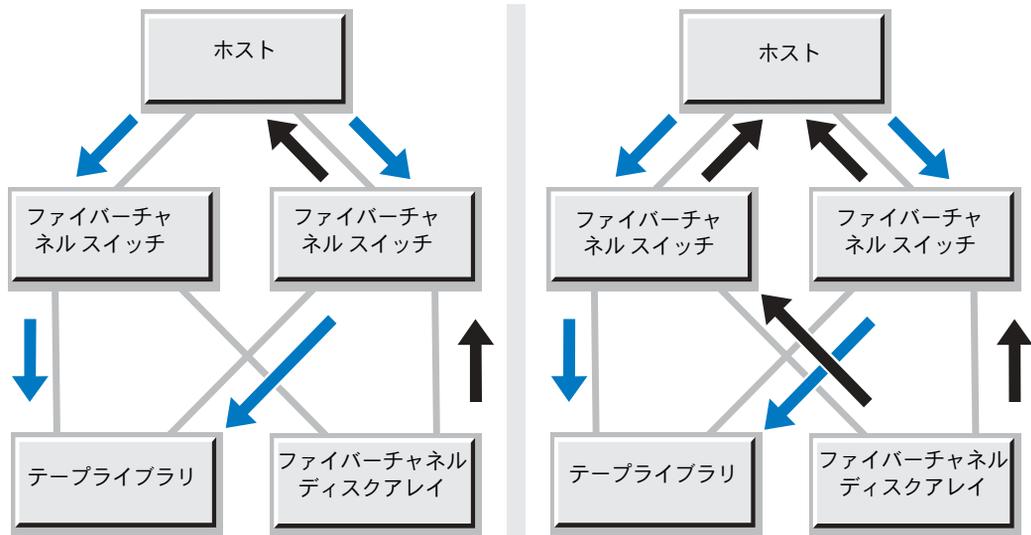
したがって、2 Gb のリンク上で 3 台以上のテープドライブを同時に使用する場合は、単一のリンクに頼るのではなく、バックアップをいくつかの接続に分散する必要があるかもしれません。このような場合に、SAN ソリューションのトポロジを理解しておく役に立ちます。バックアップ動作中にソースから読み取られ、テープに書き込まれるデータパスを追跡しておけば、管理者がボトルネックの可能性を認識するのに役立ちます。何らかのボトルネックが確認された場合は、構成に応じて対策を講じることができます。たとえば、複数のドライブを一度に操作する必要のあるバックアップソリューションの場合は、テープハードウェアを別々のファブリックに分割すると、接続の分割によってパフォーマンスが向上する可能性があります。図 1-6 を参照してください。

図 1-6 テープライブラリへのデータフロー：シングルと分割の比較



同時に、外付けの FC ディスクアレイから読み取られるデータにボトルネックが存在する場合は、EMC® PowerPath® などの I/O 管理ソフトウェアを追加のファブリック接続と共に使用すれば、データが自動的に複数のパスに負荷分散され、パスフェイルオーバーを通じて可用性が増します。図 1-7 を参照してください。図の左側は、すべてのデータが単一のリンクを通るように強制されている SAN ディスクアレイを示したものです。この場合、テープハードウェアへのデータフローを遅くするボトルネックができます。右側は、データ転送を 2 つのリンクに分割することで、負荷分散によってアレイからの I/O 帯域幅が 2 倍になる仕組みを示したものです。

図 1-7 ボトルネックの発生したデータフローと負荷分散されたディスクアレイの比較



最後に、SAN 上の FC ディスクアレイでも、「ハードドライブと RAID アレイの構成」で説明したものと同様のパフォーマンスの制約事項が発生する場合があります。したがって、ディスクアレイのパフォーマンス特性を改善することも、SAN でのバックアップ速度に対して直接的な効果があります。

ライブラリファイバーチャネルブリッジを利用した SAN 構成

一部のテープライブラリは、ファイバーチャネルブリッジモジュールを経由して SAN に接続できます。このモジュールは SCSI プロトコルと FC プロトコルの間のブリッジとして機能し、ほとんどのネイティブ FC ライブラリでは利用できない付加的な管理、セキュリティ、および操作機能を提供します。これらの機能の詳細については、ご使用のテープライブラリ用のファイバーチャネルブリッジに付属の『ユーザーズガイド』を参照してください。

テープライブラリの構成によっては、ファイバーチャネルブリッジモジュールがボトルネックとなり、テープドライブのパフォーマンスの低下を招くことがあります。これは、SCSI と FC の通信をブリッジするのに必要なファイバーチャネルブリッジモジュールの処理能力では、一部のマルチドライブ構成によって提供されるデータスループットの全体を満たすことができないためです。こうした制約があるものの、ほとんどのデータバックアップソリューションでは、ファイバーチャネルブリッジモジュールがテープパフォーマンスにおける主要な制限要因となることはありません。専用バックアップサーバーでは、データを複数のテープドライブに送ることで、ホストでの制約が悪化することがよくあります。この結果、平均ドライブパフォーマンスは、ファイバーチャネルブリッジモジュールが制約要因となる場合よりもさらに低下します。

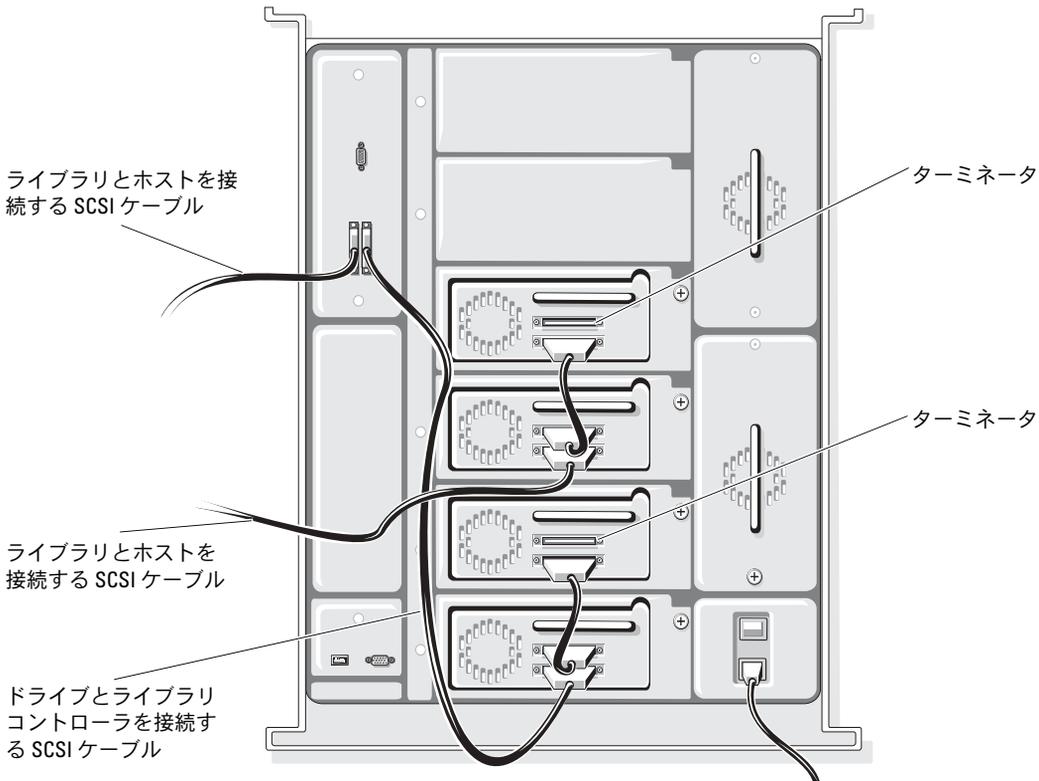
推奨するケーブル設定

ホストへの SCSI ケーブル接続

6 台までのテープドライブを備えたテープライブラリ

テープライブラリをホストに SCSI 接続する場合は、シングルバスに 3 台以上のドライブを接続しないでください。SCSI バスがスループットの最大化を阻害しないようにするには、5 台または 6 台のテープドライブを備えたライブラリには、追加の SCSI コントローラが必要です。

図 1-8 6 台までのテープドライブを備えたライブラリの SCSI ケーブル接続

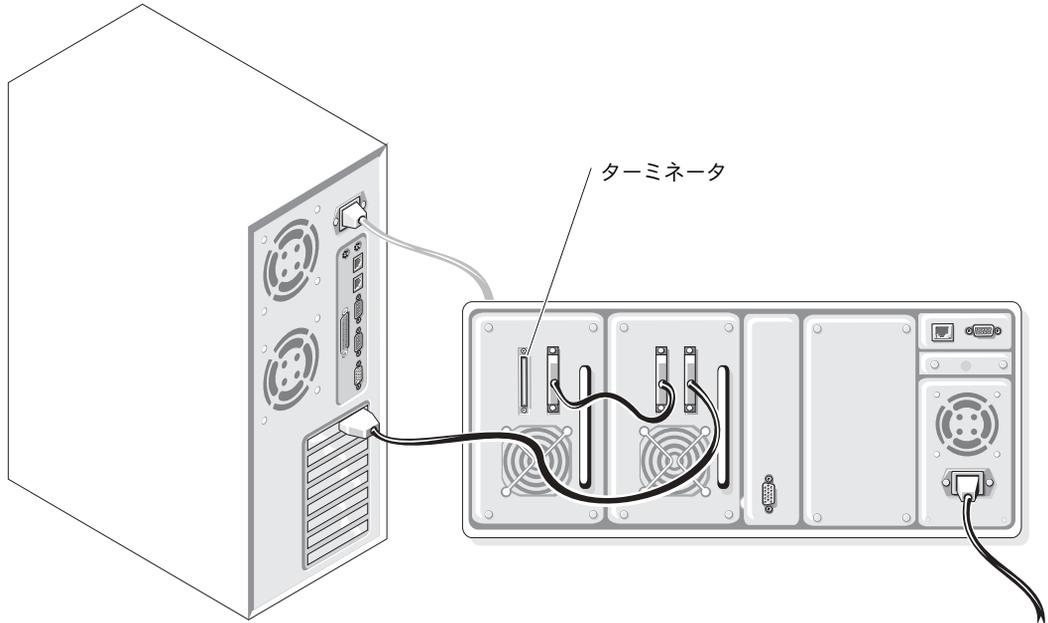


2 台までのテープドライブを備えたテープライブラリ

完全に設定されたドライブ 2 台のテープライブラリ内のドライブを同一の SCSI バス上のホストにケーブル接続し、バックアップパフォーマンスの大幅な低下を招かないようにすることは可能です。シングル SCSI バスに 2 台のドライブを接続した場合のバックアップ速度は、別々のバスに 2 台のドライブを接続した場合のバックアップ速度に匹敵します。

ただし、バックアップアプリケーションの検証機能を有効にするお客様は、2 台のドライブを 2 つの SCSI バスに分割することで、検証パフォーマンスを向上することが可能です。そうすれば、検証パフォーマンスは最大 25% 向上します。

図 1-9 2 台までのテープドライブを備えたライブラリの SCSI ケーブル接続



ファイバーチャネルブリッジへのドライブのケーブル接続

図 1-10 から 1-17 までは、ファイバーチャネルを介してテープのパフォーマンスを最適化するために、ファイバーチャネルブリッジモジュールを用いて最大 6 台のドライブを備えたテープライブラリを構成する方法を示したものです。

図 1-10 テープドライブが1台の場合のファイバーチャネルブリッジのケーブル接続

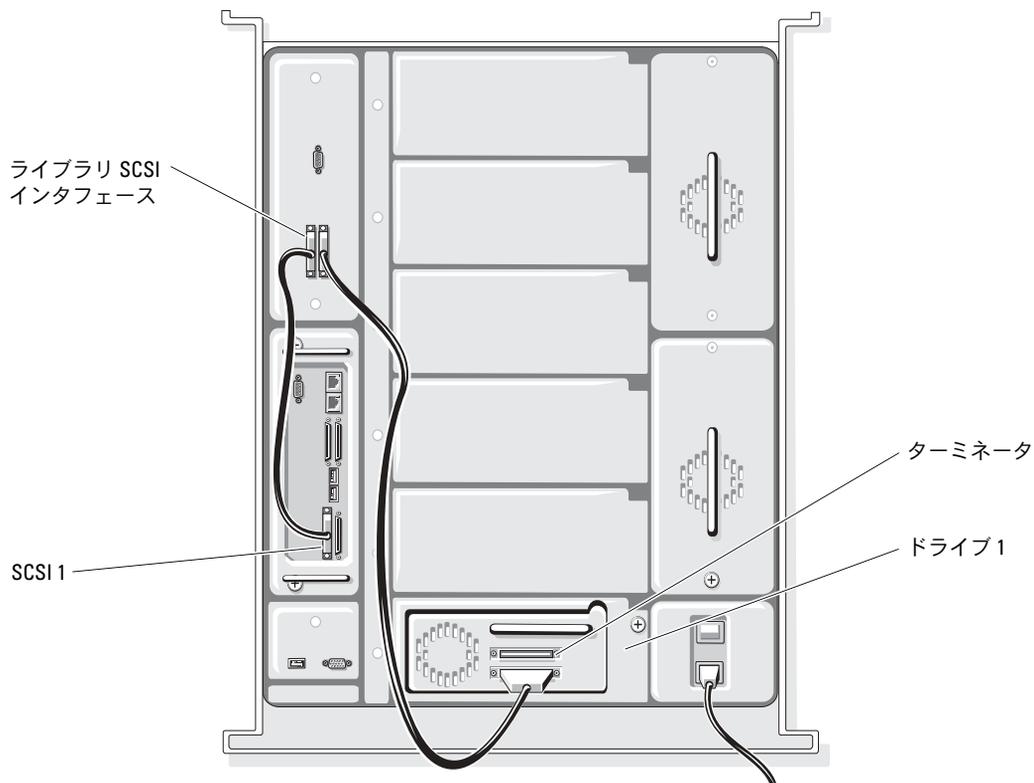


図 1-11 テープドライブが2台の場合のファイバーチャネルブリッジのケーブル接続

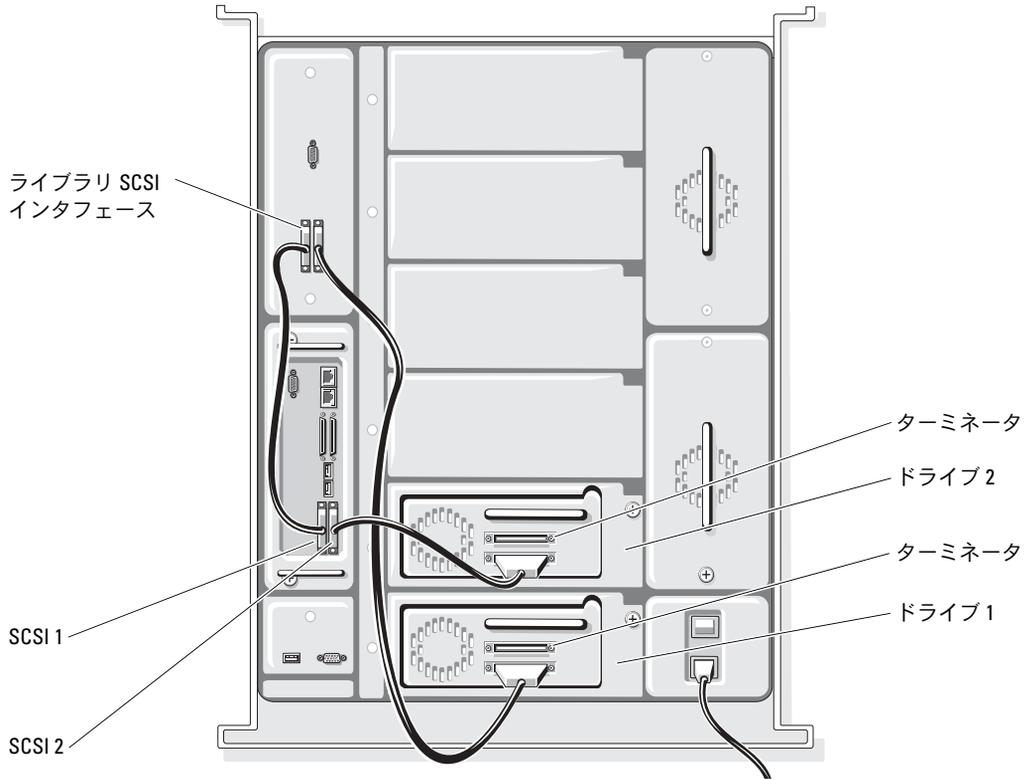


図 1-12 テープドライブが3台の場合のファイバーチャネルブリッジのケーブル接続

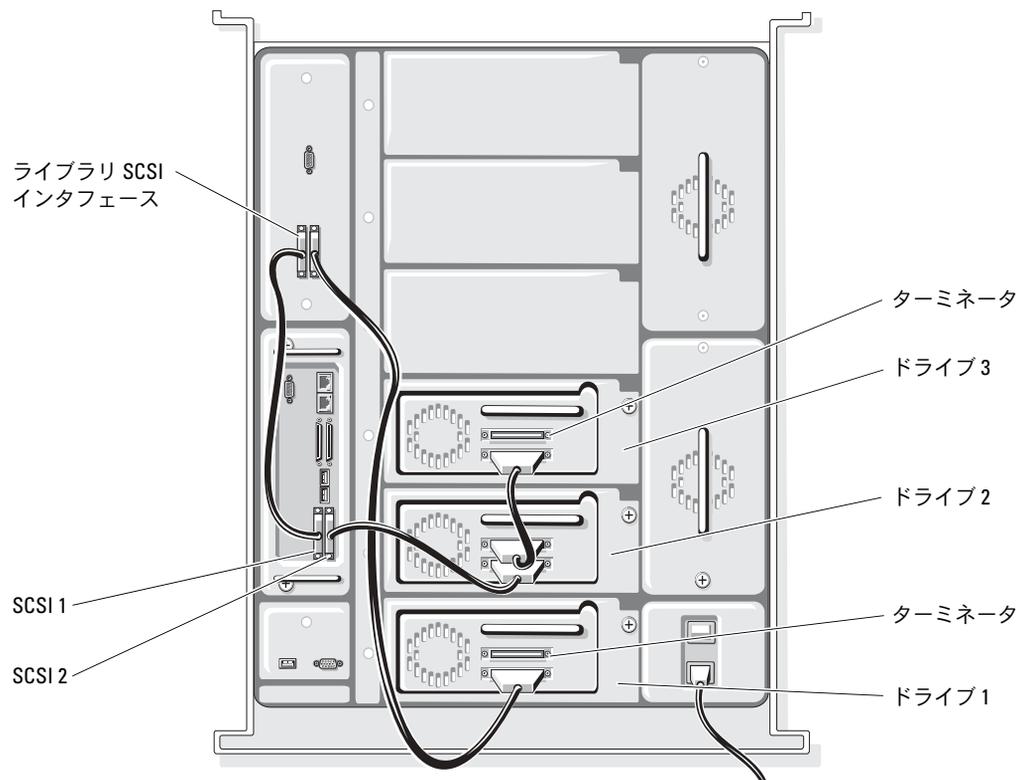


図 1-13 テープドライブが4台の場合のファイバーチャネルブリッジのケーブル接続

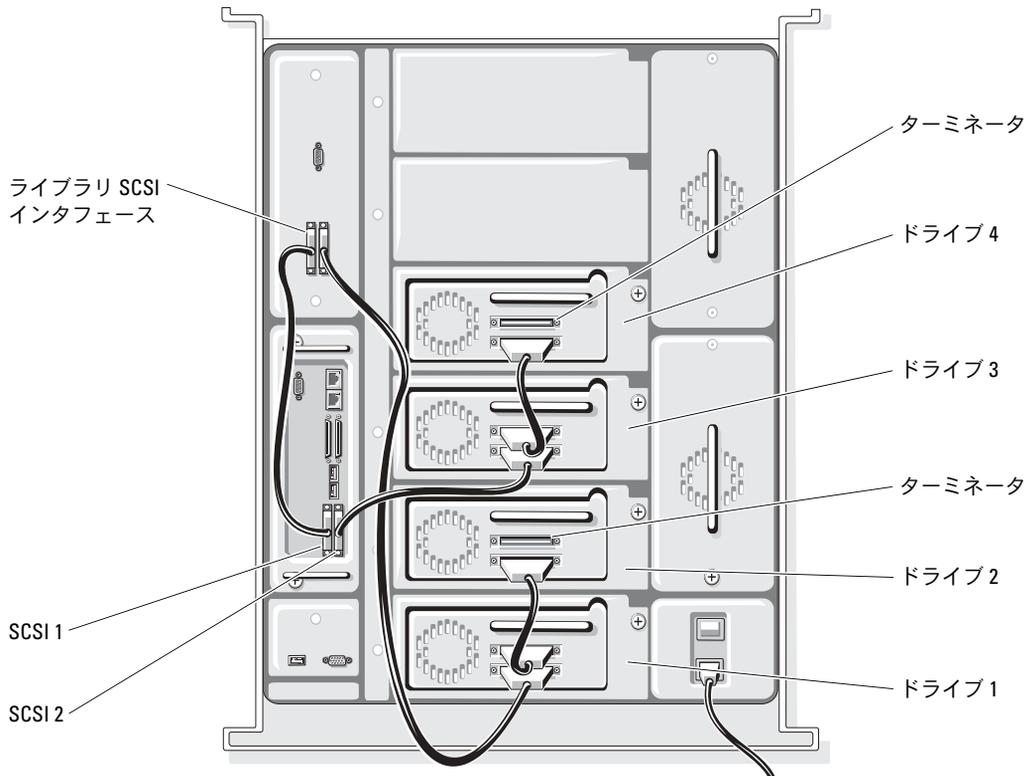


図 1-14 1台から4台のドライブを備えたテープライブラリ用のチャネルゾーニング設定

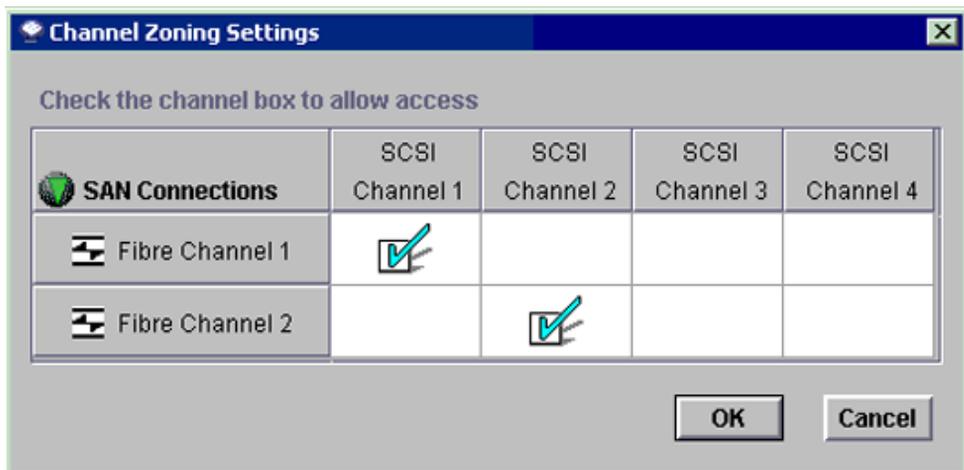


図 1-15 テープドライブが 5 台の場合のファイバーチャネルブリッジのケーブル接続

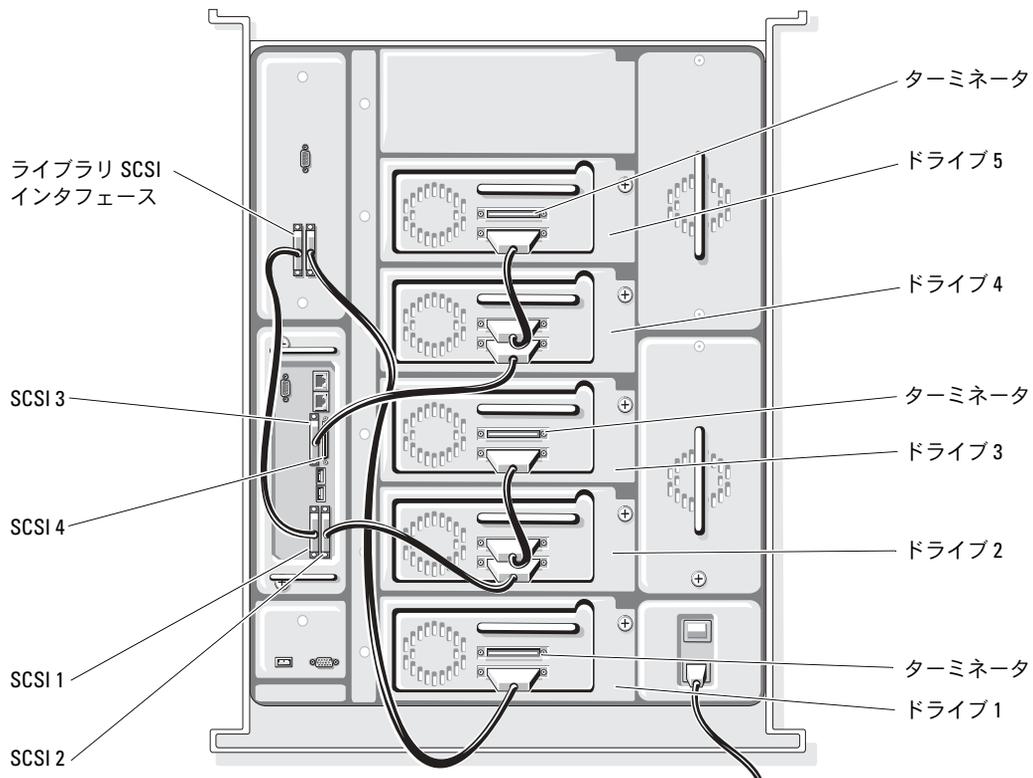


図 1-16 テープドライブが6台の場合のファイバーチャネルブリッジのケーブル接続

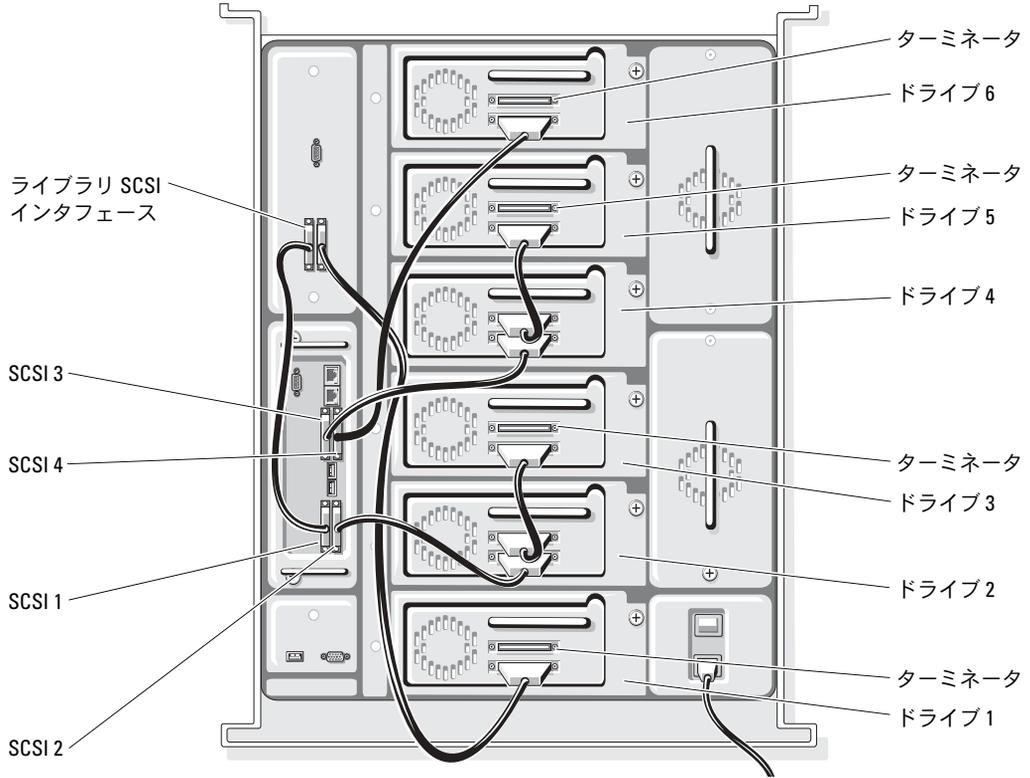


図 1-17 5台または6台のドライブを備えたテープライブラリ用のチャンネルゾーニング設定

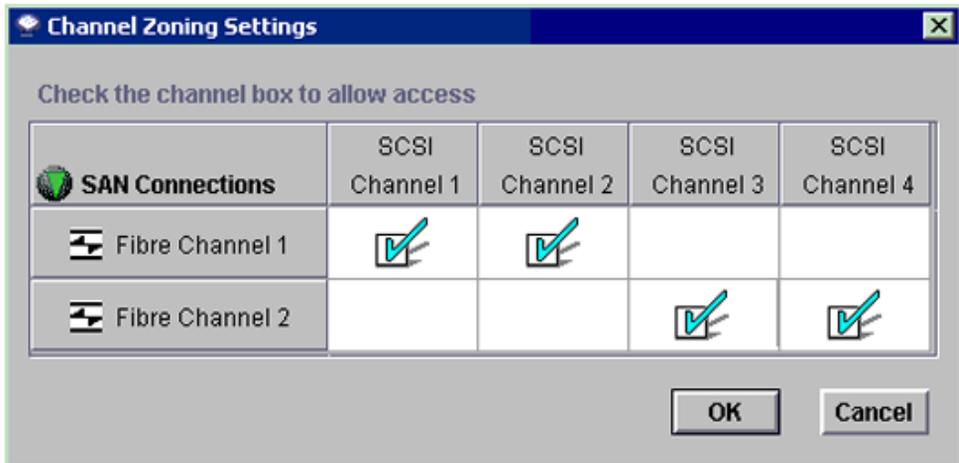


図 1-18 と 図 1-19 は、ファイバーチャネルを介してテープのパフォーマンスを最適化するために、ファイバーチャネルブリッジモジュールを用いて最大 2 台のドライブを備えたテープライブラリを構成する方法を示したものです。

図 1-18 ドライブ 1 台のテープライブラリ

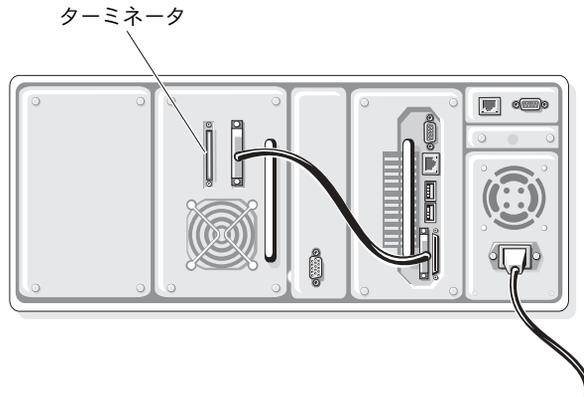
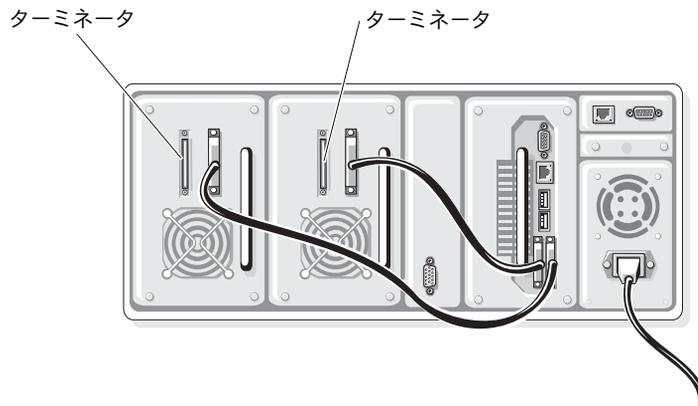


図 1-19 ドライブ 2 台のテープライブラリ



Dell™ PowerVault™ 시스템

테이프 드라이브
및 라이브러리의
성능 고려사항

본 설명서에 수록된 정보는 사전 통보 없이 변경될 수 있습니다.
© 2005 Dell Inc. All rights reserved.

어떠한 경우에도 Dell Inc.의 사전 승인 없이 무단 복제하는 행위는 엄격하게 금지되어 있습니다.

본 설명서에 사용된 상표인 *Dell*, *DELL* 로고 및 *PowerVault*는 Dell Inc.의 상표입니다.
EMC 및 *PowerPath*는 EMC Corporation의 등록 상표입니다.

본 설명서에서 특정 회사의 표시나 제품 이름을 지칭하기 위해 기타 상표나 상호를 사용할 수도 있습니다.
Dell Inc.는 자사가 소유하고 있는 것 이외에 기타 모든 등록 상표 및 상표 이름에 대한 어떠한 소유권도 보유하지 않습니다.

2005년 6월

차례

도입	131
일반적인 호스트 백업 고려사항	131
테이프 드라이브 및 데이터 고려사항	131
하드 드라이브 및 RAID 배열 구성	132
테이프 라이브러리에서 다중 드라이브 사용 시 일반적인 성능 고려사항	138
SCSI 구성	138
SAN 구성	139

도입

최근 테이프 드라이브의 전송률 향상으로 인해 호스트 서버와 테이프 드라이브에서 동일한 속도로 데이터를 처리할 수 있는지의 여부를 결정할 때 RAID 구성과 하드 드라이브 사양과 같은 많은 호스트 측 요소들이 고려되어야 합니다. 호스트 서버에서 테이프 드라이브의 처리량을 제한할 수 있는 일반적인 구성과 속성에 대해서는 "일반적인 호스트 백업 고려사항"에서 다룹니다.

여러 개의 드라이브가 테이프 라이브러리에 위치해 있으므로 다중 테이프 드라이브의 잠재적인 처리량을 수용하기 위해서는 더 큰 호스트 대역폭이 필요합니다. 권장되는 케이블 연결 구성뿐만 아니라 다중 드라이브 장치의 잠재적인 Fibre 제한은 "테이프 라이브러리에서 다중 드라이브 사용 시 일반적인 성능 고려사항"에서 논의합니다.

일반적인 호스트 백업 고려사항

여기에서 다루는 고려사항은 SCSI 및 SAN (Storage Area Network) 테이프 백업 구성 모두에 적용됩니다.

테이프 드라이브 및 데이터 고려사항

성능 평가 시 다음 사항을 고려해야 합니다.

SCSI 명령 오버헤드. SCSI 버스의 명령 오버헤드로 인해 모든 SCSI 장치가 이론상의 최대 전송 속도를 내지 못합니다. 소프트웨어 백업 소프트웨어는 이러한 오버헤드를 고려하지 않습니다. 대신 소프트웨어에서는 테이프에 대한 데이터 쓰기 속도를 측정할 뿐입니다. 예를 들어 드라이브의 데이터 처리 속도는 80MB/sec이지만 데이터 쓰기 속도는 77MB/sec가 될 수 있습니다. 이 경우 백업 소프트웨어에서 뒤쪽의 속도를 보고합니다.

테이프 블록 크기. 대부분의 테이프 드라이브의 경우 최적의 블록 크기는 64Kb입니다. 그러나 일부 백업 응용프로그램은 사용자가 블록 크기를 변경할 수 있도록 허용합니다(하지만 크기를 더 크게 해도 성능 향상에는 도움이 되지 않을 수 있습니다). 64Kb보다 작은 블록 크기를 사용할 경우 성능이 저하될 수 있습니다. 테이프 장치의 블록 크기 조정에 관한 내용은 백업 소프트웨어 사용 설명서를 참조하십시오.

백업 소프트웨어 버퍼 크기. 최적의 백업 성능을 위해 백업 소프트웨어 버퍼는 가능한 커야 합니다. 일부 응용프로그램에서는 데이터와 드라이브 간 스트림을 원활히 유지하고 (특히 소형 파일의) 전송률을 크게 높일 수 있도록 사용자가 버퍼 크기를 변경할 수 있도록 지원합니다. 버퍼 크기가 클수록 더 많은 데이터를 보유할 수 있고 디스크에서 데이터를 탐색하는 시간을 단축시킬 수 있어 메모리와 CPU 성능에 영향을 미칠 수 있습니다. 자세한 사항은 테이프 백업 응용프로그램 사용 설명서를 참조하십시오.

드라이버와 펌웨어. 항상 SCSI 또는 광 컨트롤러와 테이프 드라이브의 드라이버 및 펌웨어를 최신 상태로 유지하십시오. support.dell.com 을 방문하여 최신 Dell PowerVault 테이프 제품 드라이버 및 펌웨어를 다운로드하십시오.

테이프 드라이브와 하드 드라이브를 별도의 HBA(호스트 버스 어댑터)의 개별 컨트롤러(내부 또는 외부)에 연결하십시오. 이 사항은 컨트롤러의 성능에 다소 영향을 미치지만 가장 좋은 방법은 처리량의 극대화를 위해 테이프 드라이브 HBA를 하드 드라이브 HBA로부터 분리하는 것입니다. 대부분의 온보드 이중 모드 SCSI/RAID 컨트롤러는 하나의 프로세서를 공유하며 이로 인해 RAID 배열과 테이프 드라이브 간 대역폭을 공유해야 합니다. 따라서 하나의 컨트롤러에서 하드 디스크와 테이프 드라이브 사이의 읽기 및 쓰기를 처리하고 필요한 패리티 정보를 계산하여 하드 드라이브에 쓰는 작업을 수행합니다. RAID 배열과 패리티 바이트에 관한 자세한 정보는 "하드웨어 RAID 구성 고려사항"을 참조하십시오.

지저분한 드라이브 헤드 또는 오래된 매체. 지저분한 테이프 드라이브나 오래된 매체는 에러율을 높이고 쓰기/읽기 속도를 떨어뜨릴 수 있습니다. 드라이브에서 테이프의 블록에 다시 쓰기 또는 다시 읽기를 시도할 때마다 성능이 저하됩니다. 특정 읽기/쓰기 오류의 임계값에 도달하면 대개 드라이브에서 청소를 수행할 것을 요청합니다. 정기적으로 또는 요청 시 드라이브 헤드를 청소해야 합니다.

매체가 노후화되거나 과도하게 사용될 경우 불량 블록이 발생할 가능성이 커집니다. LTO 매체의 일반적인 수명 주기는 약 75회의 완전 테이프 쓰기/읽기입니다.

속도 일치. 신형 LTO 드라이브는 최고 비압축 데이터 전송 속도의 약 1/2 정도 낮추어, 드라이브에 제공되는 데이터 속도에 일치시킵니다. 데이터가 최저 속도 일치 한계에 비해 저속으로 드라이브에 제공될 경우 드라이브는 정지하고 버퍼가 채워져 되감기 되도록 기다립니다. 그런 다음 버퍼 쓰기를 시도합니다(이 과정을 "백 히칭(back hitching)"이라 함).

예를 들어 Dell PowerVault 110T(LTO2 및 LTO3) 테이프 드라이브는 LTO-3 매체에 쓰기 할 때 속도를 30 MB/sec로 하향 일치시킵니다. 호스트 서버에서 20MB/sec의 속도만으로 데이터를 제공할 수 있다면 드라이브는 "백 히치"되어 버퍼가 채워지도록 기다립니다. 이 상황에서 효과적인 처리량은 20MB/sec 미만(아마 15MB/sec 내외)이 될 것입니다.

테이프 드라이브의 성능 확인

일부 테이프 드라이브 제조업체는 처리량을 확인할 수 있도록 드라이브에 성능 진단 모드를 내장합니다. PowerVault 110T LTO-2 및 LTO-3(펌웨어 53XX 이상)에서는 드라이브와 매체의 빠른 읽기/쓰기 성능 테스트를 수행하는 진단 모드 "F"를 제공합니다. 성능 수치가 최대 지정 드라이브 속도의 6% 이내가 아닌 경우 테스트는 실패하고 오류 메시지를 표시합니다. 테스트에 통과하면 오류 메시지가 표시되지 않습니다. 진단 모드 "F"에 대한 자세한 내용은 테이프 드라이브 사용 설명서를 참조하십시오.

주의사항: 진단 모드 "F"를 사용하기 위해서는 진단 테스트의 일부로 안전하게 덮어쓰기가 가능한 매체가 필요합니다. 중요한 데이터가 포함된 매체를 사용하지 마십시오. 진단 테스트에 사용되는 매체에 저장된 데이터를 잃을 수도 있습니다.

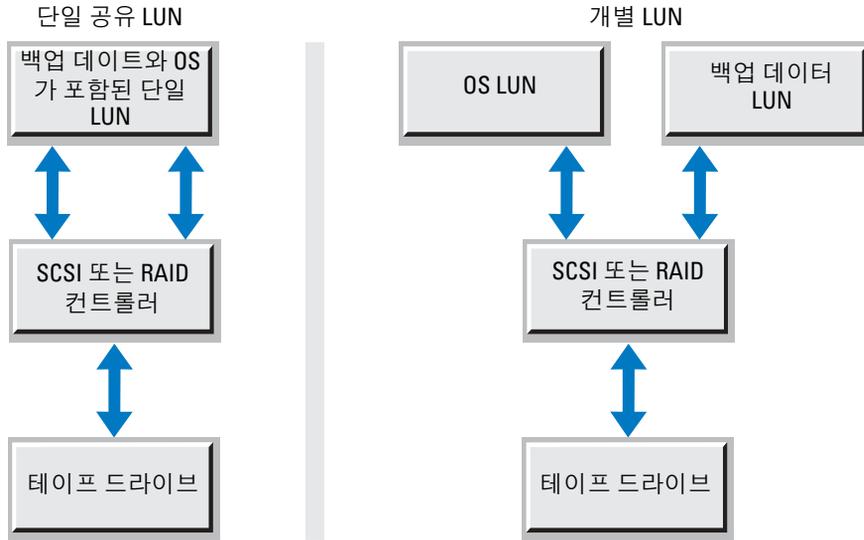
하드 드라이브 및 RAID 배열 구성

일부 하드 드라이브 및 디스크 배열(내부 및 외부 모두) 속성은 백업 또는 복원 성능에 영향을 미칠 수 있습니다. 최대 백업 및 복원 속도를 얻을 수 있도록 돕는 권장 구성과 함께 이러한 속성은 아래의 하위 절에서 다룹니다. 테이프 드라이브의 지속적인 처리량이 디스크 배열의 처리량을 초과하는 경우에는 테이프 드라이브의 최대 성능이 모두 발휘되지 않습니다.

일반적인 하드 드라이브 구성 고려사항

여러 LUN에서의 데이터/운영 체제(OS). OS LUN(Logical Unit Number)과는 별도로 LUN 상의 데이터를 백업하면 하드 드라이브에서 OS 작업과 백업 작업 간의 액세스와 오버헤드를 분할하지 않도록 합니다. 이것은 OS가 포함된 하나의 하드 드라이브 또는 디스크 배열과 백업될 데이터가 포함되는 물리적으로 분리된 하드 드라이브 또는 디스크 배열로 가능합니다.

그림 1-1. 단일 채널과 2 채널 대역폭



하드 드라이브 성능

설계상 테이프 드라이브는 순차적으로 데이터를 쓰고 드라이브 작동이 순차적으로 유지되도록(백 히칭 방지) 일정한 데이터 공급이 이루어져야 합니다. 반대로 하드 드라이브는 임의 액세스 장치입니다. 따라서 하드 드라이브는 데이터가 드라이브 플래터상에 산재해 있을 경우 순차적인 데이터를 제공하는 데 종종 어려움을 겪을 수 있습니다. 이 경우 드라이브는 끊임없이 데이터 조각들을 찾게 됩니다.

또한 다른 하드 드라이브 속성들도 테이프 드라이브의 데이터 처리량에 영향을 미칠 수 있습니다.

스핀들 속도. 보통 RPM(분당 회전수)으로 측정되는 하드 드라이브의 스핀들 속도는 드라이브 플래터 부속품이 분당 완전 회전을 수행하는 횟수를 결정합니다. 이 속성은 임의 액세스 시간과 순차 전송률 모두에 직접적인 영향을 미칩니다. 스핀들 속도가 높을수록 더 빨리 드라이브가 데이터에 액세스할 수 있습니다.

임의 액세스 시간 또는 탐색 시간. 일반적으로 밀리초로 측정되는 탐색 시간은 드라이브 헤드가 디스크상의 데이터를 찾는 데 걸리는 시간을 의미합니다. 하드 디스크의 탐색 시간은 읽기/쓰기 헤드가 플래터 표면의 트랙 사이를 이동하는 데 소요되는 시간의 양을 측정합니다. 하드 디스크는 임의 액세스 장치이기 때문에 데이터는 사실상 임의의 디스크 섹터에 저장될 수 있습니다. 데이터를 액세스하는 데 더 많은 시간이 소요되면 드라이브의 전체 처리 속도는 느려집니다. 하드 드라이브에 작은 파일이 많을 경우에 이 속성은 매우 중요해집니다. 파일 크기가 작을수록 디스크에 파일을 읽거나 쓰기 위해 드라이브를 "탐색"하는 횟수가 늘어납니다. 따라서 많은 소형 파일들을 전송할 때에는 읽기 또는 쓰기 속도가 매우 느려지는 경향이 있습니다.

STR (Sequential/Sustained Transfer Rates - 장기적 연속적으로 전송할 때의 속도). STR은 드라이브에서 실제로 얼마나 빨리 데이터를 플래터에서 읽고 플래터에 쓸 수 있는가를 측정합니다. 백업할 데이터가 인접한 단일 대형 파일일 경우 지속적인 처리량은 드라이브의 최대 STR에 근접할 것입니다. 하지만 실제 응용프로그램의 경우 데이터가 삭제되고 쓰여지기 때문에 데이터가 플래터에 산재하여 존재하게 됩니다. 하드 드라이브 조각 모음은 드라이브가 최대 STR에 근접할 수 있도록 돕습니다.

버퍼(캐쉬). 버퍼는 가장 최근에 쓰여지거나 저장된 데이터를 보유하는 드라이브상의 메모리 양입니다. 버퍼가 클수록 더 많은 데이터를 보유할 수 있고, 결과적으로 디스크상의 데이터를 찾는 데 소요되는 시간이 단축됩니다.

하드웨어 RAID 구성 고려사항

RAID의 일반 개요

이 절에서는 전형적인 RAID 구성에 대한 개요와 백업 및 복원 속도에 어떻게 영향을 미치는가에 대해 설명합니다. RAID 배열은 하드 디스크 집합으로 단일 스토리지 시스템 또는 LUN으로 작동합니다. 데이터는 동시에 각 하드 드라이브의 채널 전체에 걸쳐 전송될 수 있어 총 처리량은 배열에 있는 드라이브의 총 수를 곱한 값에다 오버헤드와 아래에서 설명하는 기타 중복을 제외한 값이 될 수 있습니다.

RAID 구성의 경우 드라이브들이 인터페이스의 대역폭을 공유하기 때문에 인터페이스 속도가 중요해집니다. 예를 들어 단일 Ultra160 드라이브는 40MB/sec만을 유지할 수 있습니다. 따라서 동일한 유형의 드라이브로 구성된 5 디스크 RAID 0 배열은 200MB/sec로 읽기/쓰기 속도를 가져야 합니다. 그러나 Ultra160 인터페이스는 최대 속도를 160 MB/sec로 제한합니다.

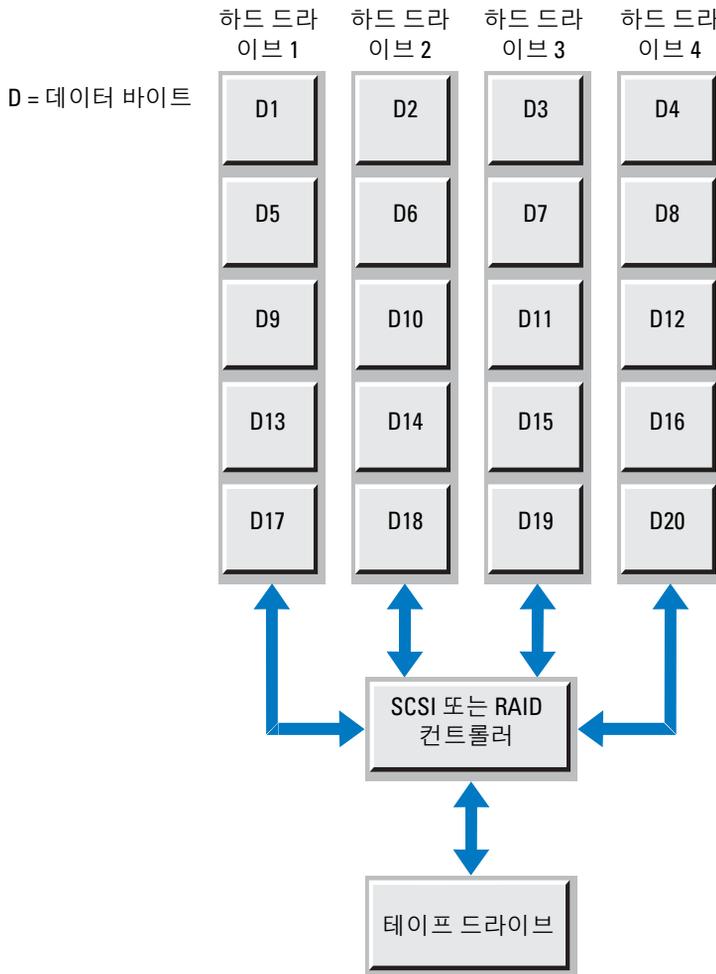
외부 디스크 배열(특히 SAN에서)은 I/O 성능을 향상시키기 위해 상당한 수준의 캐쉬 메모리를 제공할 수 있습니다. 이 캐쉬는 배열에 쓰기 작업을 수행할 때 성능을 크게 개선하며 읽기 성능 향상을 위해 자주 액세스하는 데이터를 저장할 수 있습니다. 테이프 성능에 미치는 영향과 관련하여, 배열로 데이터를 복원하거나 배열에서 데이터를 백업할 때 캐쉬는 대부분의 RAID 제한을 없애주는 역할을 합니다. 그러나 캐쉬를 사용하여 외부 배열에서 백업 작업을 수행할 때에는 여전히 디스크에서 데이터를 읽어야 하므로 RAID 구성 제한의 영향을 느낄 수도 있습니다.

RAID 0

흔히 스트라이핑이라고 하는 RAID 0은 두 개 이상의 디스크를 연결하여 단일 LUN 방식으로 하나의 가상 드라이브를 생성할 수 있습니다. RAID 0을 스트라이핑이라고 일컫는 이유는 데이터가 한번에 하나의 디스크가 아닌 배열의 모든 디스크에 걸쳐 쓰여질 수 있기 때문입니다. 따라서 처리량은 단일 하드 디스크의 단일 채널 대신에 n 채널(여기서 n 은 배열을 구성하는 하드 드라이브의 수)에 걸쳐 분산됩니다. 이로 인해 읽기/쓰기 성능이 향상되지만 결합 허용 기능은 제공하지 않습니다.

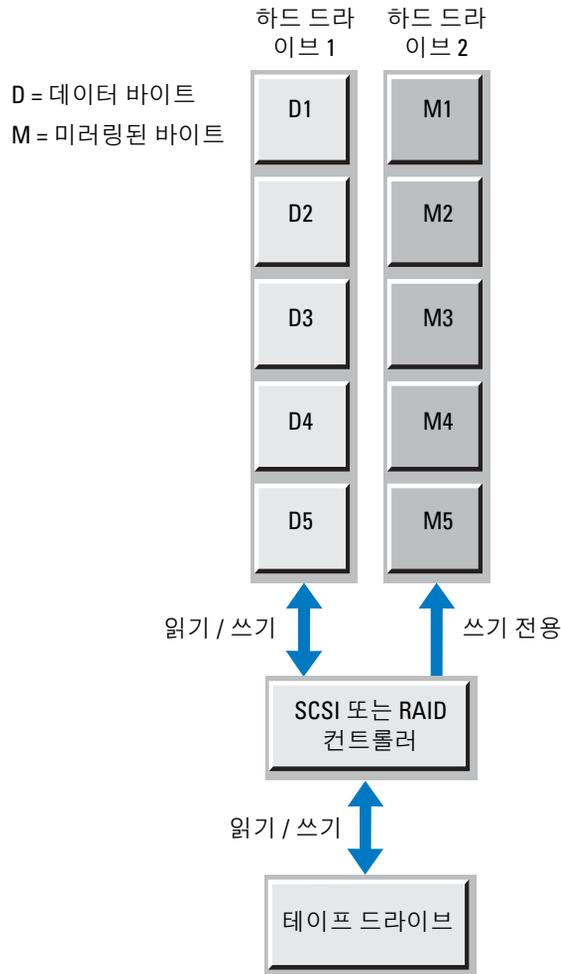
그림 1-2는 4개의 하드 드라이브로 이루어진 RAID 0 구성을 보여줍니다. 데이터는 4개의 하드 드라이브에 걸쳐 분산되어 저장되며 그로 인해 배열에 대한 읽기/쓰기가 4채널로 구현됩니다.

그림 1-2. RAID 0 구성 예



RAID 1

미러링이라고도 하는 RAID 1은 데이터가 두 개의 디스크에 동시에 두 번에 걸쳐 쓰여집니다. 하나의 드라이브에 오류가 있으면 시스템은 다른 드라이브로 전환하므로 데이터가 유실되지 않습니다. 테이프 드라이브 백업 동안 RAID 1 배열의 읽기 속도는, 주 드라이브에서 읽기를 수행하기 때문에, 단일 드라이브와 거의 동일합니다. 그러나 RAID 1 배열로 쓰기 작업을 하는 테이프 드라이브로부터의 복원 성능은 주 디스크 및 미러 디스크에 대한 쓰기 작업에 통합된 ECC(오류 확인/수정) 때문에 느려질 수 있습니다. 이러한 비효율성의 상당 부분은 미러링이 CUP나 RAID 컨트롤러에서 자주 수행되기 때문입니다. 따라서 최신 RAID 컨트롤러는 고기능의 최신 프로세서 덕분에 더욱 빨라지는 추세입니다.

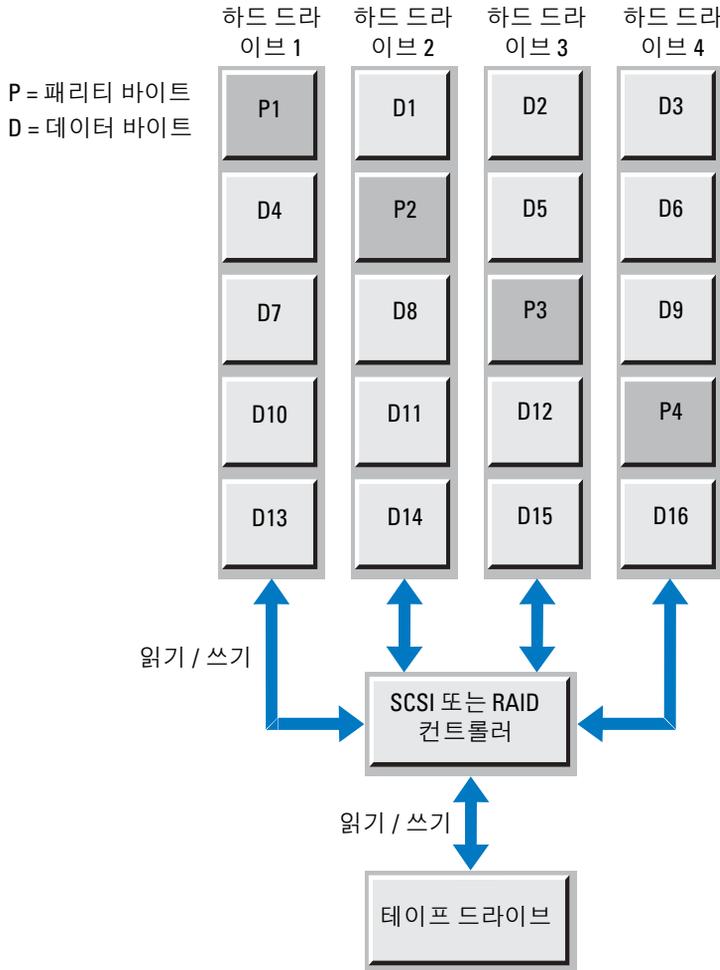
그림 1-3. RAID 1 구성 예

RAID 5

RAID 5 배열에서 데이터는 바이트 레벨에서 디스크 배열 전체에 걸쳐 저장되며 오류 수정 데이터, 즉 패리티 데이터 또한 디스크 배열 전체에 저장됩니다. RAID 5 배열은 매우 양호한 임의의 읽기 성능을 가지는 경향이 있으며 이 읽기 성능은 일반적으로 배열의 디스크 수가 증가함에 따라 향상됩니다. 규모가 더욱 큰 디스크 배열에서 읽기 성능은 RAID 0 배열보다 사실 더 뛰어난데, 이는 데이터가 추가 드라이브상에 분산되기 때문입니다. 또는 정상 읽기 중에 패리티 정보는 필요하지 않습니다.

테이프에서 RAID 5 배열로의 복원은, 패리티 정보의 계산과 쓰기 작업으로 인한 추가 오버헤드가 유발되므로, 대체로 부족합니다.

그림 1-4. RAID 5 구성 예



요약하면 RAID 0은 전체적으로 읽기 및 쓰기 성능 측면에서 최상의 구성이지만 중복성을 허용하지 않습니다. RAID 1은 전체적인 성능에서는 가장 뒤떨어지지만 배열에 기록되는 모든 데이터가 미러링되고 단일 디스크에서 읽기가 실행됩니다. RAID 5는 읽기 성능은 양호하지만 쓰기 성능은 평균 수준입니다. 그러나 배열에 더 많은 디스크가 추가될 경우 RAID 5는 향상됩니다. RAID가 상당한 수준의 캐쉬 메모리를 제공하는 인클로저 내에 위치한 경우 복원 작업 중의 성능 제한은 상당 부분 완화됩니다. 그러나 백업 작업은 계속 RAID 구성 제한의 제약을 받습니다. 또한 배열의 특성은 "하드 드라이브 성능"에 나와 있는 특정 하드 드라이브 특성에 크게 의존합니다.

테이프 라이브러리에서 다중 드라이브 사용 시 일반적인 성능 고려사항

여러 테이프 드라이브를 동시에 사용하여 데이터 백업(예: 테이프 라이브러리에서)을 수행할 경우 하드웨어 구성의 추가적인 측면을 고려해야 합니다. 하드웨어 및 케이블 연결 구성 설정 시 성능 중심의 단순한 방법을 채택함으로써 추가적인 처리량 병목 현상을 줄일 수 있습니다.

SCSI 구성

테이프 라이브러리에서 제공되는 고성능의 최신 테이프 드라이브는 SCSI 인터페이스 표준의 Ultra160 규격을 지원합니다. 그러므로 성능 최대화를 위해 SCSI를 이용하는 백업 서버에는 Ultra160 이상의 데이터 속도를 지원하는 HBA가 설치되어야 합니다. 이 요구사항을 만족하는 SCSI HBA를 사용하면 각 테이프 드라이브는 SCSI 버스에서 160MB/sec의 속도로 호스트와 통신할 수 있습니다. 테이프 드라이브 속도에 비해 SCSI 버스의 데이터 속도가 더 높으면 장치 성능을 저하시키지 않고 여러 장치를 동일한 버스에 연결할 수 있습니다. 그러나 한 지점에만 허용됩니다.

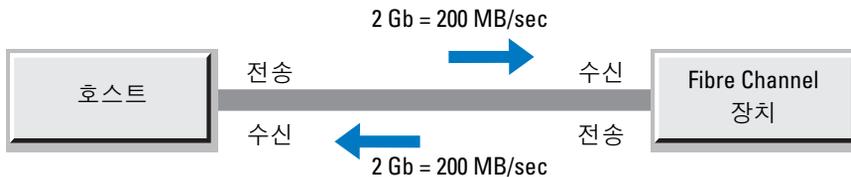
Ultra160 버스의 160MB/sec 데이터 속도가 이 버스에 연결되는 모든 장치의 최대 허용 데이터 처리량 속도입니다. 따라서 단일 테이프 드라이브는 최대 80MB/sec(기본)의 속도로 데이터를 읽고 쓸 수 있기 때문에 버스의 전체 대역폭을 소비하지 않습니다. 하지만 다중 테이프 드라이브를 버스에 의해 제공되는 전체 속도인 160MB/sec를 소비하도록 구성할 수 있는데, 이를 위해서는 각 드라이브가 각각의 최대 기본 성능으로 작동되어야 합니다. 이후에 동일한 버스에 연결되는 각각의 추가 드라이브는 각 드라이브의 평균 성능을 감소시킵니다.

따라서 테이프 라이브러리로부터 최대 성능을 이끌어내기 위해서는 각 SCSI 버스에 두 개의 테이프 드라이브만 연결할 것을 권장합니다. 자세한 내용과 해당 그림은 "권장 케이블 연결 구성"을 참조하십시오. 최소 Ultra160을 지원하는 SCSI HBA를 사용해야 하지만 테이프 드라이브의 사양이 Ultra160인 경우 Ultra320 HBA로 업그레이드하더라도 추가적인 성능 향상 효과는 얻을 수 없습니다.

SAN 구성

Fibre Channel(FC)은 SCSI에 비해 많은 이점을 제공합니다. 우선 FC는 SCSI의 거리 제약을 극복하고(LVD SCSI의 12미터 vs. 단파 2Gb FC 링크의 300미터) 고속 데이터 전송을 가능하게 합니다. SCSI와 같은 버스 기반 아키텍처가 아닌 직렬 네트워크 프로토콜로서 FC는 또한 SAN의 구현을 위한 프로토콜로 선택되어 데이터 스토리지 리소스의 통합을 허용합니다. 그리고 각 FC 연결은 전송 링크와 수신 링크로 구성되어 전이중(full-duplex) 작동이 가능합니다. 따라서 데이터는 양방향으로 동시에 전송이 가능합니다. 그러므로 단일 FC 연결을 통해 백업 작업을 수행하는 동안 데이터를 소스에서 읽어 통신 방향을 바꾸지 않고 테이프에 쓸 수 있어 연결 대역폭을 효과적으로 배가시켜 줍니다. 그림 1-5를 참조하십시오.

그림 1-5. Fibre Channel 링크 도표

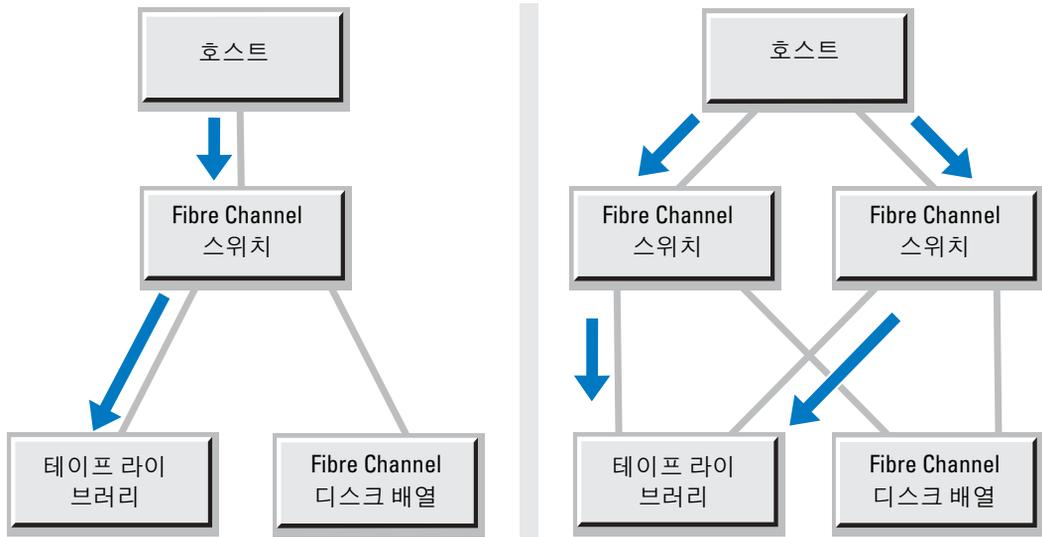


SAN에서 테이프 라이브러리를 설정할 때 성능은 다양한 요소의 영향을 받습니다. 이러한 요소에는 FC 링크 속도, 소스 및 테이프 라이브러리 간 데이터 흐름, 외부 저장 장치 배열의 성능 제한 등이 있습니다. 솔루션의 전반적인 설정 및 관리를 이해하게 되면 이러한 요소들을 상당 부분 피할 수 있습니다.

SAN에서 FC 프로토콜이 높은 데이터 대역폭을 제공하더라도 FC 링크가 성능을 제한할 수 있는 조건을 피하기 위해 테이프 드라이브에 대해 적절한 고려가 이루어져야 합니다. 2Gb FC 링크의 데이터 속도는 200MB/sec(즉, 전송 링크에서 200MB/sec, 수신 링크에서 200MB/sec)입니다. 그러므로 동일 링크로 테이프 드라이브를 작동하려고 하면 링크의 전체 대역폭을 초과할 가능성이 있습니다. 호스트가 레거시 1Gb 어댑터로 작동되는 경우 2개의 드라이브로 데이터를 백업하는 것은 현저한 성능 제한을 보여주기에 충분할 수 있습니다.

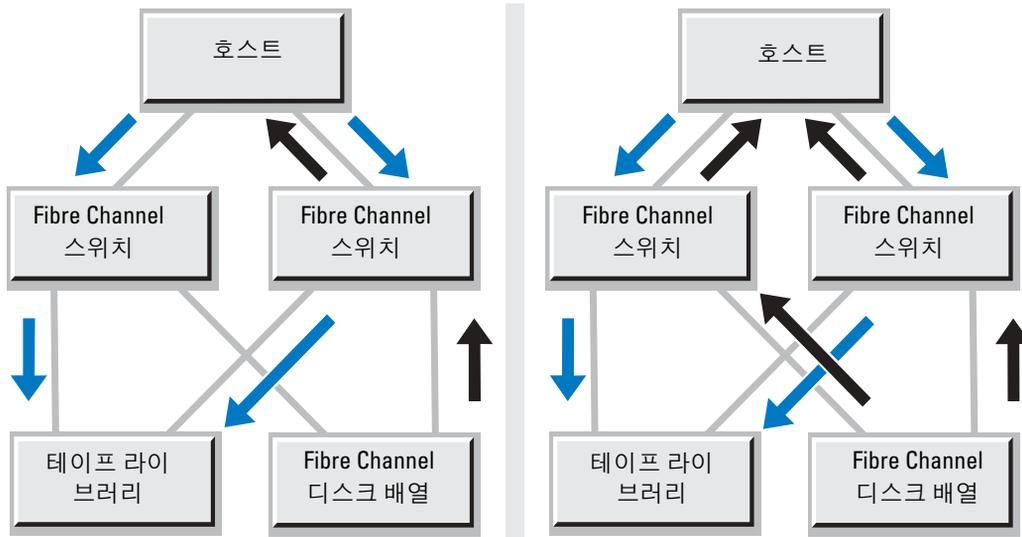
따라서 3개 이상의 테이프 드라이브를 2Gb 링크에서 동시에 사용할 때에는 단일 링크에 의존하지 말고 여러 연결로 백업을 분산해야 할 수도 있습니다. 여기서 SAN 솔루션의 토폴로지를 이해하는 것이 유익합니다. 백업 작업 중에 소스에서 읽은 다음 테이프에 쓰기를 하는 데이터 경로를 따라가다 보면 잠재적인 병목 가능성을 인식하는 데 도움이 됩니다. 병목이 확인되면 구성에 따라 조치를 취할 수 있습니다. 예를 들어 동시에 여러 드라이브가 작동되어야 하는 백업 솔루션의 경우 개별 구조에 걸쳐 테이프 하드웨어를 분할하면 연결 분할로 인해 성능이 향상될 수 있습니다. 그림 1-6을 참조하십시오.

그림 1-6. 테이프 라이브러리의 단일 및 분할 데이터 흐름



또한, 외부 FC 디스크 배열에서 읽는 데이터에 병목이 존재하는 경우, 추가적인 구조 연결과 함께 EMC® PowerPath®와 같은 I/O 관리 소프트웨어를 사용하면 다중 경로에서 데이터의 부하를 자동으로 조정하고 경로 장애 조치로 가용성을 증가시킵니다. 그림 1-7을 참조하십시오. 그림의 왼쪽은 SAN 디스크 배열을 나타냅니다. 여기에서 모든 데이터는 단일 링크를 통해서만 강제로 전송되기 때문에 테이프 하드웨어로의 데이터 흐름을 느리게 하는 병목이 발생합니다. 오른쪽 그림은 부하 조정이 데이터가 두 개의 링크를 통해 전송되도록 함으로써 배열에서 나오는 I/O 대역폭을 배가시키는 방법을 보여줍니다.

그림 1-7. 병목 상태의 데이터 흐름 vs. 부하 조정된 디스크 배열



마지막으로 SAN의 FC 디스크 배열 또한 "하드 드라이브 및 RAID 배열 구성"에서 설명하는 동일한 성능 제한 현상을 경험할 수 있습니다. 따라서 마찬가지로 디스크 배열의 성능 특성을 개선하면 SAN에서의 백업 속도에 직접적인 영향을 미칩니다.

라이브러리 Fibre Channel 브릿지를 사용하는 SAN 구성

일부 테이프 라이브러리는 Fibre Channel 브릿지 모듈이라는 방법을 통해 SAN에 연결될 수 있습니다. 이 모듈은 SCSI와 FC 프로토콜 사이의 브릿지로 작동하며 대부분의 기본 FC 라이브러리에서 제공되지 않는 추가적인 관리, 보안 및 작동 관련 기능을 제공합니다. 이러한 기능에 대한 자세한 내용은 테이프 라이브러리의 Fibre Channel 브릿지 사용 설명서를 참조하십시오.

몇몇 테이프 라이브러리 구성에서는 Fibre Channel 브릿지 모듈이 병목으로 작용하여 테이프 드라이브의 성능을 저하시킬 수 있습니다. 이러한 현상은 SCSI 및 FC 통신을 브릿지하는 데 요구되는 Fibre Channel 브릿지 모듈의 처리 용량이 특정 다중 드라이브 구성에서 제공하는 누적 데이터 처리량을 만족시키지 못하기 때문에 나타납니다. 그럼에도 불구하고 대부분의 데이터 백업 솔루션에서 Fibre Channel 브릿지 모듈은 테이프 성능의 주요 제한 요소로 작용하지 않을 것입니다. 전용 백업 서버에서는 호스트에서의 제한이 다중 테이프 드라이브로의 데이터 공급 작업에 의해 심화되는 조건이 종종 발생할 수 있습니다. 이로 인해 평균적인 드라이브 성능은 Fibre Channel 브릿지 모듈이 제한 인자로 작용할 경우의 수준보다 낮게 됩니다.

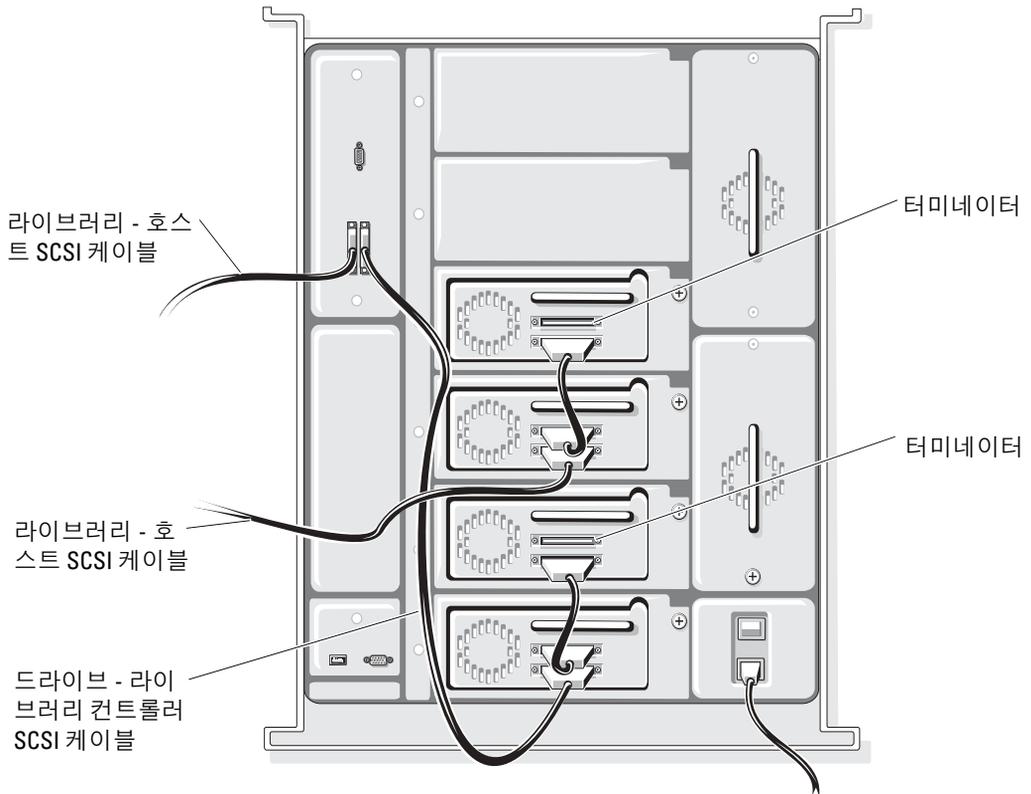
권장 케이블 연결 구성

호스트에 SCSI 케이블 연결

최대 6개의 테이프 드라이브로 구성된 테이프 라이브러리

테이프 라이브러리가 호스트에 SCSI로 연결된 경우, 단일 버스에 두 개의 드라이브만 있도록 해야 하며, 처리량을 최대화하기 위해 어떠한 SCSI 버스도 장애가 되지 않도록 5개 또는 6개 테이프 드라이브로 구성된 라이브러리에 대해 추가 SCSI 컨트롤러가 필요합니다.

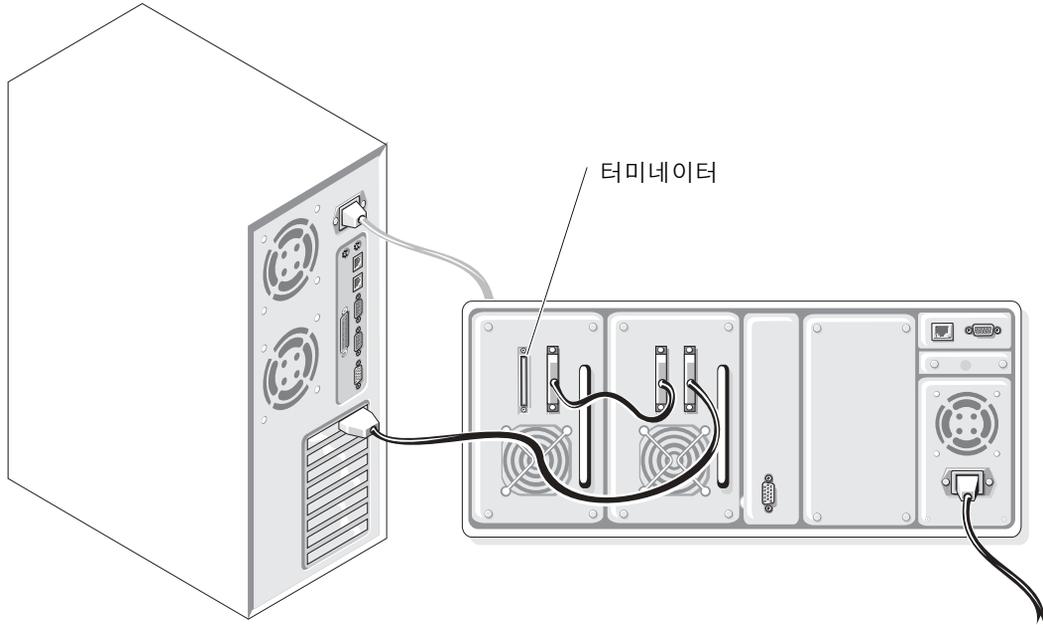
그림 1-8. 최대 6개 테이프 드라이브로 구성된 라이브러리의 SCSI 케이블 연결



최대 2개의 테이프 드라이브로 구성된 테이프 라이브러리

완전 구성된 2개의 드라이브 테이프 라이브러리의 드라이브들은 백업 성능의 심각한 제한 없이 동일 SCSI 버스의 호스트에 케이블로 연결될 수 있습니다. 단일 SCSI 버스상의 2개 드라이브의 백업 속도는 개별 버스의 2개 드라이브에 대한 백업 속도와 일치합니다. 하지만 백업 응용 프로그램의 확인 기능을 사용하는 고객은 2개의 SCSI 버스에 두 드라이브를 분할함으로써 확인 성능을 개선하기를 원할 수도 있을 것입니다. 그렇게 하면 확인 성능이 최대 25%까지 향상될 수 있습니다.

그림 1-9. 최대 2 개 테이프 드라이브로 구성된 라이브러리의 SCSI 케이블 연결



드라이브를 *Fibre Channel* 브릿지에 연결

그림 1-10부터 1-17까지는 FC 상에서의 테이프 성능을 최적화하기 위해 최대 6개의 드라이브로 구성된 테이프 라이브러리를 *Fibre Channel* 브릿지 모듈로 구성하는 방법을 보여줍니다.

그림 1-10. Fibre Channel 브릿지를 단일 테이프 드라이브에 연결

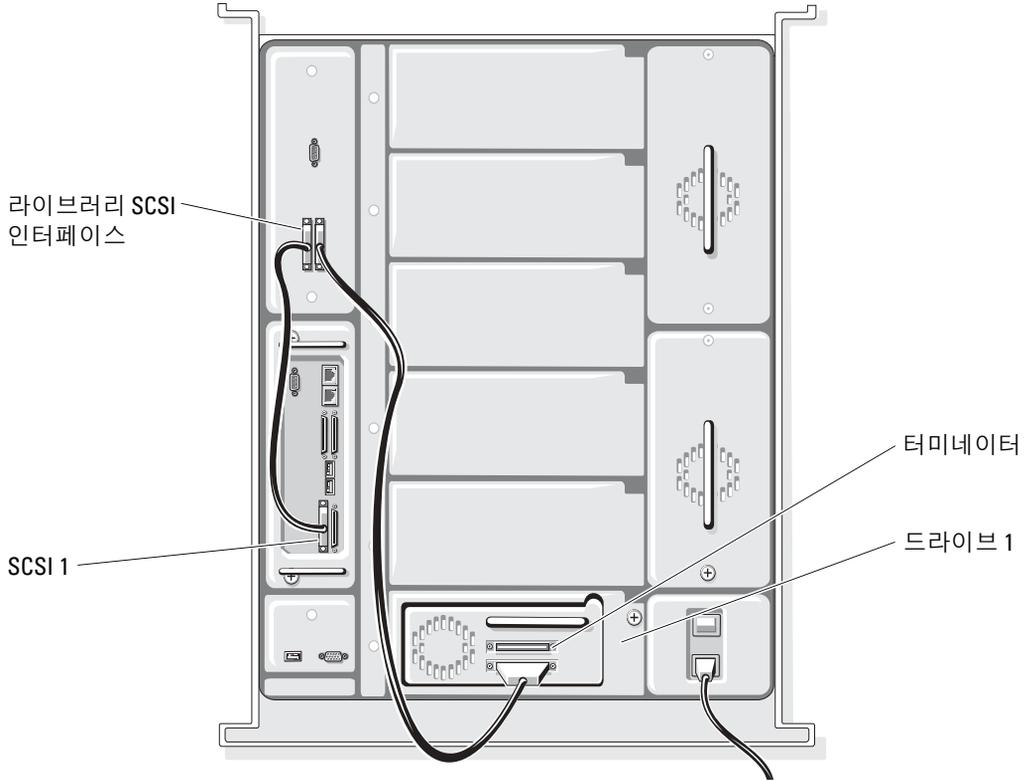


그림 1-11. Fibre Channel 브릿지를 2 개 테이프 드라이브에 연결

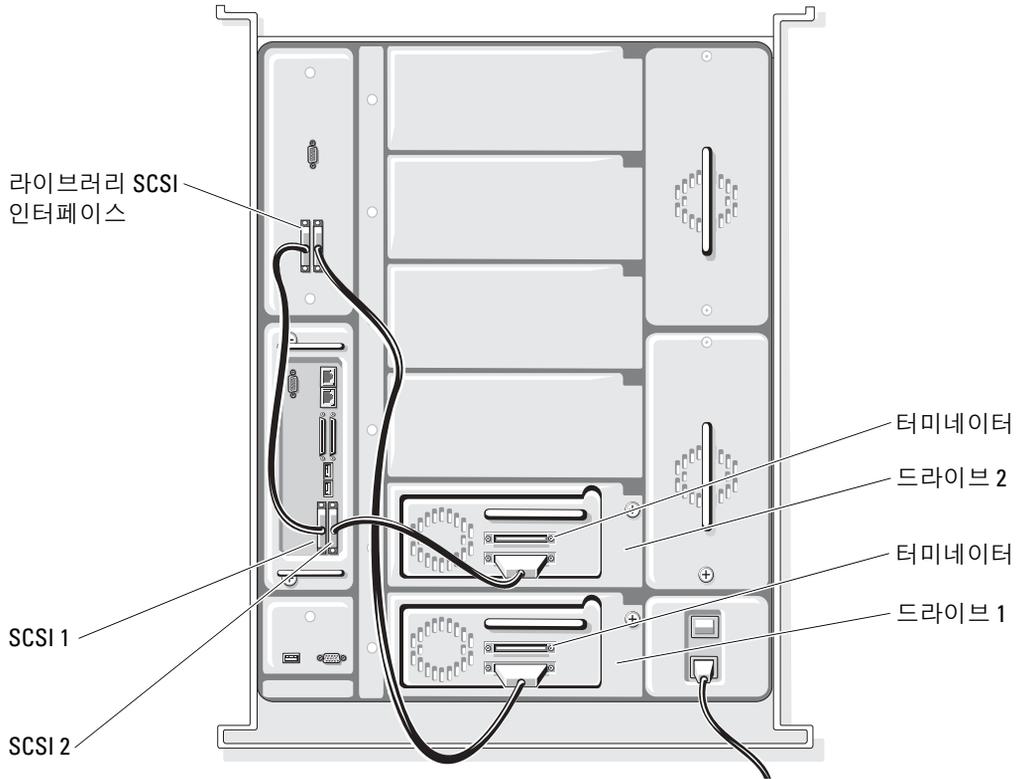


그림 1-12. Fibre Channel 브릿지를 3 개 테이프 드라이브에 연결

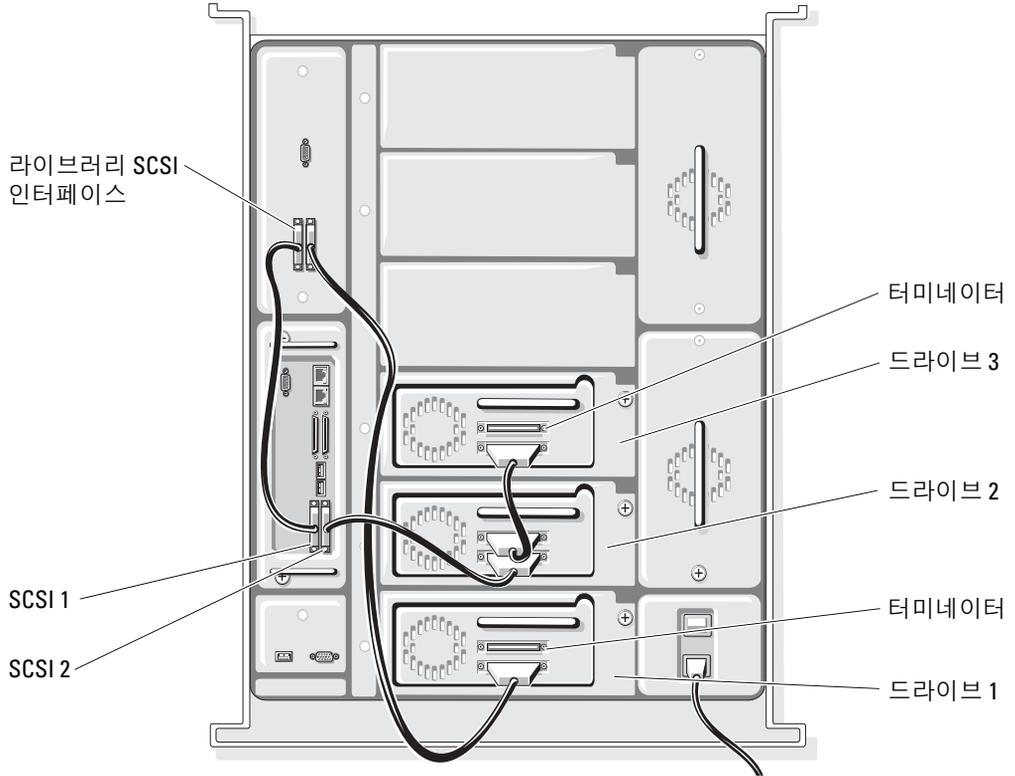


그림 1-13. Fibre Channel 브릿지를 4 개 테이프 드라이브에 연결

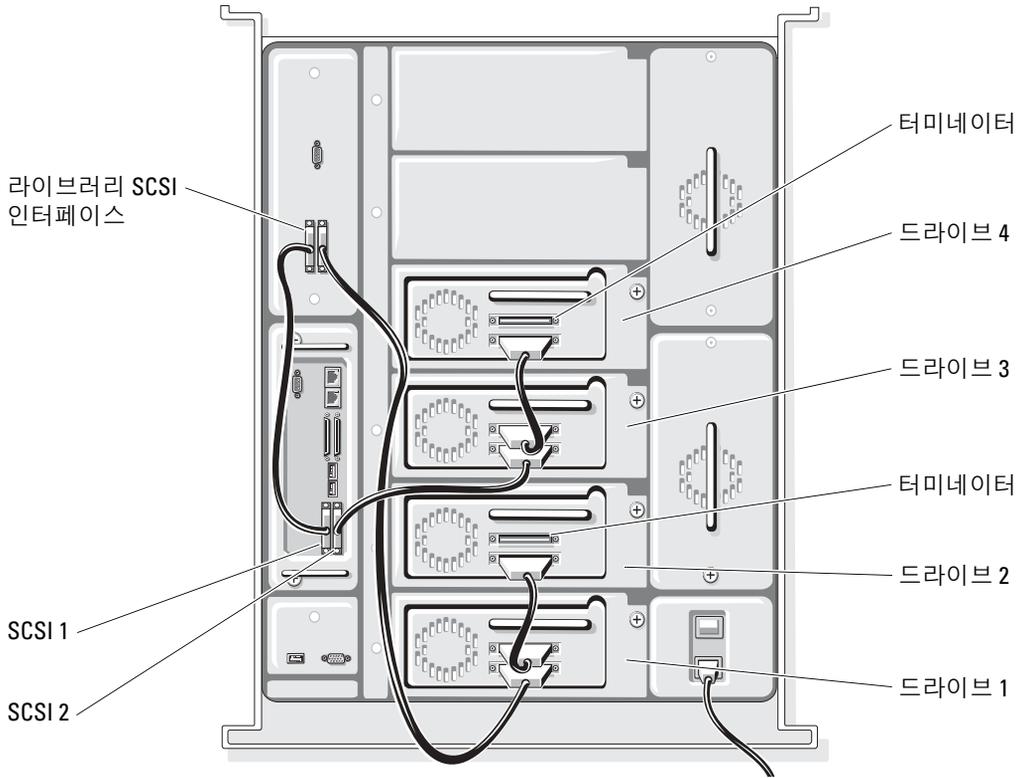


그림 1-14. 1 ~ 4 개 드라이브로 구성된 테이프 라이브러리의 Channel Zoning 설정

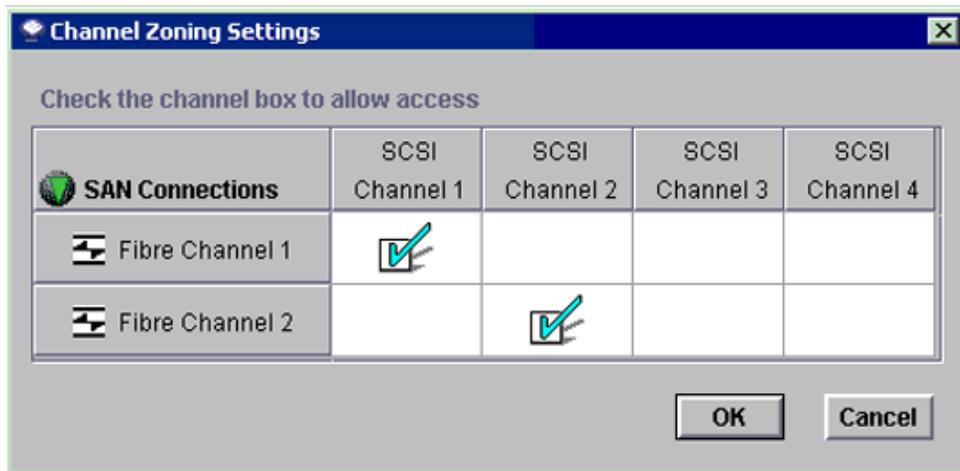


그림 1-15. Fibre Channel 브릿지를 5 개 테이프 드라이브에 연결

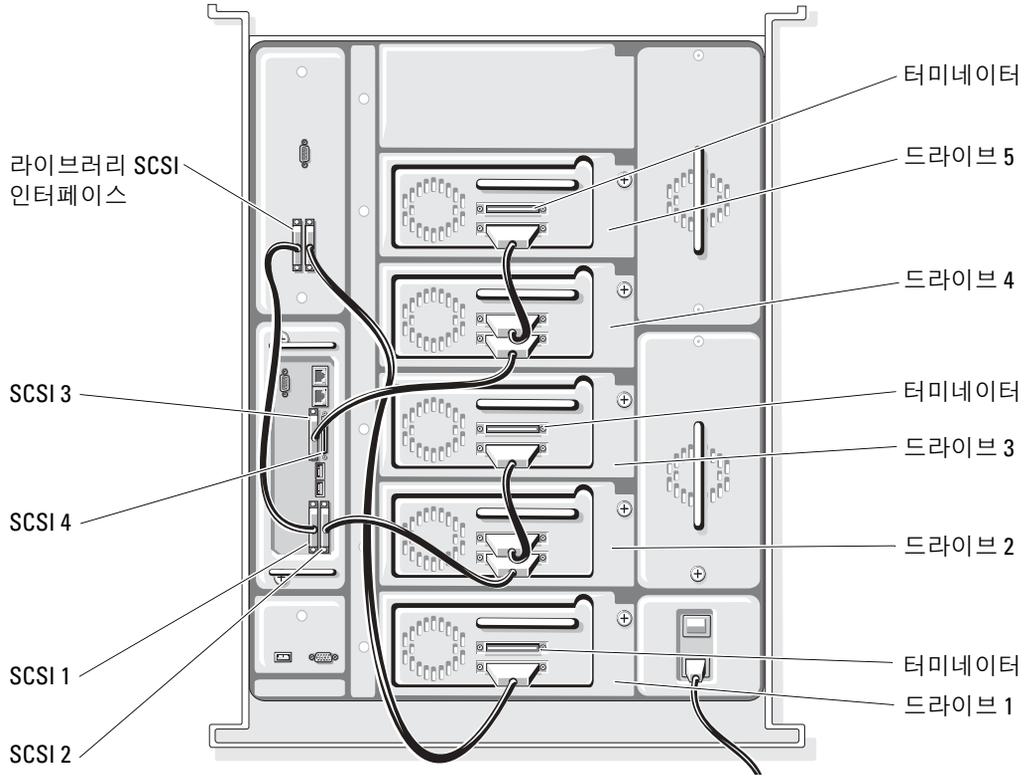


그림 1-16. Fibre Channel 브릿지를 6 개 테이프 드라이브에 연결

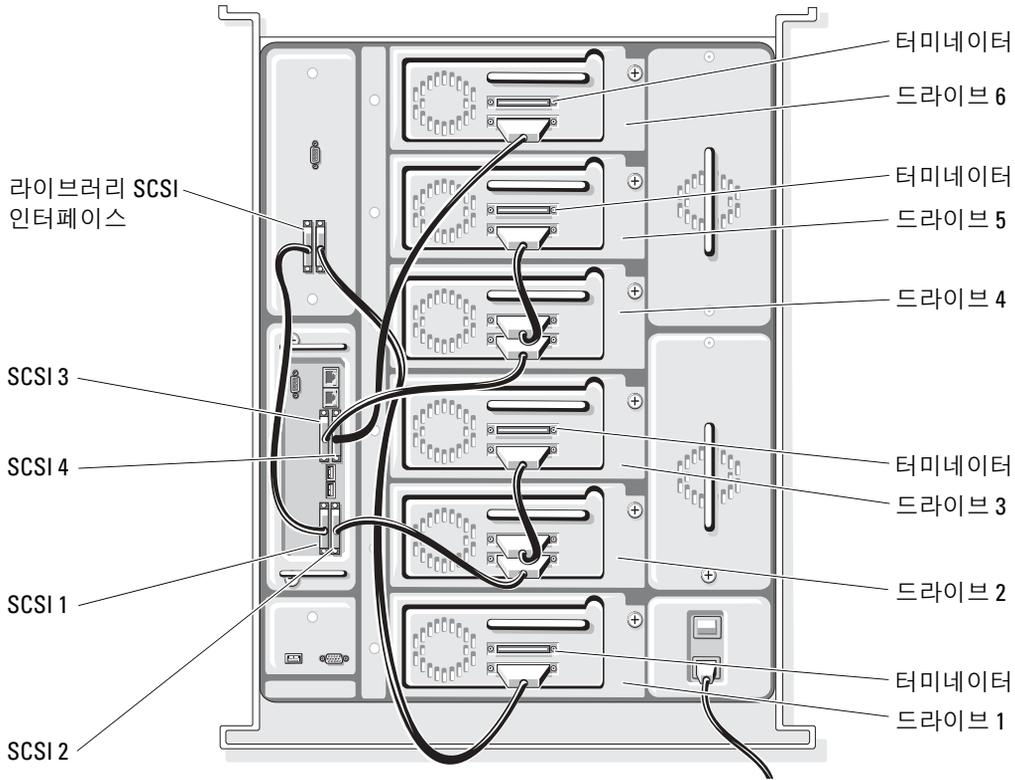


그림 1-17. 5 개 또는 6 개 드라이브로 구성된 테이프 라이브러리의 Channel Zoning 설정

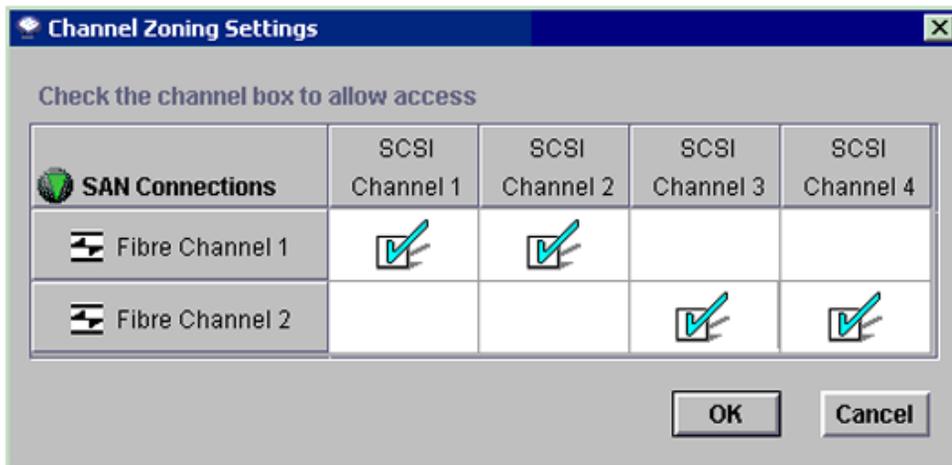


그림 1-18 및 1-19는 FC 상에서의 테이프 성능을 최적화하기 위해 최대 2개의 드라이브로 구성된 테이프 라이브러리를 Fibre Channel 브릿지 모듈로 구성하는 방법을 보여줍니다.

그림 1-18. 단일 드라이브로 구성된 테이프 라이브러리

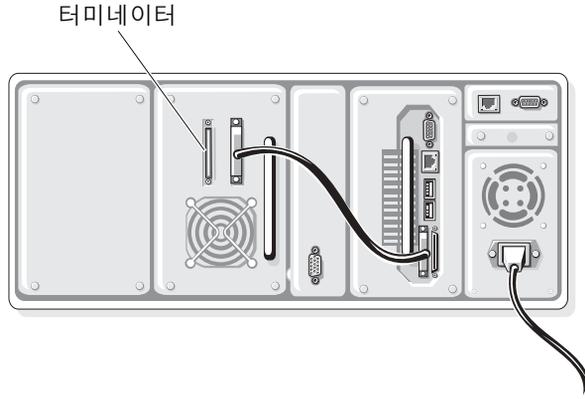
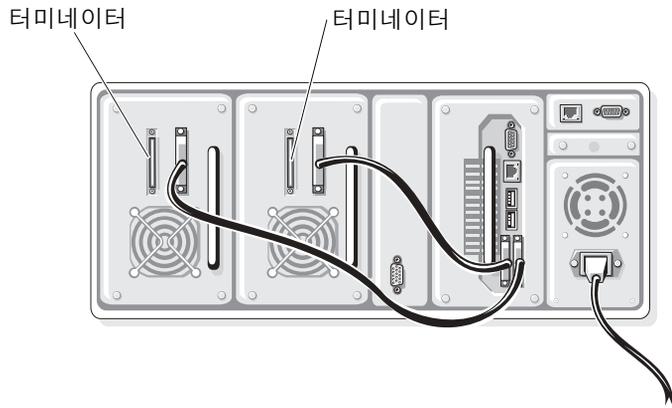


그림 1-19. 2개 드라이브로 구성된 테이프 라이브러리



Sistemas Dell™ PowerVault™

Consideraciones sobre el rendimiento de las unidades de cinta y las bibliotecas de cintas

**La información contenida en este documento puede modificarse sin previo aviso.
© 2005 Dell Inc. Reservados todos los derechos.**

Queda estrictamente prohibida la reproducción de este documento en cualquier forma sin la autorización por escrito de Dell Inc.

Marcas comerciales utilizadas en este texto: *Dell*, el logotipo de *DELL* y *PowerVault* son marcas comerciales de Dell Inc.; *EMC* y *PowerPath* son marcas comerciales registradas de EMC Corporation.

Otras marcas y otros nombres comerciales pueden utilizarse en este documento para hacer referencia a las entidades que los poseen o a sus productos. Dell Inc. renuncia a cualquier interés sobre la propiedad de marcas y nombres comerciales que no sean los suyos.

Junio de 2005

Índice

Introducción	155
Consideraciones generales sobre la copia de seguridad en el host.	155
Consideraciones sobre la unidad de cinta y los datos	155
Configuración de la unidad de disco duro y la matriz RAID	157
Consideraciones generales sobre el rendimiento al utilizar varias unidades en bibliotecas de cintas	164
Configuraciones SCSI.	164
Configuraciones de SAN	165

Introducción

Con la reciente introducción de mejoras en la velocidad de transferencia de las unidades de cinta, ahora es importante tener en cuenta algunos factores del lado del host, como la configuración RAID (matriz redundante de discos independientes [o económicos]) y las especificaciones de las unidades de disco duro, para determinar si el servidor host y la unidad de cinta pueden procesar datos a la misma velocidad. En “Consideraciones generales sobre la copia de seguridad en el host” se describen configuraciones y atributos generales que pueden limitar el rendimiento del servidor host a la unidad de cinta.

Cuantas más unidades se colocan en bibliotecas de cintas, mayores amplitudes de banda necesita el host para mantener un ritmo acorde al rendimiento potencial de varias unidades de cinta. En “Consideraciones generales sobre el rendimiento al utilizar varias unidades en bibliotecas de cintas” se describen las limitaciones potenciales de fibra para sistemas de varias unidades y las configuraciones de cableado recomendadas.

Consideraciones generales sobre la copia de seguridad en el host

Las consideraciones de esta sección se aplican a las configuraciones de copia de seguridad en cinta SCSI y SAN (red de área de almacenamiento).

Consideraciones sobre la unidad de cinta y los datos

Al evaluar el rendimiento, deben tenerse en cuenta las cuestiones siguientes:

Sobrecarga de comandos SCSI. La sobrecarga de comandos en el bus SCSI limita las posibilidades de *todos* los dispositivos SCSI de conseguir las velocidades de transferencia máximas teóricas. El software de copia de seguridad en cinta no tiene en cuenta esta sobrecarga, ya que sólo mide la velocidad a la que se graban los datos en la cinta. Por ejemplo, la unidad puede estar procesando datos a 80 MB/s, pero grabándolos a sólo 77 MB/s. El software de copia de seguridad notificará este último valor.

Tamaños de bloque en cinta. Los tamaños de bloque de 64 Kb son ideales para la mayoría de las unidades de cinta. Sin embargo, algunas aplicaciones de copia de seguridad permiten al usuario cambiar el tamaño de bloque, aunque uno mayor no ayude a mejorar el rendimiento. Utilizar tamaños de bloque inferiores a 64 Kb puede afectar negativamente al rendimiento. Consulte la Guía del usuario del software de copia de seguridad para obtener información sobre cómo ajustar el tamaño de bloque de la unidad de cinta.

Tamaño del búfer del software de copia de seguridad. Para obtener un rendimiento óptimo al realizar copias de seguridad, los búferes del software utilizado deben ser lo más grandes posible. Algunas aplicaciones permiten a los usuarios cambiar el tamaño del búfer, lo cual puede ayudar a mantener un flujo de datos continuo desde la unidad y hasta la unidad, y aumentar significativamente las velocidades de transferencia, especialmente en archivos pequeños.

Cuanto mayor sea el búfer, más datos podrá contener y menos tiempo empleará el disco en buscar los datos, aunque ello puede afectar a la memoria y al rendimiento de la CPU. Consulte la Guía del usuario de la aplicación de copia de seguridad en cinta para obtener detalles específicos.

Controladores y firmware. Asegúrese de que la unidad de cinta y la controladora SCSI o de fibra tengan siempre instalados los controladores y el firmware más recientes. Vaya a support.dell.com para descargar los controladores y el firmware más recientes para su producto de cinta Dell PowerVault.

Conecte las unidades de cinta y las unidades de disco duro en controladoras distintas (internas o externas) de adaptadores de bus de host (HBA) distintos. Aunque esto depende en cierta medida de las capacidades de rendimiento de la controladora, se recomienda mantener el HBA de la unidad de cinta separada del HBA de la unidad de disco duro para garantizar el máximo rendimiento. La mayoría de las controladoras SCSI/RAID de modo dual integradas comparten un mismo procesador, que debe compartir la amplitud de banda entre la matriz RAID y la unidad de cinta. De este modo, un controlador maneja las operaciones de lectura y escritura entre las unidades de disco duro y la unidad de cinta, y también calcula y graba la información de paridad necesaria en las unidades de disco duro. Consulte “Consideraciones sobre la configuración RAID por hardware” para obtener información específica sobre las matrices RAID y los bytes de paridad.

Cabezales de unidad sucios o soportes antiguos. Un cabezal de unidad de cinta sucio o unos soportes antiguos pueden provocar tasas de errores puede elevadas y la disminución subsiguiente de las velocidades de lectura/escritura. Cada vez que una unidad intenta volver a grabar o leer en un bloque de una cinta, el rendimiento disminuye. Cuando se alcanza un determinado umbral de errores de lectura/escritura, normalmente la unidad solicita que se efectúe una limpieza. Es importante limpiar los cabezales de la unidad regularmente o cuando se solicite.

La posibilidad de encontrar un bloque defectuoso aumenta cuanto más antigüedad tiene o cuanto más se utiliza un soporte. El ciclo de vida normal aproximado de los soportes LTO es de 75 lecturas/escrituras completas de cinta.

Coincidencia de velocidad. Las nuevas unidades LTO se adaptarán a la velocidad de los datos que se proporcionan a la unidad, hasta aproximadamente la mitad de la velocidad máxima de transferencia de datos descomprimidos. Si se proporcionan datos a la unidad por debajo del límite inferior de coincidencia de velocidad, la unidad debe parar, esperar a que el búfer se llene, rebobinar e intentar grabar en el búfer (esto se conoce como “back hitching” o demora en la grabación).

Por ejemplo, la unidad de cinta Dell PowerVault 110T (LTO2 y LTO3) adapta su velocidad hasta 30 MB/s mientras graba en soportes LTO-3. Si el servidor host sólo puede proporcionar datos a 20 MB/s, la unidad irá retrocediendo y avanzando mientras espera a que se llenen los búfers. En esta situación, el rendimiento efectivo será algo menor que 20 MB/s (probablemente cercana a los 15 MB/s).

Confirmación del rendimiento de la unidad de cinta

Algunos fabricantes de unidades de cinta incorporan un modo de diagnóstico del rendimiento en las unidades que puede utilizarse para confirmar su rendimiento. La unidad PowerVault 110T LTO-2 y LTO-3 (firmware 53XX o posterior) incluye un modo de diagnóstico “F” que realiza una prueba rápida de rendimiento de lectura/escritura en la unidad y en el soporte. Si la tasa de rendimiento no se encuentra dentro del 6 por ciento de la velocidad máxima para la unidad especificada, la prueba falla con un mensaje de error. Si la prueba es correcta, no se muestra ningún mensaje de error. Consulte el Manual del usuario de la unidad de cinta para obtener detalles específicos sobre el modo de diagnóstico “F”.



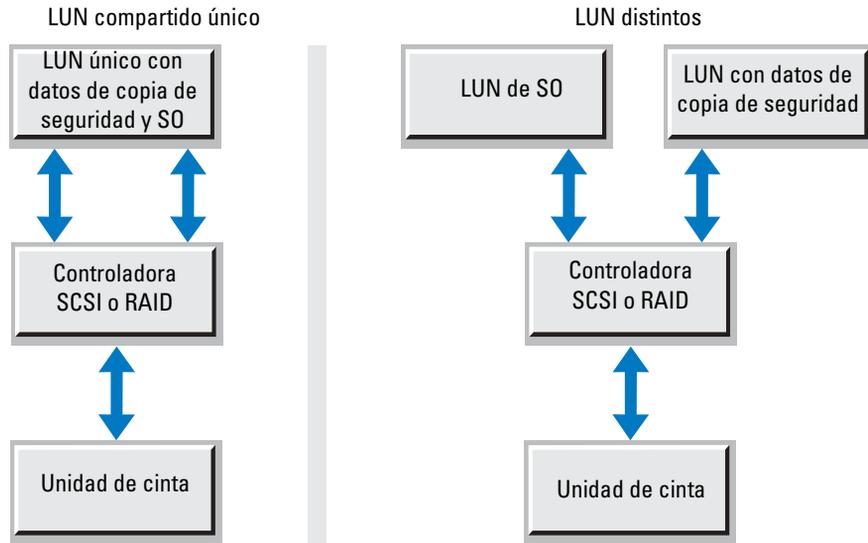
AVISO: el modo de diagnóstico “F” requiere utilizar soportes que puedan sobrescribirse de forma segura como parte de la prueba. No utilice soportes que contengan datos importantes. Los datos que contenga el soporte que se utiliza para la prueba de diagnóstico se perderán.

Configuración de la unidad de disco duro y la matriz RAID

Hay varios atributos de la unidad de disco duro y la matriz de disco (interna o externa) que pueden afectar al rendimiento de la copia de seguridad o la restauración. En los apartados siguientes se describen estos atributos, así como las configuraciones recomendadas para alcanzar las velocidades máximas de copia de seguridad y restauración. Si el rendimiento sostenido de la unidad de cinta excede el de la matriz de disco, nunca se alcanzará el rendimiento máximo de la unidad de cinta.

Consideraciones generales sobre la configuración de unidades de disco duro

Datos y sistemas operativos (SO) en LUN distintos. La copia de seguridad de datos en un número de unidad lógica (LUN) distinto del LUN del SO asegura que la unidad de disco duro no esté dividiendo el acceso y la sobrecarga entre las operaciones del SO y las operaciones de la copia de seguridad. Esto puede conseguirse haciendo que una unidad de disco duro o matriz de discos contenga el SO y que otra unidad de disco duro o matriz de discos separada físicamente contenga los datos para la copia de seguridad.

Figura 1-1. Amplitud de banda de un canal frente a dos canales**Rendimiento de la unidad de disco duro**

Por diseño, las unidades de cinta graban datos secuencialmente y requieren un suministro constante de datos para mantener el funcionamiento secuencial (y evitar así la demora en la grabación). En cambio, las unidades de disco duro son dispositivos de acceso aleatorio, por lo que a veces pueden tener problemas para suministrar datos secuenciales a las unidades de cinta si éstos se encuentran dispersos en el soporte físico de la unidad. Esto fuerza a la unidad a buscar continuamente bloques pequeños de datos.

También hay otros atributos de la unidad de disco duro que pueden afectar a la velocidad de transferencia de datos a la unidad de cinta.

Velocidad de rotación. La velocidad de rotación de la unidad de disco duro, medida generalmente en RPM (revoluciones por minuto), determina las veces por minuto que el conjunto del soporte físico de la unidad puede efectuar un giro completo. Esto tiene un efecto directo en los tiempos de acceso aleatorio y en la velocidad de transferencia secuencial. Cuanto mayor es la velocidad de rotación, más rápidamente puede acceder a los datos la unidad.

Tiempo de acceso aleatorio o tiempo de búsqueda. El tiempo de búsqueda, medido generalmente en milisegundos, es el tiempo que tardan los cabezales de la unidad en encontrar un dato que está en el disco. El tiempo de búsqueda de una unidad de disco duro mide el tiempo necesario para que los cabezales de lectura/escritura se muevan entre las pistas en el soporte físico del disco. Como los discos duros son dispositivos de acceso aleatorio, los datos pueden almacenarse en prácticamente cualquier sector del disco. Cuanto más se tarda en acceder a los datos, más lento es el rendimiento general de la unidad.

Este atributo es muy significativo cuando una unidad de disco duro contiene muchos archivos pequeños. Cuanto menores son los archivos, más “búsquedas” debe hacer la unidad para leer o grabar el archivo en el disco; por lo tanto, los discos suelen leer o grabar muy lentamente cuando se están transfiriendo muchos archivos pequeños.

Velocidad de transferencia sostenida/constante (STR). STR mide la velocidad real a la que una unidad lee o graba datos en el soporte físico del disco. Si los datos para la copia de seguridad constan de un solo archivo grande contiguo, la velocidad sostenida será cercana a la STR máxima de la unidad. Sin embargo, en las aplicaciones reales, datos están más distribuidos por todo el soporte físico a medida que se van eliminando y grabando. La desfragmentación de una unidad de disco duro puede ayudar a que la unidad alcance su STR máxima.

Búfer (caché). El búfer es la parte de la memoria de la unidad que contiene los datos que se han grabado o almacenado más recientemente. Cuanto mayor sea el búfer, más datos podrá contener y menos tiempo necesitará para buscar datos en el disco.

Consideraciones sobre la configuración RAID por hardware

Información general de RAID

En esta sección se ofrece información general sobre las configuraciones RAID típicas y se describe su efecto en las velocidades de copia de seguridad y de restauración. Una matriz RAID es un conjunto de discos duros que actúan como un solo sistema de almacenamiento o LUN. Los datos pueden transferirse simultáneamente a través del canal de cada unidad de disco duro, lo que permite que el rendimiento total sea un múltiplo del número total de unidades de la matriz, menos la sobrecarga y la redundancia, como se describe en las secciones siguientes.

En el caso de una configuración RAID, la velocidad de la interfaz es importante porque las unidades comparten la amplitud de banda de la interfaz. Por ejemplo, una única unidad Ultra160 sólo puede mantener una velocidad de 40 MB/s. Por lo tanto, una matriz RAID 0 de cinco discos compuesta por el mismo tipo de unidad debería poder leer/grabar a 200 MB/s. Ahora bien, la interfaz Ultra160 limitará la matriz a una velocidad máxima de 160 MB/s.

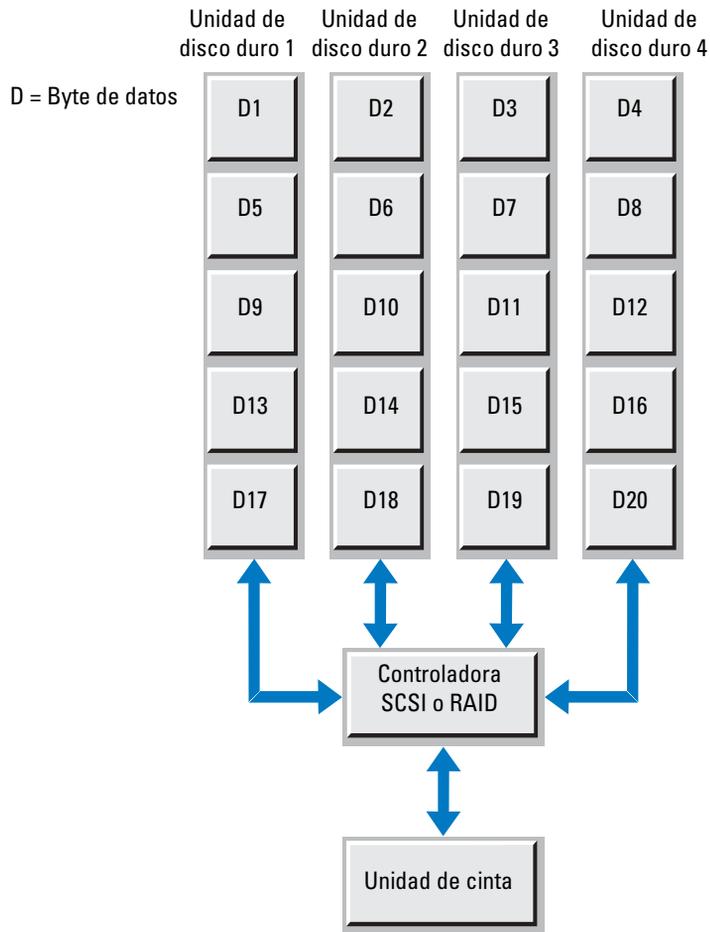
Las matrices de disco externas, especialmente en las SAN, pueden ofrecer niveles significativos de memoria caché para mejorar el rendimiento de E/S. Esta caché mejorará en gran medida el rendimiento al grabar en la matriz y puede almacenar datos a los que se accede con frecuencia a fin de mejorar el rendimiento de lectura. En cuanto al impacto en el rendimiento de la cinta, la caché paliará la mayoría de las limitaciones de RAID al restaurar datos en la matriz o al efectuar copias de seguridad de datos desde la matriz. Sin embargo, las operaciones de copia de seguridad desde matrices externas con caché pueden seguir viéndose afectadas por las limitaciones de la configuración RAID debido a que sigue siendo necesario leer los datos desde los discos.

RAID 0

RAID 0, conocida también como configuración por bandas, permite que dos o varios discos se unan para crear una unidad virtual al estilo de una única LUN. Se denomina configuración por bandas porque los datos se graban en todos los discos de la matriz, no sólo en uno cada vez. De esta manera, el rendimiento se reparte entre n canales (siendo n el número de unidades de disco duro de la matriz) en lugar de un solo canal para un solo disco duro. Con ello se consigue un excelente rendimiento de lectura/escritura, pero sin tolerancia a fallos.

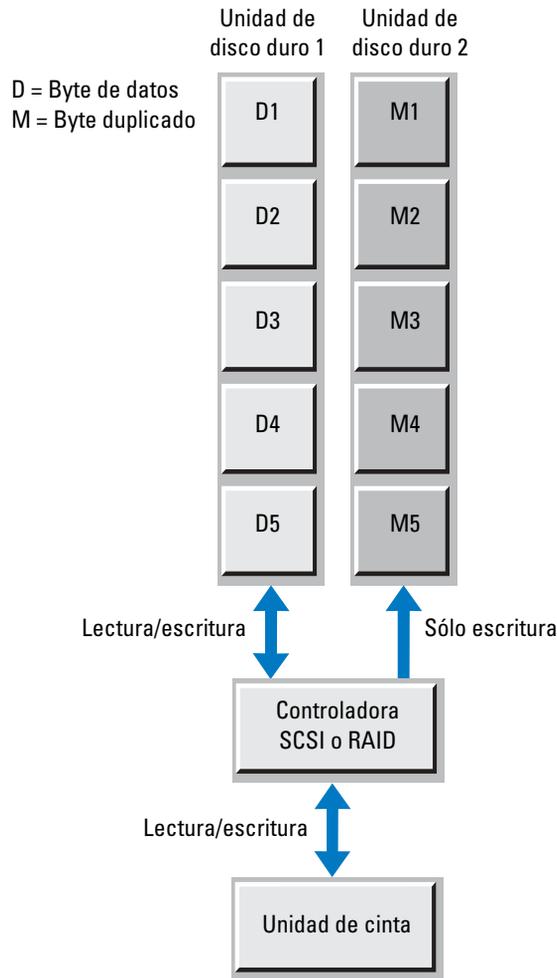
En la figura 1-2 se muestra una configuración RAID 0 con cuatro unidades de disco duro. Los datos se dividen entre las cuatro unidades de disco duro, con lo que hay cuatro canales para la lectura y la escritura en la matriz.

Figura 1-2. Ejemplo de configuración RAID 0



RAID 1

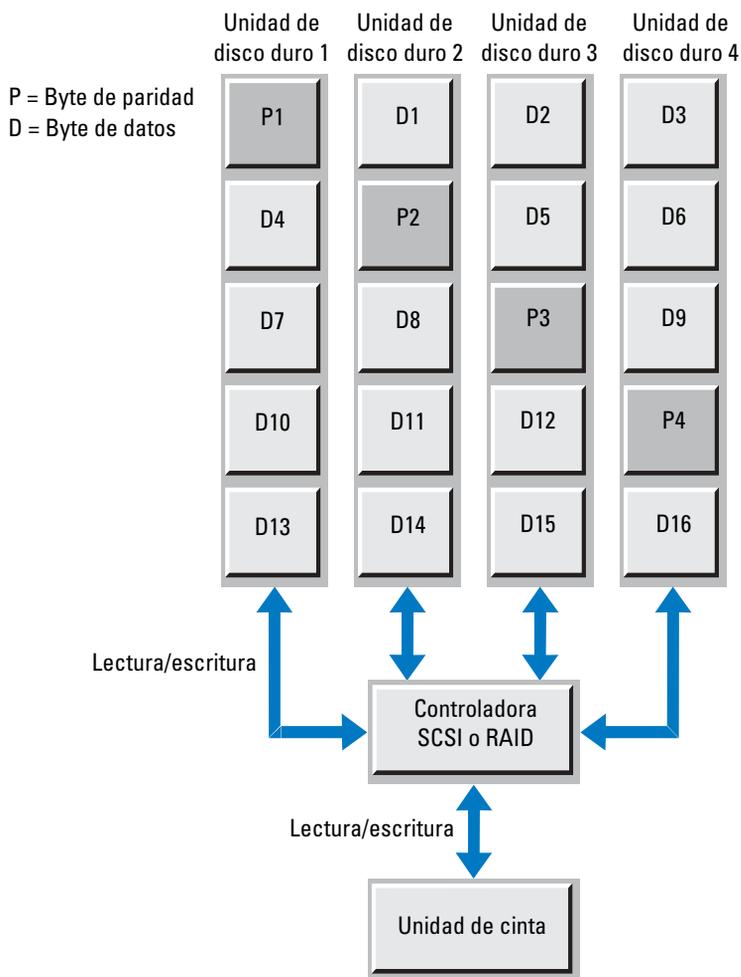
RAID 1, conocida también como duplicación, consiste en que los datos se graban dos veces en dos discos simultáneamente. Si una unidad falla, el sistema cambia a la otra unidad sin que se pierdan datos. Durante las copias de seguridad en unidades de cinta, la velocidad de lectura desde la matriz RAID 1 es aproximadamente la misma que la de una única unidad, porque la lectura se realiza desde la unidad principal. Sin embargo, el rendimiento de la restauración desde la unidad de cinta que graba en la matriz RAID 1 puede ser más lento debido a la verificación y corrección de errores (ECC) que se realiza al grabar en los discos principales y duplicados. Esta ineficiencia se debe en gran medida a que la duplicación suele llevarse a cabo en la CPU o en la controladora RAID. Por esto, las controladoras RAID más nuevas suelen ser más rápidas debido a que emplean procesadores más nuevos y con más funciones.

Figura 1-3. Ejemplo de configuración RAID 1**RAID 5**

En una matriz RAID 5, los datos se configuran por bandas en la matriz de discos en el nivel de byte y los datos de corrección de errores, o datos de paridad, también se configuran por bandas en la matriz de discos. Las matrices RAID 5 presentan generalmente un rendimiento de lectura aleatoria muy bueno, que suele mejorar a medida que aumenta el número de discos que componen la matriz. Con estas matrices de discos más grandes, el rendimiento de lectura puede superar el de las matrices RAID 0 porque los datos se distribuyen en una unidad adicional. Además, la información de paridad no es necesaria durante las lecturas normales.

Las restauraciones de una cinta a una matriz RAID 5 tienden a ser nominales, porque implican una sobrecarga adicional para calcular y grabar la información de paridad.

Figura 1-4. Ejemplo de configuración RAID 5



En resumen, RAID 0 tiende a ser la mejor configuración global en cuanto al rendimiento de lectura y escritura, pero no permite la redundancia. RAID 1 es la que ofrece un rendimiento global más pobre, ya que todos los datos que se graban en la matriz se duplican y las lecturas provienen de un solo disco. RAID 5 suele ofrecer un buen rendimiento de lectura, pero el de escritura es mediano, aunque mejora si se añaden más discos a la matriz. Si la RAID se encuentra en una carcasa con una cantidad importante de memoria caché, las limitaciones de rendimiento durante las operaciones de restauración pueden verse reducidas.

Las operaciones de copia de seguridad seguirán estando sujetas a las limitaciones de la configuración RAID. Además, las características de la matriz seguirán dependiendo en gran medida de las características específicas de las unidades de disco duro que se indican en “Rendimiento de la unidad de disco duro”.

Consideraciones generales sobre el rendimiento al utilizar varias unidades en bibliotecas de cintas

Cuando se utilizan varias unidades de cinta simultáneamente para realizar copias de seguridad de datos (como en el caso de las bibliotecas de cintas), deben tenerse en cuenta otros aspectos de la configuración del hardware. Si se utilizan métodos sencillos para agilizar el rendimiento en las configuraciones de hardware y de cableado, puede paliarse el efecto de otros cuellos de botella.

Configuraciones SCSI

Las unidades de cinta de alto rendimiento más modernas que se incluyen en las bibliotecas de cintas cumplen la especificación Ultra160 del estándar de interfaz SCSI. Por lo tanto, para conseguir el máximo rendimiento, los servidores de copia de seguridad que utilizan SCSI deben tener instalado un HBA que admita velocidades de datos de Ultra160 o superiores. Un HBA SCSI que cumpla este requisito permitirá que todas las unidades de cinta se comuniquen con el host a una velocidad de 160 MB/s en el bus SCSI. La mayor velocidad de datos del bus SCSI comparada con las velocidades de la unidad de cinta permite conectar varios dispositivos al mismo bus sin sacrificar el rendimiento de la unidad, aunque sólo hasta cierto punto.

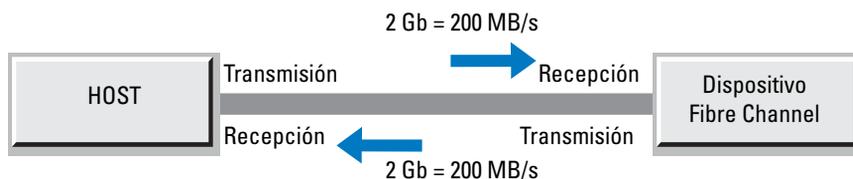
La velocidad de datos de 160 MB/s de un bus Ultra160 es el rendimiento máximo posible para todos los dispositivos conectados al bus. Por lo tanto, una sola unidad de cinta no consumirá toda la amplitud de banda del bus, ya que puede leer o grabar datos en cintas a una velocidad de hasta 80 MB/sec (nativa). Sin embargo, varias unidades de cinta pueden consumir conjuntamente los 160 MB/s que ofrece el bus si todas funcionan a su rendimiento máximo nativo. Cada unidad adicional que se conecte al mismo bus a partir de entonces provocará la reducción del rendimiento medio de las unidades.

Por consiguiente, para obtener el rendimiento máximo de una biblioteca de cintas, se recomienda no conectar más de dos unidades de cinta a cada bus SCSI. Consulte “Configuraciones de cableado recomendadas” para obtener detalles específicos e ilustraciones. Debe utilizarse un HBA SCSI que admita al menos Ultra160, pero la actualización a un HBA Ultra320 no producirá mejoras en el rendimiento si la especificación de la unidad de cinta es Ultra160.

Configuraciones de SAN

Fibre Channel (FC) presenta muchas ventajas respecto a SCSI. En primer lugar, supera las limitaciones de distancia de SCSI (12 metros para SCSI LVD frente a los 300 metros para un enlace FC de onda corta de 2 Gb) y permite la transmisión de datos a mayor velocidad. A diferencia de SCSI, que es una arquitectura basada en bus, FC es un protocolo de red serie que, además, se ha convertido en el favorito para las implementaciones de SAN, lo que permite la consolidación de los recursos de almacenamiento de datos. Asimismo, cada conexión FC está compuesta por un enlace de transmisión y otro de recepción, lo que permite un funcionamiento en modo de dúplex completo. Esto significa que los datos pueden transmitirse en dos direcciones simultáneamente. Por lo tanto, durante una operación de copia de seguridad en una sola conexión FC, los datos pueden leerse de un origen y grabarse en una cinta sin tener que realizar la comunicación por turnos, lo que permite doblar la amplitud de banda de una conexión. Consulte la figura 1-5.

Figura 1-5. Diagrama del enlace Fibre Channel



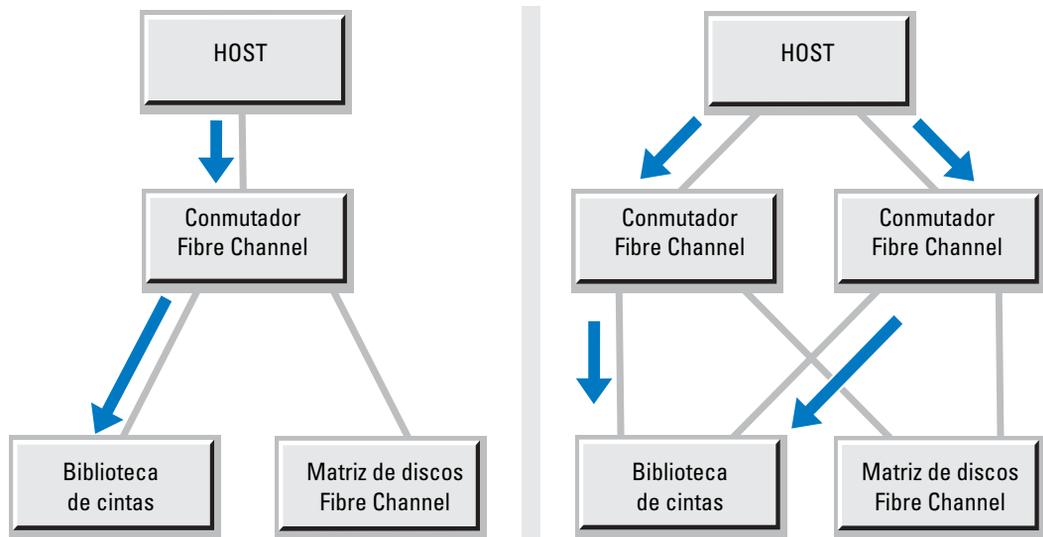
Al configurar bibliotecas de cintas en una SAN, el rendimiento aún puede verse afectado por varios factores. Entre estos factores se incluyen la velocidad del enlace FC, el flujo de datos entre el origen y la biblioteca de cintas, y las limitaciones de rendimiento de las matrices de almacenamiento externas. Muchos de estos factores pueden evitarse gracias a un conocimiento general de la instalación y la administración de la solución.

Incluso con la gran amplitud de banda que ofrece el protocolo FC en SAN, deben tenerse en cuenta algunas consideraciones para las unidades de cinta a fin de evitar situaciones en que el enlace FC pueda limitar el rendimiento. La velocidad de transmisión de datos de un enlace FC de 2 gigabits (Gb) es de 200 MB/s (es decir, 200 MB/s en el enlace de transmisión y 200 MB/s en el enlace de recepción). Por lo tanto, si se intenta utilizar varias unidades de cinta en el mismo enlace, puede excederse la amplitud de banda total de éste. Si el host funciona con un adaptador de 1 Gb antiguo, efectuar copias de seguridad de los datos en dos unidades puede ser suficiente para poner al descubierto limitaciones de rendimiento importantes.

Por lo tanto, al utilizar tres o más unidades de cinta simultáneamente en un enlace de 2 Gb, es posible que deba distribuir las copias de seguridad entre varias conexiones, en lugar de basarse en un único enlace. En casos como éste, resulta útil conocer la topología de la solución SAN. Si se sigue la ruta de los datos durante una operación de copia de seguridad desde que se leen del origen hasta que se graban en una cinta, los administradores podrán detectar mejor cualquier posible cuello de botella.

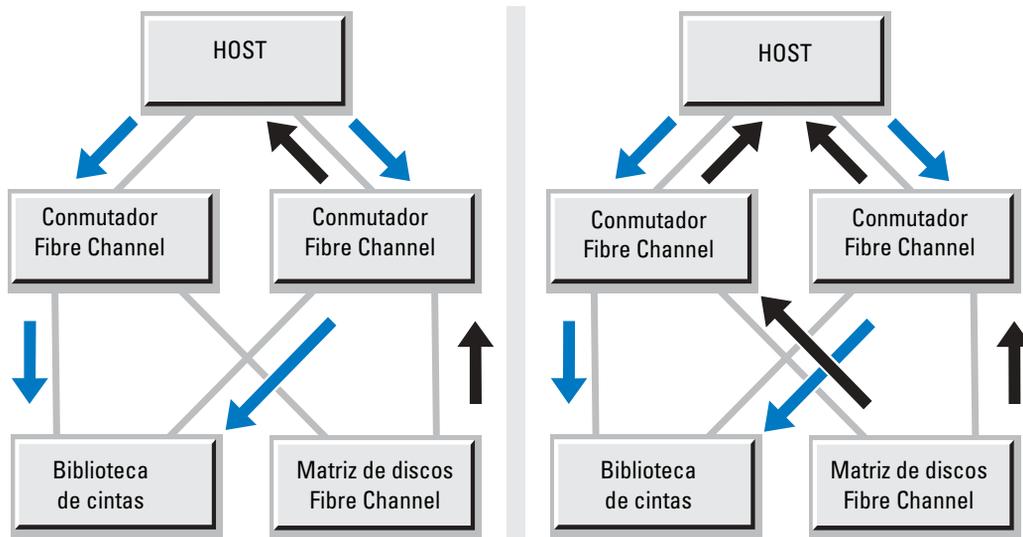
Cuando se identifica un cuello de botella, pueden tomarse medidas de acuerdo con la configuración. Por ejemplo, si la solución de copia de seguridad requiere que haya varias unidades en funcionamiento a simultáneamente, la división del hardware de cinta en varias redes Fabric puede mejorar el rendimiento al dividir las conexiones. Consulte la figura 1-6.

Figura 1-6. Flujo de datos único y dividido a la biblioteca de cintas



Asimismo, si se produce un cuello de botella en los datos que se están leyendo de una matriz de discos FC externa, la utilización de software de administración de E/S, como por ejemplo EMC® PowerPath®, con una conexión Fabric adicional equilibrará automáticamente la carga de los datos entre las distintas rutas y aumentará la disponibilidad mediante la sustitución tras error. Consulte la figura 1-7. En la parte izquierda de la figura se muestra una matriz de discos SAN en la que se fuerza que todos los datos pasen por un mismo enlace, lo que crea un cuello de botella que ralentiza el flujo de datos al hardware de cinta. En la parte derecha se muestra cómo el equilibrado de carga dobla la amplitud de banda de E/S procedente de la matriz al permitir que los datos se transmitan por dos enlaces.

Figura 1-7. Flujo de datos con cuello de botella frente a una matriz de discos con equilibrado de carga



Por último, las matrices de discos FC de la SAN también pueden experimentar las mismas limitaciones de rendimiento que se describen en “Configuración de la unidad de disco duro y la matriz RAID”. Por lo tanto, la mejora de las características de rendimiento de las matrices de discos también tendrá un efecto directo en la velocidad de copia de seguridad a través de la SAN.

Configuraciones de SAN mediante el puente Fibre Channel de la biblioteca

Algunas bibliotecas de cintas pueden conectarse a una SAN mediante un módulo de puente Fibre Channel. El módulo actúa como puente entre los protocolos SCSI y FC, y ofrece funciones adicionales de administración, seguridad y funcionamiento que no están disponibles en la mayoría de las bibliotecas FC nativas. Para obtener detalles sobre estas funciones, consulte la Guía del usuario del puente Fibre Channel correspondiente a su biblioteca de cintas.

En algunas configuraciones de biblioteca de cintas, el módulo de puente Fibre Channel puede actuar como cuello de botella y reducir el rendimiento de las unidades de cinta. Esto se debe a que la capacidad de procesamiento del módulo de puente Fibre Channel que se requiere para conectar las comunicaciones SCSI y FC no puede satisfacer la velocidad de datos agregados que ofrecen algunas configuraciones con varias unidades. A pesar de ello, en la mayoría de las soluciones de copia de seguridad de datos el módulo de puente Fibre Channel no representará el principal factor de limitación en el rendimiento de la cinta. Los servidores de copia de seguridad dedicados se encontrarán a menudo con situaciones en que las limitaciones en el host procederán de la acción de suministrar datos a varias unidades de cinta. Esto produce un rendimiento medio de la unidad inferior al nivel en que el módulo de puente Fibre Channel se convierte en un factor significativo.

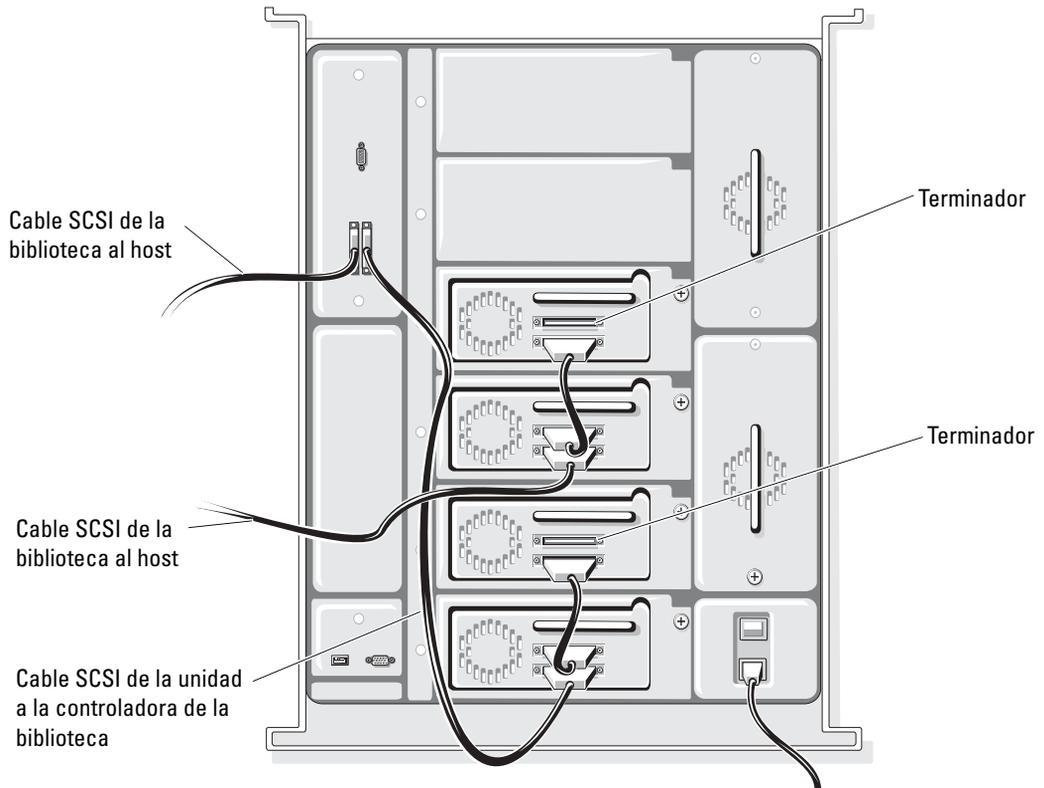
Configuraciones de cableado recomendadas

Cableado SCSI al host

Biblioteca de cintas con hasta seis unidades de cinta

Si la biblioteca de cintas está conectada mediante SCSI a un host, procure no haya más de dos unidades en un solo bus. En las bibliotecas con cinco o seis unidades de cinta, se requieren controladoras SCSI adicionales para garantizar que ningún bus se convierta en un obstáculo para maximizar el rendimiento.

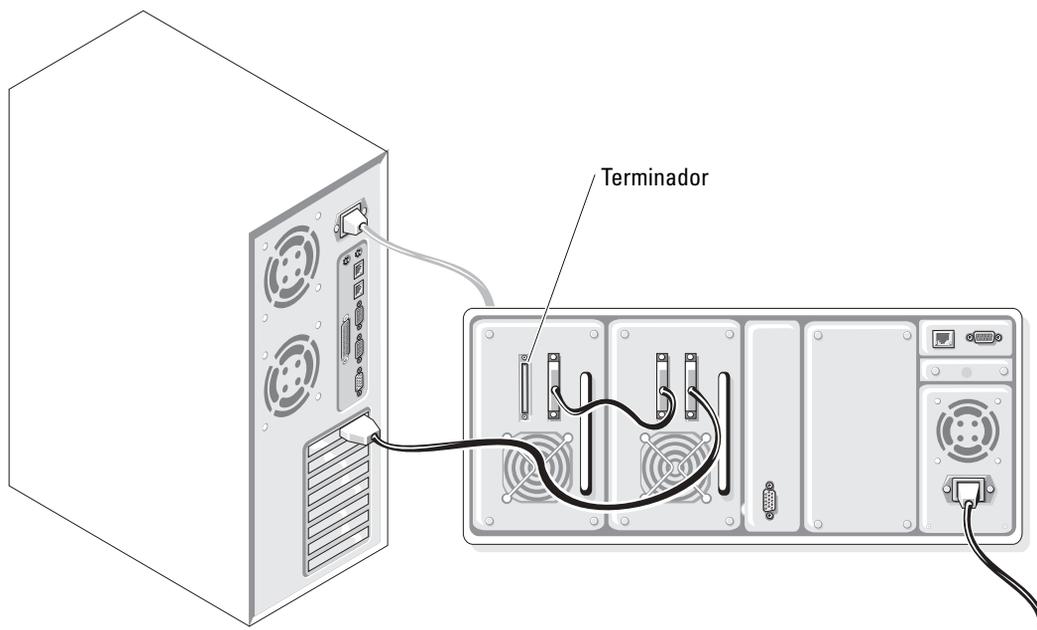
Figura 1-8. Cableado SCSI para bibliotecas con hasta seis unidades de cinta



Biblioteca de cintas con hasta dos unidades de cinta

Las unidades de una biblioteca de cintas con dos unidades completamente configurada pueden cablearse a un host del mismo bus SCSI sin que el rendimiento de la copia de seguridad encuentre limitaciones importantes. La velocidad de copia de seguridad correspondiente a dos unidades en un único bus SCSI coincidirá con la de dos unidades en buses distintos. Sin embargo, los clientes que activen la función de verificación en las aplicaciones de copia de seguridad pueden mejorar el rendimiento de la verificación dividiendo las dos unidades en dos buses SCSI. Si se hace esto, el rendimiento de la verificación puede llegar a mejorar un 25 por ciento

Figura 1-9. Cableado SCSI para bibliotecas con entre una y dos unidades de cinta



Cableado de las unidades al puente Fibre Channel

En las figuras de la 1-10 a la 1-17 se muestra cómo debe configurarse una biblioteca de cintas con hasta seis unidades de cinta y un módulo de puente Fibre Channel para optimizar el rendimiento de la cinta en FC.

Figura 1-10. Cableado del puente Fibre Channel con una unidad de cinta

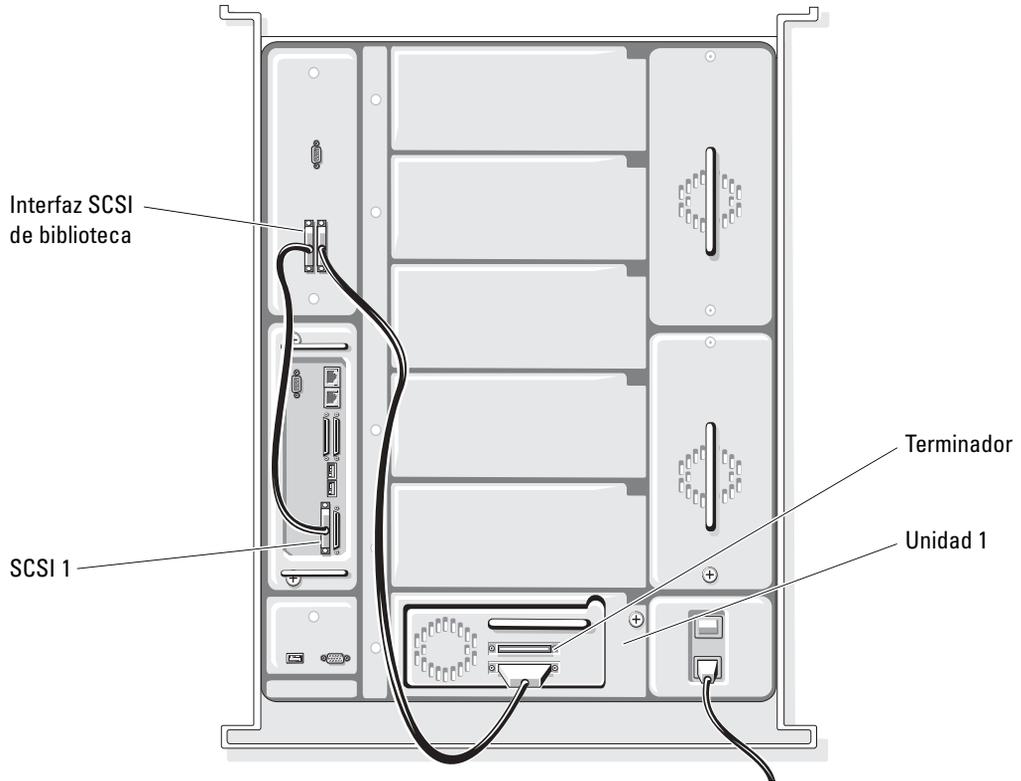


Figura 1-11. Cableado del puente Fibre Channel con dos unidades de cinta

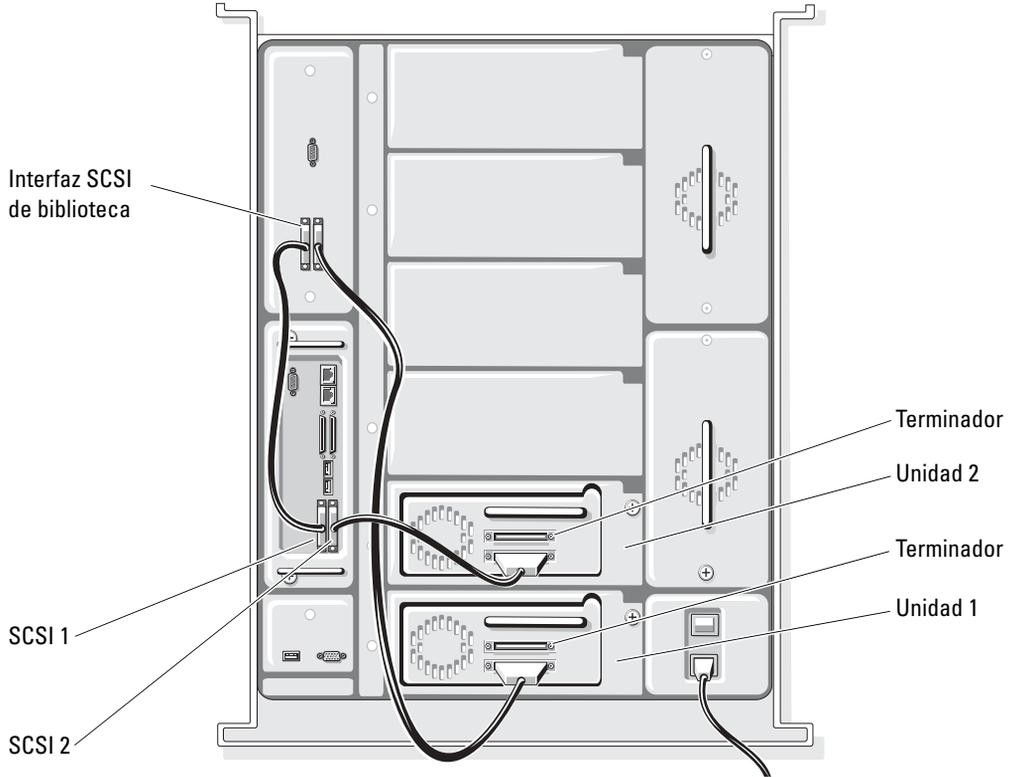


Figura 1-12. Cableado del puente Fibre Channel con tres unidades de cinta

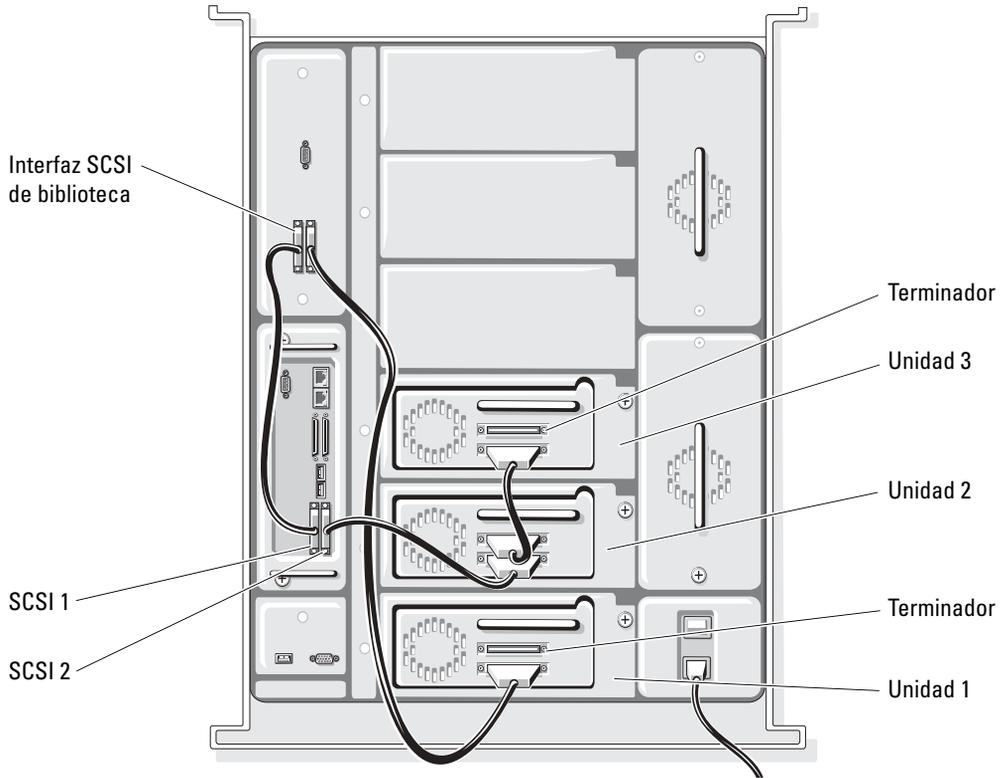


Figura 1-13. Cableado del puente Fibre Channel con cuatro unidades de cinta

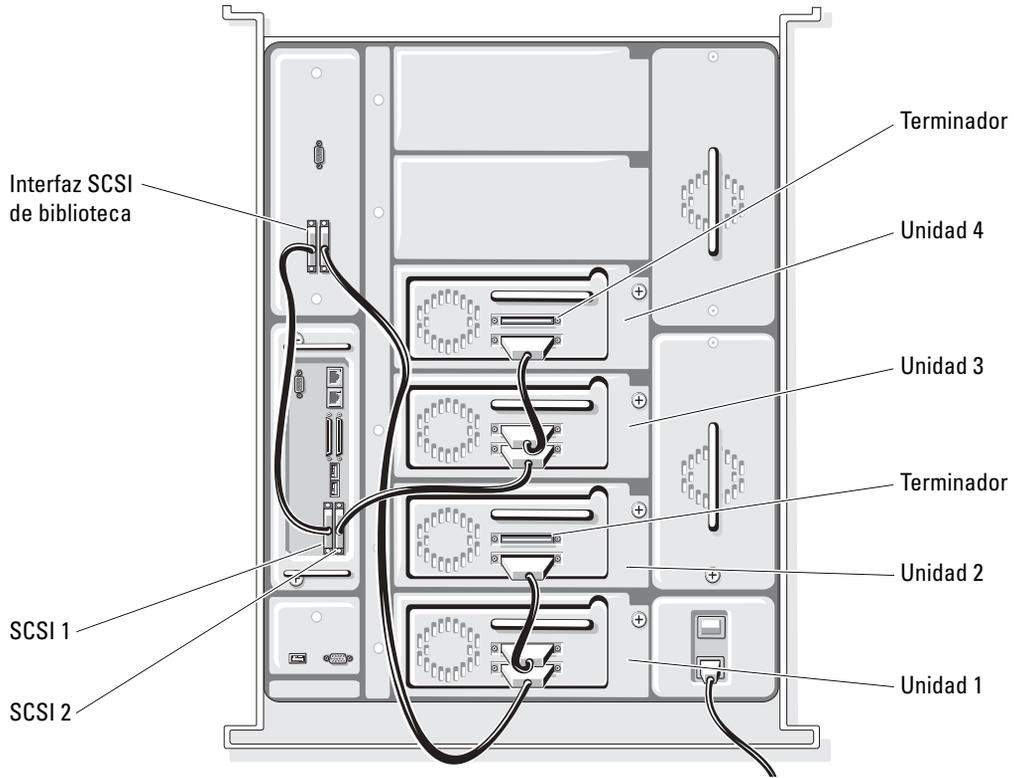


Figura 1-14. Configuración de la agrupación por zonas Fibre Channel para una biblioteca de cintas con entre una y cinco unidades

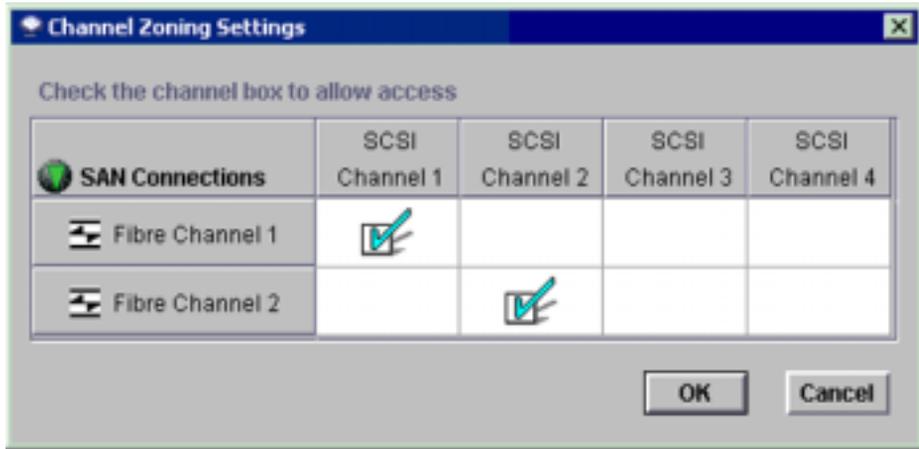


Figura 1-15. Cableado del puente Fibre Channel con cinco unidades de cinta

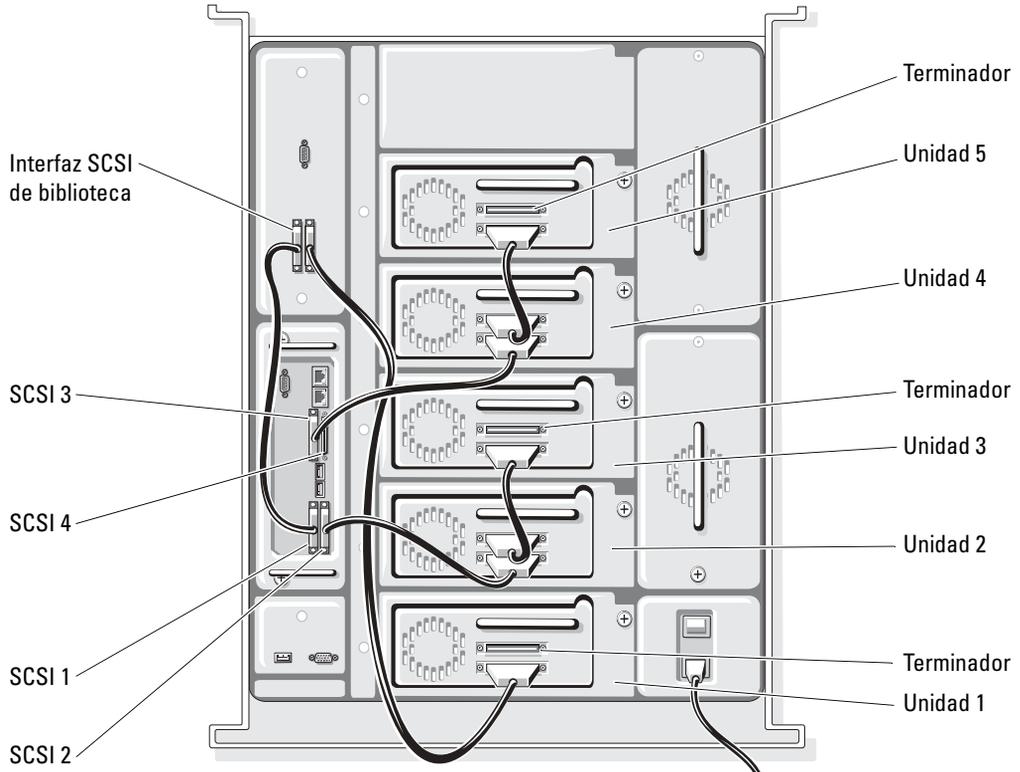


Figura 1-16. Cableado del puente Fibre Channel con seis unidades de cinta

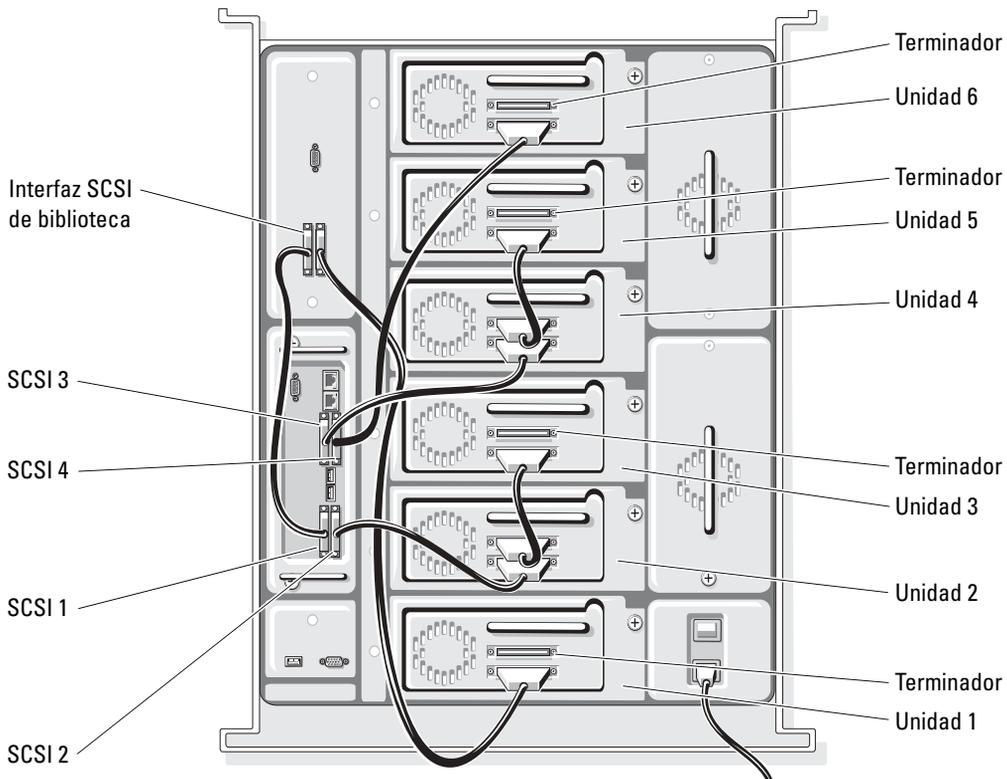
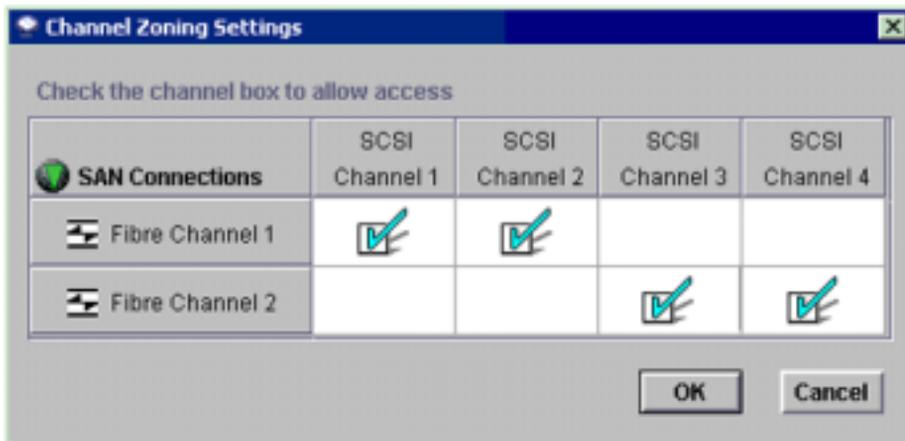


Figura 1-17. Configuración de la agrupación por zonas Fibre Channel para una biblioteca de cintas con cinco o seis unidades



En las figuras 1-18 y 1-19 se muestra cómo debe configurarse una biblioteca de cintas con entre una y dos unidades de cinta y un módulo de puente Fibre Channel para optimizar el rendimiento de la cinta en FC.

Figura 1-18. Biblioteca de cintas con una unidad

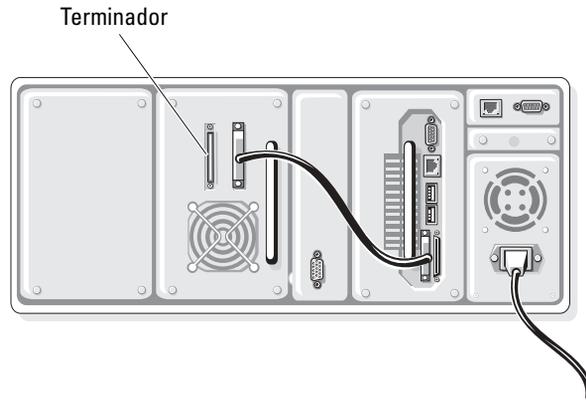


Figura 1-19. Biblioteca de cintas con dos unidades

