

# EMC CLARiiON Backup Storage Solutions: Backup-to-Disk Guide with IBM Tivoli Storage Manager

*Best Practices Planning*

---

**Abstract**

This white paper describes how to configure EMC® CLARiiON® CX Series storage systems with IBM Tivoli Storage Manager for optimal performance when backing up to disk.

October 2006

---

---

Copyright © 2006 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com

All other trademarks used herein are the property of their respective owners.

Part Number H711.2

---

<b>Table of Contents</b>	
<b>Executive summary .....</b>	<b>4</b>
<b>Introduction .....</b>	<b>4</b>
Audience .....	4
<b>Disk backup.....</b>	<b>4</b>
<b>Backup-to-disk performance factors .....</b>	<b>5</b>
RAID types .....	5
Striping and hot spares .....	5
Data alignment.....	5
File system block size .....	6
Windows file system tuning.....	6
Solaris file system tuning.....	6
HBA queue depth.....	7
<b>CLARiiON CX Series configuration settings .....</b>	<b>7</b>
Cache page size .....	7
Cache memory allocation .....	7
Watermarks.....	8
Front-end fibre port speed .....	8
Disk type and speed .....	8
Other considerations.....	8
LUN assignment .....	9
<b>Recommendations for disk-based backups.....</b>	<b>9</b>
<b>IBM Tivoli Storage Manager .....</b>	<b>9</b>
Basic architecture .....	9
Server storage .....	10
Disk-based backups.....	13
Sequential-access storage pools .....	13
Random-access disk storage pools .....	13
Performance considerations .....	16
Ethernet adapter.....	16
Host bus adapters .....	16
Server/client operating system.....	16
Disk-array type .....	16
RAID group type.....	16
Disk type.....	16
File size .....	16
Type of storage pool.....	17
Block size .....	17
Data source .....	17
LAN/LAN-free .....	17
<b>CLARiiON backup and restore performance.....</b>	<b>17</b>
Compression.....	17
Performance.....	17
<b>Conclusion .....</b>	<b>18</b>

---

## Executive summary

For many years, backing up data to tape was the industry standard. In recent years, disks have surpassed tape as the premier backup medium, due to disks' greater value proposition. Storage backup solutions that incorporate backup-to-disk yield significant benefits over traditional backup to tape. These benefits include:

- Near-term recovery of mission-critical data
- Rapid restore from disk
- Reduced backup time
- Greater reliability and predictability of the backup medium
- Significantly enhanced business uptime

Backup-to-disk (B2D) has emerged as a powerful solution for the rapid recovery of mission-critical data. Leading-edge and mainstream customers alike are increasingly using disk as the target for primary or operational backups, and quite often for many longer-term retention requirements as well. Tape complements a disk-based backup solution, and is most often used to provide longer-term data retention requirements for data whose time value to the business is no longer critical.

## Introduction

This white paper summarizes implementing a backup-to-disk solution with EMC® CLARiiON® storage systems and IBM Tivoli Storage Manager (TSM). The information in this paper provides backup administrators and technical consultants with specific configuration and parameter settings for both CLARiiON storage and TSM software to ensure optimum performance of disk backups. The solution outlined in this paper has been tested and is endorsed by EMC and Symantec.

## Audience

This white paper is intended for IT planners, storage architects, and storage and backup administrators who are involved in the direct implementation and configuration of backup-to-disk solutions.

## Disk backup

*Backup-to-disk (B2D)*, for the purpose of this paper, refers to writing backup data to a file system on disk. This capability is available in TSM. Alternate ways of performing backup-to-disk are also available in the form of tape emulation; however, this paper focuses on what we will refer to as “classic” B2D.

EMC recommends that all primary backups be written to disk—enabling significantly faster backups, and even more importantly, providing nearly instantaneous restores should the need arise. Under TSM Backup policy control, as business service levels dictate, data can be duplicated or migrated from disk to tape or other longer-term or archival media.

In the solution defined in this paper, disk storage resources are used for the primary or initial backup target often referred to as the operational backup, while tape resources are reallocated to the task of performing longer-term and offsite data retention needs.

---

## Backup-to-disk performance factors

This section describes the specific parameters that can be altered to improve the performance of backup-to-disk. EMC varied these parameters to determine recommended settings for optimum performance when backing up to a file system on disk.

### ***RAID types***

CLARiiON CX storage systems use RAID technology to combine disks into a logical unit (LUN) to improve reliability and/or performance. CLARiiON CX arrays support RAID 5, RAID 3, RAID 1, RAID 0, and RAID 1/0.

EMC specifically recommends RAID 3 for all ATA-based backup-to-disk-specific applications with CX arrays. RAID 3 has been optimized with the CLARiiON storage system to provide excellent read/write performance with ATA technology when used in backup specific applications (large sequential I/Os).

---

Configure RAID 3 for all ATA-based B2D configurations.

---

### ***Striping and hot spares***

RAID group striping of 4+1 and 8+1 provides very comparable overall storage system performance—especially when comparing performance based on total available storage capacity. RAID 4+1 groups are recommended in most customer environments. However, 8+1 can also be used and provides an easy way to optimize performance and hot spare requirements.

Within a single disk-array enclosure (DAE), 15 drives can be configured. Three 4+1 RAID groups fit nicely into this configuration. As additional DAEs are configured, a 4+1 and an 8+1 RAID group can be configured, permitting a hot spare drive to be configured in the second DAE. This provides a hot spare drive for each group of 30 drives. Alternating back to the original three 4+1 groups, 4+1 group, and 8+1 group (with two LUNs) will maintain this ratio of hot spares for all backup drives that are configured. This is not a hard requirement, but rather a rule of them to use when determining an appropriate number of hot spares to configure. Configure 4+1 single LUN RAID groups wherever possible. Utilize 8+1 RAID groups with two LUNs to help accommodate hot sparing.

### ***Data alignment***

Data alignment refers to the alignment of user data on the LUN. The mechanism for ensuring alignment is usually related to the operating system. Misalignment of data affects both read and write performance – the additional I/O caused by disk crossings slows down the overall system, while multiple accesses to disk make individual I/Os slower. In the case of cached writes, the misalignment causes slower flushing, which leads to reduced performance.

File system fragmentation over time is almost inevitable. Performing defragmentation regularly keeps performance optimal. There are a number of host-based utilities that can perform defragmentation in place to accomplish this. Other ways to defragment include copying data from one LUN to another at the file level (SnapView™, SAN Copy™ and LUN Migration will not defragment file systems) or even performing a backup of the file system, and then restoring the backed-up file system to the same LUN.

---

Perform regular defragmentation of the file system to ensure optimal performance.

---

---

## File system block size

Operating systems allocate space for files in blocks. The file system block size varies depending on the operating system being used. A larger file system block size provides a large I/O size, which can increase the bandwidth of a backup or restore operation. Larger block sizes may also reduce file system fragmentation, which can impact the backup/restore bandwidth.

The default file system block size for UFS, the native Solaris file system, is 8 KB. UFS can be configured with either 4 KB or 8 KB block size. EMC recommends that the defaults be used.

## Windows file system tuning

Windows systems (Intel) create a 63-sector hidden area on disk to house the master boot record (MBR). This causes all I/O in the file system made on the subsequent partition to be misaligned (not on a 256 KB boundary), and therefore I/O performance suffers. Because the reserved area is not an even multiple of 4 KB in size, at least one block in each element of the LUN will cause a disk crossing when accessed. The impact is particularly severe when writing.

---

This issue is not caused by Windows itself, but by the Intel architecture's use of an MBR at the beginning of disk. Linux systems, and other operating systems using the Intel architecture, will be similarly affected.

---

### Diskpar utility

One way to correct misalignment is by using `diskpar`, a program available on the Windows 2000 Resource Kit. Using `diskpar`, the user can alter the number of hidden blocks to a larger number so that the storage-system I/Os are aligned and better optimized. Use the value of 512 (blocks) for this setting.

### Diskpart utility

Microsoft's `diskpart` utility can also be used to create a partition at an explicit disk offset. The Disk Management snap-in places the partition at the end of any occupied area or on the first sufficiently large area. On MBR disks, the partition offset and the size are rounded to preserve the required cylinder alignment. Offsets are rounded to the closest valid value, and the size is always rounded up to the next valid value. Use the value of 1024 (KB) for this setting.

### CLARiiON LUN offset

An alternate way to correct misalignment if the `dispar/dispart` utilities are not readily accessible is to enter a value of 63 in the offset field when creating a CLARiiON LUN for use in a Windows system. This will correct the alignment.

It should be noted that if this LUN is used as the source for a clone or SAN Copy operation, the destination LUN should also have this offset value, otherwise it will be misaligned.

## Solaris file system tuning

There are two parameters you can tune with Solaris. The first parameter is `maxcontig` for the file system. It is important to verify that this is set high enough; otherwise, rotational delays will be incurred when laying down all the sectors in the write. Set `maxcontig` to the maximum number of sectors in a write. EMC recommends setting `maxcontig` to 256 (equivalent to 128 KB since there are 512 bytes in a sector). The `maxcontig` value may be set when making a file system (such as with `newfs`) or later with `tunefs`.

The second value to tune is the maximum number of bytes that can be written in a single I/O. This default value is usually 128 KB or 256 KB, depending on the hardware on which Solaris is running. There is little reason to go to 256 KB because the backup software would have to change to take advantage of the larger I/O size. For completeness, this line may be added to allow 1 MB I/O:

```
set maxphys=1048576
```

---

## ***HBA queue depth***

HBA queue depth refers to the number of HBA SCSI command buffers that can be allocated by an HBA. Default settings typically do not need to be reconfigured.

Typical settings should be:

Max LUNs	255
QDepth	32
Translate	QFULL --> QBUSY

## **CLARiiON CX Series configuration settings**

CLARiiON systems allow a wide range of parameters to be configured by the end user. The default settings are generally acceptable, but some additional tuning may be required to maximize performance in B2D environments. These include:

- Cache page size
- Cache memory allocation (read and write cache)
- Watermarks
- Front-end port fibre speed
- Type and speed of disks

### ***Cache page size***

The CX array supports a variety of cache page sizes including 2 KB, 4 KB, 8 KB (default), and 16 KB. 16 KB is recommended for very large I/Os such as backup-to-disk. This is particularly important with large LUN stripe sizes.

The cache page size setting is global; it is used by read and write cache operations on both SPs. It may be adjusted on the fly. Cache pages are allocated to individual I/Os as the request are made. A page size setting that is much larger than the I/O request size is inefficient. Specific environments may require you to tune this parameter to match the typical request size.

---

Use cache page size of 16 KB for backup-to-disk applications for optimal performance.

---

### ***Cache memory allocation***

The following are cache memory variables that can be set to maximize performance:

- Read Cache
  - Useful for sequential access when not buffered by host
- Write Cache
  - Large contributor to CLARiiON performance
  - Maximize!

Most applications will have data buffered by the file system that they run on top of or by specific application buffering. As a result, prefetching is not usually efficient for these applications. Therefore the size of the read cache may be set to a very low value. All remaining memory may then be used for write cache.

In a backup application, however, due to the sequential nature of backup data, prefetching is often beneficial. Restore operations (reads) will profit from the use of prefetching, provided that the file system is not too badly fragmented and that the memory available in the read cache can be increased. As an aside, CLARiiON allocates the larger proportion of write cache to the busiest SP, and recalculates the allocation every 10 minutes.

---

Set the size of the write cache to the maximum limit (3 GB in most cases) in backup-to-disk applications and split the remaining available cache to the read cache. A good rule of thumb is to allocate 100 MB to 200 MB for the read cache and the remaining to write cache.

---

## **Watermarks**

Watermark values manage write cache flushing, and are global for a CLARiiON. Default values typically perform well for most environments, including B2D applications. The default values are:

- Low watermark (LWM) set to 60 percent
- High watermark (HWM) set to 80 percent

The HWM determines when write cache flushing starts under normal operating conditions, and the LWM determines when flushing ends. These values should differ by approximately 20 percent. It is a requirement that the HWM be equal to or larger than the LWM. If watermark processing is disabled, the HWM and LWM are set to 100 percent.

---

CLARiiON default settings are optimal for backup-to-disk applications.

---

## **Front-end fibre port speed**

Front-end fibre port speed is configured manually and on a per-port basis. This value should be maximized whenever possible.

## **Disk type and speed**

The following disk variables can impact overall performance:

- FC or ATA
- Disk rotational speed
- Bandwidth (MB/s)

The rotational speed of a disk affects its random performance (for example, throughput) more than its sequential performance. A rule of thumb currently used for a 10K rpm FC disk is that it can perform 130 I/Os per second, while an ATA drive may provide about 60 (random) IO/s. However, bandwidth is almost identical however for both categories.

ATA drives do not typically implement command tag queuing, a process that handles requests from the host to the drive and allows the drive to accept multiple concurrent commands. This contributes to ATA drives not performing as well as FC drives with random I/O patterns, especially with smaller I/Os. Sequential performance is fairly good, though, making ATA drives ideal for backup-to-disk applications. With large sequential I/Os such as backup-to-disk applications, ATA performance will equal or in some cases surpass FC drives.

---

ATA drives should be configured for the majority of backup-to-disk applications.

---

## **Other considerations**

If it becomes necessary to configure array parameters differently than what is recommended in this paper, some of the implications that should be considered and taken into account include:

- Number of disks in RAID group
  - Factors to consider are:
    - Capacity requirements
    - Bandwidth and throughput requirements
    - Efficiency/level of protection
    - Rebuild time impact in case of failure
    - Likelihood of a two-disk failure
- Location of LUNs on physical disks

- 
- Outer cylinders are faster than inner cylinders
    - More sectors per track
  - LUNs nearest bust LUN get better service
    - Lower seek distances

In some environments, capacity may be less important than performance. Remember that there are different rules for IO/s and bandwidths for the different disk types.

## ***LUN assignment***

Exclusive LUNs within RAID groups must be assigned to a specific backup system/server. Multiple backup systems or servers can share CLARiiON storage systems as long as each server has exclusive use of one or more LUNs in the storage system for backup operations. LUNs can not be shared among different backup systems.

All LUNs created on a single RAID group should be owned by the same CLARiiON storage processor. This will minimize any negative performance implications that might arise from two different storage processors accessing the same LUN at the same time. This situation could arise when an 8+1 group is created with two separate LUNs.

## **Recommendations for disk-based backups**

The following disk-based backup guidelines will provide optimal performance for CLARiiON backup-to-disk configurations:

- Configure RAID 3 striping for backup-to-disk with CLARiiON ATA arrays.
- Configure all RAID groups as 4+1 stripes where possible.
- Bind one LUN per RAID group (best practice).
  - Bind two LUNS for a 8+1 RAID group (same SP) when used.
- All LUNs on a RAID group should be owned by the same storage processor.
- Use multiple backup/restore streams per RAID group – 4 to 5 works the best.
- Ensure source file defragmentation.
- Enable CLARiiON write cache.
- Set the page size to 16 KB for optimum performance.
- Set the read cache between 100 MB to 200 MB; use remaining available memory for the write cache.
  - If additional memory is available after maximizing the size of the write cache, allocate the rest to the read cache.
- Configure file system block and element sizes to their default settings,
- Configure operating system parameters as outlined in the “Performance considerations” section of this paper.

## **IBM Tivoli Storage Manager**

### ***Basic architecture***

IBM Tivoli Storage Manager (TSM) 5.3.3 is based on a client/server architecture. A Tivoli Storage Manager backup system consists of one or more Tivoli Storage Manager servers, one or more Tivoli Storage Manager clients, and optionally, one or more Tivoli Storage Manager storage agents. The Tivoli Storage Manager *server* controls and directs all Tivoli Storage Manager operations. Backup management information is stored or managed by the Tivoli Storage Manager server, including configuration

---

information, parameters for client systems to be backed up, the index of files backed up from those clients, and a database of media volumes where backed up data is stored.

Typically, a server and/or a Tivoli Storage Manager storage agent hosts one or more backup devices (disk or tape). When a backup occurs, a Tivoli Storage Manager client sends data to the Tivoli Storage Manager server, which directs the data to the appropriate storage. The server receives metadata information about the files being backed up separately and makes appropriate entries in the TSM database.

Tivoli Storage Manager manages client backup data in increments. The backup data originates from a single client and may consist of a single file, a directory, a file system, or a partition. Tivoli Storage Manager maintains a database of details about the files that are backed up, and a mapping to the volumes that contain the backup data. This database facilitates the ability to locate and recover data rapidly.

A Tivoli Storage Manager *client* is any system with data to be backed up, from desktop systems to large database servers. A client runs client software tailored to the operating system it uses. Normally, a client operates under the control of the Tivoli Storage Manager server according to the policies an administrator establishes. However, users on the client system can initiate backup and recovery actions as necessary. Select files can be backed up manually without waiting for regularly scheduled backups to occur.

Tivoli Storage Manager *storage agents* increase the parallelism and resiliency of a Tivoli Storage Manager configuration by allowing for LAN-free backups. LAN-free backups are used when the backup data is to be transferred over a SAN rather than the traditional LAN. This method has the advantages of offloading LAN traffic, increasing throughput in certain environments, lessening CPU overhead for the TSM server, and reducing restore times. A storage agent is connected to one or more storage devices. The client is configured to send its data to a storage agent for backup (for true LAN-free backups, the storage agent and client software must be installed on the same machine, although options also exist to install the storage agent on a third machine other than the client or the TSM server). Although client data is sent to the storage agent, client metadata is always directed to the Tivoli Storage Manager server. The server maintains the file index and media database information for all data on the storage agent. Thus, the storage agent acts as a slave to the Tivoli Storage Manager server and lacks the server's bookkeeping functions. Multiple SAN storage agents can share a single storage device (with additional software).

## **Server storage**

Tivoli Storage Manager was designed to take advantage of disk-based backup storage. The default setup for a Tivoli Storage Manager installation targets disk rather than tape as the first place backup data goes in the storage hierarchy. As the primary storage pool becomes full, or a specified number of days have passed, the primary storage pool can then be emptied into another storage pool.

Tivoli Storage Manager uses rules to determine how the client data is stored and managed. These rules are called *policies*, and include where the data is initially stored, how many backup versions are kept, how long archive copies are kept, and so on. Many different policies can exist and are assigned as needed to specific clients and even specific files.

A policy assigns a location in *server storage* where data is initially stored. Server storage is divided into *storage pools* that contain one or more *volumes* of a specific media type. The policy also specifies the initial pool into which data goes, and how many versions of the data to store. Thresholds are set up to determine when the data will move on to the *next* storage pool. Migration delays can also be set that will keep the data in the storage pool for a specified amount of time even if the threshold has been met.

The hierarchy of storage pools is defined by the system administrator and can be updated or changed as requirements change. Storage pools can be one of three types:

- Sequential access storage pools, either disk or tape
- Random-access disk storage pools
- Copy storage pools

---

## Sequential storage pools

Sequential-access storage pools are generally the preferred form of storage pools for use by TSM with disk or tape storage. Disk storage pools are used as the primary storage pool for backup and archive operations. Fibre Channel disks will generally yield the greatest performance. For storage pools with less demanding performance characteristics, EMC recommends ATA disk pools.

## Random-access disk storage pools

These are considered random-access devices and are recommended only for specific applications that benefit from their unique characteristics.

## Copy storage pools

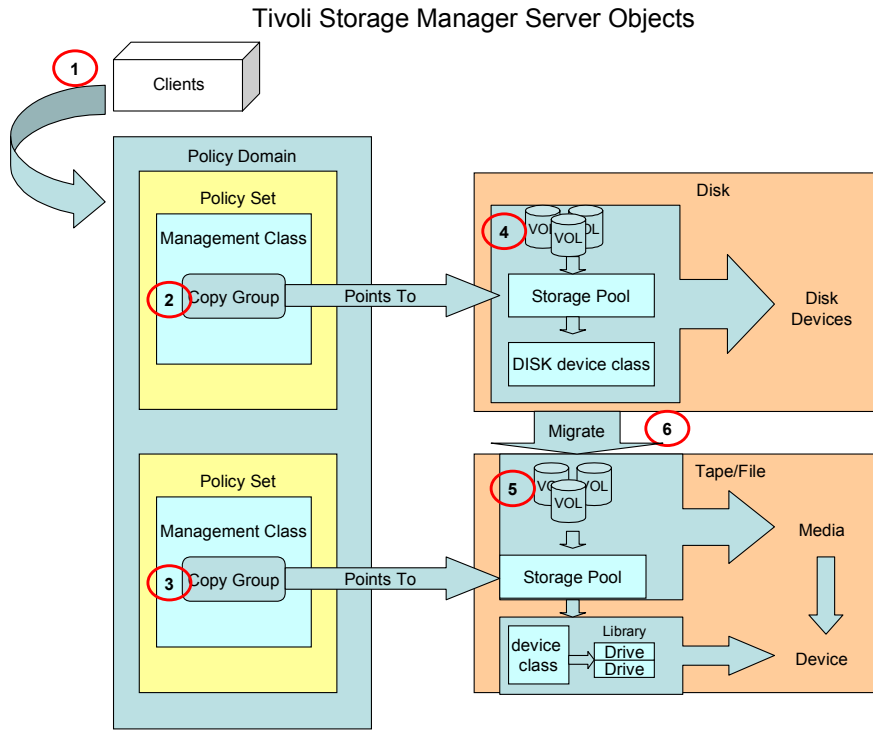
Copy storage pools are used to duplicate data to volumes that can be moved to an offsite location for safekeeping. Tape devices are useful here due to the portability of the media.

There are other types of client/server operations that can be managed with Tivoli Storage Manager:

- **Backup** — The backup process copies data from client workstations to server storage using the designated policy. The policy is version-based and determines how many copies are kept and how long “inactive” versions are kept.
- **Archive** — The archive process copies data from client workstations to server storage for long-term storage, and optionally deletes the client’s copy of the data. This is time- or event-based, rather than version-based.
- **Instant archive and rapid recovery** — A backup set is created using a complete set of recently backed-up files already in server storage. The backup set can then be taken to a client and restored directly to the client without the use of a backup server.
- **Migration** — The migration process frees server storage by moving the data from a client to server storage, leaving only a stub on the client pointing to the data’s new location. This is not to be confused with the concept of migration of data between storage pools (when data is moved from one pool to another).

The remainder of this paper focuses on the backup process.

Figure 1 shows the relationship between the different server-based objects used by Tivoli Storage Manager.



**Figure 1. Storage objects and hierarchy**

The following steps outline the backup process used by TSM:

1. When a Tivoli Storage Manager client is registered, it is associated with a policy domain. The other Tivoli Storage Manager policy objects are defined within the policy domain.
2. When the Tivoli Storage Manager client backs up or archives files, the data is bound to one or more management class in the domain's active policy set. Backup and archive copy groups within the management class and then route the data to an appropriate storage pool.
3. The storage pool specifies where and how the client data will be stored.
4. Storage pools are mapped to device classes, which represent devices. The storage pool contains volumes as indicated in the device type associated with the device class. This could be for a specific tape type or disk device.
5. All devices require a device class that specifies a device type. Removable media devices also require library and drive definitions, which enable Tivoli Storage Manager to mount and manage media.
6. Files that are initially stored in one storage pool can move to other storage pools if the pools are set up in a storage hierarchy.

---

The movement of backup data between storage pools is different than the movement when using copy storage pools. Copy storage pools allow you to make a copy of the data while writing to one of the designated storage pools. The copied data is meant for offline storage but can be used during a restore if other copies of the data are inaccessible.

## ***Disk-based backups***

There are two options for disk-to-disk backups:

- Sequential-access storage pools
- Random-access disk storage pools

### **Sequential-access storage pools**

These are considered sequential-access devices. Sequential-access pools are organized into volumes, which are associated with drives contained in libraries. When used as primary storage pools, they can be shared by one or more servers and storage agents in a backup configuration using the LAN-free option. Additional software from IBM (SANergy or SAN FS) is needed to share the disk devices. SANergy or SAN FS allow the server and storage agents to share the volumes to which backups are stored. Since the storage agent has direct access to the storage, the data does not need to flow through the LAN to the Tivoli Storage Manager server and then out to the storage. *Paths* are created so the Tivoli Storage Manager server knows to equate the storage agent's view of the volumes with its own view. Since the storage volumes are created and owned by the Tivoli Storage Manager server, the Tivoli Storage Manager server must also be running SANergy and configured along with the storage agents and clients.

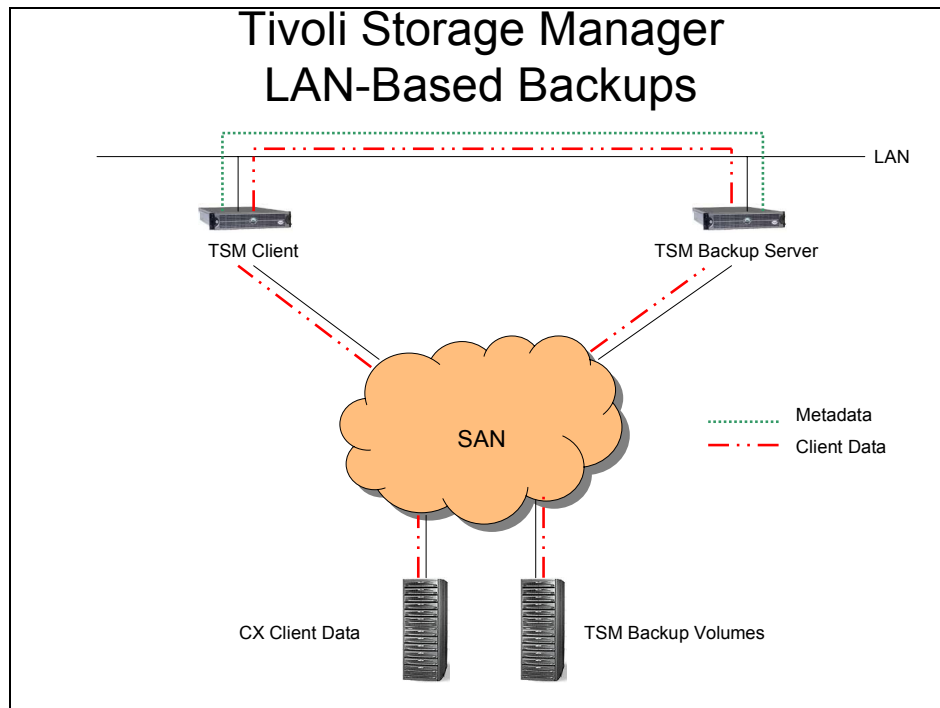
EMC recommends that you use ATA disk drives for these file library storage pools when they are used as secondary storage pools; however, Fibre Channel disks are recommended for primary pools. Secondary storage pools do not need Tivoli SANergy and LAN-free options since the Tivoli Storage Manager server owns the volumes in the primary storage pool and secondary storage pool. The Tivoli Storage Manager server manages movement of the data between these volumes.

### **Random-access disk storage pools**

Random-access storage pools are somewhat shareable because multiple clients can store data to the same disk storage pools; however, they can not be accessed in a LAN-free environment using Tivoli device sharing. The Tivoli Storage Manager server exclusively owns the volumes in these pools. Disk storage pools contain one or more volumes. Creating more and smaller volumes on separate LUNs will yield the greatest amount of backup concurrency.

Random-access storage pools are not recommended for general configurations but only for certain specialized applications of TSM.

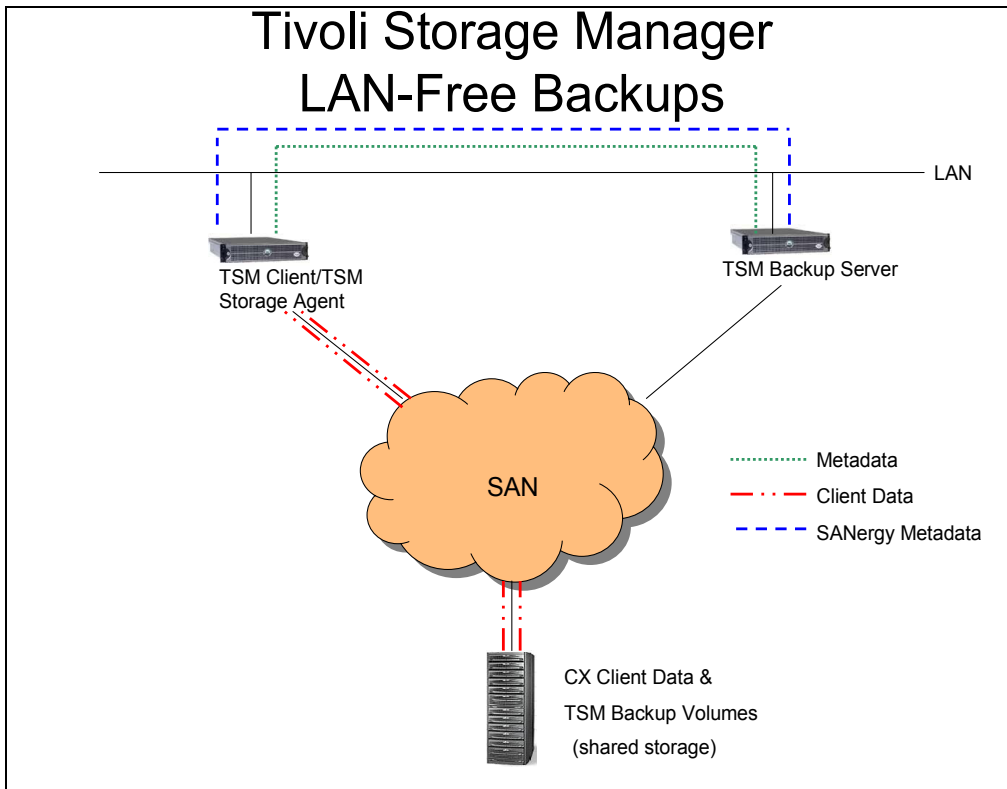
To fully understand disk backups, you need to understand how the data flows across the LAN and SAN. Figure 2 and Figure 3 represent sample backup-to-disk configurations. For illustrative purposes, the figures focus on primary storage pools.



**Figure 2. Tivoli Storage Manager LAN configuration**

In this example, client data (data to be backed up) is stored on the CX array on the left. When a backup is initiated, the client data flows from that CLARiiON array through the SAN to the client. From there, the data flows over the LAN to the backup server. Finally, the backup data flows to the storage pool owned by the backup server on the CX array on the right. Backup metadata is also passed between the client and the backup server.

The source and target arrays do not need to be separate devices. A single CLARiiON array is capable of hosting the client data as well as the backup data.



**Figure 3. Tivoli Storage Manager LAN-free configuration**

In this example, client data is stored on the CX array on disks shared by the storage agent. Each of the client systems and servers are running the Tivoli SANergy software. When a backup is initiated, the backup data flows from that CLARiiON array through the SAN to the storage agent/client. From there, the data flows back over the SAN to the storage pool owned by the backup server, which is also shared by the storage agent/client. Backup metadata is also passed between the client and the backup server. Tivoli SANergy metadata is passed between the server and storage agent to maintain volume sharing.

The primary storage pool in this case is a file library. While it is still disk-based, it is of a different device class than disk storage pools and is SAN-shareable.

---

## **Performance considerations**

There are many factors to consider when analyzing the performance of a Tivoli Storage Manager backup system. A good understanding of each of these is paramount to setting up a successful backup scenario.

### **Ethernet adapter**

When doing backups over the LAN, a GigE LAN provides far superior performance than a 100 Mb/s LAN, particularly if the LAN is not dedicated for Tivoli Storage Manager use. This factor above all others can be the biggest bottleneck in the system in a LAN-based configuration. GigE LANs approach the speed of SANs, which today usually run at 1 Gb/s or 2 Gb/s. Additionally, there is less installation overhead since the target storage devices are completely owned and managed by the Tivoli Storage Manager server. No other software such as Tivoli SANergy needs to be used in a totally LAN-based solution.

In a LAN-free solution, the importance of the network adapter choice diminishes since only metadata passes through the LAN. SANergy adds metadata traffic over the LAN above and beyond the backup metadata, but this is a relatively small amount of data.

### **Host bus adapters**

Current Fibre Channel adapters can run at 1 Gb/s or 2 Gb/s. The faster the selection, the faster data will move from the client to the Tivoli Storage Manager server in a LAN-free configuration. Without the LAN-free option, the choice is less significant since the LAN will probably become the bottleneck, especially if it is slower than a GigE LAN.

### **Server/client operating system**

Since the backups are dealing with “files” and the server stores data in “file” volumes, the speed and robustness of the operating system involved is significant. Lab testing has shown that Solaris and AIX may be slower than other operating systems due to their file system overhead.

### **Disk-array type**

Aggregate throughput will be affected by the array model and amount of cache configured.

### **RAID group type**

The grouping of the disks greatly impacts the performance of the array as a disk storage device. Performance data included in this document will demonstrate the performance gains possible with various RAID configurations.

### **Disk type**

Fibre Channel disks by nature will outperform ATA disks most of the time and are recommended for primary storage pools. ATA disks are recommended for subsequent storage pools.

### **File size**

The size of the files that are being backed up can play a big role in how fast the backups occur. Disk arrays perform best when larger chunks of data are being written at one time. The impact to the backup server will also be less with metadata since there are fewer files flowing through the backup server at a given time. Testing showed a clear distinction in performance when 1 GB files were compared to 1 MB files during the initial backup process. Moving the data from the primary storage pool to a disk-based secondary storage pool showed a less-significant performance difference.

---

## Type of storage pool

The client's policy determines the destination storage pool to use for backup. The storage pool defines what the next storage pool is, and uses thresholds to mark the point at which data empties from the destination pool to a next pool. The high migration threshold attribute determines the point at which moving should begin. This attribute specifies the percentage of available space that is used. The low threshold specifies when the moving should end. This is also specified as a percentage of available space that is used. Migration delays specify the number of days data must remain in the storage pool. The types of devices used for storage pools will greatly impact performance. EMC recommends using fibre drives for the primary storage pools, and ATA drives for secondary storage pools. Tape drives can be added to the mix for long-term or archival storage.

## Block size

Tivoli Storage Manager does not allow block sizes to be modified. Block size changes occur internally and are controlled by the Tivoli Storage Manager backup server.

## Data source

File servers and workstations should be backed up to disk storage. Multiple backup/restore sessions should be configured for file servers. The client option **resourceutilization** should be set to 10.

## LAN/LAN-free

LAN-based backups should generally use sequential storage pools. Multiple disk volumes should be configured for best performance (one for each large client that is backing up simultaneously). Only one disk volume per storage array or physical disk (LUN) should be configured.

LAN-free backups should use sequential access file or tape storage pools. SANergy or SAN FS must be installed for LAN-free backup-to-disk.

## CLARiiON backup and restore performance

Performance comparisons between disk- and tape-based solutions must take into account the total time to back up and restore data. This time is often dependant on more than just the raw data transfer speed of the storage device. You must consider overhead normally associated with tape, including things like tape positioning, file seek times, library mechanical load/unload times, and so forth. This overhead is not found in a disk-based solution, which contributes to disk solutions providing overall faster backup/restore times.

## **Compression**

TSM can compress data for disk-based backup using software. Testing found that performance suffers when this option is enabled. In cases where conserving disk space is critical, compression may be enabled.

---

Disable software compression when trying to optimize B2D performance

---

## **Performance**

Total aggregate performance of any backup or restore operation varies depending upon many variables. These variables include the type of CLARiiON storage system, backup host, and file system, and on the nature of the data and the bandwidth that is available. Single stream performance approaching 150 MB/s may be obtained in select environments where the fastest array subsystems are used. As more backup or recovery streams, RAID groups and LUNs/RAID group are added and running simultaneously, total aggregate performance of the storage system can substantially increase. Of course, smaller arrays will

---

provide slower performance levels. Appropriate EMC technical resources should be consulted for information on performance expectations for your specific operating environment.

## Conclusion

Backup-to-disk has emerged as a technology that offers significant benefits over the traditional tape backup process. With the changed economics of disk technology, backup-to-disk solutions are now affordable. Leading-edge and mainstream customers are implementing backup-to-disk solutions to overcome the limitations of traditional tape-based solutions, particularly for operational backup and restore requirements.

Many variables can be set in the CLARiiON storage system and IBM TSM to optimize backup-to-disk performance. To maximize performance of backup-to-disk and restore operations using CLARiiON and IBM TSM follow the guidelines below:

- Configure RAID 3 striping for backup-to-disk with CLARiiON ATA arrays.
- Configure all RAID groups as 4+1 stripes where possible
- Bind one LUN per RAID group (best practice).
  - Bind two LUNS for a 8+1 RAID group (same SP) when used.
- All LUNs on a single RAID group should be owned by the same SP.
- Use multiple backup/restore streams per RAID group – 4 to 5 works best.
- Ensure source file defragmentation,
- Enable CLARiiON write cache
- Set the page size to 16 KB for optimum performance.
- Set the read cache between 100 MB to 200 MB; use the remaining available memory for the write cache.
  - If additional memory is available after setting the write cache to the upper limit, add additional memory to the read cache.
- Configure file system block and element sizes to their default settings.
- Configure operating system parameters as in the “Performance considerations” section of this paper.
- Disable backup software compression where possible.

---

Note: When configuring CLARiiON storage systems for backups, create five-disk RAID 3 configurations. Exclusive LUNs within the RAID group must be assigned to each backup system/server. Multiple systems/servers can share CLARiiON storage systems as long as each server has exclusive use of one or more LUNs in the storage system for backup operations.

---