Chapter 3 Configuration Options

The advantages of RAID are: Availability, Capacity and Performance. Choosing the right RAID level and drive failure management can increase Availability, subsequently increasing Performance and Capacity. The DA-3000 RAID controller provides complete RAID functionality and enhanced drive failure management.

3.1 RAID Management

RAID stands for Redundant Array of Inexpensive Drive. The advantages of using a RAID storage subsystem are:

- Provides disk spanning by weaving all connected drives into one single volume.
- Increases disk access speed by breaking data into several blocks when reading/writing to several drives in parallel. With RAID, storage speed increases as more drives are added.
- Provides fault-tolerance by mirroring or parity operation.

What are the RAID levels?

RAID Level	Description	Minimum Drives	Data Availability	Performance Sequential	Performance Random
NRAID	Non-RAID	1		Drive	Drive
RAID 0	Disk Striping	N	==NRAID	R: Highest W: Highest	R: High W: Highest
RAID 1 (0+1)	Mirroring Plus Striping (if N>1)	N+1	>>NRAID ==RAID 5	R: High W: Medium	R: Medium W: Low
RAID 3	Striping with Parity on dedicated disk	N+1	>>NRAID ==RAID 5	R: High W: Medium	R: Medium W: Low
RAID 5	Striping with interspersed parity	N+1	>>NRAID ==RAID 5	R: High W: Medium	R: High W: Low

NRAID

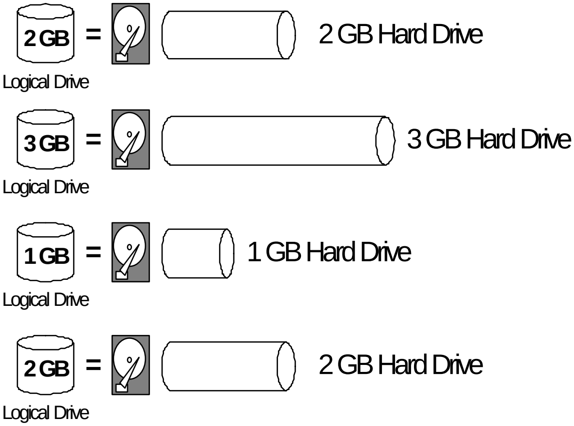
Disk Spanning

NRAID	
Minimum Disks required	1
Capacity	N
Redundancy	No

NRAID stands for Non-RAID. The capacity of all the drives are combined to become one logical drive (no block striping). In other words, the capacity of the logical drive is the total capacity of the physical drives. NRAID does not provide data redundancy.

JBOD

Single Drive Control

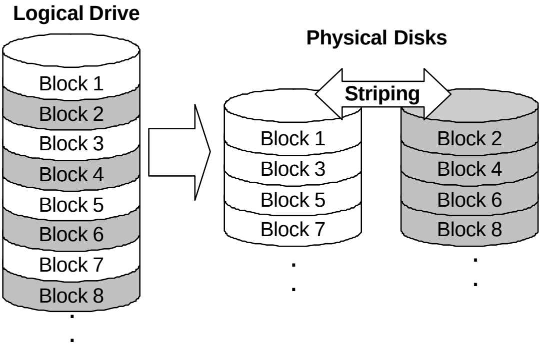


JBOD	
Minimum Disks required	1
Capacity	1
Redundancy	No

JBOD stands for Just a Bunch of Drives. The controller treats each drive as a stand-alone disk, therefore each drive is an independent logical drive. JBOD does not provide data redundancy.

RAID 0

Disk Striping

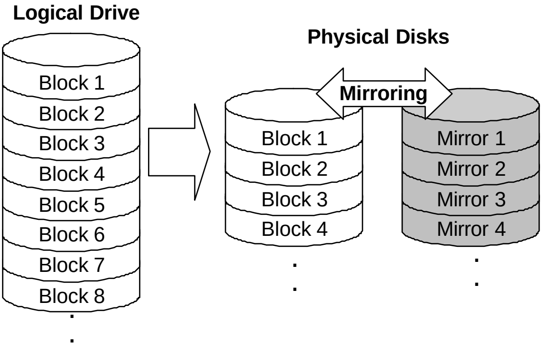


RAID 0	
Minimum Disks required	2
Capacity	N
Redundancy	No

RAID 0 provides the highest performance but no redundancy. Data in the logical drive is striped (distributed) across several physical drives.

RAID 1

Disk Mirroring

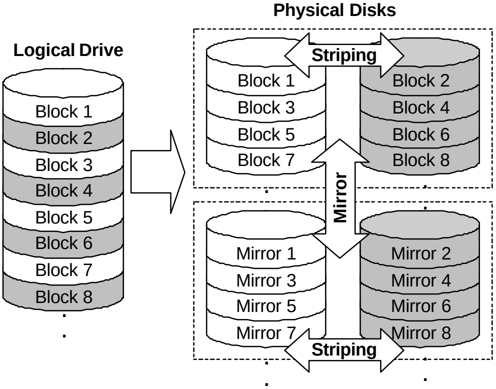


RAID 1	
Disks required	2
Capacity	N/2
Redundancy	Yes

RAID 1 mirrors the data stored in one hard drive to another. RAID 1 can only be performed with two hard drives. If there are more than two hard drives, RAID (0+1) will be performed automatically.

RAID (0+1)

Disk Striping with Mirroring



RAID (0+1)	
Minimum Disks required	4
Capacity	N/2
Redundancy	Yes

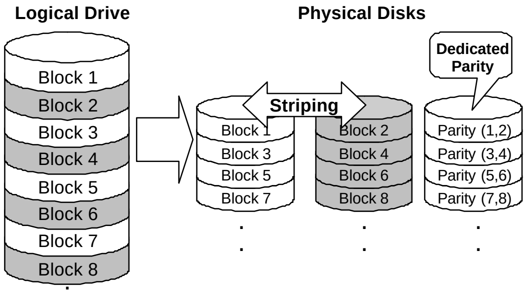
RAID (0+1) combines RAID 0 and RAID 1 - Mirroring and Striping. RAID (0+1) allows multiple drive failure because of the full redundancy of the hard drives. If there are more than two hard drives assigned to perform RAID 1, RAID (0+1) will be performed automatically.

IMPORTANT:

☞ *"RAID (0+1)" will not appear in the list of RAID levels supported by the controller. If you wish to perform RAID 1, the controller will determine whether to perform RAID 1 or RAID (0+1). This will depend on the drive number that has been selected for the logical drive.*

RAID 3

Disk Striping with Dedicated Parity Disk

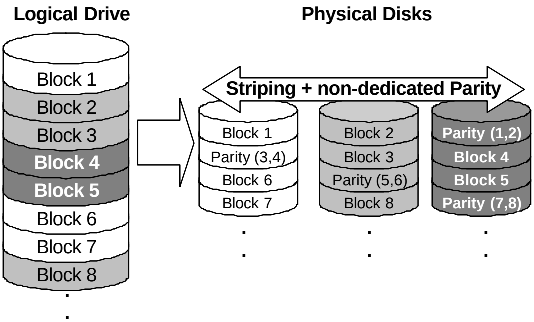


RAID 3	
Minimum Disks required	3
Capacity	N-1
Redundancy	Yes

RAID 3 performs Block Striping with Dedicated Parity. One drive member is dedicated to storing the parity data. When a drive member fails, the controller can recover/regenerate the lost data of the failed drive from the dedicated parity drive.

RAID 5
Striping with
Interspersed Parity

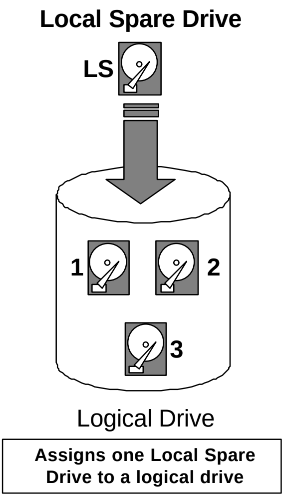
RAID 5	
Minimum Disks required	3
Capacity	N-1
Redundancy	Yes



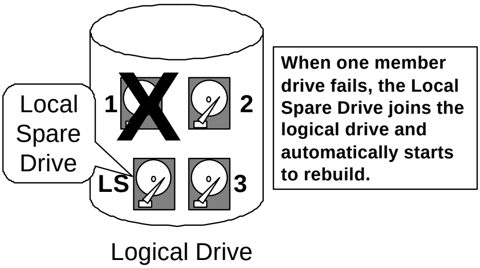
RAID 5 is similar to RAID 3 but the parity data is not stored in one dedicated hard drive. Parity information is interspersed across the drive array. In the event of a failure, the controller can recover/regenerate the lost data of the failed drive from the other surviving drives.

3.2 Drive Failure Management

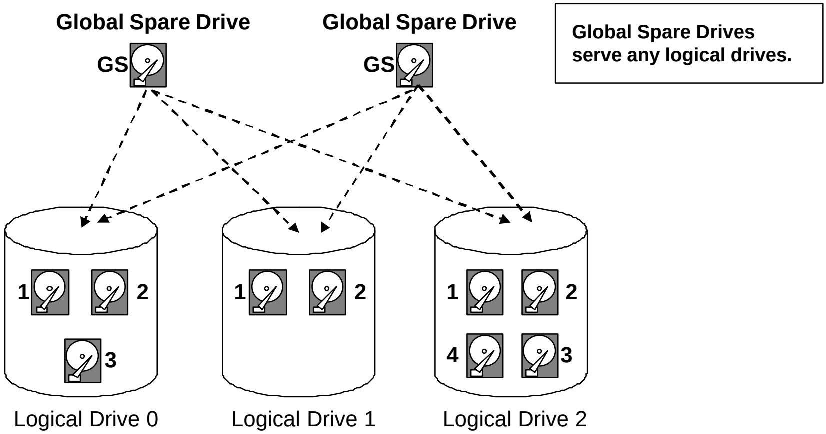
3.2.1 Global and Local Spare Drive

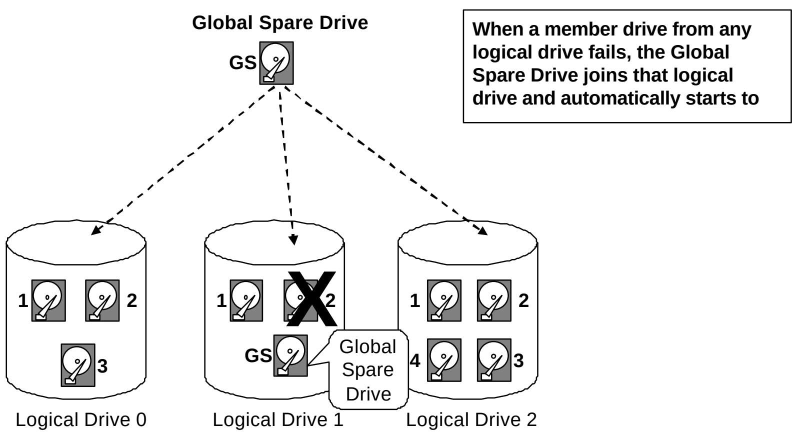


Local Spare Drive is a standby drive assigned to serve one specified logical drive. When a member drive of this specified logical drive fails, the Local Spare Drive becomes a member drive and automatically starts to rebuild.



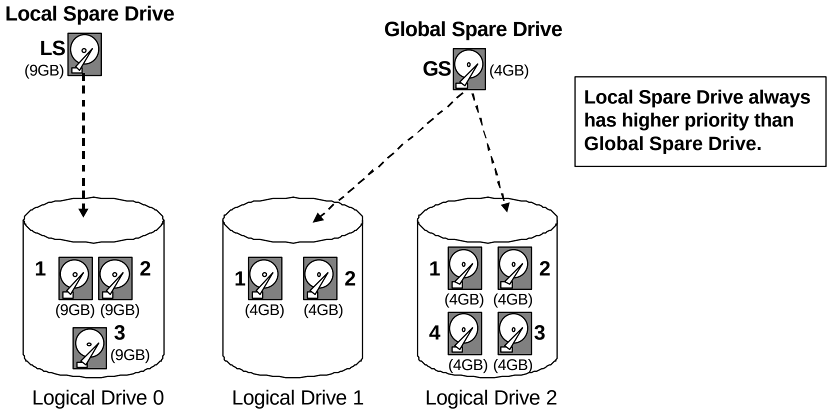
Global Spare Drive does not only serve one specified logical drive. When a member drive from any of the logical drive fails, the Global Spare Drive will join that logical drive and automatically starts to rebuild.





The DA-3000 RAID controller provides both Local Spare Drive and Global Spare Drive functions. On certain occasions, applying these two functions together will better fit various needs. Take note though that the **Local Spare Drive always has higher priority than the Global Spare Drive.**

In the example shown below, the member drives in Logical Drive 0 are 9 GB drives, and the members in Logical Drives 1 and 2 are all 4 GB drives. It is not possible for the 4 GB Global Spare Drive to join Logical Drive 0 because of its insufficient capacity. However using a 9GB drive as the Global Spare drive for a failed drive that comes from Logical Drive 1 or 2 will bring huge amount of excess capacity since these logical drives require 4 GB only. In the settings below, the 9 GB Local Spare Drive will aid Logical Drive 0 once a drive in this logical drive failed. If the failed drive is in Logical Drive 1 or 2, the 4 GB Global Spare drive will immediately give aid to the failed drive.



3.2.2 Identifying Drives

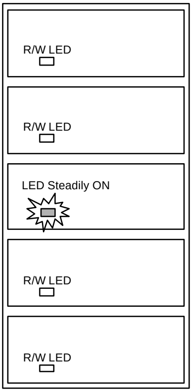
Assuming there is a failed drive in the RAID 5 logical drive, make it a point to replace the failed drive with a new drive to keep the logical drive working.

When trying to remove a failed drive and you mistakenly removed the wrong drive, you will no longer be able to read / write the logical drive because the two drives may have already failed.

To prevent this from happening, the controller provides an easy way of identifying for the failed drive. That is, the read / write LED of the failed hard drive will light. This LED will prevent you from removing the wrong drive, and is also helpful when locating for a drive.

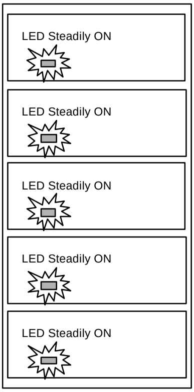
Flash Selected SCSI Drive

The Read / Write LED of the drive you selected will light steadily for about one minute.



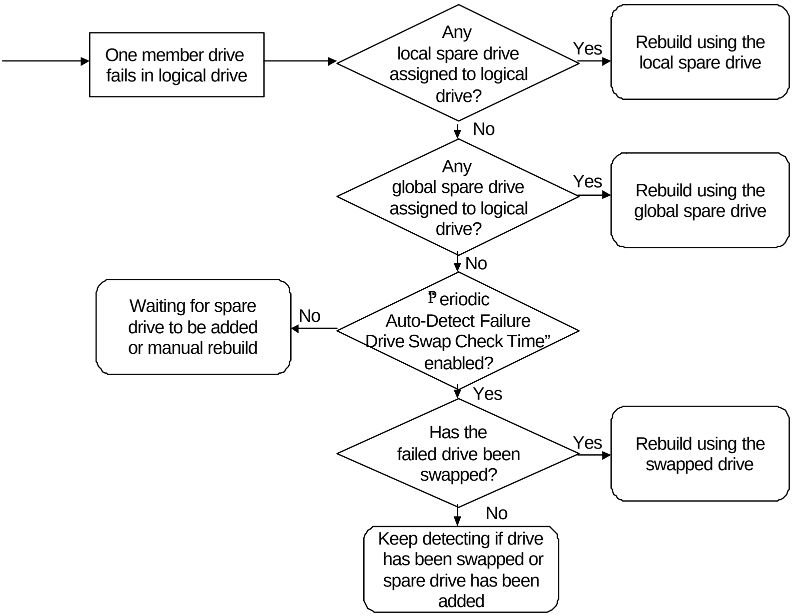
Flash All SCSI Drives

The Read / Write LED of all connected drives will light for about one minute. If the LED of the defective drive did not light on the "Flash Selected SCSI Drive" function, use "Flash All SCSI Drives". The "Flash All SCSI Drives" function will light LEDs of all the drives except the defective one.



3.72.3 Automatic Rebuild and Manual Rebuild

Automatic Rebuild



When a member drive in the logical drive failed, the controller will first check whether there is a Local Spare Drive assigned to this logical drive. If yes, it will automatically start to rebuild.

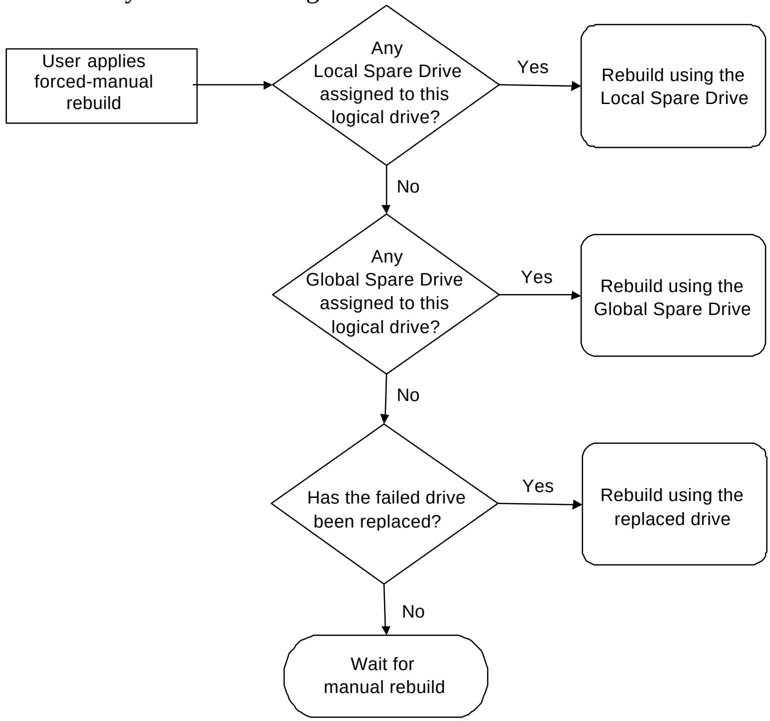
If there is no Local Spare Drive available, the controller will search for a Global Spare Drive. If there is a Global Spare Drive, it will automatically rebuild the logical drive.

If neither Local Spare Drive nor Global Spare Drive is available, the controller will not try to rebuild unless the user applies a forced-manual rebuild.

Manual Rebuild

When a user applies forced-manual rebuild, the controller will first check whether there is any Local Spare Drive assigned to this logical drive. If yes, it will automatically start to rebuild.

If there is no Local Spare Drive available, the controller will search for a Global Spare Drive. If there is a Global Spare Drive, it will automatically rebuild the logical drive.



If neither Local Spare Drive nor Global Spare Drive is available, the controller will detect the SCSI channel and ID of the failed drive. Once the failed drive has been replaced by a new drive / used drive, it starts to rebuild using the replaced drive. If there is no available drive for rebuilding, the controller will not try to rebuild again until the user applies another forced-manual rebuild.

3.2.4 Concurrent Rebuild in RAID (0+1)

RAID (0+1) allows multiple drive failure and concurrent multiple drive rebuild. Newly replaced drives must be scanned and set as Local Spare Drives. These drives will be rebuilt at the same time (you do not need to repeat the rebuilding process for each drive).

3.3 Disk Array Parameters

3.3.1 Rebuild Priority

Rebuilding time will depend on the capacity of the logical drive. The DA-3000 RAID controller provides background rebuilding ability. Meaning, the controller is able to serve other I/O requests while rebuilding the logical drives. The rebuilding process is totally transparent to the host computer or the operating system.

The background rebuild process has four priority options:

- Low
- Normal
- Improved
- High

The default priority is "Low" which uses the controller minimum resources to rebuild. Choosing "Normal" or "Improved" will speedup the rebuilding process and choosing "High" will use the controller maximum resources to complete the rebuilding process at the shortest time.

Rebuild priority can be configured through the RS-232C Terminal Interface, GUI RAID Manager or the front panel.

3.3.2 Verify-after-Write

The controller has the ability to force the hard drives to verify after data has been written to the media of the HDD. There are three selectable methods:

- Verification on LD Initialization Writes
  Performs Verify-after-Write while initializing the logical drive.
- Verification on LD Rebuild Writes
  Performs Verify-after-Write during the rebuilding process.
- Verification on LD Normal Drive Writes
  Performs Verify-after-Write during normal I/O requests.

Each method can be enabled or disabled individually. Hard drives will perform Verify-after-Write according to the selected method.

IMPORTANT:
The "Verification on LD Normal Drive Writes" method will affect "Write" performance during normal use.

3.4 Cache Parameters

3.4.1 Optimization for Sequential or Random I/O

When using RAID with applications such as video or image oriented applications, the application reads/writes from the drive using large-block, sequential files instead of small-block, random access files. The DA-3000 RAID controller provides the options to optimize for large-sequential I/O or optimize for small-random I/O access.

"Optimization for Sequential I/O" provides larger stripe size (block size, also known as Chunk size) than "Optimization for Random I/O". A lot of the controller internal parameters will also be changed to optimize for sequential or random I/O. The change will take effect after the controller reboots.

If the existing logical drives were built with "Optimization for Random I/O", these logical drives will not read/write when using "Optimization for Sequential I/O" (shows "INVALID") and vice versa because the stripe size is different. Change it back to the original setting and reset the controller to make available the logical drive data again.

IMPORTANT:
Changing the setting to "Optimization for Sequential I/O" or "Optimization for Random I/O" should be performed only when no logical drive exists. Otherwise, you will not be able to access the data in the logical drive later on.

3.5 Drive-Side SCSI Parameters

3.5.1 SCSI Motor Spin-up

When the power supply is unable to provide sufficient current for all the hard drives and controllers that are powered-up at the same time, spinning-up the hard drives serially is one of the best way of consuming lower power-up current.

By default, all hard drives will spin-up when powered-on. These hard drives can be configured so that all of them will not spin-up at power-on. There are 3 methods of spinning-up the hard drive's motor: Spin-up at power-on, Spin-up serially in random sequence or Spin-up by SCSI command. Please refer to the hard drive's user's manual for instructions on configuring the hard drive using the "Spin-up by SCSI command". The procedure for each brand/model of hard drive should vary.

Configure all the hard drives as above and enable "SCSI Motor Spin-Up" in Drive-Side SCSI Parameters. Power off all hard drives and controller, and power them on again. All the hard drives will not spin-up at this time. The controller will then spin-up the hard drives one by one at four seconds interval.

☞ **IMPORTANT:**
If the drives are configured as "Delay Motor Spin-up" or "Motor Spin-up in Random Sequence," some of these drives may not be ready yet for the controller to access when the system powers up. Increase the disk access delay time so that the controller will wait a longer time for the drive to be ready.

3.5.2 SCSI Reset at Power Up

By default, when the controller is powered up, it will send a SCSI bus reset command to the SCSI bus. When disabled, it will not send a SCSI bus reset command on the next power-up.

When connecting dual host computers to the same SCSI bus, the SCSI bus reset will interrupt all the read/write requests that are being performed. This may cause some operating systems or host computers to act abnormally. Disable the "SCSI Reset at Power-up" to avoid this situation.

3.5.3 Disk Access Delay Time

Sets the delay time before the controller tries to access the hard drives after power-on. The default is 15 seconds.

3.5.4 SCSI I/O Timeout

The "SCSI I/O Timeout" is the time interval that the controller waits for a drive to respond. If the controller attempts to read data from or write data to a drive, but the drive does not respond within the SCSI I/O timeout value, the drive will be judged to be a failed drive.

When the drive itself detects a media error while reading from the drive platter, it will retry the previous reading or recalibrate the head.

When the drive has encountered a bad block on the media, it has to reassign the bad block to another spare block. However, all of this takes time. The time to perform these operations can vary between different brands and models of drives.

During SCSI bus arbitration, a device with higher priority can utilize the bus first. A device with lower priority will sometimes get a SCSI I/O timeout when higher priority devices keep utilizing the bus.

The default setting for “SCSI I/O Timeout” is 7 seconds. It is highly recommended not to change this setting. Setting the timeout to a lower value will cause the controller to judge a drive as failed a drive is still retrying or while a drive is unable to arbitrate the SCSI bus. Setting the timeout to a greater value will cause the controller to keep waiting for a drive, and it may sometimes cause a host timeout.

3.5.5 Maximum Tag Count

The maximum numbers of tags that can be sent to each drive at the same time. A drive has a built-in cache that is used to sort all of the I/O requests (”Tags”) which are sent to the drive, allowing the drive to finish the requests faster. The cache size and maximum number of tags varies between different brands and models of drive. Using the default setting – “32” – is highly recommended. Changing the maximum tag count to “Disable” will cause the internal cache of the drive to be ignored (i.e., not used).

3.5.6 Periodic Drive Check Time

The “Periodic Drive Check Time” is an interval for the controller to check all of the drives that were on the SCSI bus at controller startup (a list of all the drives that were detected can be seen under “View and Edit SCSI Drives”). The default value is “Disabled”. “Disabled” means that if a drive is removed from the bus, the controller will not be able to know – so long as no host accesses that drive. Changing the check time to any other value allows the controller to check – at the selected interval – all of the drives that are listed under “View and Edit SCSI Drives.” If any drive is then removed, the controller will be able to know – even if no host accesses that drive.

3.5.7 SAF-TE Enclosure Monitoring

What is SAF-TE?

SAF-TE stands for SCSI Accessed Fault-Tolerant Enclosures. It is an enclosure management technology that uses the SCSI bus to interact with the controller. A SAF-TE-compliant enclosure monitors the fan temperature, power supply, UPS and also provides drive status LED

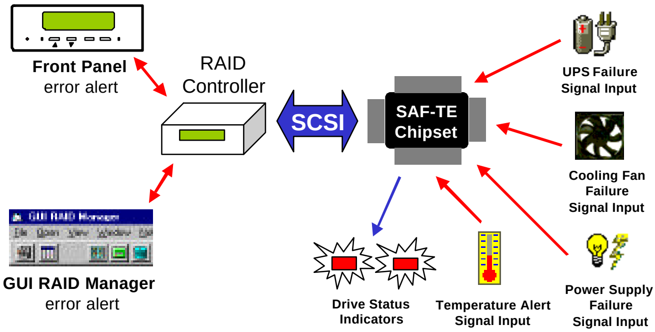
.

How does it work?

The SAF-TE device, which is often a back-plane within a drive-bay enclosure, must occupy a connector on one of the drive channels' SCSI cables. The presence of a SAF-TE device will be detected and its presence will be displayed in the BIOS configuration utility, Text RAID Manager and the GUI RAID Manager programs. The RAID controller communicates with the SAF-TE enclosure with standard SCSI commands, polling the device in order to get SAF-TE information.

The default value for "Periodic SAF-TE Device Check Time" is "Disabled". If the enclosure does have a SAF-TE device and features, enable the controller to poll the device by selecting a time interval. The RAID controller will then check the SAF-TE device status at that interval.

SAF-TE Support



- SAF-TE chipset connects to the drive channel of the controller together with the other SCSI drives.

3.5.8 Periodic Auto-Detect Failure Drive Swap Check Time

The "Drive-Swap Check Time" is the interval at which the controller checks to see whether a failed drive has been swapped. When a logical drive's member drive fails, the controller will detect the failed drive (at the selected time interval). Once the failed drive has been swapped with a drive that has adequate capacity to rebuild the logical drive, the rebuild will begin automatically.

The default setting is "Disabled," meaning that the controller will not Auto-Detect the swap of a failed drive. To enable this feature, select a time interval.

3.6 Dynamic Logical Drive Expansion

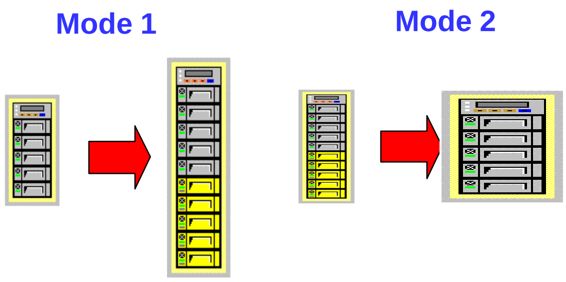
3.6.1 What Is It and How Does It Work?

Before Dynamic Logical Drive Expansion, increasing the capacity of a RAID system using traditional methods meant backing up, re-creating and then restoring. Dynamic Logical Drive Expansion allows users to add new SCSI hard disk drives and expand a RAID 0, 3 or 5 Logical Drive without powering down the system.

3.6.2 Two Modes of Dynamic Logical Drive Expansion

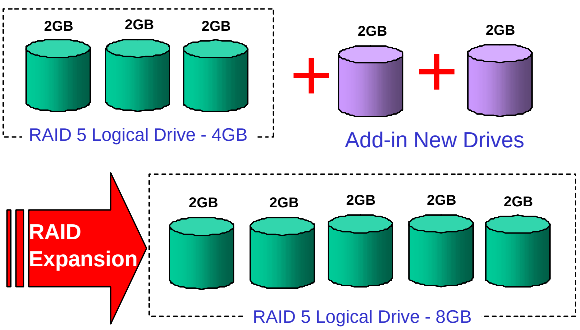
There are two modes of Dynamic Logical Drive Expansion: Mode 1 and Mode 2.

Dynamic Logical Drive Expansion



Mode 1 Expansion involves adding more SCSI hard disk drives to a logical drive, which may require that the user obtain an enclosure with more drive bays. The data will be re-striped onto the original and newly added disks.

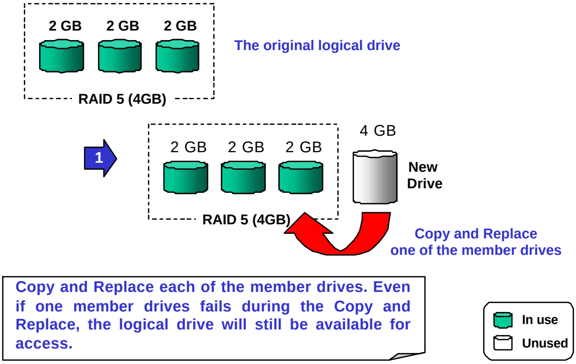
RAID Expansion - Mode 1



In the figure above, new drives are added to increase the capacity of a 4-Gigabyte RAID 5 logical drive. The two new drives increase the capacity to 8 Gigabytes.

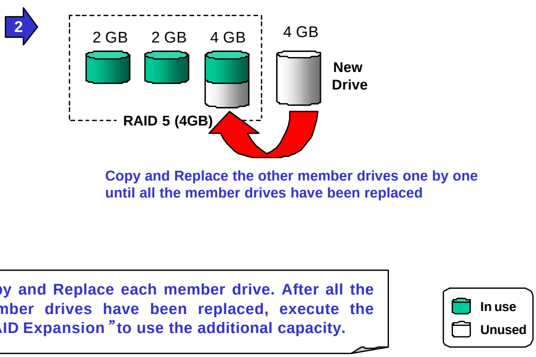
Mode 2 Expansion, on the other hand, requires the same number of higher-capacity SCSI hard disk drives for a given logical drive.

RAID Expansion - Mode 2 (1/3)

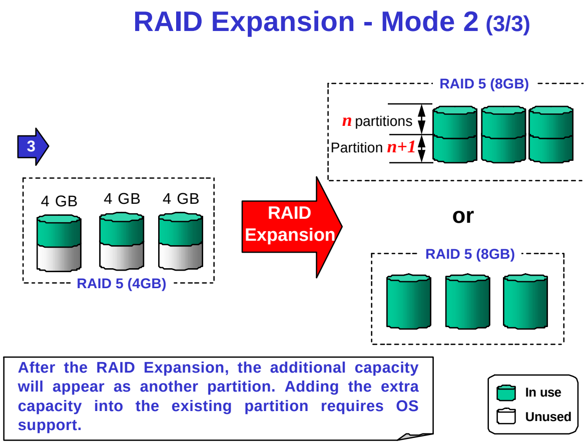


The figure above illustrates expansion of the same 4-Gigabyte RAID 5 logical drive using Mode 2 Expansion. Drives are copied and replaced, one by one, onto three higher-capacity drives.

RAID Expansion - Mode 2 (2/3)



This results in a new 4-Gigabyte, RAID 5 logical drive composed of three physical drives. The 4 Gigabytes of increased capacity is in a new partition.



IMPORTANT:

- *The increased capacity from Mode 1 Expansion of a logical drive will be a new partition.*
- *At the time of this printing, the firmware does not support the "Copy and Replace" function that is required for Mode 2 Expansion. Third-party hard disk utilities may be used for Mode 2 Expansion of logical drives. Later versions of the firmware will support "Copy and Replace."*

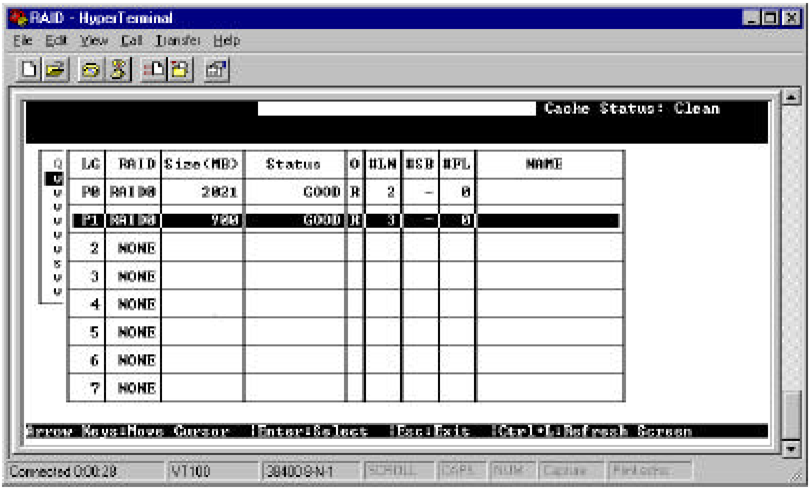
3.6.3 Example: RAID Expansion in Windows NT® Server

Limitations When Using Windows NT® 4.0

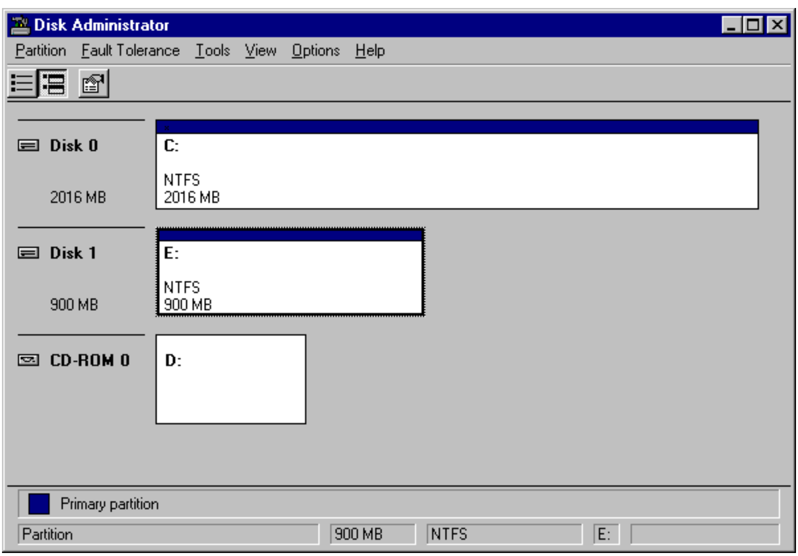
1. Only the Windows NT® Server Disk Administrator includes the Extend Volume Set function; Windows NT® Workstation does not have this feature.
2. The system drive (boot drive) of a Windows NT® system cannot be extended.
3. The drive that will be extended should be using the NTFS file system.

The Example:

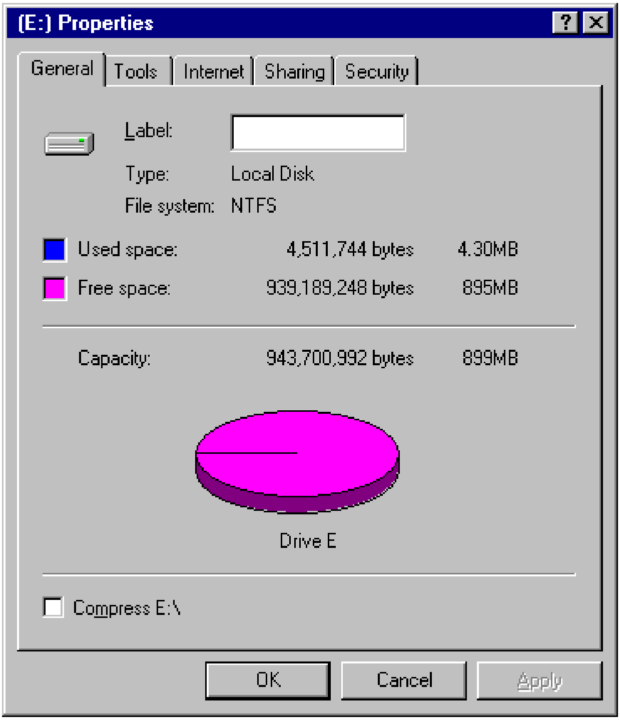
The following example demonstrates the expansion of a 900MB RAID 0 logical drive. The HyperTerminal emulation software that comes with Windows 95®/ Widnows NT® is used to connect to the RAID controller via RS-232.



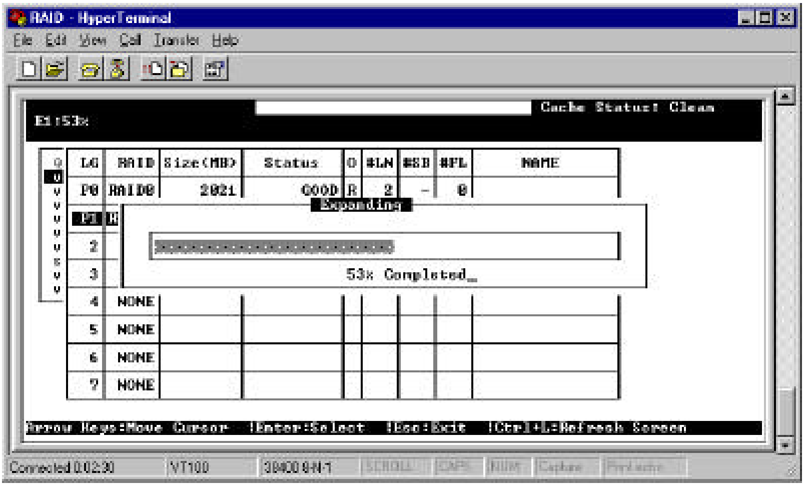
You can view information about this drive in the Windows NT® Server's Disk Administrator.



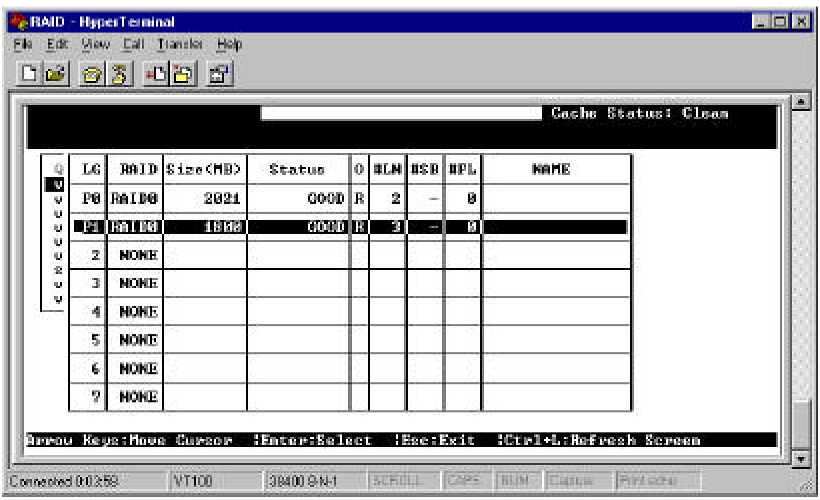
Place the cursor on Disk 1, right-click your mouse, and select "Properties." You will see that the total capacity for the Drive E: is just under 900MB.



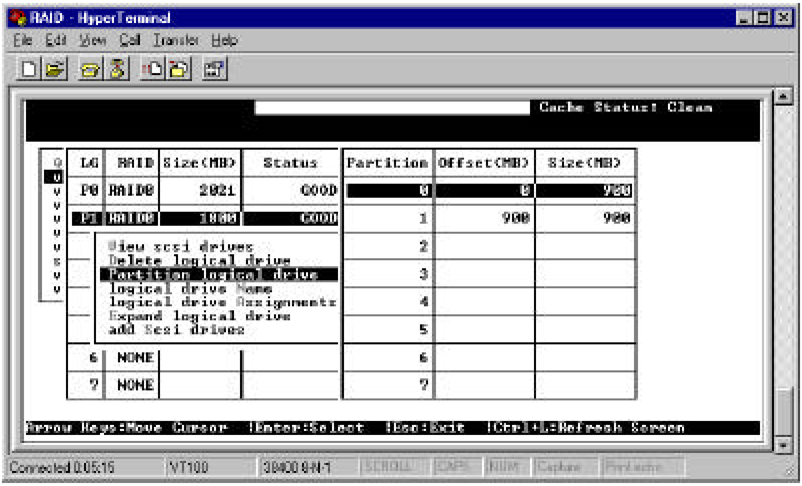
Follow the steps described in section 7.2.8 to add SCSI disk drives and perform Mode 1 Dynamic Logical Drive Expansion.



The 900MB logical drive has become a 1800MB logical drive. Place the cursor on that logical drive, and then press <Enter>.

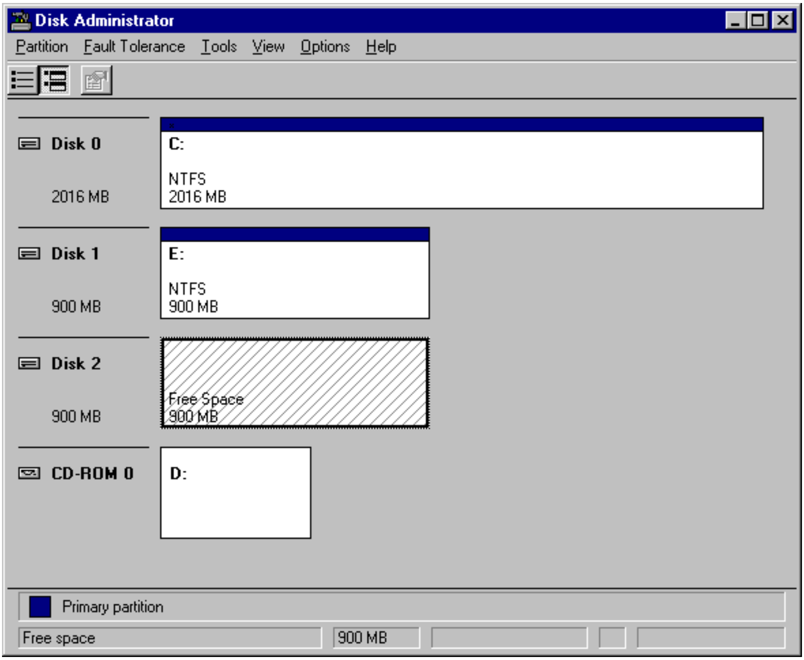


From the menu, select Partition Logical Drive. You will see that the 1800MB logical drive is composed of two 900MB partitions.

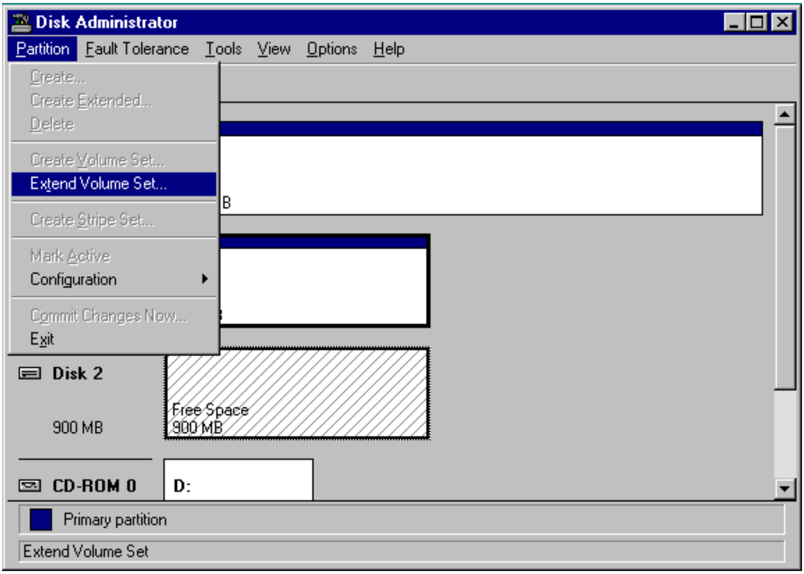


Follow the directions in section 7.3.1 to map the new partition to a Host LUN. The new partition must be mapped to a host LUN in order for the HBA (host-bus adapter) to see it. Once you have mapped the partition, reboot Windows NT®. The HBA should be able to detect an additional "disk."

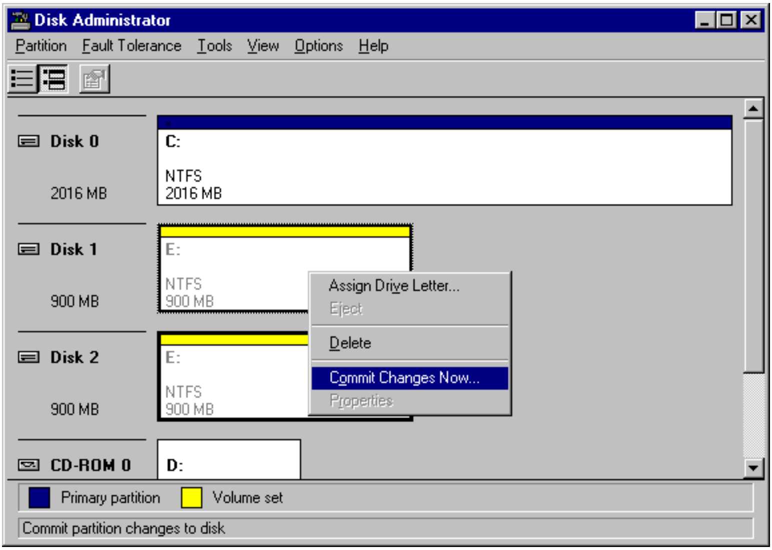
Return to Windows NT® Server's Disk Administrator. There now exists a Disk 2 with 900MB of free space. Click on Disk 2 to select it.



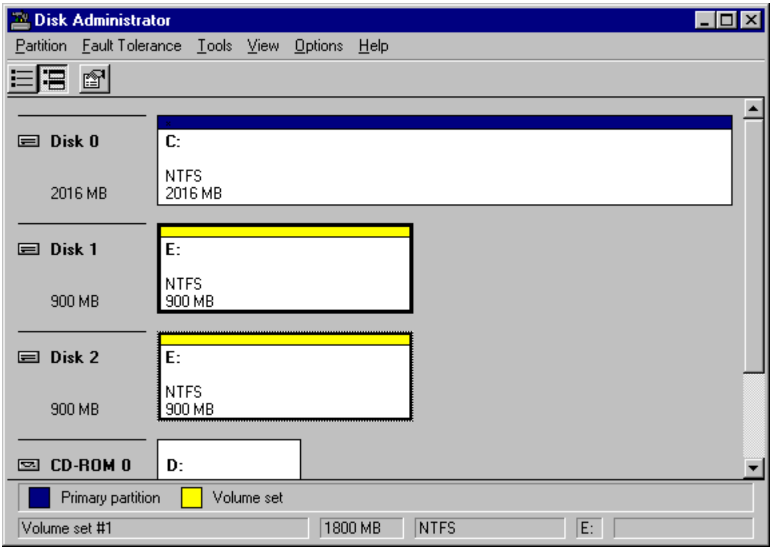
From the "Partition" menu, select "Extend Volume Set."



The screen will display that volume set of Drive E: has been extended by the 900MB in Disk2. Move the cursor to “Commit Changes Now” to confirm that you want the free space to become a part of the same logical drive.



Logical Drive E: is now composed of two 900MB partitions with a total volume of 1800MB. To see this, hold down on the <Ctrl> key and select both Disk 1 and Disk2; then right-click your mouse and select "Properties."



Drive E: now has a capacity just under 1800MB.