

1 Volume Management

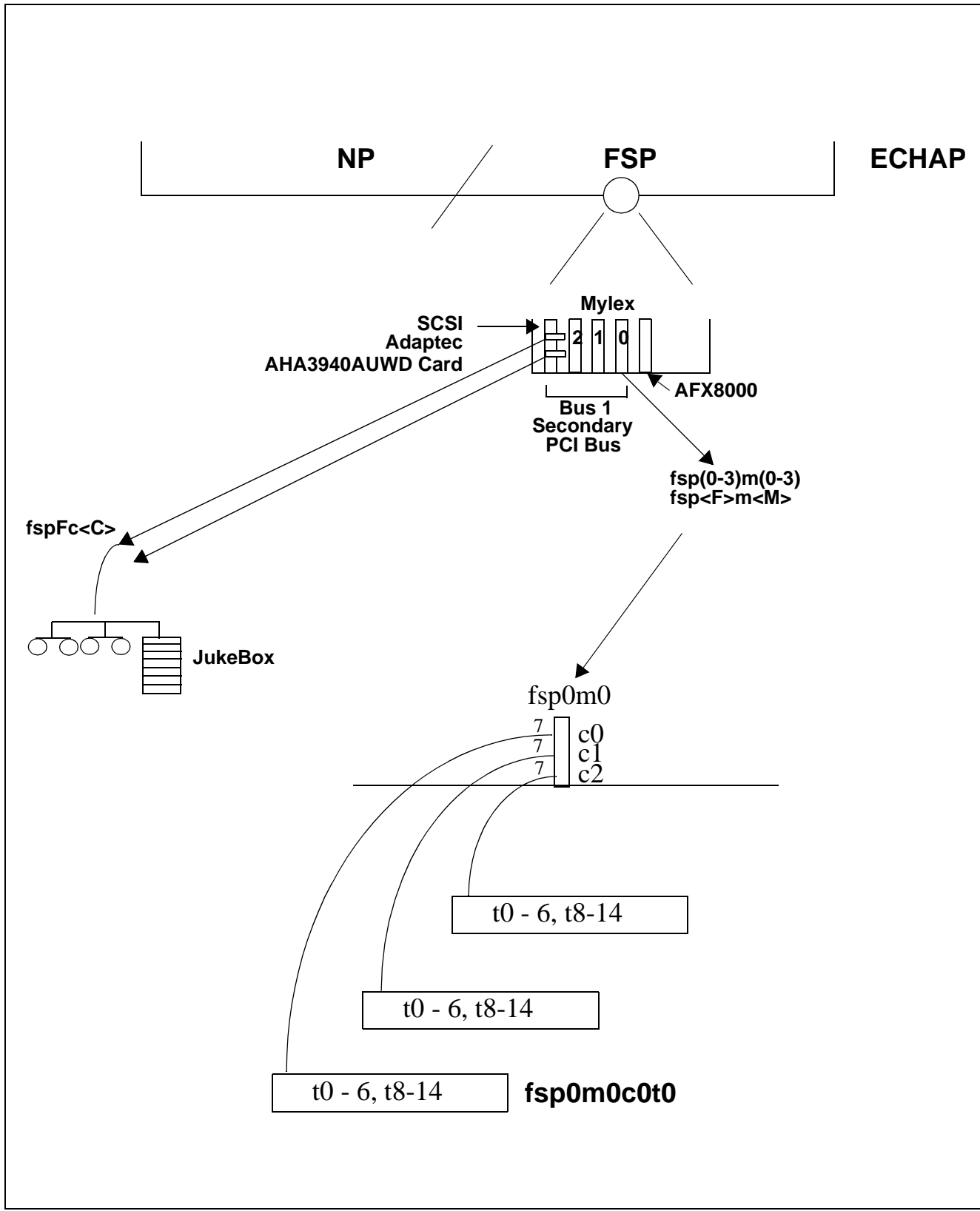
Introduction

This module covers raid array creation and management, and virtual partition creation and management.

Objectives

At the conclusion of this chapter, you will be able to do the following:

- ▲ Identify an Auspex SCSI device by its FSP, Mylex Controller, SCSCI Controller and drive number
- ▲ Use Auspex commands to create and manage raid arrays
- ▲ Use Auspex commands to create and manage virtual partitions
- ▲ Display online information about raid arrays



Component Names

Component Names

With use of the Mylex Controller, there is a change in how to access drives. The Mylex card will not export any drives that you could use directly. The nomenclature for various components in Mylex are as follows:

Mylex Controller:

`fsp<fsp_num>m<controller>`

Mylex Controller's SCSI channel:

`fsp<fsp_num>m<controller>c<scsi_channel>`

Mylex SCSI channel attached target

`fsp<fsp_num>m<controller>c<scsi_channel>t<target_ID>`

Examples:

A fully configured 10K system (a 3 node system) will have maximum of 3 FSPs and 9 Mylex Controllers. Numbering is zero based.

To access the 1st Mylex controller on the 1st FSP, the convention would look like this:

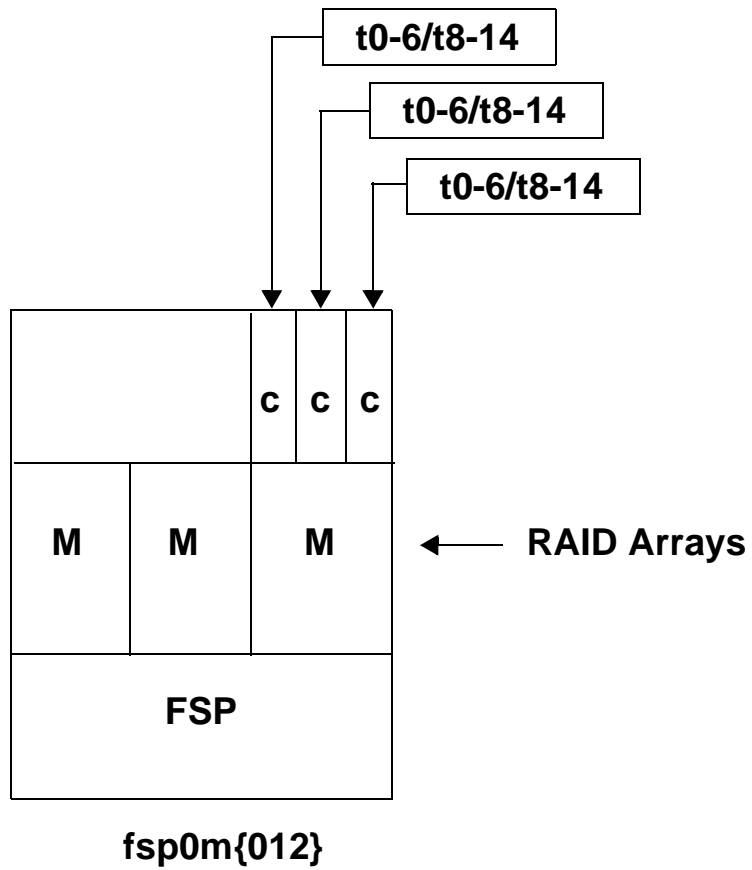
`fsp0m0`

To access the 3rd Mylex controller on the 1st FSP, the convention would look like this:

`fsp0m2`

To access the 3rd Mylex controller on the 3rd FSP, the convention would look like this:

`fsp2m2`



Volume Overview

Volume Overview

A volume is a named container for file systems or other forms of organized data.

A usable volume may be a RAID slice or a virtual partition comprised of RAID slices.

The first unit of a volume is a RAID array.

The utility to manage RAID array volumes is `ax_storage`.

RAID arrays are managed by the Mylex Controller, sometimes referred to as the RAID Controller.

RAID arrays are resident on the FSP.

The ax_storage Utility

This is the management utility for FSP resident RAID arrays

Used to create RAID0 (striped) arrays, RAID1 (mirrored) arrays or RAID5 arrays

A special RAID7(single drive) volume can be created

Must be root or super-user to run this command

The ax_storage Utility

Some of the options for the ax_storage command are as follows:

ax_storage scan <fsp>

ax_storage ldrv <fspm>

ax_storage lcdrv <fspm>

ax_storage create <fspm> <RAID0|1|5|7> [wb|wt] <cCtT>...

ax_storage create <fspm> <SPARE> <cCtT>

ax_storage stat [fsp|fspm]

ax_storage init <fspmrd>

ax_storage setmode <fspmrd> [mode=(wb|wt)]

ax_storage setparams <fspm> [stripe=8|16|32|64] [cache=8|16|32|64]

ax_storage rdparams <fspm>

ax_storage defaults <fspm>

ax_storage delete <fspmrd>

ax_storage verify <fspmrd> [pri=(lo|med|hi)]

ax_storage rebuild <fspmrd> [pri=(lo|med|hi)]

ax_storage abort <fspmrd>

ax_storage whatis <fspmct>

ax_storage erase <fspm>

Where:

fsp is the FSP number

m is the Mylex Controller number

rd is the raid array number

c is the SCSI controller number

t is the target device number

Listing Drives

Use the following command to list all available drives on controller fsp0m0:

```
ax_storage ldrv fsp0m0
```

This lists all drives, irrespective of the fact that it may be a member of an already defined RAID array

To list drives that are not part of an array and can potentially be used for building one, use the option "ldrv"

Listing Drives

When `ax_storage ldrv fsp0m0` is executed, the following output will be displayed:

```
# ax_storage ldrv fsp0m0
fsp0m0c0t0 [SEAGATE] [ST19171W S09C] [8.7G] [Configured]
fsp0m0c0t1 [SEAGATE] [ST19171W S09C] [8.7G] [Available]
fsp0m0c0t2 [SEAGATE] [ST19171W S09C] [8.7G] [Available]
fsp0m0c0t3 [SEAGATE] [ST19171W S09C] [8.7G] [Off-line]
fsp0m0c0t4 [SEAGATE] [ST19171W S09C] [8.7G] [Available]
fsp0m0c0t5 [SEAGATE] [ST19171W S09C] [8.7G] [Available]
fsp0m0c0t6 [SEAGATE] [ST19171W S09C] [8.7G] [Hot Spare]
```

Where:

fsp0m0c0t0 refers to the target disk

SEAGATE refers to the manufacturer of this device

ST19171WS09C refers to the manufacturer's part number

8.7G refers to the storage capacity of this disk

Configured means the disk is active(online) and is part of a RAID array

Available means the disk is available to be configured

Offline means the disk is configured but the drive is dead

Hot Spare means the disk is configured and part of a hot spare

This command lists all disk devices of a specified controller. Drives that are used in RAID arrays and active will be reported as Online. Drives that are not used and which can be used for defining a RAID array OR drives that are used in an array but down because of errors will be reported as Offline.

Listing Configurable Drives

To list all available drives on a specified controller, execute the following command:

```
ax_storage lcdrv fsp0m0
```

This lists all available devices for configuration on the Mylex Controller of the fsp in the base node

The drives listed can be used to create RAID arrays or spare devices

Listing Configurable Drives

When `ax_storage ldrv fsp0m0` is executed, the following output will be displayed:

```
# ax_storage ldrv fsp0m0  
fsp0m0c0t2 [QUANTUM] [QM318200] [17.4G] [Available]
```

Where:

fsp0m0c0t2 refers to target disk 2 on SCSI channel 0 on Mylex controller 0 of FSP0

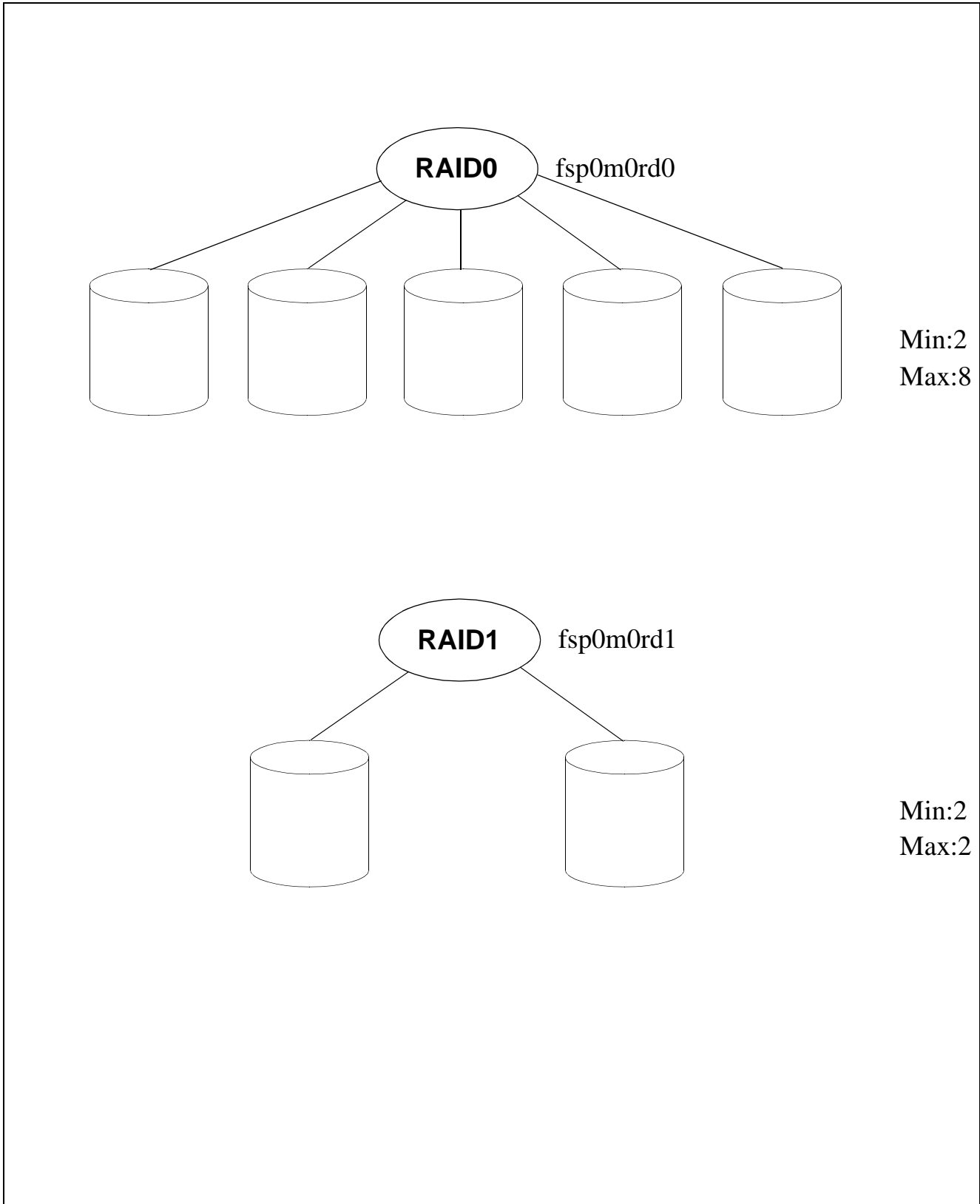
QUANTUM refers to the manufacturer of this device

QM318200 refers to the manufacturer's part number

17.4G refers to the storage capacity of this disk

Available means this disk is available for configuration

This command lists all configurable disk devices of a controller. The devices listed can actually be used to create system devices (RAID arrays) or spare devices. All drives will be reported as being Available, ready to be configured into RAID arrays or spares.



RAID Arrays

RAID Arrays

A maximum of 32 raid arrays are allowed per Mylex Controller on an FSP. RAID arrays are numbered from 0 through 31 on each controller.

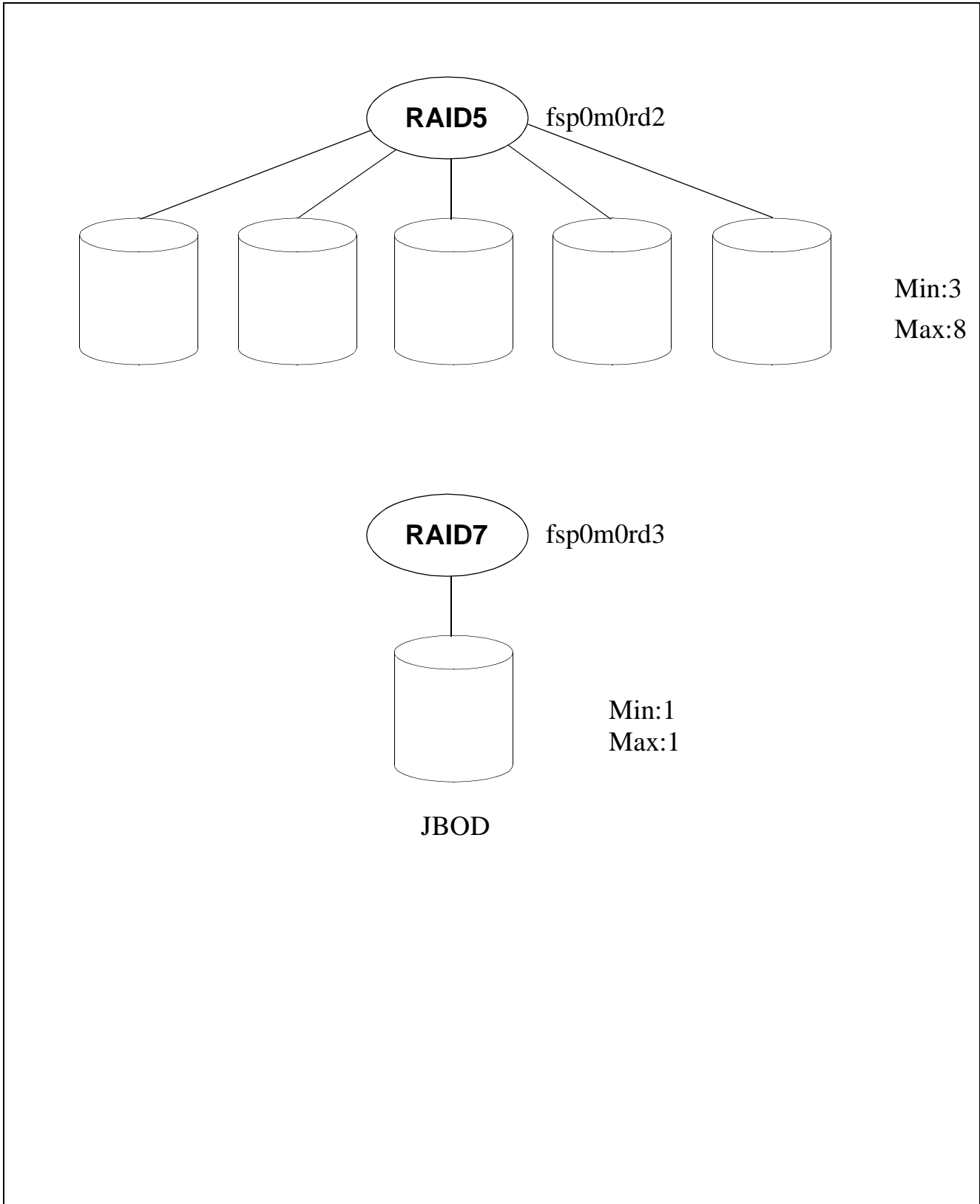
The types of RAID arrays supported are RAID0, RAID1, RAID5 and RAID7. RAID arrays reside on the Mylex Controller level and are resident on an FSP.

The following are the minimum and maximum members allowed to create a raid array:

Raid Level	Min	Max
0	2	8
1	2	2
5	3	8
7	1	1

RAID0 works with a minimum of 2 drives and features data striping. Striping data across disks increasing throughput with no capacity overhead. Typically used in applications where performance not fault tolerance is required.

RAID1 mirrors data across disks. If a drive fails on one side of the mirror, the data is available on the mirrored side, thus providing fault tolerance. RAID1 gives excellent read performance but poor write performance. Generally, RAID1 is used in smaller read intensive applications.



RAID Arrays

RAID Arrays

RAID5 features striping data and parity information. It provides fault tolerance. RAID5 gives excellent read and write performance with minimal overhead of one drive per drive group or RAID array. RAID5 is the most popular industry RAID and is optimized for network server applications.

RAID5 arrays must be initialized for usage, all other arrays do not need to be initialized.

RAID7 (JBOD or single drive) arrays are for single drives. The only way to use a plain disk is to define a RAID7 volume.

Accessing RAID Arrays:

To access the 1st RAID array on the 3rd Mylex controller of the 1st FSP, the convention would look like this:

`fsp0m2rd0`

Creating RAID Arrays

To create a RAID array, execute the following command:

```
ax_storage create fsp0m0 RAID0 wb c0t0 c1t0
```

This creates a RAID0 (stripe) array using drive c0t0 and c1t0 with the default write back caching on

This first array will be named as fsp0m0rd0

The option wb (default) is to enable write back caching and wt is for write through

Write back option will cache writes to disk and determines when to flush the data to disk

Write through will cache writes to disk and immediately flushes the data to disk

RAID arrays are stored in the file
/usr/AXbase/etc/raidtab

Creating RAID Arrays

Examples of ax_storage create:

```
% ax_storage create fsp0m1 RAID5 wb c0t0 c1t0 c2t0
```

Creates a RAID5 array using three drives; sets write-back mode meaning all i/o's will be cached on NVRAM using a write behind caching policy.

To build a RAID5 array, a minimum of 3 drives are required. Spreading the array across channels (c0t0, c1t0 and c2t0) will reduce SCSI channel contention. Using 1 drive per channel will prevent array failure due to loss of a single channel. As this is the first RAID array, it will be named as fsp0m0rd0.

Note: If at a later point in time, write through caching is desired, the option "setmode" can be used.

A RAID5 array will require initialization using the ax_storage init command.

```
% ax_storage create fsp0m0 RAID7 c2t2
```

Creates a RAID7(JBOD) to use a plain disk drive, c2t2.

Note: The only way to use a plain disk is to define a RAID7 array. This is different from the NS7000 systems, where a drive may directly be used after partitioning it.

```
% ax_storage create fsp0m0 SPARE c0t2
```

Marks c0t2 a hot SPARE - when a RAID array defined on fsp0m0 goes into degraded mode, this hot spare will be used automatically to rebuild the array.

Listing RAID Arrays

To display the status of all defined RAID arrays, execute:

```
ax_storage stat
```

Listing RAID Arrays

The output of the command:

```
# ax_storage stat
```

Current RAID configuration:

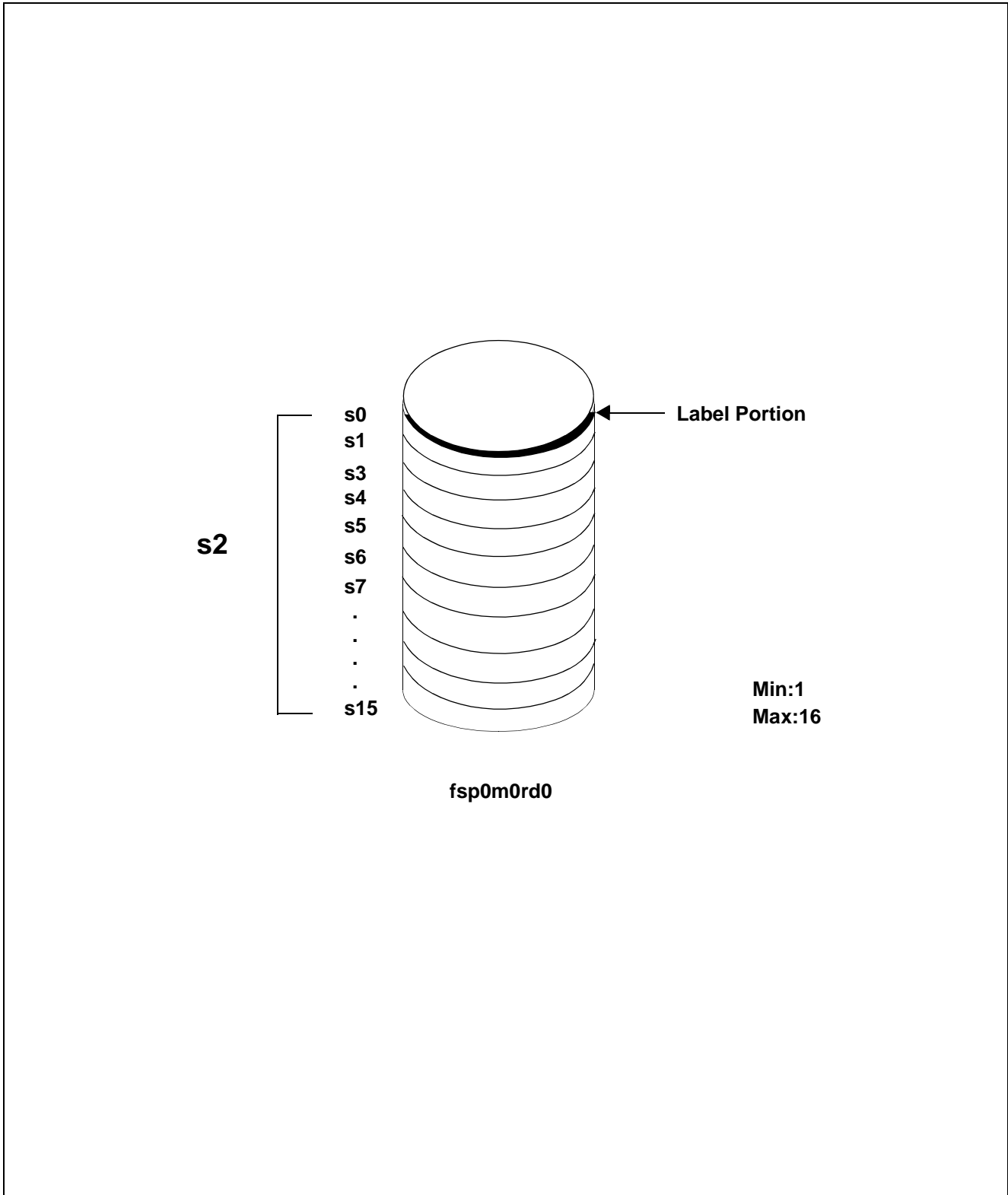
	Type	Main			Array			
		Prod	Camp	Size	State	Cache	Members	State
fsp0m0rd0	RAID0	Idle	0%	17.4G	Online	WB	fsp0m0c0t0	Online
							fsp0m0c1t0	Online
							fsp0m0c2t0	Online
fsp0m0rd1	RAID 1	Idle	0%	17.4G	Online	WB	fsp0m0c0t3	Online
							fsp0m0c1t3	Online
fsp0m0rd2	RAID 5	Init	45%	17.4G	Online	WB	fsp0m0c0t4	Online
							fsp0m0c1t4	Online
							fsp0m0c2t4	Online
fsp0m0rd3	RAID 7	Idle	0%	17.4G	Online	WB	fsp0m0c2t2	Online

Note that this command also lists any maintenance processes, completion percentage of that maintenance process, size of the array, state, caching policy and members of that array/volume.

The `ax_storage stat` output indicates four active RAID arrays, `fsp0m0rd0` through `fsp0m0rd3`.

Now that there are 4 RAID arrays/volumes, these arrays need to be partitioned in order to be available for use. Partitioning the RAID arrays is achieved using the `ax_label` command.

To view the status of SPARE arrays, use the command `ax_storage lsdrv`.



Labelling (using ax_label)

Labelling RAID Arrays

After RAID arrays have been defined, they need to be partitioned or sliced.

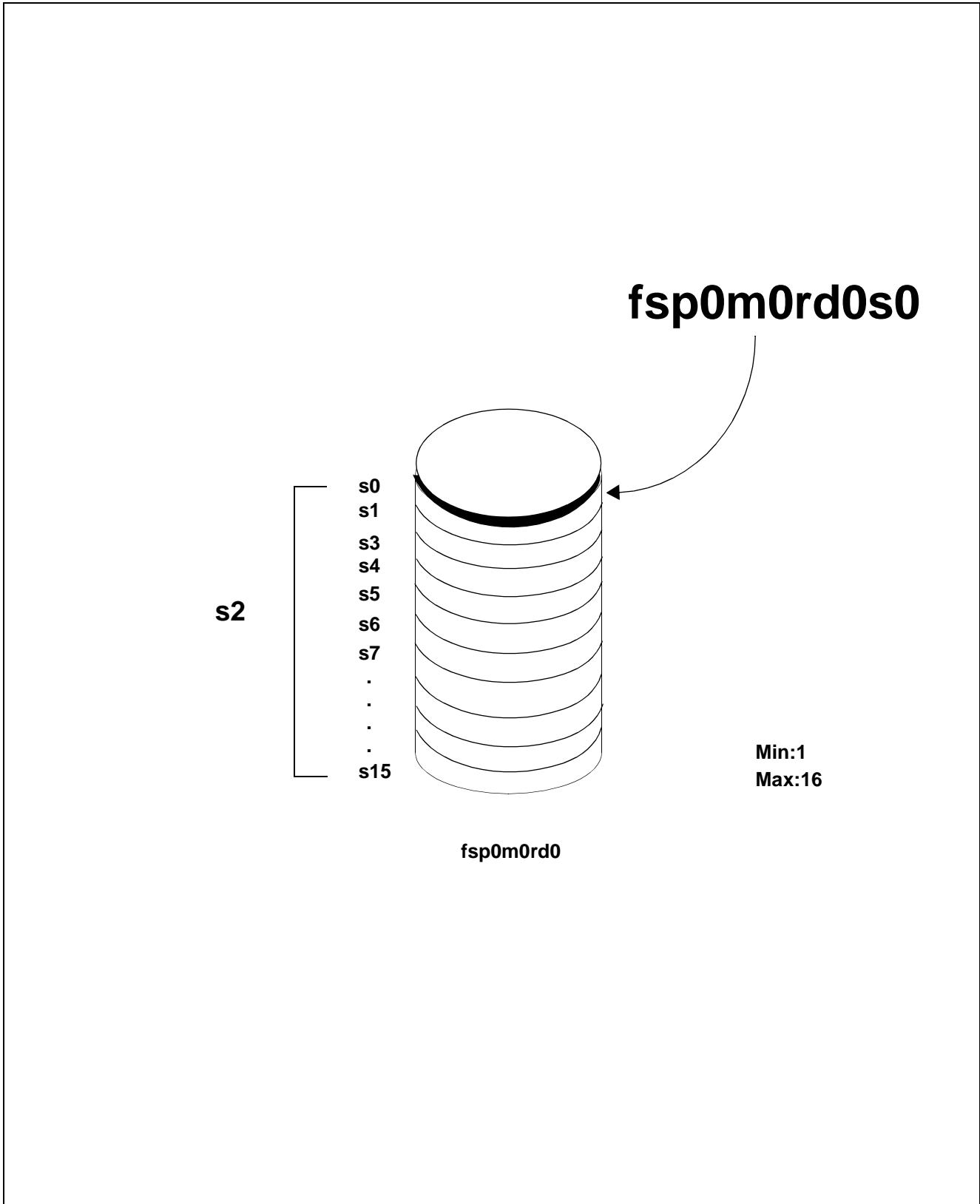
To define slices on a RAID array and write a label for a drive or a RAID array, execute the command:

```
ax_label fsp0m0rd0 -2 all
```

where rd0 refers to the first RAID array and -2 refers to the entire disk and all refers to all remaining sectors in the drive

The ax_label command can define up to 16 slices on a RAID array and writes the information onto the label sector (sector 0) of the drive or array. These changes are recorded in the FSP. After labeling, the slices are available for use.

The basic slicing unit is the MB, which is defined as 2048 512-byte sectors. Each physical slice, 0-15, may contain some integral number of MBs. Slice 2, by convention, is the entire volume, and is defined automatically. Use the ax_lslabel utility to view the label, after using ax_label.



Labelling RAID Arrays

Labelling Raid Arrays

Examples:

```
ax_label fsp0m0rd0 -0 1000 -1 1000 -3 1000 -4 1000 -A 8000
```

Partitions the RAID array fsp0m0rd0 into 5 slices, s5 being the largest slice of size 8000 MB

```
ax_label -0 400 -1 400 -3 400 -4 400 -A 400 -B 400 fsp0m0rd0
```

Partitions the RAID array fsp0m0rd0 with 6 400 MB slices. After labelling, these slices may be available for use under /dev/raxmrd (raw) and /dev/axmrd (block). For example, slice 0 may be accessed by opening /dev/raxmrd/fsp0m0rd0s0 and slice A may be accessed by opening /dev/raxmrd/fsp0m0rd0s10.

```
ax_label -R fsp0m0rd1
```

For RAID arrays only. Reads the /usr/AXbase/etc/raidlabel.map and sets an Auspex-recommended default partition map on it.

The device fsp0m0rd0s1 is a usable, mountable device.

Device Nodes

When you label any RAID arrays, device nodes are created under:

```
/dev/axmrd - block
```

```
/dev/raxmrd - raw
```

After creation of labels on fsp0m0rd0, the block nodes look like the following:

```
/dev/axmrd/fsp0m0rd0s0
```

```
/dev/axmrd/fsp0m0rd0s1
```

The raw nodes look like the following:

```
/dev/raxmrd/fsp0m0rd0s0
```

```
/dev/raxmrd/fsp0m0rd0s1
```

Listing Labels

To list the partition/slicing information on fsp0m0rd0 execute the command:

```
ax_lslabel fsp0m0rd0
```


Listing Labels

Examples:

```
ax_lslabel -w fsp0m0rd0
```

Legend: m - mounted or part of a mounted VP

v - part of a VP

u - slice unusable: slice 2 is in use

x - slice 2 excluded: slices in use

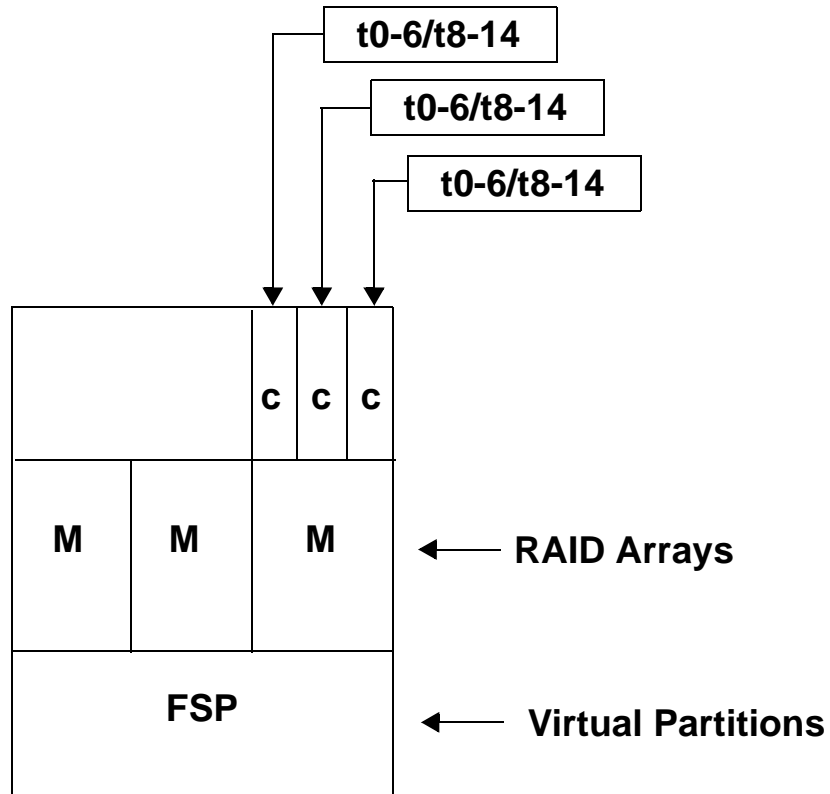
```
fsp0m0rd0s{2mv}
```

Where:

2 refers to the slice number

m refers to a slice as a mounted virtual partition

v refers to the slice as part of a virtual partition



Virtual Partition Overview

Virtual Partition Overview

Virtual Partitions reside on the FSP.

Virtual partitions are constructed either by concatenating or striping together one or more slices to make a larger partition, or by mirroring concatenated or striped virtual partitions.

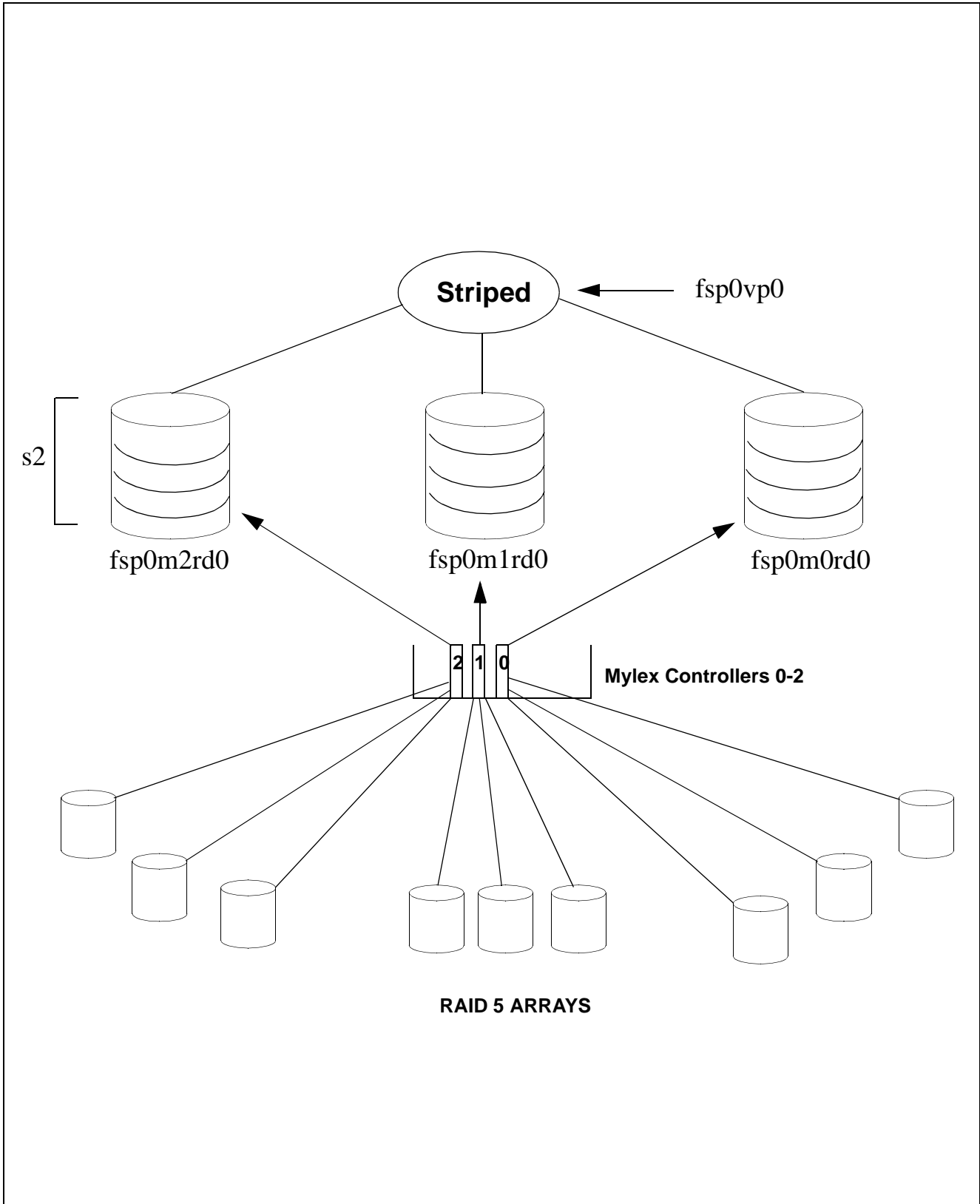
The maximum number of virtual partitions per FSP is 256.

Virtual partitions on each FSP may be numbered from 0 to 255.

All members of a virtual partition must be on the same FSP.

A virtual partition can have a maximum of 16 member disks/slices.

The path to the virtual partition table is `/usr/AXbase/etc/vpartab`.



Virtual Partitions

Virtual Partitions

A concatenated virtual partition can have a minimum of 1 member disk

A striped virtual partition can have a minimum of 2 member disks

A mirror can have a minimum of 1 member and a maximum of 2 member disks

A mirror can NOT have a physical disk and/or RAID as its direct member

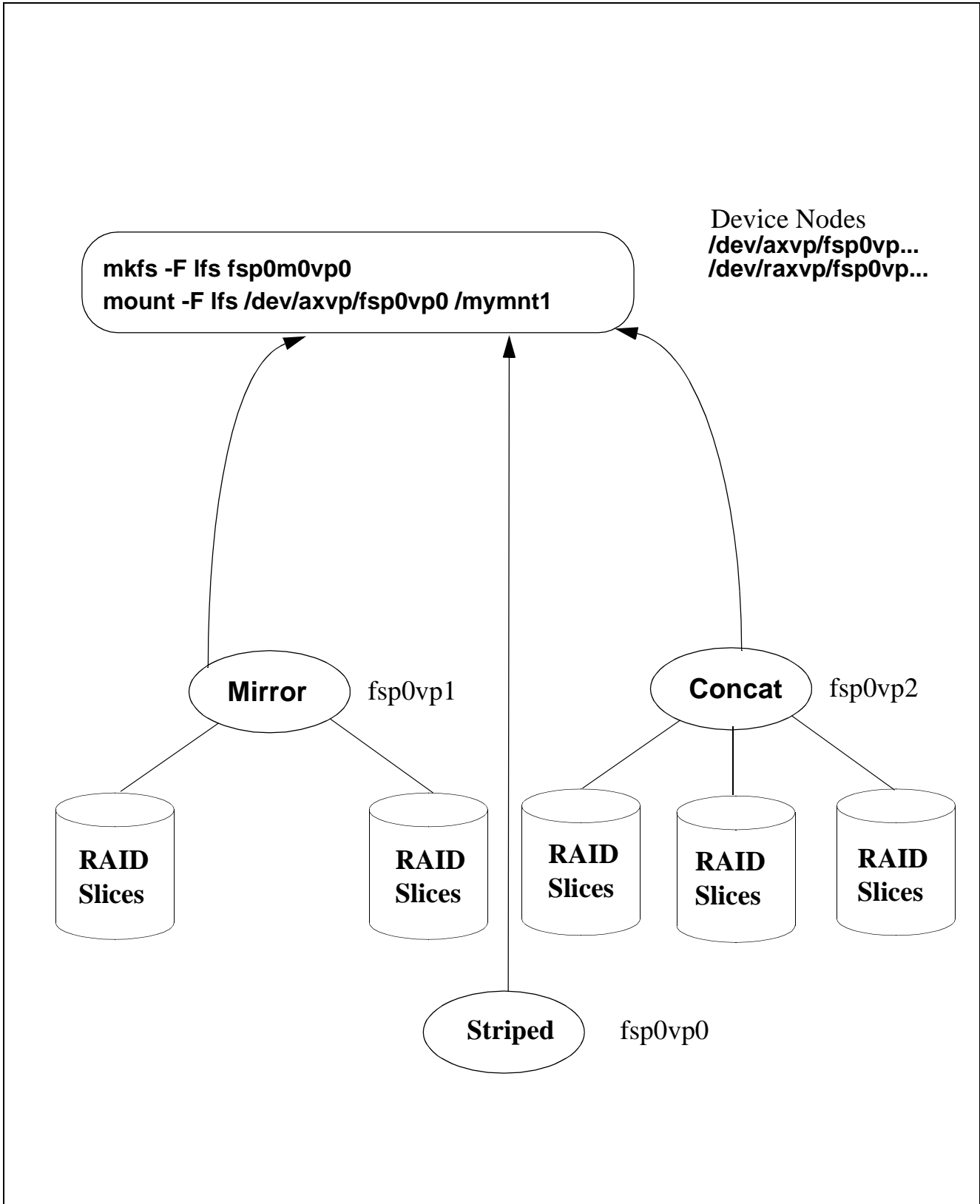
A mirror must have a virtual partition disk as its child member

A concatenated or striped virtual partition can be made up of physical disk slices or RAID slices

To access the 1st virtual partition on the 1st FSP, the convention would look like this:

fsp0vp0

Note: No Mylex Controller is referenced in the above example because virtual partitions are resident on the FSP.



Virtual Partitions

The Virtual Partition Table

Examples of entries in the /usr/AXbase/etc/vpartab file:

fsp1vp0 striped,size=128K fsp1m0rd0s2, fsp1m1rd0s2, fsp1m2rd0s2

Vp0 stripes three RAID arrays defined across three controllers, m0, m1 and m2 on FSP number 1. This is an example of a really large volume spanning across three controllers.

fsp0vp3 mirrored fsp0vp5 fsp0vp6

Vp6 defines a Mirrored VP - This mirrors two RAID5 arrays defined on two different controller cards. This is the ultimate fault-tolerance that one can get - RAID5 guards against disk failure. The mirror guards against controller failure.

fsp0vp2 concat fsp0m0rd1s2, fsp0m0rd2s0, fsp0m0rd3s2

Vp2 defines a concatenated virtual partition consisting of RAID5 slices and full partitions.

The `ax_loadvpar` Command

To download the virtual partition definitions to the FSP, execute:

```
ax_loadvpar
```


The ax_loadvpar Command

The ax_loadvpar reads the virtual partition table from /usr/AXbase/etc/vpartab and transmits that information via messaging to the appropriate FSP on which the defined virtual partitions are to be created

The command is in the form:

ax_loadvpar

To confirm that a virtual partition exists, execute the ax_vpstat command. This command will display information about all virtual partitions across all FSPs.

Adding and Removing SCSI Devices

To add, remove or replace a device, execute the commands:

```
ax_hotplug add
```

```
ax_hotplug remove
```

```
ax_hotplug replace
```

```
ax_hotplug recable
```

Adding and Removing SCSI Devices

The `ax_hotplug` command quiesces an FSP resident SCSI bus, so no operations are in progress on the particular SCSI bus and prompts the user to insert, remove, or replace a device on the SCSI bus, or recable the SCSI bus. It may spin down the device if it is to be removed.

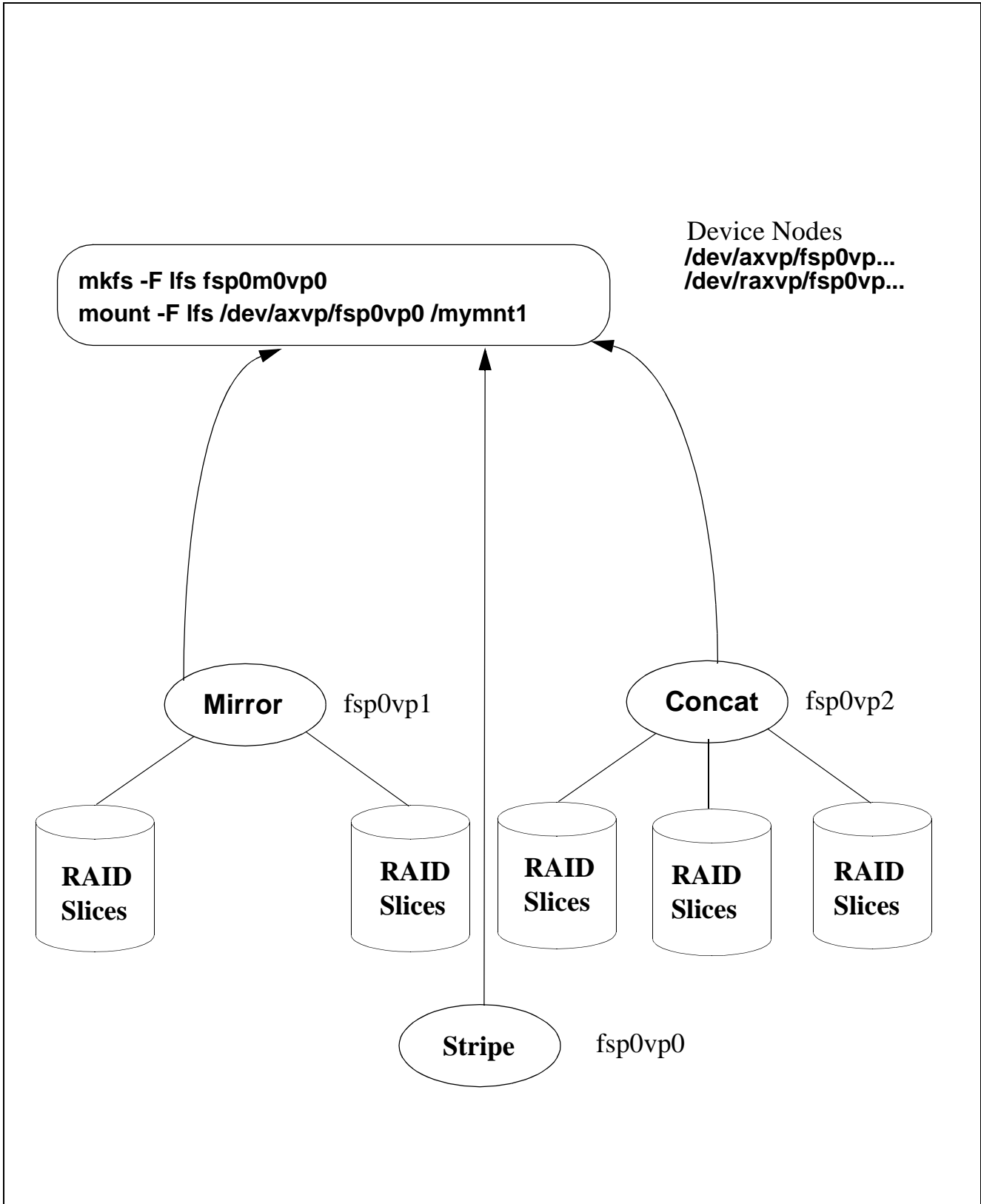
Example:

```
#ax_hotplug remove fsp0m0c0
```

You may remove fsp0m0c0 {0 1 2}

```
#ax_hotplug add fsp0m0c0
```

You may add fsp0m0c0 {3 4 5 6 8 9 10 11 12 13 14 15}



Slidetitle