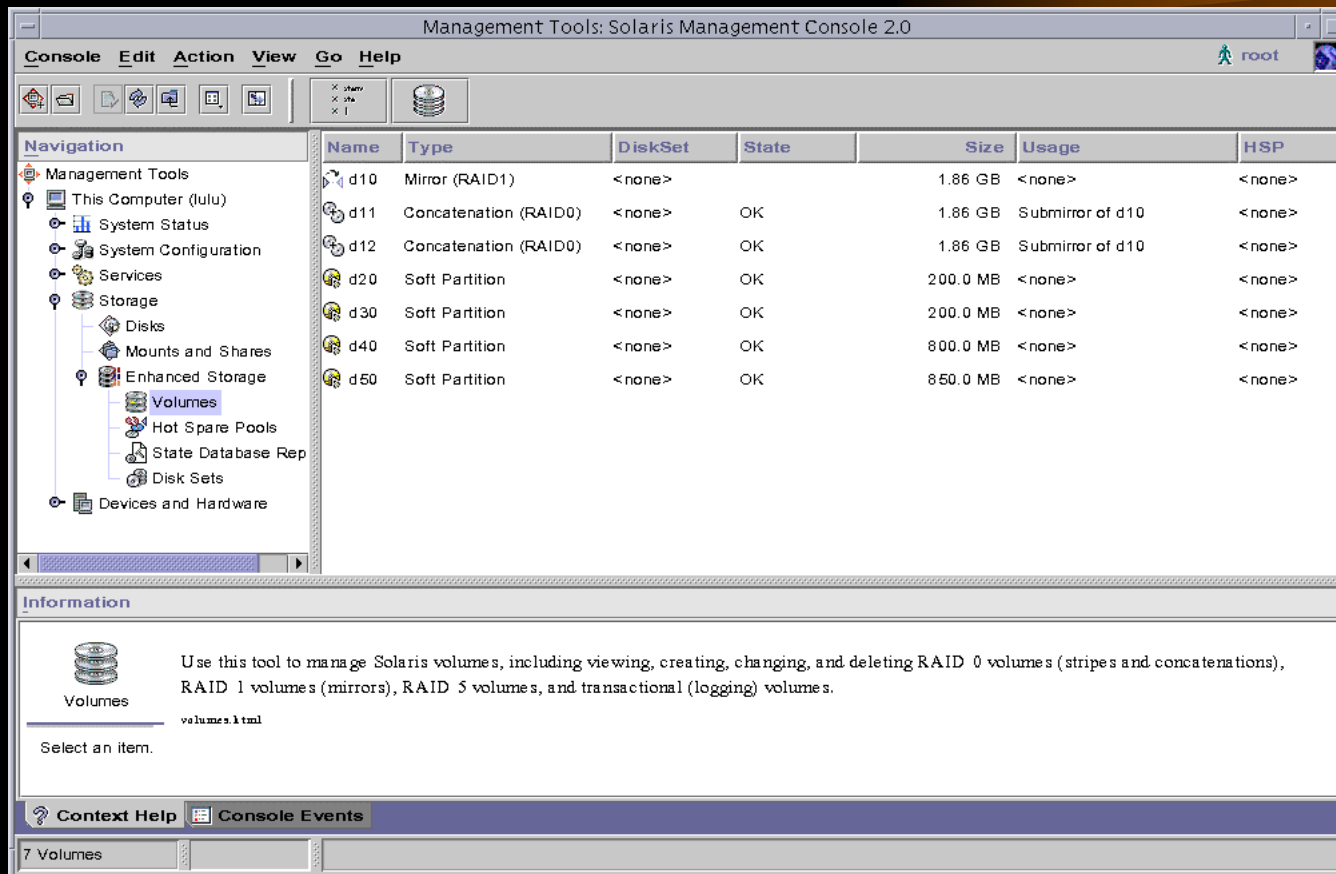


Solaris Volume Manager



The screenshot displays the 'Management Tools: Solaris Management Console 2.0' window. The interface includes a menu bar (Console, Edit, Action, View, Go, Help), a toolbar, and a navigation pane on the left. The navigation pane shows a tree structure with 'Volumes' selected under 'Enhanced Storage'. The main pane displays a table of volumes with columns: Name, Type, DiskSet, State, Size, Usage, and HSP.

Name	Type	DiskSet	State	Size	Usage	HSP
d10	Mirror (RAID1)	<none>		1.86 GB	<none>	<none>
d11	Concatenation (RAID0)	<none>	OK	1.86 GB	Submirror of d10	<none>
d12	Concatenation (RAID0)	<none>	OK	1.86 GB	Submirror of d10	<none>
d20	Soft Partition	<none>	OK	200.0 MB	<none>	<none>
d30	Soft Partition	<none>	OK	200.0 MB	<none>	<none>
d40	Soft Partition	<none>	OK	800.0 MB	<none>	<none>
d50	Soft Partition	<none>	OK	850.0 MB	<none>	<none>

Below the table, the 'Information' pane provides instructions: 'Use this tool to manage Solaris volumes, including viewing, creating, changing, and deleting RAID 0 volumes (stripes and concatenations), RAID 1 volumes (mirrors), RAID 5 volumes, and transactional (logging) volumes.' It also includes a 'Select an item.' prompt and a 'Context Help' button.

Solaris Volume Manager



- SVM uses virtual disks to manage physical disks.
 - SVM virtual disk called volume (metadevice).
- Manage large number of disks and the data on those disks.
 - Increasing storage capacity.
 - Increasing data availability.
 - Improve I/O performance.
 - Easing administration of large storage device.
 - SMC – Solaris Management Console (Java)
- Solstice DiskSuite 4.2.1

State Databases (Replicas)

- Keep track of configuration and status information about volumes.
- SVM maintains multiple copies (replicas) of the state database (redundant).
- Half+1 rule.
 - Minimum three state replicas.
- Use dedicated area of disk.
- Setup using metadb command.

Solaris Volume Manager Features



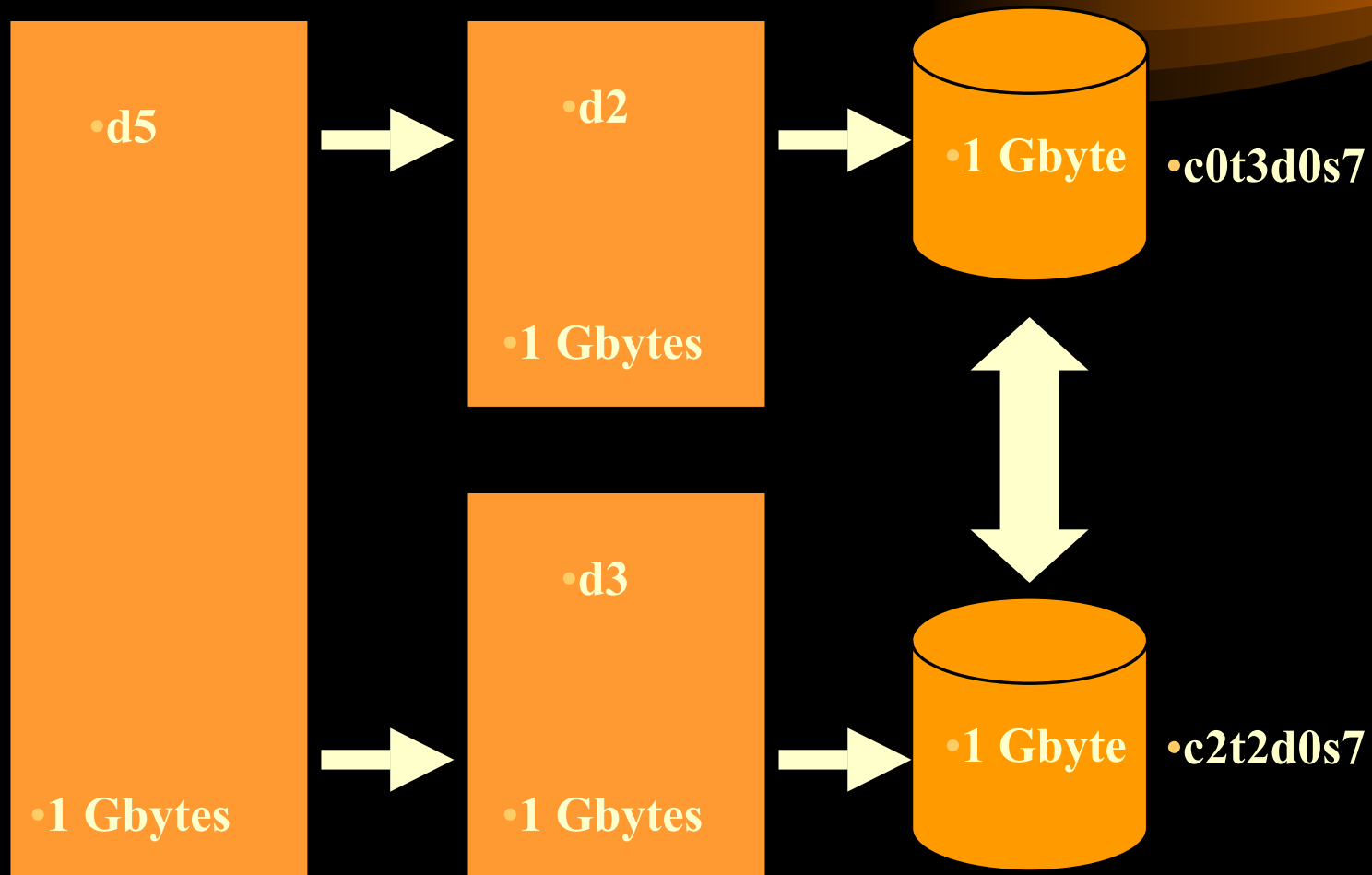
- Disk Mirroring
- Disk Concatenation
- Disk Striping
- RAID 5
- Disksets
- Hot Spares
- Soft Partition
- growfs

Disk Mirroring



- A feature used to guard against component failure.
- Any file system can be mirrored, including root, swap, and /usr.
- Required twice disk capacity.

Mirroring

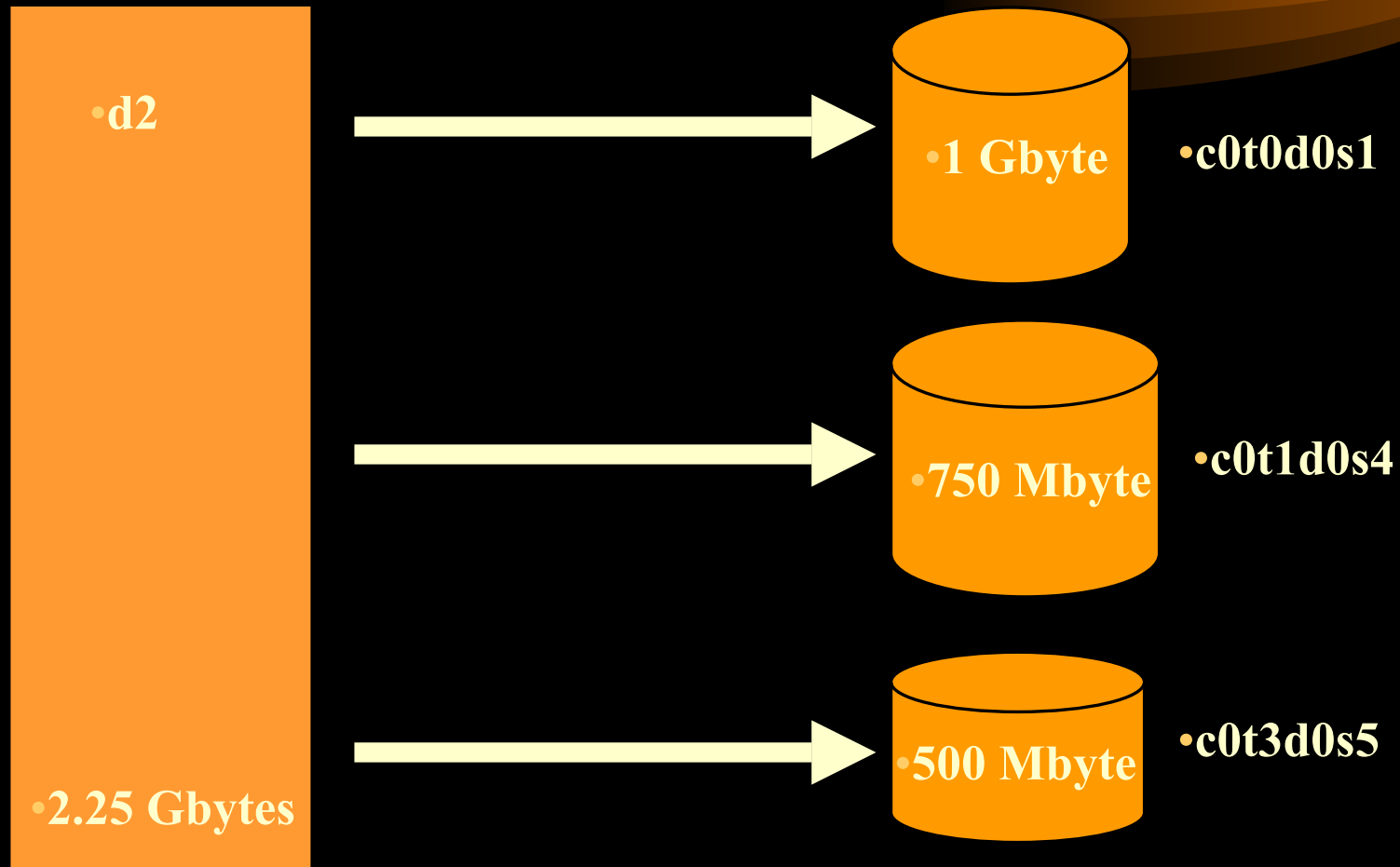


Disk Concatenation



- It combines multiple physical partitions to create a large volume.
- The data is then spread over multiple disks and multiple partitions.

Concatenation

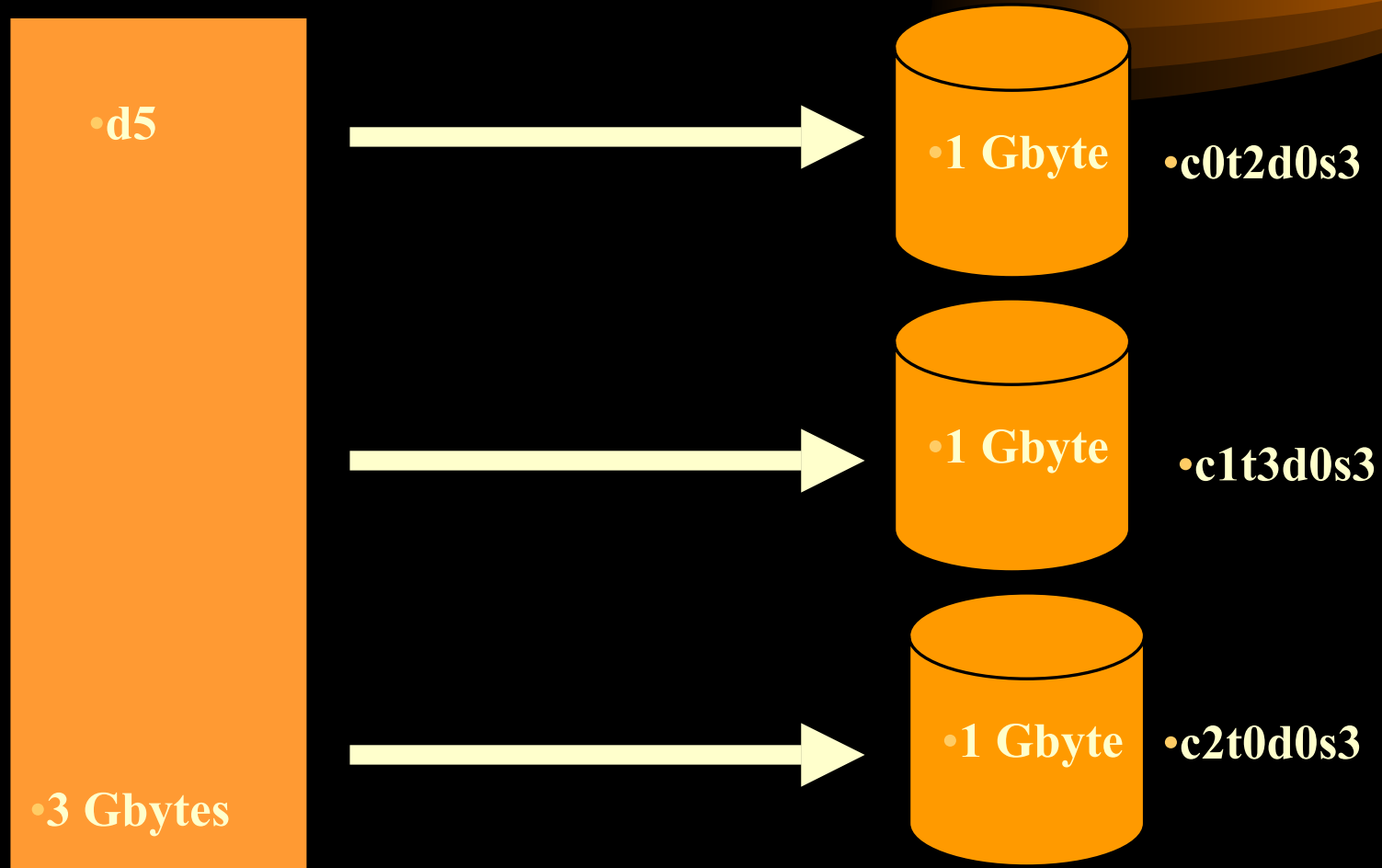


Disk Striping



- A feature used to increase read and write performance by spreading data requests over multiple components.

Striping

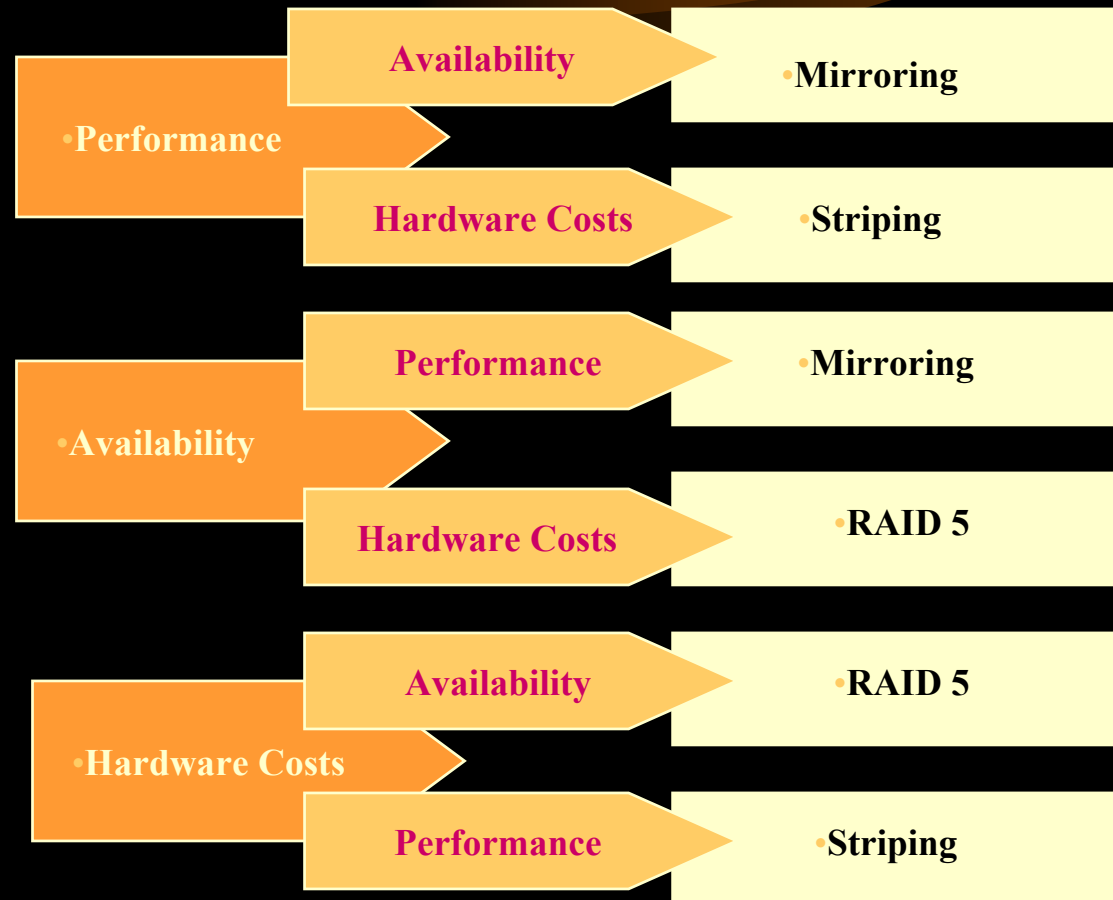


RAID Levels

- RAID 0 stripe
- RAID 1 mirror
- RAID 0+1 stripes that are then mirrored
- RAID 1+0 mirrors that are then striped
- RAID 5 stripe with parity

High Level Decision Ttree

- Use the following decision tree to determine the best fit for SVM in your environment.
- Step 1.
 - Choose the factor most important to you.
- Step 2.
 - Choose the next most important factor.
- Step 3.
 - For your needs, the best SVM feature is ...



Diskset

- A grouping of two hosts and disk drives in which all the drives are accessible by both hosts.
- This feature provides facilities for hosts to share disks in a high availability environment.



Setting Up DiskSet



- Two Steps:
 - Adding hosts to a diskset (this create the diskset).
 - Adding drives to a diskset.
- Before creating a diskset:
 - Configure disk drive device names so they have the same name on each host in the diskset.
 - Connect drives to be shared to both hosts.
 - Configure local metadvice state replicas on each host.

Setting Up DiskSet



- Create the diskset by defining the host(s) that have access to the devices.
 - # metaset -s share-dg -a -h red blue
 - where -s **setname** specifies the diskset name
 - a adds hosts to the named disksets (max 2 hosts).
 - h host... one or more hosts to be add to a diskset.
- Adding drives to a diskset
 - # metaset -s share-dg -a c1t2d0 c1t3d0 c1t4d0

DiskSet Status



- Display diskset information
 - metaset
 - Set name = share-dg, Set number = 1
- Host
 - red
 - blue
- Drive
 - c1t2d0
 - c1t2d0
 - c1t3d0
- Owner
 - Yes
- Dbase
 - Yes
 - Yes
 - Yes

Releasing and Taking a DiskSet



- Release (deport)
 - `metaset -s share-dg -r`
- Take (Import)
 - `metaset -s share-dg -t [-f]`
 - `f` taking a diskset forcibly

Hot Spare Pool



- A collection of slices that SVM uses to provide increased data availability for RAID 1 (mirror) and RAID 5 volumes.
- A hot spare reserved by SVM to be automatically substituted in case of a slice failure.

Working with Hot Spare Pool



- Create hot spare pool
 - metainit hsp001 c2t2d0s2 c3t2d0s2
 - hsp001: Hotspare pool is setup
- Associate hot spare pool with RAID 5
 - metaparam -h hsp001 d10
- Associate hot spare pool with sub mirrors
 - metaparams -h hsp001 d21
 - metaparams -h hsp001 d22

Soft Partitions



- A feature used to subdivide disks or logical volumes into more than 8 sections.
- Can be placed directly above disk slice, or on top of a mirror, stripe or RAID5 volume.

Create a Soft Partitions

- `metainit soft-partition -p [-e] component size`
 - `p` soft partition
 - `e` entire disk
 - `slice 0` data
 - `slice7` state data base replica
 - `soft-partition` the name of the soft partition.
 - `Component` disk, slice or logical volume.
 - `Size` space to take for the soft partition

growfs



- Expands a UFS file system without loss of service or data.
- It is used on line (hot) and it buffers or caches write requests as it running and completes them after the growfs utility complete.

SVM Configuration Files



- `/etc/lvm/md.tab`
 - manually configuration file that you edit,
 - read by metainit.
- `/etc/lvm/md.cf`
 - configuration information file,
 - when any SVM configuration changed, SVM automatically updates the md.cf.
- `/etc/lvm/mddb.cf`
 - record the locations of state database replicas.

SVM Configuration Files



- `/kernel/drv/md.conf`
 - read by SVM at the startup.
 - `nmd` : number of volumes (metadevices)
 - default 128 up to 8192.
 - `md_nsets` : number of disk sets.
 - default 4 up to 32.
- DO NOT EDIT OR REMOVE `md.cf` & `mddb.cf` !
- Backup the `md.cf` & `mddb.cf` on regular basis !

SVM Guidelines (Tips)



- State Database Replica guidelines
- Striping guidelines
- Concatenation guidelines
- Mirror guidelines
- RAID5 guidelines
- Hot Spares guidelines
- DRP guidelines

State Databases Replicas Guidelines

- Use the following estimates for stat database replicas based on the number of hard drive.

<u>Number of Hard Drive</u>	<u>Number of Replicas</u>
– One	Three, all on on slice
– Two-Four	Two on each drive
– Five or more	One on each drive
- Use dedicated slice,
 - when a state database replica is placed on a slice that become part of a metadvice, the capacity of the metadvice is reduced by the space occupied by the replica(s).

Striping Guidelines



- Do not use slices from disk with different disk geometry.
- Use slices on the same controller but different disks.
- Using stripes that are each on different controllers can increase the number of simultaneous reads and writes that can be perform.
- Do not stripe partitions that are on a single disk.
- Use the same size disk components. Striping different size disk components result in unused disk space.
- If you use slices of different sizes when striping, then disk capacity is limited to a multiple of the smallest.

Concatenation Guidelines



- Avoid using slice different disk geometry.
- When possible, distribute the components of a concatenated metadvice across different controllers and buses.
- Concatenation uses less CPU cycles than striping, It performs well for small random I/O distribution.

Mirror Guidelines



- Disk and controllers
 - keep the slices of different sub mirrors on different disks and controllers, redundant.
- Read/Write performance
 - Mirroring may improve read performance, but write performance is always degraded.
- Same-size sub mirrors
 - use the same size sub mirrors, different size sub mirrors result in unused disk space.
- Same-type disks and controllers
 - use the same type of disks and controllers in a single mirror. Mixing the different performance levels in a single mirror can cause performance to degrade significantly.

Mirror Guidelines



- Setting read and write policies for sub mirrors
 - experimenting with the mirror read policies can improve performance.
 - Using the **geometric** read option improves performance by minimizing head motion and access time.
- Mounting mirrors
 - only mount the mirror device directly.
 - Do not try and mount a sub mirror directly, unless it is offline and mounted read-only.
 - Do not mount a slice that is part of a sub mirror, this could destroy data and crash the system.

RAID5 Guidelines



- 20-percent write rule
 - because of the complexity of parity calculation, metadevices with greater than 20 percent writes should probably not be RAID5 metadevice.
- Drawbacks to a “slice-heavy” RAID5 metadevice
 - the more slices a RAID5 metadevice contains, the longer read and write operations will take when a component fails.
- Raid5 metadevices and striping guidelines
 - striping guidelines also apply to RAID5 metadevice configuration.

RAID5 Guidelines



- Interlace value
 - it is configured at the time the metadvice created and cannot be modified after.
 - If the different slices in the RAID5 metadvice reside on deferent controllers and the access to the metadvice are primarily large sequential access, than an interlace value of 32 Kbytes might have better performance.
- Concatenating to a RAID5 metadvice
 - concatenating a new slice to an existing RAID5 will have an impact on the overall performance.

Hot Spare Guidelines



- For hot spares as temporary fixes
 - hot spares are not designed to remain a permanent part of your configuration.
 - They need to be replaced with repaired or new slices.
- For cross-controller assigning
 - ideally, slices added to the hot spare pool should be attached to different controllers.
- For wrong size hot spares
 - do not associate hot spares of the wrong size with submirrors or RAID5 metadevices.

Hot Spare Guidelines



- For hot spare marked In Use
 - make sure that all hot spares within a hot spare pool are not marked In-Use.
- For one-way mirrors and hot spares
 - do not assign hot spare pool to a sub mirror in one-way mirror.
- Hot spares are used on first-fit basis
 - when adding hot spares of different sizes to a hot spare pool, add the smaller first.

DRP Guidelines



- Secure SVM configuration files.
 - `/etc/lvm/md.cf`
 - `/etc/lvm/mddb.cf`
 - `/kernel/drv/md.conf`
 - `metadb -i > metadb.out`
- Secure `/etc/vfstab` file.
- Secure `/etc/system/file`.
- Keep Disk VTOC for each disk component.
 - `prtvtoc /dev/dsk/c0t1d0s2 > c0t1d0s2.vtoc`
- Use `fmthard` to rebuild disk VTOC on a new replace disk.
 - `fmthard -s c0t1d0s2.vtoc /dev/rdisk/c0t1d0s2`