

Re: HBVS, shutdown procedures, dismounting disks, SHADOW_MBR_TMO

Source: <http://unix.derkeiler.com/Newsgroups/comp.os.vms/2004-05/0805.html>

From: Tony Arnold (tony.arnold_at_man.ac.uk)

Date: 05/11/04

Date: Tue, 11 May 2004 16:07:06 +0100

Phillip,

I'm not sure what problem you are trying to solve here! Let us consider a few scenarios.

Firstly, let us assume no shadow sets. We will have some disks that are local to the system being rebooted which are either mounted locally and possibly by other systems in the cluster. There will also be some disks mounted by this system but served by some other system in the cluster.

Disks that are local and not mounted by any other system will be dismounted cleanly by SHUTDOWN.COM.

Disks that are local and also mounted by other system in the cluster, will be dismounted by the local system, but remain mounted by the others. Access to this disk by the other systems will hang until the system has been rebooted. If the system is down for long enough, the disks will go into mount verify timeout on the other systems, but if you are just rebooting, this is usually not a problem.

Disks that are served by other systems but mounted locally should also be dismounted cleanly by SHUTDOWN.COM. These disks should continue to be served and available to the rest of the cluster during the reboot.

OK, now let us consider shadow sets, which is more complicated. I think the complication arises when you have one shadow set member on the rebooting system and another shadow set member on another node.

I'm not too sure about this, but I believe that if another node has a file open or is accessing the shadow set at the time you reboot, then a shadow merge or copy will be required when you reboot as one member of the set may have changed and the other not. To avoid this you would have to dismount all shadow set members at least on the nodes they are local to and possibly cluster wide. To do so, you would also have to kill of any application that is accessing these disks. But then if there are no files open on the shadow set during the reboot, then no shadow

copy/merge will be needed and you therefore would not have to dismount the disks in this situation.

Tony.

Phillip Helbig—remove CLOTHES to reply wrote:

- > *Hobbyist cluster, all disks are physical SCSI disks (or, in one case, an*
- > *RF disk on a DSSI bus) MSCP served to all nodes. No SAN, no*
- > *controller-based RAID etc. Most of the disks are in shadow sets which,*
- > *unless it is the system disk, has members on more than one node. All*
- > *disks, whether shadowed or not, are mounted by all nodes in the cluster.*
- >
- > *Occasionally, I want to reboot just one node in the cluster. I'm trying*
- > *to figure out what disks to dismount before doing a reboot. There are*
- > *the following categories of disks:*
- >
- > *A. non-shadowed disks on the node to be rebooted*
- >
- > *B. non-shadowed disks on other nodes*
- >
- > *C. shadow sets with all members on the node to be rebooted*
- >
- > *D. shadow sets with all members on other nodes*
- >
- > *E. shadow sets with some members on the node to be rebooted and some on*
- > *other nodes*
- >
- > *There is also the question whether the dismount needs to be done on the*
- > *node shutting down or on the other nodes.*
- >
- > *I've had a look at SYS\$SYSTEM:SHUTDOWN.COM, but the code is not very*
- > *easy to follow. :-|*
- >
- > *Here is my gut feeling: B is not necessary, in E: I need to dismount the*
- > *members with a direct connection to the node to be rebooted, A is*
- > *probably necessary, as are C and D.*
- >
- > *SHUTDOWN.COM does dismount some disks, of course. As the code is not*
- > *very easy to follow, does anyone know which of the A--E above it does?*
- >
- > *How is any of this related to the value of SHADOW_MBR_TMO?*
- > *SHADOW_MBR_TMO is probably more relevant to an unexpected reboot without*
- > *a clean shutdown caused by a power outage, crash of a machine etc. The*
- > *default is 120 seconds. What, if any, disadvantages are there in*
- > *setting it to a very high value?*
- >
- > *It seems that SHUTDOWN.COM invokes the site-specific shutdown procedure,*
- > *where my own dismount commands would logically be located, before*
- > *stopping the audit server and the smi server. This is unfortunate,*
- > *since if the corresponding files are not on the system disk (as in my*
- > *case, so that they can be common throughout the cluster and available to*

comp.os.vms: Re: HBVS, shutdown procedures, dismounting disks, SHADOW_MBR_TMO

- > *all members which are up), this disk cannot be dismounted. A workaround*
- > *would be to shut down the audit server etc in my site-specific shutdown*
- > *procedure before dismounting the disk, but that seems rather ugly.*
- >