Manual
CSM Management Server Failover

March 10, 2006

Overview

CSM offers a separately orderable feature called HA MS which provides a high availability management server capability. Basically, you set up two management servers and a shared disk. One management server is dubbed the primary, and the other is the backup. If the primary fails, the backup automatically kicks in and takes over. With HA MS you can also force a failover to create a maintenance window for migrating the primary management server to a newer version of CSM without disrupting cluster operations. HA MS is a great solution for some customers.

However, not everyone needs automatic, instantaneous failover. Some customers can't justify the added expense of a shared disk and CSM's HA MS feature. And for some customers, their hardware, software or configuration is simply not supported by CSM's HA MS feature.

So, for some customers, CSM's HA MS is not an option. Yet for some of these customers, a manual CSM MS failover procedure is perfectly acceptable. This paper provides a sketch of such a procedure. You'll probably need to work with this sketch and flesh it out a bit to work in your situation.

The Procedure

To manually failover a CSM management server, you need to do a few preparation steps, and then a few steps at the time of failover. *The following steps seem to do the trick, but have not been exhaustively tested and are not officially supported by IBM at this time. They were tried on a rather simple Linux cluster; so you may need to adjust or expand these steps based on your configuration. Before using these steps or any variation of them in production, you are advised to test them until you are satisfied that they do what you expect, and that all CSM management server function that you depend on actually works on the server you've failed over to!*

The conditions under which this procedure was tried:

1. The primary and backup CSM management servers and all managed nodes were x335s with Red Hat Linux AS3 QU4 installed.
2. The primary and backup CSM management servers and all managed nodes were on the same subnet, each had a single hostname and IP address, and all could resolve the others' hostname and IP address.
3. There were no CSM install servers in the picture.
4. Kerberos was not used.
5. Hardware Control was not configured and thus not tried before or after failover.
6. Implications of Least Prividege commands and the user ids involved was not considered.
7. CSM 1.4.1.12 was used.

Preparation steps:

1. Install your two management servers identically, up to and including running **csmconfig -L** to accept the CSM license. One will be considered your primary management server (MS-A) and the other your backup management server (MS-B).
2. Patch the **mgmtsvr** command if necessary. (This is necessary for CSM 1.4.1.12. Other CSM 1.4.x.x versions will need to be checked. This patch is not necessary for CSM 1.5.x.x.)
   a. Edit **/opt/csm/bin/mgmtsvr** on both MS-A and MS-B to change the one bad instance of "@::MSIPsPs" to "@::MSIPs".
   b. On MS-A, copy **/opt/csm/bin/mgmtsvr** to **/cfmroot/opt/csm/bin/mgmtsvr**
   (This patched version of **mgmtsvr** will get pushed out to all the nodes in step 3 when **installnode** or **updatenode** is run, because both commands invoke CFM under the covers.)
3. On MS-A define and install your managed nodes as per the *CSM Planning and Installation Guide*.
4. Once your cluster is all set up, do the following on MS-A:
   a. **csmbackup**
   b. **tar -Pzcvf /tmp/csmbackup.tar.gz /cfmroot /csminstall /var/opt/csm/csmdata**
   (Note: the resulting tarball can be well over 100MB, due mainly to all the rpms in **/csminstall**... so if you manually maintain **/csminstall** identically on MS-A and MS-B on your own - a BIG task - you don't need to include **/csminstall** in the tarball.)
   c. Save **csmbackup.tar.gz** in a safe place (not on MS-A!).
   d. Repeat steps a,b, and c whenever you use CSM or RSCT commands on MS-A to define new things, change things, delete things, start monitoring, stop monitoring, etc. etc. This is the step that requires the most discipline. If you're not good about generating an up-to-date tarball routinely and when warranted, you'll

back-level yourself during failover… which is not so good… but then again, not nearly as bad as having to completely rebuild and reconfigure your CSM cluster from scratch!

Failover steps:

When you want MS-B to take over, perform the following steps on MS-B:

1. Copy the most recent MS-A csmbackup tarball to MS-B, and run **tar -zxvf csmbackup.tar.gz**
2. **csmrestore**
   (And stick around to answer 'y' to prompts as **csmrestore** proceeds)
   Among other things, the ManagementServer attribute of each managed node will be set properly by the time **csmrestore** finishes.
3. **updatenode -ak**.
   (You may be prompted to provide the root password of the nodes to get ssh set up between MS-B and the nodes.)
4. **csmsetupks -ax or csmsetupyast -ax**
   These commands build the /**tftproot** directory and configure the files within it, and configure files in /**csminstall/csm**. The purpose of this step is to prepare MS-B to reinstall one or more of your nodes with **installnode** should it become necessary for some reason.
5. **chrconsolecfg -a**
   This will regenerate the /**etc/opt/conserver/conserver.cf** file and ensure that in the **default csm** configuration block in this file lists MS-B's hostname as **master**.
6. **rconsolerefresh -r**
   This reinitializes the Conserver daemon based on the new conserver.cf file.

Comments:

Once you've failed over to MS-B you should begin performing the csmbackup steps on MS-B (as described above in Preparation Step 4). You should also repair MS-A so that it can serve as your backup management server.

If you've failed over to MS-B, but MS-A is still alive and well, certain actions will still work from MS-A, but not all, because the managed nodes will no longer recognize MS-A as their management server, and because MS-A will see the Status of the nodes as 'unknown'. For instance, you can still run **lsnode** on MS-A, *but* not all attribute values will be correct! Same is true of **lsrsrc**. **dsh** will still work. **cfmupdatenode** and **smsupdatenode** will *not* work. If you had node monitoring running, commands like **lscondition** and **lscondresp** will tell you that monitoring is still running… and in a sense it is… *but* no node events will reach MS-A! The point is: don't be fooled into thinking you can have a cluster with two, fully functional CSM management servers. You can't. At least not yet.