

Hi there!

This link explains the combination of VZ+DRBD+HA:

http://wiki.openvz.org/HA_cluster_with_DRBD_and_Heartbeat

However, as some other people have pointed in this forum, it would be nice if VEs could be shared between both machines so when both are operative the load would be balanced between them but when one of them is off-line the other took over its VEs.

How would one go about this issue? The biggest problem I see is how to merge both VZ configurations. Symbolic links are a possibility, or a combination filesystem like unionfs or aufs, provided one doesn't use the same VEIDs in both machines.

Some musings:

Consider HNs named cluster11 and cluster12.

There would be two DRBD resources, /vz11 and /vz12 and HA makes them primary in their respective HN and secondary in the other. If one of the HN fails, both will be primary in the remaining node.

/vz11 and /vz12 are configured as in the aforementioned wiki page. /vz is a symbolic link to either /vz11 or /vz12 respectively.

If, e.g., cluster11 fails, at cluster12:

- For every VE in /vz11/cluster/etc/sysconfig/vz-scripts/ that doesn't exist in /vz12/cluster/etc/sysconfig/vz-scripts/:
- Link its conf, quota, private and root to /vz
- If onboot=yes, start it.

When cluster11 comes back to life the procedure would be:

- Stop cluster11's VEs.
- Undo the links.
- Keep /vz11 as primary in cluster12 until it is synced at cluster11.
- Give back /vz11 as primary to cluster11.
- Start cluster11's VEs.

This could probably be automated with some HA mojo, I'd have to look at it.

Obvious drawbacks of this approach are:

- Load balancing is manual.
- Must keep close tabs on VEs so there are no collisions (e.g., name them 11xx and 12xx respectively, separate IP address range, etc.)

- Downtime of the VEs during failback (could cook up some live migration scheme, the key point is waiting until /vz11 is back in sync.)

My questions for the forum are:

- Is the merging of VEs via symbolic links a sound approach? It seems to work well here, but perhaps there is some potential pitfall lurking in the darkness.
- I am rather new to VZ+DRBD+HA. Is there a much better solution for this issue that I have overlooked? There is this LBVM project lbvm.sourceforge.net but I haven't heard of anyone using it, so I haven't investigated it yet.

Cheers,

L
