Subject: VEs on NFS Share
Posted by kuene on Wed, 19 Aug 2009 13:07:53 GMT
View Forum Message <> Reply to Message

hi all

we're running debian 5.0.2 hardware nodes with the repository's kernel (2.6.26-2-openvz-amd64) on 2x Quad Xeon(R) CPU E5440 2.83GHz machines with 64GB of memory. we have two machines of those, which mount the same NFS export, where all VEs of both host nodes are located.

this seemed to have worked well so far, but from time to time we experienced kernel oopses. most of the time, they occur randomly and cannot be related to a specific action. when it happens, the system will freeze eventually.

when it happens, while logged in, I see the following messages in the terminal. unfortunately, we couldn't figure out what they really mean and we are not able to trigger the error.

Message from syslogd@virtue2 at Aug 6 15:55:01 ...
 kernel:[295131.883672] Code: Bad RIP value.

Message from syslogd@virtue2 at Aug 6 15:55:01 ...
 kernel:[295131.883672] CR2: 0000000600000001

Message from syslogd@virtue2 at Aug 6 15:55:01 ...
 kernel:[295131.887671] Oops: 0010 [28] SMP

Message from syslogd@virtue2 at Aug 6 15:55:01 ...
 kernel:[295131.887671] Code: Bad RIP value.

Message from syslogd@virtue2 at Aug 6 15:55:01 ...
 kernel:[295131.887671] CR2: 0000000600000001

Does anybody have experience with storing VEs on NFS shares?
Or could it be a problem that both hostnodes use the same NFS share?
Or do we have to mount an NFS share for each VE? is it related to having VEs installed on an NFS share at all?

regards kuene

Subject: Re: VEs on NFS Share

Posted by maratrus on Wed, 19 Aug 2009 14:49:01 GMT
View Forum Message <> Reply to Message

Hi,

I'd strongly recommend you to create a new bug report.
http://bugzilla.openvz.org/

Meanwhile a new RHEL5 based kernel released which contains some NFS related fixes.
http://wiki.openvz.org/Download/kernel/rhel5/028stab064.4

Subject: Re: VEs on NFS Share
Posted by reduzent on Tue, 25 Aug 2009 18:26:57 GMT
View Forum Message <> Reply to Message

Hi Maratrus,

thanks for your answer. I'm working at the same place, where we have that problem. I would really like to contribute by reporting the issue, but we don't know how do it correctly. We couldn't find a reliable way to deliberately trigger the kernel messages. Also they sometimes appear without any user interaction. The only what we're left with is those kernel messages.

Would it be sufficient for a reasonable bug report to give the specs of the machine (kernel version, distro/version, arch etc) and the kernel messages we posted here?

Thanks in advance

[edit]

I forgot to mention, that we haven't experienced the problem, since we moved the VEs out of the NFS share to a local partition. Probably I'm wrong, but it seems as the problem was related to the VEs being on a NFS export.

Subject: Re: VEs on NFS Share
Posted by seanfulton on Wed, 26 Aug 2009 21:42:33 GMT
View Forum Message <> Reply to Message

I am surprised you are able to store VEs on an NFS share. We did this two or three years ago but were unable to get quotas to work. We were strongly encouraged by the OpenVZ dev team not to do it as NFS has "lots of issues." So we abandoned the idea.

A few months ago I tried to put /vz on an NFS mount and couldn't even get vz to start. I think the error message was something along the lines of, "You can't run /vz on an NFS mount" or something equally as clear

Is this changed? Putting /vz on an HA fail-over NFS mount is a sweet way to set up redundancy.

Head and shoulders over the much, much more complicated GFS setup I tried to configure.

sean

Subject: Re: VEs on NFS Share
Posted by reduzent on Thu, 27 Aug 2009 12:26:05 GMT
View Forum Message <> Reply to Message

What's 'lots of issues' more specifically?

What kind of issues would we have to expect when going this route? So far, for what we need it seems to work except the fact that we get kernel oopses from time to time, which is a blocker.

Maratrus' post makes me assume, that the 'VEs on NFS' setup is 'supposed' to work, i.e. if there are some issues, they should be reported and are tried to be fixed.

Now, is my impression wrong? Was it never the idea to make that work? I found several resources saying that NFS with openVZ should be avoided completely, while others say, what we experience is a bug (a.k.a. when fixed, it should work). So what is the current state on the issue?

Thanks in advance

Subject: Re: VEs on NFS Share
Posted by seanfulton on Thu, 27 Aug 2009 13:27:33 GMT
View Forum Message <> Reply to Message

I'm not sure. Check out this thread:
http://forum.openvz.org/index.php?t=tree&goto=13946& &srch=sean+nfs#page_top

I was able to set up an HA/NFS situation with some issues. After this thread, I tried GFS and a couple of alternatives but all were way too complicated/not mature enough to be of use.

We finally opted for a configuration with mirrored VEs on two separate servers with HA running within the VE to fail-over a shared IP address in case the VE fails.

For our purposes, this is more reliable because it will transition on the failure of an individual VE, as well as the HN, but it is *much* more complicated to manage.

sean

Subject: Re: VEs on NFS Share
Posted by reduzent on Thu, 27 Aug 2009 14:14:44 GMT
View Forum Message <> Reply to Message

Hi Sean

Thanks for the reply. HA is not of so big importance right now, but the VE-on-NFS setup would save a lot of time.

In the meanwhile we discussed the approach you just proposed: Mirroring VEs on the HNs. I think, we'll stick to this route - at least for the near future.

Roman

Subject: Re: VEs on NFS Share
Posted by maratrus on Sun, 30 Aug 2009 09:25:25 GMT
View Forum Message <> Reply to Message

Hello guys,

Quote:
Hi Maratrus,

thanks for your answer. I'm working at the same place, where we have that problem. I would really like to contribute by reporting the issue, but we don't know how do it correctly. We couldn't find a reliable way to deliberately trigger the kernel messages. Also they sometimes appear without any user interaction. The only what we're left with is those kernel messages.

Would it be sufficient for a reasonable bug report to give the specs of the machine (kernel version, distro/version, arch etc) and the kernel messages we posted here?

If you haven't a reliable way to trigger a problem, please don't hesitate to file a bug, just mentioned this fact while creating a new bug report.
The most important points are the specs of the machine and the Oops messages itself. Perhaps, you'll have to set up a serial console to gather all kernel messages.
http://wiki.openvz.org/Remote_console_setup#Serial_console
I would recommend you to use the latest RHEL5 based kernel for your systems because this kernel is stable and still being developed.

/vz over nfs should work. Unfortunately I didn't catch the problems with quotas you were talking about. I know two problems that you may talking about.
- it's impossible to set up nfs exports located under /vz directory
http://bugzilla.openvz.org/show_bug.cgi?id=1086
- vzquota should work on /vz located on nfs according to
http://bugzilla.openvz.org/show_bug.cgi?id=219

Subject: Re: VEs on NFS Share
Posted by Gabor_Laszlo on Fri, 21 May 2010 10:56:21 GMT
View Forum Message <> Reply to Message

I have a similar setup, /var/lib/vz mounted from nfs share on NAS to two OpenVZs (2.6.26-2-openvz-686 Lenny), and a script that pushes control of the VEs to the other 'node' on demand. It works fine most of the time, but I've seen the symptoms you describe just recently following a router crash that cut the link to the NAS.

My problem is rather that at startup sometimes the nfs doesn't mount by the time vz tries to start, so vz just gives up and dies. If I do a
mount -a && /etc/init.d/vz start
it comes up just fine.

Any ideas?