Subject: [solved] nfs mount vs. mount --bind
Posted by yahbluez on Mon, 21 Jan 2008 23:25:14 GMT
View Forum Message <> Reply to Message

Gents,

my question is.

What is better mounting an NFS share into each VE

or

mounting it on the host and bind it into the root after starting the VE?

Better means here where will the servers scale to more IO performance. I don't know where the buffering happens always on the host or in the VE when mountig inside the VE?

I think mounting inside VE has more sex but when the mount --bind gives more performnce or saves RAM i will prefer that.

Any idea?

with best regards
yahbluez

Subject: Re: nfs mount vs. mount --bind
Posted by vaverin on Wed, 23 Jan 2008 05:55:30 GMT
View Forum Message <> Reply to Message

IMHO it's better to use one external mount om the host system and use bindmounts for VEs. I think it should save the memory.

thank you,
Vasily Averin

Subject: Re: nfs mount vs. mount --bind
Posted by freedev on Sun, 28 Jun 2009 00:08:44 GMT
View Forum Message <> Reply to Message

Hi all,

I have a hundred of containers that must mount the same identical NFS share list, so I have for

example: 5 NFS mount on hardware node and 5 hundreds of bind inside the containers.
But I'm experiencing a strange behavior when I stop a container: during the execution of unmount script (in /vz/private/$VEID/scripts/umount), not only the bind mount is unmounted but also the related NFS Volumes is unmounted.

So for example: if we have a NFS volume:

IPADDRESS:/VOLUME/share01 --> /usr/local/storage/share01

and a bind mount:

/usr/local/storage/share01 --> /vz/root/$VEID/usr/local/mnt/share01

when we stop the container and umount /vz/root/$VEID/usr/local/mnt/share01
also the NFS volume /usr/local/mnt/share01 is unmounted.

This problem is critical because if only one Container is stopped all remaining bind inside the containers become stale.
I found a workaround using the umount command with -l (lazy) parameter, so the bind is unmounted correctly without affect the NFS Volumes.

I apologize for a so long introduction, but the question is: could someone tell me if this behavior is related to OpenVZ?
The workaround we found it seems work well, someone think any counterindication or any suggestion?

Thanks in advance,
Vincenzo

Subject: Re: nfs mount vs. mount --bind
Posted by maratrus on Mon, 29 Jun 2009 10:12:04 GMT
View Forum Message <> Reply to Message

IMHO, it's abnormal behavior.

Suppose container is running. Does the command
"#umount /vz/root/$VEID/usr/local/mnt/share01" also umounts
/usr/local/storage/share01?

Could you please show "#cat /proc/mounts" output before the above command and after that command? Could you also show mount and umount scripts?

Subject: Re: nfs mount vs. mount --bind
Posted by freedev on Mon, 29 Jun 2009 10:59:42 GMT

Hi maratrus,

I executed the following commands:

[root@svilpar2 ~]# mount /usr/local/storage/disco1/

[root@svilpar2 ~]# vzctl start 1006
Starting Container ...
mount --bind /usr/local/storage/disco1 /vz/root/1006/usr/local/sitipersonali/disco1
Container is mounted
Setup slm memory limit
Setup slm subgroup (default)
Setting devperms 20002 dev 0x7d00
Adding port redirection to Container(1): 4643 8443
Adding IP address(es) to pool:
Adding IP address(es): XXX.XXX.XXX.XXX XXX.XXX.XXX.XXX
Hostname for Container set: linweb01.xoom.it
File resolv.conf was modified
Container start in progress...
[root@svilpar2 ~]#

[root@svilpar2 ~]# cat /proc/mounts/dev/root / ext3 rw,data=ordered 0 0
/dev /dev tmpfs rw 0 0
/proc /proc proc rw 0 0
/proc/bus/usb /proc/bus/usb usbfs rw 0 0
devpts /dev/pts devpts rw 0 0
/dev/sda5 /vz ext3 rw,noatime,data=ordered 0 0
sunrpc /var/lib/nfs/rpc_pipefs rpc_pipefs rw 0 0
/etc/auto.misc /misc autofs rw,fd=6,pgrp=7159,timeout=300,minproto=5,maxproto=5,indirect 0 0
-hosts /net autofs rw,fd=12,pgrp=7159,timeout=300,minproto=5,maxproto=5,indirec t 0 0
/dev/sda2 /home ext3 rw,data=ordered 0 0
tmpfs /dev/shm tmpfs rw 0 0
sysfs /sys sysfs rw 0 0
XXX.XXX.XXX.XXX:/BASEDIR/DISCO1 /usr/local/storage/disco1 nfs
rw,vers=3,rsize=8192,wsize=8192,soft,proto=tcp,timeo=600,ret
rans=2,sec=sys,addr=151.99.197.39 0 0
XXX.XXX.XXX.XXX:/BASEDIR/DISCO1 /vz/root/1006/usr/local/sitipersonali/disco1 nfs
rw,vers=3,rsize=8192,wsize=8192,soft,proto=tcp,timeo=600,ret
rans=2,sec=sys,addr=151.99.197.39 0 0

[root@svilpar2 ~]# vzctl stop 1006
Stopping Container ...
Container was stopped
umount /vz/root/1006/usr/local/sitipersonali/disco1
Container is unmounted

[root@svilpar2 ct-scripts]# cat /proc/mounts

```
/dev/root / ext3 rw,data=ordered 0 0
/dev /dev tmpfs rw 0 0
/proc /proc proc rw 0 0
/proc/bus/usb /proc/bus/usb usbfs rw 0 0
devpts /dev/pts devpts rw 0 0
/dev/sda5 /vz ext3 rw,noatime,data=ordered 0 0
sunrpc /var/lib/nfs/rpc_pipefs rpc_pipefs rw 0 0
/etc/auto.misc /misc autofs  rw,fd=6,pgrp=7159,timeout=300,minproto=5,maxproto=5,indirect 0 0
-hosts /net autofs  rw,fd=12,pgrp=7159,timeout=300,minproto=5,maxproto=5,indirec t 0 0
/dev/sda2 /home ext3 rw,data=ordered 0 0
tmpfs /dev/shm tmpfs rw 0 0
sysfs /sys sysfs rw 0 0
```

as you can see also the NFS is disappeared!!!

```
[root@svilpar2 ct-scripts]# vi /vz/private/1006/scripts/umount
#!/bin/bash

BASEDIR=/usr/local/sitipersonali
HWDIR=/usr/local/storage

for i in `ls $HWDIR`; do
        if ! [ -d /vz/root/$VEID/$BASEDIR/$i ];
        then
                echo
        fi
        echo umount /vz/root/$VEID$BASEDIR/$i
        umount /vz/root/$VEID$BASEDIR/$i
done

exit 0
~

[root@svilpar2 ct-scripts]# vi /vz/private/1006/scripts/mount
#!/bin/bash

BASEDIR=/usr/local/sitipersonali
HWDIR=/usr/local/storage

for i in `ls $HWDIR`; do
        if ! [ -d /vz/root/$VEID$BASEDIR/$i ];
        then
                echo mkdir -p /vz/root/$VEID$BASEDIR/$i;
                mkdir -p /vz/root/$VEID$BASEDIR/$i;
        fi
        echo mount --bind $HWDIR/$i /vz/root/$VEID$BASEDIR/$i;
        mount --bind $HWDIR/$i /vz/root/$VEID$BASEDIR/$i;
done
```

~

Subject: Re: nfs mount vs. mount --bind
Posted by maratrus on Mon, 29 Jun 2009 14:57:47 GMT
View Forum Message <> Reply to Message

Hi,

first of all, it seems that you have Virtuozzo installed, not OpenVZ, so you'd better ask such questions to virtuozzo support team instead of asking them here because it is an OpenVZ support forum.
But in this particular situation it might be said the following:
you don't need an umount script.
After "mount --bind" is invoked the information about this bind mount is available inside VE. Have a look at /proc/mounts inside VE. Thus, when a VE is stopping appropriate init scripts umount the necessary mount points properly.
The main problem is that the information about bind mounts is written to /etc/mtab so to the moment when a VEs umount script is invoked bind mount is already umounted. /bin/umount fails to do umount and umount the whole nfs mount. Hence, this is not an OpenVZ/Virtuozzo problem but the behavior of /bin/umount binary.
As an additional workaround you can perform bind mounts using "-n" option i.e. "mount -n --bind"

Subject: Re: nfs mount vs. mount --bind
Posted by freedev on Mon, 29 Jun 2009 16:52:01 GMT
View Forum Message <> Reply to Message

Hi maratrus,

I appreciate your support.
Support don't received by Parallels, because they say "the behavior is not related to the product", and reading your answer you confirm their response.

I checked your suggestion and it works...

anyway I receive an error when the umount is executed:

umount /vz/root/1006/usr/local/sitipersonali/disco1
umount: /vz/root/1006/usr/local/sitipersonali/disco1: not mounted

Thanks again,
Vincenzo

Subject: Re: nfs mount vs. mount --bind
Posted by maratrus on Tue, 30 Jun 2009 04:51:13 GMT
View Forum Message <> Reply to Message

Hi,

Quote:
anyway I receive an error when the umount is executed:

Yes, I explained the reason in the previous post. So, to get rid of the error message you mustn't create a VEs umount script. VE umounts this bind mount in one of the init scripts.