Subject: glusterfs/moosefs/etc as openvz backend?
Posted by Aleksandar Ivanisevic on Fri, 29 Apr 2011 10:46:56 GMT
View Forum Message <> Reply to Message

Hi,

Is it possible to use a distributed file system like gluster or
moosefs as a backend for openvz? I don't mean mounting it in a VE, but
running the complete VE on it, i.e. mounting /vz/root on a HN on it.

I've googled a bit and all I see is problems, hangs, etc. Has anyone
successfuly done this?

Subject: Re: glusterfs/moosefs/etc as openvz backend?
Posted by Toby Burress on Fri, 29 Apr 2011 17:36:03 GMT
View Forum Message <> Reply to Message

On Fri, Apr 29, 2011 at 12:46:56PM +0200, Aleksandar Ivanisevic wrote:
>
> Hi,
>
> Is it possible to use a distributed file system like gluster or
> moosefs as a backend for openvz? I don't mean mounting it in a VE, but
> running the complete VE on it, i.e. mounting /vz/root on a HN on it.
>
> I've googled a bit and all I see is problems, hangs, etc. Has anyone
> successfuly done this?

Our disaster recovery scenario, which has been invoked three or four
times, involves running a VE via NFS until a local copy exists, and it
works pretty well. However, they've never run like that for more than
about a day.

Subject: Re: glusterfs/moosefs/etc as openvz backend?
Posted by Angel Bosch Mora on Fri, 29 Apr 2011 17:54:45 GMT
View Forum Message <> Reply to Message

> Our disaster recovery scenario, which has been invoked three or four
> times, involves running a VE via NFS until a local copy exists, and it
> works pretty well. However, they've never run like that for more than
> about a day.
>

can you please develop a little bit more this setup?

Subject: Re: glusterfs/moosefs/etc as openvz backend?
Posted by Toby Burress on Fri, 29 Apr 2011 18:11:29 GMT
View Forum Message <> Reply to Message

On Fri, Apr 29, 2011 at 07:54:45PM +0200, Angel Bosch Mora wrote:
> > Our disaster recovery scenario, which has been invoked three or four
> > times, involves running a VE via NFS until a local copy exists, and it
> > works pretty well. However, they've never run like that for more than
> > about a day.
> >
>
> can you please develop a little bit more this setup?

We back up our VEs with rsnapshot. Usually for restores we just mount the NFS shares for each VE within the VE itself, so that users can grab their files. However if, as sometimes happens, a RAID controller decides that it doesn't want to be a RAID controller anymore, and the whole box goes down, we mount the most recent backup via NFS on another server and start the VEs there.

You have to disable ACLs, and don't squash root, but other than that it seems to work fine. It's not a distributed file system, but my guess is if the file system itself is working properly, a VE will probably run on top of it with little fuss. There might be issues with POSIX semantics, though, that these file systems (which I've never used) may not honor.

Subject: Re: glusterfs/moosefs/etc as openvz backend?
Posted by Aleksandar Ivanisevic on Mon, 02 May 2011 07:27:34 GMT
View Forum Message <> Reply to Message

Toby Burress <tdb@dimins.com> writes:

> You have to disable ACLs, and don't squash root, but other than that it
> seems to work fine. It's not a distributed file system, but my guess
> is if the file system itself is working properly, a VE will probably
> run on top of it with little fuss. There might be issues with POSIX
> semantics, though, that these file systems (which I've never used)
> may not honor.

Thats what I'm trying to do too, and those were exactly my worries. I know that it might work with NFS, but I want something more flexible, so I'm looking into distributed filesystems.

The problem is that I'm not satisfied with daily or hourly rsnapshot backup, I want to have the replica as close to the original as possible. We now replicate using DRBD but it introduces high latency

and it becomes really messy once you have more than 2 nodes (ganeti might help with this, but it was pretty imature on RHEL last time I looked).

One alternative is also lsyncd but the real deal would be a real distributed file system where everything happens in the background.