Subject: Re: [RFC][PATCH 3/4]: Enable multiple mounts of /dev/pts
Posted by Oren Laadan on Wed, 06 Feb 2008 19:25:09 GMT
View Forum Message <> Reply to Message

Serge E. Hallyn wrote:
> Quoting Pavel Emelyanov (xemul@openvz.org):
>> Serge E. Hallyn wrote:
>>> Quoting Pavel Emelyanov (xemul@openvz.org):
>>>> sukadev@us.ibm.com wrote:
>>>>> From: Sukadev Bhattiprolu <sukadev@us.ibm.com>
>>>>> Subject: [RFC][PATCH 3/4]: Enable multiple mounts of /dev/pts
[SNIP]

>>>> That stuff becomes very very similar to that in proc :)
>>>> Makes sense to consolidate. Maybe...
>>> Yeah, and the mqns that Cedric sent too. I think Cedric said he'd
>>> started an a patch implementing a helper. Cedric?
>> Mmm. I wanted to send one small objection to Cedric's patches with mqns,
>> but the thread was abandoned by the time I decided to do-it-right-now.
>>
>> So I can put it here: forcing the CLONE_NEWNS is not very good, since
>> this makes impossible to push a bind mount inside a new namespace, which
>> may operate in some chroot environment. But this ability is heavily
>
> Which direction do you want to go? I'm wondering whether mounts
> propagation can address it.
>
> Though really, I think you're right - we shouldn't break the kernel
> doing CLONE_NEWMQ or CLONE_NEWPTS without CLONE_NEWNS, so we shouldn't
> force the combination.
>
>> exploited in OpenVZ, so if we can somehow avoid forcing the NEWNS flag
>> that would be very very good :) See my next comment about this issue.
>>
>>> Pavel, not long ago you said you were starting to look at tty and pty
>>> stuff - did you have any different ideas on devpts virtualization, or
>>> are you ok with this minus your comments thus far?
>> I have a similar idea of how to implement this, but I didn't thought
>> about the details. As far as this issue is concerned, I see no reasons
>> why we need a kern_mount-ed devtpsfs instance. If we don't make such,
>> we may safely hold the ptsns from the superblock and be happy. The
>> same seems applicable to the mqns, no?
>
> But the current->nsproxy->devpts->mnt is used in several functions in
> patch 3.
>
>> The reason I have the kern_mount-ed instance of proc for pid namespaces
>> is that I need a vfsmount to flush task entries from, but allowing

>> it to be NULL (i.e. no kern_mount, but optional user mounts) means
>> handing all the possible races, which is too heavy. But do we actually
>> need the vfsmount for devpts and mqns if no user-space mounts exist?
>>
>> Besides, I planned to include legacy ptys virtualization and console
>> virtualizatin in this namespace, but it seems, that it is not present
>> in this particular one.
>
> I had been thinking the consoles would have their own ns, since there's
> really nothing linking them, but there really is no good reason why
> userspace should ever want them separate. So I'm fine with combining
> them.

If you want to run something like an X server inside each container
(eg each container holds a desktop session of a different user), then
you need a separate virtual-console namespace for each container.

(yes, X per-se needs to provide remote display as opposed to use
local hardware; see http://www.ncl.cs.columbia.edu/research/thinc/)

[SNIP]

Oren.

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers