Subject: Re: [PATCH 0/16] Pid namespaces
Posted by Pavel Emelianov on Tue, 10 Jul 2007 06:59:35 GMT
View Forum Message <> Reply to Message

Cedric Le Goater wrote:
>>>>> 3. Suka assumes that pid namespace can live without proc mount
>>>>> and tries to make the code work with pid_ns->proc_mnt change
>>>>> from NULL to not-NULL from times to times.
>>>>> My code calls the kern_mount() at the namespace creation and
>>>>> thus the pid_namespace always works with proc.
>>>> shouldn't that be done by userspace instead?
>>> It can be. But when the namespace is being created there's no
>>> any userspace in it yet.
>> I'm not talking about the 'userspace inside the space'
>> I'm talking about the userspace creating the space
>> (what if I do not want to have any proc mount?)
>
> yes, can't we let the user doing the unshare or clone decide whether
> it needs to mount /proc or not in the new pid namespace ?
>
> that's already optional on the host.

If we admit that the pid_ns->proc_mnt pointer can change from NULL
to some valid mount, then we should track all the possible races
with its dereference in the code. We have already caught one of
such "tricky" places with Suka's patches and sent the OOPs to him.

My opinion is that having kern_mount-ed proc makes no discomfort but
protects from all the possible races with this pointer dereferencing.

Thanks,
Pavel

> C.
> -
> To unsubscribe from this list: send the line "unsubscribe linux-kernel" in
> the body of a message to majordomo@vger.kernel.org
> More majordomo info at http://vger.kernel.org/majordomo-info.html
> Please read the FAQ at http://www.tux.org/lkml/
>

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