Subject: Side effects of enabling CAP_SYS_TIME inside VE
Posted by piavlo on Mon, 17 Sep 2007 11:08:12 GMT
View Forum Message <> Reply to Message

Hi,
I get a strange behaviour inside VE then i enable CAP_SYS_TIME
for a specific VE with "vzctl set <VEID> --capability sys_time:on"

Then CAP_SYS_TIME is disabled inside VE and i try to run ntpd
with -u ntp:ntp to drop root privileges, ntpd fails to start with error message:
ntpd[8176]: cap_set_proc() failed to drop root privileges: Operation not permitted

Then i enable the CAP_SYS_TIME inside the VE so that ntpd
could change system time, the side effect is that ntpd
also succeedes to drop root privileges and runs as ntp user.

Why is this so? Why enabling CAP_SYS_TIME also allows dropping root privileges?

Thanks
Alex

Subject: Re: Side effects of enabling CAP_SYS_TIME inside VE
Posted by piavlo on Thu, 20 Sep 2007 08:49:22 GMT
View Forum Message <> Reply to Message

Can one of the devs comment on this?

Subject: Re: Side effects of enabling CAP_SYS_TIME inside VE
Posted by vaverin on Thu, 20 Sep 2007 13:07:22 GMT
View Forum Message <> Reply to Message

Alex,
some action inside kernel (for example like setting system time) checks process permissions.
However kernel do not checks process or user names, it checks process's capability for this purpose.
(you can read man 7 capabilities for more details)

Ususaly capabilites are inherited from parent to children but parent process is able to restrict the children's permissions too.

By default cap_sys_time capability is not allowed inside VE, therefore even VE root cannot change system time.
As far as I understand ntpd tries to set cap_sys_time capability for its working thread and fails

because it does not have this capability.

When cap_sys_time is allowed inside VE this operation finished with success, unprivileged working thread gets the capability allowing it to change the system time.

thank you,
Vasily Averin