Subject: Re: [PATCH][RFC] Cleanup in namespaces unsharing
Posted by Cedric Le Goater on Fri, 08 Jun 2007 12:01:32 GMT
View Forum Message <> Reply to Message

Pavel Emelianov wrote:
> Cedric Le Goater wrote:
>> Pavel Emelianov wrote:
>>> Currently we have two funtions to copy the namespaces:
>>> copy_namespaces() and unshare_nsproxy_namespaces(). The
>>> second one checks for unsupported functionality with
>>>
>>> #ifndef CONFIG_IPC_NS
>>> if (unshare_flags & CLONE_NEWIPC)
>>>   return -EINVAL;
>>> #endif
>>>
>>> -like constructions, while the first one does not. One
>>> of the side effects of this is that clone() with the
>>> CLONE_NEWXXX set will return 0 if the kernel doesn't
>>> support XXX namespaces thus confusing the user-level.
>>>
>>> The proposal is to make these calls clean from the ifdefs
>>> and move these checks into each namespaces' stubs. This
>>> will make the code cleaner and (!) return -EINVAL from
>>> fork() in case the desired namespaces are not supported.
>>>
>>> Did I miss something in the design or this patch worth merging?
>> I've sent a more brutal patch in the past removing CONFIG_IPC_NS
>> and CONFIG_UTS_NS. Might be a better idea ?
>
> In case namespaces do not produce performance loss - yes.
>
> By that patch I also wanted to note that we'd better make
> all the other namespaces check for flags themselves, not
> putting this in the generic code.

yep. let's fix that in the coming ones if they have config option.

a similar issue is the following check done in
unshare_nsproxy_namespaces() and copy_namespaces() :

 if (!capable(CAP_SYS_ADMIN))
  return -EPERM;

it would be interesting to let the namespace handle the unshare
permissions. CAP_SYS_ADMIN shouldn't be required for all namespaces.
ipc is one example.

C.