
Subject: Re: [PATCH 14/20] Allow cloning of new namespace
Posted by [Pavel Emelianov](#) on Wed, 08 Aug 2007 07:29:59 GMT
[View Forum Message](#) <> [Reply to Message](#)

sukadev@us.ibm.com wrote:

> Pavel Emelianov [xemul@openvz.org] wrote:

> | When clone() is invoked with CLONE_NEWPID, create a new pid namespace
> | and then create a new struct pid for the new process. Allocate pid_t's
> | for the new process in the new pid namespace and all ancestor pid
> | namespaces. Make the newly cloned process the session and process group
> | leader.

> |
> | Since the active pid namespace is special and expected to be the first

>
> active pid ns is the *last entry* in this patchset (not first) right ?

yes. thanks.

> | entry in pid->upid_list, preserve the order of pid namespaces.

> |
> | The size of 'struct pid' is dependent on the the number of pid namespaces
> | the process exists in, so we use multiple pid-caches'. Only one pid cache
> | is created during system startup and this used by processes that exist
> | only in init_pid_ns.

> |
> | When a process clones its pid namespace, we create additional pid caches
> | as necessary and use the pid cache to allocate 'struct pids' for that depth.

> |
> | Note, that with this patch the newly created namespace won't work, since
> | the rest of the kernel still uses global pids, but this is to be fixed
> | soon. Init pid namespace still works.

> |
> | Signed-off-by: Pavel Emelyanov <xemul@openvz.org>
> | Signed-off-by: Sukadev Bhattiprolu <sukadev@us.ibm.com>

> |
> | ---

> |
> | include/linux/sched.h | 1
> | kernel/fork.c | 48 ++++++-----
> | kernel/nsproxy.c | 3 +
> | kernel/pid.c | 91 ++++++-----
> | 4 files changed, 118 insertions(+), 25 deletions(-)

> |
> | --- ./include/linux/sched.h.ve14 2007-08-06 15:00:09.000000000 +0400
> | +++ ./include/linux/sched.h 2007-08-06 15:00:09.000000000 +0400

> | @@ -27,6 +27,7 @@
> | #define CLONE_NEWUTS 0x04000000 /* New utsname group? */
> | #define CLONE_NEWIPC 0x08000000 /* New ipc */

```
> | #define CLONE_NEWUSER 0x10000000 /* New user namespace */  
> | +#define CLONE_NEWPID 0x20000000 /* New pids */  
>  
> Nit. I think "New pid namespace" would be a better comment.
```

ok.

```
> Suka  
>
```

Thanks,
Pavel
