
Subject: Re: [PATCH 0/9] namespaces: Introduction
Posted by [Alexey Kuznetsov](#) on Fri, 19 May 2006 20:52:47 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hello!

> > Migration of currently-open sockets (for example) would require storing of
> > a lot of state, wouldn't it?
>
> In a word, yes. :)

Yes. But, actually, it is not "for example". Socket state is really far more complicated thing than all the rest. I would say, migration of another objects is mostly trivial thing.

Actually, what Andrew worried about:

> snapshot/restart/migration worry me. If they require complete
> serialisation of complex kernel data structures then we have a problem,
> because it means that any time anyone changes such a structure they need to
> update (and test) the serialisation.

The answer is: after user space processes referring to objects are suspended, surprisingly, not so much of places, which have trouble with serialization remain. Actually, no serialization additional to existing one is required. Sockets are the most complicated, to suspend networking state, after processes are frozen, we have to:

1. Block access from network.
2. Stop socket timers.

Only after this we can make a coherent snapshot. But it is an exception, most of objects are in coherent state (all the VM, files etc. etc), when processes are frozen.

> I don't think the networking guys from either the OpenVZ project or IBM
> were cc'd on this. Alexey, Daniel, can you elaborate, or point us to
> any existing code?

<http://git.openvz.org>

linux-2.6-openvz/kernel/cpt/. Particularly, kernel/cpt/cpt_socket*.c.
Hairy, but straightforward.

Alexey
