Subject: Re: TCP checkpoint/restart (Re: MCR)
Posted by Daniel Lezcano on Fri, 15 Dec 2006 10:56:45 GMT
View Forum Message <> Reply to Message

Cedric Le Goater wrote:
> Thanks Daniel for moving that thread on the containers@ list.
>
> When you have some time, could you just recap the main topics
> of this discussion on tcp stack checkpoint/restart. I'm pretty
> sure the openvz team as plenty to say.

Sure.

Actually we are working on the network isolation. There are 2 aspects:

   * Full network isolation/virtualization acting at the layer 2 (device)
   * Network isolation at IP layer, we call it layer 3

The network isolation is the mandatory mechanism to ensure the
checkpoint/restart because we must identify the network ressourcess
associated to a container and avoid these ressources to overlap with
other containers.

To be able to take a snapshot of the network container, we must ensure
it is freezed during the checkpoint, because we must ensure the
consistency in the host and with the peers network stack.

We began the checkpoint/restart discussion with this point: how do we do
container's network freeze ?

  * The first step is to drop the traffic
    - shall it be done with the sk_filter fields of the socket ?
    - or with the netfilter NF_DROP/NF_STOLEN ?

   * The second step is to stop tcp timers to avoid socket destruction
while checkpointing it

   -- Daniel