Subject: Re: [PATCH] Virtual ethernet device (tunnel)
Posted by Ben Greear on Wed, 02 May 2007 16:50:48 GMT
View Forum Message <> Reply to Message

Pavel Emelianov wrote:
> Veth stands for Virtual ETHernet. It is a simple tunnel driver
> that works at the link layer and looks like a pair of ethernet
> devices interconnected with each other.
>
> Mainly it allows to communicate between network namespaces but
> it can be used as is as well.
>
> Eric recently sent a similar driver called etun. This
> implementation is closer to the OpenVZ one and it lacks
> some unimportant features of etun driver (like ethtool_ops)
> for simplicity.
>
> The general difference from etun is that a netlink interface
> is used to create and destroy the pairs. The patch for an
> ip utility is also provided.
>
> Signed-off-by: Pavel Emelianov <xemul@openvz.org>
> Acked-By: Kirill Korotaev <dev@sw.ru>
> Acked-By: Dmitry Mishin <dim@openvz.org>
> Acked-By: Alexey Kuznetsov <kuznet@ms2.inr.ac.ru>
>

In your veth_xmit method, do you perhaps also need to initialize
skb->mark to zero?

> +static int veth_create_pair(char *name, char *peer_name)
> +{
> + struct net_device *dev;
> + struct net_device *peer;
> + struct veth_struct *dev_veth;
> + struct veth_struct *peer_veth;
> + int err;
> +
> + dev = veth_dev_start(name);
> + if (IS_ERR(dev)) {
> + err = PTR_ERR(dev);
> + goto err;
> + }
> +
> + peer = veth_dev_start(peer_name);
> + if (IS_ERR(peer)) {
> + err = PTR_ERR(peer);
> + goto err_peer;

> + }
> +
> + dev_veth = dev->priv;
> + peer_veth = peer->priv;
> +
> + dev_veth->peer = peer;
> + dev_veth->dev = dev;
> + peer_veth->peer = dev;
> + peer_veth->dev = peer;
>
It looks like you add it to netdev list before you set up the peer, so
it could race
and crash on stale/null peer pointers?

> +
> + rtnl_lock();
> + list_add(&dev_veth->list, &veth_list);
> + INIT_LIST_HEAD(&peer_veth->list);
> + rtnl_unlock();
> + return 0;
>
Seems to me you might want get all of your state and internal list setup
completed before you register it with the
netdev list, though I'm not sure it matters in this case.

--
Ben Greear <greearb@candelatech.com>
Candela Technologies Inc http://www.candelatech.com