

---

Subject: [PATCH] tun: don't hold network namespace by tun sockets  
Posted by [Stanislav Kinsbursky](#) on Mon, 12 Mar 2012 13:49:08 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

v3: added previously removed sock\_put() to the tun\_release() callback, because sk\_release\_kernel() doesn't drop the socket reference.

v2: sk\_release\_kernel() used for socket release. Dummy tun\_release() is required for sk\_release\_kernel() ---> sock\_release() ---> sock->ops->release() call.

TUN was designed to destroy it's socket on network namespace shutdown. But this will never happen for persistent device, because it's socket holds network namespace.

This patch removes of holding network namespace by TUN socket and replaces it by creating socket in init\_net and then changing it's net to desired one. On shutdown socket is moved back to init\_net prior to final put.

Signed-off-by: Stanislav Kinsbursky <[skinsbursky@parallels.com](mailto:skinsbursky@parallels.com)>

---

drivers/net/tun.c | 15 ++++++++  
1 files changed, 12 insertions(+), 3 deletions(-)

diff --git a/drivers/net/tun.c b/drivers/net/tun.c  
index 2c5d349..74d7f76 100644

--- a/drivers/net/tun.c

+++ b/drivers/net/tun.c

@@ -359,7 +359,7 @@ static void tun\_free\_netdev(struct net\_device \*dev)  
{  
 struct tun\_struct \*tun = netdev\_priv(dev);

- sock\_put(tun->socket.sk);  
+ sk\_release\_kernel(tun->socket.sk);  
}

/\* Net device open. \*/

@@ -980,10 +980,18 @@ static int tun\_recvmmsg(struct kiocb \*iocb, struct socket \*sock,  
 return ret;  
}

+static int tun\_release(struct socket \*sock)  
+{  
+ if (sock->sk)  
+ sock\_put(sock->sk);  
+ return 0;  
+}  
+

```

/* Ops structure to mimic raw sockets with tun */
static const struct proto_ops tun_socket_ops = {
    .sendmsg = tun_sendmsg,
    .recvmsg = tun_recvmsg,
+ .release = tun_release,
};

static struct proto tun_proto = {
@@ -1110,10 +1118,11 @@ static int tun_set_iff(struct net *net, struct file *file, struct ifreq *ifr)
    tun->vnet_hdr_sz = sizeof(struct virtio_net_hdr);

    err = -ENOMEM;
- sk = sk_alloc(net, AF_UNSPEC, GFP_KERNEL, &tun_proto);
+ sk = sk_alloc(&init_net, AF_UNSPEC, GFP_KERNEL, &tun_proto);
    if (!sk)
        goto err_free_dev;

+ sk_change_net(sk, net);
    tun->socket.wq = &tun->wq;
    init_waitqueue_head(&tun->wq.wait);
    tun->socket.ops = &tun_socket_ops;
@@ -1174,7 +1183,7 @@ static int tun_set_iff(struct net *net, struct file *file, struct ifreq *ifr)
    return 0;

err_free_sk:
- sock_put(sk);
+ tun_free_netdev(dev);
err_free_dev:
    free_netdev(dev);
failed:

```

---

Subject: Re: [PATCH] tun: don't hold network namespace by tun sockets  
 Posted by [Stanislav Kinsbursky](#) on Mon, 12 Mar 2012 14:28:31 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Drop this. Sorry for inconvenience - mail server problems.

--  
 Best regards,  
 Stanislav Kinsbursky

---