Subject: Re: [PATCH] Don't attach callback to a going-away netlink socket
Posted by xemul on Mon, 16 Apr 2007 11:53:17 GMT
View Forum Message <> Reply to Message

Patrick McHardy wrote:
> Pavel Emelianov wrote:
>> From: Denis Lunev <den@openvz.org>
>>
>> There is a race between netlink_dump_start() and netlink_release()
>> that can lead to the situation when a netlink socket with non-zero
>> callback is freed.
>
>
> Can you describe the race in more detail please?
>
>

Here it is:

```
CPU1:                       CPU2
netlink_release():              netlink_dump_start():

                            sk = netlink_lookup(); /* OK */

netlink_remove();

spin_lock(&nlk->cb_lock);
if (nlk->cb) { /* false */
  ...
}
spin_unlock(&nlk->cb_lock);

                            spin_lock(&nlk->cb_lock);
                            if (nlk->cb) { /* false */
                                    ...
                            }
                            nlk->cb = cb;
                            spin_unlock(&nlk->cb_lock);
                            ...
sock_orphan(sk);
/*
 * proceed with releasing
 * the socket
 */
```

The proposal it to make sock_orphan before detaching the callback
in netlink_release() and to check for the sock to be SOCK_DEAD in
netlink_dump_start() before setting a new callback.