

## Re: Packet corruption in re0

---

*Source:* <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/stable/2008-03/msg00680.html>

---

- *From:* Pyun YongHyeon <pyunyh@xxxxxxxxxx>
  - *Date:* Fri, 28 Mar 2008 10:39:23 +0900
- 

On Thu, Mar 27, 2008 at 10:41:48AM -0700, Abdullah Ibn Hamad Al-Marri wrote:

----- Original Message -----

From: Pyun YongHyeon <pyunyh@xxxxxxxxxx>  
To: Ian FREISLICH <ianf@xxxxxxxxxx>  
Cc: FreeBSD Current <freebsd-current@xxxxxxxxxx>; Robert Backhaus <robbak@xxxxxxxxxx>  
Sent: Monday, March 17, 2008 8:12:03 AM  
Subject: Re: Packet corruption in re0

On Fri, Feb 22, 2008 at 10:43:22AM +0200, Ian FREISLICH wrote:

> Pyun YongHyeon wrote:

>> On Thu, Feb 21, 2008 at 01:18:18PM +0200, Ian FREISLICH wrote:

>>> Pyun YongHyeon wrote:

>>>> On Thu, Feb 21, 2008 at 02:47:43PM +1000, Robert Backhaus wrote:

>>>>> On Thu, Feb 21, 2008 at 1:50 PM, Pyun YongHyeon wrote:

> ote:

>>>>>> On Thu, Feb 21, 2008 at 11:03:02AM +1000, Robert Backhaus wrote:

>>>>>>> I am experiencing roughly 15% packet corruption on the interface

> face

>>> on

>>>>>>> my freebsd 7/amd64 box.

>>>>>>>

>>>>>>> FreeBSD gw.flexi.robbak.com 7.0-PRERELEASE FreeBSD 7.0-PRERELEASE

> SE #8

>>>> :

>>>>>>> Tue Feb 5 09:49:55 EST 2008

>>>>>>> root@xxxxxxxxxxxxxxxxxxxxxx:/usr/obj/usr/src/sys/GW amd64

>>>>>>>

>>>>>>> Just to make troubleshooting difficult, this problem only shows

> up

>>>>>>> after the system has been up for roughly 36 hours, depending

## Re: Packet corruption in re0

on  
> the  
>>>>> amount of traffic.  
>>>>>>  
>>>>>>  
>>>>> I didn't take a look attached tcpdump files but I guess the  
>>>>> instability issue was fixed in HEAD. It's not yet MFCed but  
>>>>> I'll handle it in a week.  
>>>>>>  
>>>>> Would you try re(4) in HEAD?  
>>>>>>  
>>>>>>  
>>>>> OK, I'll do that. What is the best way to do that? csupping to "."  
se  
> ems a  
>>>>> bit drastic, and I don't do much with cvs proper. I take it that I  
sh  
> ould  
>>> use  
>>>>> anon-cvs to grab the directory, but I don't quite know how.  
>>>>>>  
>>>>>>  
>>>> Copy sys/dev/re/if\_re.c, sys/pci/if\_rlreg.h in HEAD to your box.  
>>>> Due to lack of m\_defrag(9) in 7-PRERELEASE/RC, you also have  
to add  
>>>> that function to if\_re.c(Copy m\_defrag() in sys/kern/uipc\_mbuf.c on  
>>>> HEAD/RELENG\_7 to if\_re.c). That would make it build on your  
box.  
>>>>  
>>>> This doesn't solve the problem that I'm seeing on re(4) interfaces.  
>>>> It basically shows up as quagga establishing OSPF neighbours as  
>>>> "Exchange/DR" when VLAN hardware tagging is enabled. I'm running  
>>>> OSPF over 802.1Q vlans. Neighbours are correctly negotiated once  
>>>> VLAN hardware tagging is disabled on the interface.  
>>>>>>  
>>>> I'll do more debugging.  
>>>>>>  
>>>>>>  
>>>> Hmm. That sounds like different issue to me. I guess I din't change  
>> any semantics in VLAN H/W tagging. Do you still the same VLAN H/W  
>> tagging related issues on RELENG\_7?  
>>>>  
>>>> To narrow down the issue it would be even better to know which parts  
>> of H/W assistance was broken. For example,  
>> - Disable checksum offload for VLAN interface first and check  
>> whether quagga works.  
>>>>  
>>>> You can only disable offload on the parent interface.  
>>>>>>  
>>>> - Disable checksum offload for parent interface and check again.  
>>>> If you can post tcpdump output for broken conntection it may help a

## Re: Packet corruption in re0

> > lot to diagnose the issue.  
>  
> The only flag affecting this behaviour is vlanhwtag. Various  
> permutations of the interface flags make no difference to this  
> behaviour as long as hardware tagging is enabled.  
>  
> It seems like it's corrupting large packets on transmit when vlanhwtag  
> is enabled. From the tcpdump output it looks like a padding or  
> packet length issue.  
>  
> Here's what tcpdump on the re(4) device thinks it's transmitting:  
>  
> 00:08:a1:3c:32:9c > 00:90:fb:0c:89:7d, ethertype 802.1Q (0x8100), length  
1510: vlan 1000, p 0, ethertype IPv4, 196.22.138.92 > 196.22.138.89:  
OSPFv2,  
Database Description, length: 1472  
>  
> Here's what was actually recieved by the em(4) device on the  
> neighbour. Note the absense of the 801.1Q header:  
>  
> 00:08:a1:3c:32:9c > 00:90:fb:0c:89:7d, ethertype IPv4 (0x0800), length  
1506:  
196.22.138.92 > 196.22.138.89: OSPFv2, Database Description, length: 1472  
>  
> When vlanhwtagging is disabled, the re(4) device transmits:  
>  
> 00:90:fb:0c:89:7d > 00:08:a1:3c:32:9c, ethertype 802.1Q (0x8100), length  
1510: vlan 1000, p 0, ethertype IPv4, 196.22.138.89 > 196.22.138.92:  
OSPFv2,  
Database Description, length: 1472  
>  
> and the em(4) device recieves:  
>  
> 00:08:a1:3c:32:9c > 00:90:fb:0c:89:7d, ethertype 802.1Q (0x8100), length  
1510: vlan 1000, p 0, ethertype IPv4, 196.22.138.92 > 196.22.138.89:  
OSPFv2,  
Database Description, length: 1472  
>  
> Let me know if you need more detailed tcpdump output than I've provided.  
>

I guess I've found a VLAN hardware tagging bug in re(4).

Please try this one and let me know the result.

[http://people.freebsd.org/~yongari/re/if\\_re.c](http://people.freebsd.org/~yongari/re/if_re.c)

[http://people.freebsd.org/~yongari/re/if\\_rlreg.h](http://people.freebsd.org/~yongari/re/if_rlreg.h)

> Ian

>

> --

> Ian Freislich

>

Re: Packet corruption in re0

—  
Regards,  
Pyun YongHyeon

Pyun,

I used it, and I got no bufer space available message, I run a server with heavey http requests and named as we..

so I had to increase the buffer.

Please try re(4) in HEAD.

I've just committed one important fix to PCIe variants of RealTek chip. I guess re(4) in HEAD shall fix all known issues reported.

```
www# netstat -m
553/1862/2415 mbufs in use (current/cache/total)
279/1007/1286/65536 mbuf clusters in use (current/cache/total/max)
279/768 mbuf+clusters out of packet secondary zone in use (current/cache)
56/812/868/12800 4k (page size) jumbo clusters in use (current/cache/total/max)
0/0/0/6400 9k jumbo clusters in use (current/cache/total/max)
0/0/0/3200 16k jumbo clusters in use (current/cache/total/max)
920K/5727K/6647K bytes allocated to network (current/cache/total)
0/0/0 requests for mbufs denied (mbufs/clusters/mbuf+clusters)
0/0/0 requests for jumbo clusters denied (4k/9k/16k)
0/0/0 sbufs in use (current/peak/max)
0 requests for sbufs denied
0 requests for sbufs delayed
41261 requests for I/O initiated by sendfile
0 calls to protocol drain routines
```

Can you make a patch for the changes you made in HEAD for RELENG\_7?

I'll MFC re(4) changes in a week.

—  
Regards,  
Pyun YongHyeon

---

freebsd-stable@xxxxxxxxxxxxx mailing list  
<http://lists.freebsd.org/mailman/listinfo/freebsd-stable>  
To unsubscribe, send any mail to "freebsd-stable-unsubscribe@xxxxxxxxxxxxx"

Re: Packet corruption in re0