pfSense - Bug #4310

Limiters + HA results in hangs on secondary

01/27/2015 02:13 AM - Chris Buechler

Status:	Resolved	**Start date:**	01/27/2015
Priority:	Very High	**Due date:**	
Assignee:	Luiz Souza	**% Done:**	100%
Category:	Limiters	**Estimated time:**	0.00 hour
Target version:	2.4.3		
Affected Version:	2.2.x	**Affected Architecture:**	

Description

Configuring limiters on a firewall rule in 2.2 on a system using HA results in a kernel panic reboot loop. To replicate, on a basic HA setup with config sync and pfsync enabled, add a pair of limiters and put them on the default LAN rule. It'll panic upon applying the changes, and do so again after rebooting in an endless loop.

History

#1 - 01/29/2015 04:38 PM - Ermal Luçi

I think this happens because CARP packets are being sent to dummynet.
Before the kernel patch prevented this from happening.

Will investigate and fix it accordingly.

#2 - 02/03/2015 04:44 PM - Ermal Luçi

- Status changed from Confirmed to Feedback

Patch committed.

#3 - 02/13/2015 04:45 PM - Vitaliy Isarev

Hi, I have the same issue. I tried to update to the latest maintance version, but receive error after upgrade: "shared object libpcre.so.1 not found required by php"

#4 - 02/17/2015 04:13 AM - Vitaliy Isarev

Ermal Luçi wrote:

> Patch committed.

Can you post a link to the patch

#5 - 03/03/2015 01:03 PM - Chris Buechler

- Status changed from Feedback to Resolved

fixed

#6 - 03/25/2015 05:23 AM - Steve Wheeler

We are seeing a number of reports that this is still an issue in 2.2.1. At least one customer ticket and also:
https://forum.pfsense.org/index.php?topic=87541.0
Even if there is no longer a panic, though there is for some, the machine is no longer usable with the logs spammed with the erro message in the patch.

#7 - 04/02/2015 10:43 AM - Jim Pingle

- Status changed from Resolved to Confirmed

- Target version changed from 2.2.1 to 2.2.2

This is still a problem. Some cases still work but with TONS of console/log spam about pfsync_undefer_state rendering the console and logs practically unusable. There are still a couple people reporting panics as well, though they don't seem to be making it to the crash reporter (other reports are submitting fine, however).

Somewhat related, there appears to be a logic issue in the pfsync enable checkbox for those who upgraded from 2.0.x or before (before the HA settings were moved). The GUI shows the option unchecked, but the empty tag is in the config which is causing pfsync to be enabled. This can lead to these errors showing up on non-HA units until they enable and then disable pfsync.

#8 - 04/03/2015 02:05 PM - Ermal Luçi

- Status changed from Confirmed to Feedback

I pushed the messages under debug misc level and also another change to fix the root cause for it.

#9 - 04/06/2015 04:08 PM - Chris Linstruth

Looks good here. Not stressing it but enabling/disabling limiters on the cluster works, the limiters are doing what I ask, and the limited states are syncing. Thanks.

#10 - 04/07/2015 11:00 PM - Chris Buechler

Chris: that still working fine for you?

After running for a few hours, the secondary still hangs in one of our test setups. Console is non-responsive, ESX shows VM using 100% CPU. That happens after about 4 hours of run time. The primary is fine throughout, and the limiters work on both v4 and v6.

#11 - 04/08/2015 01:50 AM - Chris Linstruth

I haven't seen anything else but please understand that this is on a test bench not in production and I am not stressing it at all. If you look at ticket UJN-78146 you will see my description of a crash I have seen since starting to test yesterday's snapshot. Not sure if it is related. The HA pair has been up since but with no more than about 150 states.

#12 - 04/09/2015 05:38 PM - Chris Linstruth

A bit more info. See this thread:

https://forum.pfsense.org/index.php?topic=92128.0

Turning off the limiters makes that NAT translation work.

#13 - 04/10/2015 11:06 PM - Chris Buechler

- Target version changed from 2.2.2 to 2.2.3

this is better, though still the issue where the secondary may hit 100% CPU and hang in some circumstance. We'll revisit.

The issue with reflection and limiters is #4590

#14 - 05/15/2015 01:47 PM - Ermal Luçi

Patch was committed for this on tools repo and also the defer option in pfsync is now not used.
Both can be considered as the root cause of the issue here.

#15 - 05/19/2015 11:23 PM - Chris Buechler

- Subject changed from Limiters + HA results in kernel panic to Limiters + HA results in hangs on secondary

- Status changed from Feedback to Confirmed

- Affected Version changed from 2.2 to 2.2.x

no change, still hangs secondary within a couple hours

#16 - 06/08/2015 06:06 PM - Ermal Luçi

- Assignee changed from Ermal Luçi to Chris Buechler

Chris need to confirm this happens still or not.

#17 - 06/08/2015 06:11 PM - Chris Buechler

- Status changed from Confirmed to Feedback

I'm pretty sure it doesn't happen anymore, still have the test setup running to make sure. Given another ~48 hours, if it's still not an issue, this will be safe to consider fixed.

#18 - 06/17/2015 01:37 PM - Chris Buechler

- Status changed from Feedback to Confirmed

- Assignee changed from Chris Buechler to Ermal Luçi

no change, as long as you have some traffic passing through a limiter, the secondary hangs within ~1-4 hours.

#19 - 06/20/2015 07:38 PM - Chris Buechler

- File redmine4310-crash.txt added

- Target version changed from 2.2.3 to 2.3

Tried after changing both hosts to use unicast pfsync, which had no impact. It seems to alternate between hanging the secondary, and triggering a kernel panic. Thus far using unicast, the secondary has only kernel paniced, not hung consuming 100% CPU. Same kernel panic happened on occasion when using multicast pfsync so not sure that's actually changed.

crash report attached

#20 - 07/09/2015 06:38 PM - Bernardo Pádua

- File crash report 2.txt added

- File crash report 1.txt added

This is also happening to me. I though the issue with the limiters was fixed in 2.2.2 and 2.2.3, so I posted a duplicate ticket on #4823. But I've now disabled my limiters and saw that the backup/secondary firewall stopped crashing.

I'm posting my crash dumps (two different times the backup crashed) here in case they are of any help.

#21 - 08/31/2015 09:25 PM - Jim Thompson

- Assignee changed from Ermal Luçi to Luiz Souza

#22 - 12/07/2015 06:11 PM - James Starowitz

lastnight rolled out our 2.2.5 units using c2758s in HA, the units worked fine in a test lab, although once i put it into production the backup router would hang to the point that it could not access the webui, the hang occurs nearly every 5-10minutes, the backup router reboots and then crashs again soon after.

i disabled "state sync" on both the master and the backup and it stopped crashing.

because we use limiters per source ip each client ip has its own limit, as a result the HFSC work around cant really do what im doing with limiters

we rarely go into failover, but when we do seamless state transfers are a life saver.

i submitted the crash reports and opened a support ticket if you need the details.

#23 - 01/28/2016 08:47 AM - Lee Shiry

This problem seems to get worse after upgrading to 2.2.6. Now the secondary still hangs even with the limiters and state sync disabled.

#24 - 03/03/2016 04:32 AM - Dirk Bongard

- File 02.03.2016 22_18.txt added

I have the same issue.
Panic on my HA-Backup between 10 minutes and 3 hours. I have send you several crash reports via gui.
Limiter + Vlan + NAT

#25 - 03/04/2016 10:07 AM - William St.Denis

- File 02.03.2016 22_18.txt added

I have noticed this issue as well. We have to disable sync when using limiters because it's crashing the system. I have attached my log as well
I am running Limiter + Vlan + NAT as well. When we started running limiters we noticed the WEB UI started to slow down and were getting an error /rc.filter_synchronize: An authentication failure occurred then we got the kernel panic.

#26 - 03/04/2016 10:10 AM - William St.Denis

- File Crash_02.04.2016.txt added

Sorry wrong log. Here is the correct one

#27 - 03/05/2016 04:19 PM - Luiz Souza

- Target version changed from 2.3 to 2.3.1

#28 - 03/16/2016 08:48 AM - Mikhail Platonov

I have the same issue.
ha-backup crashed after 7 min

#29 - 03/17/2016 03:55 PM - William St.Denis

Does anyone have a work around to keep limiters and sync working? The only option I have come up with is to disable limiters or disable sync both aren't great.

#30 - 03/18/2016 12:20 AM - Chris Buechler

William St.Denis wrote:

> Does anyone have a work around to keep limiters and sync working? The only option I have come up with is to disable limiters or disable sync both aren't great.

Those are your only two options for the time being. ALTQ can often be used in the same way as limiters and doesn't have such issues. Post to the forum if you'd like to discuss further.

#31 - 04/15/2016 10:08 PM - Chris Buechler

- Target version changed from 2.3.1 to 2.3.2

#32 - 05/25/2016 11:22 AM - Jose Duarte

From the tests we ran for the last couple of days we saw kernel panic using limiters in multiple vlans but no impact when using different queues inside those limiters.

#33 - 07/08/2016 03:32 AM - Chris Buechler

- Target version changed from 2.3.2 to 2.4.0

#34 - 12/09/2016 11:24 PM - Luiz Souza

2.4 has a few new fixes for use-after-free pfsync states. The limiters issue is also fixed.

#35 - 12/10/2016 06:03 AM - Jim Pingle

I updated a test cluster to a snapshot from a couple hours ago, which from the timestamp looks like it should have this fix, and both nodes got stuck in a panic loop.

```
Version String: FreeBSD 11.0-RELEASE-p3 #233 8ae63e9(RELENG_2_4): Sat Dec 10 03:56:41 CST 2016
  root@buildbot2.netgate.com:/builder/ce/tmp/obj/builder/ce/tmp/FreeBSD-src/sys/pfSense
Panic String: pfsync_undefer_state: unable to find deferred state
```

Same panic string on both nodes, slightly different backtrace.

Primary:

```
db:0:kdb.enter.default>  bt
Tracing pid 12 tid 100044 td 0xfffff80003500500
kdb_enter() at kdb_enter+0x3b/frame 0xffffffe001a62e430
vpanic() at vpanic+0x19f/frame 0xffffffe001a62e4b0
panic() at panic+0x43/frame 0xffffffe001a62e510
pfsync_update_state() at pfsync_update_state+0x45b/frame 0xffffffe001a62e560
pf_test() at pf_test+0x1bcc/frame 0xffffffe001a62e7d0
pf_check_in() at pf_check_in+0x1d/frame 0xffffffe001a62e7f0
pfil_run_hooks() at pfil_run_hooks+0x8c/frame 0xffffffe001a62e880
ip_input() at ip_input+0x3eb/frame 0xffffffe001a62e8e0
netisr_dispatch_src() at netisr_dispatch_src+0xa5/frame 0xffffffe001a62e940
ether_demux() at ether_demux+0x15c/frame 0xffffffe001a62e970
ether_nh_input() at ether_nh_input+0x317/frame 0xffffffe001a62e9d0
netisr_dispatch_src() at netisr_dispatch_src+0xa5/frame 0xffffffe001a62ea30
ether_input() at ether_input+0x26/frame 0xffffffe001a62ea50
vmxnet3_rxq_eof() at vmxnet3_rxq_eof+0x708/frame 0xffffffe001a62eae0
vmxnet3_legacy_intr() at vmxnet3_legacy_intr+0x110/frame 0xffffffe001a62eb20
intr_event_execute_handlers() at intr_event_execute_handlers+0x20f/frame 0xffffffe001a62eb60
ithread_loop() at ithread_loop+0xc6/frame 0xffffffe001a62ebb0
fork_exit() at fork_exit+0x85/frame 0xffffffe001a62ebf0
fork_trampoline() at fork_trampoline+0xe/frame 0xffffffe001a62ebf0
```

Secondary:

```
db:0:kdb.enter.default>  bt
Tracing pid 12 tid 100032 td 0xfffff800034c1500
kdb_enter() at kdb_enter+0x3b/frame 0xffffffe001a3cf240
vpanic() at vpanic+0x19f/frame 0xffffffe001a3cf2c0
panic() at panic+0x43/frame 0xffffffe001a3cf320
pfsync_update_state() at pfsync_update_state+0x45b/frame 0xffffffe001a3cf370
pf_test() at pf_test+0x245b/frame 0xffffffe001a3cf5e0
pf_check_out() at pf_check_out+0x1d/frame 0xffffffe001a3cf600
pfil_run_hooks() at pfil_run_hooks+0x8c/frame 0xffffffe001a3cf690
ip_output() at ip_output+0xd8b/frame 0xffffffe001a3cf7e0
ip_forward() at ip_forward+0x36b/frame 0xffffffe001a3cf880
ip_input() at ip_input+0x6da/frame 0xffffffe001a3cf8e0
netisr_dispatch_src() at netisr_dispatch_src+0xa5/frame 0xffffffe001a3cf940
ether_demux() at ether_demux+0x15c/frame 0xffffffe001a3cf970
ether_nh_input() at ether_nh_input+0x317/frame 0xffffffe001a3cf9d0
netisr_dispatch_src() at netisr_dispatch_src+0xa5/frame 0xffffffe001a3cfa30
ether_input() at ether_input+0x26/frame 0xffffffe001a3cfa50
vmxnet3_rxq_eof() at vmxnet3_rxq_eof+0x708/frame 0xffffffe001a3cfae0
vmxnet3_legacy_intr() at vmxnet3_legacy_intr+0x110/frame 0xffffffe001a3cfb20
intr_event_execute_handlers() at intr_event_execute_handlers+0x20f/frame 0xffffffe001a3cfb60
ithread_loop() at ithread_loop+0xc6/frame 0xffffffe001a3cfbb0
fork_exit() at fork_exit+0x85/frame 0xffffffe001a3cfbf0
fork_trampoline() at fork_trampoline+0xe/frame 0xffffffe001a3cfbf0
--- trap 0, rip = 0, rsp = 0, rbp = 0 ---
```

#36 - 01/04/2017 01:14 AM - Vladimir Usov

Dear Luiz! Can we expect real fix in 2.4? We are waiting for it too long, and this is a really critical problem, since in a corporate environment you will always use both - HA and limiters. br Vladimir

#37 - 01/10/2017 08:29 PM - James Kohout

I would agree with Vladimir. Just would like to know if this will be definitely be fixed in 2.4 or pushed out further.
Thanks

#38 - 01/18/2017 06:06 AM - Jose Duarte

One more here, we always have limiters and HA and we are forced to use the queues. If someone makes a mistake of assigning a main limiter to a rule instant kernel panic...

#39 - 03/23/2017 04:31 PM - Steve Yates

We are not noticing our secondary (which is also a VM) hang. However, our one limited rule traffic ends overnight, so possibly it recovers after the messages end?

Reiterating from the referenced forum thread, there is a checkbox "No pfSync" on firewall rules, but checking that doesn't avoid the error message. Nor does setting "State type" to None.

We didn't see this issue until upgrading from 2.2.6 to 2.3.1_5.

#40 - 04/13/2017 11:22 AM - James Webb

Still Producing issues for me. Had to re-install pfSense on both devices in HA after encountering this bug.

#41 - 05/07/2017 07:22 PM - Sean Huggans

Experiencing this after updating from 2.1.5 to 2.3.4. Constant Kernel messages in system logs as: "pfsync_undefer_state: unable to find deferred state".

We had a limiter in place to limit bandwidth of our backup server when replicating through an IPSec tunnel to a backup server offsite in order to prevent packet loss caused by taking up all the bandwidth of our WAN.

Didn't actually notice until users reported not being able to access resources on the other side of the tunnel - apparently once backup replications started to the remote host, it killed the tunnel it was replicating through/being limited on somehow.

Any plans to resolve this issue? Limiters are a very useful feature, as is HA obviously.

#42 - 05/12/2017 02:47 PM - Matthew Brown

Hmmm.... this is very much no not ideal. :(I was going to do this in a new environment as we have soft limits in our datacenters. It would be very useful to simply limit our incoming and outgoing speed to a set amount. If we are over our allotted speed for more than x number of hours we will be forced to upgrade our connection Speed.

Does anyone know if this issue was fixed with the release of 2.4? I don't really want to install "bleeding edge" tech, but I also don't want to have to tell my boss our database networking costs will double because we are 1MBps's over limit. ^^;;;

#43 - 06/08/2017 04:44 PM - Scott Rosenberg

Has this had any development recently?

This is the primary reason I can't use limiters in my HA setup, and the assignee hasn't commented in 6 months.

#44 - 07/21/2017 04:19 AM - Jose Duarte

For those still with problems you can use limiters in HA with any version w/out kernel panic but for that you need additional configuration.

1. Create a new limiter for both upload and download with the bandwidth limit. Name it with the name you want and _donotuse at the end (just for safety)
2. Create a new Queue inside of each limiter (When inside of the limiter "Add New Queue" green button)
3. Name the queues with the vlan/rule name and the bandwidth you set in the limit and with _up or _down (for reference) and set the weight to 100 for that queue to use 100% of the limiter
4. Assign the queues you created to the rules you want to limit the bandwidth. MAKE SURE YOU ASSIGN THE QUEUE AND NOT THE LIMITER, IF YOU CHOOSE THE LIMITER YOU WILL HAVE THE KERNEL PANIC IN THE 2nd MEMBER. That's why it's a better practice to use the name _donotuse in the limiters.

Notes:
You still need to create 1xlimiter + 1xqueue per each flow per rule
If you assign the same queues to multiple rules they will share the same "roof" defined in the limiter
You can create multiple queues for one limiter with different weight, very useful if you want to have, for example, a top limit of 400Mbit and give rule1 guaranteed 10% of those, rule2 50% and rule3 40%. If all of the rules/queues are being maxed out you will have a perfect bandwidth balance. If for example rule 2 and 3 don't have any traffic rule1 will be able to use the 400Mbit since we only define a guaranteed minimum.

Cheers.

#45 - 07/24/2017 03:32 AM - Lars Jorgensen

Jose Duarte wrote:

> For those still with problems you can use limiters in HA with any version w/out kernel panic but for that you need additional configuration.

Thank you!

Confirmed working here. Great load off my chest as running without HA was never very fun.

Lars

#46 - 09/11/2017 03:57 PM - Renato Botelho

- Target version changed from 2.4.0 to 2.4.1

#47 - 09/15/2017 07:02 PM - Jose Duarte

Moved, yet again :(

#48 - 10/12/2017 10:09 AM - Jim Pingle

- Target version changed from 2.4.1 to 2.4.2

#49 - 10/12/2017 10:53 AM - Sander Naudts

Why not change target version to 2.9.9... sorry just little frustrating that this doesn't get fixed.

#50 - 10/12/2017 10:55 AM - Jim Pingle

We expected to have more time before 2.4.1 but we need to have it out in a week or so, there isn't time to get to this and the other things we have to address for it.

And if you read above, there is a viable workaround if you use queues/child limiters and not the limiters directly.

#51 - 10/13/2017 01:22 AM - Lars Jorgensen

Sander Naudts wrote:

> Why not change target version to 2.9.9... sorry just little frustrating that this doesn't get fixed.

It's not that much of a problem as long as you use the workaround described in comment #44. I've been running HA with limiters without any problems for three months now.

#52 - 10/23/2017 12:18 PM - Jim Pingle

- Target version changed from 2.4.2 to 2.4.3

#53 - 12/25/2017 03:48 AM - Eero Volotinen

Lars Jorgensen wrote:

> Sander Naudts wrote:
>
> > Why not change target version to 2.9.9... sorry just little frustrating that this doesn't get fixed.
>
> It's not that much of a problem as long as you use the workaround described in comment #44. I've been running HA with limiters without any problems for three months now.

still issue with 2.4.2 .. please at least add note to gui that you cannot use pfsync with limiters. It saves lot of time .. it took something like 2 days to figure, why ha units were crashing..

#54 - 01/24/2018 02:56 PM - Luiz Souza

- Status changed from Confirmed to Feedback

- % Done changed from 0 to 100

The crash is fixed on the last snapshot.

Tests are welcome.

#55 - 02/19/2018 05:29 AM - Fabrizio Pappolla

- File pfsense_crashlog.txt added

Before open a new ticket, i will try here since the error looks really similar. My pfSense got bootloop, the problem was caused by a black out, the error was: "kernel panic pfsync_undefer_state: unable to find deferred state". I have not HA on, only limiter and PRIQ. Attacched you can find the crash log.
pfSense Version 2.4.2-RELEASE-p1 (amd64)

#56 - 02/19/2018 08:08 AM - Jim Pingle

Fabrizio Pappolla wrote:

> Before open a new ticket, i will try here since the error looks really similar. My pfSense got bootloop, the problem was caused by a black out, the error was: "kernel panic pfsync_undefer_state: unable to find deferred state". I have not HA on, only limiter and PRIQ. Attacched you can find the crash log. pfSense Version 2.4.2-RELEASE-p1 (amd64)

The backtrace shows pfsync, so you must have that active. This has been fixed on 2.4.3, so additional problem reports on anything older are not helpful. Upgrade to a 2.4.3 snapshot and see if it is more stable there.

#57 - 03/12/2018 08:08 AM - Jim Pingle

- Status changed from Feedback to Resolved

Confirmed working by multiple tests and users.

Files

redmine4310-crash.txt	167 KB	06/21/2015	Chris Buechler
crash report 2.txt	162 KB	07/09/2015	Bernardo Pádua
crash report 1.txt	302 KB	07/09/2015	Bernardo Pádua
02.03.2016 22_18.txt	160 KB	03/03/2016	Dirk Bongard
02.03.2016 22_18.txt	160 KB	03/04/2016	William St.Denis
Crash_02.04.2016.txt	150 KB	03/04/2016	William St.Denis
pfsense_crashlog.txt	188 KB	02/19/2018	Fabrizio Pappolla