pfSense - Bug #8249

pid 77785 (php-fpm), uid 0, was killed: out of swap space

12/31/2017 06:25 PM - Pi Ba

Status:	Resolved	**Start date:**	12/31/2017
Priority:	Normal	**Due date:**	
Assignee:	Luiz Souza	**% Done:**	100%
Category:	Web Interface	**Estimated time:**	0.00 hour
Target version:	2.4.3		
Affected Version:	2.4.3	**Affected Architecture:**	All

Description

My pfSense box is running out of memory while it should have plenty to do what it does.

2.4.3-DEVELOPMENT (amd64)
built on Sat Dec 30 11:07:37 CST 2017

Running on a Virtualbox with 1GB memory and 1GB swap space.
Dashboard and htop shows both memory and swap being 'full'.

Also on console these messages pop up:
pid 77785 (php-fpm), uid 0, was killed: out of swap space
pid 1083 (php-fpm), uid 0, was killed: out of swap space
pid 53694 (php-fpm), uid 0, was killed: out of swap space
pid 65729 (php-fpm), uid 0, was killed: out of swap space
pid 48018 (php-fpm), uid 0, was killed: out of swap space
swap_pager: out of swap space
swap_pager_getswapspace(16): failed
pid 1082 (php-fpm), uid 0, was killed: out of swap space
swap_pager_getswapspace(7): failed
swap_pager_getswapspace(4): failed

After running on console '16 restart php-fpm' memory usage drops to 272MB memory and 216MB swap used. So about 1.5GB of memory was released..

I believe this is might be related to https://github.com/pfsense/pfsense/pull/3881

Output of some commands attached:
curl -k "https://localhost/status?full" > mem_phpfpm_status_full.log
/usr/bin/netstat -Ln > mem_netstat_Ln.log
/usr/bin/netstat -xn > mem_netstat_xn.log
/usr/sbin/swapinfo -h > mem_swapinfo_xn.log
/usr/bin/top | /usr/bin/head -n7 > mem_top.log
/bin/ps uxawwd > mem_psuxawwd.log
/usr/bin/sockstat > mem_sockstat.log

History

#1 - 01/01/2018 07:05 PM - Pi Ba

For a easy reproduction do the following:
-Configure dashboard with 1 widget like the "Interfaces" widget.
(It also happens with other widgets, but to keep it 'simple' 1 widget is enough to cause the issue..)

-Restart pfSense
After this there should be 7 fpm workers running (curl -k https://localhost/status), and memory usage should be pretty low in htop like < +-500 MB.

-Login to webgui with only the interface widget refreshing every second.
After this memory usage will slowly increase..
After about 2500 requests to each worker memory starts to become an issue, first regular memory will fill up, and then slowly it also consumes all swap space.. Eventually it starts killing processes.. Because he complete 1GB memory + 1GB swap is used up..

(To make that happen a bit faster for testing, in index.php the setTimeout for widget refreshing can be set to 100 instead of 1000, this imho does not

make any difference in the effect, only in the timeframe)

#2 - 01/04/2018 12:20 PM - Jim Pingle

I can't seem to reproduce this here. I have a box with 2GB of RAM and a 14 dashboard widgets loaded and it never goes above 2 active processes, even though it's configured to allow for 8. The ps output shows less than 25% of memory used as well. That system also has a few other packages on it (FRR, FreeRADIUS3, and some others) but no pfBlocker or anything else heavy.

When you notice it starts to use a large chunk of ram, try watching top -aS -o res and then see what shows as using the largest chunk of resources at the time.

It is entirely possible that PR is to blame though, it may need some additional adjustments, maybe another step or two for larger amounts of RAM, since with other processes running that may be too many workers for 1GB of RAM.

#3 - 01/04/2018 06:28 PM - Pi Ba

- File 20180105_top -aS -o res.txt added

The number of widgets doesn't really matter, only 1 will be refreshed every second..
Could use ab or wrk for a second with 10 connections to bump up the number of running workers.. Probably due to other parts of my config with several gateways and openvpn servers and what not that some extra workers get spawned during reboot. That is imho irrelevant to the issue that happens once they are running though. As after that a single dashboard widget will keep all workers alive and make them grow..

I've bumped up my memory to 4GB with 1GB swap.. (the amount of swap space used by a process cant really be seen as far as i figured out..) With 4GB regular memory the issue is probably more easily 'seen'. Anyhow above 1GB the number of workers and other settings used for fpm stay the same.. And it shouldn't ever need 4GB of memory with those settings that are apparently used for a 1GB machine.

As for pm.process_idle_timeout which you might think would terminate some workers during idle time (at least i did..). It is only used when pm=ondemand.. not with pm=dynamic according to php documentation..

Output of top attached in file: "20180105_top -aS -o res.txt"
Several of these 500MB php-fpm processes running take a lot of memory.. (Even 2 processes are to much for a 1GB machine).. (after 2000 requests each or so..)

```
Mem: 3319M Active, 17M Inact, 120M Laundry, 394M Wired, 264K Buf, 79M Free
  PID USERNAME        THR PRI NICE   SIZE    RES STATE   C   TIME    WCPU COMMAND
65564 root             1  52    0   756M   500M accept  3   2:25   0.00% php-fpm: pool nginx (php-fpm)
```

#4 - 01/11/2018 07:22 AM - Chris Collins

I created the original patch to increase the fpm processes, and my processes dont use anywhere near 500meg of ram resources each.

I think its a good idea to determine why your system has this unusual behaviour.

So some more information from myself.

I already use the interface stats widget, it does not update every second and its update rate is also not configurable, so it does sound like you have been manually editing src files outside of supported pfsense configuration.

If I leave my unit sitting on the dashboard, with numerous widgets enabled.

1 - only 2 fpm processes are running (you should only see more when is lots of background scripts running like pfblockerng updates, or if multiple clients are accessing the web ui at the same time).
2 - the fpm proceses are using about 50 (fifty) meg of ram each.

```
PID USERNAME       THR PRI NICE   SIZE    RES STATE   C    TIME    WCPU COMMAND
58857 root           1  20    0   294M 52368K accept  1    0:21   0.02% php-fpm
```

3 - the extra processes "do" shutdown when they no longer needed eventually as well. So 7-8 processes do not stay running forever.

Your system definitely is not normal behaviour, even on the old configuration it allowed 2 fpm processes, people run pfsense on very tiny systems, if it was normal for it to be 500meg per process the forum would be rife with complaints related to memory saturation.

I can if requested make a new patch that requires more ram, to spawn more processes, but this is the only complaint, and I think there is something weird on your system causing such high memory usage per process.

#5 - 01/11/2018 12:36 PM - Pi Ba

I might have edited some things.. (its my pfSense testmachine i also develop things on..) so i cant tell for sure if that could trigger part of the issue.. but after upgrading to latest snapshot the problem persists on my machine..

Besides the number of processes 2 to 8 there is also pm.max_requests=500 which changed to 5000 this makes it possible for the problem to manifest after a few thousand requests to each process..

Anyhow, ill install a clean machine and see whats needed exactly to replicate the issue.

#6 - 01/11/2018 12:39 PM - Pi Ba

p.s. As for widget refresh every second.. thats normal.. once a second 'a' widget is refreshed.. delete all your other widgets and you will see interfaces widget refresh every second.

#7 - 01/11/2018 04:16 PM - Pi Ba

Some fresh numbers from a clean install, wan(dhcp)/lan(static) ssh enabled and thats it..:
After reboot 27MB per process.
After 4900 requests 85MB per process.
Not a problem yet, but already 'growing'.

Will try and find what can make it grow faster like my original problem report.

#8 - 01/16/2018 01:42 PM - Chris Collins

I can make a new patch so its back to max 500 requests before a child process is forcibly restarted, although martin isnt around anymore to submit a commit, so someone else will need to do that.

On my pfsense unit most of my widgets dont update every second, or at least dont appear to update every second. The only one that seems to be every second is the ntp widget.

Widgets I have activated on dashboard.

pfblockerng
interfaces
gateways
smart status
interface statistics
traffic graphs
thermal sensors
ups status
services status
rss
openvpn
ntp status

so quite a lot

Memory consumption is as follows with 4 gig of ram

57meg active
574 meg inactive (cache immediatly available for reallocation)
781 meg wired (kernel memory usage, includes zfs ARC which is 153meg)
23 meg buffered
2432 meg free
0 swap usage

this is with 7 fpm process's running whilst on dashboard

restarting fpm only shifted 200meg from inactive to free, so on my system with lots of widgets its usage is barely relevant. The biggest chunk is allocated to the kernel which shouldnt include userland processes like fpm.

I have just applied a patch to my unit to test the following

max requests before kill 1500 (compromise between 5000 and 500), at 4900 your "memory leak" seems minimal.
max spare 4 (instead of 7), this means after fpm becomes less busy it will shutdown up to 4 processes instead of leaving them running, but is a risk of a gateway error if nginx requests a fpm backend and a child cannot start quick enough to serve it. So I will do some testing during a pfblockerng update to see if I get gateway errors.

#9 - 01/17/2018 06:06 PM - Pi Ba

Ok found at least part of the leak / function call causing it (on my development/test machine)..

Running the code below from diagnostics\command php results in a php process that takes SIZE 1600M and 1400M RES.
This same function is also called from interfaces widget through the get_interface_info() function.

My opt1 interface is a regular em2 interface, but it does have a ipv6 type of 6to4tunnel selected.
Same effect happens when using opt14 which is a bridge over em1.12 and em1.13 vlans. All 3 items have IPv6 configured as 'None'.

Other interfaces do not seem to be affected by this particular leak..

```
require_once("pfsense-utils.inc");
require_once("functions.inc");
for($callcount = 0; $callcount < 4900; $callcount++) {
    for($interfacecount = 0; $interfacecount < 10; $interfacecount++) {
    $ifdescr = "opt1";
    $chkif = get_real_interface($ifdescr);
    $ifinfo['linklocal'] = get_interface_linklocal($ifdescr);
    $ifinfo['ipaddrv6'] = get_interface_ipv6($ifdescr);
    $ifinfo['subnetv6'] = get_interface_subnetv6($ifdescr);
    $ifinfotmp = pfSense_get_interface_stats($chkif);
    }
}
sleep(30);
echo "Done\n";
```

```
last pid: 88446;  load averages:  1.04,  0.85,  0.68
                           up 0+00:24:16  00:40:23
142 processes: 2 running, 139 sleeping, 1 waiting
CPU:  0.0% user,  0.0% nice,  0.8% system,  0.0% interrupt, 99.2% idle
Mem: 1485M Active, 71M Inact, 28M Laundry, 324M Wired, 276K Buf, 40M Free
ARC: 142M Total, 64M MFU, 71M MRU, 170K Anon, 954K Header, 6718K Other
     82M Compressed, 227M Uncompressed, 2.77:1 Ratio
Swap: 1024M Total, 168M Used, 856M Free, 16% Inuse

  PID USERNAME      THR PRI NICE   SIZE    RES STATE   C   TIME   WCPU COMMAND
56825 root           1  52    0  1623M  1401M nanslp  2   1:10   0.00% /usr/local/bin/php -d log_errors=off
/tmp/user_code/U0OvFz
```

@Chris Collins, Try using interfaces widget alone on the dashboard, you will see it refreshes every second then. (one widget gets refreshed per second if you have 10 they will refresh once every 10 seconds..) (a few exceptions exist to this relatively new widget refresh mechanism, but in general thats how it works on recent pfSense versions..)

#10 - 01/18/2018 02:59 PM - Pi Ba

Memory usage of pfSense_getall_interface_addresses("em0")

```
$pid = getmypid();
for($i = 0; $i < 10; $i++) {
  echo "Mem:".`ps -p $pid -o rss= -o vsz=`;
  for($callcount = 0; $callcount < 1000; $callcount++) {
    pfSense_getall_interface_addresses("em0");
  }
}
echo "Mem:".`ps -p $pid -o rss= -o vsz=`;
echo "Done\n";


Mem:31584 248604
Mem:59680 277276
Mem:87696 305948
Mem:115728 334620
Mem:143788 363292
Mem:171904 394012
Mem:199956 422684
Mem:228048 451356
Mem:256052 480028
Mem:284200 508700
Mem:312328 537372
Done
```

#11 - 01/22/2018 10:14 AM - Luiz Souza

- Assignee set to Luiz Souza

#12 - 01/22/2018 11:09 AM - Luiz Souza

- Status changed from New to Feedback

- % Done changed from 0 to 100

Hm.. I found the leak.

Should be fixed in php56-pfSense-module-0.58.

Thank you!

#13 - 01/23/2018 01:14 PM - Pi Ba

Thank you!
Confirmed fixed.

2.4.3-DEVELOPMENT (amd64)
built on Tue Jan 23 04:03:53 CST 2018
FreeBSD 11.1-RELEASE-p6

Same loop as before now results in this:

```
Mem:31576 248604
Mem:31600 248604
Mem:31600 248604
Mem:31600 248604
Mem:31600 248604
Mem:31600 248604
Mem:31600 248604
Mem:31600 248604
Mem:31600 248604
Mem:31600 248604
Mem:31600 248604
Done
```

Perfect score imho :)

#14 - 01/23/2018 02:05 PM - Luiz Souza

- Status changed from Feedback to Resolved

Files

mem_netstat_Ln.log	2.15 KB	01/01/2018	Pi Ba
mem_netstat_xn.log	24.5 KB	01/01/2018	Pi Ba
mem_phpfpm_status_full.log	3.6 KB	01/01/2018	Pi Ba
mem_psuxawwd.log	18.6 KB	01/01/2018	Pi Ba
mem_sockstat.log	17.2 KB	01/01/2018	Pi Ba
mem_swapinfo_xn.log	104 Bytes	01/01/2018	Pi Ba
mem_top.log	362 Bytes	01/01/2018	Pi Ba
20180105_top -aS -o res.txt	3.51 KB	01/04/2018	Pi Ba