Subject: Re: OOM didn't save the machine
Posted by lazy on Tue, 31 Mar 2009 12:34:21 GMT
View Forum Message <> Reply to Message

maratrus wrote on Tue, 31 March 2009 06:36Hello,

Quote:
It's possible the "leak"

Yes, it might be a leak. Moreover, you said that
Quote:
Same code runs without problems on non openvz kernels 2.6.22 25 27 for months

So, let's try to find out if it's indeed a memory leak.

Quote:
Now it's happening again, apache process is using 100% cpu, I can't enter the vps beancounters bellow

Looks like you have to adjust CPUUNITS/CPULIMITS on the HN. You may read about them in OpenVZ user's guide
http://download.openvz.org/doc/OpenVZ-Users-Guide.pdf
this VE has 300% cpulimit, and themachine has 4 cores

Quote:

oomguarpages 50345 59833 35236 9223372036854775807 0

Oomguarpages exceeded barrier value, so this is likely the reason why a process inside that VE was killed.
http://wiki.openvz.org/Oomguarpages#oomguarpages
yes, first time when memory run out oom tried to kill some processes but I don't know if the one eating all the cpu was killed

Do you have a single VE on the HN? If no, please show the full user_beancounters output.

Version: 2.5

uid	resource	held	maxheld	barrier	limit	failcnt
3001:	kmemsize	164187239	168681365	268435456	536870912	0
	lockedpages	0	0	2562	2562	0
	privvmpages	750001	863788	1048576	1048576	0
	shmpages	28	48	15374	15374	0
	dummy	0	0	0	0	0

numproc 202 247 2000 2000 0
physpages 667127 774657 0 9223372036854775807 0
vmguarpages 0 0 35236 9223372036854775807 0
oomguarpages 667127 774657 35236 9223372036854775807 0
numtcpsock 134 309 2000 2000 0
numflock 10 32 1000 1100 0
numpty 0 1 200 200 0
numsiginfo 0 101 1024 1024 0
tcpsndbuf 1125360 4048272 8388608 10485760 0
tcprcvbuf 1046080 2213216 8388608 10485760 0
othersockbuf 171680 1371728 4650461 12842461 0
dgramrcvbuf 0 8736 4650461 4650461 0
numothersock 101 140 2048 3000 0
dcachesize 0 0 11452893 11796480 0
numfile 3896 4609 20480 20480 0
dummy 0 0 0 0 0
dummy 0 0 0 0 0
dummy 0 0 0 0 0
numiptent 10 10 200 200 0
3000: kmemsize 80854359 83363201 125829120 131072000 0
lockedpages 0 0 2562 2562 0
privvmpages 135920 171574 512000 537600 0
shmpages 284 284 15374 15374 0
dummy 0 0 0 0 0
numproc 65 127 2000 2000 0
physpages 37459 59833 0 9223372036854775807 0
vmguarpages 0 0 35236 9223372036854775807 0
oomguarpages 37503 59833 30000 30000 0
numtcpsock 15 95 2000 2000 0
numflock 2 13 1000 1100 0
numpty 0 1 200 200 0
numsiginfo 8 30 1024 1024 0
tcpsndbuf 266800 1460304 9300923 17492923 0
tcprcvbuf 1296 286512 638976 1048576 0

```
        othersockbuf      13136      1270944      4650461      12842461
    0
        dgramrcvbuf           0         4368      4650461       4650461
0
        numothersock         11           33         8000          9000        0
        dcachesize            0            0     11452893      11796480        0
        numfile            2126         4801        20480         20480        0
        dummy                 0            0            0             0        0
        dummy                 0            0            0             0        0
        dummy                 0            0            0             0        0
        numiptent            10           10          200           200        0
     0:  kmemsize        74806381     74857340  9223372036854775807
9223372036854775807           0
        lockedpages           0            0  9223372036854775807
9223372036854775807           0
        privvmpages        8703        20362  9223372036854775807
9223372036854775807           0
        shmpages           1287         1303  9223372036854775807
9223372036854775807           0
        dummy                 0            0  9223372036854775807  9223372036854775807
              0
        numproc              99          109  9223372036854775807
9223372036854775807           0
        physpages          4995        15435  9223372036854775807
9223372036854775807           0
        vmguarpages           0            0  9223372036854775807
9223372036854775807           0
        oomguarpages       5002        15435  9223372036854775807
9223372036854775807           0
        numtcpsock            6            7  9223372036854775807
9223372036854775807           0
        numflock              1            7  9223372036854775807  9223372036854775807
              0
        numpty                8            8  9223372036854775807  9223372036854775807
              0
        numsiginfo            2            6  9223372036854775807  9223372036854775807
              0
        tcpsndbuf        125696       698000  9223372036854775807
9223372036854775807           0
        tcprcvbuf         98304      1312608  9223372036854775807
9223372036854775807           0
        othersockbuf       9280        24832  9223372036854775807
9223372036854775807           0
        dgramrcvbuf           0         8464  9223372036854775807
9223372036854775807           0
        numothersock         25           30  9223372036854775807
9223372036854775807           0
        dcachesize            0            0  9223372036854775807  9223372036854775807
```

```
            0
        numfile                1935             2135  9223372036854775807
9223372036854775807              0
        dummy                     0                0  9223372036854775807  9223372036854775807
            0
        dummy                     0                0  9223372036854775807  9223372036854775807
            0
        dummy                     0                0  9223372036854775807  9223372036854775807
            0
        numiptent                10               10  9223372036854775807
9223372036854775807              0
```

Quote:
Next time this issue will occur try to stop the problem VE (hope CPULIMIT/CPUUNITS adjustment will allow to do that) and look at the user_beancounters again. There should be no usage when VE is stopped. Please, also look at the slab state. If the consumption value is decreased. And show please /proc/slabinfo output after and before stopping the problem VE.

after i tried to set vzctl set 3000 --cpus 1 all other vzctl's got stuck

there was monstrous filp cache usage in slab, (taken at second event at Tue)

```
mnt_cache              44     75    256   15    1 : tunables  120   60    8 : slabdata      5      5      0
inode_cache          1665   1740    608    6    1 : tunables   54   27    8 : slabdata    290    290      0
dentry_cache       896624 896624    240   16    1 : tunables  120   60    8 : slabdata  56039  56039      0
filp              9343650 9343650    256   15    1 : tunables  120   60    8 : slabdata 622910 622910      0
names_cache           33     33   4096    1    1 : tunables   24   12    8 : slabdata     33     33      0
idr_layer_cache      102    105    528    7    1 :
```

Please look at the attatched images, there is a lot of kernel activity about 100% i guess this is runaway unkillable httpd whitch was in some kind of loop whitch was also eating filp cache thus crashing the machine,

this process had under 20 open files and was sending msgs threw af_unix socket as I mentioned bellow

memory consumption by userspace processes ware normal under 4 GB all the time, it looks like filp cache was eating the rest

Today we roll out plain apache on that vps, but later I'll try to trigger this event on a test machine.

File Attachments

1) ram.png, downloaded 335 times

2) cpu.png, downloaded 307 times