Subject: IP Conntrack /DST Cache Overflow issue
Posted by KevinH on Mon, 04 Nov 2013 21:24:11 GMT
View Forum Message <> Reply to Message

My network slowly experinces slowness, intill it eventually dies off. And i am hoping someone that has had this problem could help:

Kernel: 2.6.18-348.4.1.el5.028stab107.1
OS: CentOS release 5.8 (Final)

Nov 3 07:41:18 909l7 kernel: process `sysctl' is using deprecated sysctl (sysc$
Nov 3 09:36:09 909l7 kernel: Fatal resource shortage: privvmpages, UB 167.
Nov 3 09:36:09 909l7 last message repeated 3 times
Nov 3 21:05:01 909l7 auditd[8428]: Audit daemon rotating log files
Nov 4 00:18:56 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.
Nov 4 00:18:58 909l7 last message repeated 9 times
Nov 4 00:19:01 909l7 kernel: printk: 2 messages suppressed.

Nov 4 00:18:58 909l7 last message repeated 9 times
Nov 4 00:19:01 909l7 kernel: printk: 2 messages suppressed.
Nov 4 00:19:01 909l7 kernel: Route hash chain too long!
Nov 4 00:19:01 909l7 kernel: Adjust your secret_interval!
Nov 4 00:19:06 909l7 kernel: printk: 11 messages suppressed.
Nov 4 00:19:06 909l7 kernel: Route hash chain too long!
Nov 4 00:19:06 909l7 kernel: Adjust your secret_interval!
Nov 4 00:19:11 909l7 kernel: printk: 9 messages suppressed.
Nov 4 00:19:11 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.
Nov 4 00:19:13 909l7 kernel: venet0: 5 rebuilds is over limit, route caching d$
Nov 4 00:19:16 909l7 kernel: printk: 15 messages suppressed.
Nov 4 00:19:16 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.
Nov 4 00:19:21 909l7 kernel: printk: 13 messages suppressed.
Nov 4 00:19:21 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.
Nov 4 00:19:26 909l7 kernel: printk: 3 messages suppressed.
Nov 4 00:19:26 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.
Nov 4 00:19:31 909l7 kernel: printk: 5 messages suppressed.
Nov 4 00:19:31 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.
Nov 4 00:19:36 909l7 kernel: printk: 6 messages suppressed.
Nov 4 00:19:36 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.
Nov 4 00:19:41 909l7 kernel: printk: 6 messages suppressed.
Nov 4 00:19:41 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.
Nov 4 00:19:46 909l7 kernel: printk: 7 messages suppressed.
Nov 4 00:19:46 909l7 kernel: ip_conntrack: CT 102: table full, dropping packet.

It will then slowly DST Cache overflow.

At this point, i am a tad lost on what to do.

here is the sysctl.conf file:

net.ipv4.ip_forward = 1
net.ipv6.conf.default.forwarding = 1
net.ipv6.conf.all.forwarding = 1
net.ipv4.conf.default.proxy_arp = 0
net.ipv4.conf.all.rp_filter = 1
kernel.sysrq = 1
net.ipv4.conf.default.send_redirects = 1
net.ipv4.conf.all.send_redirects = 0
kernel.panic = 10
net.ipv4.tcp_mem = 786432 1048576 1572864
net.ipv4.netfilter.ip_conntrack_max=196608
net.ipv4.tcp_syncookies = 1
net.ipv4.tcp_no_metrics_save=1
net.ipv4.netfilter.ip_conntrack_tcp_timeout_syn_recv = 5
net.ipv4.netfilter.ip_conntrack_tcp_timeout_established = 600
net.ipv4.netfilter.ip_conntrack_tcp_timeout_syn_sent = 15
net.ipv4.netfilter.ip_conntrack_tcp_timeout_fin_wait = 3
net.ipv4.netfilter.ip_conntrack_tcp_timeout_close_wait = 3
net.ipv4.netfilter.ip_conntrack_tcp_timeout_time_wait = 3
net.ipv4.route.gc_thresh = 131072

Any help to fix this so it stops the crashes and slowness, is greatly helpful.

Thanks,

Kevin Hammett

Subject: Re: IP Conntrack /DST Cache Overflow issue
Posted by Paparaciz on Sun, 10 Nov 2013 14:12:19 GMT
View Forum Message <> Reply to Message

please paste CT "102.conf" file, and /proc/user_beancounters stats for this CT

Subject: Re: IP Conntrack /DST Cache Overflow issue
Posted by KevinH on Sun, 10 Nov 2013 15:04:38 GMT
View Forum Message <> Reply to Message

102.conf:

```
NUMPROC="999999"
PHYSPAGES="0:2147483647"
VMGUARPAGES="263168:2147483647"
OOMGUARPAGES="263168:2147483647"
NUMTCPSOCK="7999992"
NUMFLOCK="999999"
NUMPTY="500000"
NUMSIGINFO="999999"
TCPSNDBUF="214748160:396774400"
TCPRCVBUF="214748160:396774400"
OTHERSOCKBUF="214748160:396774400"
DGRAMRCVBUF="214748160:396774400"
NUMOTHERSOCK="7999992"
DCACHESIZE="2147483646"
NUMFILE="23999976"
AVNUMPROC="1000:1000"
NUMIPTENT="999999"

# Disk quota parameters (in form of softlimit:hardlimit)
DISKSPACE="524288000:524288000"
DISKINODES="262144000:262144000"
QUOTATIME="0"
```

And beancounters:

```
[root@pronto /]# cat /proc/user_beancounters
Version: 2.5
       uid  resource                     held              maxheld              barrier                limit              failcnt
      102:  kmemsize                   9670459             10723310           2147483646           2147483646
                    0
            lockedpages                 8190                 8190               999999               999999
0
            privvmpages                307642               332446              1548288              1548288
      0
            shmpages                      31                  703               263168               263168                    0
            dummy                          0                    0                    0                    0                    0
            numproc                      125                  137               999999               999999                    0
            physpages                 151112               151373                    0           2147483647
  0
            vmguarpages                    0                    0               263168           2147483647
0
            oomguarpages              151112               151373               263168           2147483647
          0
            numtcpsock                    13                   21              7999992              7999992
0
            numflock                       9                   14               999999               999999                    0
            numpty                        14                   14               500000               500000                    0
            numsiginfo                     0                    3               999999               999999                    0
```

tcpsndbuf	227968	1793056	214748160	396774400	0
tcprcvbuf	212992	926848	214748160	396774400	0
othersockbuf	21024	33120	214748160	396774400	0
dgramrcvbuf	86432	1123616	214748160	396774400	0
numothersock	38	53	7999992	7999992	0
dcachesize	0	0	2147483646	2147483646	0
numfile	3130	3292	23999976	23999976	0
dummy	0	0	0	0	0
dummy	0	0	0	0	0
dummy	0	0	0	0	0
numiptent	0	0	999999	999999	0

Hope this helps

Subject: Re: IP Conntrack /DST Cache Overflow issue
Posted by pavel.odintsov on Sun, 10 Nov 2013 22:30:08 GMT
View Forum Message <> Reply to Message

Hello, KevinH!

I will recommend you to enlarge a buffer for conntrack buckets up to 1-3 millions (it's ok for modern systems).

You can do it by a following commands:
edit /etc/sysctl.conf and fix:
net.ipv4.netfilter.ip_conntrack_max=1548576

Apply changes:

sysctl -p

You can entirely disable connnection tracking for certain CT with following commands:

iptables -t raw -A PREROUTING -d VPS_IP -j NOTRACK
iptables -t raw -A PREROUTING -s VPS_IP -j NOTRACK

Be aware about using last command, it result to completely disable RELATED/ESTABLISHED flags to working inside container.

In addition I recommend you to upgrade from pretty old kernels 2.6.18 to modern 2.6.32 kernel series. Oh, sorry, I forgot to specify link, it's here https://openvz.org/028_to_042_kernel_upgrade