Subject: ip_conntrack: CT 0: table full, dropping packet.
Posted by mavines on Tue, 27 Oct 2009 01:55:10 GMT
View Forum Message <> Reply to Message

Hi,

I have the major problem with networking - /var/log/messages reports this:
ip_conntrack: CT 0: table full, dropping packet.

And I have CentOS 5.4, kernel:

Linux 2.6.18-128.2.1.el5.028stab064.7xen #1 SMP Wed Aug 26 16:41:55 MSD 2009 x86_64 x86_64 x86_64 GNU/Linux

ip rule list
0: from all lookup 255
32766: from all lookup main
32767: from all lookup default

the iptables rules are attached, tcpdump is attached as well.

here is the lsmod of main server:

Module	Size	Used by
iptable_nat	43532	0
tun	47872	2
xt_physdev	35984	0
bridge	94384	1 xt_physdev
netloop	40324	0
netbk	129984	0 [permanent]
blktap	151460	2 [permanent]
blkbk	54712	0 [permanent]
vzethdev	47520	0
vznetdev	56848	8
simfs	38296	4
vzrst	172968	0
ip_nat	53392	2 iptable_nat,vzrst
vzcpt	150840	0
vzdquota	78320	4 [permanent]
vzmon	83864	8 vzethdev,vznetdev,vzrst,vzcpt
vzdev	36872	4 vzethdev,vznetdev,vzdquota,vzmon
xt_tcpmss	35328	0
ipt_tos	34560	0
xt_tcpudp	36224	245
xt_conntrack	36352	0
ip_conntrack_irc	41168	0
xt_state	35200	23
ip_conntrack_ftp	42192	0

```
xt_multiport          36224  4
xt_length             34944  0
xt_mac                34944  0
xt_limit              36352  46
ipt_TCPMSS            37248  1
iptable_mangle        37888  5
iptable_filter        37760  6
ipt_TOS               35200  14
ipt_ULOG              42504  0
ip_conntrack          100884  8  iptable_nat,vzrst,ip_nat,vzcpt,xt_conntrack,ip_conntrack_irc
,xt_state,ip_conntrack_ftp
ipt_recent            43404  45
ipt_ecn               35200  0
ipt_owner             34944  0
ip_tables             57440  3 iptable_nat,iptable_mangle,iptable_filter
ipt_ttl               34816  0
ipt_REJECT            39684  42
ipt_LOG               39808  40
nfnetlink             40392  2 ip_nat,ip_conntrack
x_tables              52616  21  iptable_nat,xt_physdev,xt_tcpmss,ipt_tos,xt_tcpudp,xt_conntr
ack,xt_state,xt_multiport,xt_length,xt_mac,xt_limit,ipt_TCPM
SS,ipt_TOS,ipt_ULOG,ipt_recent,ipt_ecn,ipt_owner,ip_tables,i pt_ttl,ipt_REJECT,ipt_LOG
autofs4               57480  2
hidp                  83584  2
rfcomm                104872  0
l2cap                 89216  10 hidp,rfcomm
bluetooth             118916  5 hidp,rfcomm,l2cap
lockd                 101776  0
sunrpc                201416  2 lockd
ipv6                  456124  435 vzrst,vzcpt,vzmon
xfrm_nalgo            43268  1 ipv6
crypto_api            42880  1 xfrm_nalgo
ib_iser               66936  0
rdma_cm               67092  1 ib_iser
ib_cm                 67752  1 rdma_cm
iw_cm                 43400  1 rdma_cm
ib_sa                 74760  2 rdma_cm,ib_cm
ib_mad                70820  2 ib_cm,ib_sa
ib_core               93700  6 ib_iser,rdma_cm,ib_cm,iw_cm,ib_sa,ib_mad
ib_addr               42128  1 rdma_cm
iscsi_tcp             57856  0
libiscsi              63488  2 ib_iser,iscsi_tcp
scsi_transport_iscsi    66960  5 ib_iser,iscsi_tcp,libiscsi
loop                  48656  4
dm_mirror             54280  0
dm_multipath          55192  0
scsi_dh               41600  1 dm_multipath
raid0                 40448  1
```

```
video                53004  0
hwmon                36488  0
backlight            39808  1 video
sbs                  49856  0
i2c_ec               38528  1 sbs
container            37760  0
button               40480  0
battery              43784  0
asus_acpi            50724  0
ac                   38664  0
lp                   47056  0
snd_hda_intel       481584  0
snd_seq_dummy        36996  0
snd_seq_oss          65408  0
snd_seq_midi_event   40960  1 snd_seq_oss
snd_seq              87968  5 snd_seq_dummy,snd_seq_oss,snd_seq_midi_event
snd_seq_device       41492  3 snd_seq_dummy,snd_seq_oss,snd_seq
snd_pcm_oss          77440  0
snd_mixer_oss        49920  1 snd_pcm_oss
snd_pcm             116872  2 snd_hda_intel,snd_pcm_oss
snd_timer            57224  2 snd_seq,snd_pcm
snd_page_alloc       43920  2 snd_hda_intel,snd_pcm
r8169                71172  0
snd_hwdep            43528  1 snd_hda_intel
snd                  99496  9  snd_hda_intel,snd_seq_oss,snd_seq,snd_seq_device,snd_pcm_oss
,snd_mixer_oss,snd_pcm,snd_timer,snd_hwdep
mii                  38784  1 r8169
i2c_i801             41620  0
soundcore            41760  1 snd
parport_pc           62248  1
pcspkr               36224  0
i2c_core             56064  2 i2c_ec,i2c_i801
ohci1394             67928  0
shpchp               70572  0
parport              73868  2 lp,parport_pc
ieee1394            390648  1 ohci1394
r8168                95000  0
sg                   69544  0
serial_core          56192  0
dm_raid45            98704  0
dm_message           36096  1 dm_raid45
dm_region_hash       46336  1 dm_raid45
dm_log               44800  3 dm_mirror,dm_raid45,dm_region_hash
dm_mod              100560  4 dm_mirror,dm_multipath,dm_raid45,dm_log
dm_mem_cache         39424  1 dm_raid45
ahci                 68744  4
libata              208784  1 ahci
sd_mod               56448  7
```

```
scsi_mod            197528  8  ib_iser,iscsi_tcp,libiscsi,scsi_transport_iscsi,scsi_dh,sg,l
ibata,sd_mod
ext3                168848  2
jbd                 102512  1 ext3
uhci_hcd             57496  0
ohci_hcd             55988  0
ehci_hcd             65676  0
```

So, firstly I had a problem with orphaned sockets (it persisted about 6 months, but I did not paid enough attention), then server started to fail and I somehow fixed the problem with orphaned sockets at CT300. But instead I've got the problem with ip_conntrack which lasts for 2 weeks already and I am not able to figure out why.

sysctl -A output is also attached.

So, my server for now gets overflowed at least once by day and I have to increase max value for ip_conntrack, but this helps only for few hours, then I have to reboot the server or it just hangs without response.

I was suspecting that this is an attack (like DDoS) but even when the server has very few connections, it gets overflowed. And that is really nasty.

Can you please help with that?

Yes, another detail - I have 4 OpenVZ VPSs and 2 XEN VPSs running on the same server. It has 8 GB RAM and 1.5 TB HDD - Core Quad Intel CPU.

File Attachments

1) iptables_rules.txt, downloaded 670 times
2) tcpdump.txt, downloaded 654 times
3) sysctl.txt, downloaded 667 times

Subject: Re: ip_conntrack: CT 0: table full, dropping packet.
Posted by kir on Tue, 27 Oct 2009 09:33:56 GMT
View Forum Message <> Reply to Message

This is not really openvz specific, google for ip_conntrack table full dropping packet and you will find the answers.

In short -- check why there are so many entries by analyzing /proc/net/ip_conntrack, and either fix the issues found (maybe somebody is attacking you), or increase /proc/sys/net/ipv4/ip_conntrack_max.

Subject: Re: ip_conntrack: CT 0: table full, dropping packet.
Posted by mavines on Tue, 27 Oct 2009 10:59:59 GMT
View Forum Message <> Reply to Message

Thanks for the answer. Obviously I did the google search for my problem. I thought it is attack - but really it is not. Even my own connections to server stays in the /proc/net/ip_conntrack for long time. As I told, I was increasing the /proc/sys/net/ipv4/ip_conntrack_max even 4x, but still this does not fix the issue. Further more, as soon as OpenVZ (service vz restart) restart - the /proc/sys/net/ipv4/ip_conntrack_max comes back to initial value. Even though I have it hard written in the /etc/sysctl.conf file. So, I am actually not able to fix the issue myself. I was trying to do this for 2 weeks long and still no good solution. Most of google search advices to increase max conntack number. I did so. Others tell that this is possibly attack - I've blocked almost the whole internet - still nothing - conntrack stack gets full once per day. So, I hope OpenVZ gurus can advice something here.

Thanks and regards,
Maksym.

Subject: Re: ip_conntrack: CT 0: table full, dropping packet.
Posted by mavines on Tue, 27 Oct 2009 11:13:02 GMT
View Forum Message <> Reply to Message

in addition - all connections, which are in /proc/net/ip_conntrack
are to CT 300 IP to port 80. And CT 300 is the exact OpenVZ VPS which previously had issues with orphaned sockets. So, previously it throwed this error:
Orphaned socket dropped (187,374 in CT300)
For now it throws this:
ip_conntrack: CT 0: table full, dropping packet.

But still all connections which are in /proc/net/ip_conntrack goes to CT 300 - apache. The main hardware server (CT 0) does not have apache installed at all.

Regards,
Maksym.

Subject: Re: ip_conntrack: CT 0: table full, dropping packet.
Posted by kir on Tue, 27 Oct 2009 12:54:26 GMT
View Forum Message <> Reply to Message

Do you have any UBC failures for CT300?

Subject: Re: ip_conntrack: CT 0: table full, dropping packet.
Posted by mavines on Tue, 27 Oct 2009 14:04:35 GMT
View Forum Message <> Reply to Message

Here are UBC dump for CT300:

vzctl exec 300 cat /proc/user_beancounters
Version: 2.5

uid	resource	held	maxheld	barrier	limit	failcnt
300:	kmemsize	24701900	42827273	57490800	59160657	0
	lockedpages	0	0	1024	1024	0
	privvmpages	331898	420450	524288	524288	0
	shmpages	925	1581	86016	86016	0
	dummy	0	0	0	0	0
	numproc	235	368	960	960	0
	physpages	74167	94828	0	2147483647	0
	vmguarpages	0	0	135168	2147483647	0
	oomguarpages	74167	94828	104448	2147483647	0
	numtcpsock	103	303	1440	1440	0
	numflock	4	8	752	824	0
	numpty	1	2	64	64	0
	numsiginfo	0	12	1024	1024	0
	tcpsndbuf	855824	1366528	6881280	10813440	0
	tcprcvbuf	844976	1226352	6881280	10813440	0
	othersockbuf	356496	497248	4504320	8388608	0
	dgramrcvbuf	0	70464	1048576	1048576	0
	numothersock	165	194	1440	1440	0
	dcachesize	0	0	13639680	14499840	0
	numfile	3786	5048	37248	37248	0
	dummy	0	0	0	0	0
	dummy	0	0	0	0	0
	dummy	0	0	0	0	0
	numiptent	10	10	512	512	0