
Subject: iptables fail in CT latest OVZ kernel
Posted by [locutius](#) on Sun, 19 Oct 2008 23:54:22 GMT
[View Forum Message](#) <> [Reply to Message](#)

i have a server running this kernel:

```
Linux host0 2.6.18-53.1.19.el5.028stab053.14PAE #1 SMP Thu May 8 21:09:32 MSD 2008 i686  
i686 i386 GNU/Linux
```

and apf runs ok in the HN and in the CT

now i have a new box running this kernel:

```
Linux host1 2.6.18-92.1.1.el5.028stab057.2PAE #1 SMP Mon Jul 21 21:22:20 MSD 2008 i686  
i686 i386 GNU/Linux
```

and apf will NOT run in the CT

```
iptables: No chain/target/match by that name  
iptables: No chain/target/match by that name  
iptables: No chain/target/match by that name  
etc
```

after some investigation, i tried this command:

```
modprobe ipt_REJECT ipt_tos ipt_TOS ipt_LOG ip_conntrack ipt_limit ipt_multiport iptable_filter  
iptables_mangle ipt_TCPMSS ipt_tcpmss ipt_ttl ipt_length ipt_state iptable_nat ip_nat_ftp
```

and got this error:

```
FATAL: Error inserting ipt_REJECT  
(/lib/modules/2.6.18-92.1.1.el5.028stab057.2PAE/kernel/net/ipv4/netfilter/ipt_REJECT.ko):  
Unknown symbol in module, or unknown parameter (see dmesg)
```

nothing to see in dmesg

it appears there is a fault in the kernel

please advise

UPDATE:

i installed 2.6.18-53.1.19.el5.028stab053.14PAE on the new box and booted into the kernel but the problem is the same ...

... conclusion it isn't the kernel

need help plz

Subject: Re: kernel module load fail
Posted by [khorenko](#) on Mon, 20 Oct 2008 06:21:01 GMT
[View Forum Message](#) <> [Reply to Message](#)

Hi.

i think you simply do not have enough iptables modules loaded.
And please, note: you cannot load kernel modules from inside a Container - it's prohibited.
So, try to load necessary modules on the Hardware Node and reboot the Container.

Hope this helps.

--
Konstantin

Subject: Re: kernel module load fail
Posted by [locutius](#) on Mon, 20 Oct 2008 21:12:41 GMT
[View Forum Message](#) <> [Reply to Message](#)

yes that is all understood

both installations, working and not working, are set up using this guide

http://kb.parallels.com/article_130_875_en.html

both installations, the config files are identical

Subject: Re: kernel module load fail
Posted by [khorenko](#) on Tue, 21 Oct 2008 15:54:50 GMT
[View Forum Message](#) <> [Reply to Message](#)

Can you please

- 1) start both nodes - with working APF (node A) and node where APF does not work (node B)
 - 2) run APF inside a container on the node A
 - 3) save "lsmod" output from both nodes A and B (and post it here)
 - 4) post here Containers configs from both nodes as well as global vz configs (/etc/vz/vz.conf)
- ?

Thank you,
Konstantin

Subject: Re: kernel module load fail
Posted by [locutius](#) on Fri, 24 Oct 2008 05:59:04 GMT

node A (apf working in both HN and CT)

[root]# lsmod

Module	Size	Used by
vzethdev	16136	0
simfs	9068	55
vzrst	139156	0
vzcpt	114596	0
tun	15872	2 vzrst,vzcpt
vzdquota	44308	55 [permanent]
xt_tcpudp	7040	219
ip_nat_ftp	7808	0
xt_state	6144	28
xt_length	6016	0
ipt_ttl	5888	0
xt_tcpmss	6272	0
ipt_TCPMSS	8064	2
iptables_mangle	8576	57
xt_multiport	7168	8
xt_limit	6656	0
ipt_LOG	10240	0
ipt_TOS	6272	28
ipt_tos	5760	0
ipt_REJECT	9344	4
iptables_nat	13316	110
iptables_filter	8576	57
ip_nat	22288	3 vzrst,ip_nat_ftp,iptables_nat
ip_conntrack	60356	61 vzrst,vzcpt,ip_nat_ftp,xt_state,iptables_nat,ip_nat
nfnetlink	10648	2 ip_nat,ip_conntrack
ip_tables	18760	3 iptables_mangle,iptables_nat,iptables_filter
x_tables	19204	14 xt_tcpudp,xt_state,xt_length,ipt_ttl,xt_tcpmss,ipt_TCPMSS,xt_multiport,xt_limit,ipt_LOG,ipt_TOS,ipt_tos,ipt_REJECT,iptables_nat,ip_tables
autofs4	25220	2
vznetdev	21764	110
vzmon	46984	59 vzethdev,vzrst,vzcpt,vznetdev
ipv6	262048	504 vzrst,vzcpt,vzmon
vzdev	7556	4 vzethdev,vzdquota,vznetdev,vzmon
dm_mirror	28804	0
dm_multipath	21384	0
dm_mod	58776	2 dm_mirror,dm_multipath
video	19588	0
sbs	18468	0
backlight	9984	0
i2c_ec	8960	1 sbs
i2c_core	23552	1 i2c_ec
container	8320	0

```
button          10512 0
battery         13700 0
asus_acpi       19480 0
ac              9092 0
parport_pc     29092 0
lp             16168 0
parport        37960 2 parport_pc,lp
sg             35740 0
serio_raw      10628 0
ide_cd         39968 0
bnx2           155032 0
pcspkr         7040 0
cdrom          38048 1 ide_cd
ata_piix       18436 0
libata         116280 1 ata_piix
megaraid_sas   32048 3
sd_mod         24832 4
scsi_mod       133132 4 sg,libata,megaraid_sas,sd_mod
ext3           124424 2
jbd            61736 1 ext3
uhci_hcd       25356 0
ohci_hcd       23324 0
ehci_hcd       33036 0
```

```
/etc/sysconfig/vz-scripts/101.conf
```

```
# Configuration file generated by vzsplite for 2 VEs
# on HN with total amount of physical mem 8103 Mb
# low memory 811 Mb, swap size 10001 Mb, Max treads 8000
# Resource commit level 0:
# Free resource distribution. Any parameters may be increased
# Primary parameters
NUMPROC="4000:4000"
AVNUMPROC="1039:1039"
NUMTCPSOCK="4000:4000"
NUMOTHERSOCK="4000:4000"
VMGUARPAGES="1244660:2147483647"

# Secondary parameters
KMEMSIZE="1073741824:1073741824"
TCPSNDBUF="11995682:28379682"
TCPRCVBUF="11995682:28379682"
OTHERSOCKBUF="5997841:22381841"
DGRAMRCVBUF="5997841:5997841"
OOMGUARPAGES="1244660:2147483647"
PRIVVMPAGES="1244660:1369126"

# Auxiliary parameters
```

```
LOCKEDPAGES="4157:4157"
SHMPAGES="124466:124466"
PHYSPAGES="0:2147483647"
NUMFILE="33248:33248"
NUMFLOCK="1000:1100"
NUMPTY="400:400"
NUMSIGINFO="1024:1024"
DCACHESIZE="18593056:19150848"
NUMIPTENT="1600:1600"
DISKSPACE="10485760:10485760"
DISKINODES="7469229:8216152"
CPUUNITS="399024"
VE_ROOT="/vz/root/$VEID"
VE_PRIVATE="/vz/private/$VEID"
OSTEMPLATE="centos-5-i386-default"
ORIGIN_SAMPLE="2split"
ONBOOT="no"
NAMESERVER="xxx.xxx.xxx.xxx"
IP_ADDRESS="xxx.xxx.xxx.xxx"
HOSTNAME="www.xxxxx.com"
```

```
/etc/vz/vz.conf
```

```
## Global parameters
```

```
VIRTUOZZO=yes
LOCKDIR=/vz/lock
DUMPDIR=/vz/dump
VE0CPUUNITS=1000
```

```
## Logging parameters
```

```
LOGGING=yes
LOGFILE=/var/log/vzctl.log
LOG_LEVEL=10
VERBOSE=0
```

```
## Disk quota parameters
```

```
DISK_QUOTA=yes
VZFASTBOOT=no
```

```
# The name of the device whose ip address will be used as source ip for VE.
```

```
# By default automatically assigned.
```

```
#VE_ROUTE_SRC_DEV="eth0"
```

```
# Controls which interfaces to send ARP requests and modify APR tables on.
```

```
NEIGHBOUR_DEVS=all
```

```
## Template parameters
```

```
TEMPLATE=/vz/template
```

```

## Defaults for VEs
VE_ROOT=/vz/root/$VEID
VE_PRIVATE=/vz/private/$VEID
CONFIGFILE="vps.basic"
DEF_OSTEMPLATE="fedora-core-4"

## Load vzwdog module
VZWDOG="no"

## IPv4 iptables kernel modules
#IPTABLES="ipt_REJECT ipt_tos ipt_limit ipt_multiport iptable_filter iptable_mangle ipt_TCPMSS
ipt_tcpmss ipt_ttl ipt_length"
IPTABLES="ipt_REJECT ipt_tos ipt_TOS ipt_LOG ip_contrack ipt_limit ipt_multiport iptable_filter
iptable_mangle ipt_TCPMSS ipt_tcpmss ipt_ttl ipt_length ipt_state iptable_nat ip_nat_ftp"

## Enable IPv6
IPV6="no"

## IPv6 ip6tables kernel modules
IP6TABLES="ip6_tables ip6table_filter ip6table_mangle ip6t_REJECT"

```

node B (apf working in HN but not CT)

```
[root]# lsmod
```

```

vzethdev          16136  0
vznetdev          21124  4
simfs             9068  2
vzrst            141844  0
vzcpt            114724  0
tun              15872  2 vzrst,vzcpt
vzmon            49804  6 vzethdev,vznetdev,vzrst,vzcpt
ip_nat_ftp        7808  0
iptable_nat      13316  4
ip_nat           22288  3 vzrst,ip_nat_ftp,iptable_nat
xt_length         6016  0
ipt_ttl          5888  0
xt_tcpmss         6272  0
ipt_LOG          10240  0
ipt_tos          5760  0
vzdquota         45204  2 [permanent]
xt_tcpudp        7040  122
xt_state         6144  16
ipt_TCPMSS        8064  1
iptable_mangle    8576  3
xt_multiport     7168  4
xt_limit         6656  6

```

```

ipt_TOS          6272 18
ipt_REJECT       9344 2
ip_conntrack    60356 8 vzrst,vzcpt,ip_nat_ftp,iptable_nat,ip_nat,xt_state
iptables_filter  8576 3
nfnetlink       10648 2 ip_nat,ip_conntrack
ip_tables       18760 3 iptable_nat,iptable_mangle,iptable_filter
x_tables        19204 14 iptable_nat,xt_length,ipt_ttl,xt_tcpmss,ipt_LOG,ipt_tos,xt_t
cpudp,xt_state,ipt_TCPMSS,xt_multiport,xt_limit,ipt_TOS,ipt_REJECT,ip_tables
ipv6            269824 25 vzrst,vzcpt,vzmon
xfrm_nalgo      13700 1 ipv6
crypto_api      11904 1 xfrm_nalgo
vzdev           7556 4 vzethdev,vznetdev,vzmon,vzdquota
dm_mirror       29188 0
dm_multipath    22024 0
dm_mod          62108 2 dm_mirror,dm_multipath
video           21640 0
sbs             18468 0
backlight       9984 1 video
i2c_ec          8960 1 sbs
i2c_core        23552 1 i2c_ec
container       8320 0
button          10512 0
battery         13700 0
asus_acpi       19480 0
ac              9092 0
parport_pc      29092 0
lp              16168 0
parport         37960 2 parport_pc,lp
sg              35868 0
ide_cd          39968 0
pcspkr          7040 0
bnx2            138780 0
i5000_edac      12416 0
edac_mc         26192 1 i5000_edac
cdrom           38048 1 ide_cd
serio_raw       10628 0
ata_piix        22276 0
libata          144700 1 ata_piix
mptsas          37512 3
mptscsih        36864 1 mptsas
mptbase         75812 2 mptsas,mptscsih
scsi_transport_sas 30464 1 mptsas
sd_mod          24832 4
scsi_mod        135180 6 sg,libata,mptsas,mptscsih,scsi_transport_sas,sd_mod
ext3            124552 2
jbd             61736 1 ext3
uhci_hcd        25356 0
ohci_hcd        23324 0

```

ehci_hcd 33676 0

/etc/sysconfig/vz-scripts/101.conf

```
# Configuration file generated by vzsplite for 2 VEs
# on HN with total amount of physical mem 2020 Mb
# low memory 872 Mb, swap size 4000 Mb, Max treads 8000
# Resourse commit level 0:
# Free resource distribution. Any parameters may be increased
# Primary parameters
NUMPROC="4000:4000"
AVNUMPROC="1117:1117"
NUMTCPSOCK="4000:4000"
NUMOTHERSOCK="4000:4000"
VMGUARPAGES="340736:2147483647"

# Secondary parameters
KMEMSIZE="91529625:100682587"
TCPSNDBUF="279131136:410203136"
TCPRCVBUF="279131136:410203136"
OTHERSOCKBUF="279131136:410203136"
DGRAMRCVBUF="279131136:410203136"
OOMGUARPAGES="340736:2147483647"
PRIVVMPAGES="255999744:255999744"

# Auxiliary parameters
LOCKEDPAGES="4469:4469"
SHMPAGES="340736:340736"
PHYSPAGES="0:2147483647"
NUMFILE="35744:35744"
NUMFLOCK="1000:1100"
NUMPTY="400:400"
NUMSIGINFO="1024:1024"
DCACHESIZE="19988877:20588544"
NUMIPTENT="200:200"
DISKSPACE="3145728:3145728"
DISKINODES="1572864:1572864"
CPUUNITS="155175"
VE_ROOT="/vz/root/$VEID"
VE_PRIVATE="/vz/private/$VEID"
OSTEMPLATE="centos-5-i386-default"
ONBOOT="yes"
NAMESERVER="xxx.xxx.xxx.xxx"
HOSTNAME="www.xxxx.com"
IP_ADDRESS="xxx.xxx.xxx.xxx"
MEMINFO="pages:255999744"
CPULIMIT="400"
CPUS="4"
```

/etc/vz/vz.conf

Global parameters

VIRTUOZZO=yes
LOCKDIR=/vz/lock
DUMPDIR=/vz/dump
VE0CPUUNITS=1000

Logging parameters

LOGGING=yes
LOGFILE=/var/log/vzctl.log
LOG_LEVEL=10
VERBOSE=0

Disk quota parameters

DISK_QUOTA=yes
VZFASTBOOT=no

The name of the device whose ip address will be used as source ip for VE.
By default automatically assigned.
#VE_ROUTE_SRC_DEV="eth0"

Controls which interfaces to send ARP requests and modify APR tables on.
NEIGHBOUR_DEVS=all

Template parameters

TEMPLATE=/vz/template

Defaults for VEs

VE_ROOT=/vz/root/\$VEID
VE_PRIVATE=/vz/private/\$VEID
CONFIGFILE="vps.basic"
DEF_OSTEMPLATE="fedora-core-4"

Load vzwdog module

VZWDOG="no"

IPv4 iptables kernel modules

#IPTABLES="ipt_REJECT ipt_tos ipt_limit ipt_multiport iptable_filter iptable_mangle ipt_TCPMSS
ipt_tcpmss ipt_ttl ipt_length"
IPTABLES="ipt_REJECT ipt_tos ipt_TOS ipt_LOG ip_conntrack ipt_limit ipt_multiport iptable_filter
iptable_mangle ipt_TCPMSS ipt_tcpmss ipt_ttl ipt_length ipt_state iptable_nat ip_nat_ftp"

Enable IPv6

IPV6="no"

IPv6 ip6tables kernel modules

IP6TABLES="ip6_tables ip6table_filter ip6table_mangle ip6t_REJECT"

thanks for the help. i also will compare these values and see what is different

EDIT:

i have compared the modules for A and B. i cannot see a problem

Subject: Re: kernel module load fail
Posted by [khorenko](#) on Fri, 24 Oct 2008 12:32:46 GMT
[View Forum Message](#) <> [Reply to Message](#)

well, i also did not see any significant difference...
Ok, let's go another way: APF is a `_script_`.
Could you please add at the beginning "set -x" line and start APF?
In that way you'll find a iptables command that fails.

Hope to hear from you soon...

--
Konstantin

Subject: Re: kernel module load fail
Posted by [locutius](#) on Sun, 26 Oct 2008 09:52:32 GMT
[View Forum Message](#) <> [Reply to Message](#)

node B:

```
[root]# apf --start 2> apflog.txt
```

apflog.txt is without set -x
apflog1.txt is with set -x

everything iptables is failing

iptables is running in the CT. it is almost as if there is a problem with permissions

i have another HN to set up today. i cross my fingers for no problems

UPDATE:

i have installed OVZ on a new server node C. bad news: node C is behaving exactly the same as node B, apf will not run in a CT

i made an experiment and copied over the CT with working apf from node A to node C. when i started the previously good CT on node C then apf failed to work with all the same errors.
conclusion: the problem is not in the CT config

there were other problems with node B that i did not report because i did not think were interesting. however, i see the same errors on node B and node C (both are many CT running httpd and mysqld):

1. the HN will not serve webpages longer than 24 hours before they stop. restarting the CT does not fix the problem. restarting vz service does not fix the problem. rebooting the HN solves it for another 24 hours before it stops again

2. stopping a CT on both B and C gives this error:

```
Message from syslogd@ at Mon Oct 27 10:08:44 2008 ...
host1 kernel: unregister_netdevice: device f5e56000 marked to leak
Message from syslogd@ at Mon Oct 27 10:08:44 2008 ...
host1 kernel: free_netdev: device venet0=f5e56000 leaked
Message from syslogd@ at Mon Oct 27 10:08:47 2008 ...
host1 kernel: unregister_netdevice: device f5e56800 marked to leak
Message from syslogd@ at Mon Oct 27 10:08:47 2008 ...
host1 kernel: free_netdev: device lo=f5e56800 leakedVE was stopped
VE is unmounted
```

3. stopping a CT on node C gives the broken pipe error

today i will make an experiment and stop iptables on the HN and the CT to test if the CT will serve for longer than 24 hours with iptables stopped

i appreciate the help you give. i have just completed a migration of 100+ vps from one datacentre to another, from older reliable OVZ that stays up for 6 months without intervention to new OVZ that cannot stay up for 24 hours

i am pointing the finger at the kernel

while building template cache on node C vzpkgcache failed with the error "syslog file not found" when syslogd is running

if you want me to raise bugs on any of these issues please tell me. i will work on this until a solution is found

UPDATE:

node B and C running for 4 hours with iptables disabled give the same leak error when stopping a CT. iptables is not the culprit. it is the OVZ network service causing the leak

File Attachments

- 1) [apflog.txt](#), downloaded 521 times
 - 2) [apflog1.txt](#), downloaded 546 times
-

Subject: Re: kernel module load fail
Posted by [locutius](#) on Tue, 28 Oct 2008 17:17:09 GMT
[View Forum Message](#) <> [Reply to Message](#)

the failure of iptables in the CT with the latest OVZ kernel is a major problem for this project. the CT need control, of their own iptables rules. scripting tools like apf are not easy to deploy on a per CT basis from the HN

after using OVZ for 18 months i am forced to abandon OVZ. OVZ is not fit for a production environment running latest CentOS 5 on new Dell 1950iii

i am now experimenting with VMware. i will return and update you with a comparison between OVZ and VMware

thanks for the effort Konstantin. nothing more to do, OVZ is removed from the new servers. sad day for this project

Subject: Re: kernel module load fail
Posted by [khorenko](#) on Wed, 29 Oct 2008 12:21:18 GMT
[View Forum Message](#) <> [Reply to Message](#)

locutius,

looks like you are in a bad mood at the moment.
Please, do not waste nerves for these problems.

Quote:OVZ is not fit for a production environment...

And talking about this...

Look... You do want to run something in a production, right? And do it without any support of those programs you are using? Well.. i think you'll agree with me - it's a bit risky... And now you faced a problem that you cannot solve without support - that happens sometime! There are too

many ways to configure incorrectly something or on the other hand - any software contains bugs or might miss some functionality you need...

i'm not going to persuade you to return back to using OVZ, but please, think about my words - this place is just a _forum_, we try to help people but really not in a priority... I'd just say - if you run some serious production system, its support is really desired (this is true for ANY system IMHO, not only OVZ), i know for sure that issues of the people that have support for VZ or OVZ (yes, you can buy support for OVZ!) are being resolved with much higher priority; and if you need some additional functionality, your request will be of course more significant.

But in any case - good luck and don't be so worried!

--

Konstantin
