Subject: openvpn and tap/bridging with a "dummy" network adapter?
Posted by snowdeal on Thu, 23 Feb 2006 03:29:05 GMT
View Forum Message <> Reply to Message

i'm trying to get a bridged (tap) network running within a vps using the same techniques outlined in a tutorial for doing the same on a jvds virtual server [1]. i understand i need to get bridging support compiled into the kernel, which i'm working on [2] but while i'm working on that i'm also having problems starting the "dummy" network adapter. if i issue the command in the tutorial [3] within a vps i get an error [4].

is it possible to create a "fake" lan nic as described in the jvds howto with openvz?

[1]http://www.jvds.com/guide/bridging.php
[2] http://forum.openvz.org/index.php?t=tree&th=288&mid= 1624&&rev=&reveal=
[3]ifconfig dummy0 hw ether de:ad:be:ef:00:00
[4]SIOCSIFHWADDR: Operation not permitted

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by dev on Thu, 23 Feb 2006 10:20:36 GMT
View Forum Message <> Reply to Message

1. have you got dummy network device in VPS?
vps# ifconfig -a

Did you simply move it from host via vzctl set --netdev_add?

2. VPS have no rights to change MAC addresses by default.
This is a security limitation.

So you need to give CAP_NET_ADMIN capability to your VPS:
host# vzctl set VPSID --capability net_admin:on

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by snowdeal on Fri, 24 Feb 2006 01:45:14 GMT
View Forum Message <> Reply to Message

dev wrote on Thu, 23 February 2006 05:201. have you got dummy network device in VPS?
vps# ifconfig -a

Did you simply move it from host via vzctl set --netdev_add?

i do now thanks for the hint. i created it on the hardware node and moved it to the vps.

dev wrote on Thu, 23 February 2006 05:202. VPS have no rights to change MAC addresses by default.
This is a security limitation.

So you need to give CAP_NET_ADMIN capability to your VPS:
host# vzctl set VPSID --capability net_admin:on

thanks for that tip, but since i'm creating the dummy device on the hw node, i don't need to change the mac address in the vps.

but! i now have a new problem. after i create the dummy device and move it to the vps, inside the vps i then create the tap interface [1] but when i try to create br0 (obviously i've installed bridge-utils in the vps), i get an error [3]. perhaps i don't have another "capability option" turned on that's off by default? perhaps sys_admin or ve_admin?

is there a "capability option" that would allow me to create bridges in the vps?

is there a more detailed explanation of various capability options beyond the vzctl man page [4]?

[1] openvpn --mktun --dev tap0
[2] brctl addbr br0
[3] add bridge failed: Operation not permitted
[4] http://openvz.org/documentation/mans/vzctl.8

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by dev on Fri, 24 Feb 2006 09:02:27 GMT
View Forum Message <> Reply to Message

from what I see from sources:

```
case BRCTL_ADD_BRIDGE:
case BRCTL_DEL_BRIDGE:
{
        char buf[IFNAMSIZ];

        if (!capable(CAP_NET_ADMIN))
                return -EPERM;

        if (copy_from_user(buf, (void __user *)args[1], IFNAMSIZ))
                return -EFAULT;

        buf[IFNAMSIZ-1] = 0;
```

```
        if (args[0] == BRCTL_ADD_BRIDGE)
                return br_add_bridge(buf);

        return br_del_bridge(buf);
}
```

CAP_NET_ADMIN (net_admin) should be enough...
you can add all the capabilities to your VPS at first, we can resolve this later. But `strace -f brctl addbr br0` would help to do it now))

We don't have description of capabilities in the man page, since this is a standart security model of Linux kernel...

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by snowdeal on Fri, 24 Feb 2006 19:38:49 GMT
View Forum Message <> Reply to Message

from the vps config file, i can verify that net_admin is set [1]. results from strace [2].

[1] CAPABILITY="NET_ADMIN:on"
[2]
```
strace -f /usr/sbin/brctl addbr br0
execve("/usr/sbin/brctl", ["/usr/sbin/brctl", "addbr", "br0"], [/* 19 vars */]) = 0
uname({sys="Linux", node="foo.bar.com", ...}) = 0
brk(0)                                  = 0x804e000
access("/etc/ld.so.preload", R_OK)      = -1 ENOENT (No such file or directory)
open("/etc/ld.so.cache", O_RDONLY)      = 3
fstat64(3, {st_mode=S_IFREG|0644, st_size=20708, ...}) = 0
old_mmap(NULL, 20708, PROT_READ, MAP_PRIVATE, 3, 0) = 0x40017000
close(3)                                = 0
open("/usr/lib/libsysfs.so.1", O_RDONLY) = 3
read(3, "\177ELF\1\1\1\0\0\0\0\0\0\0\0\0\3\0\3\0\1\0\0\0\234#\0"..., 512) = 512
fstat64(3, {st_mode=S_IFREG|0755, st_size=38652, ...}) = 0
old_mmap(NULL, 37484, PROT_READ|PROT_EXEC, MAP_PRIVATE|MAP_DENYWRITE, 3, 0)
= 0x4001d000
old_mmap(0x40026000, 4096, PROT_READ|PROT_WRITE,
MAP_PRIVATE|MAP_FIXED|MAP_DENYWRITE, 3, 0x9000) = 0x40026000
close(3)                                = 0
open("/lib/tls/libc.so.6", O_RDONLY)    = 3
read(3, "\177ELF\1\1\1\0\0\0\0\0\0\0\0\0\3\0\3\0\1\0\0\0\20O\1\000"..., 512) = 512
fstat64(3, {st_mode=S_IFREG|0755, st_size=1451366, ...}) = 0
old_mmap(NULL, 4096, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_ANONYMOUS, -1,
0) = 0x40027000
old_mmap(NULL, 1219772, PROT_READ|PROT_EXEC, MAP_PRIVATE|MAP_DENYWRITE, 3,
0) = 0x40028000
```

```
old_mmap(0x4014c000, 16384, PROT_READ|PROT_WRITE,
MAP_PRIVATE|MAP_FIXED|MAP_DENYWRITE, 3, 0x124000) = 0x4014c000
old_mmap(0x40150000, 7356, PROT_READ|PROT_WRITE,
MAP_PRIVATE|MAP_FIXED|MAP_ANONYMOUS, -1, 0) = 0x40150000
close(3)                                = 0
old_mmap(NULL, 4096, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_ANONYMOUS, -1,
0) = 0x40152000
mprotect(0x4014c000, 4096, PROT_READ)   = 0
mprotect(0x40015000, 4096, PROT_READ)   = 0
set_thread_area({entry_number:-1 -> 6, base_addr:0x401526c0, limit:1048575, seg_32bit:1,
contents:0, read_exec_only:0, limit_in_pages:1, seg_not_present:0, useable:1}) = 0
munmap(0x40017000, 20708)               = 0
socket(PF_INET, SOCK_STREAM, IPPROTO_IP) = 3
brk(0)                                  = 0x804e000
brk(0x806f000)                          = 0x806f000
open("/proc/mounts", O_RDONLY)          = 4
fstat64(4, {st_mode=S_IFREG|0444, st_size=0, ...}) = 0
mmap2(NULL, 4096, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_ANONYMOUS, -1, 0) =
0x40017000
read(4, "simfs / simfs rw 0 0\nproc /proc "..., 1024) = 85
read(4, "", 1024)                       = 0
close(4)                                = 0
munmap(0x40017000, 4096)                = 0
ioctl(3, 0x89a0, 0xbffffc6f)            = -1 EPERM (Operation not permitted)
ioctl(3, SIOCSIFBR, 0xbffff0e0)         = -1 EPERM (Operation not permitted)
write(2, "add bridge failed: Operation not"..., 43add bridge failed: Operation not permitted
) = 43
exit_group(1)                           = ?
Process 3987 detached
```

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by dev on Fri, 24 Feb 2006 21:31:04 GMT

View Forum Message <> Reply to Message

ok, resolved. can you please apply the following patch?

```
--- ./net/core/dev.c.tst 2006-01-11 22:13:22.000000000 +0300
+++ ./net/core/dev.c 2006-02-25 00:29:45.000000000 +0300
@@ -3036,10 +3036,6 @@
  /* When net_device's are persistent, this will be fatal. */
  BUG_ON(dev->reg_state != NETREG_UNINITIALIZED);

- ret = -EPERM;
- if (!ve_is_super(get_exec_env()) && ve_is_dev_movable(dev))
-  goto out;
-
```

```
spin_lock_init(&dev->queue_lock);
spin_lock_init(&dev->xmit_lock);
dev->xmit_lock_owner = -1;
```

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by snowdeal on Sat, 25 Feb 2006 01:28:26 GMT
View Forum Message <> Reply to Message

it works!

thanks for all your help.

will this patch make it into future releases?

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by dev on Sun, 26 Feb 2006 08:49:33 GMT
View Forum Message <> Reply to Message

have you succeeded making bridged VPN?
I'm not sure this patch will get into OVZ kernel, since it has security implications.
I will try to think how to make it clean and correct way.

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by snowdeal on Sun, 26 Feb 2006 14:56:01 GMT
View Forum Message <> Reply to Message

dev wrote on Sun, 26 February 2006 03:49have you succeeded making bridged VPN?
I'm not sure this patch will get into OVZ kernel...

yes, thanks for all your help, the bridged vpn inside the vps is working perfectly.

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by dev on Sun, 26 Feb 2006 19:26:06 GMT
View Forum Message <> Reply to Message

huh, nice! you helped me very much, because I was totally out of time!!! Thanks, thanks, thanks!
can you post brief HOWTO for this?

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by snowdeal on Mon, 27 Feb 2006 00:40:09 GMT
View Forum Message <> Reply to Message

dev wrote on Sun, 26 February 2006 14:26huh, nice! you helped me very much, because I was totally out of time!!! Thanks, thanks, thanks!
can you post brief HOWTO for this?

i can certainly post a brief howto - it's the least i can do in exchange for all your help.

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by dev on Mon, 27 Feb 2006 08:22:05 GMT
View Forum Message <> Reply to Message

I will be very obliged

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by snowdeal on Sun, 12 Mar 2006 04:31:15 GMT
View Forum Message <> Reply to Message

apologies for taking so long to post the HOWTO. i've been testing the setup and it's working well, but recently i've tried to add multiple "dummy" interfaces (one for each vps) and am running in problems.

as you might recall, i'm bridging a dummy interface using a procedure similar to that outlined here [1]. i can add dummy0 [2] which i can move to a vps [3] but if i try to add a second dummy interface (for a second vps) i get an error [4]. so it would seem to make sense that i have to insmod more than 1 instance of the dummy module using a scheme similar to that outline here [5]. but when i try to load an additional dummy interface i get another error [6].

so, can anyone give my any hints on how to create multiple dummy interfaces on the hardware node so that each vps can have its own dummy interface?

[1] http://www.jvds.com/guide/bridging.php
[2] ifconfig dummy0 hw ether de:ad:be:ef:00:00
[3] /usr/sbin/vzctl set 101 --netdev_add dummy0 --save
[4] ifconfig dummy1 hw ether de:ad:be:ef:00:01
SIOCSIFHWADDR: No such device
[5] http://groups.google.com/group/comp.os.linux.networking/brow
se_thread/thread/f08148fe2a7d5747/291d1410df166fe6?lnk=st&am
p;q=multiple+dummy+interfaces&rnum=10&hl=en#291d1410 df166fe6

[6] /sbin/insmod -o dummy1 dummy.o
insmod: can't read '-o': No such file or directory

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by Jason Stubbs on Sun, 12 Mar 2006 08:07:00 GMT
View Forum Message <> Reply to Message

You need to use the -o option to modprobe in order to give the module a different name. The you given it is arbitrary and just needs to be unique.

```
# modprobe dummy ; ifconfig dummy0 up
# modprobe -o dummy1 dummy ; ifconfig dummy1 up
# modprobe -o dummy3 dummy ; ifconfig dummy3 up
dummy3: unknown interface: No such device
# ifconfig dummy2 up
# modprobe -o foobar dummy ; ifconfig dummy3 up
# ifconfig | grep dummy
dummy0    Link encap:Ethernet  HWaddr EA:2A:12:AE:C3:A3
dummy1    Link encap:Ethernet  HWaddr 96:3E:2B:F2:39:AC
dummy2    Link encap:Ethernet  HWaddr EA:BD:69:9F:54:8F
dummy3    Link encap:Ethernet  HWaddr EA:DD:D2:DE:F2:54
# lsmod | grep -E '(dummy|foobar)'
foobar                3400  0
dummy3                3400  0
dummy1                3400  0
dummy                 3400  0
```

Subject: Re: openvpn and tap/bridging with a "dummy" network adapter?
Posted by bensig on Fri, 21 Dec 2012 03:27:53 GMT
View Forum Message <> Reply to Message

Trying to do the same thing as the above user... adding a dummy device to VPS and get bridge going on Openvpn running on centos 5:

```
[root@system /]# brctl addbr br0
add bridge failed: Inappropriate ioctl for device

[root@us /]# strace -f brctl addbr br0
execve("/usr/sbin/brctl", ["brctl", "addbr", "br0"], [/* 17 vars */]) = 0
brk(0)                                  = 0x8ff000
mmap(NULL, 4096, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_ANONYMOUS, -1, 0) =
0x7f7d59fba000
mmap(NULL, 4096, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_ANONYMOUS, -1, 0) =
0x7f7d59fb9000
access("/etc/ld.so.preload", R_OK)      = -1 ENOENT (No such file or directory)
```

```
open("/etc/ld.so.cache", O_RDONLY)      = 3
fstat(3, {st_mode=S_IFREG|0644, st_size=21438, ...}) = 0
mmap(NULL, 21438, PROT_READ, MAP_PRIVATE, 3, 0) = 0x7f7d59fb3000
close(3)                                = 0
open("/usr/lib64/libsysfs.so.2", O_RDONLY) = 3
read(3, "\177ELF\2\1\1\0\0\0\0\0\0\0\0\0\3\0>\0\1\0\0\0\320&\0\0\0\0\0\0 "..., 832) = 832
fstat(3, {st_mode=S_IFREG|0755, st_size=43952, ...}) = 0
mmap(NULL, 2139280, PROT_READ|PROT_EXEC, MAP_PRIVATE|MAP_DENYWRITE, 3, 0) =
0x7f7d59b94000
mprotect(0x7f7d59b9e000, 2097152, PROT_NONE) = 0
mmap(0x7f7d59d9e000, 4096, PROT_READ|PROT_WRITE,
MAP_PRIVATE|MAP_FIXED|MAP_DENYWRITE, 3, 0xa000) = 0x7f7d59d9e000
close(3)                                = 0
open("/lib64/libc.so.6", O_RDONLY)      = 3
read(3, "\177ELF\2\1\1\0\0\0\0\0\0\0\0\0\3\0>\0\1\0\0\0\220\332\1\0\0\0\0\0 "..., 832) = 832
fstat(3, {st_mode=S_IFREG|0755, st_size=1712648, ...}) = 0
mmap(NULL, 3498328, PROT_READ|PROT_EXEC, MAP_PRIVATE|MAP_DENYWRITE, 3, 0) =
0x7f7d5983d000
mprotect(0x7f7d5998b000, 2093056, PROT_NONE) = 0
mmap(0x7f7d59b8a000, 20480, PROT_READ|PROT_WRITE,
MAP_PRIVATE|MAP_FIXED|MAP_DENYWRITE, 3, 0x14d000) = 0x7f7d59b8a000
mmap(0x7f7d59b8f000, 16728, PROT_READ|PROT_WRITE,
MAP_PRIVATE|MAP_FIXED|MAP_ANONYMOUS, -1, 0) = 0x7f7d59b8f000
close(3)                                = 0
mmap(NULL, 4096, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_ANONYMOUS, -1, 0) =
0x7f7d59fb2000
mmap(NULL, 4096, PROT_READ|PROT_WRITE, MAP_PRIVATE|MAP_ANONYMOUS, -1, 0) =
0x7f7d59fb1000
arch_prctl(ARCH_SET_FS, 0x7f7d59fb16e0) = 0
mprotect(0x7f7d59b8a000, 16384, PROT_READ) = 0
mprotect(0x7f7d59fbb000, 4096, PROT_READ) = 0
munmap(0x7f7d59fb3000, 21438)           = 0
socket(PF_INET, SOCK_STREAM, IPPROTO_IP) = 3
lstat("/sys/class/net", {st_mode=S_IFDIR|0755, st_size=0, ...}) = 0
brk(0)                                  = 0x8ff000
brk(0x920000)                           = 0x920000
ioctl(3, 0x89a0, 0x7fffe7952ca9)        = -1 ENOTTY (Inappropriate ioctl for device)
ioctl(3, SIOCSIFBR, 0x7fffe7951f60)     = -1 ENOTTY (Inappropriate ioctl for device)
write(2, "add bridge failed: Inappropriate"..., 50add bridge failed: Inappropriate ioctl for device
) = 50
exit_group(1)                           = ?
```

not sure why it's failing here... any help?