
Subject: Bridge inside VE???

Posted by [bobomagic](#) on Thu, 30 Aug 2007 19:04:45 GMT

[View Forum Message](#) <> [Reply to Message](#)

Is there really no way to have bridging inside a VE?

What would it take to make this possible?

Thanks!

Subject: Re: Bridge inside VE???

Posted by [khorenko](#) on Fri, 31 Aug 2007 13:16:23 GMT

[View Forum Message](#) <> [Reply to Message](#)

To achieve the ability to create bridges inside a VE you have at least:

1) add CAP_NET_ADMIN capability to this VE (via vzctl)

2) patch the kernel to add NETIF_F_VIRTUAL flag to bridge device

```
--- ./net/bridge/br_device.c.br 2007-06-25 13:36:46.000000000 +0400
```

```
+++ ./net/bridge/br_device.c 2007-08-31 17:12:39.000000000 +0400
```

```
@@ -216,5 +216,6 @@ void br_dev_setup(struct net_device *dev
```

```
    dev->priv_flags = IFF_EBRIDGE;
```

```
    dev->features = NETIF_F_SG | NETIF_F_FRAGLIST | NETIF_F_HIGHDMA |  
-    NETIF_F_TSO | NETIF_F_NO_CSUM | NETIF_F_GSO_ROBUST;  
+    NETIF_F_TSO | NETIF_F_NO_CSUM | NETIF_F_GSO_ROBUST |  
+    NETIF_F_VIRTUAL;  
}
```

at the glance that should be enough.

But could you please tell me why do you want this???

Thanks!

Subject: Re: Bridge inside VE???

Posted by [bobomagic](#) on Fri, 31 Aug 2007 16:45:49 GMT

[View Forum Message](#) <> [Reply to Message](#)

I'm trying to create a test environment for an existing network setup. My existing environment has typical network nodes behind a security bridge. There are several of these, and the security bridges connect in a star topology, to allow the typical network nodes to communicate.

I want a test environment for pre-production burn-in of new code for the security bridge, and it should exactly mimic the production environment.

I'll give your patch a try!

Subject: Re: Bridge inside VE???

Posted by [bobomagic](#) on Thu, 06 Sep 2007 15:47:20 GMT

[View Forum Message](#) <> [Reply to Message](#)

I was patching against kernel-2.6.16-1.2133_FC5.026test015 and had to modify your patch a little, to:

```
--- linux-2.6.16/net/bridge/br_device.c.orig 2007-09-05 22:04:10.000000000 -0700
+++ linux-2.6.16/net/bridge/br_device.c 2007-09-05 22:10:29.000000000 -0700
@@ -185,6 +185,7 @@
     dev->set_mac_address = br_set_mac_address;
     dev->priv_flags = IFF_EBRIDGE;

-    dev->features = NETIF_F_SG | NETIF_F_FRAGLIST
-    | NETIF_F_HIGHDMA | NETIF_F_TSO | NETIF_F_IP_CSUM;
+    dev->features = NETIF_F_SG | NETIF_F_FRAGLIST | NETIF_F_HIGHDMA |
+    NETIF_F_TSO | NETIF_F_NO_CSUM |
+    NETIF_F_VIRTUAL;
 }
```

I then copied brctl binary and sysfs libraries to VE, ran
% brctl addbr vbr0

and got a kernel dump:

hipvz kernel: -----[cut here]-----

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: kernel BUG at fs/sysfs/bin.c:187!

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: invalid opcode: 0000 [#1]

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: CPU: 0, VCPU: 101:0

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: EIP is at sysfs_create_bin_file+0x1b/0x36

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: eax: e2bfbb01 ebx: e0b929e0 ecx: c2979214 edx: e0b92900

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: esi: 00000001 edi: c2979000 ebp: c2979214 esp: c2980edc

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: ds: 007b es: 007b ss: 0068

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: Process brctl (pid: 12300, veid=101, threadinfo=c2980000 task=cf7af2f0)

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: Stack: <0>00000000 e0b8c227 00000000 00000001 c2979744 c2979000 e0b8969f 00000001

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: dead4ead ffffffff ffffffff 00000001 dead4ead ffffffff ffffffff 000089a0

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: bfe02d02 c2980f40 de692d00 e0b8a3c6 c6dd0180 000005dc c2977008 b7f04000

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: Call Trace:

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: [<e0b8c227>] br_sysfs_addbr+0x48/0xe2 [bridge] [<e0b8969f>]
br_add_bridge+0x19a/0x1c5 [bridge]

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: [<e0b8a3c6>] br_ioctl_deviceless_stub+0x18d/0x1b3 [bridge] [<c0151635>]
free_pgtables+0x69/0x78

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: [<e0b8a239>] br_ioctl_deviceless_stub+0x0/0x1b3 [bridge] [<c029b15b>]
sock_ioctl+0x127/0x232

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: [<c029b034>] sock_ioctl+0x0/0x232 [<c01706ce>] do_ioctl+0x16/0x48

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: [<c01708ff>] vfs_ioctl+0x1ff/0x216 [<c017095e>] sys_ioctl+0x48/0x62

Message from syslogd@hipvz at Thu Sep 6 08:26:40 2007 ...

hipvz kernel: [<c0102c5d>] syscall_call+0x7/0xb <0>Code: 00 74 0a 8b 41 30 8b 12 e8 8f df ff 31 c0 c3 53 89 c1 89 d3 85 c0 74 12 83 78 30 00 0f 94 c0 85 d2 0f 94 c2 09 d0 a8 01 74 0b <0f> 0b 66 b8 bb 00 b8 44 cf 31 c0 8b 41 30 b9 08 00 00 00 89 da

Also, CAP_NET_ADMIN is ON for this VE.

Any other ideas?

Thanks!!!
