Subject: [Devel] 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by mhw on Sat, 03 Mar 2007 17:21:05 GMT

View Forum Message <> Reply to Message

Hello,

 Introduction... I've been on the users list for a while and just
joined the developers list. Excuse the cross post but I'm not sure
which would be most appropriate.

 I work extensively with IPv6. All of my devices are configured for
IPv6 and I actively participate in the global IPv6 network and have for
over 5 years. I'm a member of the North American v6 Task Force,
NAv6TF.org. My name servers, ntp servers, smtp servers, and web servers
all show active IPv6 traffic and have for years. I'm also a big user of
VMware and I'm also the author of an IPv6 paper that's up on their site
somewhere.

 So, I'll prefix this with a bit of a rant.

 <Rant>

 I don't use venet. The way it's implemented, especially vis-a-vis
IPv6, it makes no sense to me at all. the venet0 devices have no
hardware / mac layer addresses so how is IPv6 neighbor discovery and
autoconfiguration protocols suppose to work? These are really
fundamental to the nature of IPv6 and, unless you want to use IPv6
purely like it just IPv4 with fat addresses (which it is NOT) and live
entirely in an IPv4 mind-think cocoon when working with IPv6, how can
you expect this to work with fully implemented IPv6?

 The bridged networking of the veth devices seem to be the only viable
way to deal with this. I HAVE tried using venet for IPv6 and failed
miserably.

 This is also how I deal with IPv6 on VMware. I never use the local
networking (host local or NAT) devices and only use the bridged devices
there. In the case of VMware, the bridged devices are the default and
you can add the routed or nat'ed devices. That makes a lot more sense
to me.

 So IPv6 HAS to work for me. Consequently, I have to work with
everything running over the veth devices. The venet devices just do not
cut it at all with IPv6, so are totally useless to me for both IPv4 as
well as IPv6 (why split the network two ways).

 Funny thing is, if you "turn on" IPv6 in the vz configuration files, it
dicks things up royally by trying to force IPv6 routes out through the

venet0 interface, which can't work, and interferes with stateless autoconfiguration and router discovery configuration. Leaving IPv6 off in the config, then it doesn't insert bogus routes and everything configures properly on the veth devices and IPv6 works like a charm. Interestingly enough, the veth devices DO get their proper link-local and global-unicast addresses while vnet0 sits there with not even a link-local address (of course - it has no mac address so how could it possibly have a link-local address - duh...).

 </Rant>

 Now... Onto the problem. IPv6 is working sweet with 2.6.18-ovz028test010 on Fedora Core 6. It would be nice to get it up to the FC6 2.6.19 (soon to be the 2.6.20) rebase but that's ok for now. When I tried 2.6.18-ovz028test015 and subsequently 2.6.18-ovz028test018, I found that IPv6 was totally broken. While all my interfaces properly autoconfigure on test10, stateless autoconf appears to be totally broken on test15 and test18 and even manual configuration doesn't work (hints strongly at broken neighbor discovery there).

 This may not be an OpenVZ problem, per se, though. Some time in the very later part of the 2.6.18 kernel.org release updates, a problem was introduced that broke IPv6. Some patch or another caused nodes to fail to join the "all nodes multicast" (ff02::1) address. This is critical to IPv6 router discovery and advertisement, neighbor discovery, and autoconfiguration. All fall down go boom.

 This actually wasn't "discovered" until 2.6.19 but it seems to have been introduced in a couple of late clicks of 2.6.18 and is not fixed in the kernel.org tree until 2.6.20-rc5 or so. Fedora Core seems to have retrofitted the fix into the last couple of clicks of the FC6 2.6.19 (currently 2.6.19-2911) but the stock kernel.org kernels have been broken in a couple of 2.6.18 releases and, AFAICT, all of the 2.6.19 releases.

 This is re-enforced by checking the group memberships on test010 vs test018 in both host and guest (netstat -g).

Host test010:

IPv6/IPv4 Group Memberships
Interface   RefCnt Group
--------------- ------ --------------------
lo   1   ff02::1
eth0   1   ff02::fb
eth0   1   ff02::1:ff10:f84d
eth0   1   ff02::1
veth0   1   ff02::fb

```
veth0        2      ff02::1:ff10:f84d
veth0        1      ff02::1
veth1        1      ff02::fb
veth1        1      ff02::1:ff3a:e3e2
veth1        1      ff02::1
veth1006.0   1      ff02::fb
veth1006.0   1      ff02::1:ff01:6
veth1006.0   1      ff02::1
veth1010.1   1      ff02::fb
veth1010.1   1      ff02::1:ff01:100a
veth1010.1   1      ff02::1
veth1026.0   1      ff02::fb
veth1026.0   1      ff02::1:ff01:1a
veth1026.0   1      ff02::1
```

Host test018:

```
IPv6/IPv4 Group Memberships
Interface       RefCnt Group
--------------- ------ --------------------
eth0         1      ff02::fb
eth0         1      ff02::1:ff10:f84d
veth0        1      ff02::fb
veth0        1      ff02::1:ff10:f84d
veth1        1      ff02::fb
veth1        1      ff02::1:ff3a:e3e2
veth1006.0   1      ff02::fb
veth1006.0   1      ff02::1:ff01:6
veth1010.1   1      ff02::fb
veth1010.1   1      ff02::1:ff01:100a
```

 Yeah, that would be a problem and is the symptoms reported on lkml.

 Here's one of the guest systems (1006 above):

Guest 1006 test010:

```
IPv6/IPv4 Group Memberships
Interface       RefCnt Group
--------------- ------ --------------------
lo           1      ff02::1
eth0         2      ff02::1:ff01:106
eth0         1      ff02::1
eth1         1      ff02::1:ff01:1106
eth1         1      ff02::1
```

Guest 1006 test018:

```
IPv6/IPv4 Group Memberships
Interface       RefCnt Group
-------------- ------ --------------------
eth0          1      ff02::1:ff01:106
eth1          1      ff02::1:ff01:1106
```

 Same problem. Missing the ff02::1 address on every interface. Bad
juju.

 I suspect that test15 and test18 picked up the bad IPv6 patch and broke
IPv6. We need a rebase to a working kernel or a retrofit of the patch
from 2.6.20. I have not been able to get the oz patch to patch into the
2.6.19 FC6 kernels, but that would be a good place to start since they
are no have IPv6 working. You might examine those retrofit patches.
Better yet would be to rebase the test series up to 2.6.20.

 Mike
--
Michael H. Warfield (AI4NB) | (770) 985-6132 | mhw@WittsEnd.com
   /\/\|=mhw=|\/\ | (678) 463-0932 | http://www.wittsend.com/mhw/
   NIC whois: MHW9 | An optimist believes we live in the best of all
 PGP Key: 0xDF1DD471 | possible worlds. A pessimist is sure of it!

_____
Devel mailing list
Devel@openvz.org
https://openvz.org/mailman/listinfo/devel

Subject: Re: 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by Wolfgang Schnerring on Sun, 04 Mar 2007 06:30:33 GMT
View Forum Message <> Reply to Message

* "Michael H. Warfield" <mhw@WittsEnd.com>:
> This may not be an OpenVZ problem, per se, though. Some time in the
> very later part of the 2.6.18 kernel.org release updates, a problem was
> introduced that broke IPv6. Some patch or another caused nodes to fail
> to join the "all nodes multicast" (ff02::1) address. This is critical
> to IPv6 router discovery and advertisement, neighbor discovery, and
> autoconfiguration. All fall down go boom.

I've raised that same issue here: http://bugzilla.openvz.org/show_bug.cgi?id=476

But even though Den Lunev found and backported the upstream patch (I
think from 2.6.20) that fixes the problem, I don't think this ever got into the
OpenVZ repository... any update on this?

Wolfgang

Subject: Re: Re: 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by dev on Mon, 05 Mar 2007 08:24:39 GMT
View Forum Message <> Reply to Message

Wolfgang Schnerring wrote:
> * "Michael H. Warfield" <mhw@WittsEnd.com>:
>
>> This may not be an OpenVZ problem, per se, though. Some time in the
>>very later part of the 2.6.18 kernel.org release updates, a problem was
>>introduced that broke IPv6. Some patch or another caused nodes to fail
>>to join the "all nodes multicast" (ff02::1) address. This is critical
>>to IPv6 router discovery and advertisement, neighbor discovery, and
>>autoconfiguration. All fall down go boom.
>
>
> I've raised that same issue here: http://bugzilla.openvz.org/show_bug.cgi?id=476
>
> But even though Den Lunev found and backported the upstream patch (I
> think from 2.6.20) that fixes the problem, I don't think this ever got into the
> OpenVZ repository... any update on this?
As I wrote in the bug, it is fixed in 028test019
028test018 was released earlier and was in testing for some time,
so could not accumulate this patch.

Thanks,
Kirill

Subject: [Devel] Re: 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by dev on Mon, 05 Mar 2007 08:26:47 GMT
View Forum Message <> Reply to Message

Michael,

Can you please check that the patch from bug
http://bugzilla.openvz.org/show_bug.cgi?id=476
helps you?

Thanks,
Kirill

> Hello,
>

> Introduction... I've been on the users list for a while and just
> joined the developers list. Excuse the cross post but I'm not sure
> which would be most appropriate.
>
> I work extensively with IPv6. All of my devices are configured for
> IPv6 and I actively participate in the global IPv6 network and have for
> over 5 years. I'm a member of the North American v6 Task Force,
> NAv6TF.org. My name servers, ntp servers, smtp servers, and web servers
> all show active IPv6 traffic and have for years. I'm also a big user of
> VMware and I'm also the author of an IPv6 paper that's up on their site
> somewhere.
>
> So, I'll prefix this with a bit of a rant.
>
> <Rant>
>
> I don't use venet. The way it's implemented, especially vis-a-vis
> IPv6, it makes no sense to me at all. the venet0 devices have no
> hardware / mac layer addresses so how is IPv6 neighbor discovery and
> autoconfiguration protocols suppose to work? These are really
> fundamental to the nature of IPv6 and, unless you want to use IPv6
> purely like it just IPv4 with fat addresses (which it is NOT) and live
> entirely in an IPv4 mind-think cocoon when working with IPv6, how can
> you expect this to work with fully implemented IPv6?
>
> The bridged networking of the veth devices seem to be the only viable
> way to deal with this. I HAVE tried using venet for IPv6 and failed
> miserably.
>
> This is also how I deal with IPv6 on VMware. I never use the local
> networking (host local or NAT) devices and only use the bridged devices
> there. In the case of VMware, the bridged devices are the default and
> you can add the routed or nat'ed devices. That makes a lot more sense
> to me.
>
> So IPv6 HAS to work for me. Consequently, I have to work with
> everything running over the veth devices. The venet devices just do not
> cut it at all with IPv6, so are totally useless to me for both IPv4 as
> well as IPv6 (why split the network two ways).
>
> Funny thing is, if you "turn on" IPv6 in the vz configuration files, it
> dicks things up royally by trying to force IPv6 routes out through the
> venet0 interface, which can't work, and interferes with stateless
> autoconfiguration and router discovery configuration. Leaving IPv6 off
> in the config, then it doesn't insert bogus routes and everything
> configures properly on the veth devices and IPv6 works like a charm.
> Interestingly enough, the veth devices DO get their proper link-local
> and global-unicast addresses while vnet0 sits there with not even a

> link-local address (of course - it has no mac address so how could it
> possibly have a link-local address - duh...).
>
> </Rant>
>
> Now... Onto the problem. IPv6 is working sweet with
> 2.6.18-ovz028test010 on Fedora Core 6. It would be nice to get it up to
> the FC6 2.6.19 (soon to be the 2.6.20) rebase but that's ok for now.
> When I tried 2.6.18-ovz028test015 and subsequently 2.6.18-ovz028test018,
> I found that IPv6 was totally broken. While all my interfaces properly
> autoconfigure on test10, stateless autoconf appears to be totally broken
> on test15 and test18 and even manual configuration doesn't work (hints
> strongly at broken neighbor discovery there).
>
> This may not be an OpenVZ problem, per se, though. Some time in the
> very later part of the 2.6.18 kernel.org release updates, a problem was
> introduced that broke IPv6. Some patch or another caused nodes to fail
> to join the "all nodes multicast" (ff02::1) address. This is critical
> to IPv6 router discovery and advertisement, neighbor discovery, and
> autoconfiguration. All fall down go boom.
>
> This actually wasn't "discovered" until 2.6.19 but it seems to have
> been introduced in a couple of late clicks of 2.6.18 and is not fixed in
> the kernel.org tree until 2.6.20-rc5 or so. Fedora Core seems to have
> retrofitted the fix into the last couple of clicks of the FC6 2.6.19
> (currently 2.6.19-2911) but the stock kernel.org kernels have been
> broken in a couple of 2.6.18 releases and, AFAICT, all of the 2.6.19
> releases.
>
> This is re-enforced by checking the group memberships on test010 vs
> test018 in both host and guest (netstat -g).
>
> Host test010:
>
> IPv6/IPv4 Group Memberships
> Interface   RefCnt Group
> --------------- ------ ---------------------
> lo   1   ff02::1
> eth0   1   ff02::fb
> eth0   1   ff02::1:ff10:f84d
> eth0   1   ff02::1
> veth0   1   ff02::fb
> veth0   2   ff02::1:ff10:f84d
> veth0   1   ff02::1
> veth1   1   ff02::fb
> veth1   1   ff02::1:ff3a:e3e2
> veth1   1   ff02::1
> veth1006.0   1   ff02::fb

```
> veth1006.0    1     ff02::1:ff01:6
> veth1006.0    1     ff02::1
> veth1010.1    1     ff02::fb
> veth1010.1    1     ff02::1:ff01:100a
> veth1010.1    1     ff02::1
> veth1026.0    1     ff02::fb
> veth1026.0    1     ff02::1:ff01:1a
> veth1026.0    1     ff02::1
>
> Host test018:
>
> IPv6/IPv4 Group Memberships
> Interface       RefCnt Group
> --------------- ------ --------------------
> eth0          1     ff02::fb
> eth0          1     ff02::1:ff10:f84d
> veth0         1     ff02::fb
> veth0         1     ff02::1:ff10:f84d
> veth1         1     ff02::fb
> veth1         1     ff02::1:ff3a:e3e2
> veth1006.0    1     ff02::fb
> veth1006.0    1     ff02::1:ff01:6
> veth1010.1    1     ff02::fb
> veth1010.1    1     ff02::1:ff01:100a
>
>  Yeah, that would be a problem and is the symptoms reported on lkml.
>
>  Here's one of the guest systems (1006 above):
>
> Guest 1006 test010:
>
> IPv6/IPv4 Group Memberships
> Interface       RefCnt Group
> --------------- ------ --------------------
> lo            1     ff02::1
> eth0          2     ff02::1:ff01:106
> eth0          1     ff02::1
> eth1          1     ff02::1:ff01:1106
> eth1          1     ff02::1
>
> Guest 1006 test018:
>
> IPv6/IPv4 Group Memberships
> Interface       RefCnt Group
> --------------- ------ --------------------
> eth0          1     ff02::1:ff01:106
> eth1          1     ff02::1:ff01:1106
>
```

> Same problem. Missing the ff02::1 address on every interface. Bad
> juju.
>
> I suspect that test15 and test18 picked up the bad IPv6 patch and broke
> IPv6. We need a rebase to a working kernel or a retrofit of the patch
> from 2.6.20. I have not been able to get the oz patch to patch into the
> 2.6.19 FC6 kernels, but that would be a good place to start since they
> are no have IPv6 working. You might examine those retrofit patches.
> Better yet would be to rebase the test series up to 2.6.20.
>
> Mike
>
>
> --- -----------
>

_____

Devel mailing list
Devel@openvz.org
https://openvz.org/mailman/listinfo/devel

Subject: Re: Re: 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by kir on Mon, 05 Mar 2007 08:35:58 GMT

View Forum Message <> Reply to Message

Wolfgang Schnerring wrote:
> * "Michael H. Warfield" <mhw@WittsEnd.com>:
>
>> This may not be an OpenVZ problem, per se, though. Some time in the
>> very later part of the 2.6.18 kernel.org release updates, a problem was
>> introduced that broke IPv6. Some patch or another caused nodes to fail
>> to join the "all nodes multicast" (ff02::1) address. This is critical
>> to IPv6 router discovery and advertisement, neighbor discovery, and
>> autoconfiguration. All fall down go boom.
>>
>
> I've raised that same issue here: http://bugzilla.openvz.org/show_bug.cgi?id=476
>
> But even though Den Lunev found and backported the upstream patch (I
> think from 2.6.20) that fixes the problem, I don't think this ever got into the
> OpenVZ repository... any update on this?
As Kirill said in comment #4 to the bug, this patch goes to 028test019
-- which is not yet released.
So you either have to wait till we release 019 (or later) kernel, or use
the patch from the bugreport.

Subject: Re: Re: 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by Wolfgang Schnerring on Mon, 05 Mar 2007 09:19:31 GMT
View Forum Message <> Reply to Message

* Kirill Korotaev <dev@sw.ru> [2007-03-05 11:35]:
> As I wrote in the bug, it is fixed in 028test019
> 028test018 was released earlier and was in testing for some time,
> so could not accumulate this patch.

Sorry for being dense, but the last entry at
<http://git.openvz.org/?p=linux-2.6.18-openvz;a=shortlog> is
"linux-2.6.18-028test018 released", with no mention at all of 019.
Am I looking in the wrong place?

Wolfgang

Subject: [Devel] Re: 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by dev on Mon, 05 Mar 2007 10:50:13 GMT
View Forum Message <> Reply to Message

Michael,

Some more comments below from Alexey Kuznetsov (ipv6 maintainer),
he was not subscribed, so I added his answers below and put him on CC.

> Introduction... I've been on the users list for a while and just
> joined the developers list. Excuse the cross post but I'm not sure
> which would be most appropriate.
>
> I work extensively with IPv6. All of my devices are configured for
> IPv6 and I actively participate in the global IPv6 network and have for
> over 5 years. I'm a member of the North American v6 Task Force,
> NAv6TF.org. My name servers, ntp servers, smtp servers, and web servers
> all show active IPv6 traffic and have for years. I'm also a big user of
> VMware and I'm also the author of an IPv6 paper that's up on their site
> somewhere.
>
> So, I'll prefix this with a bit of a rant.
>
> <Rant>
>
> I don't use venet. The way it's implemented, especially vis-a-vis
> IPv6, it makes no sense to me at all. the venet0 devices have no
> hardware / mac layer addresses so how is IPv6 neighbor discovery and
> autoconfiguration protocols suppose to work?
Obviously, autoconfiguration does not make sense with venet.
If you want to see autoconfiguration working you use veth, not venet.

In venet approach all the address handling is made in host system.
Virtial environment itself are not _permitted_ to intervene, hence
autoconfiguration just does not make sense. All the addresses and routing
are made by host system and setup by host system.
This is not "IPv4 mind-think cocoon", it is "scalability, efficiency
and security cocoon". If you are not paranoid about these issues,
but prefer to have real IPv6, you should forget about venet.

> These are really
> fundamental to the nature of IPv6 and, unless you want to use IPv6
> purely like it just IPv4 with fat addresses (which it is NOT) and live
> entirely in an IPv4 mind-think cocoon when working with IPv6, how can
> you expect this to work with fully implemented IPv6?
>
> The bridged networking of the veth devices seem to be the only viable
> way to deal with this. I HAVE tried using venet for IPv6 and failed
> miserably.
>
> This is also how I deal with IPv6 on VMware. I never use the local
> networking (host local or NAT) devices and only use the bridged devices
> there. In the case of VMware, the bridged devices are the default and
> you can add the routed or nat'ed devices. That makes a lot more sense
> to me.
>
> So IPv6 HAS to work for me. Consequently, I have to work with
> everything running over the veth devices. The venet devices just do not
> cut it at all with IPv6, so are totally useless to me for both IPv4 as
> well as IPv6 (why split the network two ways).
>
> Funny thing is, if you "turn on" IPv6 in the vz configuration files, it
> dicks things up royally by trying to force IPv6 routes out through the
> venet0 interface, which can't work, and interferes with stateless
> autoconfiguration and router discovery configuration. Leaving IPv6 off
> in the config, then it doesn't insert bogus routes and everything
> configures properly on the veth devices and IPv6 works like a charm.
> Interestingly enough, the veth devices DO get their proper link-local
> and global-unicast addresses while vnet0 sits there with not even a
> link-local address (of course - it has no mac address so how could it
> possibly have a link-local address - duh...).
>
> </Rant>
>
> Now... Onto the problem. IPv6 is working sweet with
> 2.6.18-ovz028test010 on Fedora Core 6. It would be nice to get it up to
> the FC6 2.6.19 (soon to be the 2.6.20) rebase but that's ok for now.
> When I tried 2.6.18-ovz028test015 and subsequently 2.6.18-ovz028test018,
> I found that IPv6 was totally broken. While all my interfaces properly

> autoconfigure on test10, stateless autoconf appears to be totally broken
> on test15 and test18 and even manual configuration doesn't work (hints
> strongly at broken neighbor discovery there).
>
> This may not be an OpenVZ problem, per se, though. Some time in the
> very later part of the 2.6.18 kernel.org release updates, a problem was
> introduced that broke IPv6. Some patch or another caused nodes to fail
> to join the "all nodes multicast" (ff02::1) address. This is critical
> to IPv6 router discovery and advertisement, neighbor discovery, and
> autoconfiguration. All fall down go boom.
Yes, you are right.
It is repaired by commit d88ae4cc97b24783ee4480697fbdcc02ab4133a6
(what is interesting is that it was unnoticed for so long in mainstream)

Alexey

_____

Subject: Re: Re: 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by dev on Mon, 05 Mar 2007 13:19:47 GMT
View Forum Message <> Reply to Message

Wolfgang,

> * Kirill Korotaev <dev@sw.ru> [2007-03-05 11:35]:
>
>>As I wrote in the bug, it is fixed in 028test019
>>028test018 was released earlier and was in testing for some time,
>>so could not accumulate this patch.
>
>
> Sorry for being dense, but the last entry at
> <http://git.openvz.org/?p=linux-2.6.18-openvz;a=shortlog> is
> "linux-2.6.18-028test018 released", with no mention at all of 019.
> Am I looking in the wrong place?
No, you are fully right.
I just didn't released 019 kernel yet, so git tree was not pushed upstream to git.openvz.org.
I pushed git, so you can see patches going to 019 kernel.

Thanks,
Kirill

Subject: [Devel] Re: 2.6.18-ovz028test015 and 2.6.18-ovz028test018 break IPv6
Posted by mhw on Mon, 05 Mar 2007 18:44:05 GMT
View Forum Message <> Reply to Message

On Mon, 2007-03-05 at 11:38 +0300, Kirill Korotaev wrote:
> Michael,

> Can you please check that the patch from bug
> http://bugzilla.openvz.org/show_bug.cgi?id=476
> helps you?

 Confirmed. That worked.

> Thanks,
> Kirill

 Regards,
 Mike
--
Michael H. Warfield (AI4NB) | (770) 985-6132 | mhw@WittsEnd.com
   /\/\|=mhw=|\/\ | (678) 463-0932 | http://www.wittsend.com/mhw/
   NIC whois: MHW9 | An optimist believes we live in the best of all
 PGP Key: 0xDF1DD471 | possible worlds. A pessimist is sure of it!

_____
Devel mailing list
Devel@openvz.org
https://openvz.org/mailman/listinfo/devel