Subject: network namespace ipv6 perfs
Posted by Daniel Lezcano on Mon, 03 Mar 2008 14:20:16 GMT
View Forum Message <> Reply to Message

Hi,

Some performance tests was made by Benjamin to watch out the impact of the network namespace. The good news is there is no impact when used with or without namespaces. That has been checked using a real network device inside a network namespace.

These results are consistent with the ones previously made for ipv4.

http://lxc.sourceforge.net/network/bench_ipv6_graph.php

Thanks to Benjamin who did all the performance tests :)

Regards
-- Daniel

Sauf indication contraire ci-dessus:
Compagnie IBM France

Courbevoie
RCS Nanterre 552 118 465
Forme Sociale : S.A.S.
Capital Social : 542.737.118 ?
SIREN/SIRET : 552 118 465 02430

_____

Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

---

Subject: Re: network namespace ipv6 perfs
Posted by Benjamin Thery on Mon, 03 Mar 2008 14:42:03 GMT
View Forum Message <> Reply to Message

Daniel Lezcano wrote:
> Hi,
>
> Some performance tests was made by Benjamin to watch out the impact of
> the network namespace. The good news is there is no impact when used

> with or without namespaces. That has been checked using a real network
> device inside a network namespace.
>
> These results are consistent with the ones previously made for ipv4.
>
> http://lxc.sourceforge.net/network/bench_ipv6_graph.php
>
> Thanks to Benjamin who did all the performance tests :)

In these results, may be, there is one thing that should be explained.
It is the CPU utilization overhead in the 'veth' case.

Compared to physical devices or macvlan, veth interfaces don't benefit
from hardware offloading mechanisms: i.e. checksums have to be computed
by the soft. That explains the big overhead in CPU utilization when
using this kind of virtual interface.

Benjamin

>
> Regards
>   -- Daniel
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>

>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
> Sauf indication contraire ci-dessus:
> Compagnie IBM France

> Courbevoie
> RCS Nanterre 552 118 465
> Forme Sociale : S.A.S.
> Capital Social : 542.737.118 ?
> SIREN/SIRET : 552 118 465 02430
>
>

--
B e n j a m i n T h e r y - BULL/DT/Open Software R&D

http://www.bull.com

_____

Subject: Re: network namespace ipv6 perfs
Posted by Benjamin Thery on Mon, 03 Mar 2008 14:48:12 GMT
View Forum Message <> Reply to Message

One more thing about these results: the kernel.
The version used to run these tests was 2.6.25-rc1 from Dave Miller's
net-2.6 tree.

(and I included results from a vanilla 2.6.23.16 as reference)

Benjamin

Daniel Lezcano wrote:
> Hi,
>
> Some performance tests was made by Benjamin to watch out the impact of
> the network namespace. The good news is there is no impact when used
> with or without namespaces. That has been checked using a real network
> device inside a network namespace.
>
> These results are consistent with the ones previously made for ipv4.
>
> http://lxc.sourceforge.net/network/bench_ipv6_graph.php
>
> Thanks to Benjamin who did all the performance tests :)
>
> Regards
> -- Daniel
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>

>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
>
> Sauf indication contraire ci-dessus:
> Compagnie IBM France

> Courbevoie
> RCS Nanterre 552 118 465
> Forme Sociale : S.A.S.
> Capital Social : 542.737.118 ?
> SIREN/SIRET : 552 118 465 02430
>
>

--
B e n j a m i n   T h e r y - BULL/DT/Open Software R&D

http://www.bull.com

_____

Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

---

Subject: Re: Re: network namespace ipv6 perfs
Posted by Pavel Emelianov on Mon, 03 Mar 2008 14:55:11 GMT
View Forum Message <> Reply to Message

Benjamin Thery wrote:
> Daniel Lezcano wrote:
>> Hi,
>>
>> Some performance tests was made by Benjamin to watch out the impact of
>> the network namespace. The good news is there is no impact when used
>> with or without namespaces. That has been checked using a real network
>> device inside a network namespace.
>>
>> These results are consistent with the ones previously made for ipv4.
>>
>> http://lxc.sourceforge.net/network/bench_ipv6_graph.php
>>
>> Thanks to Benjamin who did all the performance tests :)
>
> In these results, may be, there is one thing that should be explained.
> It is the CPU utilization overhead in the 'veth' case.
>
> Compared to physical devices or macvlan, veth interfaces don't benefit
> from hardware offloading mechanisms: i.e. checksums have to be computed
> by the soft. That explains the big overhead in CPU utilization when

You can tune the veth devices not to account checksum when unnecessary.

> using this kind of virtual interface.
>
> Benjamin
>
>> Regards
>> -- Daniel
>>
>>
>>
>>
>>
>>
>>
>>

>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>> Sauf indication contraire ci-dessus:

>> Compagnie IBM France

>> Courbevoie
>> RCS Nanterre 552 118 465
>> Forme Sociale : S.A.S.
>> Capital Social : 542.737.118 ?
>> SIREN/SIRET : 552 118 465 02430
>>
>>
>
>

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

---

Subject: Re: Re: network namespace ipv6 perfs
Posted by Benjamin Thery on Mon, 03 Mar 2008 15:04:47 GMT
View Forum Message <> Reply to Message

On Mon, Mar 3, 2008 at 3:55 PM, Pavel Emelyanov <xemul@openvz.org> wrote:
> Benjamin Thery wrote:
> > Daniel Lezcano wrote:
> >> Hi,
> >>
> >> Some performance tests was made by Benjamin to watch out the impact of
> >> the network namespace. The good news is there is no impact when used
> >> with or without namespaces. That has been checked using a real network
> >> device inside a network namespace.
> >>
> >> These results are consistent with the ones previously made for ipv4.
> >>
> >> http://lxc.sourceforge.net/network/bench_ipv6_graph.php
> >>
> >> Thanks to Benjamin who did all the performance tests :)
> >
> > In these results, may be, there is one thing that should be explained.
> > It is the CPU utilization overhead in the 'veth' case.
> >
> > Compared to physical devices or macvlan, veth interfaces don't benefit
> > from hardware offloading mechanisms: i.e. checksums have to be computed
> > by the soft. That explains the big overhead in CPU utilization when
>
> You can tune the veth devices not to account checksum when unnecessary.

Oh. This is interesting.

You mean with ethtool -K rx/tx?
I will give it a try.

Benjamin

>
>
> > using this kind of virtual interface.
> >
> > Benjamin
> >
> >> Regards
> >> -- Daniel
> >>

>--
> To unsubscribe from this list: send the line "unsubscribe netdev" in
> the body of a message to majordomo@vger.kernel.org
> More majordomo info at http://vger.kernel.org/majordomo-info.html
>

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

---

Subject: Re: Re: network namespace ipv6 perfs
Posted by Benjamin Thery on Mon, 03 Mar 2008 17:27:50 GMT
View Forum Message <> Reply to Message

Benjamin Thery wrote:
> On Mon, Mar 3, 2008 at 3:55 PM, Pavel Emelyanov <xemul@openvz.org> wrote:
>> Benjamin Thery wrote:
>> > Daniel Lezcano wrote:
>> >> Hi,
>> >>
>> >> Some performance tests was made by Benjamin to watch out the impact of
>> >> the network namespace. The good news is there is no impact when used
>> >> with or without namespaces. That has been checked using a real network
>> >> device inside a network namespace.
>> >>
>> >> These results are consistent with the ones previously made for ipv4.
>> >>
>> >> http://lxc.sourceforge.net/network/bench_ipv6_graph.php
>> >>
>> >> Thanks to Benjamin who did all the performance tests :)
>> >

>> > In these results, may be, there is one thing that should be explained.
>> > It is the CPU utilization overhead in the 'veth' case.
>> >
>> > Compared to physical devices or macvlan, veth interfaces don't benefit
>> > from hardware offloading mechanisms: i.e. checksums have to be computed
>> > by the soft. That explains the big overhead in CPU utilization when
>>
>> You can tune the veth devices not to account checksum when unnecessary.
>
> Oh. This is interesting.
>
> You mean with ethtool -K rx/tx?
> I will give it a try.

Pavel,

I had no luck with "ethtool -K veth0 rx on tx on".
On my testbed, with these options TCP drops packets
(trying to establish a ssh connection between init and child namespace).

Then, I tested "ethtool -K veth0 rx on tx off".
This time TCP (and netperf) work, but I see no difference in
CPU load compared to the case without offloading.

Can I tune veth differently?

(BTW, I run netperf between a child namespace on host A and netserv
on host B. The stream goes through the following interface:
veth1 on A -> veth0 on A -> eth1 on A -> ("real network") -> eth1 on B)

Benjamin

>
>>
>> > using this kind of virtual interface.
>> >
>> > Benjamin
>> >
>> >> Regards
>> >>   -- Daniel
>> >>
>
>> --
>> To unsubscribe from this list: send the line "unsubscribe netdev" in
>> the body of a message to majordomo@vger.kernel.org
>> More majordomo info at http://vger.kernel.org/majordomo-info.html
>>

>
>

--
B e n j a m i n T h e r y - BULL/DT/Open Software R&D

http://www.bull.com

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

---

Subject: Re: network namespace ipv6 perfs
Posted by Rick Jones on Mon, 03 Mar 2008 19:38:49 GMT
View Forum Message <> Reply to Message

Daniel Lezcano wrote:
> Hi,
>
> Some performance tests was made by Benjamin to watch out the impact of
> the network namespace. The good news is there is no impact when used
> with or without namespaces. That has been checked using a real network
> device inside a network namespace.

The *_RR tests seem to show a drop in througput and corresponding
increases in service demand - could that be because things like TSO et
al cannot mask much of anything in the way of a path-length increase?

From the annotations, I'm ass-u-me-ing that NS was only used on the
netperf side and not both netperf and netserver side?

happy benchmarking,

rick jones

> These results are consistent with the ones previously made for ipv4.
>
> http://lxc.sourceforge.net/network/bench_ipv6_graph.php
>
> Thanks to Benjamin who did all the performance tests :)
>

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

---

Subject: Re: network namespace ipv6 perfs
Posted by Daniel Lezcano on Mon, 03 Mar 2008 20:01:18 GMT
View Forum Message <> Reply to Message

Rick Jones wrote:
> Daniel Lezcano wrote:
>> Hi,
>>
>> Some performance tests was made by Benjamin to watch out the impact of
>> the network namespace. The good news is there is no impact when used
>> with or without namespaces. That has been checked using a real network
>> device inside a network namespace.
>
> The *_RR tests seem to show a drop in througput and corresponding
> increases in service demand - could that be because things like TSO et
> al cannot mask much of anything in the way of a path-length increase?

Hmm. In fact Benjamin took the 2.6.23.16 kernel where there were no
network namespace code at all. So these differences between 2.6.23.16
and 2.6.25-rc1 does not show a performance degradation especially
related to the network namespaces. The important point is the 2.6.25-rc1
without ipv6 netns and 2.6.25-rc1 with ipv6 netns code applied, I mean
the second and the third line and we can point that the ipv6 netns code
does not degrade performances for either throughput and service demand.

> From the annotations, I'm ass-u-me-ing that NS was only used on the
> netperf side and not both netperf and netserver side?

right :)

> happy benchmarking,

Thanks Rick.

  -- Daniel

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

Subject: Re: network namespace ipv6 perfs
Posted by Benjamin Thery on Tue, 04 Mar 2008 15:59:56 GMT
View Forum Message <> Reply to Message

Daniel Lezcano wrote:
> Rick Jones wrote:
>> Daniel Lezcano wrote:

>>> Hi,
>>>
>>> Some performance tests was made by Benjamin to watch out the impact
>>> of the network namespace. The good news is there is no impact when
>>> used with or without namespaces. That has been checked using a real
>>> network device inside a network namespace.
>>
>> The *_RR tests seem to show a drop in througput and corresponding
>> increases in service demand - could that be because things like TSO et
>> al cannot mask much of anything in the way of a path-length increase?
>
> Hmm. In fact Benjamin took the 2.6.23.16 kernel where there were no
> network namespace code at all. So these differences between 2.6.23.16
> and 2.6.25-rc1 does not show a performance degradation especially
> related to the network namespaces. The important point is the 2.6.25-rc1
> without ipv6 netns and 2.6.25-rc1 with ipv6 netns code applied, I mean
> the second and the third line and we can point that the ipv6 netns code
> does not degrade performances for either throughput and service demand.

As Daniel stated, we should not compare the first bar with the other
ones directly. May be I should have arranged the chart differently and
made it more clear that the first bar is "2.6.23 vanilla" and the second
one is "2.6.25-rc1 vanilla". Many changes happened in the whole kernel
between 2.6.23 and 2.6.24 so we can't compare the first two bars to tell
if network namespace degraded performances (and only a small part of
netns is in 2.6.24).
The way I presented the chart is a bit misleading. :)

What's interesting to compare in the charts is the 2nd, 3rd and 4th
lines. It shows that on the exact same hardware (in the 4th case the
physical interface is moved into the child namespace), with or without
the patchset, using network namespace or not, performance is about the
same.

Benjamin

>> From the annotations, I'm ass-u-me-ing that NS was only used on the
>> netperf side and not both netperf and netserver side?
>
> right :)
>
>> happy benchmarking,
>
> Thanks Rick.
>
>   -- Daniel
>
>

--
B e n j a m i n T h e r y - BULL/DT/Open Software R&D

http://www.bull.com

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers

---

Subject: Re: Re: network namespace ipv6 perfs
Posted by Pavel Emelianov on Wed, 05 Mar 2008 12:39:02 GMT
View Forum Message <> Reply to Message

Benjamin Thery wrote:
> Benjamin Thery wrote:
>> On Mon, Mar 3, 2008 at 3:55 PM, Pavel Emelyanov <xemul@openvz.org> wrote:
>>> Benjamin Thery wrote:
>>> > Daniel Lezcano wrote:
>>> >> Hi,
>>> >>
>>> >> Some performance tests was made by Benjamin to watch out the impact of
>>> >> the network namespace. The good news is there is no impact when used
>>> >> with or without namespaces. That has been checked using a real network
>>> >> device inside a network namespace.
>>> >>
>>> >> These results are consistent with the ones previously made for ipv4.
>>> >>
>>> >> http://lxc.sourceforge.net/network/bench_ipv6_graph.php
>>> >>
>>> >> Thanks to Benjamin who did all the performance tests :)
>>> >
>>> > In these results, may be, there is one thing that should be explained.
>>> > It is the CPU utilization overhead in the 'veth' case.
>>> >
>>> > Compared to physical devices or macvlan, veth interfaces don't benefit
>>> > from hardware offloading mechanisms: i.e. checksums have to be computed
>>> > by the soft. That explains the big overhead in CPU utilization when
>>>
>>> You can tune the veth devices not to account checksum when unnecessary.
>> Oh. This is interesting.
>>
>> You mean with ethtool -K rx/tx?
>> I will give it a try.
>
> Pavel,

>
> I had no luck with "ethtool -K veth0 rx on tx on".
> On my testbed, with these options TCP drops packets
> (trying to establish a ssh connection between init and child namespace).
>
>
> Then, I tested "ethtool -K veth0 rx on tx off".
> This time TCP (and netperf) work, but I see no difference in
> CPU load compared to the case without offloading.
>
> Can I tune veth differently?

Yup. You may try turn tso and sg on as well.

> (BTW, I run netperf between a child namespace on host A and netserv
> on host B. The stream goes through the following interface:
> veth1 on A -> veth0 on A -> eth1 on A -> ("real network") -> eth1 on B)
>
> Benjamin
>
>>> > using this kind of virtual interface.
>>> >
>>> > Benjamin
>>> >
>>> >> Regards
>>> >>   -- Daniel
>>> >>
>>> --
>>> To unsubscribe from this list: send the line "unsubscribe netdev" in
>>> the body of a message to majordomo@vger.kernel.org
>>> More majordomo info at http://vger.kernel.org/majordomo-info.html
>>>
>>
>
>