Subject: Re: Re: [PATCHSET] 2.6.20-lxc8
Posted by ebiederm on Thu, 22 Mar 2007 20:02:38 GMT
View Forum Message <> Reply to Message

Benjamin Thery <benjamin.thery@bull.net> writes:

> My investigations on the increase of cpu load when running netperf inside a
> container (ie. through etun2<->etun1) is progressing slowly.
>
> I'm not sure the cause is fragmentation as we supposed initially.
> In fact, it seems related to forwarding the packets between the devices.
>
> Here is what I've tracked so far:
> * when we run netperf from the container, oprofile reports that the top
> "consuming" symbol is: "pskb_expand_head". Next comes
> "csum_partial_copy_generic". these symbols represents respectively 13.5% and
> 9.7% of the samples.
> * Without container, these symbols don't show up in the first 20 entries.
>
> Who is calling "pskb_expand_head" in this case?
>
> Using systemtap, I determined that the call to "pskb_expand_head" comes from the
> skb_cow() in ip_forward() (l.90 in 2.6.20-rc5-netns).
>
> The number of calls to "pskb_expand_head" matches the number of invocations of
> ip_forward() (268000 calls for a 20 seconds netperf session in my case).

Ok. This seems to make sense, and is related to how we have configured the
network in this case.

It looks like pskb_expand_head is called by skb_cow.

skb_cow has two cases when it calls pskb_expand_head.
- When there are multiple people who have a copy of the packet
  (tcpdump and friends)
- When there isn't enough room for the hard header.

Any chance one of you guys looking into this can instrument up
ip_foward just before the call to skb_cow and find out which
reason it is?

A cheap trick to make the overhead go away is probably to setup
ethernet bridging in this case...

But if we can ensure the ip_foward case does not need to do anything
more than modify the ttl and update the destination that would
be good to.

Anyway this does look very solvable.

Eric

_____

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