Kirill Korotaev wrote:
> Eric W. Biederman wrote:
>> Daniel Lezcano <dlezcano@fr.ibm.com> writes:
>>
>>
>>> Benjamin Thery and I we were looking at this.
>>> For the moment we are investigating if there is IP fragmentation between the
>>> eth0 and the pair devices.
>>> The profiling shows us "pskb_expand_head" and "csum_partial_copy_generic"
>>> functions have the bigger CPU usage when we are inside a container.
>>
>> Hmm. Interesting hypothesis. I do know I haven't finished the IP fragment
>> support so there may be an issue there. Although generally especially with
>> TCP fragments are not generated. How are you thinking fragmentation would
>> be involved?
>
> I'm not sure I fully understand the test and details of the thread,
> but still...
>
> if network device inside container has MTU higher then eth0 outside the container,
> then packets will get fragmented.
> First time to MTU1 inside container and refragmented to MTU2
> outside the container. At least this is the reason we use ethernet MTU
> on venet devices in OpenVZ - to avoid framentation.

frankly speaking, there are another points, which can call to
pskb_expand_head. Hard header size on container network device should be
the same as on real one.

Regards,
 Den