Subject: Re: [PATCH 0/5] Make nicer CONFIG_NET_NS=n case code
Posted by Eric Dumazet on Thu, 01 Nov 2007 06:58:59 GMT
View Forum Message <> Reply to Message

> Eric Dumazet <dada1@cosmosbay.com> writes:
>

>>> Eric Dumazet <dada1@cosmosbay.com> writes:
>>>
>>>
>>>> Definitly wanted here. Thank you.
>>>> One more refcounting on each socket creation/deletion was expensive.
>>> Really? Have you actually measured that? If the overhead is
>>> measurable and expensive we may want to look at per cpu counters or
>>> something like that. So far I don't have any numbers that say any
>>> of the network namespace work inherently has any overhead.
>> It seems that on some old opterons (two 246 for example),
>> "if (atomic_dec_and_test(&net->count))" is rather expensive yes :(
>
> I won't argue that atomic_dec_and_test is costly. My gut feel is that
> socket creation/destruction is sufficiently rare that such a test
> would be lost in the noise. Doing anything more sophisticated is
> likely to be less readable, and unless we can measure some overhead
> my preference right now is to keep the code stupid and simple. Which
> usually has a good icache footprint.

I agree with you that with current state, this atomic_inc/atomic_dec_and_test wont come in profiles unless a trivial bench is writen

for(;;){close(socket(....));}

If David or another dev can eliminate the atomic inc/dec on device refcount cost for each packet traveling, the socket creation/destruction would certainly raise.

Other contention points is the mnt_count (yet another refcount) in "struct vfsmount", a truly useless refcount as I never had (and nobody had) to un-mount sock_mnt :)