
Subject: Re: [PATCH 0/5] Make nicer CONFIG_NET_NS=n case code
Posted by [ebiederm](#) on Thu, 01 Nov 2007 00:51:54 GMT
[View Forum Message](#) <> [Reply to Message](#)

Eric Dumazet <dada1@cosmosbay.com> writes:

```
>> Eric Dumazet <dada1@cosmosbay.com> writes:
>>
>>
>>> Definitely 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 opteron (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 am not sure per cpu counters help : I tried this and got no speedup. (This was
> on net_device refcnt at that time)
>
> (on this machines, the access through fs/gs selector seems expensive too)
>
> Maybe a lazy mode could be done, ie only do a atomic_dec(), as done in dev_put()
> ?
>
> Also, "count" sits in a cache line that contains mostly read and shared fields,
> you might want to put it in a separate cache line in SMP, to avoid cache line
> ping-pongs.
```

As for cache lines I could reverse the order 'list' and 'work' which should split the read-only and the writable fields in practice for that part of the structure.

Eric
