Subject: Re: [PATCH 0/4] MIB: add struct net to UDP accounting macros
Posted by Daniel Lezcano on Mon, 07 Jul 2008 10:19:03 GMT
View Forum Message <> Reply to Message

Pavel Emelyanov wrote:
> David Miller wrote:
>> From: Pavel Emelyanov <xemul@openvz.org>
>> Date: Fri, 04 Jul 2008 15:58:44 +0400
>>
>>> This is the first small set of MIB statistics netnsization. The easiest
>>> case is UDP stats, so I started with this one. If this set is accepted,
>>> I will go on step-by-step with adding struct net to all the other stats'
>>> accounting macros, then SNMP_XXX ones and finish with a set than will put
>>> the stats on the struct net and fix appropriate proc files.
>>>
>>> Signed-off-by: Pavel Emelyanov <xemul@openvz.org>
>>> Acked-by: Denis V. Lunev <den@openvz.org>
>> Applied, thanks Pavel.
>>
>> Are we going to provide some way for an administrator to fetch
>> stats from the perspective of all namespaces? I know the idea
>> is seperation with this stuff, but admins are going to want
>> something like that.
>
> Well, if we want to get the stats for each namespace separately, then
> this ability is already present. Since this statistics is shown via the
> /proc/net files and the /proc/net itself is seen via the /proc/<pid>/net,
> then we can walk the init-s of all the containers in the system and dump
> this info.
>
> The problem that is to be solved with this approach is how to get these
> init-s :) But since finding any namespace by some task living in it is a
> common practice now (netdev moving, sys_hijack) this one will be solved.
>
> BTW, are there some plans about implementing some netlink-based fetcher
> of this statistics? If so, then I think it's worth making this engine
> namespaces aware from the very beginning.

Shouldn't be interesting to handle the network namespaces directly with
the iproute command ?

 * ip netns add <name>
 * ip netns del <name>
 * ip netns <oldname> set name <newname>
 * ip netns show

When a network namespace is created via clone|unshare, the name is

automatically the pid of the creator.

In order to track the namespace, we can add an entry in /proc/net (aka /proc/<pid>/net) named "name" which contains the name of the namespace. I heard Eric is thinking about a netnsfs.

From a larger perspective, the iproute command with a new "netns" subcommand can be enhanced to have more features, for example the freeze and resume for the network.

* ip netns set down <name>
* ip netns set up <name>

So having the netns binded, we can plug the known subcommand (link, ip, ...) with the netns features. For examples:

* ip addr add 1.2.3.4/24 dev eth0 netns foo

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