Subject: TCP hangs up
Posted by SÃ©kiltoyai   on Sat, 02 Jan 2010 23:08:43 GMT
View Forum Message <> Reply to Message

Hello,

Happy new year !

I have some problem with tcp on my VEs. The symptom is that some tcp (not tested with udp) connexions hang up with big files. It is very clear with ftp. And doesn't happen with http, because, i guess, apache can't reach same connexion rates (the problem only occur when i send data from the server).

I first assumed it was a netfilter problem but it occur even when policies are set to access. And it disappear when i set tcpsndbuf to really huge values (obviously it cannot be considered as a good solution).

I have seen some problems with tcpsndbuf on the bugzilla but it seems to have been solved in past kernel versions (mine is a Debian 2.6.26-2-openvz-amd64). Is it really the case ?
Shall i post a bug report or error is of my own ?

I have put in attachment the /proc/user_beancounters of the testing vps and the /etc/vz/conf/240.conf.
With this conf, the ftp transfer stop at approximately 6 or 7 megabytes, for a file of 10 megabytes.

Thanks for your help.

Best regards.

File Attachments
1) user_beancounters.txt, downloaded 450 times
2) veconf.txt, downloaded 408 times

Subject: Re: TCP hangs up
Posted by nuno on Fri, 08 Jan 2010 12:04:50 GMT
View Forum Message <> Reply to Message

The tcp buffers are in bytes. So you have 8MB for them. It's not much for fast machines on fast networks. Maybe 32MB will do the trick. Don't forget to increase kmemsize (tcpsndbuf memory is accounted in kmemsize too).

If you really need the RAM maybe if you disable tcp window scaling and other tcp performance parameters, you can get away with less than 32MB... Test it and report back.

Regards,

Subject: Re: TCP hangs up
Posted by SÃ©kiltoyai on Sat, 09 Jan 2010 01:27:48 GMT

Hello,

Thanks for your answer.

Seems to be ok with your recommend. I have put tcpsndbuf to 48MB:64MB and i can send a 8GB file on my fastest connexion.

Here is the result in ubc :
240:
kmemsize   2358102 3607954 20971520 31457280 0
...
tcpsndbuf 106952   9038280 50331648 67108864 0
...
So i think a barrier around 10MB will be suitable.

But i have two questions :
- Is it a normal behaviour that programs (or kernel network driver, i don't know where is the problem) hangs up in case of this tcp buffer shortage ? Isn't there a bug in the kernel, which should simply reject new send() system calls when its buffers are full ?
- Why a call to vzsplit (with a division in 6 vps) gives me a tcpsndbuf of approximately 5.8MB:11MB, so a barrier of 5.8MB, which is far below the needed barrier ?

Subject: Re: TCP hangs up
Posted by nuno on Sat, 09 Jan 2010 09:10:11 GMT

Sékiltoyai wrote on Sat, 09 January 2010 01:27Hello,

Thanks for your answer.

Seems to be ok with your recommend. I have put tcpsndbuf to 48MB:64MB and i can send a 8GB file on my fastest connexion.

Here is the result in ubc :

240:
kmemsize 2358102 3607954 20971520 31457280 0
...
tcpsndbuf 106952   9038280 50331648 67108864 0
...
So i think a barrier around 10MB will be suitable.

Don't forget that you can have 100 of those connections. Testing is your best bet.

Sékiltoyai wrote on Sat, 09 January 2010 01:27
But i have two questions :
- Is it a normal behaviour that programs (or kernel network driver, i don't know where is the problem) hangs up in case of this tcp buffer shortage ? Isn't there a bug in the kernel, which should simply reject new send() system calls when its buffers are full ?
- Why a call to vzsplit (with a division in 6 vps) gives me a tcpsndbuf of approximately 5.8MB:11MB, so a barrier of 5.8MB, which is far below the needed barrier ?

After faling to allocate memory the TCP connection may be FUBAR (damaged beyond repair).

I never rely on vzsplit. Maybe those numbers will be fine for a 1mbit link? Or for a VPS provider that wants to put 200 containers in a small server? It's a good question for the maintainers.

Regards,

Subject: Re: TCP hangs up
Posted by SÃ©kiltoyai   on Sun, 10 Jan 2010 17:43:09 GMT
View Forum Message <> Reply to Message

Hello,

Thanks again for your advices.

nuno wrote on Sat, 09 January 2010 10:10
Don't forget that you can have 100 of those connections. Testing is your best bet.

I've made some tests and, as you have wroten, if i make several ftp transfers in the same time, the tcpsndbuf explode.
For only four parallel transfers, i have to save at least 20M of kernel memory for tcp connexions,

it's too big for an auxiliary vps (this one is going to be dedicated to the backup, and i have only 1G RAM).

It doesn't seem to be possible to control independantly tcp parameters (via sysctl) on each vps. And I don't want to disable tcp window scale for the whole server. So, i have made another test. I have limited trafic with some simple netfilter rules :

-A FORWARD -s 192.168.2.40/32 -i venet0 -o eth0 -j LIMITER
-A LIMITER -m limit --limit 2/sec -j ACCEPT
-A LIMITER -m statistic --mode nth --every 10000 -j LOG --log-prefix
"[FOR_INTERNET]limited_packet" --log-tcp-options --log-ip-options
-A LIMITER -j DROP

And for 2 connexions, the tcp buffer falls down to 244K.
I think i'll set up a consistent trafic limiting with tc, but is it a good way of solving my problem ?

nuno wrote on Sat, 09 January 2010 10:10
After faling to allocate memory the TCP connection may be FUBAR (damaged beyond repair).

And how can we check this ? And how can we "repair" a connexion in FUBAR state ? Is it automatic ?

Thank,

Regards.

Subject: Re: TCP hangs up
Posted by nuno on Sun, 10 Jan 2010 18:43:06 GMT
View Forum Message <> Reply to Message

If it's your server and your container just increase the parameters and be happy.

If the tcp connection is FUBAR, the transfer will fail. That's not good, but it isn't too bad: you just restart the transfer.

Regards,

Subject: Re: TCP hangs up
Posted by SÃ©kiltoyai on Sun, 10 Jan 2010 19:00:59 GMT
View Forum Message <> Reply to Message

---

nuno wrote on Sun, 10 January 2010 19:43If it's your server and your container just increase the parameters and be happy.

mmh, no way !

It's my server (i administrate the host system) but i must deal with each vps on this server. So I can't assign to it 1/10 of the whole RAM just for the kernel memory !