Subject: TCP packets lost?
Posted by Max Deineko on Mon, 07 Jan 2008 07:55:43 GMT
View Forum Message <> Reply to Message

Hi,

I'm not really sure where to ask for help regarding my problem; I also have posted to the apache group some time ago, but have not found a solution. If you think this is the wrong group, I apologize. Also, any hint as to where ask for help would be highly appreciated.

I'm facing following problem: Apache running inside an OpenVZ VPS, serving multiple VirtualHosts. Now, it has regular availabiliy problems, which appears to be neither load nor request pattern related.

What I don't quite understand is this: during the periods when the server isn't available (and seems to do nothing) most of TCP sessions look like the one below - apache keeps resending the handshake SYNACK packets to the client although client's response is clearly here (tcpdump done from inside the VPS).

Now I would think that even when the Apache proccess would hang, or be waiting for some other component, the TCP stack part of the application (or the kernel, for that matter) would still behave normally - is this wrong? I.e. is such behaviour normal? Or do the incoming packets get lost somewhere?

My first guess would have been that when the server is busy serving an expensive request, the tcp rcv buffer would get full and packets would get dropped. But wouldn't tcpdump not be able to see them then?

Any help would be highly appreciated,

thanks -

Max Deineko.

Session dump:

10:14:33.666935 IP CLIENT > SERVER: S 742086248:742086248(0) win 5808 <mss 1452,sackOK,timestamp 196805971 0,nop,wscale 7>
10:14:33.666965 IP SERVER > CLIENT: S 2619430727:2619430727(0) ack 742086249 win 5792 <mss 1460,sackOK,timestamp 347617915 196805971,nop,wscale 2>
10:14:33.700161 IP CLIENT > SERVER: . ack 1 win 46 <nop,nop,timestamp 196805980 347617915>
10:14:33.708406 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp

196805980 347617915>
10:14:33.709279 IP CLIENT > SERVER: P 899:963(64) ack 1 win 46 <nop,nop,timestamp 196805980 347617915>
10:14:33.945483 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp 196806039 347617915>
10:14:34.417635 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp 196806157 347617915>
10:14:35.360801 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp 196806393 347617915>
10:14:37.249017 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp 196806865 347617915>
10:14:38.253998 IP SERVER > CLIENT: S 2619430727:2619430727(0) ack 742086249 win 5792 <mss 1460,sackOK,timestamp 347622503 196806865,nop,wscale 2>
10:14:38.287866 IP CLIENT > SERVER: . ack 1 win 46 <nop,nop,timestamp 196807127 347622503,nop,nop,sack 1 {0:1}>
10:14:41.023578 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp 196807809 347622503>
10:14:44.253514 IP SERVER > CLIENT: S 2619430727:2619430727(0) ack 742086249 win 5792 <mss 1460,sackOK,timestamp 347628503 196807809,nop,wscale 2>
10:14:44.287392 IP CLIENT > SERVER: . ack 1 win 46 <nop,nop,timestamp 196808627 347628503,nop,nop,sack 1 {0:1}>
10:14:48.581817 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp 196809697 347628503>
10:14:56.258819 IP SERVER > CLIENT: S 2619430727:2619430727(0) ack 742086249 win 5792 <mss 1460,sackOK,timestamp 347640510 196809697,nop,wscale 2>
10:14:56.292441 IP CLIENT > SERVER: . ack 1 win 46 <nop,nop,timestamp 196811629 347640510,nop,nop,sack 1 {0:1}>
10:15:03.676701 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp 196813473 347640510>
10:15:20.458057 IP SERVER > CLIENT: S 2619430727:2619430727(0) ack 742086249 win 5792 <mss 1460,sackOK,timestamp 347664713 196813473,nop,wscale 2>
10:15:20.491784 IP CLIENT > SERVER: . ack 1 win 46 <nop,nop,timestamp 196817680 347664713,nop,nop,sack 1 {0:1}>
10:15:33.876920 IP CLIENT > SERVER: P 1:899(898) ack 1 win 46 <nop,nop,timestamp 196821025 347664713>
10:15:33.876953 IP SERVER > CLIENT: . ack 899 win 1897 <nop,nop,timestamp 347678134 196821025>
10:15:33.911143 IP CLIENT > SERVER: P 899:963(64) ack 1 win 46 <nop,nop,timestamp 196821035 347678134>
...(connection recovers)

uname -a

Linux 2.6.9-023stab044.4-enterprise #1 SMP Thu May 24 17:41:23 MSD 2007 i686 i686 i386 GNU/Linux

httpd -V

Server version: Apache/2.0.52
Server built: Jul 14 2007 11:53:18
Server's Module Magic Number: 20020903:9
Architecture: 32-bit
Server compiled with....
-D APACHE_MPM_DIR="server/mpm/prefork"
-D APR_HAS_SENDFILE
-D APR_HAS_MMAP
-D APR_HAVE_IPV6 (IPv4-mapped addresses enabled)
-D APR_USE_SYSVSEM_SERIALIZE
-D APR_USE_PTHREAD_SERIALIZE
-D SINGLE_LISTEN_UNSERIALIZED_ACCEPT
-D APR_HAS_OTHER_CHILD
-D AP_HAVE_RELIABLE_PIPED_LOGS
-D HTTPD_ROOT="/etc/httpd"
-D SUEXEC_BIN="/usr/sbin/suexec"
-D DEFAULT_PIDLOG="logs/httpd.pid"
-D DEFAULT_SCOREBOARD="logs/apache_runtime_status"
-D DEFAULT_LOCKFILE="logs/accept.lock"
-D DEFAULT_ERRORLOG="logs/error_log"
-D AP_TYPES_CONFIG_FILE="conf/mime.types"
-D SERVER_CONFIG_FILE="conf/httpd.conf"

Some sysctl parameters:

net.ipv4.tcp_moderate_rcvbuf = 1
net.ipv4.tcp_rmem = 4096 87380 174760
net.ipv4.tcp_wmem = 4096 16384 131072
net.ipv4.tcp_dsack = 1
net.ipv4.tcp_fack = 1
net.ipv4.tcp_max_syn_backlog = 1024
net.ipv4.tcp_keepalive_intvl = 75
net.ipv4.tcp_keepalive_probes = 9
net.ipv4.tcp_keepalive_time = 7200
net.ipv4.tcp_window_scaling = 1

net.core.rmem_max = 131071
net.core.wmem_max = 131071
net.core.somaxconn = 128
net.core.netdev_max_backlog = 300
net.ipv4.tcp_tw_reuse = 0
net.ipv4.tcp_tw_recycle = 0
net.ipv4.tcp_max_tw_buckets = 180000
net.ipv4.tcp_max_tw_buckets_ve = 16536

VPS limits:

uid resource held maxheld barrier limit failcnt

tcpsndbuf	3024352	3189816	2147483647	2147483647	0
tcprcvbuf	4662588	4728124	2147483647	2147483647	0