FreeNAS - Bug #17272

SMB performance dropped from 90% to 50% of network speed after update

09/03/2016 01:22 AM - Alfred Schlütter

Status:	Closed: Third party to resolve		
Priority:	Nice to have		
Assignee:	Kris Moore		
Category:	OS		
Target version:	N/A		
Seen in:	9.10.1	**Needs Merging:**	Yes
Severity:	New	**Needs Automation:**	No
Reason for Closing:		**Support Suite Ticket:**	n/a
Reason for Blocked:		**Hardware Configuration:**	
Needs QA:	Yes	**ChangeLog Required:**	No
Needs Doc:	Yes		

Description

Hi Guys,

after update freenas from FreeNAS-9.10-STABLE-201606270534 (dd17351) to FreeNAS-9.10.1 (d989edd),
i have only 50% network speed over smb. Before 890 Mbit/s, after update 420 Mbit/s. I run four
different servers in production. The three servers with update have 50% network speed.
The fourth server without update is ok.

I cant't see any error.

Best regards
Alfred

History

#1 - 09/03/2016 05:35 AM - Josh Paetzel

- Subject changed from 50% network speed after update to SMB performance dropped from 90% to 50% of network speed after update

- Assignee set to Josh Paetzel

- Priority changed from No priority to Important

I'm going to be really pedantic here, but it's important. Based on the information you've provided the only thing we know is SMB is slower. Whether it's a networking issue remains to be seen.

First question to short circuit all of this:

9.10.1 has a samba fix for the badlock vulnerability. OSX interacted poorly with this. If your clients are Macs and they are NOT running 10.11.6 you either need to upgrade the macs or downgrade FreeNAS.

Otherwise, first order of business is to use iperf (included in FreeNAS) to determine if the network really has slowed down. Run iperf -s on the 9.3 and 9.10 FreeNAS systems. Install iperf2 on your clients and run iperf -c ip.of.freenas9.3 and then 9.10

If it has slowed down we'll troubleshoot that, otherwise we'll move on to eliminating ZFS being the bottleneck.

#2 - 09/03/2016 08:00 AM - Alfred Schlütter

Networking is not the problem. Network speed is OK. See the results below:

On FreeNAS-9.10.1 (d989edd)

1. iperf -s

   ```
   ------------------------------------------------------------
   Server listening on TCP port 5001
   TCP window size: 64.0 KByte (default)
   ------------------------------------------------------------
   [  4] local 192.168.100.12 port 5001 connected with 192.168.100.32 port 51951
   [ ID] Interval       Transfer     Bandwidth
   [  4]  0.0-10.0 sec  1.04 GBytes   892 Mbits/sec
   ```

On FreeNAS-9.10-STABLE-201606270534 (dd17351)

1. iperf -s

   ```
   ------------------------------------------------------------
   Server listening on TCP port 5001
   TCP window size: 64.0 KByte (default)
   ------------------------------------------------------------
   [  4] local 192.168.100.14 port 5001 connected with 192.168.100.32 port 51964
   [ ID] Interval       Transfer     Bandwidth
   [  4]  0.0-10.0 sec  1.05 GBytes   902 Mbits/sec
   ```

Is samba the problem? Any information?

#3 - 09/03/2016 08:13 AM - Alfred Schlütter

Sorry forgotten Information. My Clients are all Windows 7 64 bit with 1 GBit networkcard.

#4 - 09/03/2016 10:20 AM - Josh Paetzel

Ok. So the next thing to eliminate is the ZFS performance.

Locally on FreeNAS can you run:

iozone -r 128 -s 100G -t 1 -+n -i 0 -i 1 -+C 1 -+w 1 -+y 1

And paste me the results.

This will do a sequential write and sequential read test of a single 100GB file with 30% compression. Note if your RAM + L2ARC is larger than 100GB the results will be skewed. Adjust the -s 100G parameter in that case.

#5 - 09/03/2016 12:29 PM - Alfred Schlütter

1. iozone -r 128 -s 100G -t 1 -+n -i 0 -i 1 -+C 1 -+w 1 -+y 1
   Iozone: Performance Test of File I/O
   Version $Revision: 3.420 $
   Compiled for 64 bit mode.
   Build: freebsd

   Contributors:William Norcott, Don Capps, Isom Crawford, Kirby Collins
   Al Slater, Scott Rhine, Mike Wisner, Ken Goss
   Steve Landherr, Brad Smith, Mark Kelly, Dr. Alain CYR,
   Randy Dunlap, Mark Montague, Dan Million, Gavin Brebner,
   Jean-Marc Zucconi, Jeff Blomberg, Benny Halevy, Dave Boone,
   Erik Habbinga, Kris Strecker, Walter Wong, Joshua Root,
   Fabrice Bacchella, Zhenghua Xue, Qin Li, Darren Sawyer,
   Vangel Bojaxhi, Ben England, Vikentsi Lapa.

   Run began: Sat Sep 3 21:23:34 2016

   Record Size 128 KB
   File size set to 104857600 KB
   No retest option selected
   Dedupe within 1 percent.
   Dedup activated 1 percent.
   Dedupe within & across 1 percent.
   Command line used: iozone -r 128 -s 100G -t 1 -+n -i 0 -i 1 -+C 1 -+w 1 -+y 1
   Output is in Kbytes/sec
   Time Resolution = 0.000001 seconds.
   Processor cache size set to 1024 Kbytes.
   Processor cache line size set to 32 bytes.
   File stride size set to 17 * record size.
   Throughput test with 1 process
   Each process writes a 104857600 Kbyte file in 128 Kbyte records

   Children see throughput for 1 initial writers = 354183.12 KB/sec
   Parent sees throughput for 1 initial writers = 344749.42 KB/sec
   Min throughput per process = 354183.12 KB/sec
   Max throughput per process = 354183.12 KB/sec
   Avg throughput per process = 354183.12 KB/sec
   Min xfer = 104857600.00 KB

#6 - 09/03/2016 12:34 PM - Alfred Schlütter

Children see throughput for 1 readers = 364193.00 KB/sec
Parent sees throughput for 1 readers = 364191.72 KB/sec
Min throughput per process = 364193.00 KB/sec
Max throughput per process = 364193.00 KB/sec
Avg throughput per process = 364193.00 KB/sec
Min xfer = 104857600.00 KB

iozone test complete.

#7 - 09/03/2016 12:44 PM - Josh Paetzel

- Category changed from 20 to 57

- Status changed from Unscreened to Investigation

Ok, so that eliminates ZFS.

So we can now say this is a Samba issue. I'm driving now but I'll help you chase this down in the next couple days.

#8 - 09/04/2016 09:05 AM - Alfred Schlütter

Very Thanks!

#9 - 09/06/2016 05:55 AM - Kris Moore

- Priority changed from Important to Expected

- Target version set to 9.10.1-U1

#10 - 09/08/2016 10:10 AM - Kris Moore

- Due date set to 09/19/2016

#11 - 09/12/2016 08:17 AM - Josh Paetzel

Alfred,

What times are you available to try and get to the bottom of this? I'm US/Central timezone but have some flexibility in my schedule.

#12 - 09/13/2016 08:48 AM - Alfred Schlütter

Josh, sorry but i'm the next 5 weeks not in the Office. I write, when i'm back. I hope is OK.

#13 - 09/13/2016 08:50 AM - Josh Paetzel

- Priority changed from Expected to Nice to have

- Target version changed from 9.10.1-U1 to 49

Yes, that will be fine. Please respond to this ticket when you are ready to look at this.

#14 - 09/18/2016 11:46 AM - Alfred Schlütter

Now we are little bit closer. When samba is configured as AD Domain-Controller, smbd eats 100% cpu!
Also on other systems.

Copy large files on shares. See results from my assistent below.

AD Domain-Controller fn1 - FreeNAS-9.10.1 (d989edd) - Samba version 4.3.11-GIT-UNKNOWN: 420 Mbit/s - Not OK!
[root@fn1] ~# top
last pid: 43816; load averages: 0.91, 0.54, 0.37 up 0+05:23:58 18:23:08
67 processes: 2 running, 65 sleeping
CPU: 22.9% user, 0.0% nice, 2.6% system, 3.1% interrupt, 71.4% idle
Mem: 164M Active, 1378M Inact, 12G Wired, 2516M Free
ARC: 8797M Total, 5415M MFU, 3241M MRU, 56M Anon, 38M Header, 47M Other
Swap: 8192M Total, 8192M Free
PID USERNAME   THR PRI NICE   SIZE   RES STATE   C   TIME   WCPU COMMAND
43210 root   1 103   0   541M 65668K CPU1   1   2:21 100.00% smbd
2899 root   6 20   0   367M   162M select 0   2:05   0.10% python2.7
2999 root   12 20   0   241M 20676K nanslp 3   0:32   0.00% collectd

AD Member-Server fn2 - FreeNAS-9.10.1 (d989edd) - Samba version 4.3.11-GIT-UNKNOWN: 650 Mbit/s - OK!
[root@fn2] ~# top
last pid: 39546; load averages: 2.98, 1.26, 0.77 up 0+05:33:04 18:51:44
37 processes: 2 running, 35 sleeping
CPU: 2.0% user, 0.0% nice, 75.8% system, 9.3% interrupt, 13.0% idle
Mem: 65M Active, 500M Inact, 11G Wired, 7436K Cache, 19G Free
ARC: 10G Total, 56M MFU, 10G MRU, 174M Anon, 25M Header, 18M Other
Swap: 10G Total, 10G Free
PID USERNAME   THR PRI NICE   SIZE   RES STATE   C   TIME   WCPU COMMAND
37206 root   1 76   0   325M 28028K RUN   0   0:28 13.87% smbd
2697 root   1 21   0   233M 76260K select 0   0:15   0.10% python2.7
2724 root   6 20   0   380M   168M select 0   2:41   0.00% python2.7

AD Member-Server fn2 - FreeNAS-9.10-STABLE-201606270534 (dd17351) - Samba version 4.3.6-GIT-UNKNOWN: 664 Mbit/s - OK!
[root@fn2] ~# top
last pid: 7498; load averages: 3.77, 1.68, 0.77 up 0+00:08:17 19:51:06
36 processes: 2 running, 34 sleeping
CPU: 4.1% user, 0.0% nice, 79.7% system, 6.1% interrupt, 10.0% idle
Mem: 334M Active, 176M Inact, 9014M Wired, 6316K Cache, 22G Free
ARC: 8362M Total, 51M MFU, 8165M MRU, 105M Anon, 20M Header, 21M Other
Swap: 10G Total, 10G Free
PID USERNAME   THR PRI NICE   SIZE   RES STATE   C   TIME   WCPU COMMAND
7313 root   1 76   0   325M 28664K RUN   0   0:17 14.60% smbd
4124 root   6 20   0   358M   153M select 0   0:07   0.00% python2.7
4097 root   1 21   0   233M 75012K select 0   0:05   0.00% python2.7

AD Member-Server fn3 - FreeNAS-9.10-STABLE-201606270534 (dd17351) - Samba version 4.3.6-GIT-UNKNOWN: 750 Mbit/s - OK!
[root@fn3] ~# top
last pid: 7172; load averages: 6.13, 2.82, 1.38 up 0+00:07:20 19:13:16
38 processes: 2 running, 36 sleeping
CPU: 1.2% user, 0.0% nice, 94.0% system, 0.6% interrupt, 4.2% idle
Mem: 208M Active, 264M Inact, 7270M Wired, 158M Free
ARC: 6900M Total, 51M MFU, 6503M MRU, 311M Anon, 17M Header, 19M Other
Swap: 4096M Total, 4096M Free
PID USERNAME   THR PRI NICE   SIZE   RES STATE   C   TIME   WCPU COMMAND
6997 root   1 82   0   325M 28716K RUN   3   0:36 34.28% smbd
4002 root   6 20   0   354M   153M select 3   0:08   0.10% python2.7
3974 root   1 26   0   217M 57096K select 1   0:02   0.00% python2.7

Here an other system - AD Domain-Controller dc1 openSUSE 42.1 - Samba version 4.4.5:
dc1:~ # top
top - 20:03:30 up 7:04, 1 user, load average: 0.28, 0.09, 0.02
Tasks: 161 total,   2 running, 159 sleeping,   0 stopped,   0 zombie
%Cpu(s): 24.2 us, 1.9 sy, 0.0 ni, 73.5 id, 0.3 wa, 0.0 hi, 0.1 si, 0.0 st
KiB Mem: 4030176 total, 1082164 used, 2948012 free,   1724 buffers
KiB Swap: 2103292 total,   0 used, 2103292 free.   790256 cached Mem
PID USER   PR NI   VIRT   RES   SHR S %CPU %MEM   TIME+ COMMAND
3510 200500   20   0 484696 53120 21744 R 100.0 1.318   0:30.36 smbd
391 root   20   0   12044   5732   1496 S 0.330 0.142   0:01.66 haveged
2034 root   20   0 537988 45516 13656 S 0.330 1.129   0:57.87 samba

#15 - 09/19/2016 08:44 AM - Josh Paetzel

When it's using all the CPU can you do procstat -kk <smbd_pid> so we can try to see what it's doing?

#16 - 09/20/2016 12:17 AM - Alfred Schlütter

Here the results from AD Domain-Controller fn1.

During copy large files to share or from share: cpu 100% (smbd - 69726)
[root@sfn1] ~# procstat -kk 69726
PID TID COMM TDNAME KSTACK
69726 100752 smbd - <running>

Before and after copy: cpu 0 % (smbd - 69726)
[root@sfn1] ~# procstat -kk 69726
PID TID COMM TDNAME KSTACK
69726 100752 smbd - mi_switch+0xe1 sleepq_catch_signals+0xab sleepq_timedwait_sig+0x10 _cv_timedwait_sig_sbt+0x19e seltdwait+0xa4 kern_poll+0x464 sys_poll+0x61 amd64_syscall+0x40f Xfast_syscall+0xfb

#17 - 09/21/2016 11:22 AM - Josh Paetzel

Well, that's not very useful, it's off in userland.

#18 - 09/23/2016 11:27 AM - Alfred Schlütter

Hi Josh, possibly signing the Problem?

See CVE-2016-2118 (Badlock) samba 4.3.8 or CVE-2016-2119 samba 4.3.11.

After disbable signing in smb.conf:

client signing = No
server signing = No
client ipc signing = No

i have 930 Mbit/s and cpu 30% for the smbd pid during copy.

A challenge for the samba team?

#19 - 09/23/2016 11:31 AM - Josh Paetzel

Can you try out a nightly? That has samba 4.4 which boasts better signing performance.

#20 - 09/23/2016 11:54 AM - Alfred Schlütter

I hope i have a little bit time on sunday and can test it in my private network.

#21 - 09/24/2016 12:10 AM - Alfred Schlütter

Ups, when i want to change the train to nighlies freeNAS shows following warning:

Are you sure you want to change trains?
WARNING: Changing to a nightly train is a one way street. Changing back to stable is not supported!

I thought i can went back to STABLE after test. The servers in production, so i wont't leaving stable Train forever. What should i do?

#22 - 09/24/2016 12:22 AM - Jordan Hubbard

Just boot back into the boot environment you were last running -STABLE from, and continue forward on -STABLE from there.

#23 - 09/24/2016 04:49 AM - Alfred Schlütter

Now FreeNAS-9.10-MASTER-201609240510 (c5dc7d1) with samba 4.4.5

During copy large files from or to share:
cpu use for smbd pid: 85% - 95%
but only 400 MBit/s (TX) and 360 MBit/s (RX)

#24 - 12/08/2016 07:14 AM - Kris Moore

- Target version changed from 49 to 9.10.3

Just a FYI. We will be pushing 4.5.X of samba into the nightlies here soon. If you want to re-test with that version any additional data would be appreciated to determine if this is still an issue.

#25 - 12/29/2016 12:13 AM - xing yu ye

Josh Paetzel wrote:

> I'm going to be really pedantic here, but it's important. Based on the information you've provided the only thing we know is SMB is slower. Whether it's a networking issue remains to be seen.
>
> First question to short circuit all of this:
>
> 9.10.1 has a samba fix for the badlock vulnerability. OSX interacted poorly with this. If your clients are Macs and they are NOT running 10.11.6 you either need to upgrade the macs or downgrade FreeNAS.
>
> Otherwise, first order of business is to use iperf (included in FreeNAS) to determine if the network really has slowed down. Run iperf -s on the 9.3 and 9.10 FreeNAS systems. Install iperf2 on your clients and run iperf -c ip.of.freenas9.3 and then 9.10
>
> If it has slowed down we'll troubleshoot that, otherwise we'll move on to eliminating ZFS being the bottleneck.

Hi Josh, Did fix your problem? because I also have the same problem. but just copy back from the server not transfer to server. Also can you please share your system-tuning options to me? may be its about the tuning parameter. Thanks a lot.

#26 - 12/31/2016 07:21 PM - Josh Paetzel

- Status changed from Investigation to Unscreened

- Assignee changed from Josh Paetzel to Kris Moore

#27 - 01/05/2017 06:59 AM - Alfred Schlütter

A short summary. What i (we?) knew:

1) FreeNAS/FreeBSD, ZFS, network are ok and not the bottleneck. See my tests above.

2) High CPU usage (smbd pid) during copying over SMB comes up with CVE-2016-2118 (Badlock) samba 4.3.8 or CVE-2016-2119 samba 4.3.11 on maschines that configured as DC (Domaincontroller).

3) If i disable signing in smb.conf with following parameters, i become full performance back (over 900 Mbit/s):
client signing = No
server signing = No
client ipc signing = No

4) On my private DC on openSUSE with now samba 4.5.3 still the same issue. Slow copying over SMB
and full permance, when signing was disabled - see 3).

5) My clients are all on OS: Windows 7 64bit Professional with Service Pack 1

6) It's not recommended to use samba DC as fileserver … but my budget is small … ;-)

#28 - 01/10/2017 09:08 AM - Kris Moore

- Status changed from Unscreened to Closed: Third party to resolve

Ok, closing this out since it clearly seems like a samba issue. Hopefully if/when they fix it upstream we can pull in those changes.

#29 - 02/16/2017 07:08 AM - Kris Moore

- Target version changed from 9.10.3 to N/A