

Re: Still getting NFS client locking up

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/current/2003-11/0635.html>

From: Robert Watson (rwatson_at_freebsd.org)

Date: 11/10/03

Date: Mon, 10 Nov 2003 11:28:40 -0500 (EST)

To: Matt Smith <matt@xtaz.co.uk>

On Mon, 10 Nov 2003, Matt Smith wrote:

- > *With a current build from november the 9th I am still getting exactly*
- > *the same NFS lockups. I assume soren is as well. NFS has basically been*
- > *pretty unusable now for over a month.*
- >
- > *As only a couple of people have complained about this from what I can*
- > *see I assume it is something related to something specific such as a*
- > *network card?*

I'm fairly baffled. I tried for many hours to reproduce the problem in two separate sets of systems here, and completely failed. I left buildworlds, cvs updates, blah blah blah, running for 96 hours across pools of clients and servers and no hint of the problem. I also use NFS daily on my primary workstation at work, as well as in my normal development setup with diskless crashboxes. So indeed, there must be some very specific piece of the picture that I'm not reproducing, such as a specific network card, or there's a race condition that requires very specific timing, etc.

How fast are your systems, speaking of which? I live in the world of 300-500 mhz machines at work, and 300-800 mhz boxes at home. If you're using multi-ghz boxes, that could well be the distinguishing factor between our configurations...

- > *From my testing I only get this lockup when writing to the server.*
- > *Reading from the server works perfectly all the time. So luckily I can*
- > *still manage an NFS mounted installworld/kernel.*
- >
- > *I just got the lockup again now whilst it downloaded p5-Net-DNS to*
- > *portupgrade into /usr/ports/distfiles. This is a very small file but it*
- > *was enough to trigger it off. So it doesn't look like a size related*
- > *issue either as I can download around 4% of mysql before it locks up.*
- >
- > *Obviously we should really try and find the cause of this before 5.2. I*
- > *am willing to try any patches/debug on my systems. But I just have zero*

freebsd-current: Re: Still getting NFS client locking up

> *clue about what to look for myself.*
>
> *As a start here is the relevent parts of my dmesg to show the NIC's I'm*
> *using. I wonder if this corresponds to sorens?*
>
> *NFS CLIENT (xl1 would be the card it's using to talk to the server):*
> *xl0: <3Com 3c905B-TX Fast Etherlink XL> port 0xe400-0xe47f mem*
> *0xea000000-0xea00007f irq 12 at device 15.0 on pci0*
> *xl0: Ethernet address: 00:a0:24:ac:e1:b4*
> *miibus0: <MII bus> on xl0*
> *xlphy0: <3Com internal media interface> on miibus0*
> *xlphy0: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, auto*
> *xl1: <3Com 3c905-TX Fast Etherlink XL> port 0xe800-0xe83f irq 11 at*
> *device 17.0 on pci0*
> *xl1: Ethernet address: 00:60:08:6d:1e:3b*
> *miibus1: <MII bus> on xl1*
> *nsphy0: <DP83840 10/100 media interface> on miibus1*
> *nsphy0: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, auto*
>
> *NFS SERVER:*
> *xl0: <3Com 3c905C-TX Fast Etherlink XL> port 0x1000-0x107f mem*
> *0xfc304800-0xfc30487f irq 10 at device 7.0 on pci5*
> *xl0: Ethernet address: 00:04:76:8d:c5:fd*
> *miibus0: <MII bus> on xl0*
> *xlphy0: <3c905C 10/100 internal PHY> on miibus0*
> *xlphy0: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, auto*

My server:

xl0: <3Com 3c905B-TX Fast Etherlink XL> port 0xd880-0xd8ff mem
0xff202000-0xff20207f irq 11 at device 17.0 on pci0
xl0: Ethernet address: 00:b0:d0:29:ec:ce
miibus2: <MII bus> on xl0
xlphy0: <3Com internal media interface> on miibus2
xlphy0: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, auto

My client1:

xl0: <3Com 3c905B-TX Fast Etherlink XL> port 0xdc00-0xdc7f mem
0xff000000-0xff00007f irq 11 at device 17.0 on pci0
xl0: Ethernet address: 00:c0:4f:0d:6b:bc
miibus0: <MII bus> on xl0
xlphy0: <3Com internal media interface> on miibus0
xlphy0: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, auto

My client2:

xl0: <3Com 3c905B-TX Fast Etherlink XL> port 0xd880-0xd8ff mem
0xff202000-0xff20207f irq 11 at device 17.0 on pci0
xl0: Ethernet address: 00:b0:d0:2b:76:d5
miibus2: <MII bus> on xl0

Re: Still getting NFS client locking up

freebsd-current: Re: Still getting NFS client locking up

xlphy0: <3Com internal media interface> on miiibus2
xlphy0: 10baseT, 10baseT-FDX, 100baseTX, 100baseTX-FDX, auto

> *Both connected to a 100meg full duplex switch.*

Ditto.

> *Any ideas? As I have said I'm happy to enable some major debugging etc.*
> *But I just need somebody to give me a step by step guide for what to do*
> *and look for.*
> *In case this thread is too old now and nobody remembers anything about*
> *it the previous email regarding it is at*
>

<http://docs.freebsd.org/cgi/getmsg.cgi?fetch=1183410+0+archive/2003/freebsd-current/20031102.freebsd-current>

Ok, here's the strategy I was planning to take once I could reproduce it:

- (1) Attempt to further narrow down responsibility to client/server. In particular, see if an apparent hang on one client affects the other clients.
- (2) Investigate Soren's report that killing and restarting nfsd on the server would clear the hang.
- (3) Look at stack traces of involved processes on both the client and server: in particular, look at traces for any client blocked in NFS, any nfsiod processes on the client, and the nfsd processes on the server. Also look at the wait channels on clients and servers for these processes. Particularly interested in whether nfsd processes are blocked trying to grab locks.
- (4) Look at netstat information for NFS sockets, in particular, if the buffers are full, or not being drained. In particular, on the server, is the input queue not being drained by nfsd worker threads?
- (5) Try backing out src/sys/nfsserver/nfs_serv.c:1.137, which removed another deadlock problem, but did change locking behavior in the NFS server.
- (6) Look at packet traces between the client and server with ethereal, which has pretty good NFS decoding. Is the client retransmitting an RPC to the server and the server just isn't responding, or is the client failing to transmit? At the point of the hang, what sorts of RPCs are outstanding to the server? In the past, we've seen "apparent hangs" when some or another more obscure unusual error case on the NFS server fails to respond to an RPC, which causes the client to "wait forever".

To do all this, you'll want to compile DDB into your kernel, and make sure you have a copy of the kernel on-disk with debugging symbols. Do this on both the client and the server. To generate stack traces, you can break

freebsd-current: Re: Still getting NFS client locking up

to the debugger on the console of each system, use the debugger "ps" command to identify victims, then "trace pid" replacing "pid" with the pid number of interest. Ideally, you'll use a serial console and use serial break (requires BREAK_TO_DEBUGGER) to do this so you can copy and paste traces rather than having to hand-transcribe.

Things to look for: normally, idle nfsd and nfsiod processes have a WCHAN of "-" (ps -lax), which indicates they're blocked waiting for some event to kick them off. If you see nfsd processes "hung" in another state, it's a good sign we've identified a server problem. In the nfs client processes, "nfsrvlk" typically indicates a process has sent out an RPC and is now waiting on a response.

If you need any help getting debugging stuff up and running, let me know.

Robert N M Watson FreeBSD Core Team, TrustedBSD Projects
robert@fledge.watson.org Network Associates Laboratories

freebsd-current@freebsd.org mailing list

<http://lists.freebsd.org/mailman/listinfo/freebsd-current>

To unsubscribe, send any mail to "freebsd-current-unsubscribe@freebsd.org"