

Re: NFS Problems/Questions

Source: <http://unix.derkeiler.com/Mailing-Lists/FreeBSD/questions/2007-06/msg01624.html>

- *From:* Jason Morgan <jwm-freebsd-questions@xxxxxxxxxxxxxxxxxxxxxx>
 - *Date:* Sat, 23 Jun 2007 19:42:24 -0400
-

On Sat, Jun 23, 2007 at 12:46:27PM -0700, Michael Smith wrote:

Hello Jason:

On Jun 23, 2007, at 9:34 AM, Jason Morgan wrote:

I've been having some trouble with NFS performance for some time and now that class is out, I've had a bit of time to investigate but I'm stuck. Below are the details of my investigation. Hopefully, someone here can give me some advice.

The basic problem is that my NFS performance is very slow. Right now, I am connecting two workstations to a NFS server, which has my home directory, etc, mounted. They are connected over a gigabit network (right now with mtu set to 7000, which is supported by all hardware -- changing it to 1500 has no effect on performance, which is strange). Each system is running 6.2-RELEASE or -STABLE. Each system is also using the following network card:

```
# ifconfig sk0
sk0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST>
mtu 7000
options=b<RXCSUM,TXCSUM,VLAN_MTU>
inet 10.0.0.2 netmask 0xfffff00 broadcast 10.0.0.255
ether 00:17:9a:bb:05:87
media: Ethernet autoselect (1000baseTX <full-
duplex,flag0,flag1>)
status: active
```

```
# dmesg | grep sk
skc0: <D-Link DGE-530T Gigabit Ethernet> port 0xec00-0xecff mem
0xfdf8000-0xfdfbfff irq 18 at device 10.0 on pci0
skc0: DGE-530T Gigabit Ethernet Adapter rev. (0x9)
sk0: <Marvell Semiconductor, Inc. Yukon> on skc0
sk0: Ethernet address: 00:17:9a:XX:XX:XX
```

```
## Server /etc/rc.conf settings
```

Re: NFS Problems/Questions

```
rpcbind_enable="YES"
rpc_lockd_enable="YES"
rpc_statd_enable="YES"
nfs_server_enable="YES"
nfs_server_flags="-u -t -n 12"
nfs_bufpackets="32"
mountd_flags="-r"
```

```
## Client /etc/rc.conf settings
```

```
nfs_client_enable="YES"
nfs_bufpackets="32"
nfsiod_enable="YES"
nfsiod_flags="-n 6"
rpc_lockd_enable="YES"
rpc_statd_enable="YES"
rpcbind_enable="YES"
```

```
## /etc/exports
```

```
/usr -alldirs,maproot=root client1 client2
```

For performance benchmarking, I am using dd. Locally from the server, this is a representative result when writing a 1GB file:

```
## Local write test (for an upper-bound on what to expect).
```

```
# dd if=/dev/zero of=./nfs.dat bs=1024k count=1000
1000+0 records in
1000+0 records out
1048576000 bytes transferred in 19.580184 secs (53552919 bytes/sec)
```

Connecting from a client (both clients get approximately the same results).

```
## Remote connection (UDP), mounted in /etc/fstab as with flags:
## rw,-U,-3,-r=32768,-w=32768
```

```
# dd if=/dev/zero of=./nfs.dat bs=1024k count=1000
1000+0 records in
1000+0 records out
1048576000 bytes transferred in 101.151139 secs (10366428 bytes/sec)
```

```
## Remote connection (TCP), mounted in /etc/fstab as with flags:
## rw,-T,-3,-r=32768,-w=32768
```

```
# dd if=/dev/zero of=./nfs.dat bs=1024k count=1000
1000+0 records in
1000+0 records out
```

Re: NFS Problems/Questions

1048576000 bytes transferred in 59.668585 secs (17573334 bytes/sec)

As can be seen above, TCP is much faster than UPD. I have tried many different mount settings and these are the best results I could get. To test whether or not I have having network issues, I transferred the same nfs.dat file via a http connection and got ~32MB/sec --- almost 2x the speed of the TCP NFS connection. 32MB/sec is about what I would expect given that my fastest write speed is ~50MB/sec.

At this point I am stumped. I have tried increasing/changing the number of nfsiod servers as well as nfs_bufpackets. No matter what settings I change, the results are always the same. I get only two errors, first on /var/log/messages on the server I have just begun seeing:

```
Jun 22 21:13:47 crichton routed[666]: sendto(dc1, 224.0.0.2):  
Operation not permitted  
Jun 22 21:13:47 crichton routed[666]: sendto(sk0, 224.0.0.2):  
Operation not permitted  
Jun 22 21:13:50 crichton routed[666]: sendto(dc1, 224.0.0.2):  
Operation not permitted  
Jun 22 21:13:50 crichton routed[666]: sendto(sk0, 224.0.0.2):  
Operation not permitted
```

This appeared after I added a route; however, I added the route after many of the tests were done. I get the same results now as before the new route. On one of the clients (the one running 6.2-RELEASE-p1), I also get a nasty error:

```
nfs/tcp clnt: Error 60 reading socket, tearing down TCP connection
```

This cropped up last night after I tweaked some settings. They have now been changed back, but I still get this error. The other client is unaffected.

I appreciate any help people can provide on tracking down the issues. Sorry about the long email --- just trying to be thorough. Of course, I've searched the Internet and can't find any clear assistance on these issues.

Cheers,
~Jason

We use the following settings on a mail cluster that's pushing about 50 MB/sec sustained.

```
10.211.1.213:/m0 /mail/m0 nfs  
rw,tcp,intr,noatime,nfsv3,-w=65536,-r=65536
```

Re: NFS Problems/Questions

```
# NFS Server
rpcbind_enable="YES"
rpc_lockd_enable="YES"
rpc_statd_enable="YES"
nfs_server_enable="YES"
nfs_server_flags="-u -t -n 16 -h 10.211.1.213"
mountd_flags="-r"
```

I would imagine the larger read/write values above would be fine for you as well, given you have Gigabit links. The 'noatime' setting may be problematic depending on your application. You might want to Google specifics on what applications use atime to see if that's a good flag for you or not.

I'd love to see your results if you decide to test those settings offline.

Regards,

Mike

Mike, thank you for the advice. I followed it, but didn't make much progress. I am still pushing the same ~17M/sec. I did mess with net.inet.tcp.sendspace and net.inet.tcp.recvspace some and managed to get a couple more MB/sec out of the connection, but not much else. It's quite possible I am missing something obvious.

Thanks again,

~Jason

freebsd-questions@xxxxxxxxxxx mailing list

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

To unsubscribe, send any mail to "freebsd-questions-unsubscribe@xxxxxxxxxxx"