Re: nfs server overload (nfsd)

Source: http://unix.derkeiler.com/Mailing−Lists/FreeBSD/isp/2005−12/msg00062.html

- *From*: Eric Anderson <anderson@xxxxxxxxxxxx>
- *Date*: Thu, 29 Dec 2005 07:04:33 −0600

```
Angel Blazquez wrote:


      Hello,


      We are expecting incredible overload in a NFS server. A top shows nfsd
      consuming most of the CPU:


      PID USERNAME PRI NICE SIZE RES STATE C TIME WCPU CPU COMMAND
      6000 root −8 0 1204K 660K biord 1 124:15 27.88% 27.88% nfsd
      6002 root 4 0 1204K 660K *Giant 0 124:18 17.58% 17.58% nfsd
      6006 root 4 0 1204K 660K *Giant 0 123:38 10.21% 10.21% nfsd
      6005 root 4 0 1204K 660K *Giant 0 123:36 7.47% 7.47% nfsd
      6003 root 4 0 1204K 660K *Giant 0 123:08 4.15% 4.15% nfsd
      6001 root 4 0 1204K 660K *Giant 0 123:16 2.83% 2.83% nfsd



During these loads, can you run nfsstat −s −w 1 on the server and see
what is going on?


      Memory looks fine:


      Mem: 27M Active, 910M Inact, 136M Wired, 51M Cache, 112M Buf, 1828K Free
      Swap: 2048M Total, 72K Used, 2048M Free


      Typing in the nfs server (console/ssh) becomes terrible, the server does
      not reply well.
```

We are running this nfs server in FreeBSD 5.3-RELEASE-p23 on a Compaq Proliant server with a Compaq Smart Array 5300 that comunicates with a array of disks:

```
/dev/da0s1d 164G 124G 27G 82% /data0
/dev/da1s1d 131G 80G 41G 66% /data1
```

You may want to also look at gstat and see how busy your disks look.

We have /data0 and /data1 exported:

```
/data0   -maproot=root -alldirs -network 192.168.62.0 -mask 255.255.255.0
/data1   -maproot=root -alldirs -network 192.168.62.0 -mask 255.255.255.0
```

so a couple of incoming SMTP servers we have can deliver e-mail to those filesystems.
We are running exim 4.60.0 in those other servers, 4.10-RELEASE-p5 in one of them, and FreeBSD 6.0-RELEASE #0 in the other one.

If we stop exim delivering e-mail, nfs server does well, the cpu gets
free, and the nfs server works fine (replies to user interaction, etc).

You may try mounting the filesystems on the server with the 'noatime' option to reduce disk writes.

You can also try setting the gatherdelay down from 10000 to 1000 on the server and see if that helps.

Also - are you sure you want:
vfs.nfsrv.async: 1

I fairly certain default is 0, which is safer.

FreeBSD 6.0 sysctl output (nfs related):

```
vfs.nfs4.access_cache_timeout: 60
vfs.nfs4.nfsv3_commit_on_close: 0
vfs.nfs.downdelayinitial: 12
vfs.nfs.downdelayinterval: 30
vfs.nfs.realign_test: 1294030
vfs.nfs.realign_count: 0
vfs.nfs.bufpackets: 4
vfs.nfs.reconnects: 2
vfs.nfs.iodmaxidle: 120
vfs.nfs.iodmin: 4
vfs.nfs.iodmax: 20
vfs.nfs.defect: 0
vfs.nfs.nfs_ip_paranoia: 1
vfs.nfs.diskless_valid: 0
vfs.nfs.diskless_rootpath:
vfs.nfs.access_cache_timeout: 2
vfs.nfs.nfsv3_commit_on_close: 0
vfs.nfs.clean_pages_on_close: 1
vfs.nfs.nfs_directio_enable: 0
vfs.nfs.nfs_directio_allow_mmap: 1
vfs.nfsrv.nfs_privport: 0
vfs.nfsrv.async: 0
vfs.nfsrv.commit_blks: 0
vfs.nfsrv.commit_miss: 0
vfs.nfsrv.realign_test: 0
vfs.nfsrv.realign_count: 0
vfs.nfsrv.gatherdelay: 10000
vfs.nfsrv.gatherdelay_v3: 0
```

FreeBSD 4.10 sysctl output (nfs related):

```
vfs.nfs.nfs_privport: 0
vfs.nfs.async: 0
vfs.nfs.commit_blks: 0
vfs.nfs.commit_miss: 0
vfs.nfs.realign_test: 84602323
vfs.nfs.realign_count: 99713
vfs.nfs.bufpackets: 4
vfs.nfs.gatherdelay: 10000
vfs.nfs.gatherdelay_v3: 0
vfs.nfs.defect: 0
vfs.nfs.nfs_ip_paranoia: 1
vfs.nfs.diskless_valid: 0
vfs.nfs.diskless_rootpath:
vfs.nfs.diskless_swappath:
vfs.nfs.access_cache_timeout: 2
vfs.nfs.nfsv3_commit_on_close: 0
```

This couple of servers mounts the filesystems with this options:

192.168.62.54:/data1 /mail nfs
rw,nfsv3,intr,dumbtimer,rdirplus,nosuid,nodev 0 0
192.168.62.54:/data0 /data0 nfs
rw,nfsv3,intr,dumbtimer,rdirplus,nosuid,nodev 0 0

On the server, sysctl nfs related output looks like this:

```
vfs.nfs.downdelayinitial: 12
vfs.nfs.downdelayinterval: 30
vfs.nfs.realign_test: 2694
vfs.nfs.realign_count: 0
vfs.nfs.bufpackets: 4
vfs.nfs.reconnects: 2
vfs.nfs.iodmaxidle: 120
vfs.nfs.iodmin: 4
```

```
vfs.nfs.iodmax: 20
vfs.nfs.defect: 0
vfs.nfs.nfs_ip_paranoia: 1
vfs.nfs.diskless_valid: 0
vfs.nfs.diskless_rootpath:
vfs.nfs.access_cache_timeout: 2
vfs.nfs.nfsv3_commit_on_close: 0
vfs.nfs4.access_cache_timeout: 60
vfs.nfs4.nfsv3_commit_on_close: 0
vfs.nfsrv.nfs_privport: 0
vfs.nfsrv.async: 1
vfs.nfsrv.commit_blks: 579238
vfs.nfsrv.commit_miss: 413059
vfs.nfsrv.realign_test: 88269083
vfs.nfsrv.realign_count: 11961
vfs.nfsrv.gatherdelay: 10000
vfs.nfsrv.gatherdelay_v3: 0
debug.hashstat.nfsnode: 65536 5 1 0
```

Thanks in advance,

Best regards,
Angel Blazquez

_____
freebsd-isp@xxxxxxxxxxx mailing list
http://lists.freebsd.org/mailman/listinfo/freebsd-isp
To unsubscribe, send any mail to

"freebsd-isp-unsubscribe@xxxxxxxxxxx"

_____
freebsd-isp@xxxxxxxxxxx mailing list
http://lists.freebsd.org/mailman/listinfo/freebsd-isp
To unsubscribe, send any mail to "freebsd-isp-unsubscribe@xxxxxxxxxxx"