

My Sun Storage 7310 perf limits

27 May 2009

I originally posted this at

http://blogs.sun.com/brendan/entry/my_sun_storage_7310_perf.

As part of my role in Fishworks, I push systems to their limits to investigate and solve bottlenecks. Limits can be useful to consider as a possible upper bound of performance, as it shows what the target can do. I previously posted my [results](#) for the Sun Storage 7410, which is our current top performing product. Today the Sun Storage 7310 has been [launched](#), which is an entry level offering that can be clustered for high availability environments like the 7410.

The following summarises the performance limits I found for a single 7310 head node, along with the current 7410 results for comparison:

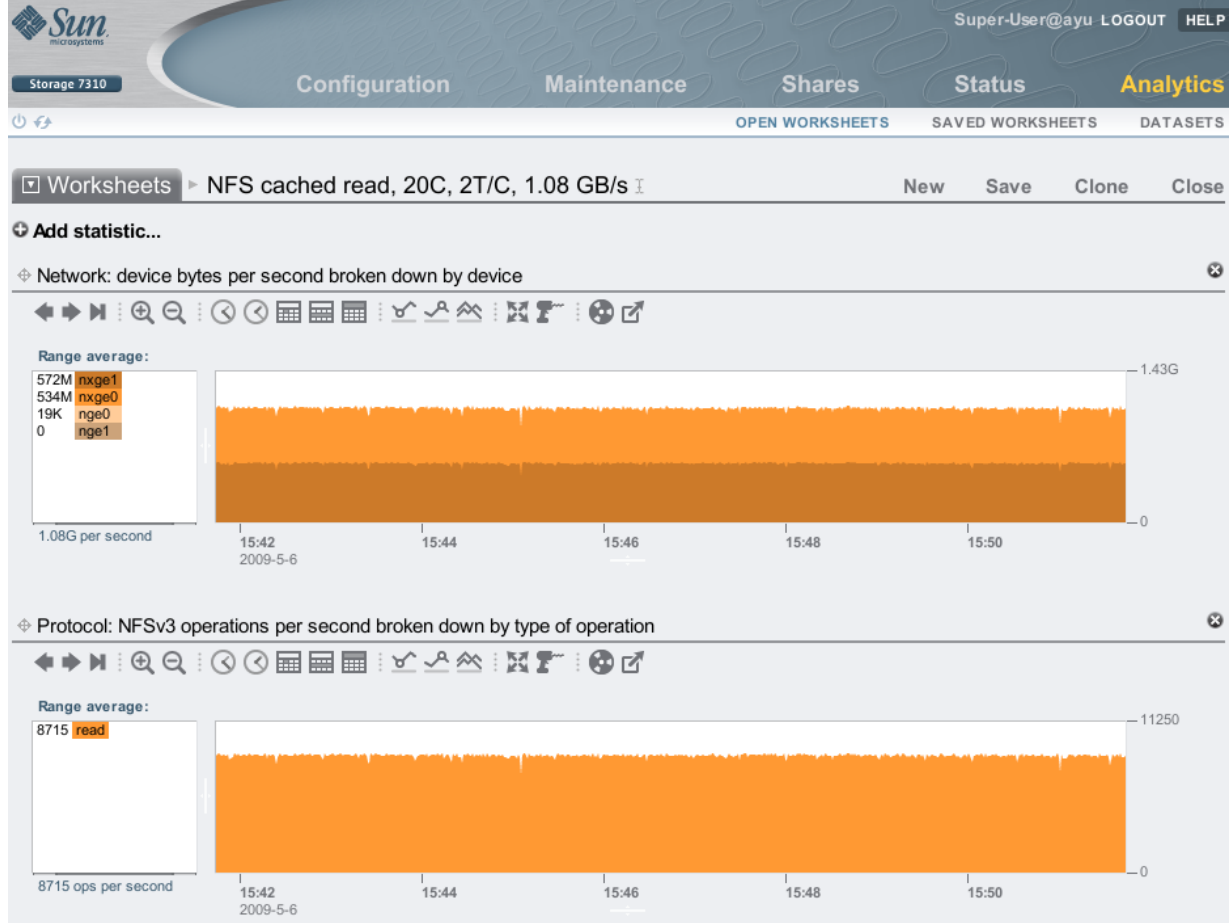
Workload	7310 Result	7410 Result
NFSv3 streaming read from DRAM	~1.08 Gbytes/sec	~2.02 Gbytes/sec
NFSv3 streaming read from disk	780 Mbytes/sec	1.20 Gbytes/sec
NFSv3 streaming write to disk	~477 Mbytes/sec	~595 Mbytes/sec
NFSv3 read IOPS from DRAM	180,000 IOPS	281,000 IOPS

As shown above, the 7310 has very reasonable performance in comparison to the high end 7410. (As of this blog post, I've yet to update my [7410 results](#) for the [2009.Q2 software update](#), which gave about a 5% performance boost).

Next I'll show the source of these results, screenshots taken from the 7310 using [Analytics](#), to measure how the target NFS server actually performed. I'll finish by describing the 7310 and clients used to perform these tests.

NFSv3 streaming read from DRAM

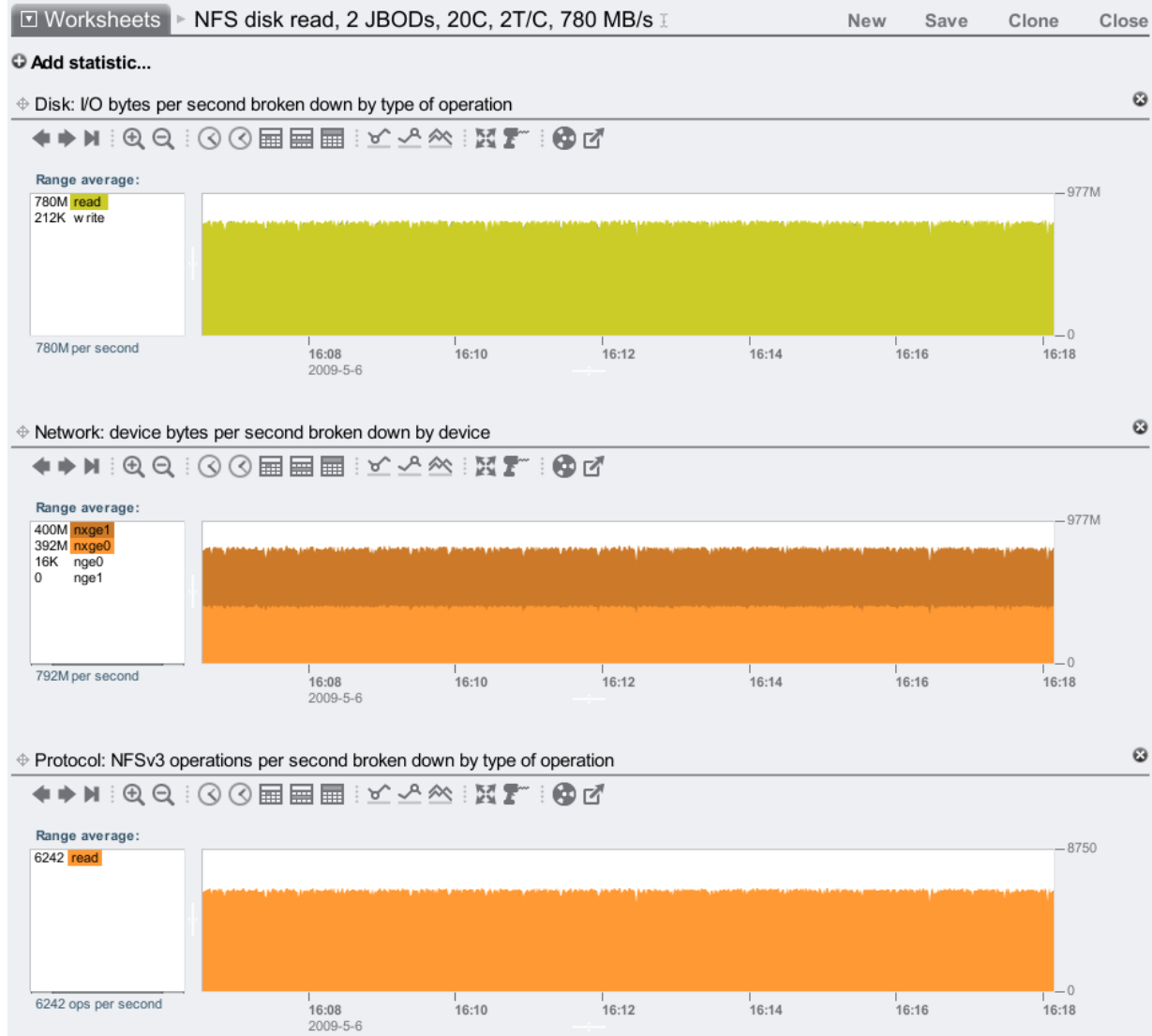
To find the fastest read throughput possible, I used 10 Gbytes of files (working set) and had 20 clients run two threads per client to read 1 Mbyte I/O through them repeatedly:



Over this 10 minute interval, we've averaged 1.08 Gbytes/sec. This includes both inbound NFS requests and network protocol headers, so the actual data transferred will be a little less. Still, breaking 1 Gbyte/sec of delivered NFS for an entry level server is a great result.

NFSv3 streaming read from disk

This time 400 Gbytes of files were used as the working set, to minimize caching in the 7310's 16 Gbytes of DRAM. As before, 20 clients and 2 threads per client read through the working set repeatedly:



This screenshot from Analytics includes the disk bytes/sec, to confirm that this workload really did read from disk. It averaged 780 Mbytes/sec – a solid result. The 792 Mbytes/sec average on the network interfaces includes the NFS requests and network protocol headers.

NFSv3 streaming write to disk

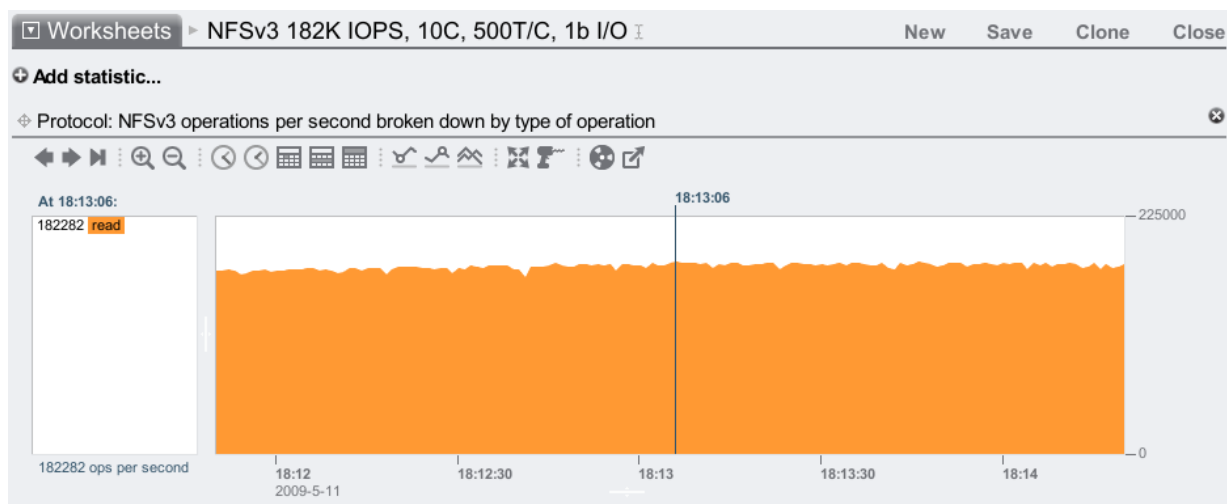
To test streaming writes, 20 clients ran 5 threads per client performing writes with a 32 Kbyte I/O size:



While there were 477 Mbytes/sec on the network interfaces (which includes network protocol headers and ACKs), at the disk level the 7310 has averaged 1.09 Gbytes/sec. This is due to software mirroring, which has doubled the data sent to the storage devices (plus ZFS metadata). The actual data bytes written will be a little less than 477 Mbytes/sec.

NFSv3 max read IOPS

I tested this with the 7410, more as an experiment than of practical value. This is the most 1 byte reads that NFS can deliver to the clients, with the 10 Gbyte working set entirely cached in DRAM. While 1 byte I/O isn't expected, that doesn't render this test useless – it does give the absolute upper bound for IOPS:



The 7310 reached 182,282 reads/sec, and averaged around 180K.

NFSv3 4 Kbyte read IOPS from DRAM

For a more realistic test of read IOPS, the following shows an I/O size of 4 Kbytes, and a working set of 10 Gbytes cached in the 7310's DRAM. Each client runs an additional thread every minute (stepping up the workload), from one to ten threads:



This screenshot shows %CPU utilization, to check that this ramping up of workload is pushing some limit on the 7310 (in this case, what is measured as CPU utilization). With 10 threads running per client, the 7310 has served a whopping 110,724 x 4 Kbyte read ops/sec, and still has a little CPU headroom for more. Network bytes/sec was also included in this screenshot to double check the result, which should be at least be 432 Mbytes/sec (110,724 x 4 Kbytes), which it is.

Configuration

As a **filer**, I used a single [Sun Storage 7310](#) with the following config:

- 16 Gbytes DRAM
- 2 JBODs, each with 24 x 1 Tbyte disks, configured with mirroring
- 1 socket of quad-core AMD Opteron 2300 MHz CPU
- 1 x 2x10 GbE card, jumbo frames
- 1 x HBA card
- noatime on shares, and database size left at 128 Kbytes

It's not a max config system. The 7310 can currently scale to 4 JBODs and have 2 sockets of quad-core Opteron.

The **clients** were 20 blades, each:

- 2 sockets of Intel Xeon quad-core 1600 MHz CPUs
- 3 Gbytes of DRAM
- 1 x 1 GbE network ports
- Running Solaris, and NFS mounted `forcedirectio` for read testing to avoid client caching

These are great, apart from the CPU clock speed, which at 1600 MHz is a little low.

The network consists of multiple 10 GbE switches to connect the client 1 GbE ports to the filer 10 GbE ports.

Conclusion

We've spent years working on maximizing performance of our highest end offering, the Sun Storage 7410. New family members of the Sun Storage 7000 series inherit most of this work, and really hit the ground running. When [Adam Leventhal](#) designed the 7310 product, we weren't thinking of a system capable of 1 Gbyte/sec, but as the pieces came together we realized the performance could in fact be very good, and was. And it's worth emphasizing – the results above are all from a single socket 7310; the 7310 can be configured with two sockets of quad core Opteron!