Data Transport System (DTS) Transfer Test Results

Mike Fitzpatrick
November 12, 2010

Executive Summary

During the week of 11/1/2010, several tests of DTS were conducted to measure the effective throughput that can be reliably achieved using the existing network connections from Chile to Tucson. A *Golden Standard Night* (GSN) of DES data was used as the test dataset, this is smaller (only 87 GB) than the average DES night that is expected to produce approximately 200-300 GB that must be transferred in 18 hours. Given the expected DES data volume and the 18-hour transfer window, we chose 30 Mbs as a target transfer rate since DTS should then complete a minimal DES night in approximately 15 hours. Images were transferred at 15-minute intervals so as not to burden an operational system, as well as to allow us to sample network performance at different times of the day.

Our results show that we were able to maintain an average transfer rate of 29.6 (± 2.4) Mbs for the duration of the GSN transfer test (~40 hours) using 50 threads. Further tests conducted using a range of transfer threads indicate it is possible to tune the DTS for a desired transfer rate, although we do note some greater variability when targeting higher transfer speeds. On the existing Chile-Tucson link transfer rates as high as 48 Mbs were seen, a target rate of 40 Mbs (i.e. what is needed to transfer a 300 GB DES night) *can* be achieved, but not consistently.

These results reaffirm that DTS performance tests conducted during the development phase meet the design goals. Tests will, of course, need to be repeated once the underlying network infrastructure is upgraded and/or the final system deployment is in place. We have no reason to believe that similar tests over another data path would show significant differences.

Test Descriptions

Two separate tests were conducted: the primary test was to see whether DTS could sustain an average throughput necessary to transfer an minimal night's DES data volume (~200 GB) within the 18-hour constraint required for DES. A second test sampled the throughput that could be achieved (and our ability to tune the DTS for given network performance) over a range in the number of threads used to transport data.

Because we could not suspend normal operations to conduct this test, it was designed to be *representative* of the results we can expect and not an exact measure. In the GSN transfer test, images were transferred at 15-minute intervals until all 142 files were moved. The source machine at CTIO was dtsct-hs.ctio.noao.edu (139.229.13.187) and the destination machine was tucana.tuc.noao.edu (140.252.1.86) in Tucson. The intercontinental network link out of Chile is the slowest, highest-latency connection and will be the bottleneck regardless of data path so we do not expect results using other source/destination machines to significantly affect results. Transfers were done using 50 parallel threads since we expected this to produce the desired throughput, however we note

that this puts a significant load on the machine at each end of the connection, essentially making the machine unusable to interactive use during transfer periods. Dedicated machines for DTS transfers are recommended. Detailed timings for each transfer are listed in Appendix I.

We were not able to monitor other network usage during the test period, but there were no significant prolonged drops in throughput to indicate that transfers during the nighttime hours would be any worse than daytime use. Transfer of all 142 images completed successfully, however there were two interruptions during the test period: In one case the source machine was rebooted in the middle of image 99, in another the destination disk filled up while transferring image 140. In both cases the DTS recovery was as expected.

A second test was conducted which transferred the same file (chosen to be about the size of the average DECam image in a DES night) sequentially 3 times, but for each set we changed the number of transfer threads used. As expected, the effective throughput varies with the number of threads and reaches a peak when the load on the machine to service each thread cannot keep up with the requests. Further tests would need to be conducted to see whether a more capable machine (e.g. a faster quad-core system) could extend this peak to allow more transfer threads to be used.

Since this test was designed to run the system at full capacity we chose a small number of transfers (3 at each thread setting) to minimize the impact for other users of the network. When tuning the DTS for final deployment, more extensive testing at each setting will be required to choose an optimal number of transfer threads. Detailed timings for each transfer are listed in Appendix II.

Sample Data Set vs. A DES Night

The Golden Standard Night (GSN) data set consists of a total of 142 images including 60 calibration data frames (10 Bias and 50 Flats). When FITS-compressed, these amount to approximately 87 GB in total, of which ~37 GB is calibration data. A typical DES night is expected to produce up to ~360 object images, however the calibration data should amount to roughly the same data volume for each night (i.e. ~100 GB, perhaps less if dark frames are not required); variability in the DES data volume will come from the object frames where field density, filter used and noise characteristics all affect compression efficiency and the number of object frames depends on observing conditions.

Assuming a compression factor of 2 for all data, we are still left with the prospect of transferring ~300 GB/night within the required 18-hour timeframe. This will require a minimum sustained throughput of ~37 Mbs for 300 GB. Keeping up with the data at the rate at which it is acquired will require a speed on the order of 50+ Mbs, but this has the benefit that the network will be available for other uses (e.g. PI data retrieval, archive replication, etc) during non-observing times of the day. While DTS is capable of saturating a network connection using many transfer threads, there is currently not enough bandwidth on the shared 50 Mbs link out of Chile to adequately test whether 37 Mbs could be sustained without interfering with normal use of the network for prolonged periods. Note also, these numbers do *not* take into account the bandwidth required to process a backlog of transfers due to extended network outages.

Appendix I: Detailed GSN Transfer Timings

File Name	Size (MB)	Speed (Mbs)	Time (sec)
decam--20-0-Y-0.fits.fz	570.647	27.07	168.610
decam--20-0-Y-1.fits.fz	550.384	28.43	154.842
decam--20-0-g-0.fits.fz	485.443	34.87	111.350
decam--20-0-g-1.fits.fz	478.343	30.22	126.616
decam--20-0-i-0.fits.fz	548.098	23.86	183.703
decam--20-0-i-1.fits.fz	570.526	28.50	160.115
decam--20-0-r-0.fits.fz	510.746	30.68	133.174
decam--20-0-r-1.fits.fz	518.852	29.94	138.628
decam--20-0-z-0.fits.fz	627.427	30.67	163.602
decam--20-0-z-1.fits.fz	617.284	29.22	168.960
decam--25--42-Y-1.fits.fz	594.102	31.51	150.801
decam--25--42-Y-2.fits.fz	593.843	31.24	152.059
decam--25--42-Y-3.fits.fz	595.633	31.67	150.424
decam--25--42-Y-4.fits.fz	606.562	30.34	159.927
decam--25--42-Y-5.fits.fz	605.521	32.17	150.535
decam--25--42-Y-6.fits.fz	627.843	28.36	177.096
decam--25--42-g-1.fits.fz	519.741	33.51	124.064
decam--25--42-g-2.fits.fz	508.540	19.08	213.226
decam--25--42-g-3.fits.fz	535.115	33.17	129.047
decam--25--42-g-4.fits.fz	526.197	33.83	124.413
decam--25--42-g-5.fits.fz	528.070	23.78	177.586
decam--25--42-g-6.fits.fz	504.045	32.65	123.477
decam--25--42-g-7.fits.fz	525.473	32.44	129.576
decam--25--42-g-8.fits.fz	541.855	29.75	145.664
decam--25--42-i-1.fits.fz	604.579	31.78	152.150
decam--25--42-i-10.fits.fz	596.454	30.37	157.100
decam--25--42-i-11.fits.fz	636.155	29.66	171.525
decam--25--42-i-12.fits.fz	629.488	30.50	165.068
decam--25--42-i-13.fits.fz	628.683	29.16	172.419
decam--25--42-i-14.fits.fz	586.249	31.03	151.130
decam--25--42-i-15.fits.fz	636.918	23.18	219.805
decam--25--42-i-2.fits.fz	608.625	28.19	172.671
decam--25--42-i-3.fits.fz	609.109	30.15	161.614
decam--25--42-i-4.fits.fz	586.944	22.97	204.382
decam--25--42-i-5.fits.fz	593.931	30.52	155.654
decam--25--42-i-6.fits.fz	591.984	29.52	160.380
decam--25--42-i-7.fits.fz	596.487	30.79	154.930
decam--25--42-i-8.fits.fz	656.300	28.66	183.146
decam--25--42-i-9.fits.fz	620.856	29.23	169.876
decam--25--42-r-1.fits.fz	585.514	31.29	149.678
decam--25--42-r-2.fits.fz	564.753	29.49	153.192
decam--25--42-r-3.fits.fz	582.060	29.21	159.395
decam--25--42-r-4.fits.fz	587.591	31.58	148.823
decam--25--42-r-5.fits.fz	563.470	31.85	141.511
decam--25--42-r-6.fits.fz	595.875	29.17	163.377
decam--25--42-r-7.fits.fz	575.098	31.64	145.377
decam--25--42-r-8.fits.fz	592.237	30.25	156.601
decam--25--42-z-1.fits.fz	686.623	28.55	192.350
decam--25--42-z-10.fits.fz	660.428	30.39	173.833
decam--25--42-z-11.fits.fz	651.284	30.71	169.638
decam--25--42-z-12.fits.fz	649.403	30.01	173.059
decam--25--42-z-13.fits.fz	674.940	31.35	172.219
decam--25--42-z-14.fits.fz	666.986	31.25	170.720
decam--25--42-z-15.fits.fz	666.143	29.89	178.238
decam--25--42-z-2.fits.fz	644.732	30.02	171.796
decam--25--42-z-3.fits.fz	661.260	30.74	172.039
decam--25--42-z-4.fits.fz	670.630	29.02	184.839
decam--25--42-z-5.fits.fz	660.684	24.19	218.480

decam--25--42-z-6.fits.fz	641.435	31.10	164.957
decam--25--42-z-7.fits.fz	642.622	29.48	174.339
decam--25--42-z-8.fits.fz	656.181	30.60	171.493
decam--25--42-z-9.fits.fz	665.444	27.88	190.924
decam-25-0-Y-0.fits.fz	598.379	30.29	158.004
decam-25-0-Y-1.fits.fz	550.676	30.29	145.436
decam-25-0-g-0.fits.fz	469.331	30.81	121.833
decam-25-0-g-1.fits.fz	450.470	33.21	108.507
decam-25-0-i-0.fits.fz	551.946	29.02	152.140
decam-25-0-i-1.fits.fz	550.448	31.18	141.185
decam-25-0-r-0.fits.fz	505.045	31.75	127.233
decam-25-0-r-1.fits.fz	529.020	25.71	164.589
decam-25-0-z-0.fits.fz	599.941	30.07	159.604
decam-25-0-z-1.fits.fz	612.547	28.78	170.260
decam-34-0-Y-0.fits.fz	583.041	30.80	151.408
decam-34-0-Y-1.fits.fz	564.310	30.05	150.191
decam-34-0-g-0.fits.fz	472.275	27.89	135.441
decam-34-0-g-1.fits.fz	471.714	28.31	133.265
decam-34-0-i-0.fits.fz	568.945	29.48	154.377
decam-34-0-i-1.fits.fz	550.335	25.87	170.144
decam-34-0-r-0.fits.fz	539.219	32.18	134.009
decam-34-0-r-1.fits.fz	516.946	30.87	133.927
decam-34-0-z-0.fits.fz	610.831	31.85	153.381
decam-34-0-z-1.fits.fz	629.339	29.89	168.394
decam-bias-0.fits.fz	371.235	33.74	88.001
decam-bias-1.fits.fz	371.246	31.65	93.813
decam-bias-2.fits.fz	371.064	30.78	96.430
decam-bias-3.fits.fz	371.158	35.05	84.705
decam-bias-4.fits.fz	371.103	34.01	87.285
decam-bias-5.fits.fz	371.051	32.48	91.361
decam-bias-6.fits.fz	371.059	32.24	92.065
decam-bias-7.fits.fz	371.216	30.92	96.013
decam-bias-8.fits.fz	371.089	31.28	94.905
decam-bias-9.fits.fz	371.230	34.05	87.196
decam-flat-Y-0.fits.fz	615.725	30.42	161.878
decam-flat-Y-1.fits.fz	615.714	30.75	160.162
decam-flat-Y-2.fits.fz	615.714	30.44	161.778
decam-flat-Y-3.fits.fz	615.706	28.77	171.147
decam-flat-Y-4.fits.fz	615.711	27.69	177.875
decam-flat-Y-5.fits.fz	615.733	27.67	178.006
decam-flat-Y-6.fits.fz	615.717	21.26	231.663
decam-flat-Y-7.fits.fz	615.722	27.02	182.286
decam-flat-Y-8.fits.fz	615.706	27.16	181.312
decam-flat-Y-9.fits.fz	615.722	25.57	197.394
decam-flat-g-0.fits.fz	630.997	28.05	179.891
decam-flat-g-1.fits.fz	630.964	28.67	176.052
decam-flat-g-2.fits.fz	630.972	27.63	182.671
decam-flat-g-3.fits.fz	630.961	26.28	192.004
decam-flat-g-4.fits.fz	630.958	27.18	185.699
decam-flat-g-5.fits.fz	630.953	26.08	193.473
decam-flat-g-6.fits.fz	630.961	30.88	163.438
decam-flat-g-7.fits.fz	630.969	28.17	179.126
decam-flat-g-8.fits.fz	630.947	28.79	175.319
decam-flat-g-9.fits.fz	630.972	28.63	172.798
decam-flat-i-0.fits.fz	618.531	28.10	176.054
decam-flat-i-1.fits.fz	618.446	30.41	162.656
decam-flat-i-2.fits.fz	618.504	30.94	159.877
decam-flat-i-3.fits.fz	618.507	28.68	172.483
decam-flat-i-4.fits.fz	618.523	29.59	167.208
decam-flat-i-5.fits.fz	618.509	27.98	176.787
decam-flat-i-6.fits.fz	618.498	28.63	172.815
decam-flat-i-7.fits.fz	618.518	29.90	165.435
decam-flat-i-8.fits.fz	618.496	27.77	178.158
decam-flat-i-9.fits.fz	618.529	28.02	177.252
decam-flat-r-0.fits.fz	620.947	30.49	162.881

decam-flat-r-1.fits.fz	620.925	30.24	164.249
decam-flat-r-2.fits.fz	620.963	31.81	156.130
decam-flat-r-3.fits.fz	620.908	28.74	172.819
decam-flat-r-4.fits.fz	620.939	28.19	176.164
decam-flat-r-5.fits.fz	620.944	28.54	174.039
decam-flat-r-6.fits.fz	620.947	29.19	170.136
decam-flat-r-7.fits.fz	620.908	30.73	161.597
decam-flat-r-8.fits.fz	620.947	29.60	167.776
decam-flat-r-9.fits.fz	620.928	27.11	181.733
decam-flat-z-0.fits.fz	616.042	31.07	158.607
decam-flat-z-1.fits.fz	616.033	28.34	173.845
decam-flat-z-2.fits.fz	616.009	30.28	162.720
decam-flat-z-3.fits.fz	616.022	30.43	161.942
decam-flat-z-4.fits.fz	616.028	30.52	161.428
decam-flat-z-5.fits.fz	616.039	28.43	173.318
decam-flat-z-6.fits.fz	616.050	26.84	183.588
decam-flat-z-7.fits.fz	616.028	30.61	160.967
decam-flat-z-8.fits.fz	616.028	28.64	172.040
decam-flat-z-9.fits.fz	616.014	30.30	162.554

Appendix II

Table of sample transfer times achieved using differing numbers of transfer threads. Using more threads means that each thread is transferring a smaller portion of the file; ideally, processing on other threads compensates for any inefficiency in network throughput. In reality, context-switching in the CPU to service multiple threads can begin to degrade performance after a certain point.

Test file: decam-25-0-i-1.fits.fz
Size: 563 MB

	Time (sec)	Speed (Mbs)	Speed (MBs)
20 Threads:	202.621	21.79	2.72
	190.039	23.23	2.90
	220.723	20.00	2.50
30 Threads:	188.692	23.40	2.92
	193.385	22.83	2.85
	173.576	25.43	3.18
40 Threads:	149.760	29.48	3.68
	151.956	29.05	3.63
	159.320	27.71	3.46
50 Threads:	148.440	29.74	3.72
	144.892	30.47	3.81
	134.641	32.79	4.10
60 Threads:	119.506	36.94	4.62
	141.142	31.28	3.91
	142.453	30.99	3.87
70 Threads:	141.445	31.21	3.90
	136.911	32.25	4.03
	133.654	33.03	4.13
80 Threads:	120.730	36.57	4.57
	131.313	33.62	4.20
	107.280	41.15	5.14
90 Threads:	126.067	35.02	4.38
	132.980	33.20	4.15
	146.147	30.21	3.78
96 Threads:	128.475	34.36	4.30
	128.117	34.46	4.31
	124.487	35.46	4.43