STK 9840 Tape Drives in an SGI/Cray Environment

Installation - Performance - Stability

Ulrich Detert
U.Detert@fz-juelich.de

Institute for Applied Mathematics
Research Center Jülich

Abstract

In February 1999 StorageTek ® (STK) announced the new tapes and tape drives 9840 (unofficially called "Eagle"). The following paper reports on experiences gained during the installation and early usage phase of these tapes in an SGI/Cray environment. Specific emphasis is put on performance and stability in this report. Some questions concerning DMF configuration and tuning are also discussed.

System Configuration in Jülich

The SGI/Cray systems at the Research Center in Jülich are configured around a J90se system operating as a file server for all SGI/Cray systems (see Fig. 1). The file server hosts the home file systems for all users and a specific file system "/arch" used for automatic data migration onto tapes. The tape hardware comprises formerly two - now three - STK Automatic Cartridge System silos (ACS) (two 4410, one 9310). SGI/Cray's Data Migration Facility (DMF) is used for automatic migration of data to and from tapes.

The two 4410 ACS systems are equipped with 16 tape drives 4490 (Silverton) and are connected to the J90 via BMX channels. The newer 9310 ACS is connected via Fast/Wide SCSI-2 and is currently equipped with four tape drives 9840 (Eagle). Additionally, four tape drives 9490 (Timberline) were temporarily used as a migration path from Silverton to Eagle.

For a selection of 9840 tape drive specifications see Tab. 1.

Capacity	20 GB
Transfer rates	10 MB/s head-to-tape 18 MB/s ESCON 40 MB/s Wide-Ultra-SCSI 20 MB/s Wide SCSI-2
Compression	enhanced LZ-1, 4:1 ratio
Data buffer	8 MB per drive
Block size	256 KB (max.)
Search time	16 sec. (avg.)
Rewind time	16 sec. (max.)

Tab. 1 STK 9840 Tape Specifications

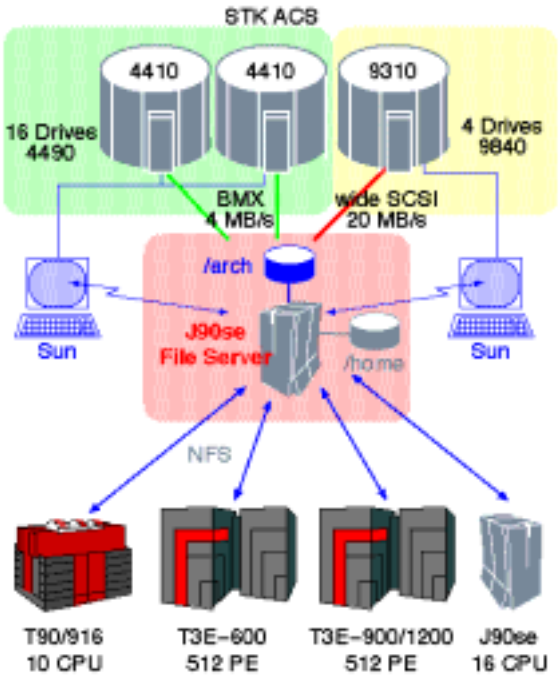

Fig. 1 System Configuration in Jülich

Migration from Silverton to Eagle

The Migration from Silverton tapes to a combination of Silverton and Eagle tapes was done in several steps:

1. Installation of a third silo with 4 9490 Timberline drives and 1000 3490E tape cartridges (interim solution until the availability of the Eagle tapes).

2. Reconfigure DMF such that files larger than 250 MB are migrated to Timberline tapes (use two Media Specific Processes (MSP) and select them depending on file size).
3. Add 4 Eagle drives 9840 to third silo.
4. Reconfigure DMF such that files larger than 100 MB are migrated to Eagle tapes (use three MSPs, where the 2nd MSP is used for recalling existing Timberline data only).
5. Move remaining data from Timberline to Eagle tapes.
6. Delete the 2nd MSP from DMF.
7. Remove Timberline drives and tapes.

Fig. 2 shows the distribution of tape data among the involved DMF MSPs during the migration phase. The lower growth rate of migrated data during the intermediate phase is due to a restrictive file quota policy.

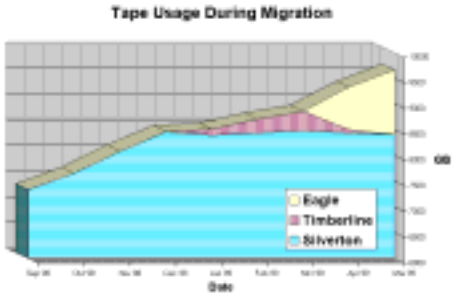

Fig. 2 Data Volume on DMF MSPs

Performance

The overall performance of tape hardware is influenced by several factors:

- head-to-tape data transfer rate of the drive
- transfer rate of the I/O channel
- maximum block size of the drive
- size and number of files on tape
- properties of compression hardware and software.

For almost all of these items the 9840 tape drives are superior to both 4490 and 9490. Tab. 2 compares data transfer rates for all three drives. Note that with compression switched on, the effective data transfer rate for Eagle drives is higher than the theoretical head-to-tape peak rate. This results from the fact that the data used for the test could easily be compressed by a factor of 12. Thus, the actually achieved transfer rate was more determined by the channel bandwidth and the speed of the compression algorithm than by the maximum head-to-tape transfer rate. In all tests only one file was written onto tape. The file size was 285 MB and the block size 64 KB for 4490 and 9490, and 256 KB for 9840, respectively.

		4490 Silverton	9490 Timberline	9840 Eagle
uncom-pressed	rd	2.6	4.6 - 4.8	8.1 - 8.2
	wr	2.2 - 2.5	3.8 - 4.0	7.8
com-pressed	rd	3.1 - 3.5	6.3 - 6.6	11.3 - 11.6
	wr	2.9 - 3.3	4.7 - 5.1	11.6 - 12.3

Tab. 2 Data Transfer Rates (MB/s)

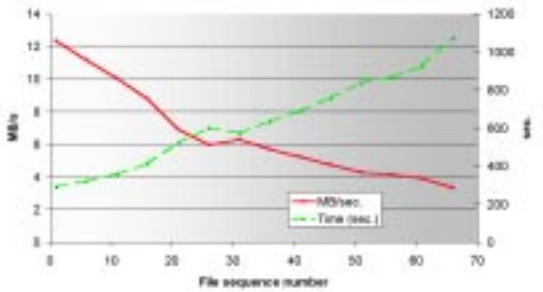

Fig. 3 Transfer Rate Depending on File Position

If multiple files are stored on tape, the tape drive performance is not only determined by the data transfer rate in streaming mode, but also by search and positioning times. Fig. 3 shows the elapsed time and effective data transfer rate for a write operation on an Eagle tape depending on the file sequence number. According to Tab. 1 the average search time is 16 sec. This time matches well with the results shown in Fig. 3. The transfer rate drops from over 12 MB/s for the first file on tape to below 4 MB/s for the 66th file. Adversely, the elapsed time raises from 292 sec. for the first file to 1074 sec. for the last. (Note that using the *-Q* option of the Unicos *tpmnt* command does not improve the performance of this test case.)

DMF

The results from Fig. 3 imply that it is reasonable not to have many small files on 9840 tapes. At Research Center Jülich this has been achieved by:

- Configuring DMF such that 9840 tapes are used for files larger than 100 MB, only.
- Using a zone size of 1 GB for Eagle tapes. (Within DMF zones are handled like tape files and multiple user data files are combined into zones.) Thus, there are typically 20 zones per tape and 10 files per zone, at maximum.

Due to the large zone size, no DMF caching is used for Eagle tapes in Jülich. Instead, data that is to be migrated is moved directly from disk to tape (via named pipes controlled by DMF). Tab. 3 and Tab. 4 show performance data extracted from DMF statistics. Important results are:

1. The average number of mounts per tape increased only under-proportionally from Silverton to Eagle.
2. The average position time did not increase.
3. The average transfer rate increased by a factor of almost 4 for reads and by a factor of 7 for writes from Silverton to Eagle.

Note that the data in Tab. 3 and Tab. 4 are only very preliminary, since the statistics taken for Eagle tapes include significant amounts of data movement due to tape media inspection and other housekeeping activities.

Migrate	Silverton Nov 98	Eagle Apr 99
tapes touched	805	77
requests	173720	5176
mounts / month	2660	519
requests / mount	65	10
data stored (GB)	854	1504
avg. transf. rate	1.0	7.2
mounts / tape	3.3	6.7

Tab. 3 DMF Statistics (Migrate)

Recall	Silverton Nov 98	Eagle Apr 99
tapes touched	3806	106
requests	125023	5322
mounts / month	15971	1854
requests / mount	7.8	2.9
data recall. (GB)	1392	2362
avg. transf. rate	1.8	6.75
position time (s.)	60	58
mount time (s.)	52	48
mounts / tape	4.2	17.5

Tab. 4 DMF Statistics (Recall)

Data Compression

The new 9840 tape drives are equipped with a data compression engine located in the drive itself (not in the controller). Data compression is done "on the fly" during data transmission to or from the tape. As already seen in Tab. 2 the compression of data can significantly increase the data transfer rate due to the reduced amount of data that has to pass the drive's read/write head. Tab. 5 shows DMF statistics data that compares data compression for Silverton and Eagle tapes. It should be noted that the compression ratio found in regular production data is typically below the maximum compressions rate. This is in part due to the fact that most production data on the Cray systems in Jülich are of "binary" nature (as compared to ASCII data) and in part to the fact that many users use data compression functions by themselves in order to reduce the amount of data under quota limitations.

	Silverton (MB)	Eagle (MB)
nominal capacity	800	20000
max. data on tape	2200	30516
avg. max. on tape	1080	22118
avg. data on tape	769	18118
max. # of files	26949	323
avg. # of files	220	86

Tab. 5 DMF Compression Statistics

The best compression results for Eagle tapes could be achieved with a "synthetic" test case. A file of size 3.6 GB could be written to tape in 67 copies, which gives a total of 244 GB on one tape. Compared with the nominal tape capacity of 20 GB, this gives a compression ratio of 12.2. The total time for writing this tape was 12h 40m which yields an average transfer rate of 5.5 MB/s (including mount and search times). For comparison: the UNIX command *compress* required 2h 19m for the compression of just one file (on a J90se) and achieved a compression ratio of 11.3.

Stability

The installation of the new ACS silo and the Timberline and Eagle drives was performed by STK without any technical problem. The Eagle tape drives were available on schedule and could be taken into production soon after delivery. The first week of production (with reduced workload on the Eagle tape drives) showed no serious problems. With increasing workload, however, tape media errors showed up in the error logs with increasing frequency.

One source of bad tape media could early be identified by IMATION (STK's supplier of tape media). 12 (out of a total of 500) tapes belonging to a bad production series were replaced by new tapes.

Another possible source of (maybe falsely diagnosed) tape media errors seems to be the microcode in the tape drives. The release currently used in Jülich is 1.23.105, the new release 1.23.202 is in beta-testing. It is expected that this release will improve the stability of the 9840 tape drives significantly.

Conclusion

With respect to capacity and performance, the new STK 9840 tape drives fully meet the expectations. The installation of the hardware and the configuration of the tape software and the DMF is not significantly different from previous tape drives. With respect to stability, the Eagle tapes still seem to be in there infancy. It is expected, however, that this situation is improved soon, since the problem areas are clearly identified.