Some Parameters in the design of Streaming Tape Drives

V. Chinnaswamy
SHR1-4/F2
Quantum Corporation
333, South St. , Shrewsbury, MA 01545
Fax: 508-770-2869
Tel: 508-770-2441
Email: Av.chinnaswamy@quantum.com

Abstract:

Several parameters play an important role in the design of tape subsystems. This paper analyzes some of these parameters and their relationships in determining the capacity and throughput of the tape subsystem. The pros and cons of trade offs among these parameters are also offered in choosing the right values for these parameters.

1. Introduction

There has been an enormous growth in the amount of data storage in computers in recent years. The explosion in the amount of data is due to the graphical and multimedia nature of data caused by the growth in internet. The value of data is becoming so important that the survival of many businesses depend on the preservation and integrity of their data. The vulnerability of data loss due to equipment malfunction, human error or natural catastrophes have made backup and archival of data very important.

The massive amount of data and the shrinking time to backup have made capacity and throughput of tape drives very important. Virtual storage, Hierarchical Storage Management, near-line and/or on-line storage of data have also made tape drives, their throughput and capacity more important. This paper discusses several parameters that effect the throughput and capacity of tape drives.

There are two types of tape drives predominantly in the market today. One is serpentine and the other is helical scan. DLT™ tape drives are serpentine and Exabyte or Sony AIT drives are helical scan. Although much of the discussion in this paper applies to both types, the focus here is only on serpentine drives.

The performance of a streaming tape drive depends on several parameters and their relationships. Some of the parameters that determine the capacity and throughput are:

- Tape length
- Tape speed
- Bit density
- Total number of tracks
- Number of tracks written simultaneously
- Formatting overhead

- Host input rate
- Tape drive transfer rate
- Compression ratio
- Cache size of the drive
- Tape reposition time
- Tape streaming time

Almost all the tape drives nowadays are streaming tape drives and they use some kind of compression algorithm. A streaming tape drive consists of a controller and a drive. During backup, the controller receives data from the CPU, formats and compresses the data and sends it to the drive. If the data transfer rate from the disk, CPU and the interconnecting busses (hereinafter to be referred as host transfer) equals or exceeds the tape drive transfer rate, the drive will be streaming. If the host transfer rate equals or exceeds the specified maximum drive transfer rate, the effective throughput of the drive will be the specified maximum throughput from the drive. If the host transfer rate falls below the transfer rate of the tape drive, a condition known as "data underrun" occurs and the drive goes into reposition. To receive data from the host when the drive is in reposition, some amount of buffering or cache is built into the controller. Cache helps to compensate for the differences in disk, cpu and tape transfer rates and keep the tape drive streaming as much as possible. Once the drive goes into reposition, the controller can receive data, but it cannot supply data to the tape drive until the drive completes its reposition. This reposition time may vary from several milliseconds to a few seconds depending on the technology of the drive. Since the differences between different host transfer rates and the tape drive transfer rates can vary so much, it is a difficult decision to balance the size of cache, reposition time, host transfer rates, drive transfer rates, cost and architecture of the drive.

The loss in throughput due to disparities in the host and drive transfer rates can be prevented by doing one or more of the following:

- Increase cache size
- Reduce reposition time
- Start writing fillers
- Use cache watermark
- Turn compression on and off
- Change the speed of the tape drive

This paper analyzes the relationships of the above variables, and present some solutions with a discussion of the pros and cons of these techniques.

Section 2 develops several mathematical models and formulas. Section 3 summarizes the conclusions from the analysis and some future directions.

2. Mathematical Models

Let

- L denote the length of tape in feet,
- d, bit density per inch per track,
- N, total number of tracks,
- n, number of tracks written simultaneously,

- S, the tape speed under write head while writing or reading,
- O, Formatting and other overhead,
- λ, the host transfer rate,
- μ, the drive transfer rate,
- C, the cache size,
- c, the compression ratio,
- t_r , the reposition time,
- t_s, the streaming time,
- a, acceleration or deceleration of the tape speed.

2.1 Capacity of cartridge and Throughput

The performance of a tape drive can be divided into two parts: capacity of cartridge or cassette and throughput. Capacity of cartridge depends on three parameters:

- Tape length
- Total number of tracks
- Bit density/inch/track

The throughput of the drive depends on:

- Tape speed
- Number of tracks written simultaneously
- Bit density/inch/track

The capacity of cartridge can be increased by increasing any number of the parameters mentioned. In the same manner, throughput can be increased by increasing tape speed, number of tracks written at the same time and/or bit density. The bit density is the common factor in both. By increasing bit density per inch per track, both capacity and throughput can be increased. Capacity of cartridge can be increased by increasing the length of tape and the total number of tracks without affecting throughput. In the same manner, throughput can be increased by increasing the tape speed and/or the number of tracks written simultaneously without affecting the capacity of cartridge.

The capacity of a cartridge is determined by L, d, N and O. Overhead is due to error correction blocks, CRCs, EDCs (error detection blocks), preample, postamble, etc. This overhead may account for about 25 percent or more. The capacity of the cartridge is given by,

$$\text{Capacity} = \frac{12 \cdot L \cdot N \cdot d}{8 \cdot 10^9 \cdot O} \text{ GB.} \qquad (1)$$

This is the user perceived capacity. Users are not interested in the overhead. In a serpentine recording technology such as Quantum's DLT™ tape drive DLT7000, the number of tracks per half inch is 208, density is 86000 bits per inch and length of tape is 1800 feet. Products to be released soon will have densities in the range of 100K bits per inch. Some of the products to be announced next year will have native capacities of 100 GB. In the next two to three years, these numbers are projected to increase enormously with new technologies such as MR heads, PRML, Optical Servo, etc. Some lab tests indicate that the capacity of cartridges could reach one terabyte with 2 : 1 compression in about two to three years.

Tape drives are designed to be backwards compatible. The present generation of tape drives should read the past generation cartridges. So, it is not possible to change the size of a cartridge from one generation to the next. When the length of tape is increased, it has to fit in the old cartridge. So the thickness of tape has to go down. Tapes are made up of three or four layers of coatings. The thickness of these coatings has to be reduced. This poses several problems. Increasing the bit density poses several read and write signal problems. Increasing the total number of tracks creates head and track alignment problems. It is not the purpose of this paper to discuss these details.

The throughput of the tape drive is given by

$$\text{Throughput} = \frac{S \cdot n \cdot d}{8 \cdot 10^6 \cdot O} \quad \text{MB/s} \qquad (2)$$

Again in Quantum’s DLT7000, the speed of tape is 160 inches/sec. and four tracks are written simultaneously. By increasing the number of tracks written simultaneously, increasing the bit density, and/or the speed of the tape, the throughput can be increased. Future generations of tape drives will increase any or all of these parameters. By just doubling the number of tracks written and increasing the bit density about 50 percent, throughput can be increased three fold. In a few years, tape drives with capacities of terabytes per cartridge and throughputs in the range of 100 MB/s would be available.

2.2 Reposition Time

The reposition time is a function of the speed, S, the acceleration and the decelaration, a. Suppose the tape is moving at a speed S and it runs out of data at point A. See Fig. 1. The drive decelarates and comes to a stop at B. From B, it accelerates backwards until it reaches a speed of S at point C. From C, it decelerates backwards and comes to a stop at D. From D, it accelerates forward and reaches a speed of S at A. If data is available the drive will start writing from A. If data is not available, the drive stops at D. Normally, the acceleration and decelaration rates are the same. Using Newton’s law of motion

$$V = u + at$$

Where V is the final velocity, u the initial velocity, a the acceleration and t the time. If t is the time for going from A to B, it is given by

$$0 = S - at$$

$$\text{i.e.,} \quad t = \frac{S}{a}$$

The reposition time is then given by

$$t_r = \frac{4S}{a} \qquad (3)$$

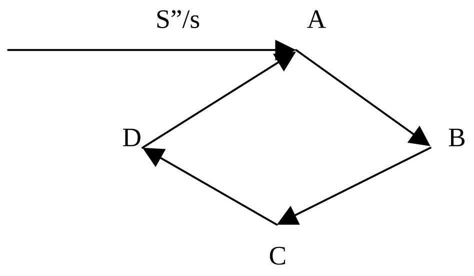

Fig. 1

In actual practice, it could be a little longer than this because of some time delays during stop and start. Also, a few blocks known as margin blocks are written when the drive runs out of data. It is not necessary to go into those details for our purpose here. The point is that the tape drive transfer rate and the reposition times depend on the speed of the tape drive. When the tape speed goes down, the reposition time goes down and transfer rate goes down. For a given host transfer rate, a reduced transfer rate of the tape drive will make the drive stream longer. Since the reposition time is also short, the drive will be ready to receive data before cache is filled up.

2.3 Cache Analysis

Six different relationships can be identified with the host transfer rate, drive rate, reposition time, cache and compression ratio.

Case i: $\lambda > \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} < \frac{C}{t_r}$

Without compression, the host transfer rate exceeds the drive rate and the maximum throughput one can get is the spec rating of the tape drive. The tape drive is the bottleneck without compression.

With compression, the host rate is less than the drive rate. It is also such that it does not fill up cache during reposition. Hence, there is no loss in throughput. But there will be many repositions causing mechanical wear and tear which may lead to reduced MTBF. A solution to this problem is to use some cache watermark to reduce the number of repositions. Reducing the speed of the tape drive will also help. Cache watermark and reducing the speed of the drive will be discussed later.

Case ii: $\lambda > \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} > \frac{C}{t_r}$

In this case, without compression the drive rate is unable to keep up with the host rate. Without compression, the throughput will be the maximum throughput of the drive. The tape drive is the bottleneck without compression.

With compression, the host rate is such that it does fill up cache during reposition. Hence, there will be loss in throughput. This case will be discussed later in greater detail. Cache watermark is not helpful here since cache is filled up during reposition.

Case iii: $\lambda > \mu, \frac{\lambda}{c} > \mu$

In this case, host transfer before and after compression exceeds the drive transfer rate. Cache will be always full. The maximum specified throughput of the tape drive will be achieved. Since host transfer rate exceeds the tape drive transfer rate, the tape drive is the bottleneck.

Case iv: $\lambda < \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} < \frac{C}{t_r}$

Here, the host rate before and after compression is less than the drive rate. Cache does not get filled up during reposition. There is no loss in throughput. But there will be many repositions causing tape wear. A cache watermark or a reduction in the write/read speed will help reduce the number of repositions. Section 2.8 discusses the cache watermark. Speed variation will be discussed in Section 2.10.

Case v: $\lambda < \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} > \frac{C}{t_r}$

Here, the host rate before and after is less than the drive rate. But cache will be filled up during reposition. There will be loss in throughput. The loss in throughput can be avoided if the write speed of the tape drive can be reduced to an acceptable level. Cache watermark can also be used.

Case iv: $\lambda < \mu, \frac{\lambda}{c} > \mu$

Compression expands data. This case might happen if compressed data is compressed again or random numbers are used in the compression. This case not important to our discussion.

2.4 Loss in Throiughput due to Inadequate Cache or Reposition Time

Case ii: $\lambda > \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} > \frac{C}{t_r}$; Case v: $\lambda < \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} > \frac{C}{t_r}$

The only cases of interest where the effective throughput is less than the host transfer rate and the tape drive transfer rate are cases ii and v. It can be shown that the loss in throughput, T, is given by [1,2]

$$T = \begin{cases} \lambda & \text{if } t_r \dfrac{\lambda}{c} \leq C \\ \dfrac{c\mu}{1 + \dfrac{t_r(\mu - \frac{\lambda}{c})}{C}} & \text{if } t_r \dfrac{\lambda}{c} > C \end{cases} \tag{4}$$

Wthout compression, the above formula changes to

$$T = \begin{cases} \lambda & \text{if } \lambda t_r \leq C \\ \dfrac{\mu}{1 + \dfrac{\mu}{\dfrac{t_r(\mu - \lambda)}{C}}} & \text{if } \lambda t_r > C \end{cases} \tag{5}$$

The loss in throughput, when there is no compression, is given by

$$L_t = \lambda - \frac{\mu}{1 + \dfrac{t_r(\mu - \lambda)}{C}} \tag{6}$$

Differentiating with respect to λ, we can show that the maximum loss occurs when

$$\lambda = (\frac{C}{t_r} + \mu) - \sqrt{\frac{C * \mu}{t_r}} \tag{7}$$

Suppose, C = 4 MB, λ = 3 MB/s, μ = 4 MB/s, then loss in throughput will be

$$L_t = 3 - \frac{4}{1 + \dfrac{2(4 - 3)}{4}} = 3 - \frac{4}{1.5} = 3 – 2.67 = 0.33 \text{ MB/s}$$

The effective throughput will be 2.67 MB/s instead of the actual host rate 3 MB/s.

2.5 Increased Cache Size

Case ii: $\lambda > \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} > \frac{C}{t_r}$; Case v: $\lambda < \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} > \frac{C}{t_r}$

How do we determine the size of cache if the tape drive rate, μ , and reposition time, t_r are fixed and we want to achieve the maximum throughput for all possible compression ratios? $C = \mu t_r$ should be right cache size to get the best possible throughput when host transfer after compression is less than the drive transfer. For instance, if μ = 6 MB/s and t_r = 2 seconds, then C must be equal to 12 MB. With 12 MB cache, there will be no loss in throughput with compression ratios 2 : 1 or higher. If there are no bottlenecks in the compression chip or any other hardware or software, throughput can be at least 12 MB/s with files giving at least 2 : 1 compression ratios, higher throughputs possible with higher compression ratios.

2.6 Reduced Reposition Time

If C and μ are fixed, these two values determine the reposition time. When $t_r = \frac{C}{\mu}$, the drive will stream all the time or there will be no loss in throughput. For instance, if C = 8 MB and μ = 5 MB/s, then the reposition time should be at most 1.6 sec.

2.7 Streaming Fillers

Case $\lambda > \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} < \frac{C}{t_r}$

A streaming tape drive goes into reposition when there is no data in cache. The repositioning can be avoided if the tape drive is allowed to write fillers instead of real data until data is available for writing. This technique will prevent the loss in throughput, but there will be a loss in actual data capacity of the cartridge. Part of the cartridge now contains fillers instead of real data.

It can be shown that the percentage loss in capacity of the cartridge is given by

$$L_c = 1 - \frac{\lambda}{\mu} \qquad (8)$$

when there is no compression. If the drive rate is 6 MB/s and the host rate is 3 MB/s, fifty percent of the capacity of the cartridge will be lost due to fillers.

If there is compression, the loss is given by

$$L_c = 1 - \frac{\lambda}{c\mu} \qquad (9)$$

Fillers are not needed when the host transfer is such that the cache does not get filled up during reposition. Fillers also are not needed when host transfer exceeds the transfer rate of the tape drive. Fillers are helpful when the host transfer rate is within a certain range. But there is a penalty in the capacity of the cartridge.

2.8 Cache watermark: $\lambda > \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} < \frac{C}{t_r}$

If cache is not filled up during reposition either with or without compression, the drive will go into a number of repositions. To reduce the number of repositions, a cache watermark can be used. For instance, in Fig 2 say the data is filled up to level C3 during reposition. C1 is the full cache size.

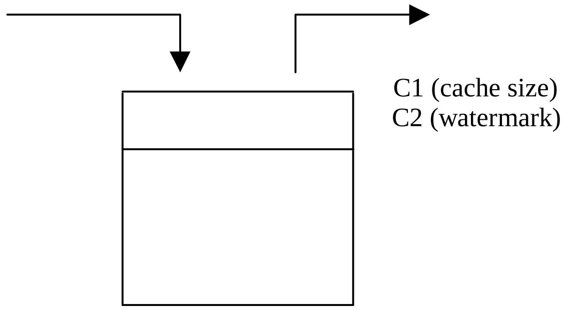

Fig. 2

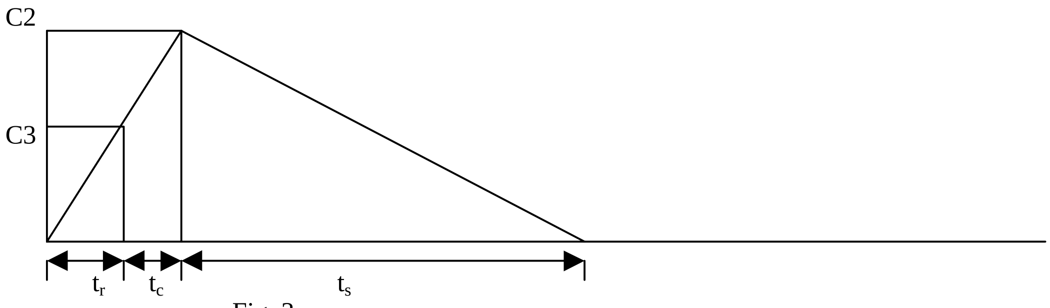

Fig. 3

Cache fills up to C3 during reposition. The drive waits for t_c time units after reposition until cache fills up to C2, then starts transferring data from cache. So, we have

$$\lambda (t_r + t_c + t_s) = \mu t_s$$

Solving for t_s, we get

$$t_s = \frac{\lambda (t_r + t_c)}{\mu - \lambda} \tag{10}$$

Throughput is given by

$$T = \frac{\mu t_s}{(t_r + t_c + t_s)}$$

Substituting the value of t_s, we get

$$T = \lambda$$

an obvious answer. Introducing cache watermark has not changed the effective throughput. It is an expected answer because the host transfer was not blocked. But the streaming time has increased since

$$t_s = \frac{\lambda\,(t_r + t_c)}{\mu - \lambda} > \frac{\lambda t_r}{\mu - \lambda} \tag{11}$$

The number of repositions will go down. It will reduce the mechanical wear and tear of the drive and would increase the MTBF of the drive.

2.9 Compression on and off:

Assume for a moment that the cache is full and compression on. When compression on, the host rate is less than the drive rate. So, cache will be depleted. When data in cache goes to the low watermark level as in Fig. 4, we know the drive will soon go into reposition and we will lose throughput. At this point, we turn compression off. When compression is off, the host rate is higher than the drive rate and data in cache increases. It will soon reach the high watermark. At this point, we turn compression on again. The data in cache goes up and down between the two watermark levels. Throughput is not lost. But capacity in the cartridge will go down because we do not compress the data all the time.

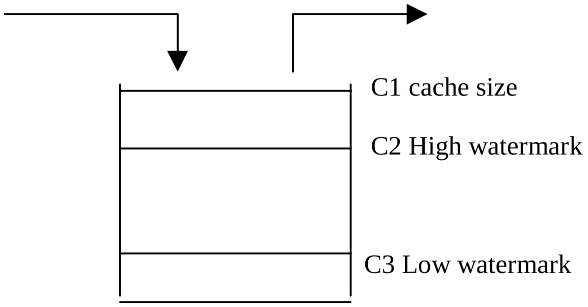

Fig. 4

2.10 Variable Speed: Case ii: $\lambda > \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} > \frac{C}{t_r}$; Case v: $\lambda < \mu, \frac{\lambda}{c} < \mu, \frac{\lambda}{c} > \frac{C}{t_r}$

If the host is unable to supply data at the spec rating of the tape drive, a “data underrun” condition would occur. During a restore operation, if the host is unable to receive data fast enough, “data overrun” condition occurs. Cache (also called buffer) is being used to compensate for the disparity between the transfer rate of the tape drive and the host transfer rate. Cache is often limited by cost, space and system architecture. When tape drive transfer rate or capacity is increased from one generation of tape drives to another by increasing the bit density or tape speed, changing the size of cache may not be a viable option. In such instances, adjusting the speed of the tape drive in synchronism with the host transfer rate is the best solution.

As pointed out earlier, the reposition time of the drive is a function of the tape drive speed. If the drive speed increases, the reposition time will increase. If the speed goes down, the reposition time will go down. When the host is unable to supply data fast enough to keep up with the speed of the drive, it can be shown that the throughput can be increased by reducing the speed of the drive.

During a backup operation, the host input may not remain constant. The rate would depend on the nature of files in the disk. A disk may contain small files and large files. In addition, there may be fragmentation of files. If small and fragmented files are read from disk, the host input rate would go down. It would increase when reading large unfragmented files. So, the drive transfer rate should be changed up and down in sync with the host transfer rate.

3. Conclusion

The analysis presented here gives several relationships with the parameters of interest in the design of tape subsystems. One area where further research may be needed is to synchronize dynamically on the fly the tape drive transfer rate with changes in the host transfer rate and compression ratios. Compression ratios may change depending on the incoming data pattern. Host transfer rates may change depending on the file sizes in the disk, fragmentation of files, other workloads in the system, etc. Some actual implementations change the transfer rate of the drive only during reposition or track turn around time. Further work is needed to change the speed on the fly instantly. This will reduce the need for large caches and the cost of the drive.

References:

[1] V. Chinnaswamy, Analysis of Cache for Streaming Tape Drive, Proc. Goddard Conference on Mass Storage Systems and Technologies, pages 299-310, Greenbelt, MD, Sep. 22, 1992.

[2] V. Chinnaswamy, Mathematical Models for the Design of a Streaming Tape Drive, Proc. of The IASTED International Conference on Modeling, Simulation and Optimization, Gold Coast, Australia, May 6-9, 1996.