STRIPED TAPE ARRAYS

Ann L. Drapeau
Randy H. Katz
Computer Science Division
University of California
Berkeley, California

ABSTRACT

A growing number of applications require high capacity, high throughput tertiary storage systems [1] [2]. We are investigating how data striping ideas apply to arrays of magnetic tape drives. Data striping increases throughput and reduces response time for large accesses to a storage system. Striped magnetic tape systems are particularly appealing because many inexpensive magnetic tape drives have low bandwidth; striping may offer dramatic performance improvements for these systems. There are several important issues in designing striped tape systems: the choice of tape drives and robots, whether to stripe within or between robots, and the choice of the best scheme for distributing data on cartridges. One of the most troublesome problems in striped tape arrays is the synchronization of transfers across tape drives. Another issue is how improved devices will affect the desirability of striping in the future. We present the results of simulations comparing the performance of striped tape systems to non-striped systems.

INTRODUCTION

Striping has been widely used in arrays of magnetic disk drives [3] [4] [5]. In striped disk arrays, a single file is striped or interleaved across several disk devices. Because a striped file can be accessed by several disks in parallel, the sustained bandwidth to the file is greater than in non-striped systems, where accesses to the file are restricted to a single device. As a result, latency is reduced for large accesses with long periods of data transfer.

Applying striping ideas to magnetic tape drive arrays is appealing for several reasons. A growing number of inexpensive tape technologies is available, but these tape drives provide low bandwidth, and potentially would benefit from the throughput advantages offered by striping. More expensive tape drives offer higher throughput, but could still make use of striping to meet the bandwidth requirements of demanding applications. One such application is the NASA Earth Observing System, which anticipates collecting and processing data at a sustained bandwidth of 100 MBytes/sec [1]. Many robot devices are now available for handling tape cartridges automatically, making it possible to treat large collections of tapes as if they are almost online. Finally, it is convenient to add redundancy information to striped storage systems to improve the reliability and availability of data stored in the array. Reliability concerns for high capacity tape systems include media and head lifetimes, as well as the occurrence of drive, robot and supporting hardware failures.

Table 1 compares several magnetic tape drives. The inexpensive helical scan drives (Exabyte 8mm and DAT 4mm) have high capacity but low bandwidth and long access times. Higher performance, higher capacity helical scan drives like the D-1 and D-2 have better bandwidth, but are very expensive and still suffer from long positioning times. The inexpensive serpentine drives (1/4") have only moderately improved bandwidth over the helical scan drives. And, the linear 3490 drives have fast positioning time and moderate bandwidth, but are low capacity. There is no clearly superior choice for a drive to be used in tape arrays.

Striping offers potential benefit for all these drives. For the low bandwidth drives, striping offers obvious advantages of much higher throughput to individual files than could be provided by single readers, and thus greatly reduced response time. Even the higher bandwidth drives can potentially be used in a striped configuration to multiply the available bandwidth and satisfy demanding applications.

We argue that data striping in magnetic tape arrays will improve performance for a range of workloads. First, we describe considerations in applying striping to magnetic tape arrays, including a discus-

Drive	Capacity Per Cartridge GBytes	Sustained Bandwidth MB/sec	Average Seek sec	Approximate Drive Cost $	Approximate Media Cost $/MByte
Exabyte EXB8500	5	0.5	40	$3,000	$0.008
DAT	1.3	0.18	20	$1,000	$0.025
Metrum 1/2"	14.5	2	45	$40,000	
19mm D-1	90	45	N/S	$300,000	$0.0012
19mm D-2	25	15	15	$150,000	
1/2" D-3	20	12	N/S	N/S	N/S
3490 1/2"	.40	6	N/S	$20,000	$0.025
1/4"	2	3	43	$1,000	$0.018

Table 1: *Compares the cartridge capacity, sustained bandwidth, average seek and approximate drive and media cost for different magnetic tape drives. N/S indicates not specified.*

sion on available tape drives and robots, types of striping, synchronization and configuration issues. We present simulation results showing the effect of striping on response time and throughput for two tape array configurations. Finally, we discuss our plans for future work.

TAPE STRIPING ISSUES

In this section, we discuss design choices and difficult issues for striped tape arrays. A tape array designer must choose tape drive and robot technologies, and must pick from among several striping options. The designer must also decide on the size of the interleave unit, a choice complicated by the long access times in magnetic tape arrays, and on the amount of buffering required to mask troublesome synchronization problems. This section concludes with a discussion of predicted device improvements in this decade, and how these changes are likely to affect striped tape systems.

Striping Options

There are two main options for striping in storage systems composed of some combination of magnetic tape drives and robotic tape libraries. Data may be striped within an individual robot, or across several robots.

Table 2 shows our classification of cartridge-handling tape robots, which is based primarily on cost and number of cartridges in the system. Table 3 describes an example of each type of robot. Large libraries generally contain hundreds or thousands of cartridges, several drives and one or two robot arms for picking and placing cartridges. The cartridges are often arranged in a rectangular array. Other "large library" configurations include a hexagonal "silo" with cartridges and readers along the walls [6], and a library consisting of several rotating cylindrical columns [7]. Usually these large libraries are quite expensive ($500,000 or more). They have the highest cost per reader, but often the lowest cost per MByte compared to other robots. Large libraries also take up the least amount of machine room floor space for a given capacity, giving them the highest ratio of MBytes/square foot, a practical consideration for many massive storage facilities. A disadvantage of large tape libraries is the low ratio of readers and robot arms to cartridges. In a heavily-loaded system, there is likely to be contention for both arms and readers.

There are less expensive, smaller, often slower tape libraries in the range of $50,000 which hold fewer cartridges than the large, expensive libraries. Also in the moderate price range are carousels, which cost about $30,000 and hold approximately 50 cartridges. The carousel rotates to position the cartridge over a drive, and a robot arm pushes the cartridge into the drive. In most cases, there are one or two drives per carousel.

Finally, the least expensive robotic device ($10,000 or less) is a stacker, which holds approximately 10 cartridges in a magazine and loads a single reader. Stackers generally have the lowest cost per reader but the highest cost per megabyte and the lowest MB/square foot compared to other robots. For a given capacity, a storage system composed of stackers would have the highest ratio of robot arms and readers to cartridges.

The most obvious application of striping is within a large robotic library. Striping this way is convenient, since it is easy to keep the tape cartridges that are

Type	No. Cartridges	No. Readers	No. Robot Arms	Cost
Large Library	100s to 1000s	several	one or two	high ($100,000+)
Small Library	around 100	several	one	moderate
Carousel	around 50	one or two	one per reader	moderate
Stacker	around 10	one	one (magazine or arm)	low (under $10,000)

Table 2: *Classification of storage robots.*

	Metrum RSS-600	Exabyte EXB-120	Spectra Logic 8mm Carousel	Exabyte EXB-10
Classification	Large Library	Small Library	Carousel	Stacker
Number drives	up to 5	4	1 or 2	1
Number cartridges	600	116	45	10
Number robot arms	1	1	1 or 2	1
Cartridge format/capacity	1/2" 14.5 GB	8mm 5 GB	8mm 5 GB	8mm 5 GB
Total capacity (GBytes)	over 6000	580	225	50
Approximate robot cost	$540,000 (2 drives)	$61,965	$27,500 (1 drive)	$8798
Robot cost/MB	$.09	$.10	$.12	$.17
Avg. robot access time (sec)	8	18	10	less than 20

Table 3: *Comparison of four available robotic devices: the Metrum RSS-6000 VLDS system, the Exabyte EX-120 Library, the Spectra Logic STL-8000H carousel, and the Exabyte EXB-10 Stacker. Prices indicated are list prices.*

logically connected in a stripe together physically when they are in a single physical enclosure. The disadvantages of using a single large library are the small ratios of readers and robot arms to cartridges. In a heavily-loaded system, there is likely to be contention for robot arms and readers that will delay cartridge switches and increase response times. The penalty will be more severe for striped systems, since they generate more cartridge switches.

Striping across independent robotic libraries has the advantage of allowing each robot arm to operate independently. Requests for cartridges in a stripe will not be serviced by the same robot arm. At low loads, this should reduce the number of requests queued on a particular robot arm, reducing latency for the operations. One of the disadvantages of striping between physically separate robots is the complexity of managing cartridges that are logically grouped into a stripe set, but which are stored in physically separate libraries. This administrative problem is alleviated if cartridges never leave the library, or if there is a standard procedure for moving stripe sets between the library and the shelf.

An interesting configuration for striping between robots is a system composed of inexpensive stackers. Such a system would have the highest proportion of arms and readers to cartridges, about 1:10 in each case, compared to 1:100s in the case of large libraries. A striped system composed of stackers would likely experience the least contention under heavy loads. The main disadvantage of such a system is the relatively high cost per megabyte. For example, for about the same price, one could purchase a Metrum RSS-600 robot or about 60 EXB-10 stackers. The stacker system would have 60 drives capable of 0.5 MBytes/sec each, compared to at most five drives in the library capable of 2 MBytes/sec each. But, the RSS-600 library would have about twice the storage capacity (6 TBytes vs. 3 TBytes).

When designing a striped system, the likelihood of contention is an important question. In traditional archival systems, contention has not been much of an issue, since there has not been much concurrent activity. When interactive systems with good response times are available, contention for resources will likely increase.

Access Time and Configuration Issues

The access time for a request includes time spent setting up the access as well as time to transfer data. Data striping can provide greater throughput for individual accesses, but does not change the access time characteristics of the devices and robots.

If an access is intended for a cartridge that is already loaded in a tape drive, the access time consists of time to position the heads and transfer data. For accesses that require a cartridge switch, we model access time as the sum of the times to rewind the drive to a place where an eject is allowed, the eject operation, the time for the robot to shelve the old cartridge and grab a new one, the drive load operation (which wraps the tape around the reels and reads format and servo information), and finally the data transfer time.

Table 4 shows our measurements for the tape drive components of access time for three drives: an Exabyte EXB8500 8mm drive, a WangDAT DAT drive and a Metrum 1/2" drive. Load and eject times are the means of twenty measurements each; variance was low in each case. In each case, the combination of a drive eject and load takes at least 30 seconds. Figures 1 and 2 show measured rewind and search behavior for the Exabyte EXB8500 drive; these measurements were made for tapes written entirely with 10 MByte files, so that filemarks (48 KBytes each) are a small fraction of the total tape. The graphs show one set of measurements; the tests were run several times, and little variance was exhibited. We observe that rewind and search times scale linearly with the number of bytes passed over, after a constant startup time. Table 4 shows the startup time and linear rewind and search rates for each of the drives.

Depending on the speed of the robot, its contribution to tape access time in operations requiring a cartridge switch ranges from about 5 to 50 seconds. Table 5 shows measurements of grab time for an Exabyte EXB-120 robot, a simple rectangular array of 116 cartridges and four tape drives. We also measured robot arm movement time, and found that it varied between 1 and 2 seconds. Thus, arm movement is a relatively insignificant component of overall access time, while pick time is significant for this robot.

Average access time for the three devices measured here is several minutes. For example, Table 6 shows

Time to pick cartridge from drive	19.2 sec
Time to put cartridge into drive	21.4 sec

Table 5: *Measured times for robot to grab a cartridge from a drive and push a cartridge into a drive for the EXB-120 robot system.*

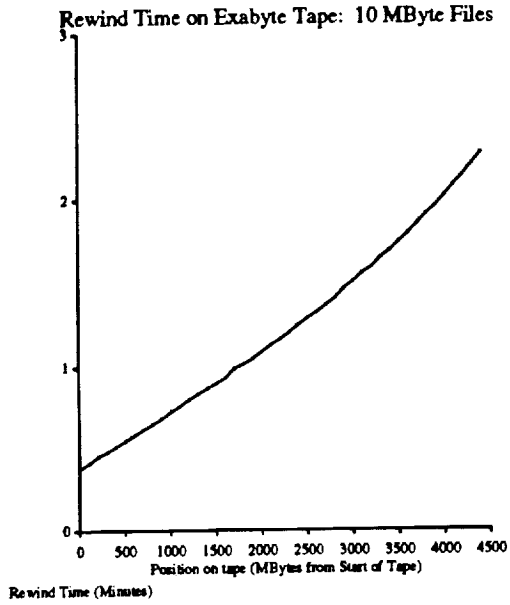

Figure 1: *Measured rewind behavior for Exabyte EXB8500 drive. Tape written entirely with 10 MByte files.*

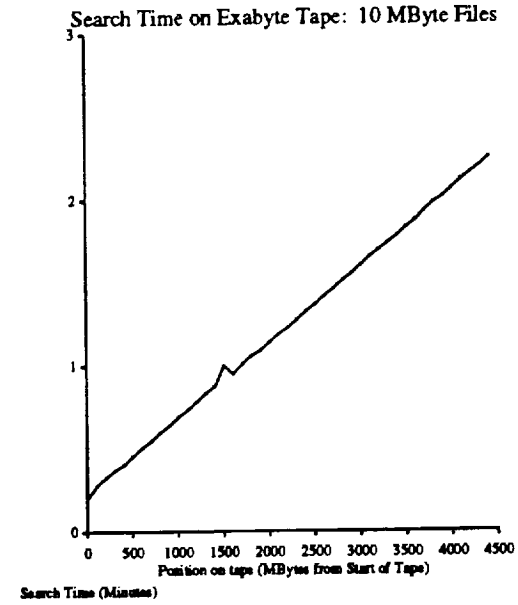

Figure 2: *Measured search behavior for Exabyte EXB8500 drive. Tape written entirely with 10 MByte files.*

Operation	4mm DAT	8mm Exabyte	0.5" Metrum
Mean drive load time (sec)	16	35.4	28.3
Mean drive eject time (sec)	17.3	16.5	3.8
Constant rewind startup time (sec)	15.5	23	15
Rewind rate (MB/sec)	23.1	42.0	350
Constant search startup time (sec)	8	12.5	28
Search rate (MB/sec)	23.7	36.2	115
Read transfer rate (MB/sec)	0.17	0.47	1.2
Write transfer rate (MB/sec)	0.17	0.48	1.2

Table 4: *Measurements of 4mm, 8mm and 0.5" helical scan magnetic tape drives.*

Operation	Time (sec)
Rewind time (1/2 tape)	75
Eject time	17
Robot unload	21
Robot load	22
Device load	35
Search (1/2 tape)	84
Total	254

Table 6: *Components of cartridge switch time for Exabyte EX120 Robot.*

that the cartridge switch time for the EXB-120 robot (not including data transfer) takes four minutes. Even the expensive, high-bandwidth drives (D-1 and D-2) and robots, with faster robot arms and drive mechanics, may take up to a minute for a cartridge rewind, switch and positioning.

Because the penalty for switching cartridges is so severe, and striped systems generate additional cartridge switches compared to non-striped systems, striped systems must be carefully designed so that the penalties of cartridge switching are offset by the response time gains striping offers. One obvious rule-of-thumb is that the "stripe width" or number of cartridges across which data are interleaved should not exceed the number of readers in the system. For example, if a library contains four robots, but the files are interleaved across eight cartridges at a time, then many accesses would require two cartridge loads per reader. Because not all the cartridges in a stripe are loaded at once, this striped system gets less benefit from locality of reference on subsequent accesses, since cartridge switches will still be required. Such a system will likely have poor performance.

Another configuration issue is the interleave factor, or the amount of data written on one cartridge before switching to another cartridge in the stripe. If the interleave factor is too small, requests of moderate size will require access to several cartridges. As a result, cartridge switch penalties won't be offset by the throughput gains normally offered by striping, since the amount transferred from each cartridge is relatively small. If the interleave unit is too large, then most requests will be limited to one or a small number of cartridges, and the throughput benefits of striping will be lost. We plan to identify optimum interleave units for a variety of workloads.

Synchronization Issues

Synchronization is one of the most troublesome issues for striped tape. It is impossible to operate several tape drives involved in a striped access in true synchronization. After writing data to a tape, the drive immediately reads back the data to be sure that the write completed correctly. This read-after-write checking typically encounters a high rate of errors [8]. The drive responds to errors by retrying write operations on subsequent regions of the tape until data is written correctly. As a result, particular data blocks do not reside at known locations on the tape. Synchronization like that in disk arrays, based on known sector positions, is impossible. Large flawed sections of the tape that are unwritable will cause many retry attempts, delaying both write and read operations on those sections.

Another synchronization problem arises from competition among requests for use of robot arms. If a single robot arm is loading several drives sequentially, or if there is contention for the use of several robot arms, the load time for the tape drives involved in a stripe access will vary.

If the array is configured using large block interleaving (a RAID level 5 configuration [9]), another synchronization problem can occur. Requests smaller than the interleave factor can be satisfied by ac-

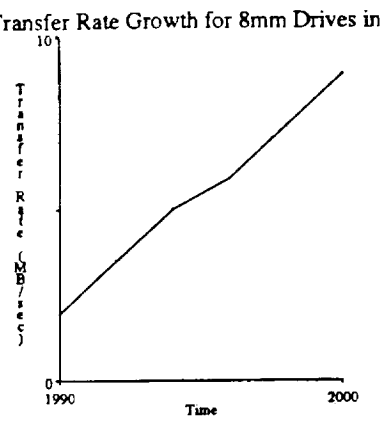

Figure 3: *Predictions for bandwidth improvements for 8mm tape drives in this decade. Source: Harry Hinz, Exabyte Corp.*

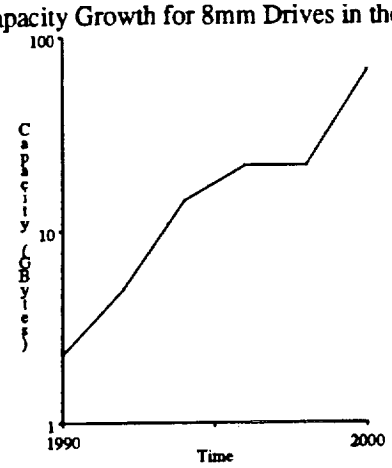

Figure 4: *Predictions for capacity improvements for 8mm tape drives in this decade. Source: Harry Hinz, Exabyte Corp.*

cessing a single cartridge. The drives may act independently on several such requests, loading and accessing unrelated cartridges. Later, a larger request requiring access to several logically-grouped cartridges may see widely different latencies on the separate cartridge accesses, since each new access will require a cartridge reload or reposition operation.

Buffering can be used to coordinate unsynchronized tape operations. One of the issues not yet explored in striped tape systems is the amount of buffer space that will be required to maintain reasonable performance. Buffer space requirements increase with the amount of concurrency in the system, and with size of the interleave unit. Available buffer space is also important to keeping tape drives streaming, so that they perform at their maximum possible throughput.

Future Devices

One interesting question is how future tape drives and robots will affect the need for and effectiveness of striping. Figures 3 and 4 show predicted improvements in cartridge capacity and bandwidth for 8mm tape drives in this decade [10]. It shows both capacity and throughput doubling approximately every two years, reaching 67 GBytes per cartridge and 6 MBytes/sec by the end of the decade, compared to 5 GBytes per cartridge and 0.5 MBytes/sec today. Improvements required to reach these goals include increasing track density, decreasing track width and pitch, reducing tape thickness and increasing rotor speed.

Besides data transfer rate, other components of access time should also improve. Several drive manufacturers are reducing rewind and search times by implementing periodic zones on the tapes where eject and load operations operations are allowed, rather than requiring that a tape must always be ejected and loaded at the start of the tape [7]. As access times become more of a concern, it is likely that mechanical operations like load, eject and robot grab and insert will become faster. Robot arms are being improved so that they are lighter and faster [11].

All these drive and robot improvements should make striping more attractive. No matter what the throughput available on readers, striping is a valid technique for multiplying throughput seen by a single file. And, since the main penalty of striped versus non-striped systems is that suffered when additional cartridge switch operations are required, any improvements that reduce cartridge switch penalties will only improve the performance of striped systems.

PERFORMANCE SIMULATIONS

We have written an event-driven tape array simulator. The models for tape device and robot performance used in the simulator are based on the device and robot measurements described in the previous sections. We currently simulate tape arrays as closed systems, keeping the number of outstanding requests constant. In response to input files that contain device and robot specifications, striping

configuration (interleave unit, stripe group width, redundancy scheme), request size and position distributions, the simulator calculates the mean response time and sustained bandwidth of the array. The following results compare striped and non-striped performance for two tape arrays: a small tape library, the EXB120, and a large library, similar to the Ampex DST600.

EXB-120 Library

The first array we simulated is the EXB-120 library from Exabyte. This library has a single robot arm, four Exabyte EXB-8500 readers, and 116 cartridges. Performance of the Exabyte robots and readers has already been described. The striping configuration for these simulations is rotated single-bit parity with an interleave unit of 100 MBytes. Files are striped over groups of three data cartridges plus one parity cartridge. The workload for these systems is 25% write operations, which require accesses to the parity cartridges, and 75% read operations, which do not require parity accesses. These simulations do not include error recovery operations. Our simulations assume that enough buffer space is present in the system that the drives always operate in streaming transfer mode. This simplifying assumption will be removed in future simulations.

Figures 5 and 6 compare the response times and bandwidth of striped and non-striped accesses for average request sizes up to 1 GByte. When there is a single request in the system (concurrency = 1), striping improves response time for requests over 80 MBytes in size. It is not surprising that striping has little effect on requests of average size less than 80 MBytes, since with an interleave unit of 100 MBytes, smaller requests are usually handled by a single cartridge access. Thus, for requests smaller than the interleave unit, the performance of striped and non-striped systems will look very similar. Striped systems will see some performance penalty for the extra accesses required to write redundancy information. When average request size reaches 1 GByte, response time for striping with concurrency=1 is about half the response time for a non-striped system: around 20 minutes compared to 40 minutes per access.

When there are two requests in the system, striping improves performance for requests of average size over 200 MBytes. At a request size of 1 GByte, striped system response time is about 25% better than non-striped. For higher concurrencies, striping hurts response time versus non-striped systems.

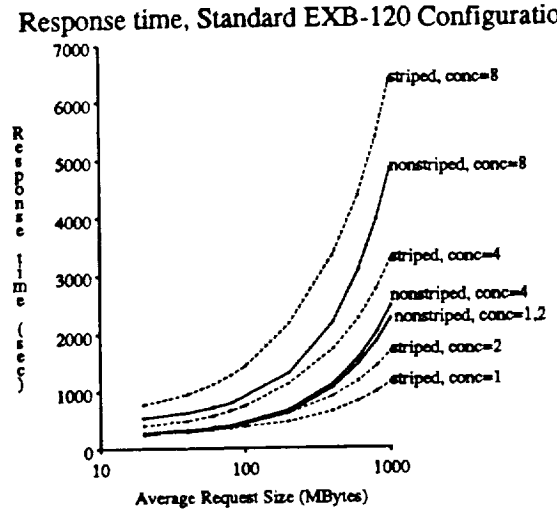

Figure 5: *Response time comparison of striped system with interleave unit of 100 MBytes to a non-striped system, for concurrencies of 1, 2, 4 and 8. Standard Exabyte configuration: four EXB-8500 drives, 116 cartridges, one robot arm. Nonstriped performance curves for concurrency of 1 and 2 overlap, and are represented with a single label.*

Consider a 300 MByte request. Usually, in a striped system, this request will be spread over three data cartridges and a parity cartridge. 100 MBytes will be read or written from each cartridge. At a transfer rate of 0.5 MBytes/sec on the drive, data transfer takes 200 seconds. Not including data transfer, cartridge switches take about 250 seconds in this system, on average. So, each drive is unable to transfer data more than half the time due to cartridge switches. If there are few outstanding requests in the system, the throughput benefits of striping outweigh the penalty suffered for additional cartridge switches. But at higher concurrencies, the large number of cartridge switches creates long queues for drives and the robot arm, increasing response times.

This contention is alleviated to some degree with the addition of extra drives. Figure 7 shows the effect of a hypothetical EXB-120 system where the number of drives is doubled to eight. We have kept the other parameters of the simulation identical to the earlier case, so files are still being striped across groups of three data cartridges plus one parity cartridge. In such a system, at a concurrency of four, striping improves response time for requests over average size 400 MBytes. At a concurrency of eight, the number of cartridge switches penalizes striped

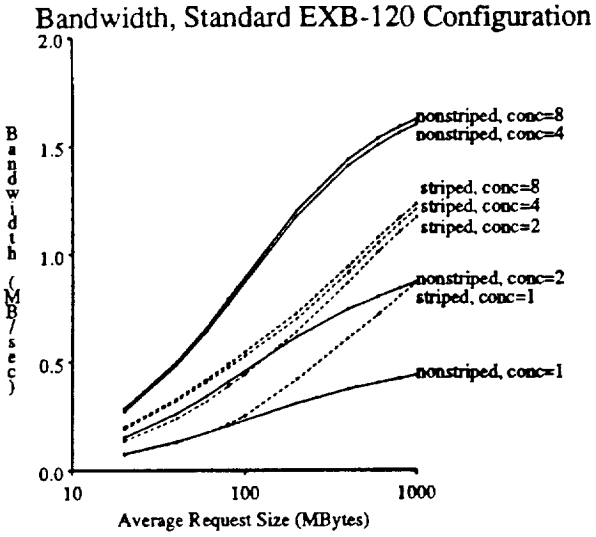

Figure 6: *Bandwidth comparison of striped system with interleave unit of 100 MBytes to a non-striped system, for concurrencies of 1, 2, 4 and 8. Standard Exabyte configuration.*

systems severely.

To explore the effect of reducing cartridge switch time on the original four-drive system, we simulated the robot with a new set of mechanical parameters, shown in Table 7. These hypothetical mechanical improvements reduce cartridge switch time to 60 seconds on average. As shown in Figure 8, while these mechanical improvements reduce response times for all configurations and concurrencies, striped systems still perform worse than non-striped systems at higher concurrencies. The additional accesses required in the striped system cause contention for the four readers. For good response time at high concurrencies, more readers must be added to the system.

Large, High Performance Library

The other storage array we simulated is similar to the Ampex DST800 Library. The library holds 600 cartridges and four tape drives, with a single robot arm loading the drives. Each cartridge holds 25 GBytes. Table 8 shows simulation parameters for the drive and robot. These parameters are loosely based on product literature for the Ampex DST600 drive and DST800 robot [12]. However, we don't accurately simulate the Ampex robot, since our drive model currently does not account for the periodic eject zones available on the DST600 drive that im-

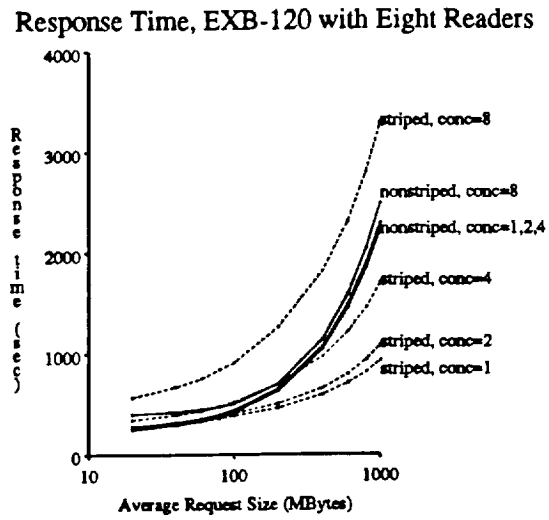

Figure 7: *Response time comparison of striped system with interleave unit of 100 MBytes to a non-striped system, for concurrencies of 1, 2, 4 and 8. Exabyte configuration including eight readers.*

Mechanical operation	Time or Rate
Eject	5 sec
Load	5 sec
Search startup time	5 sec
Search rate after startup	1 GB/sec
Rewind startup time	5 sec
Rewind rate after startup	1 GB/sec
Robot move time	2 sec
Robot pick time	3 sec
Robot place time	3 sec

Table 7: *Hypothetical parameters for improved drive mechanics; new average cartridge switch time is 60 seconds.*

prove rewind and search performance. The striping configuration simulated is similar to that used in the EXB-120 simulations: striping over three data cartridges and one parity cartridge, with an interleave unit of 100 MBytes.

Striping is not particularly effective in this robot except for very large requests (over 1 GByte for concurrency of 1, and even larger requests for higher concurrencies). The reason for this is the high transfer rate provided by the drive. For a stripe unit of 100 MBytes, data transfer takes approximately 7 seconds, while a cartridge switch takes 64 seconds. For requests smaller than 1 GByte, transfer time is far outweighed by cartridge switch time. At high concurrencies, the drives will spend the majority of

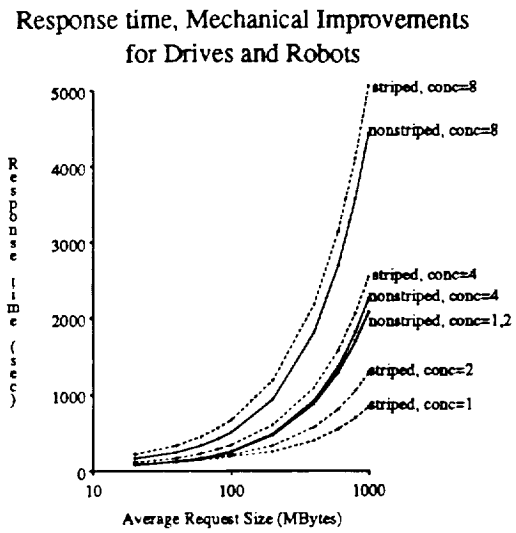

Figure 8: *Response time comparison of striped system with interleave unit of 100 MBytes to a non-striped system, for concurrencies of 1, 2, 4 and 8. Readers and robots have mechanical improvements to reduce average cartridge switch time to 60 seconds.*

Operation	Time or Rate
Data transfer rate	15 MB/sec
Eject	5 sec
Load	5 sec
Search startup time	5 sec
Search rate after startup	750 MB/sec
Rewind startup time	5 sec
Rewind rate after startup	750 MB/sec
Robot move time	2 sec
Robot pick time	3 sec
Robot place time	3 sec

Table 8: *Simulation parameters for large, high performance library.*

their time on cartridge switches, and striping will perform poorly. It is likely that this system will be a poor striping candidate except for workloads where concurrency is low and average request sizes are in the GByte range.

Simulation Summary

We have shown that for a small library composed of many cartridges, few drives and a single robot arm, contention for the small number of readers limits the value of striping for workloads with high concurrency. For a large library with fast robots and high throughput drives, the cartridge switching time pre-

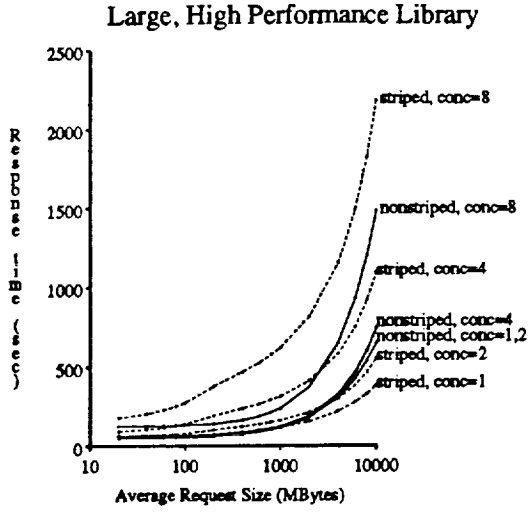

Figure 9: *Response time comparison of striped system with interleave unit of 100 MBytes to a non-striped system, for concurrencies of 1, 2, 4 and 8. Large, high performance library: four drives, 600 cartridges, one robot arm.*

dominates and penalizes striped systems for all but the largest accesses at low concurrencies. We will be exploring many more robot configurations, including combinations of low-performance, inexpensive readers and robots, which are likely to get the greatest benefit from striping.

FUTURE WORK

We plan to validate the accurateness of our simulator by comparing simulation results with measured robot performance on comparable workloads. We will continue simulations to better understand how to configure striped tape systems, and what workloads can benefit from striping. The simulator will be improved to add subtlety to the models of tape and robot behavior. For example, we currently model transfer time using aggregate measured bandwidth from the devices. The model does not take into account the difficulty of keeping devices streaming, or the penalty paid when the drive pauses. In addition, our performance simulations currently don't model the effect of bit errors or drive failures, or the difficulty of synchronizing drives.

We will extend our array simulator to examine arrays of optical and magneto-optical disks.

We will model the reliability characteristics of tape media, drives and robots. Tape media suffer from a relatively high rate of bit errors, and tapes that are frequently read or written wear out after a few

hundred or thousand passes [13] [8] [14]. Tape heads undergo considerable wear, and last for only a few hundred or thousand hours of contact with the media [15]. Tape drives suffer from mechanical and electronics failure as well as head wear-out. In addition, robot mechanics and supporting hardware may fail. We are studying these reliability issues, and will apply some of the techniques used by Gibson [5] to determine the mean time to data loss of the tape array given various redundancy schemes. We intend to determine the amount of external error correction that must be added to the array to maintain adequate reliability.

Finally, we plan to implement a striped tape system and to measure its effectiveness under real workloads.

SUMMARY

Data striping in magnetic tape arrays offers the potential of greatly improved response time on large accesses. Our simulations have demonstrated, however, that the storage system designer must carefully configure the tape array to ensure that striping helps rather than hurts storage system performance. Issues that affect the success of striped systems include the throughput and access times of readers, the number of readers in a robot, the speed and number of robot arms, the ratio of cartridges to readers and robot arms, the choice of striping configuration (within or between robots), the striping interleave unit, the difficulty in synchronizing the tape drives and the concurrency of the workload. We will continue to explore these issues with simulations, and with striped tape array implementations.

REFERENCES

[1] Ben Kobler and John Berbert. NASA earth observing system data information system (EOSDIS). In *Digest of Papers.* Eleventh IEEE Symposium on Mass Storage Systems, October 1991.

[2] Wiliam M. Callicott. Data management in NOAA. In *Goddard Conference on Mass Storage Systems and Technologies.* NASA/Goddard Space Flight Center, September 1992.

[3] K. Salem and H. Garcia-Molina. Disk striping. In *Proceedings IEEE Data Engineering*, pages 336–342, February 1986.

[4] M. Y. Kim. Synchronized disk interleaving. *IEEE Transactions on Computers*, C-35:978–988, November 1986.

[5] Garth Alan Gibson. *Redundant Disk Arrays: Reliable, Parallel Secondary Storage.* PhD thesis, U. C. Berkeley, April 1991. Technical Report No. UCB/CSD 91/613.

[6] David D. Larson, James R. Young, Thomas J. Studebaker, and Cynthia L. Kraybill. Storagetek 4400 automated cartridge system. In *Digest of Papers.* Eighth IEEE Symposium on Mass Storage Systems, May 1987.

[7] Metrum Information Storage. RSS-600 Rotary Storage System product literature.

[8] C. Denis Mee and Eric D. Daniel, editors. *Magnetic Recording, Volume II: Computer Data Storage.* McGraw-Hill, New York, 1988.

[9] David A. Patterson, Garth Gibson, and Randy H. Katz. A case for redundant arrays of inexpensive disks (RAID). In *Proceedings ACM SIGMOD*, pages 109–116, June 1988.

[10] Harry C. Hinz. Magnetic tape technology in the 1990s. In *Digest of Papers.* Tenth IEEE Symposium on Mass Storage Systems, May 1990.

[11] Storage Technology Corporation. Powderhorn product literatrue.

[12] Ampex Corporation. DST800 and DST600 product literature.

[13] H. Goto, A. Asada, H. Chiba, T. Sampei, T. Noguchi, and M. Arakawa. A new concept of data/DAT system. *IEEE Transactions on Consumer Electronics*, 35(3), August 1989.

[14] Bharat Bhushan. *Tribology and Mechanics of Magnetic Storage Devices.* Springer-Verlag, New York, 1990.

[15] C. Denis Mee and Eric D. Daniel, editors. *Magnetic Recording, Volume III: Video, Audio, and Instrumentation Recording.* McGraw-Hill, New York, 1988.