Evolving Requirements for Magnetic Tape Data Storage Systems

S32-15

John J. Gniewek
IBM Corporation
9000 S. Rita Road
Tucson, Arizona 85744
jgniewek@vnet.ibm.com
+1-520-799-2390
Fax: +1-520-799-3665

Introduction

Magnetic tape data storage systems have evolved in an environment where the major applications have been back-up/restore, disaster recovery, and long term archive. Coincident with the rapidly improving price-performance of disk storage systems, the prime requirements for tape storage systems have remained: 1) low cost per MB, and 2) a data rate balanced to the remaining system components. Little emphasis was given to configuring the technology components to optimize retrieval of the stored data. Emerging new applications such as network-attached HSM, and digital libraries, place additional emphasis and requirements on the retrieval of the stored data. It is therefore desirable to consider the system to be defined by both STorage And Retrieval System (STARS) requirements. It is possible to provide comparative performance analyses of different STARS by incorporating parameters related to A) device characteristics, and B) application characteristics in combination with queuing theory analysis. Results of these analyses are presented here in the form of response time as a function of system configuration for two different types of devices and for a variety of applications.

STARS (STorage And Retrieval System)

A) Performance Model

Some simplifying assumptions will be necessary to be able to provide comparative performance analyses for two different tape devices. A list of the required input parameters for both A) Hardware characteristics and B) Application characteristics is given in Table 1. The output of the model will be given as an average response time for various combinations of these parameters. In order to directly compare device types only, the assumption is made that both devices are serviced by a robot with identical characteristics, i.e. a fixed robot average cycle time with no allowance for queuing delays at the robot level. Device queuing delays are calculated using the methodology in reference [1]. For this analysis* an M/M/C queue is used.

* The model and graphical output are developed using MathCad® 6.0 (®-MathSoft...)

Table 1
Input Parameters Required for Performance Model

Hardware Characteristics			Application Characteristics		
Device data rate	(MB/sec)	(D)	Service Request Rate	(#/Hr)	(I)
Cartridge Capacity	(MB)	(C)	Object Size	(MB)	(O)
Recording Density	(MB/m)	(K)	Library Size	(MB)	(L)
Search/Rewind Velocity	(M/sec)	(V)	Random Retrieval Factor	(#)	(A)
Load Time	(sec)	(LD)	Drive Number	(#)	(N)
Unload Time	(sec)	(ULD)			
Robot Cycle Time	(sec)	(AS)			

It is necessary to distinguish between: a) a base cycle time which is a necessary input for calculating queuing delay times, and b) a base service time which is defined as the response time in the absence of any queuing delays. Additionally, service times are defined for the cases where: 1) the required cartridge is already mounted in a drive, or 2) the cartridge is in a library bin and must be transported to and loaded into a drive. The assumed sequences of operations for both cycle time and service time are listed in Table 2. These sequences assume that there is no preemptive unloading of cartridges upon completion of a service request. It is additionally assumed that at the completion of each service request, there is a rewind to the starting point prior to servicing the next request.

Table 2
Sequence of Operations for Cycle Time and Service Time

Cycle Time		Service Time	
Unmounted	Mounted	Unmounted	Mounted
Unload Cartridge		Unload Cartridge	
Robot Put Cartridge		Robot Put Cartridge	
Robot Get Cartridge		Robot Get Cartridge	
Load Cartridge		Load Cartridge	
Search	Search	Search	Search
Read	Read	Read	Read
Rewind	Rewind		

In order to calculate an average system response time, it is necessary to be able to estimate the probability, P, that the next incoming request will be serviced by an already mounted cartridge as opposed to requiring a robotic Put and Get operation. The expression given in Equation (1) is used to estimate this probability as a function of the library size, the cartridge capacity, the number of drives in the robotic system, and an application dependent adjustable parameter, A, which characterizes the degree to which the requested objects are randomized within the library cartridges [2].

$$P = \frac{1}{100}\left\{100^A - \left(100 - 100\left(\frac{N \cdot C}{L}\right)\right)^A\right\}^{1/A} \qquad (1)$$

where: A = application dependent random factor
C = cartridge capacity (MB)
L = library size (MB)
N = number of drives in library
P = probability of next request being serviced by a mounted volume

Figure 1 shows P as a function of the cumulative percent of volumes in the library ordered by activity from highest to lowest. Equation (1) is derived empirically, however its relevance to a statistical analysis is covered in reference [2].

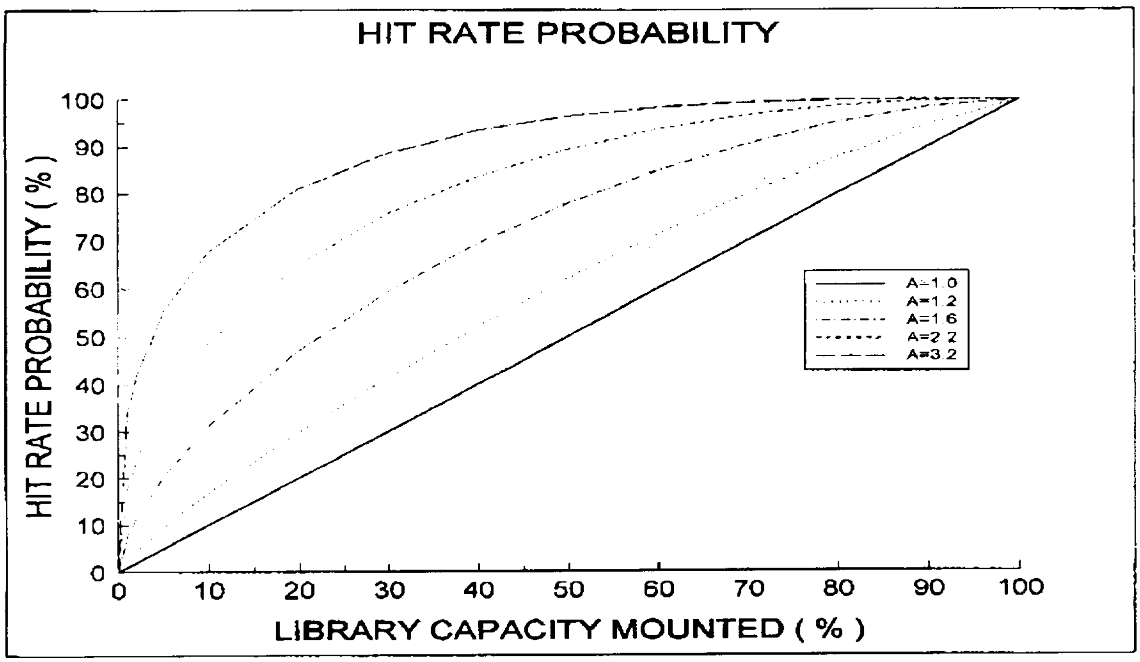

Figure 1. **Probability, P, of next request being serviced by a mounted volume as a function of percent of library capacity mounted. Plotted from Equation (1) for A = 1.0, 1.2, 1.6, 2.2, and 3.2.**

The output of the model is presented in terms of an average System Response Time (SRT), (to complete the service request) as a function of various combinations of the other input parameters. Intermediate output parameters include, in addition to P, an average cycle time, CT, (appropriately weighted by the fraction of requests serviced by mounted and unmounted volumes), a drive utilization factor, U, and an average device queuing delay, QD.

The expression for the queuing delay time for multiple servers in an M/M/C queue is adapted from reference [3] as:

$$QD = CT\left(\frac{(NU)^N \cdot P_o}{N!\ N(1-U)^2}\right) \tag{2}$$

$$\text{where } P_o = \left[\frac{(NU)^N}{N!(1-U)} + \sum_{i=0}^{N-1} \frac{(NU)^i}{i\ !}\right]^{-1} \tag{2a}$$

The M/M/C queue designates an exponential interarrival time request distribution and an exponential service time distribution with means designated by the chosen values for the input parameters.

Average cycle times, utilization factors, and service times are calculated in a straight forward manner from the parameters designated in Tables 2 and 3. The average system response time is the sum of the average service time and the average queuing delay time.

Table 3
Hardware Characteristics of Two Prototype Devices

Parameter			Device I	Device II
Data Rate	(D)	(MB/sec)	9	2.2
Cartridge Capacity	(C)	(MB)	10000	5000
Recording Density	(K)	(MB/m)	34	34
Velocity	(V)	(m/sec)	5	10*
Load Time	(LD)	(sec)	15	2
Unload Time	(ULD)	(sec)	11	2
Robot Cycle Time	(AS)	(sec)	fixed**	fixed**

* Actual Search Velocity 5 m/sec. Midpoint cartridge load translates this parameter into an effective 10 m/sec consistent with the model formulation.

** Assumed common robotic device to highlight contrasts in device characteristics. See Figures for values used.

B) Hardware Characteristics

A description of two different types of prototype devices that could be developed from advanced technology recording components has previously been presented [4]. Their characteristics are designed to complement each other for different application requirements. For the purpose of this performance analysis, devices are assumed with characteristics similar to those previously described. The specific values used in the comparative performance analysis are given in Table 3. STARS are defined by specifying:

a) a robot system with a fixed cycle time**, b) the type of device (I or II) defined in Table 3, c) the number of devices in the library system, and d) the library capacity (in MB).

Analysis

In addition to the large number of STARS hardware variables, the analysis must accommodate the application variables; I, the service request rate, O, the object size being retrieved, and A, the random retrieval factor. In the first analysis, the system response time is calculated as a function of the service request rate for both types of devices and with the number of drives as an independent parameter. All other parameters are fixed. This is done for both a large object size, 300 MB, and a small object size, 1 MB, to highlight the differences in device characteristics.

A helpful way to assess appropriate operating domains for the different types of devices is to plot the average system response time as isochrons over the domain space of object size, O, and service request rate, I, with the remaining parameters fixed. Most of the calculated results are presented in this format. System configuration variables such as the number of drives in the library represent another dimension and these data can be represented parametrically in separate, individual plots.

Results

A. System Response Time as a Function of Request Rate

Figures 2A and 2B show the system response time as a function of request rate for the two different types of devices in system configurations of 1, 2, or 4 drives, and for a library capacity of 10 TB, with a value of the randomness factor, A, equal to 3.2 Figure 2A plots results for a 1 MB object size and Figure 2B is calculated for 0 = 300 MB. Figures 2C and 2D repeat these calculations with all values the same except the library capacity is set to 0.5 TB. This has the effect of modifying the mounted/unmounted service request ratio via the probability function given in Figure 1. The effect this has on system response time is a complex function of device characteristics, robotic cycle time and the specific application parameters. The data in Figure 2 illustrate one possible scenario. In general the response time is improved at smaller library capacities as a result of a higher probability of the request being satisfied by an already mounted volume, thereby eliminating the need to invoke a robotic move to satisfy that parameter request.

** The performance model presented here does not provide for robotic queuing delays in order to emphasize the different characteristics of the two types of devices.

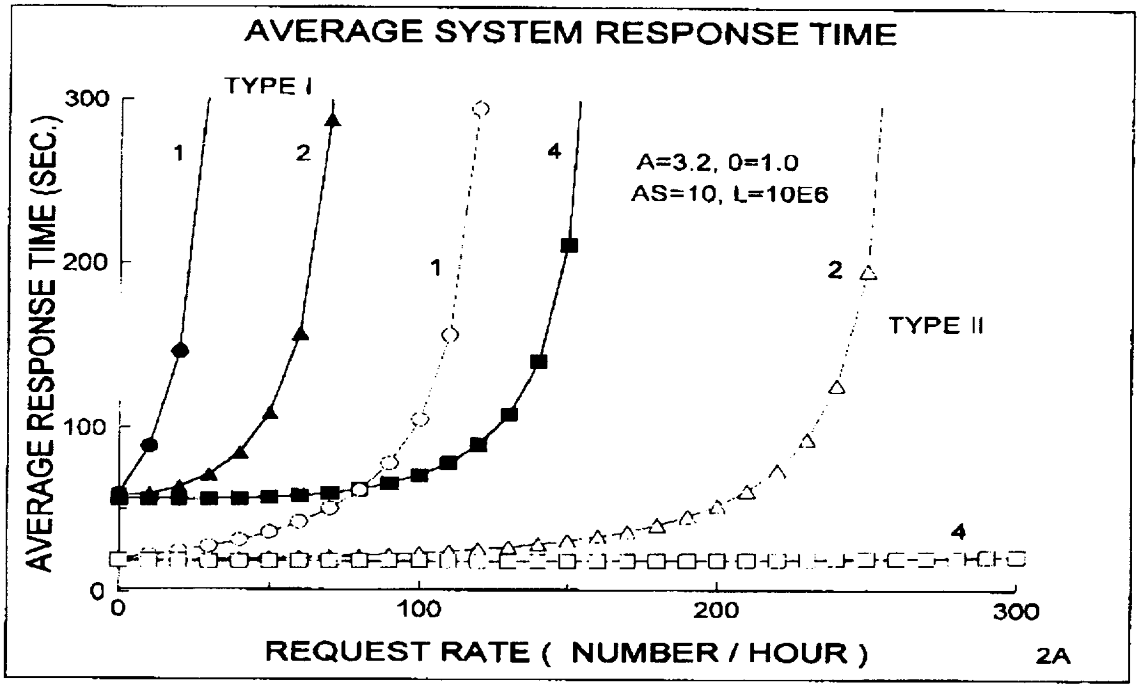

Figure 2A. Average system response time, SRT, as a function of service request rate (#/hour), for N = 1, 2, 4 drives of Device Type I and Device Type II. A = 3.2, AS = 10 seconds, L = 10 TB, O = 1 MB.

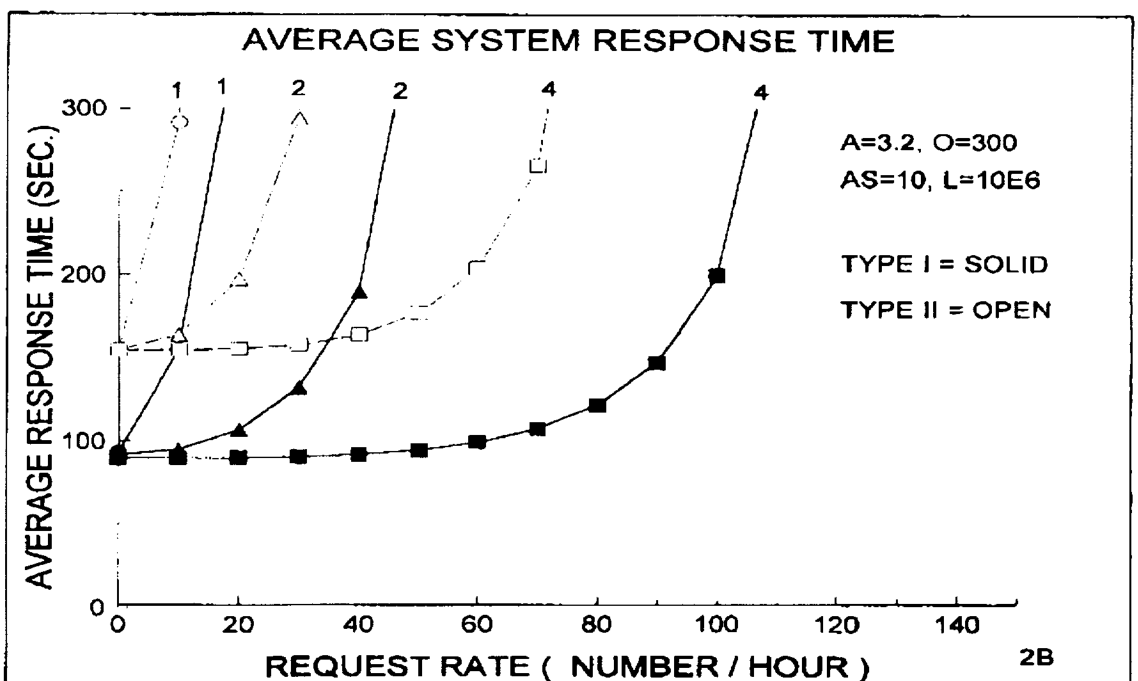

Figure 2B. Average system response time, SRT, as a function of service request rate (#/hour), for N = 1, 2, 4 drives of Device Type I and Device Type II. A = 3.2, AS = 10 seconds, L = 10 TB, O = 300 MB.

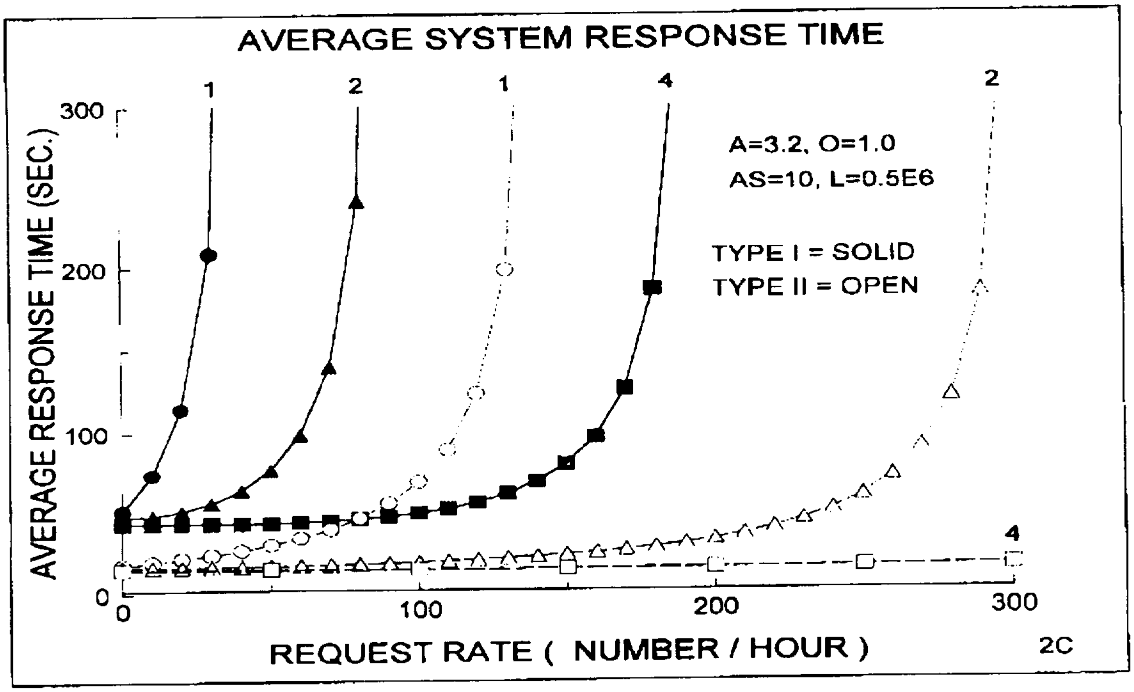

Figure 2C. Average system response time, SRT, as a function of service request rate (#/hour), for N = 1, 2, 4 drives of Device Type I and Device Type II. A = 3.2, AS = 10 seconds, L = 0.5 TB, O = 1 MB.

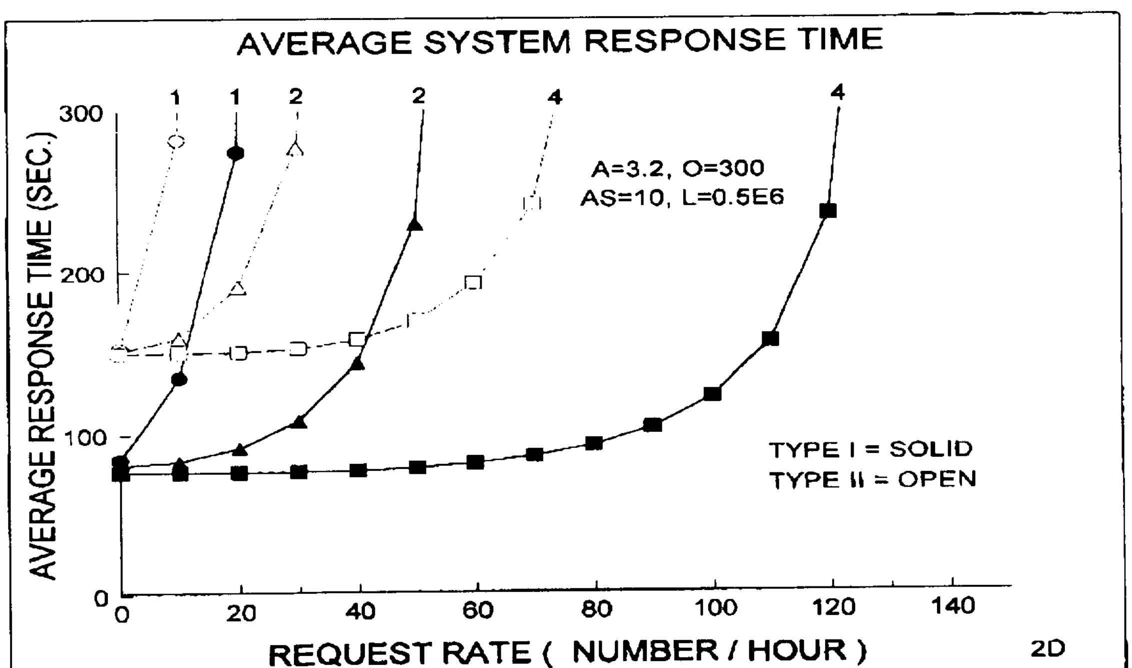

Figure 2D. Average system response time, SRT, as a function of service request rate (#/hour), for N = 1, 2, 4 drives of Device Type I and Device Type II. A = 3.2, AS = 10 seconds, L = 0.5 TB, O = 300 MB.

It is readily apparent from the data presented in Figure 2 that Device Type I provides superior performance (on a per drive basis) for those applications requiring large object sizes and modest request rates. Conversely, Device Type II excels for those applications characterized by modest object sizes and high request rates. In order to get a better perspective of the preferred operating domain for these two different types of devices, the data is next presented as a series of isochrons (lines of constant system response time) over the domain space of object size X request rate. The information is presented with number of drives as one of the parameters thus resulting in performance comparisons on a "per drive" basis.

B. System Response Time Isochrons

Figures 3A and 3B display isochrons of average system response time for Device Types I and II respectively. These figures display the operational range of object size (MB) X request rate (#/hour) that can be satisfied within 60 seconds, 90 seconds, or 120 seconds, for a system configuration of 4 drives, a library capacity of 10 TB, and a random retrieval factor, A = 3.2. In Figure 3B, as a result of the faster response times (for small to medium object sizes) of the Type II Device, an isochron at 30 seconds is also shown. Figure 4 shows only the 90 second isochron for both types of devices in overlay fashion to better illustrate the operational domain where each type of device excels, as well as the range of overlap.

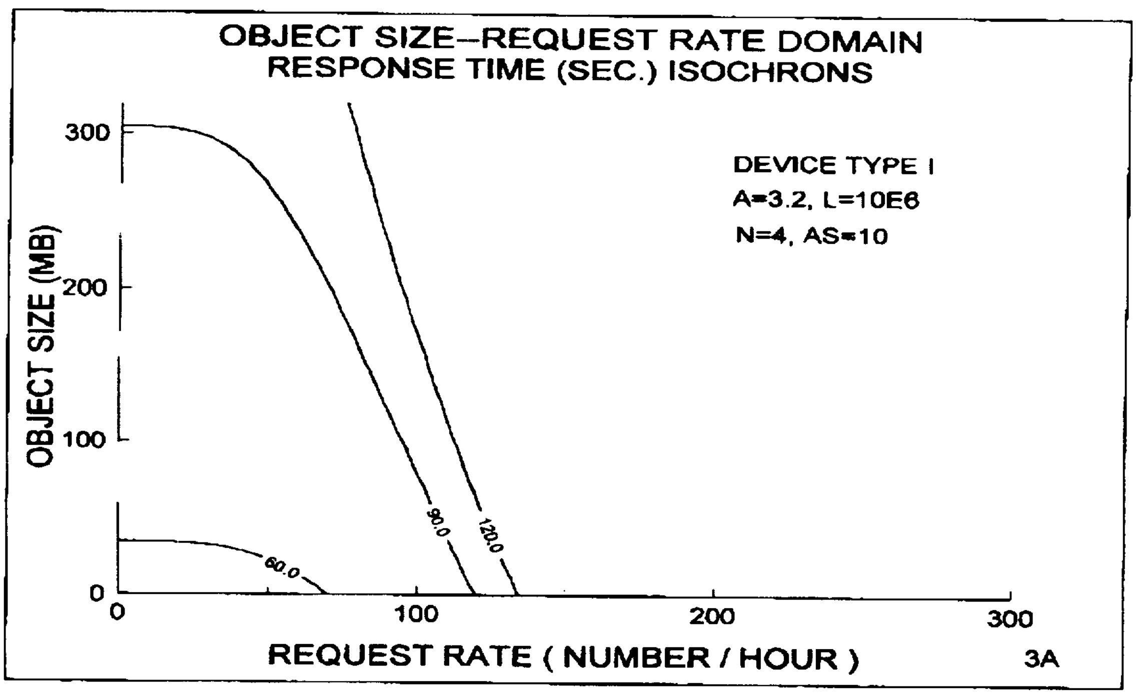

Figure 3A. Isochrons of average system response time, SRT, in seconds in the operating domain of object size (MB) X request rate (#/hour). A = 3.2, N = 4, L = 10 TB, AS = 10 seconds. Device Type I.

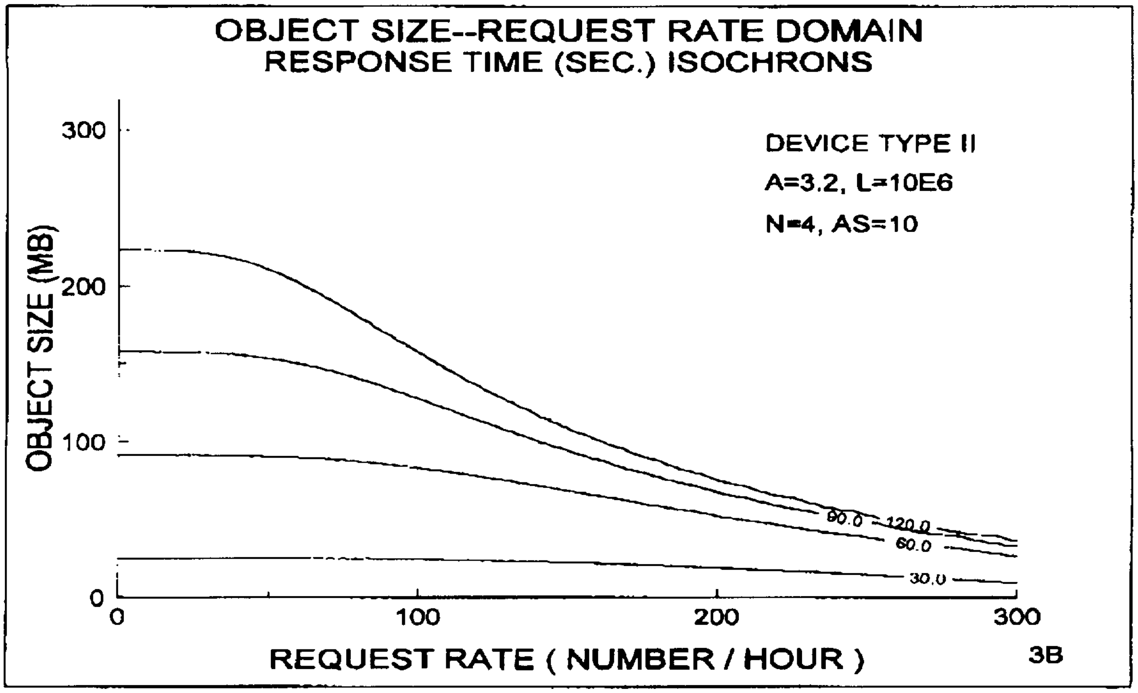

Figure 3B. Isochrons of average system response time, SRT, in seconds in the operating domain of object size (MB) X request rate (#/hour). A = 3.2, N = 4, L = 10 TB, AS = 10 seconds. Device Type II.

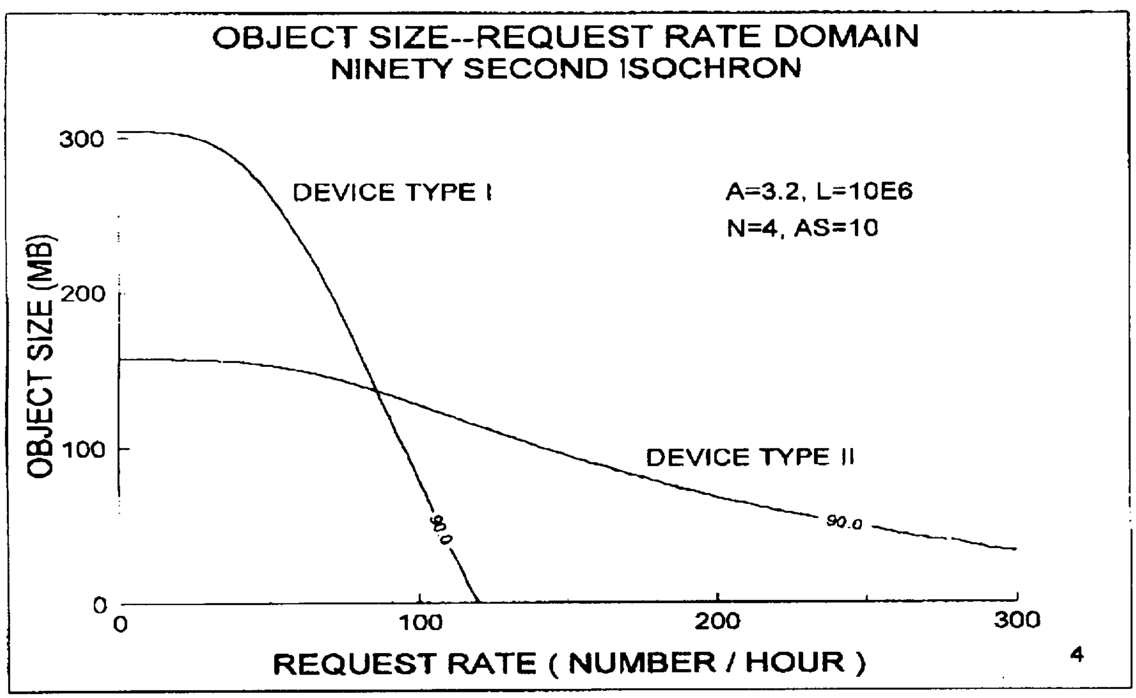

Figure 4. Ninety second isochron of average system response time in the operating domain of object size (MB) X request rate (#/hour). Overlay of Device Type I and Device Type II. Same conditions as defined in Figure 3.

Discussion

These analyses provide a means of quantifying expected performance for many variables characterizing different types of devices, different system configurations, and different application parameters. It is obvious that a higher data rate device will perform better than a lower data rate device for large object sizes. However, because of the non-linear nature of the queuing delays, it is not obvious where the crossovers occur in the operating space of object size X request rate. A comparison of the data in Figures 3A and 3B illustrates that for object sizes up to slightly in excess of 100 MB, the Type II device equals or exceeds the performance of the Type I device for all values of the request rate. At 150 MB, the Type II device becomes superior only for request rates greater than approximately 100 per hour for the configurations assumed. In order to convert this type of analysis to a price-performance analysis (rather than on a per drive basis) it would be necessary to first convert the configurations to equal dollar configurations and then compare performance over the operational space of interest for the given application requirements.

The performance results are specific for the defined assumptions made for the cycle time and service time components. Performance enhancements via software control algorithms are possible. For example, the algorithm used in this analysis assumes that following a read operation, the device rewinds the tape to the beginning of the tape and the tape remains mounted until a service request arrives that requires a new cartridge mount whereupon a drive is unloaded. If a request arrives for an object on a cartridge that is mounted, the drive searches from the beginning of tape to the new object location without invoking a robotic action. Depending upon the application, two possible alternative cycle sequences may provide better performance. In one situation it may be preferable to search for an incoming request from the stop point of reading the previous request rather than automatically rewinding to the beginning of tape. This could result in shorter average search distances. Another scenario could provide preemptive drive unloads [5] which might shorten robotic service times under some application conditions. An early pre-analysis of the specific use conditions would permit "tuning" the system for optimized performance.

Cartridge Capacity Considerations

In passive tape storage applications that very infrequently retrieve stored data objects, the emphasis has been on higher data rates and higher cartridge capacity. Increases in capacity can be achieved by either an increase in areal density or by way of a longer length, thinner tape. The K value, given in MB/M, is reflective of the areal density capability for a given tape width. For a fixed value of K, capacity is linearly dependent upon tape length, which in turn affects search and rewind times. An analysis of the average system response time as a function of cartridge capacity (i.e. length) is shown in Figures 5 and 6 as isochrons of average response time over the domain space of request rate X capacity for the fixed conditions listed.

Device Type II is unique in its design for type storage devices to be able to economically provide solutions for active applications such as HSM and digital libraries. However, for wide acceptance, it must also be capable of meeting the needs of the passive applications.

Hence, it was necessary to provide a carefully considered balance between capacity and response time under multi-user loaded conditions. The performance target was set to be in the range of 15-30 second average response time for request rates of a few hundred per hour and with an economical number of devices. There are many variables that affect the design space over which these objectives may be achieved. These include the randomization factor (A), the library size, the robot cycle time, and the size of the object being retrieved. Using the technology recording density (K = 34) and the design of Device Type II, a capacity in the range of 5-10 GB provides a reasonable balance to meet the wide range of application characteristics. This is illustrated in Figures 5 and 6. These analyses are analogous to assessments of the trade-offs made between disk storage capacity and number of actuator arms. The result has been smaller physical disk sizes as the technology advanced to provide higher recording densities.

Figure 5A presents 20, 25, and 30 second isochrons over the domain space of request rate X cartridge capacity for the system parameters stated. Of note, randomization factor, A, is set at 2.2, the accessor cycle time = 15 seconds and the number of drives = 2. For a 5 GB cartridge capacity an average response time of $\leq$ 25 seconds is maintained up to approximately 100 requests per hour. Figure 5B illustrates the improved performance and the enlarged acceptable operating domain resulting from the addition of a third drive. Alternatively, improvements may be obtained by using a faster accessor. The results obtained with an accessor cycle time, AS of 10 seconds (and with 2 drives), is shown in Figure 5C. Figure 6A illustrates the effect (with 3 drives and AS = 15 seconds) of an application that has a highly non-random recall pattern (A = 3.2). This results in a high 'hit'

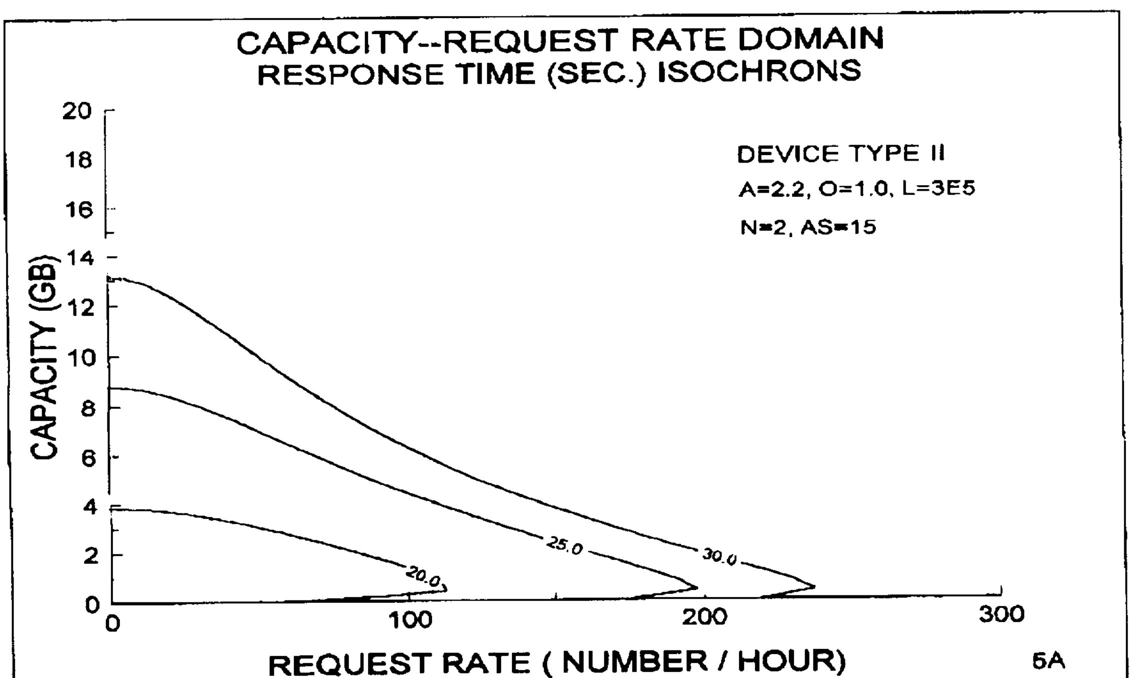

Figure 5A. Isochrons of average system response time, SRT, (seconds) in the design space of request rate X capacity (as determined by tape length) for Device Type II. A=2.2, 0=1 MB, L=3x10^5 MB, N=2, AS=15 seconds.

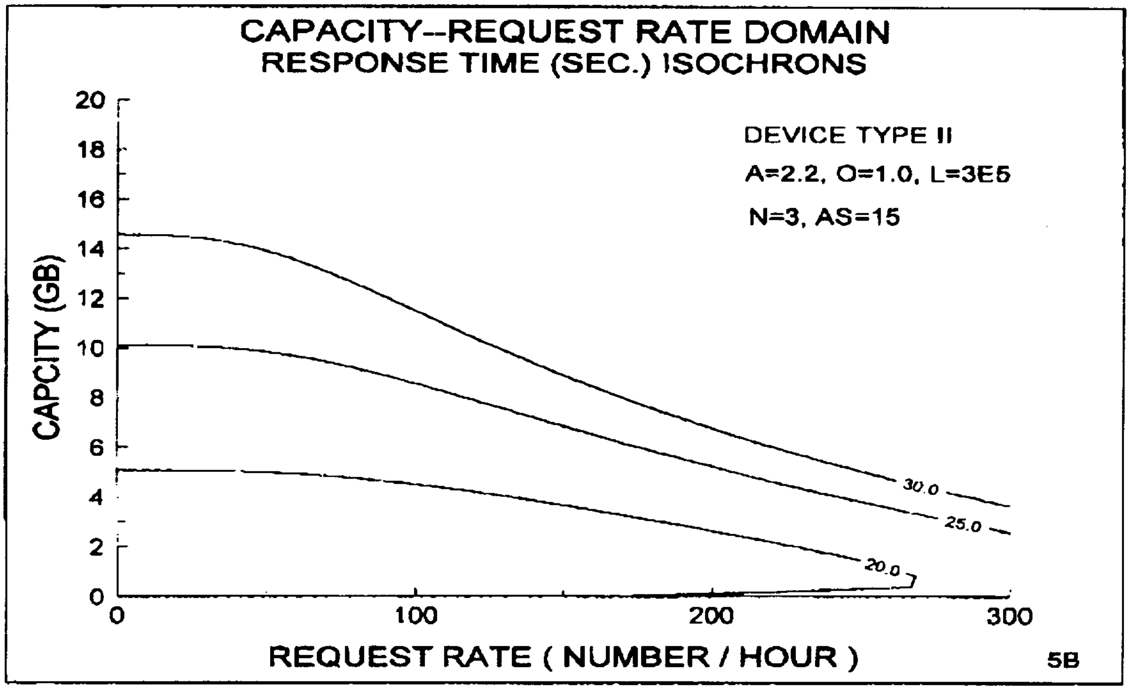

Figure 5B. Isochrons of average system response time, SRT, (seconds) in the design space of request rate X capacity (as determined by tape length) for Device Type II. A=2.2, 0=1 MB, L=3 x 10^5 MB, N=3, AS=15 seconds.

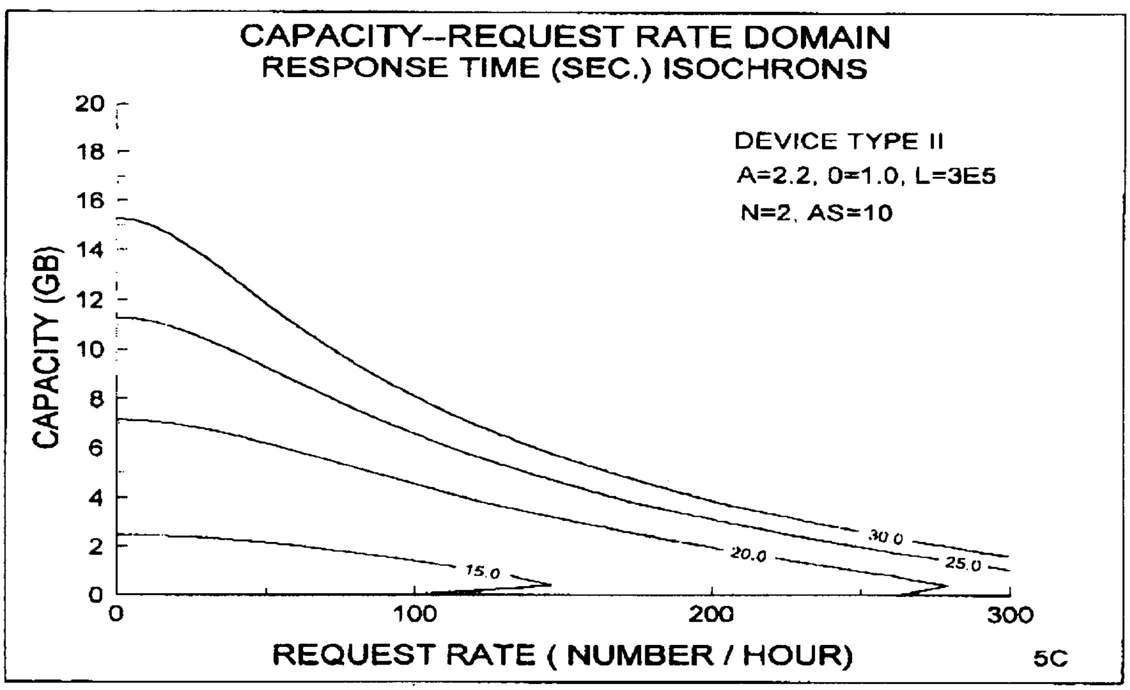

Figure 5C. Isochrons of average system response time, SRT, (seconds) in the design space of request rate X capacity (as determined by tape length) for Device Type II. A=2.2, 0=1 MB, L=3x10^5 MB, N=2, AS=10 seconds.

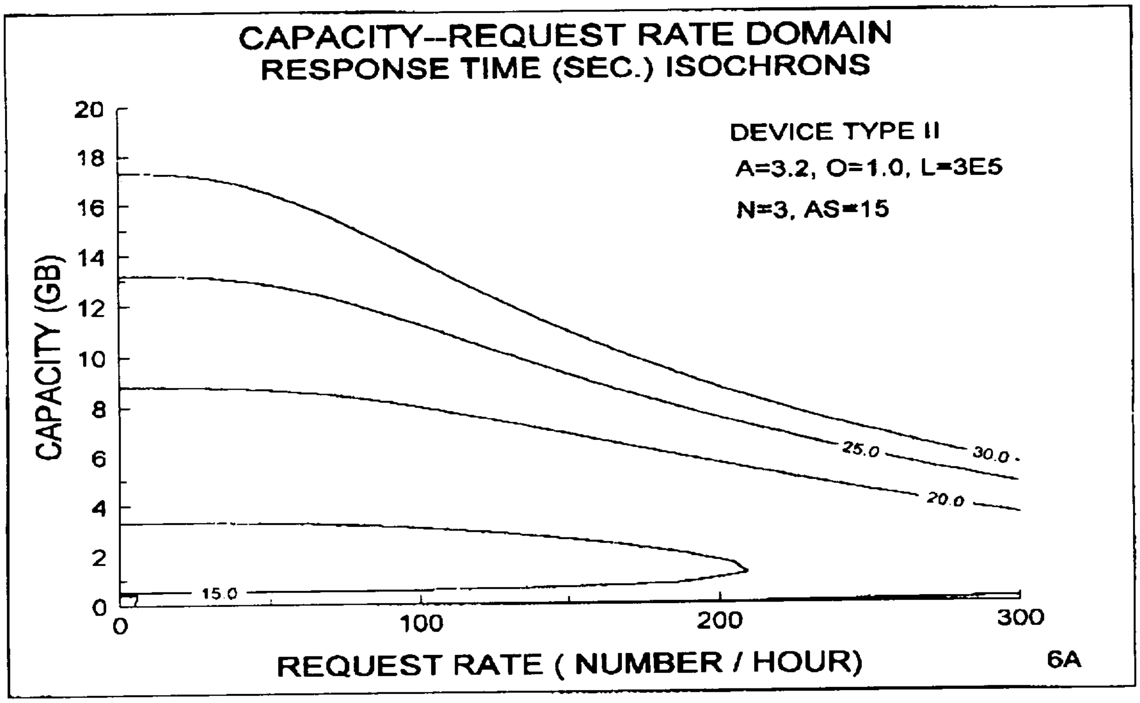

Figure 6A. Isochrons of average system response time (seconds) in the design space of request rate X capacity (as determined by increased tape length) for Device Type II. N=3, AS=15 seconds, O=1 MB, L=3x10^5 MB. A=3.2.

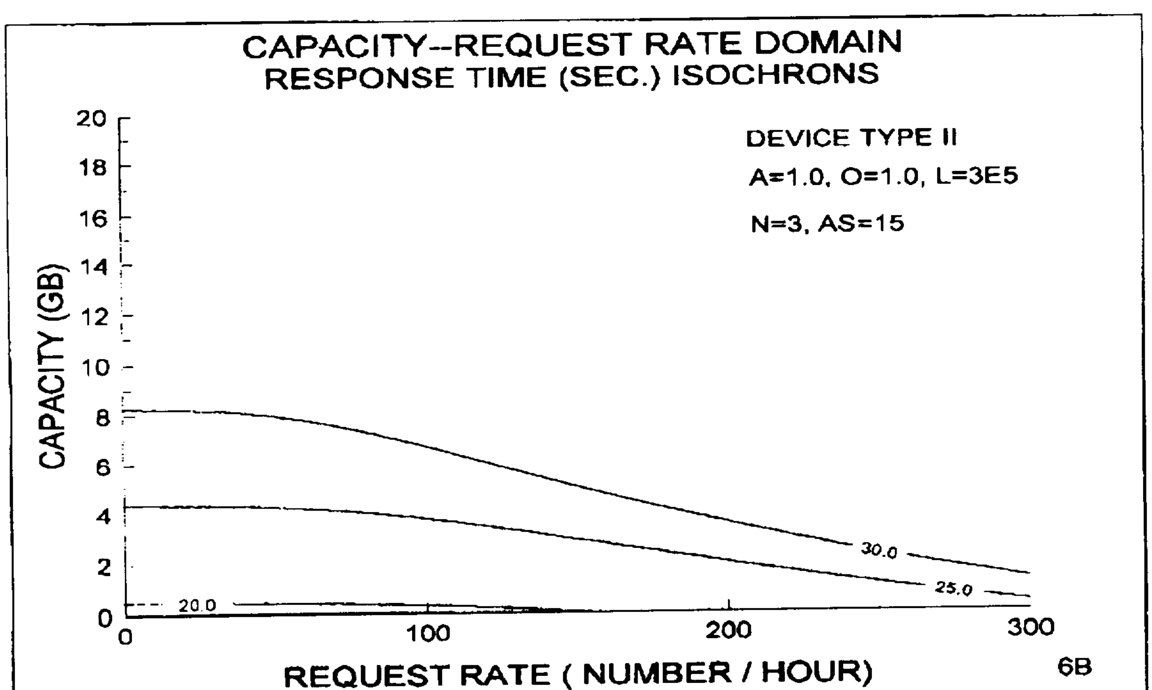

Figure 6B. Isochrons of average system response time (seconds) in the design space of request rate X capacity (as determined by increased tape length) for Device Type II. N=3, AS=15 seconds, O=1 MB, L=3x10^5 MB. A=1.0.

rate to mounted cartridges and hence performance improvement as a result of fewer required robot movements, and load-unload cycles. Figure 6B shows the other extreme of a totally random retrieval pattern (A=1) for the same system configuration parameters. Figures 6A and 6B should be compared to Figure 5B to see the effect of the application retrieval pattern on the response time for an otherwise fixed set of system parameters. Clearly, optimizing for one application will result in sub-optimization for other applications. The values chosen for the parameters for Device Type II in Table 3 provide a good balance, are achievable with current technology, and provide a basis for possible future technology enhancements.

Conclusions

The type of analysis presented here may be useful to the application engineer in comparing different STARS as to suitability for different application requirements. Likewise, this analysis has been used by developers to guide the development of device characteristics to meet existing or anticipated application requirements. The diversity of applications precludes the possibility of a single device doing all jobs equally well. A comparison of the preferred operating domains for two different types of devices which have been developed from a common advanced recording technology has been presented.

References

1. J.J. Gniewek, "Application of Queuing Theory to Performance Analysis of Automated Removable Media Storage Subsystem Response Time - Bounding the Problem," submitted for publication in IBM Journal of Research and Development.

2. J.J. Gniewek, "Factors Affecting Response Time Performance of Removable Media Storage Subsystems." Internal IBM Tucson Technical Report (1996). (This report is available from the author by e-mail request).

3. J. White, J. Schmidt, and G. Bennett, "Analysis of Queuing Systems," Academic Press, New York (1975).

4. J.J. Gniewek and S.M. Vogel, "Influence of Technology on Magnetic Tape Storage Device Characteristics," NASA Conference Publication 3295, Fourth NASA Goddard Conference on Mass Storage Subsystems and Technologies, 237-251 (March, 1995).

5. J.C. Hartung, et. al., "Preemptive Demount in an Automated Storage Library," U.S. Patent 5,239,650, issued 8/24/1993.