SDS

MEMORANDUM

TO R. Spinrad DATE 13 September 68

FROM J. Shemer

SUBJECT: Performance Investigation of Mass Memories REFERENCE

ANALYSIS OF "IBM 2314" VERSUS "7232 RAD AND CRAM" FOR FILE STORAGE IN TSU

I INTRODUCTION

This memo investigates the performance of two hardware organization schemes for file storage in the TSU time-sharing environment — 1) a mass storage device identical to the IBM 2314 disc and 2) a storage heirarchy of a 7232 RAD coupled with a lower level CRAM mass memory in which files migrate between the two devices as determined by file promotion/demotion activity with the majority of file I/O to and from the 7232 (i.e., this latter structure is analogous to that currently proposed for TSU.) In this analysis, it is assumed that the principal criterion of performance is response* to file access requests since rapid response is a principal goal of time-shared computer operation. Therefore, the analysis focuses attention upon the derivation of the expected response time experienced by individual file I/O requests as a function of device performance characteristics and device loading. In addition, estimates of core memory requirements for file I/O buffering are obtained for each of the two implementations.

The mathematical analysis is presented in Section II. Then utilizing the models of Section II together with example parameters constructed from typical TSU applications, a set of graphical studies, comments, and recommendations are presented in Section III.

II MATHEMATICAL MODELS OF THE 2314 AND THE 7232/CRAM FILE SYSTEMS

A. Model of 2314 Seek and Latency Servicing

Assume requests for file access occur at Poisson rate λ and are equally distributed

---

* Here, response is defined as that period of time elapsing between requesting file I/O from a device and completion of that file I/O request by the appropriate device, neglecting any software time necessary to set up or acknowledge such I/O transfer.

among n storage modules which comprise the 2314 system (i.e., $1 \le n \le 8$). Furthermore, with the 2314 storage system, assume response to file requests is achieved subsequent to service in two queues. First, requests enter a seek (or positioning) queue in which service is first-come-first-served, and each of the n storage modules can concurrently serve distinct requests. Then upon completion of positioning service, the request enters a latency queue where requests for data channel service are treated first-come-first-served, yet on a one at a time basis by each data channel.

Thus to study response characteristics of a 2314 type of device, the problem is to examine the random variables w_1, t_1, w_2, and t_2; where w_1 and t_1 denote the waiting time and service time in the seek queue, respectively; and w_2 and t_2 denote the waiting time and service time in the latency queue, respectively.

For notational purposes, let $p_x(t)$ represent the probability density function for a random variable x with respect to the independent variable t (where t denotes time), and let X(s) be the Laplace-Stieltjes transform of the distribution for x. That is

$$X(s) = \int_0^\infty e^{-st} p_x(t)\, dt$$

Employing this notation, consider the variables w_1 and t_1 are independent, then the time m which a request spends waiting for and accomplishing cylinder positioning is distributed as the sum of w_1 and t_1; whereby

$$M(s) = W_1(s)\, T_1(s) \tag{1}$$

Now since the input to the positioning queue is Poisson with mean request rate λ/n per storage module, the transform $W_1(s)$ is obtained from the Pollaczek-Khintchine formula*

---

* See "Elements of Queueing Theory by T. L. Saaty, McGraw-Hill, 1962.

$$W_1(s) = \frac{(1-p_1)s}{s - \lambda/n\ (1-B_1(s))} \tag{2}$$

where p_1 denotes the utilization factor of each storage module, and $B_1(s)$ is the Laplace-Stieltjes transform of the distribution for the access mechanism servicing period b_1 (i.e., b_1 equals the sum of the seek positioning time t_1, latency queue waiting time w_2, and latency queue service time t_2).

Here from the properties of the Laplace-Stieltjes transform, the expectation* of b_1 is

$$E\left[b_1\right] = \left.\frac{-dB_1(s)}{ds}\right|_{s=0} = -B'(0)$$

$$= E\left[t_1\right] + E\left[w_2\right] + E\left[t_2\right] \tag{3}$$

and the device load p_1 is given by

$$p_1 = \lambda/n\ E\left[b_1\right] \tag{4}$$

Substitution of these results into equation (1) yields

$$E\left[m\right] = E\left[w_1 + t_1\right] = -M'(0)$$

$$= \underbrace{\frac{\lambda\ E\left[b_1^{\,2}\right]}{2n\ (1-p_1)}}_{E\left[w_1\right]} + \underbrace{(-T_1'(0))}_{E\left[t_1\right]} \tag{5}$$

* Using standard notation, $E\left[\ \right]$ represents the expectation of the variable or function enclosed within the brackets.

where

$$E\left[b_1^{\,2}\right] = \left.\frac{d^2 B_1(s)}{ds^2}\right|_{s=0} \tag{6}$$

Extending this reasoning, consider the variables t_1, w_2, and t_2 are independent. Then the time d spent waiting for request recognition and accomplishing data transfer is distributed as the sum of w_2 and t_2; hence

$$D(s) = W_2(s)\ T_2(s)$$

where

$$B_1(s) = T_1(s)\ W_2(s)\ T_2(s) \tag{7}$$

and

$$W_2(s) = \frac{(1-p_2)\, s}{s - \lambda/c\,(1 - T_2\,(s))} \tag{8}$$

Here, all n devices supply cumulative input λ to the latency queue; thus if there are c data channels, then the utilization of each channel is expressed

$$p_2 = \frac{\lambda}{c}\ E\left[t_2\right] = \frac{\lambda}{c}\ \left(-T'_2\,(0)\right) \tag{9}$$

Thus, proceeding as before

$$E\left[d\right] = -D'(0) = \underbrace{\frac{\lambda\, E\left[t_2^{\,2}\right]}{2c\,(1-p_2)}}_{E\left[w_2\right]} + \underbrace{\left(-T'_2\,(0)\right)}_{E\left[t_2\right]} \tag{10}$$

where

$$E\left[t_1^2\right] = \left.\frac{d^2 T_2(s)}{ds^2}\right|_{s=0} \tag{11}$$

Hence, employing the foregoing results, the expectation of the 2314's response time r_f to file requests is

$$\begin{aligned} E\left[r_f\right] &= E\left[w_1 + t_1 + w_2 + t_2\right] \\ &= \underbrace{\frac{\lambda E\left[b_1^2\right]}{2n\,(1-p_1)} + (-T_1'(0))}_{E[m] \text{ seek service}} + \underbrace{\frac{\lambda E\left[t_2^2\right]}{2c\,(1-p_2)} + (-T_2'(0))}_{E[d] \text{ data channel service}} \end{aligned} \tag{12}$$

This expression can be further simplified by noting

$$\begin{aligned} E\left[b_1^2\right] = {} & E\left[t_1^2\right] + E\left[t_2^2\right] + E\left[w_2^2\right] \\ & + 2E\left[t_1\right]\left(E\left[t_2\right] + E\left[w_2\right]\right) + 2E\left[w_2\right]E\left[t_2\right] \end{aligned} \tag{13}$$

The result of equation (12) is of particular interest since it not only provides an indication of response speed, but also an estimate of core I/O buffer requirements for file referencing. Let n_f be a random variable denoting the number of pages which must be transferred for each file request (i.e., n_f is a function of the record size). Then the expected number of core pages which must be dedicated to sustain this I/O activity is

$$E\left[p_f\right] = \lambda E\left[n_f\right] E\left[r_f\right] \tag{14}$$

Note, however, that this result does not include the storage required for the 2314 I/O software and the data bases which this software must use to achieve file management.

B. Model of 7232/CRAM Servicing

Consider the 7232/CRAM system currently contemplated for TSU. In this system, the 7232 RAD is reserved for storage of active files, or portions of such files; whereas the CRAM is the mass storage reservoir for dormant files and large system files. Here it is assumed that files (or portions of files) are promoted and demoted between the CRAM and 7232 with Poisson mean activity rate λ' for each device. In addition to demand λ', CRAM service is requested at Poisson rate λ'' for individual page transfers (e.g., direct, single page references of a large file), and service of the 7232 RAD is demanded at Poisson file referencing frequency λ_o. Here, in order to be consistent with the 2314 model described in part A of this section, $\lambda_o = \lambda - \lambda''$ where λ equals the total file request frequency. Thus the CRAM and 7232 are demanded at Poisson request rates λ_c and λ_d, respectively, with $\lambda_c = \lambda' + \lambda''$ and $\lambda_d = \lambda' + \lambda_o$.

Assume that each device awards service on a first-come-first-served basis. Let t_o denote the time required for the 7232 to service a file request and let t_p represent the 7232 service time for a single promotion/demotion request. Furthermore, let t'' represent the CRAM service time of a single page transfer request, and let t' be the CRAM service time for a file promotion or demotion activity. (Here the service time includes both positioning time and data transfer time.) Now let t_c represent the service time of a typical CRAM request and t_d denote the service time of a typical 7232 request. Hence the expectations of t_c and t_d are given by

$$E\left[t_c\right] = \left(\frac{\lambda'}{\lambda'+\lambda''}\right) E\left[t'\right] + \left(\frac{\lambda''}{\lambda'+\lambda''}\right) E\left[t''\right] \tag{15}$$

$$E\left[t_d\right] = \left(\frac{\lambda_o}{\lambda_o+\lambda'}\right) E\left[t_o\right] + \left(\frac{\lambda'}{\lambda_o+\lambda'}\right) E\left[t_p\right] \tag{16}$$

and the loading of the 7232 is

$$p_d = (\lambda_o + \lambda') \; E\left[t_d\right] \tag{17}$$

Similarly if there are c data channels serving the CRAM, the load per channel is

$$p_c = \frac{(\lambda' + \lambda'')\, E\left[t_c\right]}{c} \tag{18}$$

Then reasoning as in Part A, since the service is first-come-first-served and the generation of requests is a Poisson process, the distribution of the waiting time w_d for 7232 service is described by the Laplace-Stieltjes transform

$$W_d(s) = \frac{(1-p_d)\, s}{s - \lambda_d\, (1 - T_d(s))} \tag{19}$$

Thus the response time r_d to file requests stored on the 7232 has mean

$$E\left[r_d\right] = \underbrace{\frac{\lambda_d\, E\left[t_d^{\,2}\right]}{2\,(1 - p_d)}}_{E\left[w_d\right]} + \underbrace{E\left[t_o\right]}_{\substack{\text{Record}\\ \text{access time}}} \tag{20}$$

Similarly the distribution of the waiting time w_c for CRAM service is characterized by

$$W_c\,(s) = \frac{(1 - p_c)\, s}{s - \lambda_c\, (1 - T_c\,(s))/c} \tag{21}$$

whereby the expected response time to a direct file access of a single page record stored on the CRAM is

$$E\left[r_c\right] = \underbrace{\frac{\lambda_c E\left[t_c^{\,2}\right]}{2c\,(1 - p_c)}}_{E\left[w_c\right]} + \underbrace{E\left[t''\right]}_{\text{Single page record access time}} \tag{22}$$

The core memory requirements p_{dc} (in pages) for I/O buffering in the 7232 CRAM system can be estimated by the expected value

$$E\left[p_{dc}\right] = \lambda'' E\left[r_c\right] + \lambda' E\left[n_p\right] \left(E\left[w_c\right] + E\left[t'\right] + E\left[w_d\right] + E\left[t_p\right]\right) + \lambda_o\, E\left[n_f\right] E\left[r_d\right] \tag{23}$$

where $E\left[n_p\right]$ denotes the mean number of pages which are transferred with each file promotion (or demotion) activity, and $E\left[n_f\right]$ is the expected number of pages which are accessed with each file reference (i.e., n_f is a function of record size). Note however that this result does not include the storage required for the I/O software and data bases which are necessary for managing the 7232 RAD and the CRAM.

Another measure of interest is the time f_p required to promote (or demote) a file in its entirety from one device to the other — the CRAM to 7232 (or visa versa). An estimate is provided by

$$E\left[f_p\right] = E\left[w_c\right] + E\left[t'\right] + E\left[w_d\right] + E\left[t_p\right] \tag{24}$$

Equations (20) and (22) provide response measures for the 7232 and CRAM, respectively, which can be compared to the 2314 result given in equation (12). Similarly, equations (14) and (23) can be utilized to compare the core memory requirements for I/O buffering in the two systems (2314 and 7232 CRAM).

III PERFORMANCE ANALYSIS

A. Graphical Results

In order to compare the two file storage design schemes (1) the 2314 and 2) the 7232/CRAM), it is imperative to select reasonable estimates of the variables for each of the mathematical models. This can be achieved, in part, by considering a number of cases for each of the systems as determined by variables such as record size, file size, and positioning time.

First, let us examine the 2314 system for the following four cases:

Case 1 - small records and non-restricted cylinder accessing.
Case 2 - small records and restricted cylinder access.
Case 3 - large records and restricted cylinder access.
Case 4 - large records and non-restricted cylinder accessing.

with

$$E\left[t_1\right] = \begin{cases} 74.25 \text{ ms if non-restricted cylinder access.} \\ 33.65 \text{ ms if restricted cylinder access.} \end{cases}$$

$$E\left[t_2\right] = \begin{cases} 19.5 \text{ ms if small records.} \\ 26.5 \text{ ms if large records.} \end{cases}$$

c = 1 or 2 data channels

where a "small record" is equivalent to one page (512 words) per file request, and a "large record" corresponds to an average of two pages per file request. Here the consideration of different seek positioning times arises because it is conceivable that storage allocation on the 2314 is designed to minimize mechanical motion (e.g., no more than 11 cylinders ($\approx$220 tracks) for 90% of the requests*. Such a situation might result from storage allocation algorithms based upon frequency of file usage and/or file contents and ownership.

In addition to the foregoing, assume the mean file request rate λ varies from 5 to

* Note, however, if limited motion is strictly adhered to, it would effectively reduce the amount of immediately addressable data.

30 requests/sec. (a range corresponding to the TSU estimates verbally obtained from W. Shultz), and let the system consist of 8 independent storage modules (i.e., n = 8) in order to comply with TSU file storage estimates.* It is also assumed that the 2314 storage module operates at 2400 rpm with 7 sectors (256 words each) recorded per track.

Now directing attention to the 7232/CRAM system, let us consider

<u>Case A</u> - small records and small files.
<u>Case B</u> - small records and large files.
<u>Case C</u> - large records and small files.
<u>Case D</u> - large records and large files.

with

$$E\left[t_o\right] = \begin{cases} 22.5 \text{ ms if small records.} \\ 28.4 \text{ ms if large records.} \end{cases}$$

$$E\left[t_p\right] = \begin{cases} 63 \text{ ms if small files.} \\ 109 \text{ ms if large files.} \end{cases}$$

$$E\left[t'\right] = \begin{cases} 515 \text{ ms if small files.} \\ 915 \text{ ms if large files.} \end{cases}$$

$$E\left[t''\right] = 165 \text{ ms.}$$

λ' = 0.5 requests/sec. (promotion/demotion activity).

λ'' = 2 requests/sec. (single page file accesses from CRAM).

c = 1 or 2 data channels.

where a "small file" is of average size 8 pages and a "large file" is of average size 16 pages. Here the attributes of a "small record" and a "large record" are the same as above. The promotion/demotion activity λ' is based upon a 200 user system with an average session time of 20 min/user and 3 file promotion/demotion activities per session. Assume that the CRAM cluster consists of two decks totaling 768 cards (or $\approx 226 \times 10^6$ bytes) in accordance with the minimum TSU

---

* Nominally, 225 mega bytes.

configuration described by Computer Sciences Corporation in their TSU architecture document. (Note this two deck cluster approximately corresponds to the storage capacity provided by the 8 (n = 8) storage modules in the 2314 system.) Again let λ range from 5 to 30 file requests/sec.

Utilizing these parameterizations, system performance is derived and graphically presented in Figures 1-5.

Figures 1 and 2 depict the expected response time $E\left[r_f\right]$ to file requests for a single data channel 2314 system and a two data channel 2314 system, respectively. As indicated, a nominal response time of $\approx$ 60 to 120 ms is obtained for a broad spectrum of the independent variables.

In Figures 3 and 4, the mean core memory requirements are presented. Figure 3 displays page requirements $E\left[p_f\right]$ and $E\left[p_{dc}\right]$ versus λ when a single data channel is used by the file storage device (the 2314 or the CRAM, respectively); whereas Figure 4 displays these requirements when two data channels are used by the file storage device. Note that (barring the case of "small files") the nominal main memory savings ranges from 10 to 25 pages (5K words to 12.5K words) when the 2314 system is used.

The expected response time to file requests for 7232 RAD service in the 7232/CRAM system is shown in Figure 5. Here, the nominal response ranges from 35 to 70 ms for a broad range of variables.

As a final set of results, the expected response time $E\left[r_c\right]$ to file requests for CRAM service and the mean promote (or demote) time $E\left[f_p\right]$ are tabulated in Figure 6 for the spectrum of conditions examined above in each system.

B. <u>Comments and Recommendations</u>

i) Parameter Assumptions and Results

The primary intent of this analysis was to examine the response characteristics

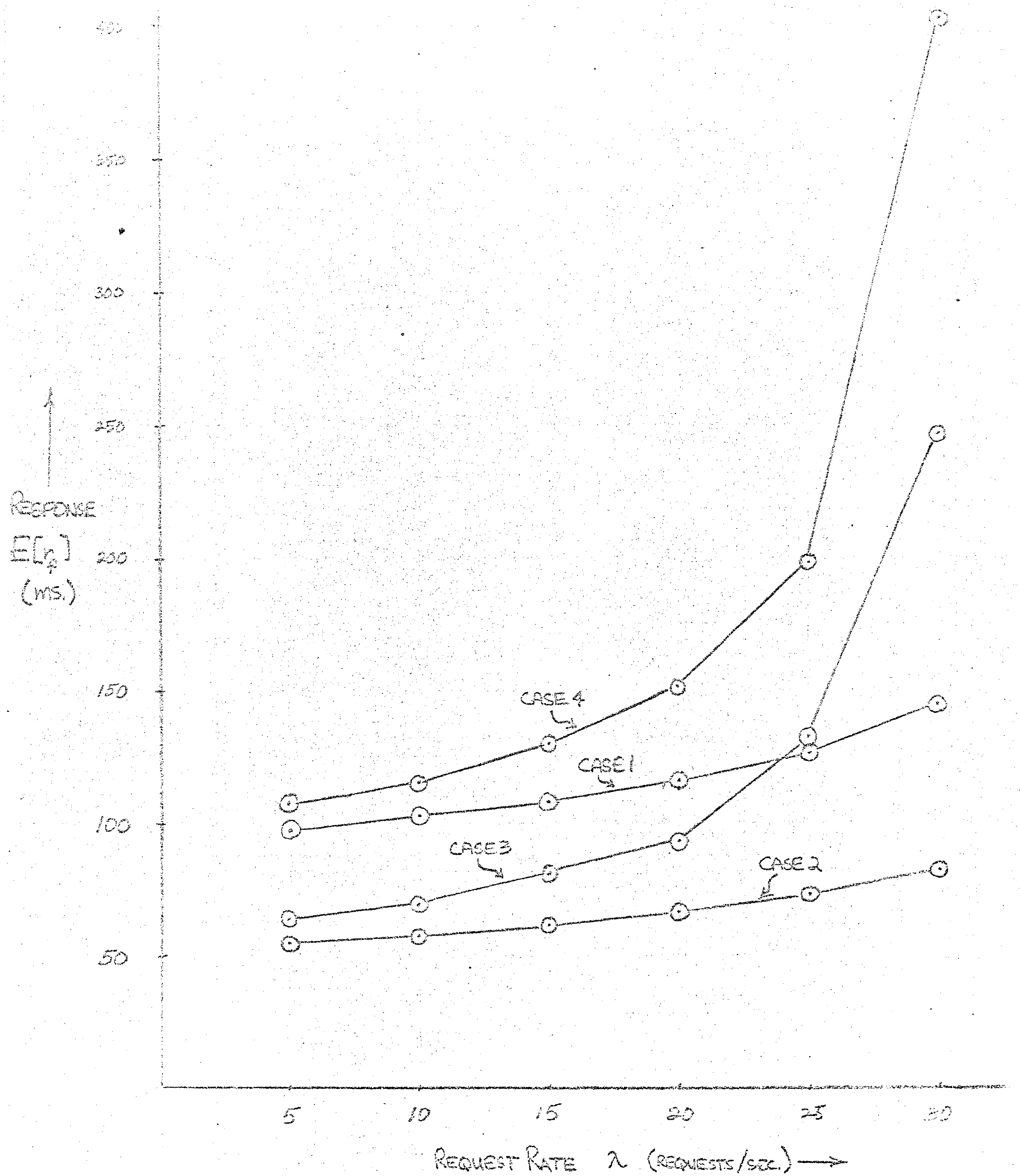

FIGURE 1. EXPECTED RESPONSE TIME FOR 2314 WITH SINGLE DATA CHANNEL

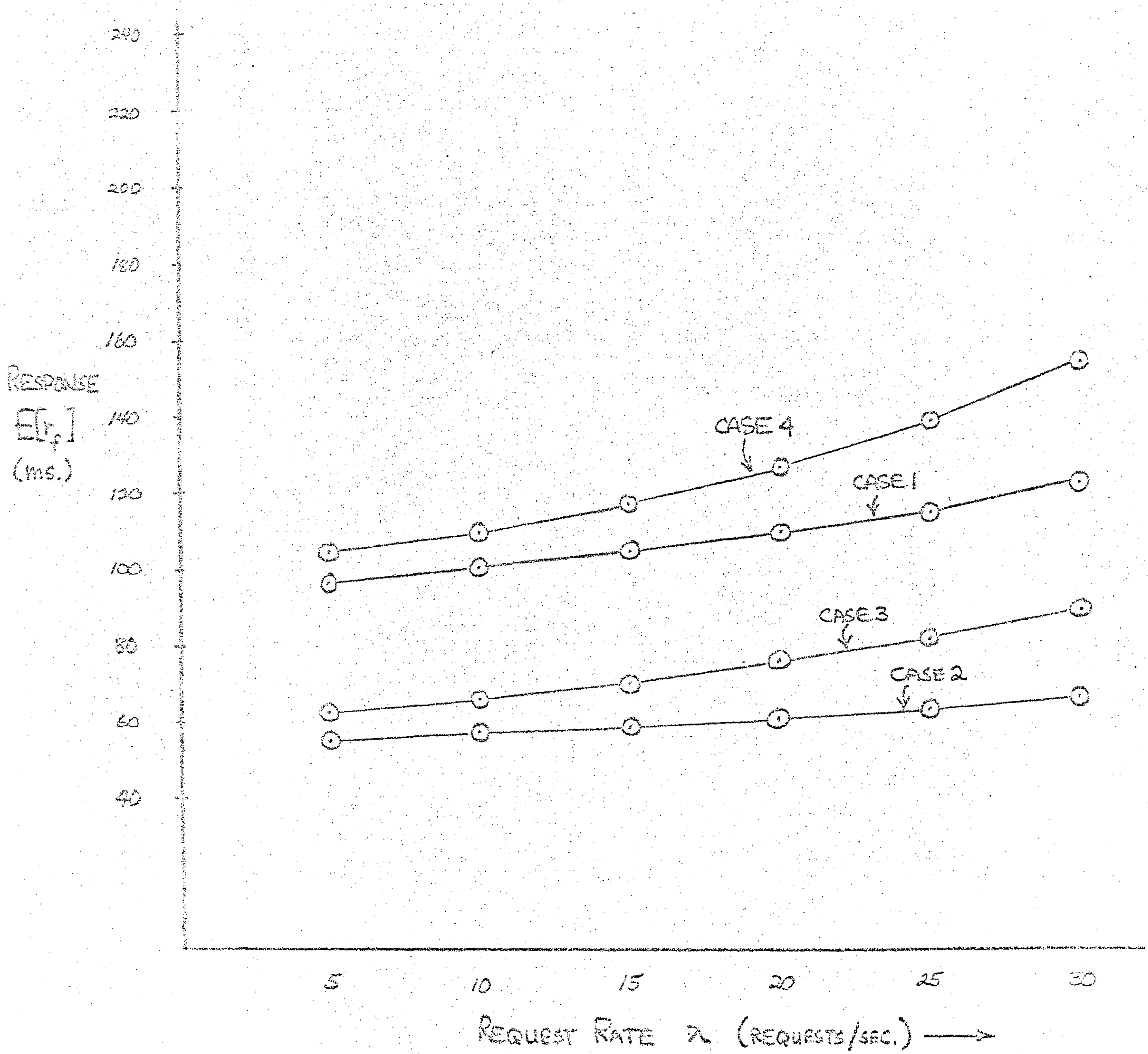

FIGURE 2. EXPECTED RESPONSE TIME FOR 2314 WITH TWO DATA CHANNELS

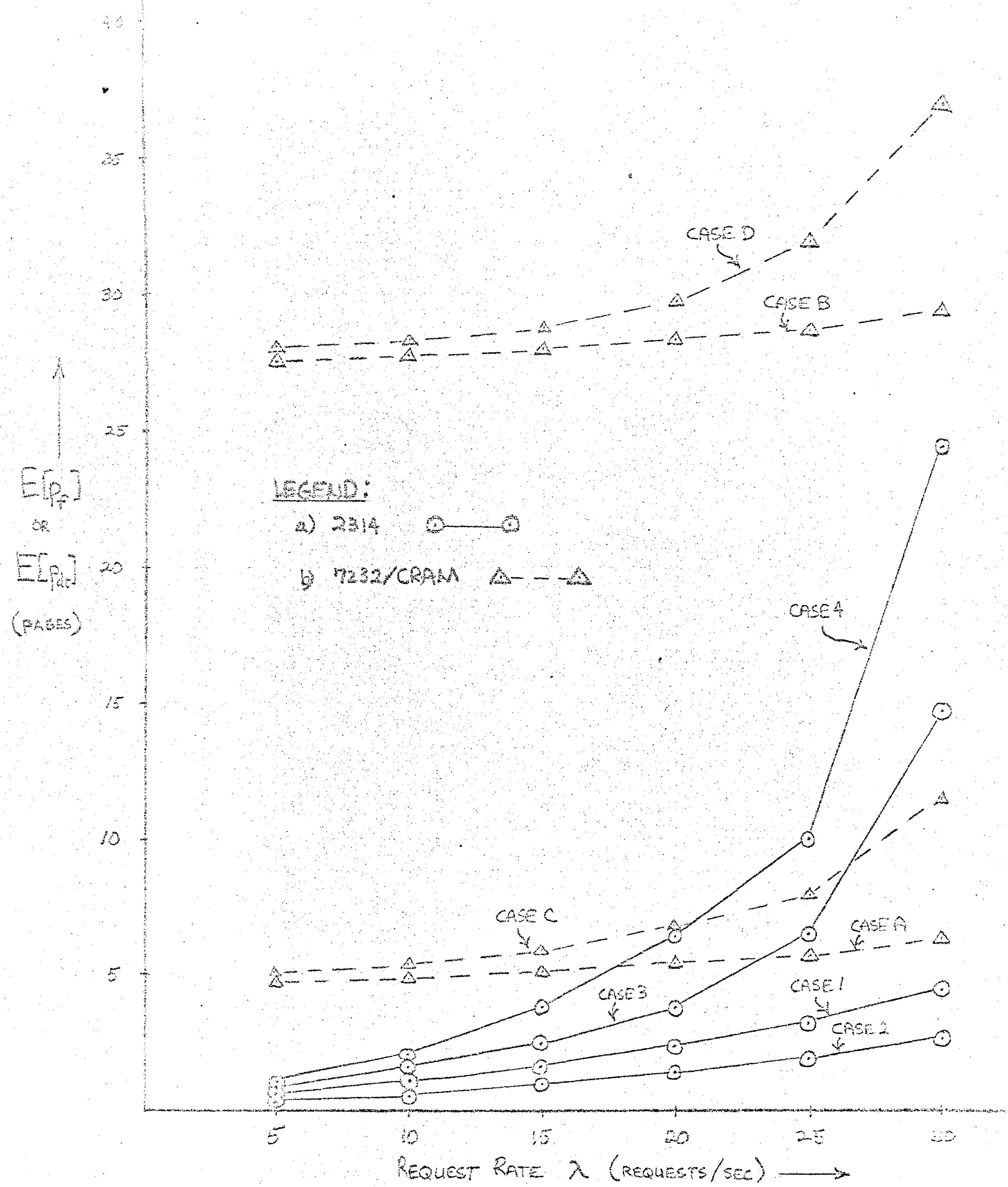

FIGURE 3. EXPECTED CORE MEMORY REQUIREMENTS FOR I/O BUFFERING WHEN SINGLE DATA CHANNEL IS USED BY FILE STORAGE DEVICE

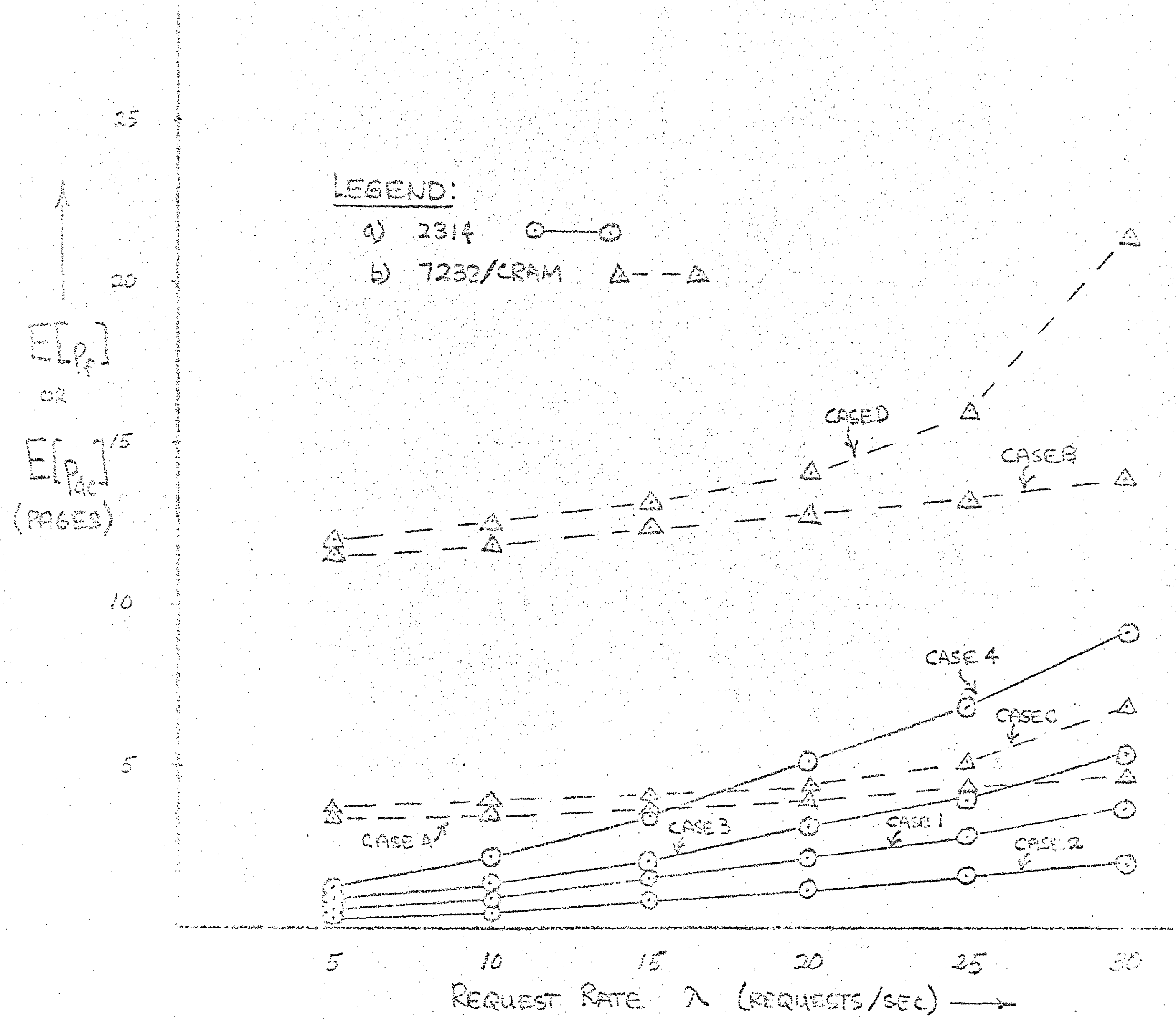

FIGURE 4. EXPECTED CORE MEMORY REQUIREMENTS FOR I/O BUFFERING WHEN TWO DATA CHANNELS ARE USED BY FILE STORAGE DEVICE

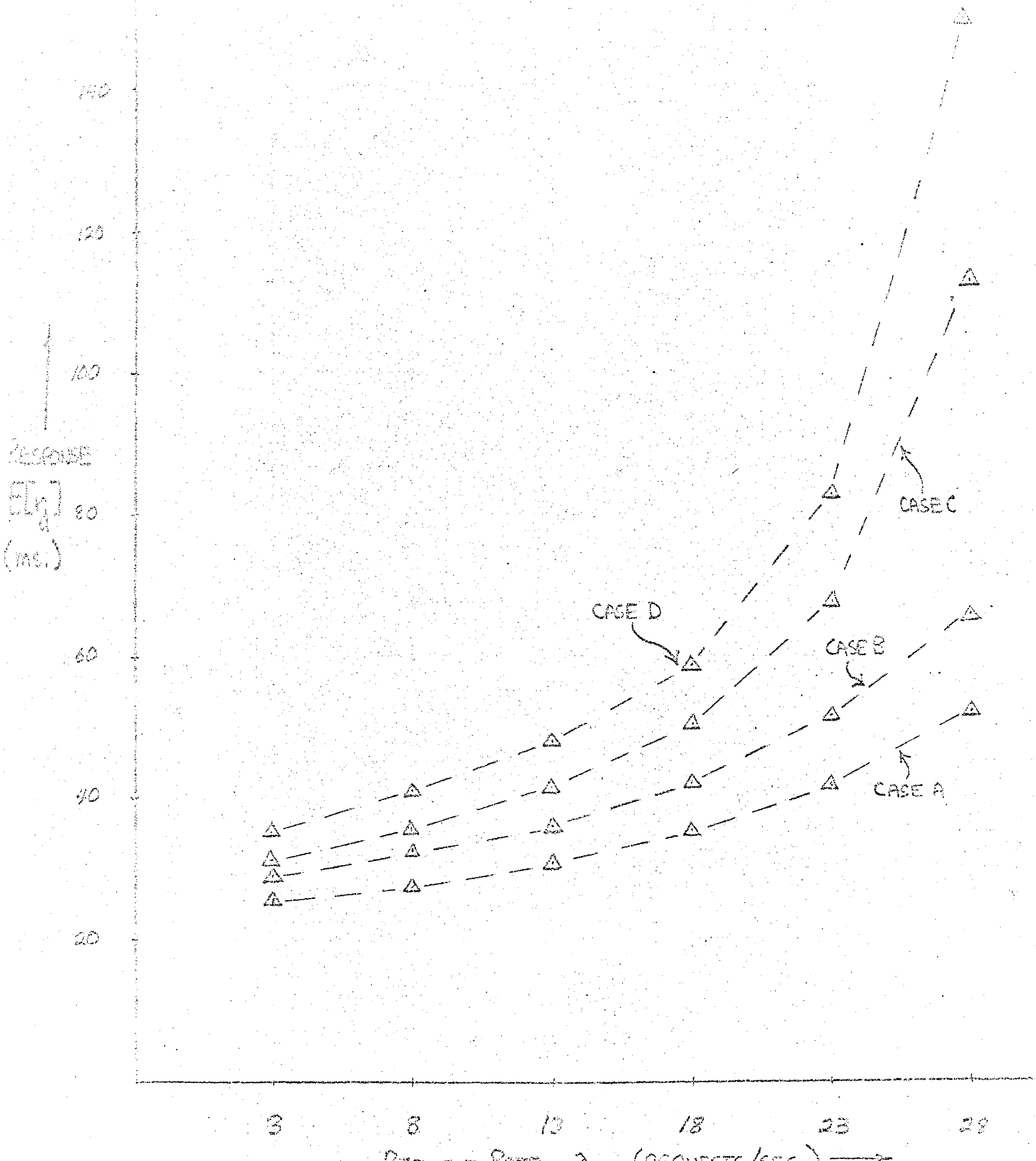

FIGURE 5 EXPECTED RESPONSE TIME FOR FILE REQUESTS FROM A 7232 RAD (NOTE HERE THE 7232 RAD FILE REQUEST RATE IS $\lambda = \lambda - \lambda'' = \lambda - 2$)

FIGURE 6. TABULAR PRESENTATION OF MENU CRAM RESPONSE AND PROMOTION (DEMOTION) TIME

	CRAM FILE REQUEST RESPONSE $E[r_c]$		7232/CRAM PROMOTION (DEMOTION) TIME $E[f_p]$			
	CASES A or C	CASES B or D	CASE A	CASE B	CASE C	CASE D
SINGLE DATA CHANNEL	498 ms.	2050 ms.	≈ 920 ms.	≈ 2920 ms.	≈ 930 ms.	≈ 2930 ms.
TWO DATA CHANNELS FOR CRAM	262 ms.	495 ms.	≈ 685 ms.	≈ 1370 ms.	≈ 695 ms.	≈ 1400 ms.

of file storage hardware for two different implementation schemes in the TSU system.* Here, the attempt was made to provide reasonable approximations of the independent variables as construed from current estimations of the TSU environment. Thus as we acquire more knowledge about TSU, other ranges of variables may prove to be relevant for a study of this sort.

On the surface, it appears as though file response in the 7232/CRAM system is better than that provided by the 2314 system (see Figures 1, 2, and 5). Yet it should be pointed out that a portion of file requests (those directly referencing the CRAM with rate λ'') experience long response times as shown in Figure 6. To further discount this performance difference, it should also be noted that the response time estimates for the 7232/CRAM system are biased towards the "best case" since it was assumed:

a) files were mapped into contiguous sectors on the 7232 RAD

b) a file resided in its entirety upon one cylinder of a single CRAM card

c) the time required by the executive to accomplish file response (interrupt handling, dictionary searches, etc.) was negligible

d) the total file request λ of the 2314 system directly translated into a corresponding $\lambda_o + \lambda''$ in the 7232/CRAM system (i.e., $\lambda = \lambda_o + \lambda''$).

In regard to the latter of these, I would suspect that identical user environments would produce different total file request rates in each system due to added requests for maintaining and utilizing device maps, dictionaries and directories in the 7232/CRAM system.

Along this same line of reasoning, the file promotion (demotion) rate λ' and the file size significantly affect the 7232/CRAM performance. Assuming the same set of variables used in Section III.A of this study, if λ' were to

* Note however that the emphasis of this study to TSU does not preclude the application of the mathematical models presented here to performance investigation in other system environments.

increase to 1 request/sec., the CRAM with a single data channel would rapidly become overloaded, thereby dramatically increasing response times. Moreover, it is obvious that file promotion (demotion) activity could easily produce additional 7232 RAD traffic (extended servicing periods) if it became a necessity to "write check" all write operations.

In short, one could easily conceive of a variation in parameters which would degrade the response characteristics of the 7232/CRAM system. In contrast, it is difficult to produce the corresponding result in the 2314 system so long as the records are typically 1 or 2 pages per file request and the file request rate λ is less than 30 requests/sec.

However, in defense of the CRAM, the 7232/CRAM system provides up to $\approx$ 888 mega-bytes of file storage capacity since six additional decks could be added to the two deck CRAM cluster with all eight utilizing the same controller; whereas if additional storage (above the $\approx$225 mega-bytes provided by the eight modules) is required in the 2314 system, then another controller must be added for each group of eight storage modules. This poses an attendant problem of this analysis — how much file storage is required in the TSU environment? Note however that 225 mega-bytes provides file storage sufficient for over 2000 users, each with files of 100K bytes.

ii) Reliability and Cost

Apart from marketing considerations of potential sales, there are at least two other major areas of system design which this analysis did not examine — 1) system reliability and 2) the total cost of each storage system. Since these latter considerations cannot be divorced from design decisions, it is appropriate to comment on these.

a) Reliability: In time-sharing environments, one key to system efficiency is a reliable and rapid access mass storage system. This memo and others have examined system response; yet to the best of my knowledge, no one at SDS has examined the CRAM reliability. Storage devices such as the CRAM are referred to by some engineers in the industry as "mechanical monsters." This label is due (in part) to the inherent reliability problems of their mechanisms and their stringent maintenance requirements.

If TSU is implemented with 7232/CRAM file storage, I won't be suprised if system reliability is a direct reflection of CRAM reliability, independent of the reliability of other devices comprising the system! In comparing the reliability of the two file storage systems, I would intuitively expect that the mean time before failure of the 2314 storage is at least a factor of 2 longer than that of the 7232/CRAM storage.

It suffices to emphasize that reliability considerations must not be overlooked, deferred, or taken lightly — we must look further into this area, NOW! In this respect, it would certainly be worthwhile to sample NCR's CRAM customers.

b) Cost: It is difficult to project cost estimates of the file storage systems considered here, not only because some of the sub-systems are not yet engineered, but also because the cost of such storage depends upon the system's storage capacity*. The tables below serve to provide a cursory comparison of the two systems. (Note the inclusion of the additional core memory required in the 7232/CRAM system for I/O buffering - see Section III.A.)

* The range and flexibility of CRAM capacity are undoubtedly its most desirable attributes.

7232/CRAM System

Device	Cost (\$ x 10^3)
One 7231 controller (with 4 byte interface)	≈ 20
Two 7232 storage units	≈ 80
Two deck CRAM	≈ 120
CRAM controller (single data channel)	≈ 25
≈ 8K core memory	≈ 50
Total:	≈ \$295 x 10^3

2314 System

Device	Cost (\$ x 10^3)
Eight storage modules	≈ 280
Controller (single data channel)	≈ 25
Total:	≈ \$305 x 10^3

Admittedly, the above estimates are subject to closer scrutiny; however, what is important to note is the relative cost differential between the two systems. On the surface, it appears as though both hardware configurations are about the same price.*

* These cost estimates were obtained in an informal discussion with Product Planning for purposes of comparison and are <u>not</u> to be assumed as specific price quotes.

iii) Recommendations

If these conclusions are correct, then I strongly favor the 2314 for TSU (barring adverse reliability performance) since for the variables chosen, the response characteristics of the two systems are not substantially different; and moreover, the above somewhat superficial cost estimates do not include the cost of core memory required by the supporting software and its data bases, nor does it account for the cost of software design and development.

One should not under estimate the design effort (and cost) necessary to build and maintain a file system with the promotion (demotion) structure currently envisioned for TSU with the 7232/CRAM; nor should one base the operational efficiency of a time-sharing system upon a basically unreliable mechanical device.

J. E. Shemer Sept. 16, 1968

J. E. Shemer

bp

cc: A. Bongarzone, G. Boyd, E. Bryan, D. Cota, F. DeKalb, V. DeVine, B. Doeppel, P. England, B. Gable, H. Johnson, E. Maso, D. McGurk, J. Mendelson, M. Phister, A. Rosenberg, E. Shifflett, W. Shultz, J. Smith, R. Spinrad, D. Stein, S. Vorgitch, S. Zasloff