Performance of Movable-Head Disk Storage Devices

C. C. GOTLIEB AND G. H. MACEWEN

University of Toronto, Toronto, Ontario, Canada

ABSTRACT. A queueing model of movable-head disk storage systems is developed so that the performance, as measured by the mean response time, can be calculated. Queue scheduling algorithms which improve the performance are considered. Single-module disk systems are analyzed, incorporating the SCAN scheduling algorithm suggested by Denning so that comparisons with the FIFO algorithm are possible. This analysis is extended to multimodule systems whereby tables of approximate mean response values time can be calculated over system parameters describing equipment characteristics, equipment configuration, system loading, file organization, and scheduling algorithm (SCAN or FIFO). The use of such tables is discussed and the applicability of the analysis to a recently marketed disk is noted.

KEY WORDS AND PHRASES: secondary storage, disk storage, queueing model, queueing analysis, scheduling, file organization, input/output, state-dependent queues, imbedded Markov chain, steady state analysis, system design

CR CATEGORIES: 4.3, 4.41, 6.2, 6.35

1. *Introduction*

Movable-head disk storage devices (simply called disks in the remainder of the paper) are commonly used for *secondary storage*, that is, for the storage of programs and data that are required relatively infrequently but which may be needed at any time. They are attractive because of their size, speed, and cost, as compared with alternatives such as magnetic tape or card devices. Many general purpose computing systems, as well as special purpose systems such as message switching systems, reservation systems, and information retrieval systems, use disks.

The overall processing efficiency and performance of computer systems depend critically on the *response time* of secondary storage [2, 3], the time from the initiation of a *data transfer request* to the completion of the transfer operation. Disk devices have been evaluated from varying viewpoints, and with varying techniques [4–9]; the *performance measure* chosen reflects the approach taken in the evaluation. Fife and Smith [4] chose the performance measure of maximum mean data transfer rate to analyze multichannel disk configurations. This led them to assume full load conditions (that is, the queue of requests is never empty) in their analysis. A more common measure is mean response time [5–8] since it does not imply any such restriction.

Different ways of scheduling data transfer requests have been suggested [6, 8, 9] as a means of reducing mean response time. In particular, two scheduling algorithms are simple enough and yet potentially effective enough to be considered as alternatives to the FIFO (first-in-first-out) queue discipline which is generally adopted. The first,

Authors' present addresses: C. C. Gotlieb, Department of Computer Science, University of Toronto, Toronto M5S 1A7, Ontario, Canada; G. H. MacEwen, Department of Computing and Information Science, Queen's University, Kingston, Ontario, Canada. This paper is based on the Ph.D. thesis of G. H. MacEwen [1].

SCAN,[1] operates by sweeping the access arm in one direction across the disk, always moving from the cylinder[2] last involved in a data transfer to the closest cylinder (in the direction the arm is being swept) for which a data transfer request is waiting in the queue; if no such request exists, the direction of sweeping is reversed. The second scheduling algorithm, SSTF (shortest-seek-time-first), operates by simply moving the arm to the closest cylinder for which a data transfer request is waiting in the queue, regardless of the direction.

Queueing theory is useful to represent the congestion of data transfer requests occurring in a heavily loaded disk system; no queueing model for multimodule[3] movable-head disks has previously been developed which incorporates a scheduling algorithm other than FIFO. In this paper, disk storage performance is analyzed by using queueing theory to calculate the mean response time over a range of parameters which determine the characteristics of the disk storage system.

A queueing model of a single-channel disk system is first developed so that mean response time, $E[T]$, can be calculated as a function of:

(a) equipment characteristics;

(b) equipment configuration (the number of disk modules is not restricted to large values, as in [7], but is variable, as in [5], up to certain limits imposed by practical rather than theoretical reasons);

(c) file organization—block lengths are taken to be uniformly distributed over a fixed range and Seaman's method [5] of channel activity analysis is generalized to incorporate this distribution;

(d) system loading;

(e) request scheduling algorithm—SCAN or FIFO. SCAN scheduling, rather than SSTF, is chosen for detailed attention on the basis of system equity; SSTF discriminates against requests accessing certain cylinders. Preliminary simulations showed that although SSTF yields lower values of $E[T]$ than SCAN, most of the improvement over FIFO gained with SSTF is obtainable with SCAN.

$E[T]$ is not obtained in a closed form, but a procedure is given whereby tables of approximate values of $E[T]$ are calculated and typical curves are produced. General conclusions about disk systems, based on calculated results, are drawn. In addition, the use of such tables of $E[T]$ for engineering design is discussed. Finally, the relevance of the work to newer types of disks is discussed.

It should be added that a mathematical analysis of an idealized single-module disk model has recently been described [23]. This analysis examines two different versions of SCAN: one in which requests arriving during a sweep are serviced in the following sweep, and one in which each sweep is performed in the same direction.

2. *The Disk Model*

The type of device under study is typified by the IBM 2314 [10] which is used for illustration in the remainder of this paper. The timing sequence occurring as a data transfer request[4] is satisfied, if software and hardware overhead times are ignored, is described by Figure 1. The data channel is needed to initiate arm motion. It can be shown [1] that the time required to wait for the channel to become free to initiate arm motion may be ignored since its mean value is small compared to the times shown in Figure 1.

The disk model is a set of queues as shown in Figure 2. Each distinct module queue is considered, since one objective is to determine the response time for a varying number

[1] SCAN, as described here, is called LOOK in [8].

[2] A *cylinder* is a set of tracks (one or more per disk surface) written onto while the access arm is stationary at one of a set of discrete mechanical positions to which it can be moved.

[3] A disk module contains one independently movable access arm.

[4] A *request*, generated by any process in the system, is received by the Input/Output Supervisor which manages the data transfer.

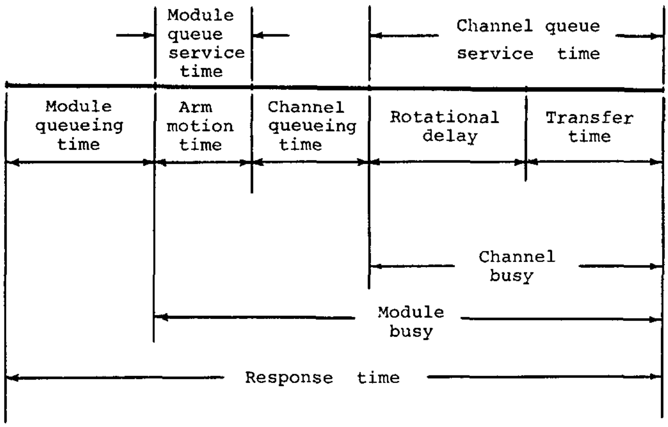

FIG. 1. Timing diagram for disk model

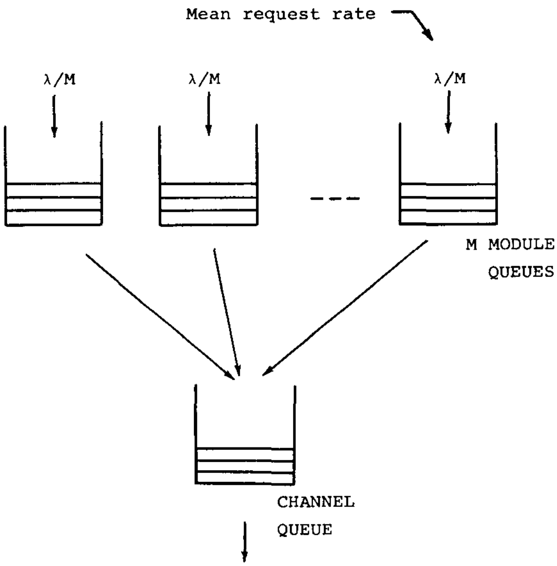

FIG. 2. Queueing model of single-channel disk

of modules. Arm motion time constitutes the module queue service time, and rotational delay plus block transfer time constitute the channel queue service time.[5] Note from Figure 1 that a module queue server remains busy as long as the request it last served occupies the channel queue. It is necessary, therefore, to specify that a module queue server remains busy after serving a request until that request has completed its channel service time.

The following assumptions complete the specification of the model.

(a) Requests are independent and arrive randomly in time at the queues, each module being equally likely to be addressed. The interarrival distribution function is given by

$$F_e(t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0, \end{cases}$$

where λ is the mean request rate. The interarrival time distribution of requests to each module queue is also exponential with a mean request rate of λ/M.

[5] *Blocks* of data are recorded on circular tracks which rotate under read/write heads mounted on the access arm.

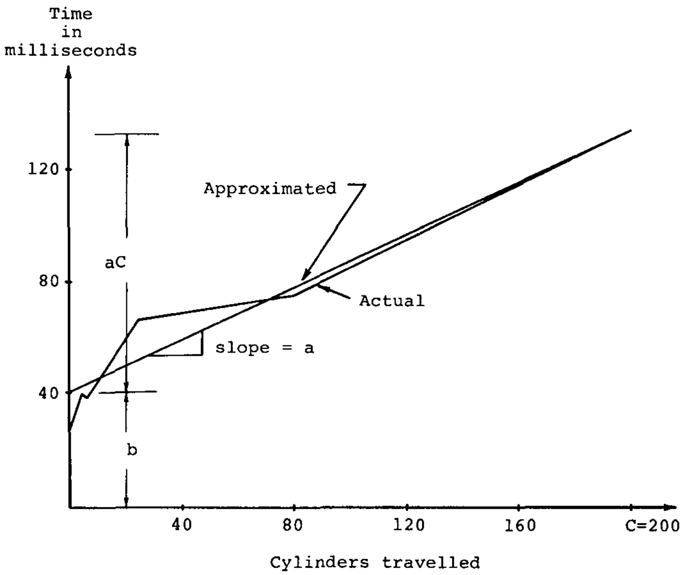

FIG. 3. IBM 2314 arm movement time

The exponential distribution is chosen for two reasons. First, the assumption of random arrivals is the safest one in the absence of other information about the interarrival distribution. (Information about arrival distributions of disk requests in actual systems is needed, and could be provided by a study analogous to that of Fuchs and Jackson [11] on computer communications traffic.) Second, a large body of useful queueing theory becomes available if one assumes an exponential distribution for interarrival times.

(b) The plot of arm motion distance against time is approximated by a straight line as shown in Figure 3. C is the number of cylinders and a and b are determined by fitting a straight line to the actual module characteristics, shown for the 2314.

(c) The channel queue scheduling algorithm is FIFO.

(d) The channel queue service time consists of the rotational delay plus the transfer time. The distribution of this random variable depends on the manner in which blocks are stored on the disk and the lengths of the blocks. (For example, some disks store blocks on fixed length physical records called sectors. The waiting time for a randomly chosen sector to rotate under the heads is uniformly distributed between 0 and r, the total time to complete one full rotation. If block lengths, hence block transfer times, are uniformly distributed between 0 and B, the channel queue service time to transfer a randomly chosen block is given by a triangular distribution, nonzero between 0 and $r + B$.) To approximate a variety of situations and for ease of analysis, channel queue service time is taken to be uniformly distributed between 0 and $r + B$. The service time distribution function is, therefore,

$$F_c(t) = \begin{cases} 0, & t < 0, \\ t/R, & 0 \le t < R, \\ 1, & t \ge R, \end{cases}$$

where $R = r + B$.

(e) Module queue scheduling is either FIFO or SCAN.

(f) Cylinder addresses, which have integer values between 1 and C, are approximated by real numbers in the interval $(0, C)$. A cylinder address is associated with each request. Each address is a random variable[6] $\mathbf{z}$ distributed according to a given continuous probability density function, $f_\mathbf{z}(z)$. The function $f_\mathbf{z}(z)$ is included in the model in an attempt to take into account the way in which disk storage space is allocated in a given

[6] Boldface letters denote random variables.

application. For this work, $f_z(z)$ is taken as the beta density function, given by

$$f_z(z) = \begin{cases} 0, & z < C, \\ \beta\binom{2\beta - 1}{\beta - 1}(z/C)^{\beta-1}(1 - z/C)^{\beta-1}, & 0 \leq z \leq C, \\ 0, & z > C, \end{cases}$$

where $1 \leq \beta < \infty$. The beta density function is uniform when $\beta = 1$ and becomes increasingly more peaked as β increases. Increasing β, then, models the effect of grouping together the most-used files on contiguous cylinders.

(g) The cylinder address of each request in the queue is independent of the addresses of other requests in the queue. This assumption is valid for FIFO scheduling since arriving requests are independent and their ordering in the queue does not change. For SCAN scheduling, however, requests are selected for service on the basis of their cylinder addresses. After such a selective procedure, the remaining requests in the queue are no longer independent. This assumption is discussed for SCAN scheduling in [1], where it is concluded that since the consequences of the assumption are negligible they can be ignored. In addition, simulation experiments have been conducted to verify the following analysis which is based partially on this assumption. Simulation results are presented in Section 3.

To summarize, the parameters of the disk model are:
(a) a—the slope of the approximated arm motion time characteristics,
(b) b—the fixed component of arm motion time,
(c) r—the time to complete one rotation,
(d) B—the maximum block transfer time,
(e) M—the number of modules,
(f) C—the number of cylinders,
(g) β—the beta distribution parameter,
(h) λ—the mean request rate,
(i) the module queue scheduling algorithm.

Assumption (g) deserves one further comment. SCAN discriminates to some extent against requests to the extreme cylinders; two full sweeps occur between servicings of an extreme cylinder, while only two half-sweeps occur between servicings of a central cylinder. Consequently there is a tendency for requests to bunch at the extreme cylinders more than is actually indicated by $f_z(z)$. It is this bunching that is ignored by assumption (g).

3. *The Single-Module System*

Two characteristics of the disk model introduce analytical complications:

(a) interdependence of requests arriving at the channel queue—requests leaving the module queues and arriving at the channel queue are no longer independent.

(b) module-channel queue interaction—a module queue server is kept busy until its last served request has left the channel queue. The operation of the module queues is therefore dependent on the operation of the channel queue.

A disk with only one module can be modeled by a single queue which combines the module and channel queue (see Figure 4), thus avoiding the complications resulting from the two characteristics noted above. For this reason, the single-module device is evaluated first. Analytical results obtained are then extended to the multimodule case.

Four analysis procedures are presented:

	β	*Scheduling algorithm*
(A)	1	SCAN
(B)	1	FIFO
(C)	>1	SCAN
(D)	>1	FIFO

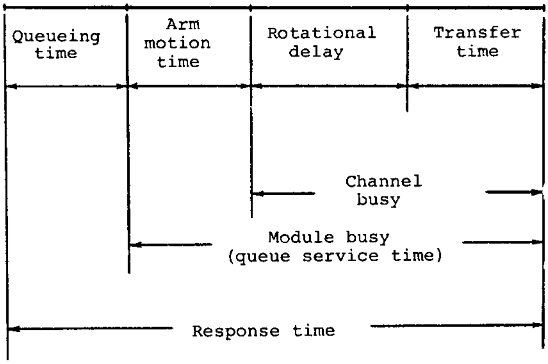

FIG. 4. Timing diagram for single-module disk model

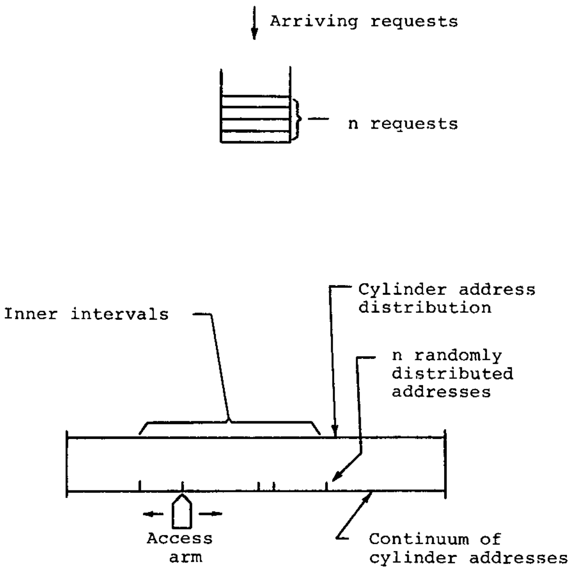

FIG. 5. The single-module disk model with uniform address distribution

(A) SCAN SCHEDULING WITH UNIFORM CYLINDER ADDRESS DISTRIBUTION ($\beta = 1$). Consider the sequence of discrete times at which service periods[7] (accesses) start. At each of these starting times the SCAN scheduling algorithm is applied and a request is selected from the queue for service. Let the number of requests in the queue at any starting time be n. From assumption (g), the n requests have n associated cylinder addresses which can be treated as a random sample of size n chosen from a population with uniform probability distribution between 0 and c. There are, then, n randomly distributed cylinder addresses on the interval $(0, C)$.

The position of the access arm at any service starting time is, *by assumption* (g), at a random point on the interval $(0, C)$ (see Figure 5). This is so because the requests in the queue at the *previous* service starting time had a set of cylinder addresses uniformly distributed on the interval $(0, C)$.

The arm position and the cylinder addresses define $n + 1$ random points on the interval $(0, C)$. It can be shown [12, p. 21] that $n + 1$ random points on the interval $(0, C)$ create $n + 2$ intervals whose lengths have the common distribution.

[7] The term service *period* is used in this section to avoid confusion with the service starting *time*.

$$\text{Prob [interval length} > L] = (1 - L/C)^{n+1}, \qquad 0 < L < C. \tag{1}$$

The distance that the arm will move to service the request selected from the queue by the SCAN algorithm is given by one of the inner n intervals (see Figure 5). This is true regardless of the direction of SCAN or whether the arm continues its SCAN direction or reverses it. The distribution of the arm movement distance $\mathbf{s}$ is therefore

$$F_{\mathbf{s}|n}(s|n) = \begin{cases} 0, & s \leq 0, \\ 1 - (1 - s/C)^{n+1}, & 0 < s < C, \\ 1, & s \geq C. \end{cases} \tag{2}$$

Note that the distribution is dependent on n, the number of requests in the queue when the request being serviced is selected.

An expression for the service period distribution is now required. The service period can be broken down into three components:

(a) the fixed portion b of arm movement time;

(b) the variable portion $\mathbf{x}$ of arm movement time;

(c) the time $\mathbf{c}$ to position and transmit the block (the channel service time).

$\mathbf{x}$ is a random variable distributed according to the function

$$F_{\mathbf{x}|n}(x \mid n) = \begin{cases} 0, & x \leq 0, \\ 1 - (1 - x/aC)^{n+1}, & 0 < x < aC, \\ 1, & x \geq aC, \end{cases}$$

where a is the slope of the arm movement time characteristic. $\mathbf{c}$, according to assumption (d), is a random variable distributed according to the function

$$F_{\mathbf{c}}(c) = \begin{cases} 0, & c \leq 0, \\ c/R, & 0 < c < R, \\ 1, & c \geq R, \end{cases}$$

where R is the maximum rotational plus transfer time.

Therefore the service period is a random variable, $\mathbf{t} = b + \mathbf{c} + \mathbf{x}$, whose distribution $F_{\mathbf{t}}(t)$ is found by taking the convolution of the three component distributions. Thus

$$F_{\mathbf{t}|n}(t \mid n) = F_b(t) * F_{\mathbf{c}}(t) * F_{\mathbf{x}|n}(t \mid n),$$

which yields [1]

$$F_{\mathbf{t}|n}(t|n) = \frac{aC}{R} g[(t - b)/aC]u[(t - b)/aC] - \frac{aC}{R} g[(t - b - R)/aC]u[(t - b - R)/aC], \tag{3}$$

where

$$g(t) = \begin{cases} t + (1 - t)^{n+2}(n + 2)^{-1} - (n + 2)^{-1}, & 0 < t \leq 1, \\ t - (n + 2)^{-1}, & t > 1, \end{cases} \tag{4}$$

and

$$u(t) = \text{the unit step function} = \begin{cases} 0, & t \leq 0, \\ 1, & t > 0. \end{cases}$$

As expected, the mean service period decreases for higher values of n, the number of requests in the queue when service commences. A queue in which the service period depends upon the queue length is called state dependent [13, 14].

The queue is now completely specified and, in the queueing theory notation of Kendall, may be described as an M/G/1: (∞, SCAN) queue [15].

State dependent queues with an infinitely large family of general service period distributions have not been treated in the literature. However, the case of exponentially distributed service periods has been examined by Harris [13, 14]. The imbedded Markov chain (IMC) approach provides the basis for Harris's method of solution. The states of the IMC of a queue are given by $S_c = \{0, 1, 2, 3, \cdots\}$ and represent the number of requests in the queue including the request undergoing service. The IMC changes state at discrete points in time, $T_c = \{t_1, t_2, t_3, \cdots\}$, where t_i is the time immediately following the completion of service of the ith request to receive service. This homogeneous IMC is completely specified by its transition probability matrix,

$$P = \begin{Vmatrix} k_{01} & k_{11} & k_{21} & k_{31} & k_{41} & k_{51} & \cdots \\ k_{01} & k_{11} & k_{21} & k_{31} & k_{41} & k_{51} & \cdots \\ 0 & k_{02} & k_{12} & k_{22} & k_{32} & k_{42} & \cdots \\ 0 & 0 & k_{03} & k_{13} & k_{23} & k_{33} & \cdots \\ 0 & 0 & 0 & k_{04} & k_{14} & k_{24} & \cdots \\ 0 & 0 & 0 & 0 & k_{05} & k_{15} & \cdots \\ 0 & 0 & 0 & 0 & 0 & k_{06} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \end{Vmatrix},$$

the elements of which are

$$P_{ij} = \text{Prob [the IMC is in state } j \text{ at time } t_n \mid \text{it was in state } i \text{ at time } t_{n-1}].$$

It can be seen that

$$P_{ij} = \begin{cases} k_{j-i+1,i}, & i \geq 1, \\ k_{j,1}, & i = 0, \end{cases}$$

where

k_{ij} = Prob [i requests arrive at the queue during a service period | there were j requests in the queue when service began].

From the assumption of exponentially distributed interarrival times, it follows that

$$k_{ij} = 1/i! \int_0^\infty (\lambda t)^i e^{-\lambda t}\, dF_{t|j}(t \mid j). \tag{5}$$

Before proceeding, it is necessary to prove that the IMC is ergodic,[8] i.e. that steady state probabilities for queue length exist. The required proof for a general service period distribution, given in [1], is a modified form of a proof for exponentially distributed service periods in [13]. The conditions for the existence of steady state are:

$$F_{t|n}(t \mid n) \text{ has finite mean } E[t \mid n] = \tau_n, \text{ such that } \lambda\tau_j < 1, \qquad j \geq 1.$$

The method of determining the steady state probabilities of an ergodic Markov chain, used by Harris and followed here, requires an analytic expression for the k_{ij}'s. First the density function of the service period, $f_{t|n}(t \mid n)$, is obtained by differentiating the function $g(t)$ given by (4). Thus

$$h(t) = g'(t) = \begin{cases} 1 - (1-t)^{n+1}, & 0 < t \leq 1, \\ 1, & t > 1, \end{cases}$$

and

$$f_{t|n}(t \mid n) = (aC/R)h[(t-b)/aC]u[(t-b)/aC] - (aC/R)h[(t-b-R)/aC]u[(t-b-R)/aC]. \tag{6}$$

[8] Loosely speaking, an ergodic Markov chain is one in which (a) every state can be reached from every other state, and (b) the mean time between occupancies of a state is finite for all states. See [14 and 16] for a more rigorous definition.

Substituting $f_{t|j}(t \mid j)\, dt$ for $dF_{t|j}(t \mid j)$ in (5) and integrating, one obtains

$$\begin{aligned} k_{ij} = {} & (\lambda^i/i!\, R(aC)^{j+1}) \sum_{n=0}^{j+1} \sum_{k=0}^{n+i} \binom{j+1}{n} \binom{n+i}{k} (k!\, (-1)^n/\lambda^{k+1}) \\ & \cdot [(aC + b + R)^{j+1-n}(b + R)^{n+i-k} e^{-\lambda}(b + R) \\ & - (aC + b)^{j+1-n} b^{n+i-k} e^{-\lambda b} + (aC + b)^{j+1+i-k} e^{-\lambda(aC+b)} \\ & - (aC + b + R)^{j+1+i-k} e^{-\lambda(aC+b+R)}] \\ & + (\lambda^i/i!\, R(aC)^{j+1}) \sum_{k=0}^{i} \binom{i}{k} (k!\, (aC)^{j+1}/\lambda^{k+1})[e^{-\lambda b} b^{i-k} - e^{-\lambda(b+R)}(b + R)^{i-k}]. \end{aligned} \tag{7}$$

Note that the derivation of (7) has assumed that $R < aC$. This is valid for most disks including those used as examples in the thesis. It is also the condition of most interest since SCAN attempts to reduce the variable portion of arm movement time; SCAN scheduling, then, is more effective for disks in which this time is large in comparison to the rotational times. This condition may not be valid in general.

Having specified the transition matrix P_{ij}, and having proved that the IMC is ergodic, the steady state queue length probabilities can be determined. Let

π_i = Prob [there are i requests in the queue immediately following a completion of service],

p_i = Prob [there are i requests in the queue, including the one undergoing service, at any random time].

The sequence of probabilities $\{\pi_i\}$ can be found [16, p. 249] from the set of linear equations

$$\pi_j = \sum_{i=0}^{\infty} p_{ij}\pi_i, \tag{8}$$

which hold for an ergodic Markov chain; iterating on the equations (8), one can determine π_i, $i = 1, 2, 3, \cdots, k$, in terms of π_0 where all π_i, $i > k$, are negligible. Such a k always exists for a queue with an ergodic IMC since the steady state mean queue length is finite. π_0 can be determined from

$$\sum_{i=0}^{\infty} \pi_i = 1 \simeq \sum_{i=0}^{k} \pi_i. \tag{9}$$

Since the object of the analysis is to determine the steady state behavior of the queue at *any* random time (as opposed to service completion times), the $\{p_i\}$ are required; unfortunately, they are much more difficult to find than the $\{\pi_i\}$. It is assumed[9] for the remainder of this work that $p_i = \pi_i$, $i = 0, 1, 2, 3, \cdots$. From (8), $\pi_0 = p_{00}\pi_0 + p_{10}\pi_1$. Therefore,

$$\pi_1 = \pi_0[(1 - k_{01})/k_{01}], \qquad \pi_2 = \pi_0[(1 - k_{01})(1 - k_{11})/k_{01}^2 - k_{11}/k_{01}],$$

and similarly all the π_i, $i = 1, 2, 3, \cdots, k$, can be determined in terms of π_0 (which is then determined from (9)). Note that it is not necessary to find an analytic expression for each π_i; the coefficient of π_0 for each π_i may be calculated iteratively. When calculating the coefficients, each value of k_{ij} should be calculated only once (when first required) and then stored.

The mean queue length, $E[n]$, is

$$E[n] = \sum_{i=0}^{\infty} iP_i \approx \sum_{i=0}^{k} i\pi_i. \tag{10}$$

[9] The two sequences, $\{p_i\}$ and $\{\pi_i\}$, are identical for the M/G/1:(∞, FIFO) nonstate dependent queue [13]. Harris proved [14] that they are not identical for the state dependent queue. Since the $\{p_i\}$ are difficult to find, simulations of the disk model using SCAN were conducted to see if the sequence $\{\pi_i\}$ is a good approximation to the sequence $\{p_i\}$. Estimates for the $\{p_i\}$ and $\{\pi_i\}$, derived from these simulations, differed by less than one percent for all values of i.

The determination of mean response time, $E[T]$, is the final step of the analysis. For this we use a theorem of Little [17] which states that for *all* queue scheduling algorthms,

$$E[T] = E[n]/\lambda. \tag{11}$$

The mean response time may now be calculated and plotted as a function of λ, the mean request rate. The response curve calculated for the IBM 2314 disk assuming the use of only one module is shown in Figure 6.

For a given set of disk model parameters, the conditions necessary for steady state establish a value of λ, denoted λ_{max}, at which the analysis is not valid. In the disk model

$$\lambda_{max} = 1/\tau_1, \tag{12}$$

where τ_1 is the mean service period assuming one request in the queue when service begins. λ_{max} is not readily calculated when following the procedure just described. It can be calculated, however, by a method described below.

(B) FIFO SCHEDULING WITH UNIFORM CYLINDER ADDRESS DISTRIBUTION ($\beta = 1$). Here, service time is not dependent on the length of the queue, and the model is an example of the well-known M/G/1: (∞, FIFO) queue for which the mean response time is given by the Pollaczek-Khintchine formula [18, p. 40]:

$$E[T] = (\lambda(\tau^2 + \sigma^2)/2(1 - \lambda\tau)) + \tau, \tag{13}$$

where τ is the mean service time, and σ^2 is the service time variance.

The service time, **t**, is distributed according to the density function $f_{t|n=1}(t \mid n = 1)$ given by (6). Therefore,

$$\tau = E[t) = \int_{-\infty}^{\infty} t f_{t|n=1}(t \mid n = 1)\, dt,$$

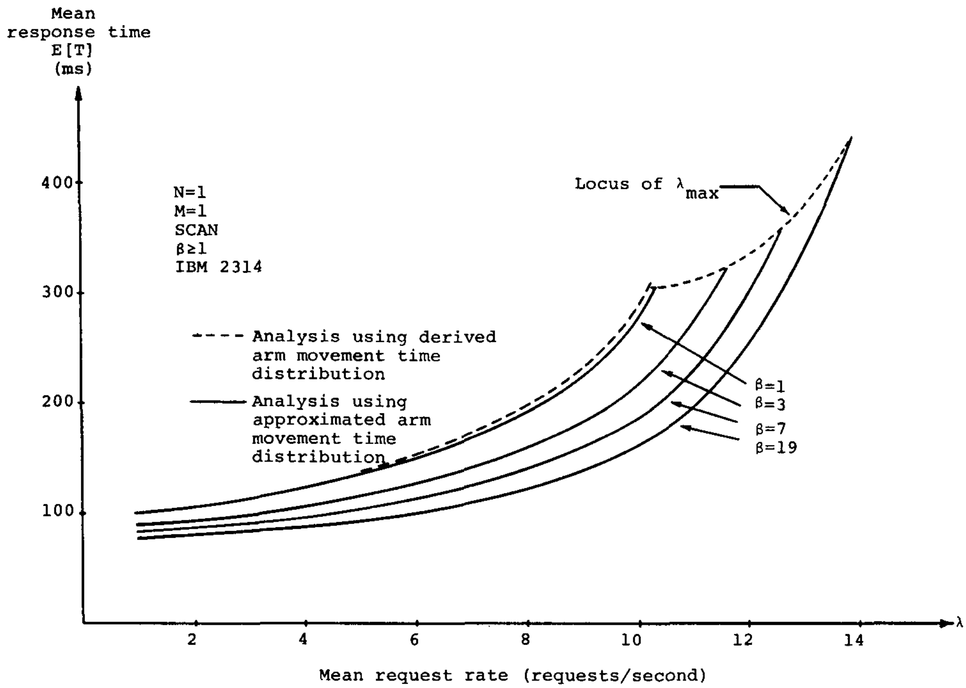

FIG. 6. Calculated mean response time for $\beta \geq 1$ under SCAN

which, on substitution, yields

$$\begin{aligned}\tau = {} & [(aC)^2(b + R)^2 - (aC)^2b^2 + (aC + b)^2b^2 - (aC + b + R)^2(b + R)^2 \\ & - (aC + b)^4 + (aC + b + R)]/2R(aC)^2 \\ & + [2(aC + b)^4 - 2(aC + b)b^3 + 2(aC + b + R)(b + R)^3 - 2(aC + b + R)^4] \\ & /3R(aC)^2 + [b^4 + (aC + b + R)^4 - (b + R)^4 - (aC + b)^4]/4R(aC)^2.\end{aligned} \tag{14}$$

Similarly,

$$\begin{aligned}E[t^2] = {} & [(aC)^2(b + R)^3 - (aC)^2b^3 + (aC + b)^2b^3 - (aC + b + R)^2(b + R)^3 \\ & - (aC + b)^5 + (aC + b + R)^5]/3R(aC)^2 \\ & + [2(aC + b)^5 - 2(aC + b)b^4 + 2(aC + b + R)(b + R)^4 \\ & - 2(aC + b + R)^5]/4R(aC)^2 \\ & + [b^5 + (aC + b + R)^5 - (b + R)^5 - (aC + b)^5]/5R(aC)^2.\end{aligned} \tag{15}$$

Calculating τ and $E[t^2]$ from (14) and (15) and σ^2 from $\sigma^2 = E[t^2] - \tau^2$, the mean response time can be calculated from (13). Figure 7 includes a response curve for the IBM 2314 disk assuming the use of only one module.

(C) SCAN SCHEDULING WITH NONUNIFORM CYLINDER ADDRESS DISTRIBUTION ($\beta > 1$). Here, as with the uniform distribution model, we must find the distribution of the lengths of the n interior intervals created by $n + 1$ points on the interval $(0, C)$ where the points are chosen as a random sample of size $n + 1$ from the distribution defined by $f_z(z)$. In the uniform case, it was sufficient to find the distribution of the ith interval, $1 \leq i \leq n$, since all intervals were identically distributed. In the present case, the distribution of a *randomly chosen* interval is required. The density function of this distribution, which may be found directly [1] or through the use of order statistics [19], is

$$g(s) = (n + 1)\int_0^{C-s} f_z(u)f_z(u + s)\left[\int_0^u f_z(x)\,dx + \int_{u+s}^C f_z(x)\,dx\right]^{n-1} du. \tag{16}$$

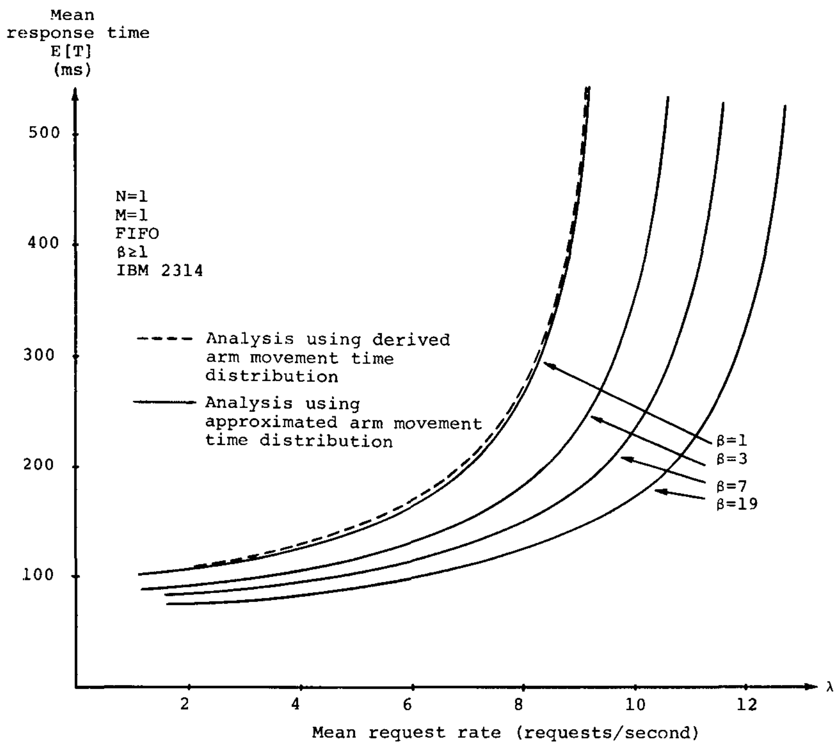

FIG. 7. Calculated mean response time for $\beta \geq 1$ under FIFO

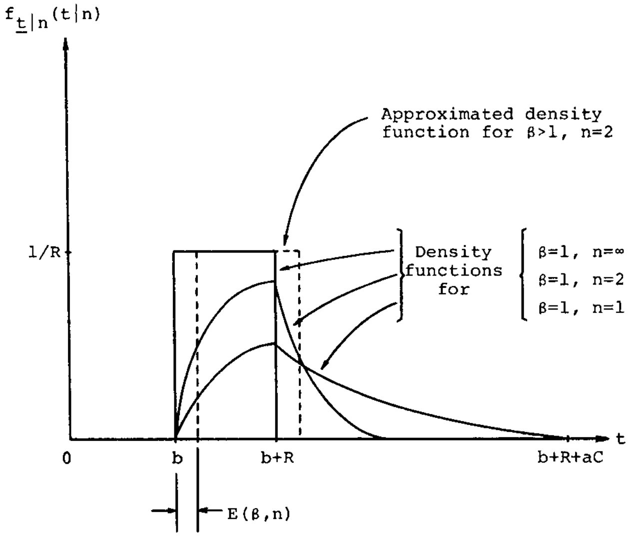

FIG. 8. Comparison of service time density functions

This function is intractable for any nontrivial $f_z(x)$.[10] Since $f_z(z)$ is taken here to be the beta distribution, $\beta > 1$, $g(s)$ cannot therefore be used analytically to determine the arm movement distance. To overcome this difficulty, an approximation is made at this point in the analysis.

The approximation is based upon the assumption that the arm movement distance is a constant for every value of n, the number of requests in the queue. This constant, $E'(\beta, n)$, is taken to be the mean arm movement distance. Thus $E'(\beta, n) = \int_0^C sg(s)\, ds$.

The mean arm movement *time* is then $E(\beta, n) = aE'(\beta, n)$. Where needed (see formula (20)), $E(\beta, n)$ is calculated numerically[11] for given values of β and n.

The approximated service time is

$$\mathbf{t} = b + \mathbf{c} + E(\beta, n) \tag{17}$$

where b is the fixed portion of arm movement time, $\mathbf{c}$ is the channel service time, $E(\beta, n)$ is the variable portion of arm movement time, and $\mathbf{t}$ is distributed according to

$$f_{\mathbf{t}|\beta,n}(t \mid \beta, n) = \begin{cases} 0, & t < b + E(\beta, n), \\ 1/R, & b + E(\beta, n) \le t < b + R + E(\beta, n), \\ 0, & t \ge b + R + E(\beta, n). \end{cases} \tag{18}$$

Figure 8 shows three of the density functions, from the family of functions derived for the $\beta = 1$ model, for $n = 1$, 2, and ∞. The approximated density function for $\beta > 1$ and $n = 2$ is also shown in dotted lines. The family of approximated density functions consists of identically shaped rectangular distributions displaced from the distribution for $\beta = 1$, $n = \infty$ by an amount $E(\beta, n)$.

The true density functions, $f_{\mathbf{t}|\beta,n}(t \mid \beta, n)$, for $\beta > 1$ are unknown. For a given value of n, however, the functions for $\beta > 1$ would be nonzero between b and $b + R + aC$ and would have a lower mean value than the function for $\beta = 1$ since mean arm movement time would be less. As n increases, the true density functions and the approximate density functions both approach the rectangular distribution shown in Figure 8. The

[10] Note that substituting $f_z(z) = 1/C, 0 \le z \le C$, yields $g(s) = (n + 1)(1 - s/C)^n$, which agrees with (1), the formula previously used for the interval distribution when $f_z(z)$ was taken to be the uniform distribution.

[11] Calculation for the results shown used an 8-point Gaussian integration method.

approximation, then, becomes better as n increases. Note that the service time dependence on n has been retained in the approximation.

As β increases, $E(\beta, n)$ decreases, so that the approximated density function approaches the rectangular distribution for $n = \infty$. The true density function also approaches this rectangular distribution as β increases. Consequently, the approximation becomes better as β increases. This argument will be referred to later in observing calculated results.

This mean service time assuming one request in the queue is given by $\tau_1 = b + R/2 + E(\beta, 1)$, $\beta \geq 1$. Therefore, from (12), the maximum value of the mean request rate is

$$\lambda_{max} = 1/(b + R/2 + E(\beta, 1)). \tag{19}$$

λ_{max} should be found before attempting to calculate a response curve so that the limits of λ for a valid computation are observed. Before applying the computation procedure for $\beta = 1$, λ_{max} can be found by first calculating $E(1, 1)$.

The mean response time may now be determined in exactly the same way as for $\beta = 1$. The elements of the transition probability matrix are, from (4),

$$\begin{aligned} k_{ij} &= 1/i! \int_0^\infty (\lambda t)^i e^{-\lambda t} f_{t|\beta,j}(t \mid \beta, j)\, dt \\ &= 1/i! \int_{b+E}^{b+E+R} ((\lambda t)^i e^{-\lambda t}/R)\, dt \\ &= 1/R \sum_{k=0}^{i} ((i-k)!\, \lambda^{k+1-i})^{-1} [e^{-\lambda(b+E)} (b+E)^{i-k} - e^{-\lambda(b+E+R)} (b+E+R)^{i-k}], \end{aligned} \tag{20}$$

where $E = E(\beta, j)$.

The calculation is carried out as described above; for efficiency, the $E(\beta, n)$ values are calculated only once (when first required) and then stored. Figure 6 shows a family of mean response curves calculated for the IBM 2314 disk and assuming the use of only one module. Note that two curves for $\beta = 1$ are included: one calculated using the approximate analysis, and one using the first analysis presented—the "exact" analysis. These curves, which differ by less than 3 percent for all values of request rate λ, indicate that the consequences of the approximation are acceptable. The comparison of these curves also gives assurance that the curves for $\beta > 1$ are acceptably accurate since, as argued previously, the approximation becomes better as β increases.

(D) FIFO SCHEDULING WITH NONUNIFORM CYLINDER ADDRESS DISTRIBUTION ($\beta > 1$). Mean response time can be calculated from (13), and the service time density function is given by

$$f_{t|\beta,n=1}(t \mid \beta, n = 1).$$

Therefore,

$$\text{mean service time} = \tau_1 = b + R/2 + E(\beta, 1), \tag{21}$$

$$\text{service time variance} = \sigma^2 = R^2/12. \tag{22}$$

Substituting (21) and (22) in (13), the mean response time is easily calculated. Figure 7 shows a family of response curves for the IBM 2314 disk assuming the use of only one module. As in Figure 6, two curves for $\beta = 1$ are inclnded: one calculated using the approximate analysis, and one using the first analysis presented. These curves differ by less than 5 percent for all values of request rate. Following the arguments given previously, it can be concluded that the curves shown for $\beta > 1$ are acceptably accurate.

Simulations of the disk model and measurement (via software) of a real disk (the IBM 2315) have been carried out to verify that the analysis procedures described adequately represent the behavior of real disks. Figure 9 compares the results of the simulation experiments with analytical results for SCAN scheduling. The simulation values tend to fall below the curve of calculated values as the load increases. This is expected since the analysis is based on the assumption that the arm must always move to serve a request,

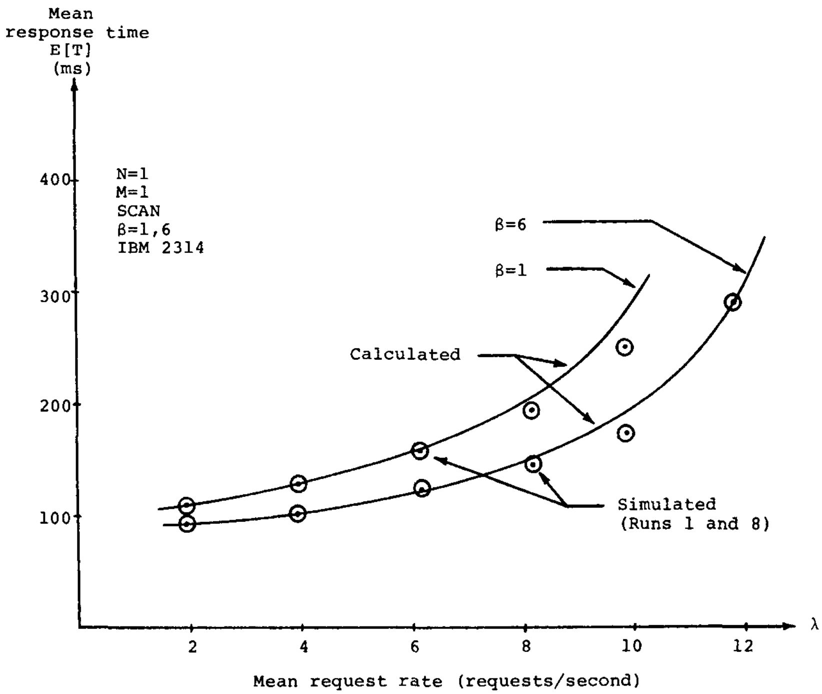

FIG. 9. Comparison of calculated and simulation results—SCAN

thus requiring at least the fixed portion of arm movement time. In practice, of course, the arm may not need to move, especially if $\beta \gg 1$ or the load is heavy. The values derived from the simulations run at low request rates are sufficiently close to give assurance that the calculations do represent the behavior of the disk model.

Figure 10 compares the results of the measurement experiments with analytical results for SCAN scheduling. Again, the measured values tend to fall below the curve of calculated values as the load increases. Nevertheless, the measurements taken at low request rates are sufficiently close to give assurance that the calculations do represent the behavior of the model of the 2315 disk operating in a steady state condition.

4. *The Multimodule System*

Consider the M-module single-channel disk model represented by the set of queues in Figure 2 and the associated timing diagram in Figure 1. This model, defined in Section 2, is analyzed using a method based partially on a technique of Seaman et al. [5]; Seaman analyzes the channel queue by applying results from the "machine repair" queueing model. In this model there is a single repairman to service M machines, each of which may break down at any time. That is, the time to the next failure for each operating machine is exponentially distributed. A queue of inoperative machines awaiting repair may form.

In the disk model, the channel may be looked upon as the repairman in the machine repair model and the M model queues may each be looked upon as the machines. The event of a request leaving a module queue to joint the channel queue corresponds to the event of a machine breaking down and joining the queue for the repairman.

Seaman's channel queue analysis assumes an exponentially distributed channel service time; this analysis is generalized to permit a uniformly distributed channel service time. The analysis of Section 2 is then applied to calculate the mean response time for FIFO and SCAN scheduling.

Recall that the module queue server remains busy until the channel service time is

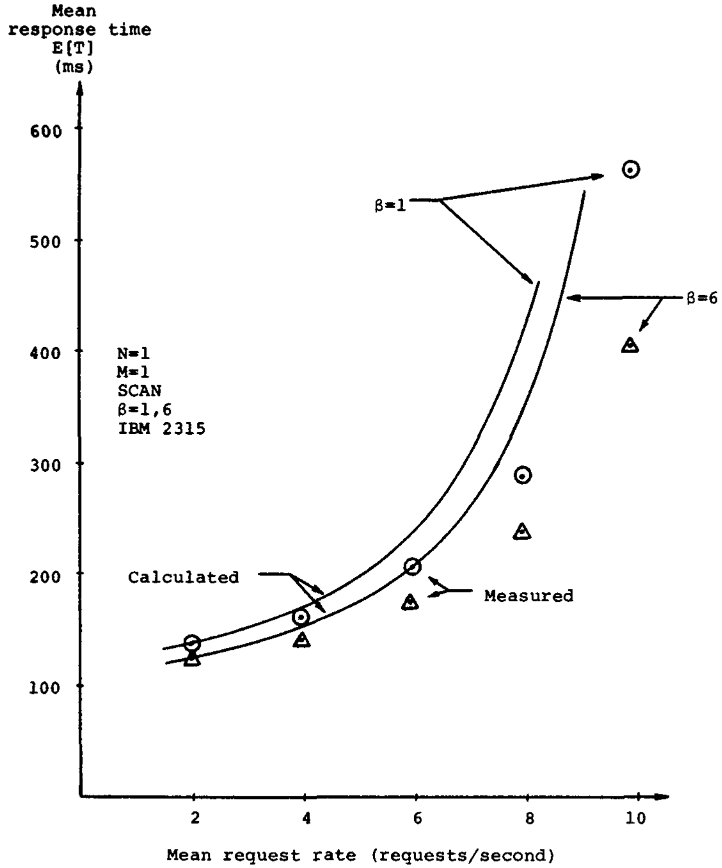

FIG. 10. Measured mean response time—SCAN

completed. Because of this interaction any one module queue, along with the channel queue, can be viewed as a single *equivalent* queue having a service time composed of: **m**, the arm movement time; **q**, the channel queueing time; and **c**, channel service time. Requests arrive at the equivalent queue with a mean request rate of λ/M. If the mean and the variance of **q** can be determined, the mean response time *for FIFO module queue scheduling* can be calculated directly from

$$E[T] = (\lambda/M(\tau^2 + \sigma^2)/2(1 - \lambda\tau/M)) + \tau, \tag{23}$$

where

$$\tau = \text{mean service time} = E[\mathbf{m}] + E[\mathbf{q}] + E[\mathbf{c}], \tag{24}$$

and

$$\sigma^2 = \text{variance of service time} = \sigma_{\mathrm{m}}^2 + \sigma_{\mathrm{q}}^2 + \sigma_{\mathrm{c}}^2. \tag{25}$$

In order to find $E[\mathbf{q}]$ and σ_{q}^2, the channel queue is analyzed as follows. From the viewpoint of the channel queue, the M module queues can be looked upon as M sources, each of which generates requests at an effective mean rate w. Assume that the requests are generated randomly in time, that is, requests arrive at the channel queue with exponentially distributed interarrival times. The M sources, because of the module-channel queue interaction described previously, do not, however, generate requests continuously; each source may contribute only one request at a time to the channel queue and therefore does not generate requests if this contribution is in the queue. Note that

(a) the channel queue contains at most M requests;

(b) $w > \lambda/M$, since the mean arrival rate to the channel queue is λ and the sources do not generate requests continuously.

The channel queue, as described here, is an example of the "machine repair" queueing model [18, p. 232]. Applying the theory which has been developed for such a queue [20] along with some general probability theory (see [1]), w is calculated as a function of λ using an iterative technique,[12] and $E[\mathbf{q}]$ and σ_q^2 are found to be

$$E[\mathbf{q}] = (M - 1) \cdot R/2 - (1 - \pi_{M-1})/w \qquad (26)$$

where

π_{M-1} = Prob [there are no requests in the channel queue immediately preceding completion of a service time],

$$\sigma_q^2 = \sum_{k=0}^{M-1} p_k[R^2/6 - R^2(1 - p_M)/9 + R^2(M - k - 1)/12] \qquad (27)$$

where

p_k = Prob [there are $M - k$ requests in the channel queue at any random time], $0 \leq k \leq M$.

Tables of π_{M-1} and p_k, as functions of λ, can be calculated as described in [1].

We now examine the four possible combinations of scheduling (FIFO and SCAN) with values of $\beta = 1$ and $\beta > 1$.

(A) SCAN, $\beta = 1$. Channel queueing time, $\mathbf{q}$, is assumed to be a constant; this approximation is necessary so that $\mathbf{q}$ can be incorporated into the previous analysis. This can be done by assuming that queueing time is part of the fixed portion, b, of arm movement time. The equivalent queue is therefore identical to the single-module queue analyzed previously, and $E(T)$ may be calculated by applying the corresponding single-module procedure with (i) b replaced by $b + E[\mathbf{q}]$, and (ii) λ replaced by λ/M.

(B) FIFO, $\beta = 1$. Mean response time, $E[T]$, is calculated directly from (23) where

$$\tau = \tau' + E[\mathbf{q}], \qquad \tau' = E[\mathbf{m}] + E[\mathbf{c}]$$

which is obtained from (14),

$$\sigma^2 = \sigma^{2\prime} + \sigma_q^2, \qquad \sigma^{2\prime} = \sigma_m^2 + \sigma_c^2 = E[t^2] - \tau'^2,$$

and $E[t^2]$ is obtained from (15).

(C) SCAN, $\beta > 1$. Following the argument of (a) above, $E[T]$ is now calculated, for SCAN and $\beta > 1$, by applying the corresponding single-module procedure with (i) b replaced by $b + E[\mathbf{q}]$, and (ii) λ replaced by λ/M.

(D) FIFO, $\beta > 1$. Mean response time is calculated directly from (23) where

$$\tau = \tau' + E[\mathbf{q}], \qquad \tau' = E[\mathbf{m}] + E[\mathbf{c}] = b + R/2 + E(\beta, 1) \qquad \text{(see (21))},$$

$$\sigma^2 = \sigma^{2\prime} + \sigma_q^2, \qquad \sigma^{2\prime} = \sigma_m^2 + \sigma_c^2 = R^2/12 \qquad \text{(see (22))}.$$

The values of parameters over which the calculations were performed are:
$M = 1, 2, 3, 4$ (see footnote 12);
$R = 30, 40, 50$ (maximum block transfer time of 5, 15, 25 ms);
$\beta = 1, 6$;
Scheduling: FIFO, SCAN.

[12] Newton's method for solving the equation $f(w) = 0$. Since $f(w)$, as it occurs here, becomes a very long formula for $m > 2$, a formula manipulation system, FORMAC, was used to generate the formulas for $f(w)$. FORMAC is unable to evaluate formulas containing exponentials; consequently, the formulas were printed out, punched, and inserted into the Newton method program. This procedure became unmanageable for $M > 4$. For present purposes, however, calculations up to $m = 4$ are sufficient and no other method of generating the formulas was pursued.

5. *Use of Results*

The improvement in mean response time resulting from SCAN, rather than FIFO, scheduling is illustrated in Figure 11. Note that the improvement is much greater for high request rates. The improvement itself is not as significant as the smaller slope retained by the SCAN curve through higher request rates; this enables SCAN to operate satisfactorily at values of λ that would cause very long queues to build up under FIFO scheduling. SCAN, then, is most useful under very heavy loading conditions.

One can observe certain basic quantities illustrated in Figure 11 by $\Delta_M E(M, S)$, $\Delta_S E(M, S)$, and $\lambda_{\lim}(M, S)$:

$\Delta_M E(M, S)$—the change in $E[T]$ produced by adding one module (thereby increasing the total amount of stored data by the capacity of a module) to an $M - 1$ module configuration, assuming that λ/M and the scheduling algorithm, denoted by S, are fixed. As shown, $\Delta_M E(M, S)$ for $M = 2$, $\lambda/M = 7.5$, and $S =$ FIFO, equals +18 ms.

$\lambda_{\lim}(M, S)$—the highest value of λ for which $E[T]$ falls below some given maximum allowable mean response time $E[T]_{\max}$. As shown, $\lambda_{\lim}(M, S)$ for $M = 1$, $E[T]_{\max} =$ 240 ms, and $S =$ FIFO, equals 7.5 requests/second.

$\Delta_S E(M, S)$—the improvement in $E[T]$ obtainable through scheduling (as opposed to no scheduling, i.e. FIFO), assuming that all other parameters are fixed. As shown, $\Delta_S E(M, S)$ for $M = 1$, $S =$ SCAN, and $\lambda = 7.5$, equals -60 ms.

As a first approximation in estimating the changes resulting from adding more modules, one expects that adding m modules would, for example, increase the permissible range of λ by some multiple of m. In other words, the quantity $\lambda_{\lim}(M, S)/M$ should not vary greatly as M increases. We find from Figure 9 the following variation:

M	$\lambda_{\lim}(M, S)$	$\lambda_{\lim}(M, S)/M$
1	7.5	7.5
2	14.5	7.3
3	20.4	6.8
4	25.2	6.3

$\lambda_{\lim}(M, S)/M$ decreases as M increases. This decrease is to be expected from increasing congestion in the channel as the data transfer load increases; increasing M, and thereby λ, raises channel utilization which, as queueing theory predicts, results in a longer channel queue.

Channel congestion affects other quantities as well: $\Delta_M E(M, S)$ indicates the cost in mean response time of connecting one additional module to a single channel. Assuming that each module has a mean request rate of 7.5 requests/second, $\Delta_M E(M, S)$ for $S =$ FIFO varies as follows:

M	$\Delta_M E(M, S)$ (ms)	$\Delta_M E(M, S)$ (%)
2	18	8
3	57	22
4	160	51

Thus the addition of one module to a three-module system increases the mean response time of each of the three modules by 51 percent.

Further discussion, with examples, of the use of the results is given in [1]. In particular, different curves are plotted that illustrate the effect on $E[T]$ of the module arm movement characteristics, the grouping of files on adjacent cylinders, and the mean block length.

6. *Extensions*

Tables such as those represented by the data plotted in Figure 11 may be useful for evaluating system configurations without resorting to time-consuming and expensive simulation or measurement. To make comparisons between devices and to consider the

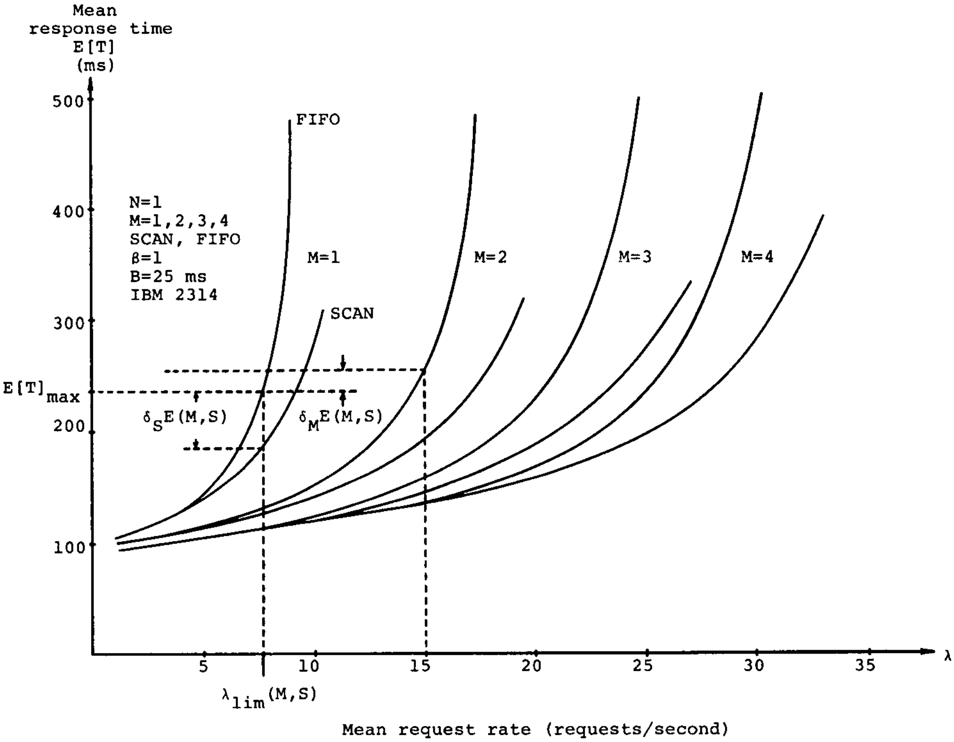

FIG. 11. Variation in $E[T]$ with M

trade-offs involved, similar tables are also needed for other devices such as drums and fixed head disks.

The system designer must, of course, be fully aware of the assumptions and approximations upon which the calculation of the tables is based. These assumptions may not, however, seriously restrict the utility of the tables; in many design situations, the estimates of system loading either are approximate or are bounds on an allowable range of load. Under such conditions tables provide: (a) a frame of reference within which the designer may work to evaluate various combinations of parameters; (b) some indication of the theoretical limits within which the system must operate (given that the assumptions of randomness apply).

An empirical study of disk performance over a wide range of operating conditions is needed. Data from such a study would provide information about the applicability of the idealized disk model and, more importantly, would provide some insight into the behavior of real systems so that the model may be appropriately refined.

Although the SCAN scheduling algorithm was chosen for attention on the basis of system equity, the SSTF algorithm is not completely rejected; an analytical procedure for accurately determining the performance of disks using SSTF scheduling would contribute to understanding this algorithm. A modified SSTF may prove to be effective. For example, recording the waiting time of each data transfer request and giving service priority to requests that have waited an excessive amount of time, may prove to be effective under certain conditions. Unless the time limit were dynamically adjusted, however, this modified SSTF would degenerate to FIFO under a very heavy load.

Both SCAN and SSTF are clearly better than FIFO with respect to mean response time. To investigate the characteristics of these as well as other algorithms [8] within a real system, experiments should be carried out by inserting a queue scheduling program

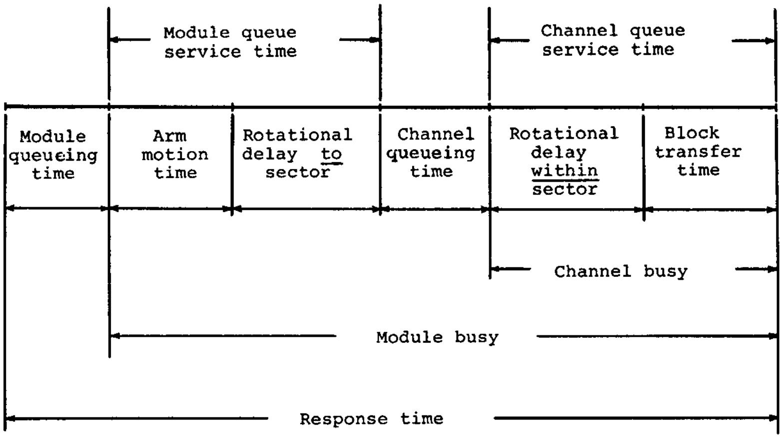

FIG. 12. Timing diagram for IBM 3330 disk

module into the Input/Output Supervisor and measuring the performance under each algorithm. It is encouraging to note in this regard that a version of SCAN scheduling has very recently been offered as an option in OS/360 [21].[13]

A hardware approach to reducing response time is taken in a recently marketed movable-head disk, the IBM 3330. This device, available with 2, 4, 6, or 8 modules (as modules are defined here) eliminates use of the channel during part of the rotational delay time. Channel utilization for a given value of λ is thereby decreased and channel congestion is reduced accordingly.

The timing diagram for a 3330 disk model is shown in Figure 12. As in Figure 1, overhead times are ignored. The queueing diagram for the 3330 is the same as for the 2314 (Figure 2).

As with the 2314 model, the module queue server remains busy when a request leaves the module queue to join the channel queue (Time 1), until that request leaves the channel queue (Time 2). Channel queueing time is either zero or a multiple of a full rotation time. Note that for a single-module configuration of such a disk, channel queueing time is always zero and the timing diagram reduces to that of Figure 4 (except for the channel busy period shown in Figure 4). The model and analysis procedure described in Sections 2 and 3 are therefore valid for a disk of this type.

A multimodule disk model of a 3330-like disk can be analyzed as in Section 4 except that a new channel queue analysis, analogous to that contained in [1], is needed. The channel queue is similar to the drum model analyzed by Abate and Dubner [22] except for two important differences: first, the queue can only contain up to M requests; second, there are M, not one, distinct rotating surfaces. Nevertheless, Abate and Dubner's analysis may provide insight into the channel queue analysis needed here.

Note that it is important, in view of the flexible configuration of the 3330 disk (2, 4, 6, or 8 modules), to be able to analyze and compare multimodule models with differing numbers of modules.

The introduction of the IBM 3330 strengthens the opinion that movable-head disks will be important for a significant time to come. Recently, other types of secondary storage devices, e.g. magnetic card devices, have become less important because they have not competed favorably with movable-head disks. At the same time, newer technologies such as optimal recording have not progressed to a state whereby useful large capacity devices can be built.

[13] No information concerning performance measurements or implementation problems is currently available.

ACKNOWLEDGMENTS. The authors wish to thank Professors J. J. Horning and J. G. C. Templeton for many useful comments and suggestions. The authors also wish to thank the National Research Council of Canada for financial assistance, and to acknowledge the programming help of Miss Mina Craig.

REFERENCES

1. MACEWEN, G. H. Performance of disk storage devices in computer systems. Ph.D. Th., Dep. of Computer Science, U. of Toronto, Toronto, Canada, 1971.
2. FREEMAN, D. N. A storage hierarchy system for batch processing. Proc. AFIPS 1968 SJCC, Vol. 32, AFIPS Press, Montvale, N. J., pp. 229–243.
3. GAVER, D. P. JR. Probability models for multi-programming computer systems. *J. ACM 14*, 13 (July 1967), 423–438.
4. FIFE, D. W., AND SMITH, J. L. Transmission capacity of disk storage systems with concurrent arm positioning. *IEEE Trans. EC-14* (Aug. 1965), 575–582.
5. SEAMAN, P. H., LIND, R. A., AND WILSON, T. L. An analysis of auxiliary storage activity. *IBM Syst. J. 5*, 3 (1966), 158–170.
6. DENNING, P. J. Effects of scheduling on file memory operations. Proc. AFIPS 1967 SJCC, Vol. 30, AFIPS Press, Montvale, N. J., pp. 9–21.
7. ABATE, J., DUBNER, H., AND WEINBURG, S. B. Queueing analysis of the IBM 2314 disk storage facility. *J. ACM 15*, 4 (Oct. 1968). 577–589.
8. TEOREY T. J., AND PINKERTON, T. B. A comparative analysis of disk scheduling policies. *Comm. ACM 15*, 3 (Mar. 1972), 177–184.
9. FRANK, H. Analysis and optimization of disk storage devices for time-sharing systems. *J. ACM 16*, 4 (Oct. 1969), 602–620.
10. IBM System/360 component description—2314 direct access storage facility. Form A26-3599, IBM Syst. Ref. Library.
11. FUCHS, E., AND JACKSON, P. E. Estimates of distributions of random variables for certain computer communications traffic models. *Comm. ACM 13*, 12 (Dec. 1970), 752–757.
12. FELLER, W. *An Introduction to Probability Theory and Its Applications, Vol. II.* Wiley, New York, 1966.
13. HARRIS, C. M. Queues with state dependent stochastic service rates. *Oper. Res. 15*, 1 (Jan. 1967), 117–130.
14. HARRIS, C. M. Queues with state-dependent stochastic service rates. Ph.D. Th., Polytechnic Inst. of Brooklyn, Brooklyn, N. Y., 1966.
15. KENDALL, D. G. Stochastic processes occurring in the theory of queues and their analysis by means of the imbedded Markov chain. *Ann. Math. Stat. 24* (1953), 338–354.
16. PARZEN, E. *Stochastic Processes.* Holden-Day, San Francisco, 1962.
17. LITTLE, J. D. C. A proof of the queueing formula: $L = \lambda W$. *Oper. Res. 9*, 3 (Mar. 1961), 383–387.
18. SAATY, T. L. *Elements of Queueing Theory with Applications.* McGraw-Hill, New York, 1961.
19. FOSTER, F. G. On stochastic matrices associated with certain queueing processes. *Ann. Math. Stat. 24* (1953), 355–360.
20. TAKACS, L. *Introduction to the Theory of Queues.* Oxford U. Press, New York, 1962.
21. IBM Installation Newsletter, Issue No. 70-07, Apr. 1970.
22. ABATE, J., AND DUBNER H. Optimizing the performance of a drum-like storage. *IEEE Trans. C-18*, 11 (Nov. 1969), 992–997.
23. COFFMAN, E. G., KLIMKO, L. A., AND RYAN, BARBARA. Analysis of scanning policies for reducing disk seek times. *SIAM J. Comp.* (to appear).

RECEIVED MARCH 1971; REVISED MAY 1972