PERFORMANCE EVALUATION OF CCD CHIP ORGANIZATIONS FROM MEMORY SYSTEM DESIGN VIEWPOINT

Satish L. Rege Burroughs Corporation Computer Systems Group Piscataway, New Jersey08854

This paper discusses the computer memory system design implications of different CCD devices, such as Circulating Shift Register, Serial Parallel Serial, and Line Addressable Organizations.

The performance of the memory using these devices is evaluated in a stand alone mode using a single server queuing model, and as a buffer between the disk and main memory in a computer system by using a two server cyclic queuing model. The performance of the stand alone mode for CCD memory system is defined as the time elapsed between the request for a record and the completion of that request. The performance comparison of stand alone mode shows that Serial Parallel Serial Organization has the worst performance as one should expect. It is shown that the Circulating Shift Register with burst mode for refreshing has better performance than the Line Addressable Organization. Circulating Shift Register Organization with cache has the best performance of all but would require a considerable amount of extra cost.

In evaluating the CCD chip organizations as a buffer between the disk and the main memory, it is shown that all the three organizations (Serial Parallel Serial, Circulating Shift Register, and Line Addressable) have a place in a memory system design depending on the user requirements. It is also shown that a three level (Bipolar, MOS and CCD) memory system using the different CCD chip organizations has a considerable advantage in performance over a two level (Bipolar, MOS) organization for the same total cost.

1. INTRODUCTION

Charge Coupled Devices (CCD's) were developed to increase the density and, consequently, decrease the cost/bit of MOS memory systems. In achieving this goal, it was necessary to make these devices serial in nature. A natural consequence was to assume that the memory systems built using these devices would be useful for replacing serially accessed stores such as disks or drums. Even though CCD's are useful for such application, they have a much higher potential. The electronic nature of CCD's make them more versatile than the mechanically controlled drums and disks. Modularity at a much smaller size with efficient synchronization between the different modules is possible with CCD. This possible modularity should be exploited to build memory systems with flexible architecture, making a single design suitable for various existing and proposed computer systems and software.

CCD's useful for computer memory system implementation have been designed and devices fabricated. These are the Circulating Shift Register (11). Serial Parallel Serial (SPS) (2), and Line Addressable Organization (LARAM) (2). The essentials of these organizations are shown in Figure 1. This paper evaluates these three organizations from a computer system designers point of view. The performance of memories built using these organizations is evaluated in two different modes of operation: (a) A stand alone mode in which the memories are operated as a disk or a drum type memory and (b) In a computer system as a gap filling technology between the disk and the main memory. The two modes are evaluated using a single server and a two server cyclic queuing model respectively. For quantitative evaluation some representative parameters as shown in Figure 1 are used.

2. SINGLE SERVER QUEUING MODEL

Optimization of stand alone memory systems such as disk and drums have been reported by various researchers in the literature: Abate (1) and Fuller (6) have analyzed drum type memory systems, while Bhandarkar (3) has evaluated magnetic bubble memories. A memory system using CCD's is essentially a service facility with a queue and hence, can be modeled as a single server queuing system. Figure 2 shows the different events and the intervals encountered in satisfying a request made to a CCD memory system.

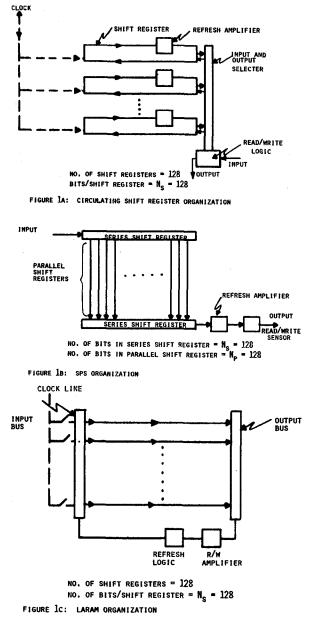


FIGURE 1: DIFFERENT CCD CHIP ORGANIZATIONS

A request arriving at the server is serviced in time A, where A is a random variable with arbitrary distribution function F_A (t). After serving a request, the server inspects the queue to see if there are any outstanding requests. If the queue is not empty, then a new service period begins and if it is empty, then the server is latent until a new service request is made.

Before analyzing the memory systems using the different chip organizations, it will be necessary to make some assumptions about the requests made to the service facility and the size of the data requested. Requests will be assumed to have a Poisson arrival process with a rate λ (5). For most computing systems such an assumption is valid. This assumption states that any two given time in-

tervals of equal lengths experience an arrival with equal probability, and the number of arrivals in disjoint time intervals are independent. The assumption implies that the probability of K arrivals in an arbitrary interval of time t is:

PROB. {K ARRIVALS IN TIME t} = $\frac{(\lambda t)^{K} e^{-\lambda t}}{K!}$

and the interarrival times have the exponential density function:

PROB. {INTERARRIVAL TIME = t} = $\lambda e^{-\lambda t}$

Memory will be assumed to be divided into pages and the size of the record requested will be assumed to be an integral multiple of the page size. Fuller (6) reports a study of the record lengths conducted by him on the drum storage units on an IBM 360/91. He has shown that the record lengths can be approximated by an exponential distribution. Thus, in a discrete

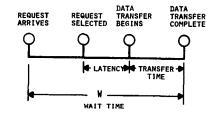


FIGURE 2: INTERVALS AND EVENTS ASSOCIATED WITH SERVING A REQUEST BY A CCD MEMORY SYSTEM.

case, the record length can be expressed as a random integer variable R, which has a geometric distribution. The probability mass function of R is:

PROB. {R = 1} =
$$\frac{(1-\alpha)}{\alpha(1-\alpha n)} \alpha^{1}$$

for i = 1, 2..., n and $\alpha < 1$

If n is large and $\alpha^n <<1$, then the function can be approximated as:

PROB.
$$\{R = i\} = (1 - \alpha) \alpha^{i}$$

 $i = 1, 2...n$

and the first and the second moment of R are given by:

$$\overline{R} = \frac{1}{(1 - \alpha)}$$
$$\overline{R}^2 = \frac{(1 + \alpha)}{(1 - \alpha)^2}$$

With the above-mentioned fundamental assumptions and first-in-first-out (FIFO) scheduling policy for the queue, the average waiting time W using the Pollaczek-Khintchine formula is given by: (Ref. (7)).

$$\overline{W} = \left\{ \frac{\lambda \ \overline{A}^2}{2 \ (1 - \lambda \ \overline{A})} + \ \overline{A} \right\}$$

The unit for the waiting time will be determined in number of periods of the access frequency and will be normalized to account for variable record length, where record length is measured in number of pages and the page size is equal to the shift register length N_S. The normalized waiting time is given by $W/N_S\bar{R}$. Physically the normalized waiting time denotes waiting time per bit of information. Figure 1 shows the values for different parameters to be used in our analysis.

3. ANALYSIS OF THE DIFFERENT ORGANIZATIONS

The service time for the circulating shift register organization is the sum of the time taken to align the bit addressed as the first bit under the read/write sensor and the time taken to transfer the whole shift register. Let N_S be the number of bits in a shift register, L the latency time to align the first bit under the read/write head, R the record size in number of pages and the page size equal to the shift register length. Then:

$$A = L + N_{S}R$$
$$\bar{A} = \frac{(N_{S}-1)}{2} + N_{S}\bar{R}$$

VAR (A) =
$$N_{s}^{2} \left\{ \frac{\alpha}{(1-\alpha)^{2}} + 0.0883 \right\}$$

$$\bar{A}^2 = VAR(A) + (\bar{A})^2 *$$

In Figure 3 these equations are used to draw the graph of traffic intensity (ρ =Arrival Rate/Service Rate) versus normalized \overline{W} . The graph shows that normalized \overline{W} improves as the average record size increases because the larger the average record size, the more the number of bits over which the aligning time is distributed.

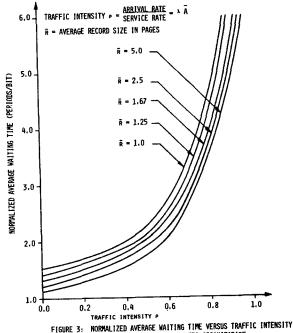
The SPS organization shown in Figure 1 requires two frequencies, f_8 and f_p , such that $f_8 = N_8 f_p$. Again we assume that the page size is N_8 . We also assume that there are p such circuits on a chip so that a parallelism of p is possible.

Then,

$$\overline{A} = N_{g} \{0.5N_{p} + \frac{1}{p(1-\alpha)}\}$$

$$VAR(A) = \{\frac{N_{g}^{2}}{p^{2}} - \frac{\alpha}{(1-\alpha)^{2}} + \frac{N_{p}^{2}N_{g}^{2}}{12}\}$$

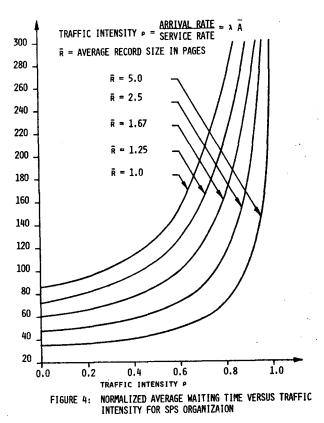
A graph of ρ versus the normalized \vec{W} for p=1 and various values of average record sizes is



FOR CIRCULATING SHIFT REGISTER ORGANIZATION

drawn in Figure 4. The calculation for p=4indicates that the normalized \overline{W} does not change much from p=1 thus showing that the latency time L is the dominant term for the SPS organization.

The graph in Figure 4 shows that SPS organization is approximately 50 times slower than the circulating shift register type organization



^{*}In the later sections \overline{A} and VAR (A) will be determined and this equation will be used to determine $\overline{A^2}$.

for $\rho=0.5$. It is practical to make the serial shift frequency (f_B) for the SPS organization 5 to 10 times that of the Circulating Shift Register organization. Later, to compare the organizations, the highest value of 10 which gives the best possible performance for SPS is used. Then the waiting time for the SPS organization will have to be divided by ten thus improving its performance.

LARAM (2) has advantages over the Circulating Shift Register because: (a) A single refresh amplifier is shared between number of different shift registers, (i.e., a block of shift registers); (b) The power dissipation is less because all the shift registers are not moving simultaneously; (c) Whenever an access is made, there is no time wasted in aligning the shift register.

The disadvantages of this organization are: (a) Whenever a shift register is being refreshed the block to which the shift register belongs cannot be accessed, and, (b) The shift registers have to be refreshed cyclically one at a time. The first requirement results in making a block of shift registers refresh busy any time a single shift register in the block is refreshed while the second requirement needs a large storage interval. Notice that if a block is refresh busy then it cannot be accessed for data. Assuming that the shift registers in a block are cyclically refreshed at equal intervals, the "effective refresh time" for a LARAM type device is given by the refresh time/ bit divided by the number of shift registers that share a common refresh amplifier.

To determine \overline{W} , see Figures 5 and 6 showing the latency time and transfer time versus the time of request. If f_r is the frequency at which the shift registers are refreshed, then

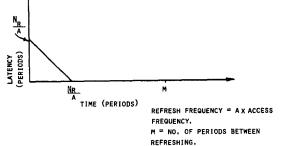


FIGURE 5: LATENCY VERSUS TIME OF REQUEST FOR LARAM ORGANIZATION

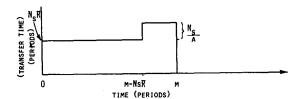


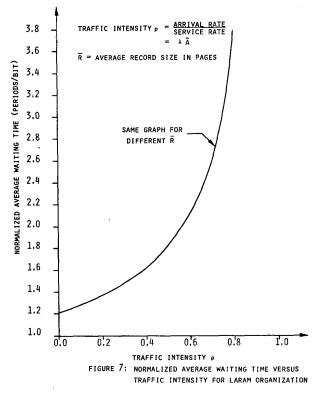
FIGURE 6: TRANSFER TIME VERSUS TIME OF REQUEST FOR LARAM ORGANIZATION

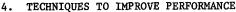
 $f_r=af_a$, where 'a' is the constant of proportionality. Let m be the number of periods between refreshing. Then,

$$\overline{A} = \frac{N_{g}^{2}}{2a^{2}m} + N_{g}\overline{R} \left(1 + \frac{N_{g}}{am}\right)$$

$$VAR (A) = \frac{N_{g}^{3}}{12ma^{3}} \left(4 - \frac{3N_{g}}{ma}\right) + \frac{N_{g}^{3}\overline{R}}{a^{2}m} \left(1 - \frac{N_{g}\overline{R}}{m}\right)$$

A graph of ρ versus normalized \overline{W} for the LARAM organization with 32 shift registers is shown in Figure 7. The graph shows that this organization has less normalized \overline{W} than the Circulating Shift Register organization. Notice that in using this organization, the advantage achieved due to the absence of a requirement for aligning the registers is lost by the inability to access data while refreshing any of the shift registers.





The performance of the different organizations can be improved by various means. Here we will consider two different ways which are: (a) Circulating shift register with burst mode of operation and, (b) Circulating shift register with a cache. The circulating shift register with burst mode of operation has been discussed earlier in (10). Such a mode of operation is very similar to LARAM, except every shift register has an individual refresh amplifier. Therefore, during the burst refresh all the shift registers are refreshed simultaneously. Hence for this case:

$$\overline{A} = \frac{N_{g}^{2}}{2a^{2}m} + N_{g}\overline{R}$$

$$WAR (A) = \frac{N_{g}^{3}}{12ma^{3}}(4 - \frac{3N_{g}}{ma}) + N_{g}^{2} \{\frac{\alpha}{(1-\alpha)^{2}}\}$$

Using a refresh frequency ten times faster than the access frequency, a graph of ρ versus normalized \tilde{W} is drawn in Figure 9. It is easily seen that this mode of operation is advantageous over both normal circulating shift register and LARAM designs.

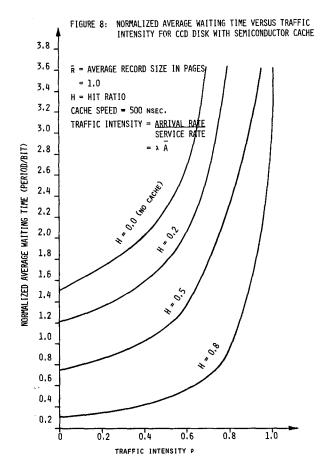
The performance of a CCD memory can be improved by using cache type systems. A MOS memory may be used as a cache to a CCD memory. Such an organization then can be characterized by a hit ratio H, and the latency and access times are a function of this ratio. Assuming a word length of B (where Ns is an integer multiple of B) and access time Tf for the MOS memory,

$$\overline{Y} = H\overline{R}T_f \frac{N_s}{B} + (1-H) \overline{A}$$

 $VAR(Y) = (1-H)^2 VAR(A)$

Where \overline{A} and \overline{Y} are the average service times for Circulating Shift Register organizations without cache and with cache respectively. A graph of ρ versus normalized \overline{W} for various values of hit ratios is drawn in Figure 8. This design clearly has a better performance than all the other organizations. The requirement of a substantial amount of cache memory will make this design more costly than other designs.

Another technique would be to modularize the CCD memory system at a small module size and

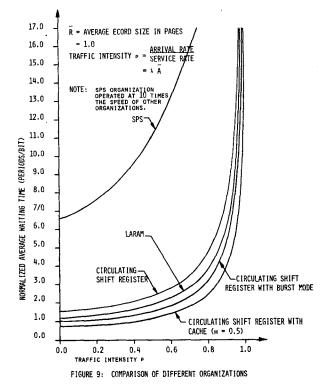


either interleave the different modules or operate them in parallel to achieve higher performance. Such a system has not been evaluated here.

5. COMPARISON OF DIFFERENT ORGANIZATIONS

To compare the different organizations, we plot ρ versus normalized \widetilde{W} for all the organizations in Figure 9. The access frequency for all organizations except the SPS is assumed to be the same. The SPS organization is operated at a 10 times faster frequency than the other organizations.

The graph in Figure 9 shows that the SPS is the worst organization of all from the standpoint of performance. The latency time in the parallel path of the SPS organization is quite high and dominates the performance. As stated earlier, parallelism does not help to improve this situation. One solution is to have record lengths very large so that the latency time is distributed over a large number of bits. Notice that one of the main advantages claimed for the SPS organization is its low power dissipation. The higher access frequency increases the power dissipation for the SPS organization. Notice that LARAM may be used for lowering the power dissipation.



The Circulating Shift Register organization has a performance advantage of at least four times that of the SPS organization and is useful because of its short shift register length. This organization is also more flexible for building a serial memory system that can emulate a number of different existing designs.

The LARAM organization is better than the Circulating Shift Register organization both in terms of power dissipation and normalized average waiting time. The disadvantage here is the requirement for large storage time as compared to other organizations.

The Circulating Shift Register with burst mode of operation is faster than the SPS or the LARAM. Power dissipation for the Circulating Shift Register with burst mode of operation can be improved over the Circulating Shift Register with continuous refresh by judiciously choosing the right type of refreshing. However, the cost/ bit for this design would be slightly higher.

Finally, the Circulating Shift Register organization with a cache that achieves a hit ratio as small as 0.5 has the best performance of all. The Circulating Shift Register with cache will perhaps have the highest cost of all the designs.

6. IMPLICATIONS OF CCD IN A MEMORY SYSTEM DESIGN

Performance expectations of paged memories has always been constrained due to thrashing. Denning (4) describes thrashing as "excessive overhead and severe performance degradation or collapse caused by too much paging." One solution suggested to overcome thrashing has been to use a slow speed bulk core storage such as used by Laur (6).

The effects of thrashing can be reduced by using Charge Coupled Devices (CCD's) whose performance lies somewhere between that of the random access memories such as, MOS RAM's and serially accessed memories such as, disks and drums. Therefore, CCD is a good candidate as a level between these two technologies in a memory hierarchy design for a computer system. Such a design should make the memory hierarchy look more homogeneous. Figure 10 shows an n level hierarchy.

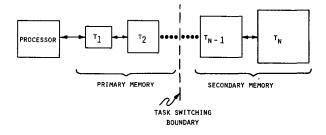
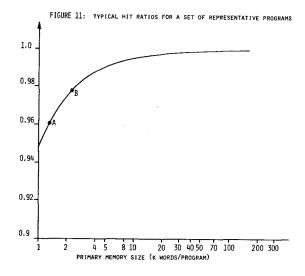


FIGURE JO: A MEMORY HIERARCHY WITH A TASK SWITCHING BOUNDARY

The performance of a hierarchy can be determined by knowing the hit ratio characteristics of the program environment in which the memory system will be used. Figure 11 shows a typical hit ratio characteristic as collected for some sample programs in a large computer system. The performance will be determined by assuming a multiprogramming mode of operation.

Multiprogramming, a pseudo parallel operation, is multiplexing of CPU over a number of different tasks to increase the CPU utilization.



Therefore, given a processor, and a memory hierarchy using different technologies T_1 , $T_2...T_n$, a boundary exists such that an access across the boundary necessitates a task switch due to the excessive amount of time required for retrieving the data. One of the main reasons for the boundary to exist between any two technologies is the disparity between the task switching time required and the access time of the technology.

In the memory hierarchy the technologies that are used on the processor side of the task switching boundary will form a part of the primary memory while the others, the secondary memory. The degree of multiprogramming is the average number of active tasks that reside in the primary memory and is usually a function of the primary memory size and working set size of the program.

7. A TWO SERVER QUEUING NETWORK MODEL

In a multiprogramming environment a task cycles through four states: The first state is the task being serviced by the processor, the second waiting for the secondary memory or I/O service in a queue, third is the task being serviced by the secondary memory or I/O, and fourth is the task waiting in a queue for processor service. This, then, can be modeled by a cyclic queuing model using two queues and two service facilities (Figure 12). The criterion used for evaluating the memory hierarchy here will be the ratio of the actual number of instructions executed by the processor to the maximum number of instructions executed provided all the memory was

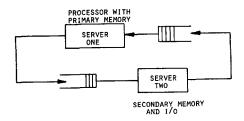


FIGURE 12: TWO SERVER QUEUING MODEL

substituted by the level having the fastest speed. All servers are assumed to have an exponental service time distribution, and a FIFO scheduling philosophy is assumed for all queues in the system. For a detailed analysis see Rege (12).

A typical processor activity is characterized as an instruction fetch, instruction decode, data fetch and data operation (Figure 13). Using this simplified processor behavior, the average time interval between the issuance of successive memory accesses can be determined.

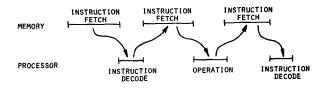


FIGURE 13: MODEL FOR PROCESSOR AND MEMORY OPERATION

If λ is assumed to be the average service rate of the processor and primary memory and h the hit ratio, then the mean execution interval $1/\lambda$ can be expressed as (3),

$$1/\lambda = \frac{h}{(1-h)} \{t (M_p) + t (P_c)\}$$

Where t(M_p) = aggregate access time for the primary memory,

 $t(P_c)$ = average processing time between successive memory accesses.

Assuming μ as the service rate for the second server the probability of CPU being busy (CPU utilization) is given by:

U = probability of CPU being busy

$$= \frac{1 - \rho^{D}}{1 - \rho^{D+1}} \quad (\text{Ref. (7)})$$

Where D = the degree of multiprogramming and $\rho = \frac{\lambda}{\mu}$

Once the CPU utilization is found, then the figure of merit (f) can be derived as:

$$= \frac{t (P_c) + t (fastest memory)}{t (P_c) + t (M_p)}$$

Where t (fastest memory) = access time of the fastest memory.

8. A MEMORY HIERARCHY DESIGN

The final outcome of a memory system design is its cost and performance. Invariably, the requirements are to minimize the cost while maximizing the performance. The cost and performance of the memory system are a function of the technologies T_1 , T_2 T_n and their sizes S_1 , S_2 ... S_n used at any level.

Reduction in the cost of the memory system necessitates a small amount of memory at the levels nearer the processor, whereas increase in performance necessitates a large amount of memory at the levels nearer the processor. Therefore, a problem encountered in the design of the memory hierarchies is that of finding a mix of memories for different levels in the hierarchy that would give an optimum performance for a given cost.

Assume that a certain cost constraint exists for the design of the primary memory. Also assume that a two level hierarchy of technologies T1 and T2 with sizes S1 and S2, respectively, satisfies the cost constraint and places the hierarchy at point A on the hit ratio characteristics. A three level hierarchy using technologies T1, T2, and T3 having the same cost as above is one that has memory sizes of S1, S2, and S₃, respectively, such that S₃=X. (S₂-S₂), and X>1 is the cost/bit ratio between the technology T2 and T3. Let these sizes place the memory hierarchy design at Point B on the hit ratio curve (Figure 11). This, then, is a constant cost conversion. Since S3>S2 the primany memory hit ratio is improved.

The analysis here will consider two cases: (a) A two level memory hierarchy of Bipolar and MOS, and (b) A three level memory hierarchy of Bipolar, MOS and CCD. For CCD's we will assume that three types of devices having fast, medium and slow speed are used. In a real world these may be Line Addressable, Circulating Shift Register, or Series Parallel Series organization. Since the cost varies as the performance, we will assume 3, 4 and 5, as the cost ratio between the MOS RAM and the different CCD memories, respectively. Even though these cost ratios may not be exact in a given situation, they are within the ranges as forecasted by Martin (9). In any particular situation, the method developed here should be used by substituting the actual values.

With the parameters as shown in Table 1, graphs of the MOS memory size/program versus the Figure of Merit for a two level (Bipolar, MOS) hierarchy and a three level (Bipolar, MOS, CCD) hierarchy for various CCD chip organizations are drawn in Figure 14. It is easily seen that for the parameters chosen the different CCD organizations are best suited for the different regions of the requirements for MOS memory size/program. For example, SPS organization gives the best performance between 8K to 20K of MOS memory/program, whereas the Circulating Shift Register type device has the highest advantage between 20K to 40K of MOS memory/program. This shows that all the three CCD chip organizations have a place in a memory system design, depending on the amount that a designer wishes to spend on the memory system.

NAME	SYMBOL	CHARACTERISTICS	REMARKS
PROCESSOR	т(Р _с)	500 NSEC.	THE PROCESSOR CHARACTERISTICS ARE FOR AVERAGE TIME BETWEEN IS- SUANCE OF MEMORY REQUESTS.
BIPOLAR MEMORY	T ₁ .	100 NSEC.	
mos Memory	T ₂	500 NSEC.	
CCD Memory	T ₃	Fast 40 usec. Med. 192 usec. Slow 400 usec.	COST RATIO (mos/ 3 ccd) 4 5
DISK AND I/O	-	10 msec.	
DEGREE OF MULT PROGRAMMING	I- D	8	

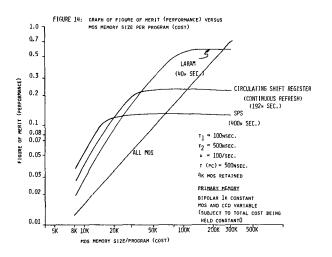


9. CONCLUSIONS

Computer Memory System design implications of different CCD chip organizations have been evaluated using single server and two server queuing network models. The performance is evaluated in a stand alone mode, and as a memory technology with an access time between a MOS memory and disk. Some suggestions have been made to improve the performance of a memory system designed using CCD's.

The performance comparison of CCD organization in a stand alone mode shows that SPS has the worst performance, as one should expect. The interesting result shown is that the Circulating Shift Register with a burst mode for refreshing has a better performance than the LARAM organization. Circulating Shift Register organization with cache has the best performance of all, but will require a considerable amount of extra cost. If a cache were to be used with the LARAM and the SPS organization, then similar performance improvements can be expected.

It is shown that the CCD as a gap filling technology has tremendous advantages. It is also shown that for the cost and performance parameters chosen, all three organizations, i.e., SPS, Circulating Shift Register, and LARAM have a place in a Computer Memory System design.



REFERENCES

- Abate, J., and Dubner, H.: "Optimizing the Performance of a Drum-Like Storage." <u>IEEETC</u>, C-18, Nov. 1969, pp. 992-996.
 Amelio, G. F.: "Charge-Coupled Devices
- (2) Amelio, G. F.: "Charge-Coupled Devices for Memory Applications," <u>NCC, AFIPS</u> <u>Conference Proceedings</u>, Vol. 44, May 1975, pp. 515-522.
- Bhandarkar, D. P., and Juliussen, J. E.: "Computer System Advantages of Magnetic Bubble Memories," <u>Computer</u>, Vol. 8, No. 11, Nov. 1975, pp. 35-40.
- (4) Denning, P. J.: "Thrashing Its Causes and Prevention," <u>AFIPS Proceedings</u>, FJCC 1968, Vol. 33, Part I, pp. 915-922.
- (5) Feller, W.: An Introduction to Probability Theory and its Application, <u>Wiley</u>, New York, 1968.
- (6) Fuller, S. H., and Baskett, F.: "An Analysis of Drum Storage Units," <u>JACM</u>, Vol. 22, No. 1, Jan. 1975, pp. 83-105.
- Hiller, F. S., and Lieberman, G. L. : Introduction to Operations Research, <u>Holden-Day</u>, San Francisco, 1967.
- Laur, H.: "Bulk Core in 360/67 Time Sharing System," <u>AFIPS Proceedings</u>, FJCC 1967, Vol. 31, pp. 601-609.
- (9) Martin, R. R., and Frankel, H. D.: "Electronic Disks in the 1980's," <u>Computer</u>, Vol. 8, No. 2, Feb. 1975, pp. 24-30.
- (10) Panigrahi, G., Woo, B., and Chu, B.: "Charge Coupled Memory Test Philosophy," <u>Digest of</u> <u>Papers, Semiconductor Test Symposium</u>, Cherry Hill, N. J., Oct. 14, 1975, pp. 9-17.
 (11) Papenberg, Bob: "Design and Application of
- (11) Papenberg, Bob: "Design and Application of Intel's 2416 16K Charge Coupled Device," Application Note AP-14, Intel Corporation, 3065 Bowers Avenue, Santa Clara, California 95051.
- (12) Rege, S. L.: "Cost, Performance and Size Tradeoffs for Different Levels in a Memory Hierarchy," <u>Computer</u>, Vol. 9, No. 4, April 1976, pp. 43-51.