Report Number: WUCS-84-5

1984-07-01

Design Analysis of a Wide-Band Picture Communication System

Chung-Dak Shum, Jerome R. Cox, and G. James Blaine

A design study of a Picture Communication Network is presented with special emphasis on those issues that differ from conventional networks. The large amount of data in an image together with the requirement for a rapid response time make a wide-band transmission medium a necessity. The wide-band medium is perceived as a collection of parallel broadcast channels. Four different logical organizations are proposed for this architecture. Design equations are developed and cost-performance curves are generated. Advantages of one organization over another under different circumstances are pointed out.

Recommended Citation

Shum, Chung-Dak; Cox, Jerome R.; and Blaine, G. James, "Design Analysis of a Wide-Band Picture Communication System" Report Number: WUCS-84-5 (1984). *All Computer Science and Engineering Research.*
https://openscholarship.wustl.edu/cse_research/863

DESIGN ANALYSIS OF A WIDE-BAND PICTURE COMMUNICATION SYSTEM

Chung-Dak Shum, Jerome R. Cox
and G. James Blaine

WUCS-84-5

To appear in the Proceedings of
MEDPACS '84, Arlington, Virginia, July 24-27, 1984.
Published by IEEE Computer Society Press.

Department of Computer Science
Washington University
St. Louis, Missouri 63130

July 1984

DESIGN ANALYSIS OF A WIDE-BAND PICTURE COMMUNICATION SYSTEM

Chung-Dak Shum*, Jerome R. Cox and G. James Blaine

Department of Computer Science, Biomedical Computer Laboratory
Washington University School of Medicine, St. Louis, MO 63130
*Current address: Department of Computer Science, UCLA

Abstract

A design study of a Picture Communication Network is presented with special emphasis on those issues that differ from conventional networks. The large amount of data in an image together with the requirement for a rapid response time make a wide-band transmission medium a necessity. The wide-band medium is perceived as a collection of parallel broadcast channels. Four different logical organizations are proposed for this architecture. Design equations are developed and cost-performance curves are generated. Advantages of one organization over another under different circumstances are pointed out.

Introduction

A picture network differs from traditional computer networks in that the basic unit of information involves several orders of magnitude more data. A typical picture with 512x512 pixels and 8-bit gray scale corresponds to 2 megabits of information. In contrast, an I/O record size from tens to thousands of bytes are common for most computers. The preponderance of these large blocks of image data and the requirement for rapid network response in radiology applications make the wide-band transmission medium a necessity.

A communication framework, based on broadband CATV technology, has been installed at the Mallinckrodt Institute of Radiology (MIR) facilitating picture networking experiments[1]. On the architecture level, this PACS workbench can be viewed as providing multiple parallel broadcast channels.

Three types of nodes[2] can be identified in a PACS, namely: picture source, picture archive and picture viewing station. The main objective of the picture network is to allow node-to-node communication. The basic problem to be addressed here is how to tie the nodes together to achieve both good performance and a cost-effective system.

We begin our discussion by developing a simplified model of the picture network. Next, we identify a few potential logical organizations for the bus architecture. Appropriate queueing models for different organizations will then be derived and analyzed. Third, a cost function in terms of the channel bandwidth is defined and an estimate is made of the network cost for each organization. Lastly, we create a set of engineering design curves which can be useful in making PACS design decisions.

Functional Classification

Let us begin by examining the information exchange between the picture archive and the picture viewing station. Viewed from a user perspective, the most important functions of a viewing station are those concerned with images. Interactive display, of course, is the one that draws most attention. A typical image retrieval protocol consists of an image request, an acknowledgement and an image delivery. Conceptually, we would like to perceive the viewing station as comprised of two functions, a user interface function and a display function. The user interface function is responsible for all user interactions with the PACS such as retrieving and updating a patient's record, making an image request or processing an image. On the other hand, the display function buffers the images received from the network and allows some simple image manipulation such as windowing, zooming and gray-scale enhancement. In a similar manner, the picture archive also consists of two functions: 1) the control/processing function which responds to all queries and updates directed to it; 2) the archive function which stores images as well as delivers them over the network. Figure 1 shows a functional partitioning of both the picture viewing station and the picture archive interconnected by a communication network.

We anticipate that this link between the picture archive and the picture viewing station will carry the heaviest traffic of all links in the communication network.

Message Network and Service Network

The communication network, obviously, provides a "message link" between the user interface function and the control/processing function. In addition it also transmits images from the archive function to the display function via the "service link". Image transmission, of course, is distinguished by its large data volume; it is not uncommon that the information flow on the service link is several orders of magnitude more than that on the message link. Another, perhaps less apparent, difference is in the basic nature of the communication. Typically, the message link supports a query-based conversation, characterized by alternating question and answer sequences. In comparison, the service link is responsible for delivering a large packet from one node to another. In short, the message link is bidirectional while the service link is unidirectional.

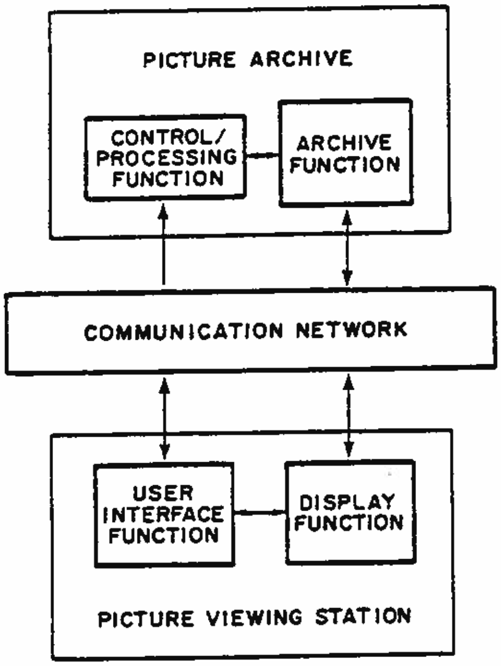

Figure 1 Functional partitioning of the picture archive and the picture viewing station

The message link and service link concepts suggest a functional division of the communication network into two parts, the message network and the service network (Figure 2). The message network, with the characteristics defined above, fits nicely into the domain of local area networks, a topic which is well documented$^{3-5}$ and not explored further here.

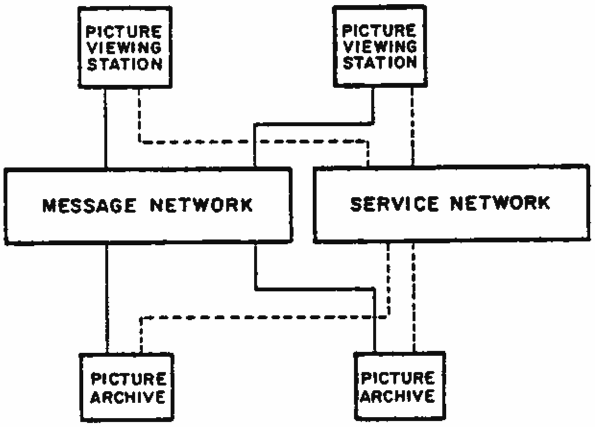

Figure 2 Functional partitioning of the communication network

Service Network Model

The rapid delivery of an image or a large data packet from its source to its destination is the main responsibility of the service network. As an overview, let us consider a scenario in which a user tries to fetch an image from a particular picture archive. Initially, a logical link between the user station and the picture archive must be set up to handle all query-based conversation. Once a picture request has been assembled, it is dispatched via the message network to a request queue where it waits for service. In general, a first-come-first-serve queueing discipline is enforced to ensure fairness. When its turn arrives, the image is delivered through the service network.

The above description, as depicted in Figure 3(a), gives only a simplified picture of the entire image retrieval process. One important time-consuming procedure we ignore is the access time for the requested image in the image database. In some cases it may be negligible compared to the transmission time. If not, taking a simplified view, we hope to gain needed insight into the picture communication problem itself.

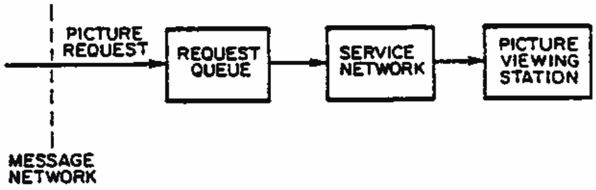

Figure 3(a) Abstract representation of an image retrieval process

A particular archive may serve more than one viewing station at the same time; this is the whole purpose for the request queue. We begin by modeling the arrivals of the picture requests from different nodes as an arrival process. The statistical behavior of the arrival process will certainly depend on the imaging modality, the patient population, and will likely involve even the time of the day. With all these complicated factors interacting with one another, any attempt to arrive at an exact representation will certainly prove fruitless. A Poisson distribution is assumed for the arrival process, mainly because of its simplicity, tractability and general appropriateness.

A straightforward translation from our image retrieval scenario to a standard queueing system model is given in Figure 3(b). Using the widely known queueing system language, we refer to each picture request arrival as a customer's arrival, the request queue simply as the queue, and the service network as the server. As will be seen, most of our studies stem from this simple, yet versatile, queueing model.

Logical Organization

The existing PACS workbench at MIR provides multiple, parallel broadcast channels on a broadband cable and makes possible a variety of logical network organizations. In this section, we will look at a few of these possibilities and present a qualitative description of each of them in subsequent paragraphs. For generality, and in this section only, we will not distinguish between archive and viewing-station nodes.

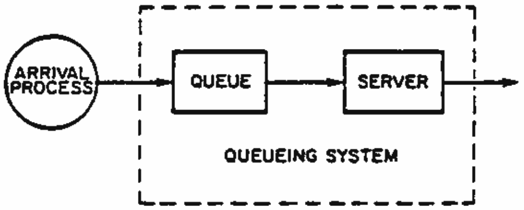

Figure 3(b) Translation of the image retrieval process into the standard queueing system model

Point-to-Point Channels

As its name suggests, a point-to-point network consists of a set of nodes interconnected by dedicated channels. In a broadband communication-environment, point-to-point channels need not be implemented by dedicated physical links, instead, frequency-division multiplexing can be utilized. In essence, each link between a pair of nodes corresponds to a unique frequency band as illustrated in Figure 4. One important observation regarding this implementation is the cost of constructing each channel. The channel capacity or bandwidth accounts for a major portion of the cost, but, unlike most other point-to-point networks, the distance between nodes contributes no additional cost beyond the initial cable installation.

An obvious question about a point-to-point organization is whether it is fully connected or partially connected. A partially connected network means fewer links and in general, a lower cost. This argument, however, is not based upon the assumption that the cost per link is constant. Most often the cost per link depends on its bandwidth, and this in turn varies with the total number of links in the network. In order to attain the same network performance, a partially connected network has to have a higher bandwidth for some of its links. A bandwidth economy-of-scale comes into play and leads to a lower cost.

From the design point of view, a partially connected network depends heavily on the available traffic data; something that is usually obtained by guessing during the design phase. A small deviation in the traffic pattern may degrade the network performance considerably. One further observation is the necessity of buffering between non-compatible channel bandwidths in a partially connected network. With the large packet size involved, the digital buffer required may account for a significant portion of the network cost. More importantly, the associated network delay may make rapid response difficult to achieve.

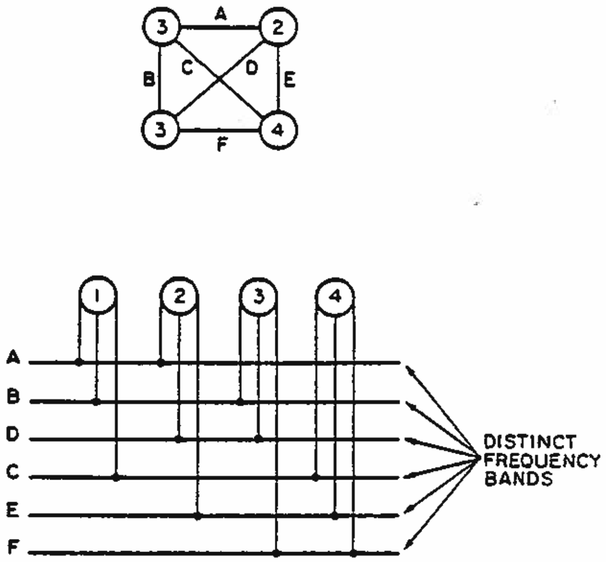

Figure 4 Implementation of Point-to-Point Channels by frequency division multiplexing

The above considerations seem to argue against a partially connected network. We will, thus, tailor our discussion towards fully connected networks only. Specifically, each picture archive is connected to each picture viewing station.

Single Shared-Transmit/Receive Channel

A fully connected point-to-point network and a single high-bandwidth shared channel are at opposite extremes organizationally. Often, a shared resource can only be allocated to a single user at any particular instant. To resolve the conflicting demands, certain time-division allocation mechanisms must be used. Two general classes of time-division multiplexing are synchronous time-division multiplexing (STDM) and asynchronous time-division multiplexing (ATDM).

In STDM, as shown in Figure 5, each node is assigned a fixed time slot on the shared channel for transmission of information. Any node on the channel can receive from every time slot and, therefore, can listen to every other node. Source addressing is usually not required, since each node is identified by its time-slot position. To accommodate an uneven traffic flow between node pairs, additional slots can be statically allocated to those nodes with high traffic.

The STDM technique has been known for many years in the field of data communication. The technology concerned with building STDM systems has already matured, and implementation is straightforward. Another less obvious merit of STDM is a deterministic reponse time. Once a node has something to send, the time the node spends waiting for its own slot does not depend on the current activities of the other nodes on the channel. This ensures more independence in the operation of the nodes.

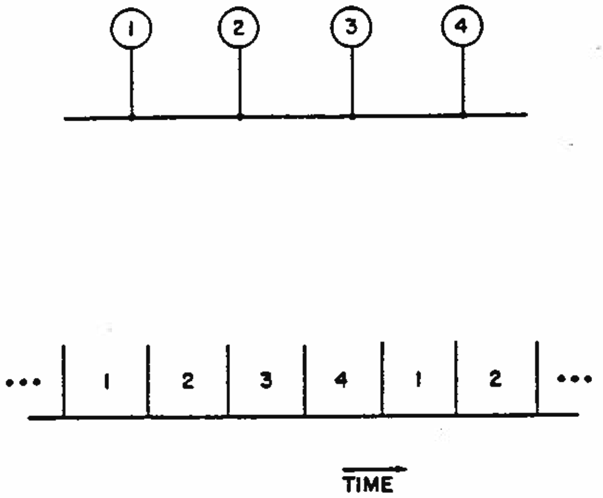

Figure 5 Example of a four-node STDM

The STDM technique also has certain disadvantages. Channel utilization is inefficient due to the fact that an assigned slot is only being used a portion of the time. In other words, a slot is left idle when its owner is not transmitting anything during that interval. Furthermore, if the STDM scheme is implemented by a centralized synchronization mechanism, as are most systems today, robustness is not upheld. Deviant behavior of the centralized component may cause the communication system as a whole to fail. Also, when the distance between the two ends of the cable becomes too long, timing may pose difficulties for system implementation. Since, the physical extent of our PACS is not expected to grow too large, STDM may still be feasible.

To increase channel utilization, the ATDM technique has been proposed[6]. The basic idea is to employ a more sophisticated channel allocation algorithm in which a time slot is acquired and released as it is needed. With such an arrangement, each node is granted access to the channel only when it has data to send. A typical segment of an ATDM data stream is shown in Figure 6. The crucial attribute of the ATDM technique is that source and destination addressing is required for each transmitted packet.

The dynamic assignment policy, of course, pays off during light traffic. Unlike the STDM technique, a node does not have to wait for its own time slot when the channel is idle. However, the operation of one node is no longer independent of the others. If a first-come-first serve discipline is enforced, a requesting node may have to wait in a queue a length of time that increases with the activity of the other nodes.

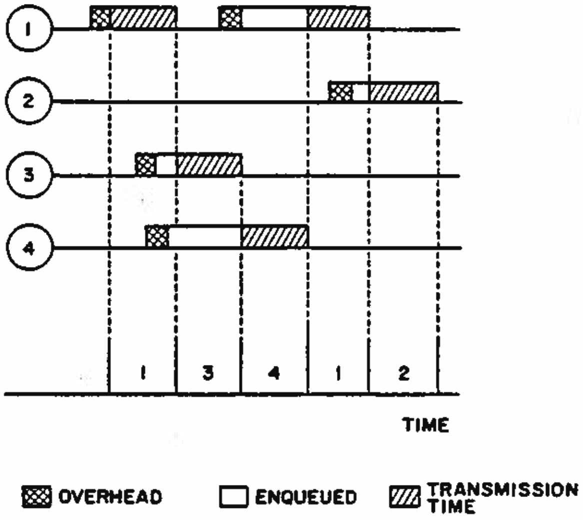

Figure 6 A time segment of a typical ATDM data stream

Multiple Shared-Receive Channels

We have already seen the two extremes, that is, the fully connected point-to-point channels and the single shared channel. Now let us look at something in between. The motivation for an intermediate configuration is twofold. First, the point-to-point dedicated link tends to exhibit low channel utilization. This is mainly due to the bursty nature of the traffic between any given pair of nodes. Second, a single high bandwidth shared channel may not be economical, since an inexpensive wide-band modem is not readily available. In order to fully exploit the capacity of a broadband system, a method utilizing multiple broadcast channels may be desirable.

The fundamental concept of Multiple Shared-Receive Channels is borrowed from that of radio broadcast. Basically, each archive, corresponding to a radio station, is assigned a distinct frequency band for broadcasting. Individual receivers, or viewing stations in our case, can select the appropriate channel by tuning their receiving frequency. The main idea of this scheme is reviewed in Figure 7.

The remaining question to be answered is the availability of a receiver with the ability to adjust its tuning to any of the frequencies allocated to the archives (henceforth called an agile receiver). In television reception, superheterodyne receivers with frequency synthesized tuning have been available for some years. Although we do not see any fundamental difference between the design of such an agile receiver for television and one for data communication, the latter are now quite costly. As the technology becomes more established, agile receivers are certain to become a more common and less expensive component. The study of Multiple Shared-Receive Channels can thus act as a guide to PACS organizations feasible in the future.

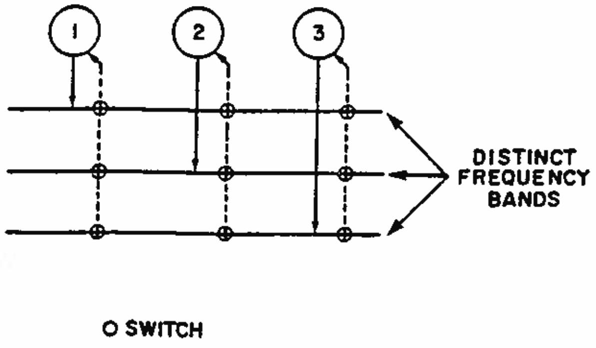

Figure 7 Conceptual view of Multiple Shared-Receive Channels

Multiple Shared-Transmit/Receive Channels

In the previous section, we discussed an organization based on the availability of agile receivers. Although a more efficient use of bandwidth is expected compared to the point-to-point channels, we may still encounter some inadequacy as the number of nodes increases. First, we have to realize that state-of-the-art technology only permits us to use a limited bandwidth, typically 400 MHz on a broadband coax. To make matters worse, a large chunk of the available bandwidth is usually reserved for other purposes, such as video and audio applications. Second, the tunable receiver may not be able to cover the entire frequency spectrum. A natural solution, of course, is to make more efficient use of the allocated bandwidth.

An obvious way to increase utilization is to assign multiple archives to a given frequency band. Within each frequency band, channel allocation is done basically the same way as in ATDM. Figure 8 describes this arrangement.

A similar, but more dynamic, scheme requires the use of frequency-agile transmitters as well as agile receivers. The basic idea is to treat the channels within the available frequency band as a pool of resources which many users share. Whenever an archive is ready to transmit a packet, a channel is allocated dynamically depending on which channel is idle at that particular moment. Figure 9 presents a conceptual view of this scheme.

The discussion about the availability of the frequency agile modem in the last section applies equally well to this scheme. The conceptual view of channel switching is implemented by frequency switching which makes the agile modem an absolute necessity.

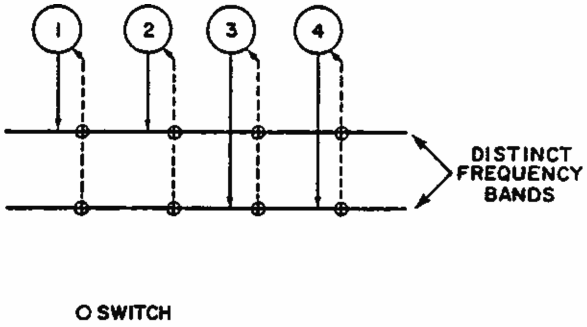

Figure 8 Conceptual view of assigning multiple archives to a given frequency band

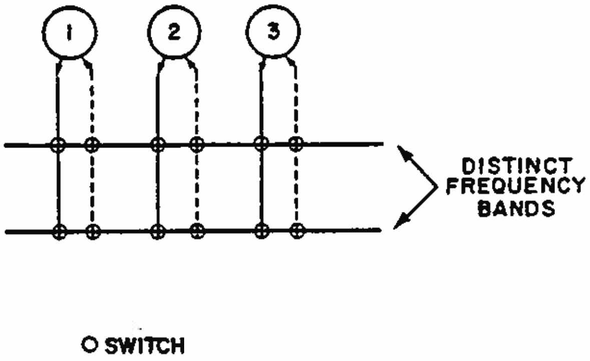

Figure 9 Conceptual view of dynamic channel allocations

Cost-Delay Analysis

The last section was devoted to the qualitative description of some possible service network organizations. In order to gain a better perspective on these organizations, we examine their cost-performance behavior. Response time is the performance criterion we choose for our attention.

Mean Delay Analysis

Instead of looking at each organization one by one, we are going to identify a general expression which will give us the solutions in all cases. To characterize the queueing system model of the service network, two probability distributions, the interarrival distribution and the service time distribution, must be specified unambiguously. An exponential distribution with 1/ as the mean interarrival time is chosen for the former. The rationale for this choice is twofold. First, the bursty nature of an exponential distribution matches well the nature of the arrival process. More important, the memoryless property of this distribution tends to produce a substantial simplification. For the present analysis, let us simplify the second probability distribution and assume that each request involves only one picture of constant size $1/\mu$ bits. For a channel capacity of C bits/sec, the transmission time portion of the service time is $1/(\mu C)$ sec. The balance of the service time depends upon the organization under consideration.

An extremely well known formula for the average waiting time W in an M/G/1 queueing system is commonly referred to as the Pollaczek-Khinchin mean-value formula. A simplified[7] version of the formula for W is

$$W = \frac{W_o}{1-\rho} \qquad (1)$$

where ρ is generally known as the utilization factor and is defined as the ratio of the rate at which work enters the system to the maximum rate at which the system can perform the work.

W_o is the average remaining service time for the customer (if any) found in service by a new arrival. From Eq. 1, it is not difficult to see that for a mean service time $\overline{x}$, the average time T a customer spends in the system is $\overline{x} + W$ or

$$T = \overline{x} + \frac{W_o}{1-\rho} \qquad (2)$$

This simple expression is the one upon which most of our analysis is based.

Cost Model

We compare the system's performance to its cost and begin by identifying the costly components of our CATV based system. The backbone of the system, of course, is a high quality coaxial cable. Interspersed along the cable are the line extender amplifiers. The head-end equipment is located at the root of the dual-cable tree structure. Although the components mentioned above account for a significant portion of the system cost, none of them are dependent upon a particular organization. We will refer to this fixed cost as the installation cost and exclude it from the cost function. Another component of the cost more directly related to the specific organization is the modem. This is true because the capacities of each channel in a particular organization are governed by the bandwidths of the associated modems. Since wide-band modems are the major organization-sensitive cost components, we define the cost function as the total modem cost.

One observation about the cost of a modem is that it tends to increase with bandwidth. However, the exhibited relationship is not linear. Economy of scale, as it always does, seems to play a role. We will model this as a power-law cost function. With D (dollars) denoting the modem cost and C (Capacity) representing the modem bandwidth, we have

$$D = C^{\alpha} \qquad (3)$$

where $0 < \alpha \leq 1$. Figure 10 shows a justification for this model. Notice the sharp increase in prices for modems of data rate beyond 1 Mbit/sec. The low demand for high speed modems and the immaturity of the modem technology might help explain this phenomenon. Of course, a different exponent should be used above 1 Mbit/sec.

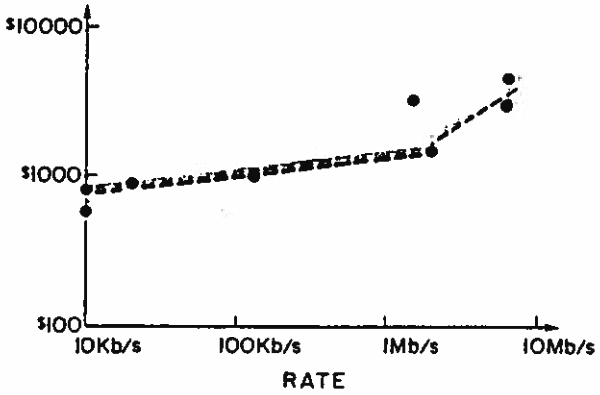

Figure 10 Relationship between cost and bandwidth for modems

Analysis of Different Organizations

Our objective here is to apply our mean delay analysis and cost model to each logical organization described above and to derive closed form expressions. First we need to know the traffic in the service network. In this analysis, the major traffic of concern is that between the archive and the viewing station and it is mainly in one direction, from the archive to the viewing station. The following two variables are introduced to describe the configuration of our service network and will be used throughout the discussion:

m = total number of picture archives

n = total number of picture viewing stations

We further assume that the average traffic from archive to viewing-station is given by λ pictures/sec. Although it is clear that this assumption is far from realistic, it gives us a stepping stone to the more general solution. Thus we infer that the total average traffic in the service network is given by $mn\lambda$, and the average traffic flowing from any particular archive adds up to $n\lambda$.

Next we use T_{avg} to denote the average response-time, the time between the acceptance of a picture request and the complete delivery of the desired image. Our design criterion T_{avg} is thus related to Eq. (2) by the following inequality:

$$T_{avg} > T = \bar{x} + \frac{W_o}{1-\rho} \qquad (4)$$

Allowing T to approach T_{avg} and replacing $\bar{x}$ by $1/(\mu C)$ we obtain the design equation

$$T_{avg} = \frac{1}{\mu C} + \frac{W_o}{1-\rho} \qquad (5)$$

The remaining work to be done is to identify the expressions for W_o and ρ for each organization. As will become apparent later, the channel capacity C will appear in each case. Solving for C and substituting it into Eq. (3), we obtain the cost of each organization.

Finally we would like to point out that the unidirectional traffic assumption implies a possible split of the digital modem into its two basic components, the modulator and the demodulator. The former is useful for the archive; the latter useful for the viewing station. In fact, some small amount of traffic from viewing station to archive is likely and would require both modem components in each unit. Thus, for simplicity, we will refer to each unit as a modem and make no distinction between the unidirectional and bidirection versions.

Point-to-Point Analysis

Specifically, the problem here is to find the total cost of implementing a "fully connected" point-to-point system; "fully connected" in the sense that every archive is directly linked to every viewing station. Since we assume a uniform traffic distribution, an analysis on any link applies equally well to every other link. From the definition of utilization ρ, we get

$$\rho = \lambda/(\mu C) \qquad (6)$$

Since Poisson arrivals take a random look at the system, the average remaining time for the customer found in service by a new arrival is given by

$$W_o = \frac{\rho\bar{x}}{2} = \frac{\lambda}{2(\mu C)^2} \qquad (7)$$

Making use of Eqs. (5), (6) and (7), we obtain

$$T_{avg} = \frac{2\mu C - \lambda}{2\mu C(\mu C - \lambda)} \qquad (8)$$

Solving for C, assuming a stable system ($\rho < 1$), we get

$$C = \frac{1}{\mu T_{avg}} + \frac{(1+\beta)\lambda}{2\mu} \qquad (9)$$

where

$$\beta = \frac{\sqrt{(\lambda T_{avg})^2 + 1^2} - 1}{\lambda T_{avg}} < 1$$

Applying our cost model, the cost D_1 for establishing one link is

$$D_1 = 2\left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)\lambda}{2\mu}\right]^{\alpha}$$

where the factor 2 accounts for the two modems in each link. With m archives and n viewing stations, the total number of links is mn. Multiplying the last equation by this factor, we finally obtain the total cost D_T of the system as

$$D_T = 2mn\left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)\lambda}{2\mu}\right]^{\alpha} \quad 0<\alpha,\beta<1 \qquad (10)$$

Single Shared-Transmit/Receive Channel Analysis

As described before, there are two subdivisions in this organization, namely, STDM and ATDM. Let us begin with the STDM model. Consider that each archive is assigned a unique fixed time slot and we ignore the negligible traffic flowing away from each viewing station. The total number of time slots in a period is then equal to m, the total number of archives. The traffic from each archive is considered as a Poisson process with mean $n\lambda$, where n is the total number of viewing stations. The length of each time slot corresponding to the transmission time of each image is given by $1/(\mu C)$. From this we figure out the period

$$P = m/(\mu C) \qquad (11)$$

Once the period is identified, the expression for utilization falls out readily as

$$\rho = n\lambda P = mn\,\lambda/(\mu C) \qquad (12)$$

Since each arrival has to wait for its own slot, the average waiting time is given by

$$W_o = \frac{P}{2} = \frac{m}{2\mu C} \qquad (13)$$

Applying Eq. (13) to Eq. (5), we have

$$T_{avg} = \frac{1}{\mu C} + \frac{m}{2(1-\rho)}\left(\frac{1}{\mu C}\right)$$

Additional simplification using Eq. (12) is possible if we assume m to be much greater than $2(1-\rho)$, in which case the channel capacity can be written as

$$C = \frac{m}{2\mu T_{avg}} + \frac{mn\lambda}{\mu} \qquad (14)$$

With a uniform traffic distribution, each archive must have a modem of channel capacity C, and in a similar manner, each viewing station must have a modem of channel capacity C. Utilizing our cost model, we obtain the total cost D_T of the system as

$$D_T = (m + n)\left[\frac{m}{2\mu T_{avg}} + \frac{mn\lambda}{\mu}\right]^{\alpha} \quad 0<\alpha<1 \quad (15)$$

In ATDM, the model will simply consist of a queue and a server. The former collects the picture requests and then lines them up for service in a first-come-first-serve manner. The latter is merely a transmission channel which serves the queue. Thus the mean arrival rate of our single-channel model is given by $mn\lambda$. Following the derivation in Point-to-Point analysis, and replacing λ by $mn\lambda$ in Eq. (9), we obtain

$$C = \frac{1}{\mu T_{avg}} + \frac{(1+\beta)mn\lambda}{2\mu} \qquad (16)$$

where

$$\beta = \frac{\sqrt{(mn\ T_{avg})^2 + 1^2} - 1}{mn\lambda T_{avg}} < 1$$

With m archives and n viewing stations, the total cost D_T of the system is

$$D_T = (m + n)\left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)mn\lambda}{2\mu}\right]^{\alpha} \quad 0<\alpha,\beta<1 \quad (17)$$

Multiple Shared-Receive Channel Analysis

The basic idea in this scheme is to allocate a unique frequency band for each archive. With our uniform traffic assumption, the average arrival-rate of requests at each archive is equal to $n\lambda$. We follow the same procedure as in Point-to-Point analysis. Substituting $n\lambda$ for λ in Eq. (9), we get

$$C = \frac{1}{\mu T_{avg}} + \frac{(1+\beta)n\lambda}{2\mu} \qquad (18)$$

where

$$\beta = \frac{\sqrt{(n\lambda T_{avg})^2 + 1^2} - 1}{n\lambda T} < 1$$

As has been pointed out, the success of this scheme depends on the availability of a frequency-agile receiver. However, one would expect the cost of such a device to be more than a conventional demodulator. We will thus introduce a factor $f<1$ to account for this additional expense. With m archives and n viewing stations, the total cost D_1 of the system based on our cost model is

$$D_T = (m + nf)\left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)n\lambda}{2\mu}\right]^{\alpha} \quad 0<\alpha\ \beta<1 \quad (19)$$

Multiple Shared-Transmit/Receive Channels Analysis

Earlier we discussed two forms of this organization. The key difference between the two is depicted in Figure 11. Consider the first alternative, that is, assigning multiple archives to a given frequency band. While this arrangement involves a smaller number of channels compared with Multiple Shared-Receive Channels, the bandwidth of each one increases. By observation, we conclude that the total cost D_T of this system, similar to Eq. (19), is

$$D_T = (m + nf)\left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)kn\lambda}{2\mu}\right]^{\alpha} \quad 0<\alpha,\beta<1 \quad (20)$$

where

$$\beta = \frac{\sqrt{(kn\lambda T_{avg})^2 + 1^2} - 1}{kn\lambda T_{avg}} < 1$$

and k denotes the number of archives associated with a channel. Notice that the total bandwidth required to implement such a system shows almost a k-fold decrease as compared to the case when k=1, namely, Multiple Shared-Receive Channels. However, a brief look at Eqs. (20) and (19) reveals that the former is more costly when k>1.

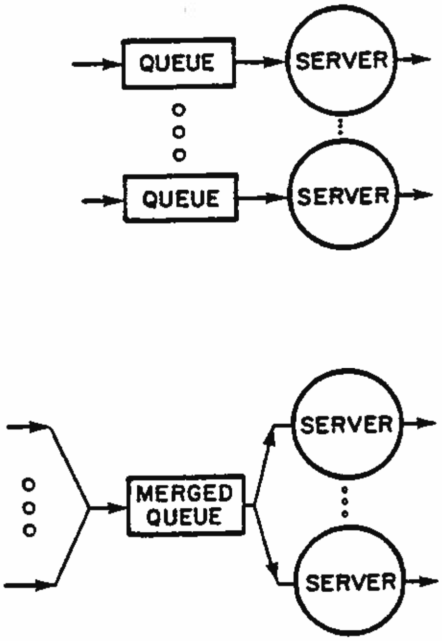

Figure 11 Difference between separate queue and merged queue

The other alternative can be modeled by a multiple-server system with c representing the total number of channels or servers. An explicit expression for this type of system is very complicated[8], nevertheless, our intuition tells us that the capacity C of each channel satisfies

$$C < \frac{1}{\mu T_{avg}} + \frac{(1+\beta)kn\lambda}{2\mu} \qquad (21)$$

where

$$\beta = \frac{\sqrt{(kn\lambda T_{avg})^2 + 1^2} - 1}{kn\lambda T_{avg}} < 1$$

and $kn\lambda = mn\lambda/c$ is the "average" traffic in each channel. The expression on the right of Eq. (21) represents the capacity necessary to satisfy the average response time constraint for an independent channel with average traffic $kn\lambda$. Since we have a merged queue in this case, work can be more evenly distributed among different servers. Thus a more efficient use of resources is expected and the capacity C of each channel can take on a smaller value. Consequently, the total cost D_T for implementing such a system is given by

$$D_T < (m + n)f \left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)kn\lambda}{2\mu}\right]^{\alpha} \quad 0<\alpha,\beta<1 \quad (22)$$

Notice that this scheme requires a frequency agile modem for both the transmitter and the receiver.

Summary and Discussion

The major findings in our analyses are summarized in Table 1. The physical meanings of the parameters associated with the equations are briefly reviewed here. The mean response time, T_{avg}, is the design constraint under which we desire our system to operate. The total number of archives and viewing stations are given by m and n, respectively. The parameter λ represents the average unidirectional traffic between any archive/viewing-station pair and $1/\mu$ denotes the constant picture size. The power index, α, introduced by the cost model, takes on values between 0 and 1. Lastly, the factor f accounts for the additional expense incurred in the utilization of frequency agile modems.

The first equation under Multiple Shared-Transmit/Receive Channels represents the case in which multiple archives are being assigned to each frequency band. A quick comparison with Multiple Shared-Receive Channels reveals that the latter are less expensive. As has been mentioned earlier, the former pays a higher price for the more efficient use of the bandwidth. However, minimizing bandwidth is hardly our objective at the present moment, as we do not foresee any bandwidth shortage in the development. We will, thus, exclude multiple assignments of archives to a frequency band from further consideration.

The other alternative for Multiple Shared-Transmit/Receive Channels involves dynamic channel allocation. Unless we do not have enough bandwidth for every archive, or the receiver's range is not broad enough to cover the entire spectrum, there is no reason for this scheme. Assuming neither of these problems apply, we will also exclude this scheme from further consideration.

In order to gain a better understanding of the remaining cost-delay equations, let us rewrite them in the following manner. First, we assume that the total number of archives m is much less than the total number of viewing stations n. Or equivalently, we replace (m + n) and (m + nf) by n and nf respectively. Next we divide each equation by n and present the logarithm of the results in Table 2.

Point-to-Point Channels

$$D_T = 2mn \left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)\lambda}{2\mu}\right]^{\alpha}; \quad \beta = \frac{\sqrt{(\gamma\lambda T_{avg})^2 + 1^2} - 1}{\gamma\lambda T_{avg}}$$

where $\gamma = 1$ and $0<\beta<1$

Single Shared-Transmit/Receive Channel

$$\text{STDM: } D_T = (m + n) \left[\frac{m}{2\mu T_{avg}} + \frac{mn\lambda}{\mu}\right]^{\alpha}$$

$$\text{ATDM: } D_T = (m + n) \left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)mn\lambda}{2\mu}\right]^{\alpha}; \quad \gamma = mn$$

Multiple Shared-Receive Channels

$$D_T = (m + nf) \left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)n\lambda}{2\mu}\right]^{\alpha}; \quad \gamma = n$$

Multiple Shared-Transmit/Receive Channels

$$D_T = (m + nf) \left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)kn\lambda}{2\mu}\right]^{\alpha}; \quad \gamma = kn$$

$$D_T < (m + n)f \left[\frac{1}{\mu T_{avg}} + \frac{(1+\beta)kn\lambda}{2\mu}\right]^{\alpha}; \quad \gamma = kn$$

Table 1 Summary of cost-delay equations

Now the left hand side of each equation acts as a measure of the system cost per viewing station. A careful study on the right hand sides reveals that the term $\alpha \log(1/\mu)$ is present in all four equations. Whereas the packet size $1/\mu$ plays an important role in the absolute cost of each station, it has nothing to do with the relative costs among different organizations. Furthermore, the parameters T_{avg} and λ appear within the square brackets of each equation. If we loosen the design constraint, that is, allowing a larger value for the mean response time, T_{avg}, it is not difficult to see that the system cost decreases. On the other hand, an increase in the average traffic parameter λ would lead to a more expensive system.

It may be worthwhile here to take a good look at each equation in Table 2. We begin with Point-to-Point Channels. Notice that the design constraint, T_{avg}, and the traffic parameter λ are incorporated respectively in the first and second term in square brackets. At light traffic, (λ small), the term $1/T_{avg}$ dominates. Figure 12(a) illustrates this

Point-to-Point Channels

$$\log(D_T/n) = \alpha \log\left[\frac{1}{T_{avg}} + \frac{(1+\beta)}{2}\lambda\right] + \alpha \log \frac{1}{\mu} + \log(2m)$$

Single Shared-Transmit/Receive Channel

$$\text{STDM:} \quad \log(D_T/n) = \alpha \log\left[\frac{m}{2T_{avg}} + mn\lambda\right] + \alpha \log \frac{1}{\mu}$$

$$\text{ATDM:} \quad \log(D_T/n) = \alpha \log\left[\frac{1}{T_{avg}} + \frac{(1+\beta)}{2} mn\lambda\right] + \alpha \log \frac{1}{\mu}$$

Multiple Shared-Receive Channels

$$\log(D_T/n) = \alpha \log\left[\frac{1}{T_{avg}} + \frac{(1+\beta)}{2} n\lambda\right] + \alpha \log \frac{1}{\mu} + \log f$$

Table 2 Simplified version of cost-delay equations. Note that the β's are defined as in Table 1

situation when λ takes on values between 0.01 and 0.001. The additional term log(2m) is mainly responsible for the high cost of this organization. This extra term is introduced because each viewing station is directly connected to each of the m archives and it requires two modems for every link.

In Multiple Shared-Receive Channels, all traffic from a particular archive to n viewing stations flows in a unique channel, explaining the factor n in the traffic term. As one would expect, the system cost is more sensitive to λ here than that of Point-to-Point Channels; this is shown in Figure 12(b). The last term log f accounts for the extra cost incurred by frequency agility.

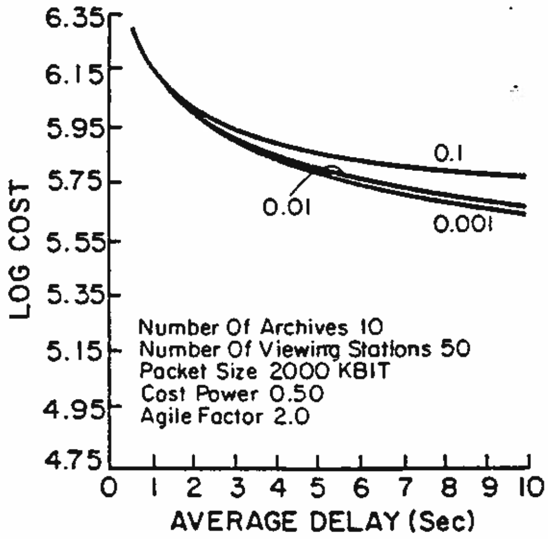

Figure 12(a) Cost-delay(mean) plot for Point-to-Point Channels

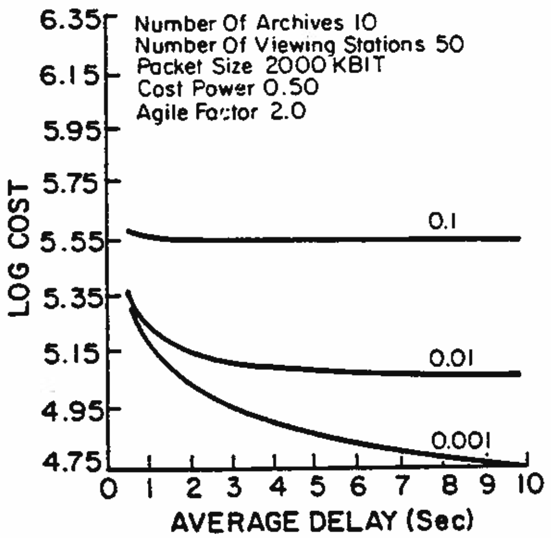

Figure 12(b) Cost-delay(mean) plot for Multiple Shared-Receive Channels

Finally we examine the two equations under Single Shared-Transmit/Receive Channel. Comparing the STDM equation with others, we immediately observe that the term associated with the design constraint T_{avg} presents itself differently. In what follows we give a physical interpretation for this. Under a very light traffic situation, it is not difficult to see that the channel capacity required to implement a non-STDM system is proportional to $1/T_{avg}$. On the contrary, this is not true for STDM. A request which arrives at an idle archive has to wait for the transmission slot of that particular archive. For a light traffic condition, we would expect on the average m/2 slots to pass by before any transmission occurs. In other words, the actual transmission only takes up a fraction, approximately 1/(m/2), of the system time. A higher channel capacity, proportional to $(m/2)(1/T_{avg})$, is therefore required to deliver the same amount of data. The additional m/2 factor makes STDM more sensitive to the change in the design constraint T_{avg}. Despite the fact that STDM directs the information flow from each archive through a distinct time slot, the traffic term is still characterized by mn λ, the average total traffic. Likewise, ATDM has all traffic merged into one channel. The extra multiplier, mn, associated with the traffic terms in STDM and ATDM, makes both system costs much more sensitive to λ ; this is illustrated separately in Figures 13(a) and 13(b).

Since all plots in Figures 12 and 13 are to the same scale, it is obvious that Point-to-Point Channels are not a cost effective solution. Apparently the lower bandwidth requirement for each channel does not justify the 2m-fold increase in the number of modems necessary for this implementation. Nevertheless, a heavy traffic situation and a relatively slow response requirement would make its cost comparable to that of Single Shared-Transmit/Receive Channel.

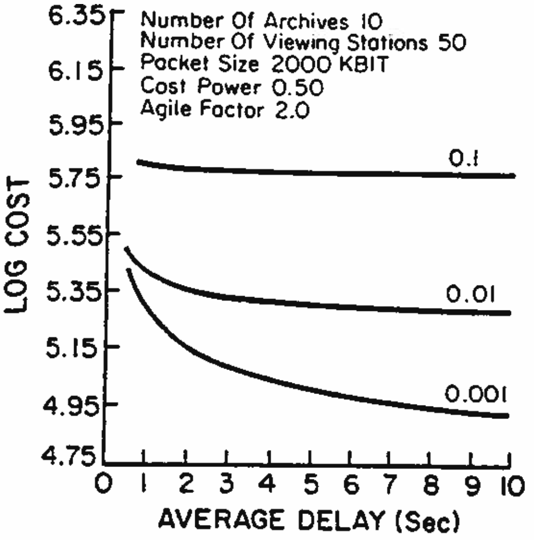

Figure 13(a) Cost-delay(mean) plot for STDM

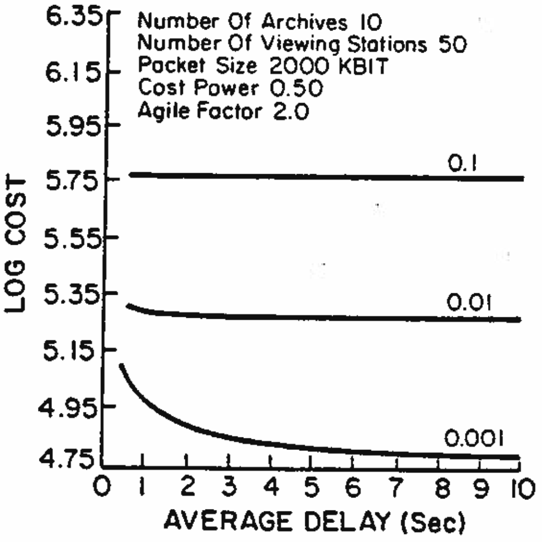

Figure 13(b) Cost-delay(mean) plot for ATDM

A comparison between STDM and ATDM is shown in Figure 14. It can be seen that ATDM always costs less than STDM. The equations listed in Table 2 clearly back up this observation. Since β lies between 0 and 1, the traffic term in ATDM is obviously smaller. The lack of an $m/2$ factor in the design-constraint term again makes ATDM less expensive. Intuitively STDM wastes more bandwidth and hence results in a relatively more expensive system under light traffic conditions. Figure 14 illustrates this point. At λ equals 0.001, the traffic term is negligible and the advantage of ATDM over STDM becomes apparent. On the other hand, the implementation costs approach one another as λ increases.

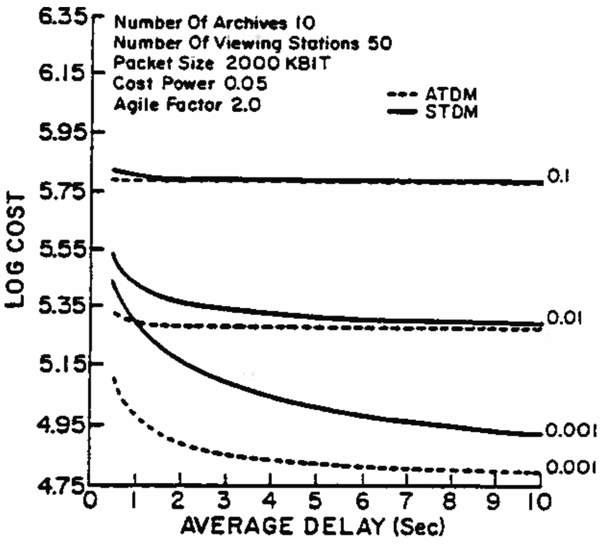

Figure 14 Comparison between STDM and ATDM

Now let us turn to the comparison between ATDM and Multiple Shared-Receive Channels. We see from Table 2 that the latter includes an extra term, $\log f$. To make the two expressions comparable, we multiply the expression in the square bracket by $f^{(1/\alpha)}$. This effectively increases the sensitivity of the system cost to λ. In spite of the extra multiplier $f^{(1/\alpha)}$, it is unlikely that the traffic term will get as large as that in ATDM. In other words, we expect $f^{(1/\alpha)}$ to be much less than m. Thus, Multiple Shared-Receive Channels appears to be more advantageous for large λ. At light traffic situations, however, the term associated with the T_{avg} dominates and the extra multiplier again exerts its effect and drives the cost up. Figure 15 shows the relationship between the two.

The foregoing discussion illustrates that no single universal winner exists in our candidate organizations. In particular, light traffic favors ATDM, but heavy traffic favors Multiple Shared-Receive Channels.

Although CATV technology was chosen as the basic communication framework, our results can apply equally well to fiber optics or multiple cable systems. In the discussions here, the wide-band medium is viewed as a collection of parallel broadcast channels. A multiple cable system, of course, automatically fits this model. Different techniques, however, are used to switch channels. In CATV systems, agile modems are used, while in multiple cable systems, physical switches are necessary. In fiber optics, a broadcast channel can be implemented by a "star" configuration with a central passive coupler. It is an easy matter to extend this technology to a collection of parallel broadcast channels by bundling multiple fibers within a single cable. As the fiber optic technology matures, Wavelength Division Multiplexing may also be utilized to implement multiple broadcast channels for this medium.

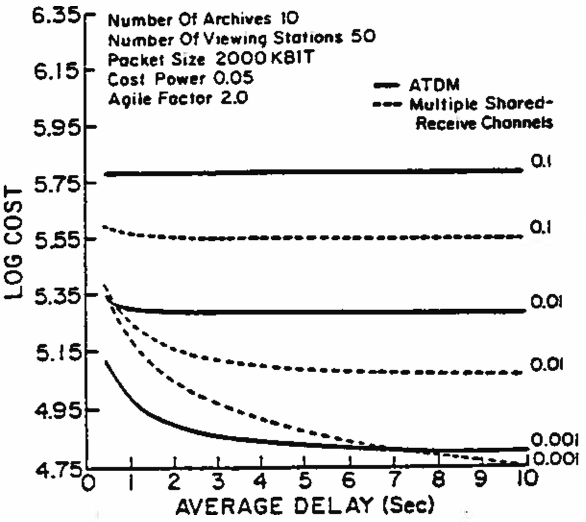

Figure 15 Comparison between ATDM and Multiple Shared-Receive Channels

ATDM has become the leading communication method utilized by local area networks. It is based on a single channel of fixed bandwidth and provides both high utilization of the channel and good response time. Because the physical characteristics of many transmission media under consideration for picture communication networks lend themselves to multiple channels, restricting design considerations to a fixed bandwidth and a single channel may be undesirable. This study compares several alternative logical organizations for a wide-band CATV network capable of supporting many channels by frequency division multiplexing. The aggregate bandwidth available on such a network is large and not a prime factor in system cost. Access to the network through modems does represent a major cost and one that is quite sensitive to the network's logical organization. Our analysis shows that such a multiple channel organization can be more cost effective than ATDM under heavy traffic conditions. This conclusion results from a favorable trade-off between cost and bandwidth of wide-band modems.

Acknowledgments

This study was supported in part by NIH grants RR00396 and RR01380 from the Division of Research Resources, by the Mallinckrodt Institute of Radiology, and by Washington University.

References

1. G.J. Blaine, R.L. Hill, J.R. Cox, and R.G. Jost, "PACS Workbench at Mallinckrodt Institute of Radiology (MIR)", Proceedings of 1983 PACS Conference. SPIE Vol. 418, pp. 80-86, Kansas City, 1983.

2. J.R. Cox, et al, "Some Design Considerations for Picture Archiving and Communication Systems", Computer Magazine, Vol. 16, No. 8, pp. 39-49, Aug. 1983.

3. A.S. Tanenbaum, COMPUTER NETWORKS, Prentice-Hall, Englewood Cliffs, New Jersey, 1981.

4. W. Stallings, LOCAL NETWORKS: AN INTRODUCTION, The MacMillan Publishing Co., Inc., New York, New York, 1983.

5. K.J. Thurber and H.A. Freeman, "Architecture Considerations for Local Area Networks", Proceedings of the First International Conference on Distributed Computing Systems, pp. 131-142, October 1979.

6. W.W. Chu, "A Study of the Technique of Asynchronous Time Division Multiplexing for Time-Sharing Computer Communications", Proceedings of the Second Hawaii International Conference on Systems Science, pp. 607-710, January 1969.

7. L. Kleinrock, QUEUEING SYSTEMS, Vol. 1, Chapter 5, John Wiley & Sons, 1975.

8. T.L. Saaty, ELEMENTS OF QUEUEING THEORY, Chapter 6, McGraw-Hill Book Company, 1961.