INVESTIGATIONS OF

INTERLIBRARY RESOURCE-SHARING NETWORKS

by

J. Francis Reintjes

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

Investigations of electronic interlibrary resource-sharing networks
have been made for the purpose of determining their applicability to library
operations and the ability of technology to meet operational and economic
requirements. It has been found that electronic networks of libraries are
especially suited to the movement of the information content of serials
between nodes. Requests involving serials are usually for individual
articles which average eight or so pages, a tractable number for document
delivery by electronic means. A system configuration is proposed. It con-
sists of an online-computer ordering subsystem and a separate digital
document-delivery subsystem. The salient elements of each subsystem are
discussed, including the combined index of document holdings, the ordering-
system software, document scanner and printers, communication links and re-
quirements for data compression. Also presented is an analysis of the
capital-equipment and operating costs for a four-node network with a maxi-
mum separation distance between nodes of ten miles and a ten-year projection
of total costs. The report concludes with a discussion of advantages of
resource-sharing networks to libraries and end users, and barriers to their
immediate adoption.

ACKNOWLEDGMENTS

CONTENTS

CHAPTER I

SYSTEM CONSIDERATIONS

SUMMARY

We report here our investigations of electronic interlibrary resource-sharing networks. Our purpose was to test the hypothesis that it is now technologically and economically feasible to move the content of documents electronically among nodes of a library network rather than to move documents themselves or photocopies thereof. Our research results confirm the validity of the hypothesis, subject to certain limitations on source materials and the geographic extent of the network.

We find that electronic interlibrary resource-sharing networks are ideally suited to situations where there are high-frequency occurences of internode requests for information contained in serials holdings, where a network can be formed with separation distances between nodes that do not exceed a few tens of miles and where the copy to be reproduced is in six-point type or larger. The body of literature included in serials is especially appropriate since requests are normally for a specific article that averages, typically, about six or eight pages. Metropolitan-area networks are presently most suitable for the application since electrical-communication costs can then be kept within reasonable bounds, and a lower bound of six-point type is desirable so that low-cost printers and solid-state scanners can be employed rather than more expensive laser scanners and printers.

BACKGROUND DATA ON INTERLIBRARY RESOURCE-SHARING

In light of modern-day photocopying machines, "photocopying in lieu of lending" has become the norm for exchange of serials holdings rather than outright borrowing and lending of the bound volumes. Typically an interlibrary request for a journal article is filled by photocopying the requested material and mailing or vehicle-delivering it to the requesting site. Photocopying for intra-library and interlibrary purposes is a well-

entrenched procedure in everyday library practice. [1]*

In order to gain insights into the magnitude of photocopying opera-
tions, we gathered information about borrowing and lending practices at
two large libraries -- the Massachusetts Institute of Technology library
system and the Library of Congress. Although these two entities obviously
do not embrace the entire spectrum of libraries in terms of size, quantity
of holdings, interlibrary traffic, and so forth, they do, nevertheless,
give insights into the nature of interlibrary activity among clusters of
research-oriented institutions.

In 1977 the Library of Congress (LC) initiated a photocopying service
for periodicals. These are supplied to federal libraries in the Washington,
D.C. area in lieu of lending. In 1978, 14,443 articles were photocopied
for this purpose; the average article length was 6.6 pages. During the
first four months of Fiscal Year 1979, photocopying was at an annual rate
of 26,000 articles per annum with an average length of 6.1 pages per
article. A comprehensive report on the characteristics of interlibrary-loan
requests at the Library of Congress is given in the Trevvett paper [2]. It
should be noted that LC appears to classify only those loans to libraries
outside the Washington, D.C. area as interlibrary loans.

The time to fill loan requests conventionally is an important con-
sideration in any comparison with electronic resource-sharing systems.
Three components of the over-all time can be identified: the time for the
borrower to deliver a request to the lending library; the time required by
the lending library to process the request, and the time to deliver the
material to the requesting library. The time to get a request into LC
averaged five days for requests within Continental United States and from
6.5 to 11 days from libraries outside the U.S. The Trevvett study provides
LC data for the second component. At LC loan-processing time varied with
the kind of material involved. Processing times from the general book
collections were 5.9 days, for other collections the time ranged from
8.9 days to 27.4 days. Requests filled from serials needed an average of
9.3 days for processing. To all these times must be added the transit

---

* The bracketed numbers refer to the bibliography at the end of the report.

time required for the processed material to reach the source of the request as well as the time required for packaging, addressing and so forth.

Interlibrary borrowing and lending data at the Massachusetts Institute of Technology libraries are documented in reports by Reekes [3] and Ferriero [4]. The following relevant items of information are drawn from these reports and from discussions with M.I.T. library personnel. In FY 1978 the M.I.T. libraries photocopied 11,670 serials items in lieu of lending; they represented seventy-eight percent of all items loaned. The average number of pages per article was 6.3. In addition, approximately 1800 photocopied articles were received from outside. M.I.T.'s turnaround times for photocopying, defined as the elapsed time between the date a user makes a request and delivery of it to the requestor's library are typically these: 86.3 percent of photocopy requests are filled within 34 days; 38 percent of all requests (books and photocopies) are filled within two weeks.

Cost figures for interlibrary borrowing and lending at M.I.T. have been computed on the basis of 1977 labor costs. The 1977 data show that lending costs were $5.30 per item and borrowing costs, $7.34 per item. If a 10 percent annual cost increment is assumed, the corresponding 1981 figures become $7.75 and $10.75 respectively. These figures are typical of those commonly cited for current borrowing and lending costs. If 1981 costs are now applied to the FY 1978 photocopy volume, we find that M.I.T.'s annual cost is close to $111,000.

The preceding data provide a basis for comparing electronic ordering and delivery systems with the photocopying method. Comparisons should be made on the basis of turnaround (response-to-request) times as well as cost factors and, of course, the quality of the copy provided to the end user. Expressed another way, the question is: Can an electronic document-ordering-and-delivery system be configured that will supply acceptable quality copy to end users more quickly and economically than is now possible with conventional procedures? This report is a response to this question.

THE RESEARCHABLE ISSUES

We found that the crucial issues relevant to electronic document-ordering-and-delivery systems centered on an interrelated set of require-

ments for document scanning and printing at each node of the network, and upon the kind of communication used to link the nodes. Demand for high-quality copy necessitates use of a large number of resolvable picture elements (pels) per page. If a short per-page transmission time is demanded, high-data-rate communication links are required to handle the data involved. Because of this constraint communication costs are tied to resolution requirements. Conversely, to constrain communication costs for high data-rate links, separation distances between nodes must be constrained since the tariffs for leased digital-data lines are essentially proportional to line length.* Similarly, one has a choice of a laser or a solid-state charge-coupled device (CCD) in the document scanner and a laser printer or an electrostatic printer for the output devices. The laser approach entails a cost which is nearly an order-of-magnitude more than the CCD approach, and to take advantage of the inherently higher resolution of laser scanners and printers, much more data must be handled by the communication links. Hence, communication-link issues are further aggravated. These issues are developed more fully in the next several sections.

SYSTEM CONFIGURATION

In our investigations we took the approach that ordering, processing and delivering journal articles should be treated as an integrated process and technology should be employed at any point of the process where its use shows promise of improving performance. A preliminary study based on this principle led us to the concept of an on-line-computer document-ordering system into which all transactions among nodes are entered, and to a separate electronic document-delivery system completely independent of the ordering system. Separating the ordering and the delivery phases allows use of low data-rate connections to the ordering systems and avoids complications that could result if all text were passed through the central computer.

The system we examined is illustrated in block diagram form in Fig. I-1 for a network of three libraries. The document-ordering system resides in

---

* Microwave links and packet-switched networks are discussed in the section on Communications as alternatives to digital-data wire service.
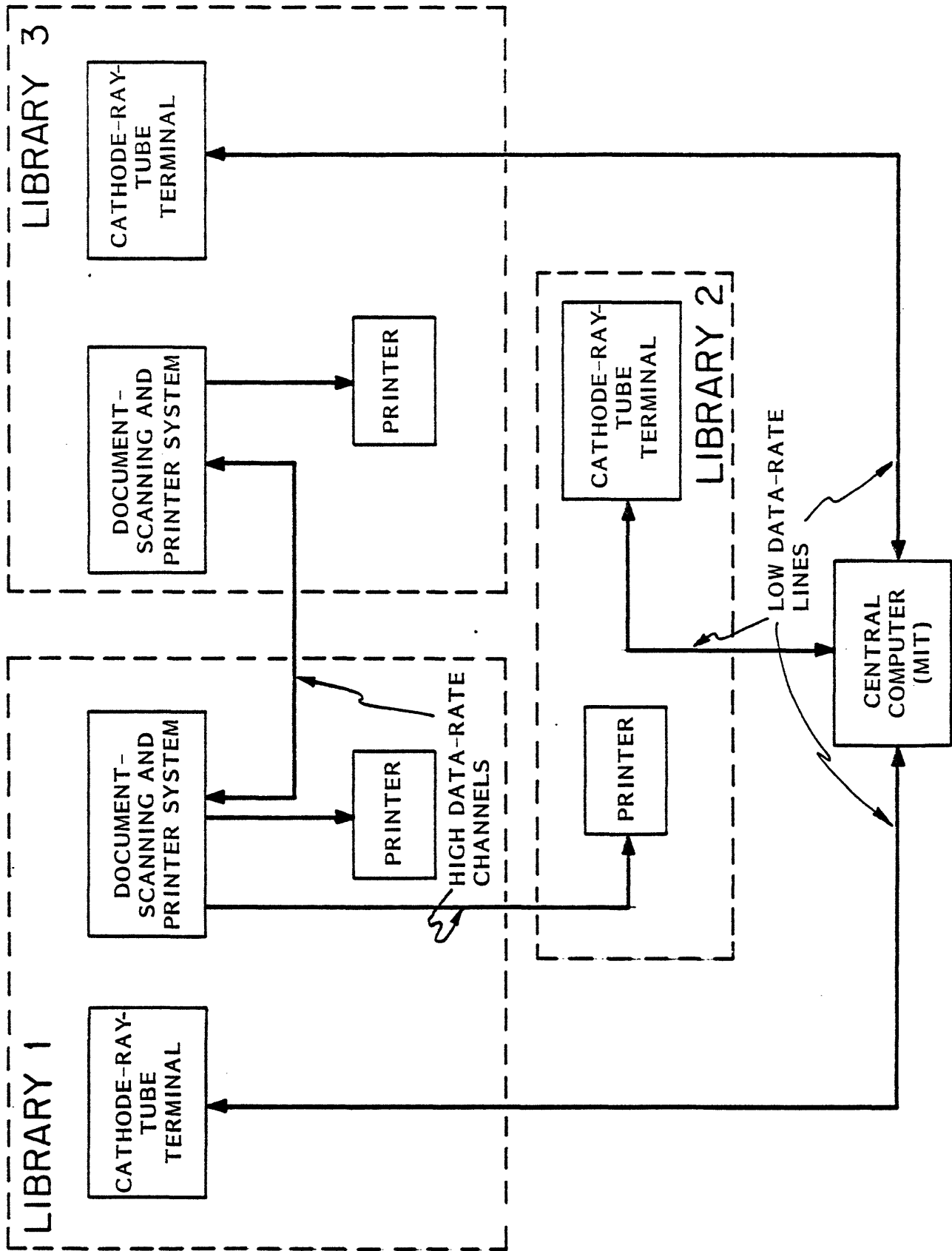
Fig. I-1. Block diagram for an electronic resource-sharing network.

an interactive time-sharing computer drawn at the bottom of the figure and is accessible from cathode-ray-tube terminals over voice-grade phone lines. Each library has a bound-document scanner and a printer which is connected to each of the other nodes by means of high data-rate channels.

Further insights into network operation can be given with the aid of Fig. I-2. The process starts at the upper-left part of the diagram. It is assumed, for library-control purposes, that all requests for serials flow through the system. In order to obtain a document, the user fills out a request form.* The form is given to a library clerk who verifies its correctness. If incorrect, it is returned for correction; otherwise, the information is used to determine whether or not the user's own library owns the document. This step is envisioned as a computer operation, since a complete index of all serials held by network members will be stored in the computer. Should the document be owned locally, the form is given to the user who seeks it in the stacks. The process ends if the document is found (see diagram); if not, the user returns to the clerk and resubmits the request form. Alternatively, it may turn out that the local library does not own the document. In either negative case, the clerk ascertains from computer-stored information which, if any, libraries in the network hold the document in question; if none does, the process ends. If one or more do, the clerk enters a request into the computer. The amount and kinds of data that must be posted remains to be decided; in general, the data will depend upon the degree of automaticity that follows in succeeding steps. Two limits are envisioned: At one extreme, the clerk transmits a request message via computer to the library most likely, in the clerk's judgment, to have the item being sought, this step being repeated by the clerk for the next-most-likely source and so on, in case of failures, until a supplier has been found. At the other extreme and in light of more elaborate computer programs, further clerical action beyond the initial entry of a request would be unnecessary other than to check on the status of a request. The computer would automatically transmit the request serially to various libraries until the request has been honored.

---

* An alternate approach is to have the user insert the request himself/herself at a computer terminal. This approach might be introduced into the system later, after experience has been gained with network procedures.

REQUESTOR'S LOCATION   SUPPLIER'S LOCATION

USER OPERATIONS

REQUESTOR OPERATIONS

USER

FILL OUT REQUEST FORM → FORM TO CLERK

FORM CORRECTED ← NO — FORM OK?

YES

FORM TO USER ← YES — USER'S LIBRARY OWNS DOC?

NO

DOC. AVAILABLE? — NO →

YES

END

AVAILABLE AT CO-OPERATING LIBRARIES? — NO — END

YES

SUPPLIER OPERATIONS

LOW-SPEED DATA LINKS

POST REQUEST DATA IN COMPUTER ⇄ CENTRAL COMPUTER → SUPPLIER'S CLERK RECEIVES REQUEST FROM COMPUTER

MAIL "REGRETS" TO USER ← NO — AVAILABILITY STATUS?

END

YES

DELIVER FORM TO XMISSION OPERATOR

POST AVAILABILITY MESSAGE IN COMPUTER ← NO — DOC AVAILABLE?

YES

DELIVER DOC AND REQUEST FORM TO XMISSION OPERATOR

RECEIVE DOC. ← — XMIT DOC.

HIGH-SPEED DATA LINK

MAIL DOC. TO USER — INFORM CLERK
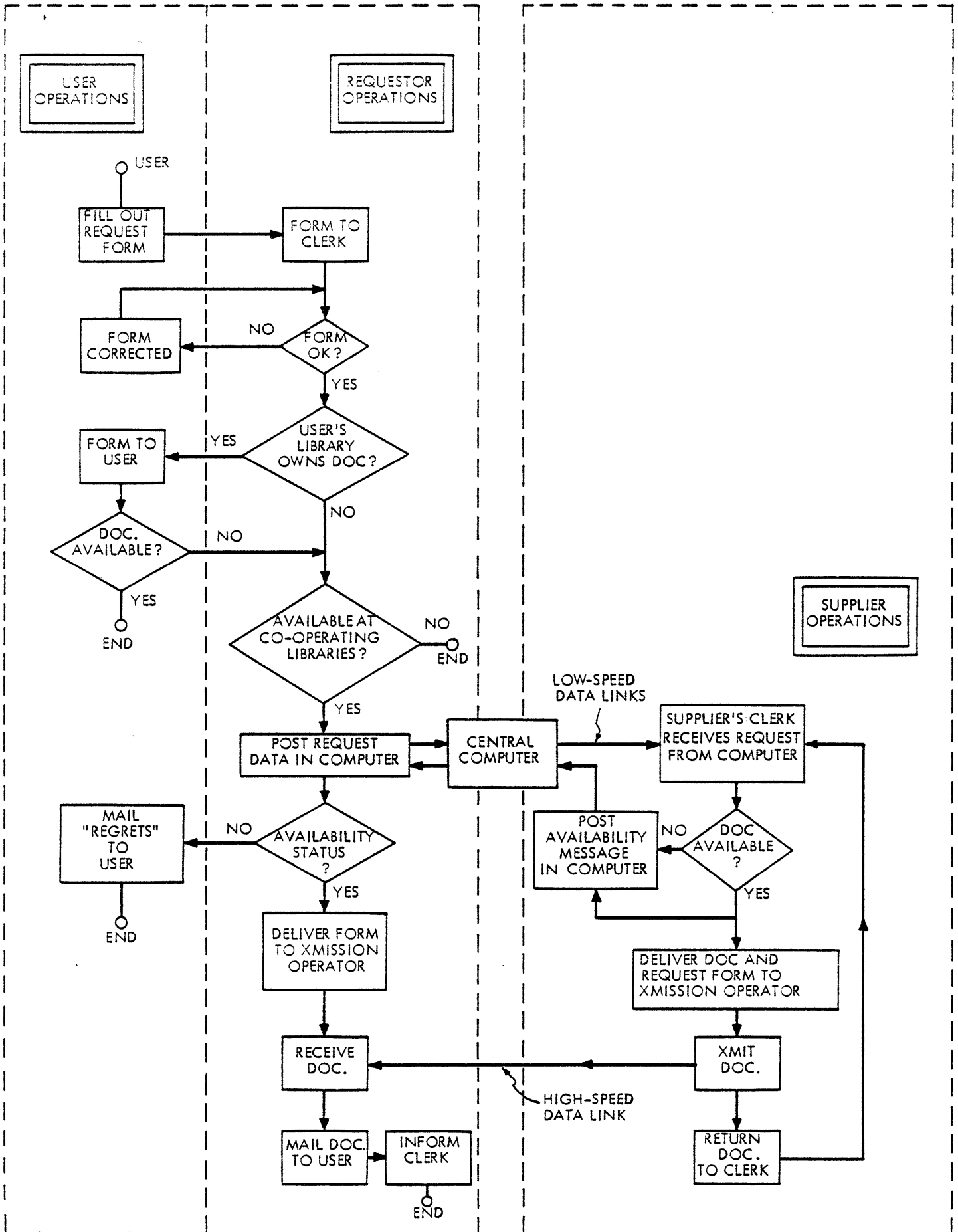
END

RETURN DOC. TO CLERK

Fig. I-2.   Information flow between two network nodes.

Suppose, now, that no supplier in the network can be found. Obviously, the process ends at this point and the conventional interlibrary-loan process takes over. On the other hand, if a supplier is located within the network, the request form is delivered to the local transmission operator by the requesting clerk for identification purposes, at which point the local clerk's task ends (except perhaps to enter billing information into the computer on receipt of notice that the document has been sent to the user).

Action is now transferred to the supplier's location (see lower right side of Fig. I-2). The supplier's clerk receives the document request from the resource-sharing computer and determines whether or not the item is available. (We know the supplier owns the document; otherwise, that particular supplier would not have received the request.) If the answer is NO, that message is inserted into the computer and the next most likely source is tried. If the item is available, it is retrieved and given to the transmission operator together with the request form. The transmission operator, in turn, electronically scans and transmits the document, after which it is returned to the clerk for replacement in the stacks.

Action now reverts to the requestor's location. The transmission operator there receives the transmitted document, informs the user by phone of its availability for pickup, or dispatches it by internal mail and returns the original request form to the clerk, who posts accounting data into the computer. Thus, the transaction ends.

DOCUMENT-TRANSMISSION SYSTEM

In our investigations of document-transmission systems we chose to confine our studies to two-level digital systems. We did this with knowledge that reducing the process to a binary, black-or-white decision process prohibits reproduction of gray-level photographs faithfully and, hence, would impose this constraint on any operational systems based on this kind of design. On the other hand, in the literature of many professional disciplines photographs play only a supplementary role to alphanumeric text and line drawings (which cause no problem in a two-level system.) Furthermore, operating in the binary rather than the analog domain allows considerable

flexibility in processing signals -- data compression and signal enhancement, for example.

Figure I-3 shows a block diagram of equipment at one station of a document delivery system. It is configured to scan opaque pages or microforms -- roll film or microfiche. The output signals of the scanner are converted to binary form in the digitizer after which the digitized signals are compressed and impressed upon a digital-data transmission link. As indicated in the diagram, digital signals arriving over the transmission line from another node are impressed on a decoder and applied to a printer capable of accepting the decoded signals. The dashed line in the diagram indicates that compressed data can be observed locally before transmission by passing a sample through the local decoder and printer. The digital controller is programmed to carry out the sequence of operations in proper sequence.

Our experimental research has centered on all of the blocks of Fig. I-3 except the printer and microform scanner. For outputting copy we chose to employ a Gould 5200 electrostatic plotter which we had on hand and to build the other elements around its characteristics. This device employs a staggered linear array of stylii to impose electrostatic charge on roll paper. The charge attracts a liquid toner to the paper. Hence, presence or absence of voltage on each stylus determines the presence of absence of a dot.

The Gould 5200 accepts serial digital data and prints the image a line at a time in raster format. Experiments with the plotter revealed that the paper-advance mechanism, driven by stepping motors, limited its output rate to approximately three milliseconds per line or about 5.1 sec per page. Since the plotter paper is eleven inches wide, we chose to line scan a page along its longer vertical dimensions and to output the scanned lines along the 11-inch dimension of the plotter paper. A programmable cutter is used to cut the roll paper after each page.

Although the plotter electrodes which transfer the image to the paper are spaced 0.125 mm apart, microscopic observation of the images shows that alternate picture elements are staggered (because the stylii are staggered)
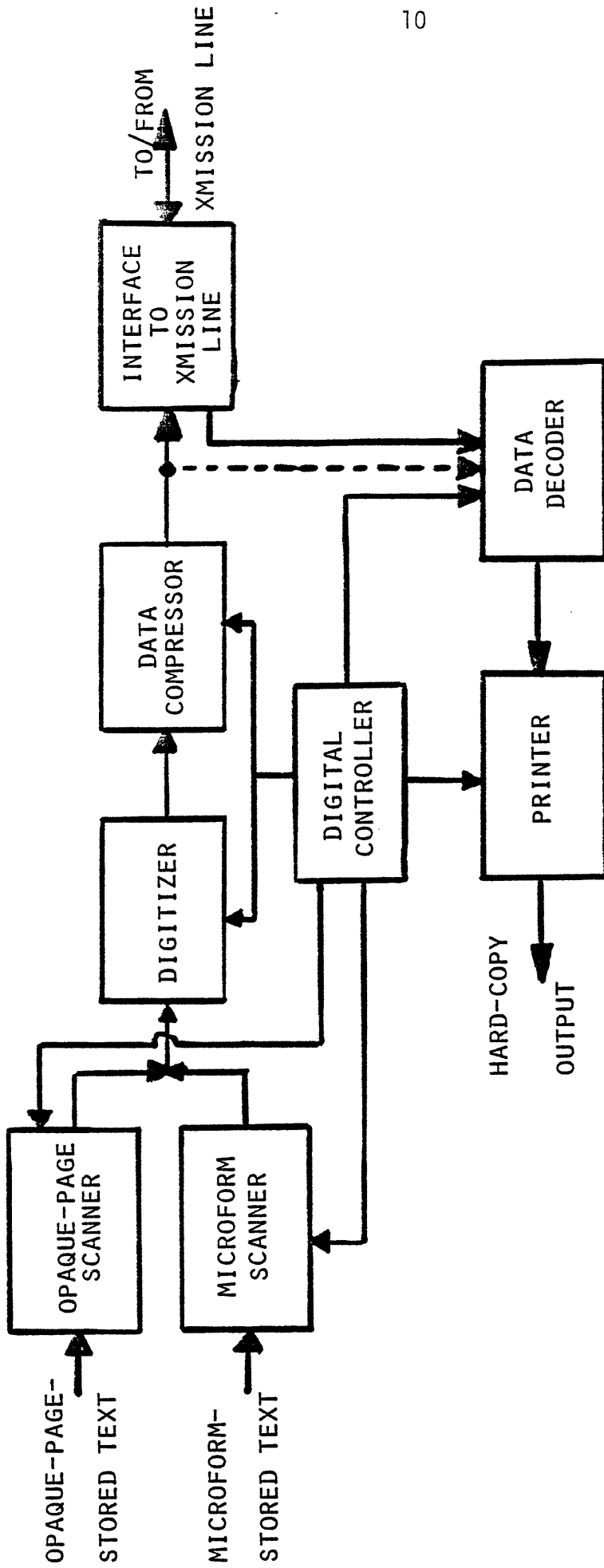
Fig. I-3. Block diagram of equipment at one station of a document delivery system

and there is considerable overlap between the dots. The output dot pattern
is illustrated in Fig. I-4. In this illustration, dots 1_1, 1_2, ... 1_5 repre-
sent picture elements for one scan line; dots 2_1, 2_2, ... 2_5 represent picture
elements from the next scan line. The overlap between dots tends to fatten
character lines from the scan data and to reduce the effective resolution of
the plotter below the 0.125 mm one would expect from the electrode spacing.
The irregular edges that result from the staggering of alternate dots is less
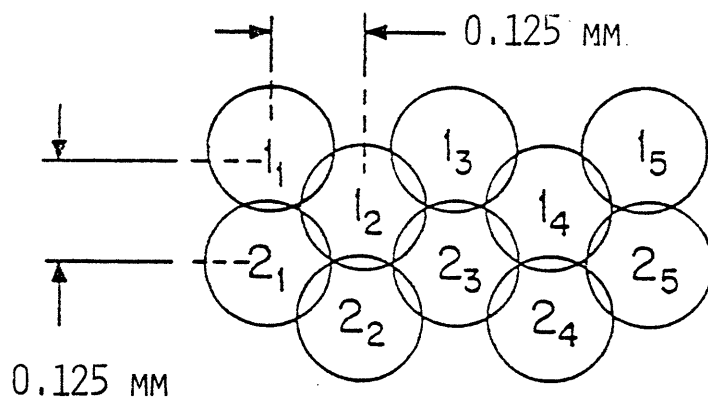pleasing than sharp, crisp edges, but since 0.125 mm is close to the maximum

Fig. I-4. Dot Geometry of Gould 5200 Plotter

resolution of the human eye at normal reading distance, edge distortion is
barely detectable without magnification. The image contrast produced by the
plotter is ordinarily high, although humidity effects on the paper reduce the
contrast considerably after long periods of idleness. This should not be a
problem in an operational system where the plotter is operated daily.

Since we adapted the characteristics of our experimental document-
transmission system to the Gould 5200 plotter, several parameters are fixed
as a result of this decision. Minimum time to scan a page is set at about
5.1 sec, and the total number of pels in an 8-in. by 10-in. format that
must be scanned in that time span to take full advantage of the Gould's
resolution is 3.5 million. Hence, the maximum real-time data-transmission
rate is 725 kilobits per sec. To make printing and scanner resolutions
equal, the scanner should have a linear resolution of at least 4 line pairs
per mm or 100 line-pairs per in.

CHAPTER II

SCANNERS

A logical first question to ask on this subject is: Why scan documents? Why not take advantage of the character redundancy that exists in text by digitally encoding character-by-character, thereby greatly reducing the amount of data acquired in the process? The answer is that when alphanumerics are interspersed with graphics we are led to raster scanning, rather than to individual character encoding, for the reason that alphanumeric codes cannot handle the graphics. As a practical matter, we are also led to raster scanning whenever printed text must be reproduced retrospectively in exactly the same type font and style as the original. In principle one could use a character-recognition technique in this case and match each character against characters in a multiplicity of type faces and sizes, but practically an excessively large digital store of characters would be required.

RESOLUTION REQUIREMENTS

Scanner-resolution requirements for a document-delivery system are tied to the quality of reproduced copy that will be demanded by end users. End users will undoubtedly accept electronically reproduced copy provided it can be easily read. Table II-1 shows the resolution quality, measured in line pairs per millimeter, of several imaging materials and devices. A line pair is defined as a black line followed by a white line of equal thickness. Several points are worth noting in Table II-1: The resolution of microfilm (and photographic film, in general) exceeds by far the resolution capabilities of the other media in the list; the resolution capability of the human eye at normal reading distances about matches that of photocopy machines; and television images have poor quality relative to the others. The present-day accepted quality for many forms of duplicated materials seems to be that provided by photocopy machines -- about 6 line pairs/mm. With this resolution one can easily read type

down to 6-point size* and it is thought by many that such resolution
is a fundamental requirement, since a small percentage of text is often
set in that type size. In the printing of scientific and technical works,

Table II-1 Resolution quality of several imaging materials
and devices

Standard microfilm	630 line pairs/mm
Photocopy machines	6 line pairs/mm
Human eye	1 minute of arc, or
	8 line pairs/mm @ 400 mm
U. S. television	262 line pairs/frame
	0.86 line pair/mm for
	a 12-inch screen

for example, superscript and subscripts are normally set in type that is
smaller than that used for the main text. Hence, five or six line-pairs
per millimeter represents the goal one should strive for in electronic
document scanner. As explained below, we fell slightly short of this
goal with our experimental solid-state scanner, but our results show
that six-point type is legible. The over-all system resolution we achieved
was about 3.6 line pairs per millimeter.

*The point is the unit of type size. One definition is that it is the
height of the slug on which the type is contained. One point is very near-
ly 1/72 inch. Another definition is that point size is the separation
distance between the top of the highest ascender and bottom of the lowest
descender in the character set. An ascender is the part of a lower case
character that extends above the x character and a descender is the part
of a character that extends below the x character.

OTHER REQUIREMENTS FOR BOUND-DOCUMENT SCANNERS

We were unable to locate on the open market a bound-document scanner that conformed to our specifications; hence, we were forced to experiment with our own configurations. Critical issues center on the following matters: the document cradle configuration; the illumination source and corresponding sensing element; and the method used to accomplish scanning in two dimensions.

Bound documents should be supported in the scanner so as to avoid excessive strain on the binding while ensuring that the entire page being scanned lies in a single plane. Fulfilling this requirement is challenging especially for pages near the beginning and end of a thick volume. In thick documents front and back pages tend to curve inward toward the spine unless the whole document is pressed firmly into place. It is considered essential, also, that both parts of documents be supported during page scanning in order to avoid stress on bindings caused by the weight of the dangling pages. We concluded that the design of a realistic <u>automatic</u> page positioner and page turner that.meets the preceding requirements was outside the scope of our project. Our experiments were therefore predicated on manual execution of these two operations. Hence, a further document-cradle requirement is that a manual operator must be able to load, position and flip pages easily, quickly and without risk of doing damage.

Possible illumination sources and corresponding sensor combinations are: a flying-spot cathode-ray-tube (CRT) scanner and photomultiplier tube (sensor); a laser beam and photomultiplier-tube; and fluorescent tubes or incandescent lamps and a solid-state photodiode-type sensor such as a charge-coupled device (CCD) or charge-injection device.

Our past experience with CRT flying-spot scanners quickly drove us away from this approach. The light source, a cathode-ray tube, is bulky and requires a high-voltage power supply. The CRT phosphor also tends to be noisy. On the other hand, raster-scanning the light spot is easily accomplished by deflecting the electron beam within the CRT in two dimensions through use of signals applied to deflection circuitry. A

laser light source should not be dismissed for this application, but if a photomultiplier tube is used as the sensor, a strong signal level must be maintained to overcome photomultiplier noise. Also, since the laser gives a point source of light, some means must be employed to generate the two-dimensional light-beam motion required for raster scanning.

For our scanner experiments we settled on a one-dimensional CCD as the sensor and two fluorescent lamps as the light source. To scan in the second dimension we chose to keep the document stationary and move the lamps, focusing lens and CCD as an ensemble across the page. Other options for motion in the second dimension are: keep the lamps, lens and CCD stationary and move the document; or, sweep the light reflected from the page across a focusing lens and CCD combination by means of a rotating mirror. Moving the document is rather unattractive because of the mass involved. Use of a rotating mirror appears to be a viable alternative to our approach but our experimentation did not include this method.

Figure II-1a illustrates our configuration for a bound-volume scanner. The document cradle (Fig. II-1b) consists of two laboratory-type scissors jacks which support both sides of an open document. The jacks are mounted on a platform attached to two pair of ball-bearing guides mounted orthogonally to each other. Mechanical motion in three dimensions is thus available. The material to be scanned is pressed upward on a 1/4 in. plate-glass window just above the jacks.

The upper part of the document scanner is the electro-optical system. See Fig. II-1a. Two fluorescent tubes suspended from the upper structure illuminate the scanning surface. Light reflected from the document passes upward through the slit in the shroud that supports the tubes. The shroud also ensures that only reflected illumination reaches the lens and sensor.

Near the top of the scanner is a focusing lens and above it are the CCD sensor and its control board. Lamps, lens, CCD and control circuitry all move as a unit from left to right under action of a constant-speed motor that drives a lead screw. All scanner operations, except document loading, are under control of a microprocessor.
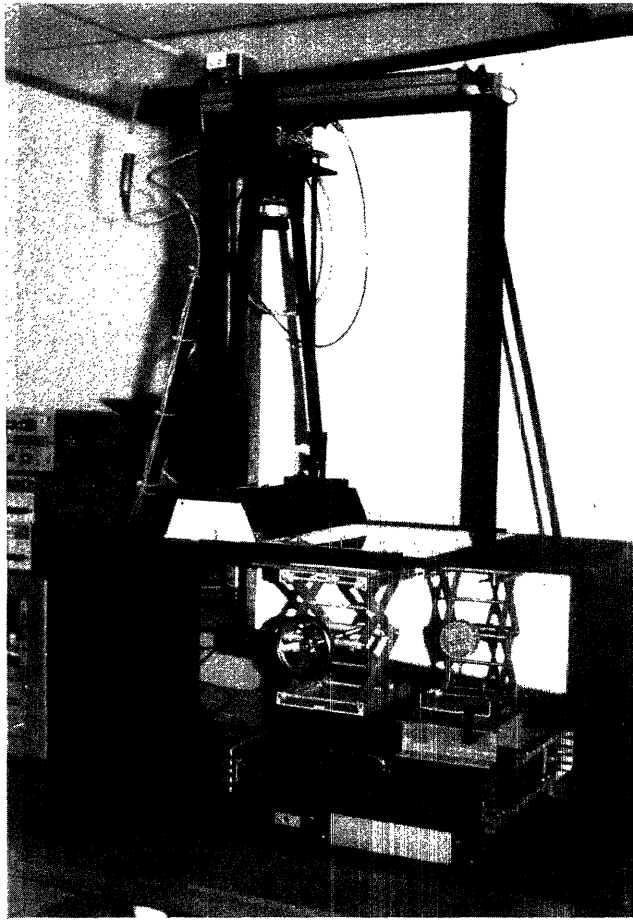
(a)   The Complete Assembly
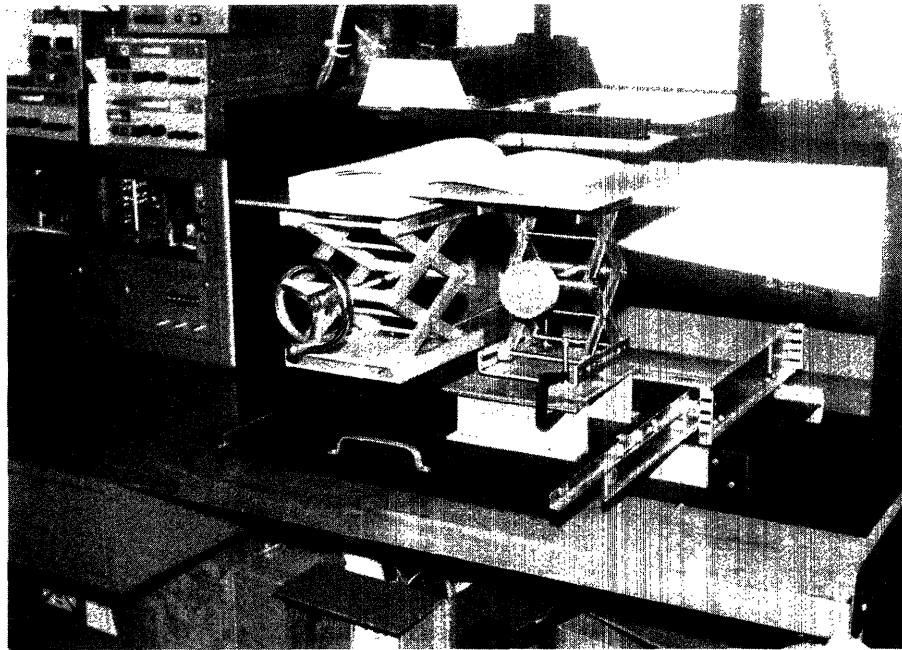
(b)   The Document Cradle

Fig. II-1.   An Experimental Bound-Volume Solid-State Scanner.

To scan, a document is placed into position beneath the glass window by an operator. When the "scan" switch is flipped, the lamps are energized and linear motion of the electro-optical assembly commences. At the end of page traversal, the lamps are extinguished and the electro-optical assembly returns to its start position.

The document cradle. Several forms of document-cradles can be suggested. The open document can be supported with pages facing upward or downward; it can be supported fully open or partially open with pages facing up or down. All these possibilities influence how the rest of the scanner is configured. In terms of minimum wear and tear on documents, the configuration of Fig. II-1 appeared highly promising, although a slight modification, to be described later, might reduce materials-handling times slightly.

To load a document into the cradle in Fig. II-1, the jacks are pulled clear of the viewing window and its frame, and the document is placed on the jacks. The assemblage is then positioned beneath the scanner window and the page to be scanned is pressed firmly against the window by raising the appropriate jack. The other jack supports the rest of the document. After a left-page scan, switch-over to a right-hand page is accomplished by lowering the left side of the document, repositioning the cradle and placing the new page under pressure. To turn a page, the assemblage is pulled sufficiently forward to clear the viewing window while the jacks are being lowered.

Since materials-handling time is an important part of the over-all scanning cycle, we measured the times required to switch from a left-hand to a right-hand page and to turn a page. Two persons (experimental subjects -- ES's) participated in the experiment, and each ES was timed while handling three different documents -- a soft-bound book measuring 8-1/2" x 11" x 1", a hardbound book 6" x 9" x 2" and a soft-cover journal 8-1/2" x 11" x 1/16".

Initially, many page positionings and page turnings were performed by each ES to overcome learning problems. When the time data reached a

reasonably steady state, the average time of several repositionings and page turnings was measured. The results are summarized in Table II-2 for page turning and repositioning, and in Table II-3 for page turning only.

Table II-2 Time to turn and reposition a page for scanning

Item	Time averages, seconds		
	ES - 1	ES - 2	Average
Soft-bound book	10.3	13.9	12.1
Hard-bound book	10.6	12.6	11.6
Unbound journal	8.4	13.5	10.9
Averages	9.8	13.3	11.5

Table II-2 indicates that the time performance of each ES was about the same for the three types of documents, but the two ES's took a substantially different amount of time to do the same jobs. Clearly, the experiment should be continued to include more ES's so that a better statistical sample can be obtained. Time data for page repositioning (Table II-3) are more uniform, but even here a larger sample of ES's might serve to establish average page-turning time more precisely.

Table II-3 Time to reposition a page for scanning

Item	Time averages, seconds		
	ES - 1	ES - 2	Average
Soft-bound book	4.5	5.5	5.0
Hard-bound book	5.9	6.0	6.0
Unbound journal	4.4	5.3	4.8
Averages	4.9	5.6	5.3

Nevertheless, the times presented in the two tables provide order-of-magnitude data and also indicate that, when page turning as well as page positioning is involved, the time required to do the former task is about twice that of the latter task. The over-all average time for a two-page combination consisting of one page of repositioning and one page of turning plus repositioning is approximately 8.4 seconds. If one includes the per-page scan time of 5.1 seconds, the average per-page scan time for a two-page sequence will be of the order of 13.5 seconds per page. Scanning-time profiles are discussed in detail in Chapter IV.

An alternate document cradle makes use of one scissors jack to raise and lower a platform to which is affixed two separate book-holding platforms that are coupled together and movable up or down about a central pivot. The movable book-holding platforms may be thought of as being similar to a weighing scale. The jack serves to press either a left-hand or right-hand page into position for scanning. Although this configuration was not implemented, it is possible the materials-handling time will be shorter and more convenient because only one jack has to be manipulated.

The electro - optical system. Included in this part of the scanner are the optical sensor, focusing lens and illumination source. A 2048-element linear charge-coupled device (CCD), manufactured by the Fairchild Camera and Instrument Company (Fairchild CCD142) was used as a sensor. Ideally, one would prefer a high-resolution two-dimensional array of at least 2,000 elements by 2,000 elements so that an electronic "snapshot" camera could be devised. However, since two-dimensional CCD arrays with 4-million sensing elements are beyond the state of current technology, one must resort to a linear array and a mechanical sweeping action in the second dimension.

The Fairchild CCD142 with its linear array of 2048 elements can provide 8 scanning elements per millimeter over a 10-inch span of text. This resolution is nicely matched to the linear array of 2,000 stylii in the electrostatic printer used in our experiments. Butting together two or more linear arrays to double sensor resolution is possible, but a printer with corresponding resolution is required to realize the CCD resolution im-

provement. We were unable to procure an electrostatic printer with the additional resolution. An electronic control board associated with the CCD provides the timing signals required to read out the generated video signals serially on a line-by-line basis.

The quantities that determine the choice of a lens are: the distances p and q from object and image to the lens center, the focal length of the lens, its F-number and diameter D (See Fig. II-2). Principles that underlie procedures for finding these quantities can be found in text books on optics (see for example, ref. 6, 7). Approximate relationships based on first-order thin-lens theory are summarized and used here as a basis for specifying approximate values for the commercially available lens system used in the scanner.

In the scanner, the object is the opaque page, assumed to be 10.25 in. (260.4 mm) in the direction of the linear CCD array. The image is the CCD array, the length of which is calculated as follows: the diameter of one CCD sensor element is 13 micrometers; hence, the length of the 2,048-element array is:

$$\text{CCD-array length} = 13 \ \mu\text{m/element} \times \frac{2048 \ \text{elements}}{1000 \ \text{mm}/ \ \mu\text{m}} = 26.6 \ \text{mm}$$

The lens demagnification $\underline{m}$ is set by the ratio of q to p:

$$m = q/p = 26.6/260.4 = 0.102. \tag{II-1}$$

Now, the resolution of a diffraction-limited lens is given by the relationship:

$$\text{F-Number} = \frac{d_{min}}{2.44\lambda} \tag{II-2}$$

where λ is the light wavelength, assumed to be 0.5 μm, and d_{min} is the diameter of the smallest spot that can be resolved by the lens. Minimum spot size corresponds to the diameter of the Airy disk (for an explanation of Airy disk see [6], p. 128). Since we are trying to resolve picture elements on CCD photodetectors with diameters of 13 μm each, $d_{min} = 13 \ \mu$m and

$$\text{F-Number} \leqq \frac{13 \ \mu\text{m}}{2.44 \times 0.5 \ \mu\text{m}} = 10.7$$
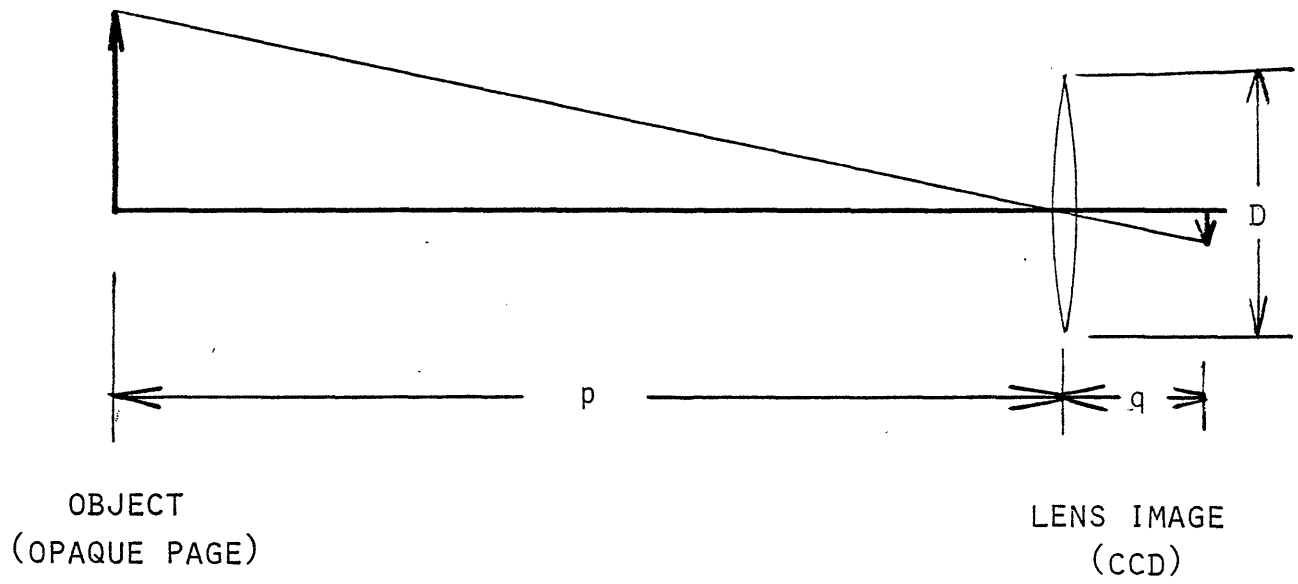
OBJECT
(OPAQUE PAGE)

LENS IMAGE
(CCD)

Fig. II-2. Thin-lens geometry

We have the further approximate relationship for thin lenses:

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q}$$

where f is the focal length of the lens.

Since $m = \frac{q}{p}$, we can write

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{mp} = \frac{m+1}{mp}$$

$$f = \frac{mp}{m+1} \tag{II-3}$$

We now pick a value for f that will yield a convenient physical separation q between lens and CCD. A value of f = 2 in. (50.8 mm) yields:

$$P = \frac{50.8 \times (0.102+1)}{0.102} = 548.8 \text{ mm} = 21.6 \text{ in.}$$

Hence, distance q = mp = 0.102 x 21.6 = 2.2 in., a reasonable value for convenient working distance.

The minimum lens diameter D is obtained from the relation D = f/F-number, where F-number is given by Eq. II-2.

Hence,
$$D \gtreqqless 50.8/10.7 = 4.75 \text{ mm}$$

The lens diameter and maximum lens speed (F-number) have thus been determined on the basis of a diffraction limitation on the lens and without consideration of light-level requirements. It turns out that if the calculated extremes of D and F-number are employed, an impractical amount of light would be required at the page surface in order to activate the CCD through the lens. Hence, in practice a smaller F-number -- of the order of 2.8 to 5.6 -- and a correspondingly larger diameter lens -- of at least 10 - 15 mm, or so -- are needed.

For the experimental scanner, a commercially available Nikkor lens system was employed in place of a simple thin lens. It has a focal length of 54.26 mm (2.14 in), a diameter of 55 mm and F-numbers that are adjustable

from 1.2 to 11. The lens speed is determined partly by the amount of illumination required at the CCD location and partly by the depth of field needed to keep the entire page in focus. Since the two requirements are in conflict with respect to F-number, a compromise setting must be struck. In practice a lens speed between f-4 and f-5.6 has been found satisfactory.

Light Source. Two GTE 14-watt 15-in. fluorescent-light tubes were used as a source of illumination. The tubes are housed in a metallic shroud, the inner surface of which is white enamel. See Fig. II-1a. Light reflected from the page being scanned emerges from the shroud through a rectangular slit approximately 1/2" x 13". During a scan, the lighting assembly, lens and CCD sensor move together across the page. The two tubes provide 800 lumens, an adequate number to actuate the CCD.

Because the analog signals that emerge from the CCD are reduced to binary form in a constant-bias threshold-comparator circuit, it is essential that CCD signals corresponding to white have constant amplitude over each scan line of the page. This condition will not prevail if the page illumination is uniform, the reason being that the distance from the page to the lens increases as one moves away from the page center toward the edges. Since the light intensity is inversely proportional to the square of the separation distances, constant illumination across the page results in contoured illumination across the CCD.

There are several possibilities for correcting this non uniform signal response. One is to illuminate the page with several incandescent lamps mounted in directional reflectors rather than with fluorescent tubes, and to contour the page illumination so that the luminous intensity is greater near the edges than at the page center. A second method is to make the necessary corrections electronically by multiplying the output of each CCD cell by an appropriate correction factor. In our search for a simple approach we experimented with the following two additional techniques. In one, we sought to reduce the light intensity emanating near the page center to the intensity level near the edges by contouring the slit in the

fluorescent-tube shroud through which the reflected light emerges. See Fig. II-3a. This scheme is satisfactory but care must be taken to stiffen the shroud suspension so as to avoid vibrations in the assembly during scanning. Otherwise, reflected light at the narrow parts of the slit may be completely blocked. A better arrangement is to reduce the light intensity near the center of the tube through use of light attenuators attached to the fluorescent tubes. We found that five or six 1-in. bands of white-bond paper wrapped around the tube are highly effective attenuators. See Fig. II-3b. Band positions can be adjusted while observing the output-signal trace of the CCD on an oscilloscope. The trace before and after application of the attenuator bands is shown in Fig. II-4. The bands should be wrapped so that they can be easily slid along the tubes during adjustments. The disadvantage of this method is that replacement fluorescent tubes may require different placements of the attenuating bands, a procedure that is most easily carried out by making CCD output-wave observations with the aid of an oscilloscope.

The fluorescent tubes should be driven by a direct-voltage rather than an alternating-voltage source to avoid flicker during scanning. Normal a-c operating parameters of each tube employed in the experimental setup are 39 volts and 0.38 ampere. In the d-c drive we provided a capability to control light intensity by varying the current through the tubes above and below 0.38 ampere. Circuitry was also included to permit a one-second warmup period at reduced current prior to the start of each scan and a means for switching the direction of current flow on alternate scans. The latter arrangement prevents the excessive buildup of mercury at one end of the tube that would occur if current flow were always in the same direction.

ELECTRIC DRIVE FOR THE ELECTRO-OPTICAL SYSTEM

For the experimental scanner a drive mechanism was needed to impart linear mechanical movement of the CCD, optics and lighting assembly. Since we had experimented with a stepping motor in the past for this purpose, we chose this time to use a constant-speed direct-motor drive in order to make comparisons. To obtain the required linear mechanical motion, a d-c motor-driven lead screw was employed.
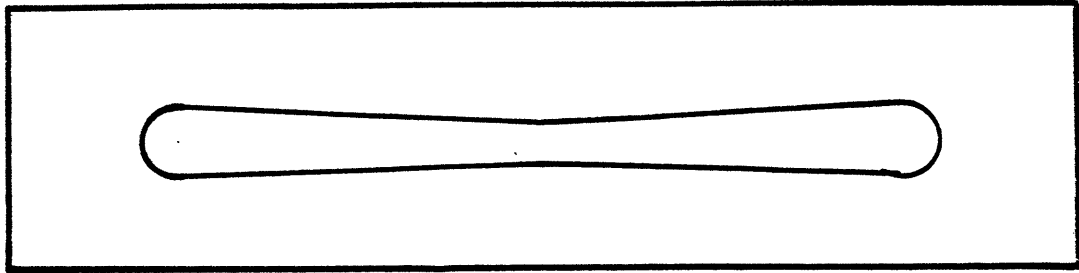
(A) TOP VIEW OF SHROUD WITH CONTOURED SLIT

TWO TURNS OF
WHITE BOND
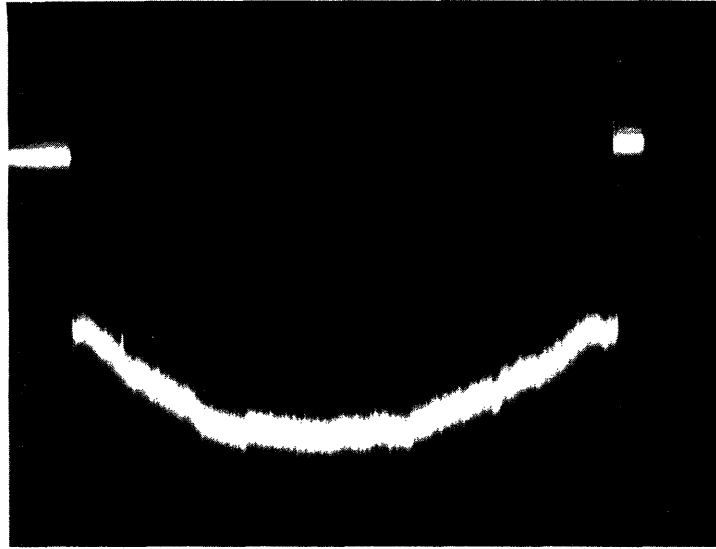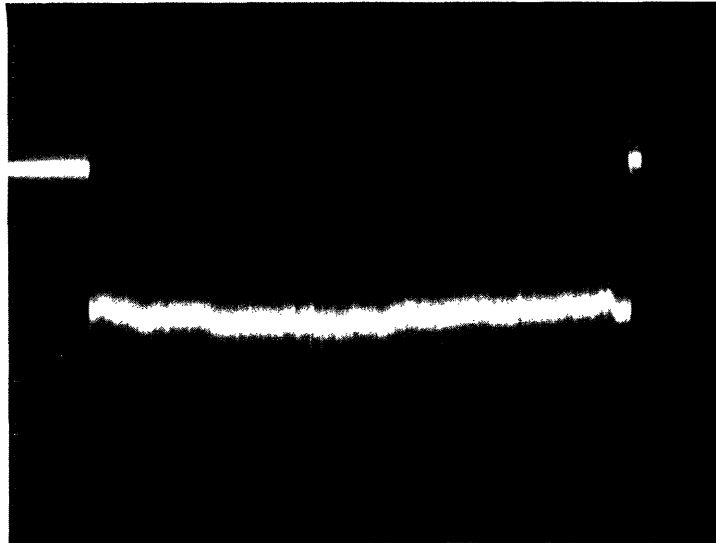PAPER
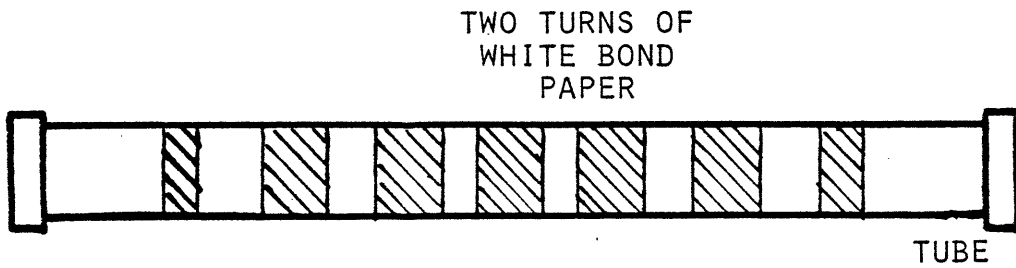
TUBE

(B) FLUORESCENT TUBE WITH PROPER LIGHT ATTENUATORS

Fig. II-3. Two methods for reducing the intensity of
reflected light emerging near the page center.

(a)   Without Light Attenuators

(b)   With Light Attenuators

Fig. II-4.   CCD Output-Voltage Waveforms before and
after Use of Light-Attenuator Bands

In the design of the motor and lead-screw arrangement several inter-related quantities must be taken into account, namely, motor speed, lead-screw linear speed and maximum lead-screw error. A good starting point for determining these quantities is the page-scanning time T_p. In our experimental scanner the lower time limit $T_{p\ min}$ is set by the maximum speed at which the electrostatic printer can operate -- 1.75 in. per sec -- or by the maximum allowable linear speed of the lead-screw nut that avoids vibration, whichever has the lower value. Since the lead screw we chose carries a specification of approximately 2 in. of linear travel per sec maximum, the printer determines $T_{p\ min}$, namely 4.9 sec for an 8-1/2 in. page. It should be noted that the slower the per-page scan time the longer the CCD light integration time, and consequently the less stringent is the light-intensity demand on the illumination source.

The upper limit on T_p is arbitrary. Long scan times per page result in long times to transmit a page, but the longer the scan time the lower is the bandwidth requirement for the transmission link. Our decision was to take a conservative position and to calculate the lead-screw speed on the basis of a per-page scan time of 5 seconds. The corresponding linear speed of the lead-screw nut then becomes: the page width of 8.5 in. divided by 5, or 1.7 in. per sec, below its rated value and about equal to the maximum printer speed.

In order to specify lead-screw requirements completely we must first set a value for mechanical-scan accuracy. If we demand that we hold the number of lines per 8.5-in. page width to 1700 $\pm$ 0.5, the maximum tolerable error is (1/1700) x 100, or 0.06 percent. We were able to locate lead screws manufactured by Velmex, Inc., having true pitch-deviation errors ranging from 0.013 percent to 0.033 percent, well within our tolerance requirement.

There remains the task of choosing a motor-drive speed that is within lead-screw specifications and which is compatible with our 5-sec page-scan time. If $\underline{d}$ is the nut linear advancement per revolution of the lead screw and $\underline{w}$ is the total linear travel, $w/d = R$ is the total number of revolutions per scan, and R/T_p is the turning speed of a directly

coupled motor drive. When $\underline{w}$ and $\underline{d}$ are in identical units and T_p is in seconds, the motor revolutions per minute, RPM, becomes:

$$RPM = \frac{R}{T_p} \times 60 = \frac{w}{dT_p} \times 60$$

In our case, $w = 8.5$ in., $d = 0.2$ in. per revolution (manufacturer's specification) and $T_p = 5$ sec. Hence, the motor RPM is:

$$\frac{8.5}{(0.2)(5)} \times 60 = 510 \text{ RPM}$$

To ensure a smooth, constant speed at the above rate, the motor and lead-screw assembly were made part of a phase-locked loop that feeds back a signal from the lead-screw motion as output to the motor input.

Drive comparisons. We compare here the performance of the phase-locked-loop, d-c motor drive used in the experimental bound-document scanner with our experience with a stepping-motor drive employed in an earlier project involving a solid-state microfiche scanner. Underlying principles are similar for both scanners, except that microfiche scanners operate from light transmitted through the fiche, whereas opaque-page scanners utilize light reflected from the page. The optical systems must accommodate different magnification requirements.

Our observation is that the stepping-motor-driven lead screw operated quite satisfactorily in the microfiche-scanning environment and the associated electronic circuitry was simpler than that required by the phase-locked-loop motor drive. Because of the inherent smoothness of the d-c motor drive, the load, consisting of the electro-optical and sensor components, exhibited no observable vibration effects. On the other hand, one might have expected vibratory effects in the stepping-motor setup because of the discrete steps in which the motor rotates, but this turned out not to be the case. Image distortion attributable to stepping-motor-induced vibrations was not discernible. One should keep in mind, however, that the total linear traversal of the stepping-motor load was only an inch or so, whereas in the d-c-motor setup there were 8-1/2 inches of linear displacement. We suspect, without proof, that stepping-motor

performance may be inferior where long travel distances are encountered because lead-screw vibrations may increase with lead-screw length. Also in very high resolution scanning (above the 200 lines per inch, for example), the inherent smoothness of the phase-locked-loop drive may be necessary in order to achieve the required resolution. This point deserves further investigation.

INTERPRETATION OF RESULTS

The configuration employed for the experimental bound-document scanner performed satisfactorily. We were able to reproduce readable copy down to 6-point type, although not with letter-press quality. The configuration was quite compatible with our requirement that the scanner must accommodate bound volumes without strain on bindings; that is, the idea of moving the electro-optical/sensor system as a unit over the page to be scanned, led to a simple, over-all structure that performed reliably without resort to high-precision fabrication. Our decision to implement the particular configuration was primarily judgmental and based on the premise that a minimum-size package was not a constraint. The experimental scanner also met our criterion that it be a low-cost device.

Nevertheless, we recognize more research on scanner configurations may well lead to improved units. Since so little work has been put into document-support mechanisms, further investigations of this aspect of bound-document scanners is especially desirable. However, such investigations should not lose sight of the fact that "librarian-friendly" document-support devices are mandatory. In our dealings with the library community we have observed great dissatisfaction with present photocopying machines in respect to their "unfriendliness" toward bookbindings. One can understand why this is so -- the photocopy machines are designed to handle principally single sheets of paper.

A feature of our experimental scanner that is noteworthy is that is operates satisfactorily so long as there is contrast between the print and its background. Hence, both color and black/white materials were scannable. However, re-adjustment of the threshhold voltage of the digitizer is needed,

not only for different colors, but also for various degrees of glossiness of white paper.* We are currently experimenting with an automatic method for adjusting voltage threshhold that would be universal.

---

*Threshhold voltage is the voltage against which scanner output voltages are compared for the purpose of assigning a zero or one to the output voltage. Output voltages that exceed the threshhold are designated a one; voltages less than threshhold value are called a zero.

CHAPTER III

COMMUNICATIONS

INTRODUCTION

Several types of communications media are available as connecting links between nodes of an electronic interlibrary resource-sharing network. They are: coaxial cable, microwave transceivers and public-utility wire-communication service. Wire-communication links operate either as packet-switched networks, conventionally switched networks or as point-to-point lines. Satellite links are also available but they will likely prove un-economical over the near term except in very special situations.

In packet-switched networks a digitized message is broken into an ensemble of "packets". Each packet is composed of a fixed number of message bits plus a destination code and message identifier. The packet is then routed under computer control to its destination, where all packets of the same message are collected and reconstructed into the original message. Tariffs for packet-switched network include several components, but a distinguishing characteristic is that the variable-rate component pertains only to the number of packets transmitted and is independent of the separation distances between nodes. The data rate is presently 56 kilobits per sec, but there is no fundamental constraint that would prevent higher data rates in the future.

The Bell System's communications network can be used for digital data transmission at several discrete data rates. Lines can be leased for point-to-point service, or the switched network can be used, subject to certain constraints on data rates. The variable component of Bell's tariff structure imposes charges that are proportional to the separation distances between terminals.

Final choice of a communication medium and its data rate is governed by over-all system operating requirements and tradeoffs that are possible among these requirements. It may be desirable, for example, to hold the per-page transmission time to a few tens of seconds; otherwise, the number of documents

that can be put through the system per-unit time may be too small. However, the number of bits of information that can be transmitted over certain types of communication channels, particularly public-utility channels, is determined by the discrete classes of service the utility provides. Furthermore, the higher the data rate, the higher is the tariff for the service. It is important to bear in mind that in order to obtain a good reproduction of a full page of text that includes line drawings and type sizes down to 6 points, of the order of 4 million bits of information are required. In order to transmit this quantity of data in a few tens of seconds, a high data-rate channel will be needed.

Two approaches can be taken to increase document throughput between terminals of a communication channel: a higher data-rate can be employed, or the lower data rate can be retained and the data that is used to describe the content of each page can be electronically compressed before transmission. There are tradeoffs involved in each approach. High data-rate channels may, under certain circumstances, entail intolerably high tariffs. On the other hand, data compression comes at the expense of extra electronic equipment at both the sending and receiving ends. These opposing factors must be reconciled. Other tradeoffs are also possible. If one is willing, for example, to accept both slow transmission time and the extra equipment that goes with data compression and decompression, it is possible simply to employ ordinary voice-grade telephone lines -- either on a leased-line or a dial-up basis. Throughput is greatly reduced, of course, in this approach.

These issues are discussed in detail in this chapter with respect to several modes of transmission.

COAXIAL CABLE

Coaxial cables are ideally suited for networks that can be constructed over private rights of way. In many situations the propagation characteristic of these lines is adequate for handling required data rates without need for data compression and it should be possible to accommodate separation distances between nodes of three-quarters of a mile or so without

need for an intermediate amplifier repeater. Pulse-reshaping circuitry
at the receiving ends may be required, however. One would certainly want
to give strong consideration to coaxial-cable transmission between nodes
of an intra-organizational network of on-site branch libraries of a uni-
versity or of an industrial or government complex. At current prices, it
should be possible to pull RG-58/U coaxial cable through existing under-
ground ducts or string the cable in corridors for less than three dollars
a foot, including the cost of the cable. No data compression would be
necessary in this transmission mode.

MICROWAVE TRANSCEIVERS

Microwave communication is a viable data-transmission mode for inter-
library resource-sharing systems in environments where public rights of
way must be crossed and where network nodes are within line-of-sight dis-
tances. Microwave transceivers, like coaxial cable, are inherently broad-
band devices and hence would not require data compression to achieve
typically required data rates. Use of a microwave carrier system does,
of course, involve extra equipment and it is likely that each transmit-
receive antenna will have to be located selectively at each node--probably
on a roof-top--in order to ensure a clear line-of-sight path between
antennas.

Unlike coaxial cable, which is a very simple passive transmission
medium and hence should require essentially no maintenance, microwave
communication systems are more complex devices which will undoubtedly re-
quire adjustment and preventive maintenance from time to time. One micro-
wave transceiver for the interlibrary resource-sharing application can be
expected to cost approximately seventy-five thousand dollars.

PUBLIC-UTILITY WIRE COMMUNICATION

Bell-system services. Several kinds of digital service are currently
available in the Bell system. They are: voice-grade telephone lines,
offered on either a leased-line or dial-up basis; 56 kilobit-per-second
service, available mostly as point-to-point leased-line service, but more

recently available as switched service between a few principal cities; and 1.544 megabit-per-second service.

Voice-grade lines can handle digital transmissions up to at least 9,600 bits per sec through use of appropriate modulators-demodulators (modems). The T-1, 1.544 megabit-per-sec digital data service is intended primarily for long-distance trunk service between telephone exchanges in major cities, and appears too expensive for this application unless there are extenuating circumstances.

Clearly, then, the rates at which the information content of documents can be put through these channels vary widely, as do their tariffs. Hence, choice of service must be based on a tradeoff between throughput and service cost. The latter factor, in turn, is dependent upon separation distance between network nodes, since leased-line rates are essentially proportional to nodal spacing. More precisely, Bell-System charges for point-to-point interstate leased-line digital service can be divided into three components regardless of the data rate. They are:

C_{FI} = one time installation charge

n = number of months over which C_{FI} is allocated

C_F = fixed monthly charge

C_{Ft} = total fixed monthly charge

C_V = variable distance charge, per month

If C_{FI} is apportioned over a period of $\underline{n}$ months, then, over this period, $C_{FI/n}$ and C_F can be combined into a total fixed-charge component C_{Ft}, and the total monthly charge C_T as a function of distance becomes $C_T = C_{Ft} + C_V$. The general shape of this cost function is shown in Fig. III-1. Specific values depend on the service data rate. Note, however, that break points occur in the cost curve at certain distances. At each point the per-mile rate of the variable-rate component C_V diminishes slightly.
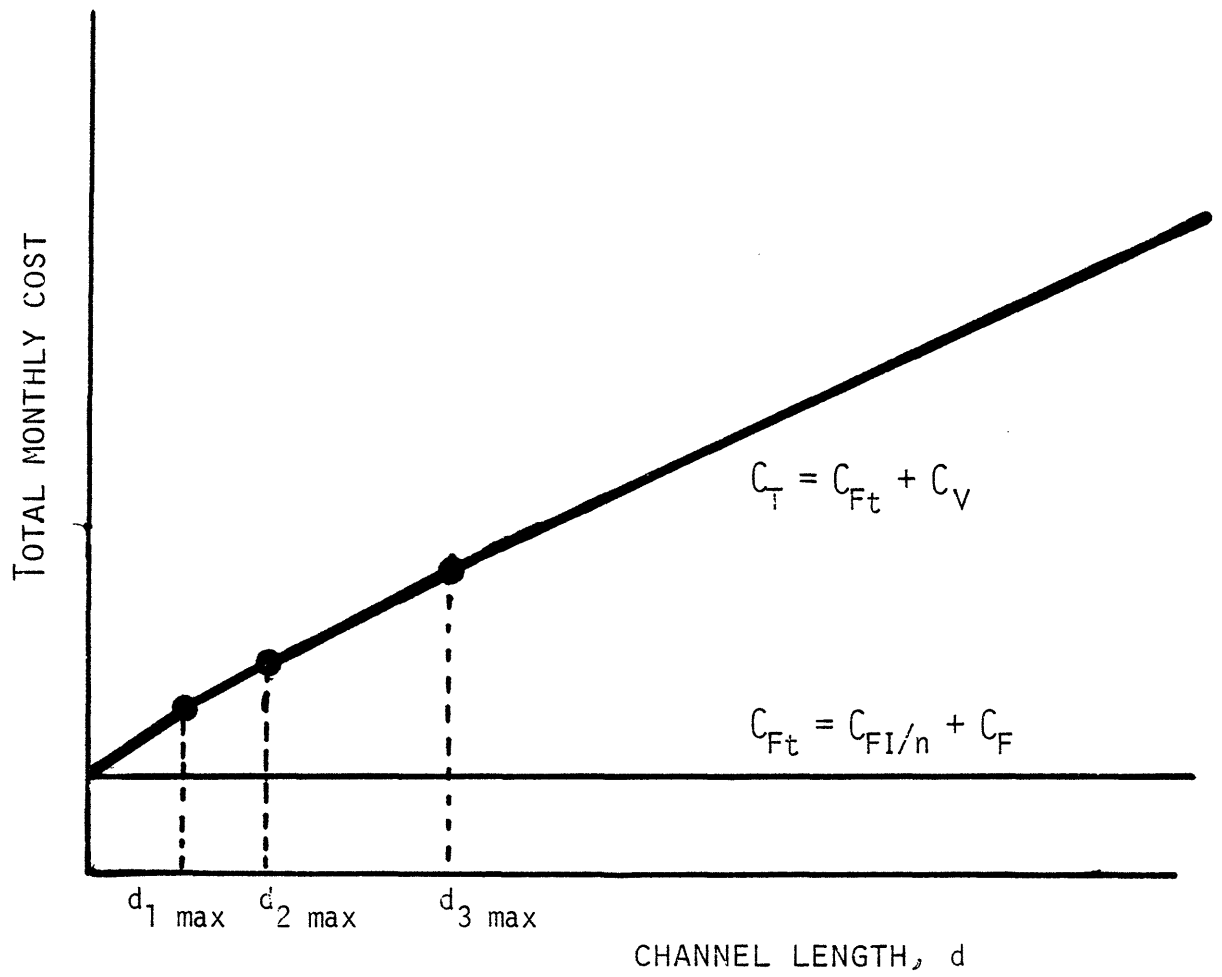
Fig. III-1. The general shape of the monthly lease function for phone-company digital data service. See text for definition of symbols.

General relationships between the variable lease-cost factor C_V and channel length can be developed as follows:

Let d_1 = channel length for $0 < d_1 \leq d_{1\ max}$ miles

r_1 = per-month monthly rate, $0 \leq d_1 \leq d_{1\ max}$ miles

d_2 = total channel length for $d_{1\ max} < d_2 \leq d_{2\ max}$ miles

r_2 = per-mile monthly rate, $d_{1\ max} < d_2 < d_{2\ max}$

d_3 = total channel length $d_3 > d_{2\ max}$ miles

r_3 = per-mile monthly rate, $d_3 > d_{2\ max}$ miles

Then,

$$C_V = r_1 \, d_1 \quad \text{when } d_1 < d_{1\ max} \qquad \text{III-1}$$

$$C_V = r_1 \, d_{1\ max} + r_2 \, (d_2 - d_{1\ max}) \quad \text{when } d_2 < d_{2\ max} \qquad \text{III-2}$$

$$C_V = r_1 \, d_{1\ max} + r_2 \, (d_{2\ max} - d_{1\ max}) + r_3 \, (d_3 - d_{2\ max}), \quad d_3 > d_{2\ max}$$

$$\text{III-3}$$

Because tariffs are set by federal or state regulatory bodies, depending upon whether the communication channel is interstate or intrastate, monthly lease costs depend on the geographic locations of the network nodes. The following illustration of a monthly cost is typical, however, and can be used as a guide. We have chosen 56-kilobit-per-sec interstate service between two nodes as an example and have assumed that the one-time installation charge will be spread over 48 months; that is $\underline{n} = 48$. Consider the following rate components (all charges are on a per-month basis):

C_F = \$1235

C_{FI} = \$443/48 = \$9.20

C_{Ft} = $C_F + C_{FI}$ = \$1244

$d_{1\ max}$ = 15 miles; $d_{2\ max}$ = 25 miles; $d_{3\ max}$ = 100 miles

r_1 = \$9.45 per mi; r_2 = \$7.90 per mi; r_3 = \$5.90 per mi.

Using Eq. III-1, 2 and 3, we obtain

$$C_T = \$1244 + \$9.45 \times d_1, \quad 0 < d_1 < 15 \tag{III-4}$$

$$C_T = \$1366 + \$7.90 \times (d_2 - 15), \quad 15 < d_2 < 25 \tag{III-5}$$

$$C_T = \$1465 + \$5.90 \times (d_3 - 25) \quad 25 < d_3 < 100 \tag{III-6}$$

These values are plotted in Fig. III-2.

Transmission time per document and maximum monthly throughput of two classes of digital-data service can be calculated. Let us assume a 30-day month and that 36 megabits of information are generated per average-length document. This number corresponds to 4.5 megabits per page and an average length of 8 pages per document.* On a 24-hour operational-day basis, therefore, we obtain the monthly throughputs for 9.6 kilobit-per-sec and 56-kilobit-per-sec lines shown in Table III-1, COL (2). Transmission time per document is given in COL (3).

Table III-1. Channel capacity and document transmission time for two different data-rate channels.

COL (1) Data rate, kilobits per sec	COL (2) Throughput capacity, documents per month	COL (3) Transmission time per doc, minutes
56	4,032	10.7
9.6	691	62.6

One can use Eq. III-4, -5, and -6 and the data in Table III-1 to calculate the monthly transmission <u>cost</u> per document over a 56-kilobit-per-sec channel, assuming the channel is busy 100 per cent of the time. The cost curve is shown in Fig. III-3 for distances up to 50 miles.

Using published tariffs, one can construct similar cost figures for other data rates such as 1.544 megabits per sec and 9.6 kilobits per sec, the latter service being obtainable either through line-leasing or dialup. The excessive cost of leasing 1.544 megabit-per-sec lines would probably

---

* 4.5 megabits per page would yield a higher resolution than the 3.5 megabits per page generated by the experimental scanner described in Chapter II. As components become available networks would likely opt for the higher resolution.
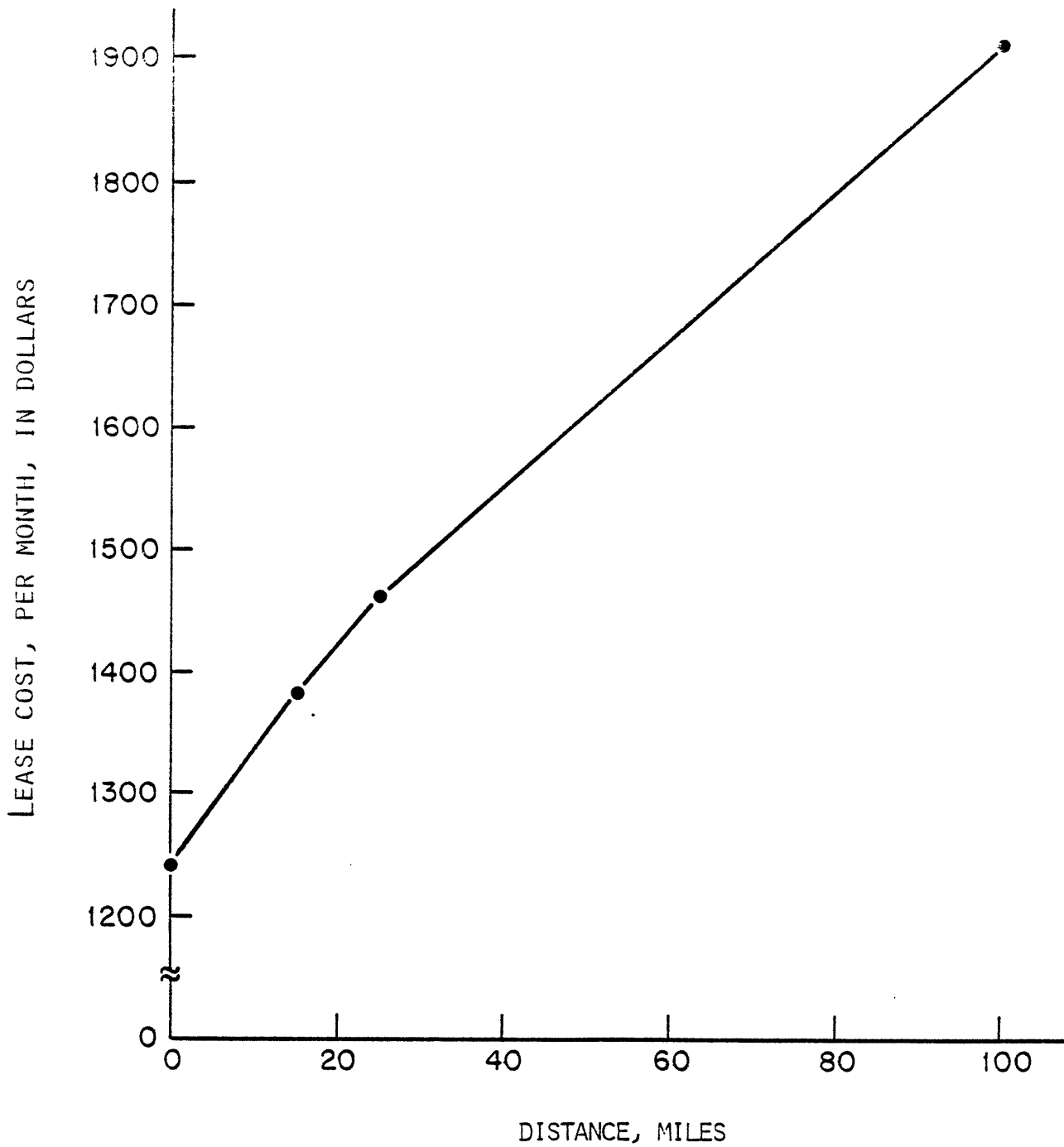
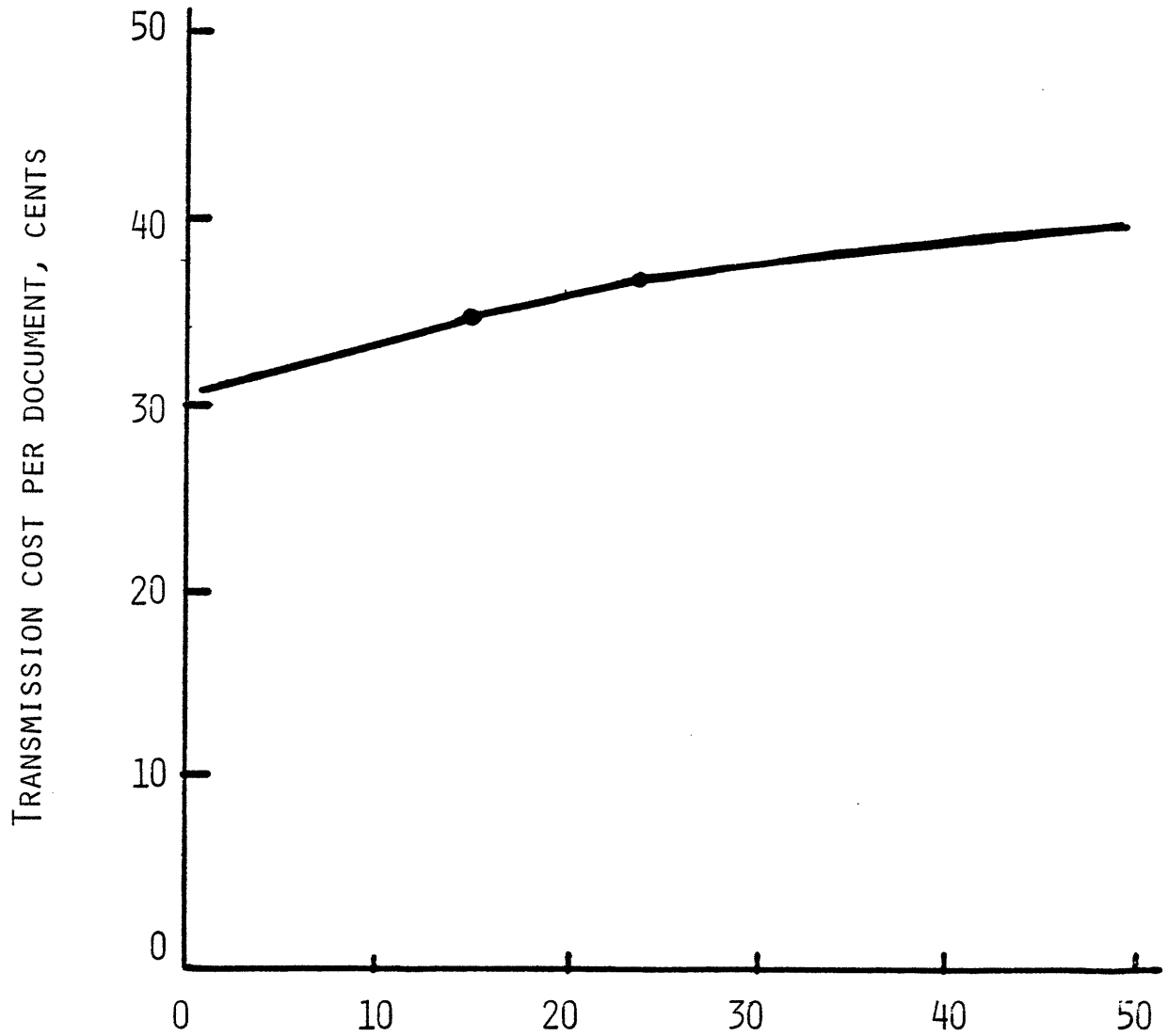Fig. III-2. Representative monthly lease cost of a 56 kilobit-per-sec digital-data-communication link.

Fig. III-3. Transmission cost per document, assuming a 56-kilobit-per-sec data channel is used at the upper limit of its capacity.

rule out this service, whereas the low throughput capacity of 9.6 kilobits per sec lines is likely to make them unattractive except in low traffic situations where overnight document delivery on a dialup basis may be satisfactory.

Still another factor that influences the choice of transmission data rate requires discussion. It is transmission speed in relation to operational speeds of other equipment in the document-delivery system. Consider the 10.7 minutes required to transmit an 8-page document over a 56 kilobit-per-sec line. These numbers equate to 1.34 minutes per page, which is far longer than the 5 seconds required by our experiment scanner/printer system to scan and print a page. One would like, at least, to match transmission time more nearly to scan time plus page-turning time. It will be shown in the next chapter that such a match is possible provided data-compression techniques are employed prior to transmission.

PACKET-SWITCHED NETWORKS

This type of service may hold some future promises for expanding the geographic extent of electronic interlibrary resourse-sharing networks since service charge is independent of distance. However, it appears that current per-packet charges that are in force will result in excessively high communication costs for the document-delivery application.

The most recent tariff schedule of one major operator of a packet-switched data-communications network includes several components that can be grouped into fixed charges and variable charges [8]. Fixed charges, in turn, have several components; they include a network-access charge, on-site network-interface-equipment charge and a monthly account charge. Where equipment is involved, thre is a one-time installation fee and a monthly rental fee. The principal variable charge is a traffic charge based on the number of packets transmitted through the network regardless of distances.* The current tariff is $1.00 per thousand packets, each packet containing 1,024 bits of user data. Hence the fee for transmitting a typical (8-1/2" x 11") page of uncompressed scanned data at a resolution of 200 pixels per inch (3.5 megabits per page) is $3.50 regardless of dis-

---

*There is also a ten-cent fee for each call connection through the network.

tance. To this amount must be added prorated fixed charges. Clearly, if this type of packet-switched network is employed in an electronic document-delivery system, it behooves one to include a substantial amount of data compression in the design. A 10-to-1 compression ratio, which is within the realm of possibility, would reduce the variable charge to 35 cents per page -- an attractive figure.

In any analysis of digital-data-link configurations, it is important to bear in mind that a uniform data-rate capability must be maintained throughout the channel regardless of the kind of service employed. If a high data rate is chosen (56 kilobits per sec, for example) nowhere in the communication channel can a low-data-rate segment of line be tolerated. In other words, the same data rate must prevail from scanner to printer.

ORGANIZATION OF HIGH-SPEED NETWORKS

In absence of a high-speed switched communication network, every node of electronic document-delivery network must be directly connected to every other node. A network of this kind, called a mesh network, is shown in Fig. III-4 for four nodes. If, on the other hand, a high-speed switching station is available, each node can reach every other node through the switch, thus offering potential for a reduced number of lines. In Fig. III-4, for example, a switched configuration (called a hub-and-spoke network) is shown for four nodes. Note that two fewer two-way lines are required than in the mesh arrangement. The number of two-way lines required in a mesh configuration is n(n-1)/2, where n equals the number of nodes. In a hub-and-spoke arrangement, however, n lines are needed. Table III-2 indicates the number of two-way data links required for each configuration as a function of n. It is evident that for networks of four nodes or less the number of communication links is a relatively unimportant issue, but above four, the mesh arrangement becomes increasingly unattractive.

Fortunately, prospects are bright for 56-kilobit-per-sec conventionally switched service in the United States. This class of service is
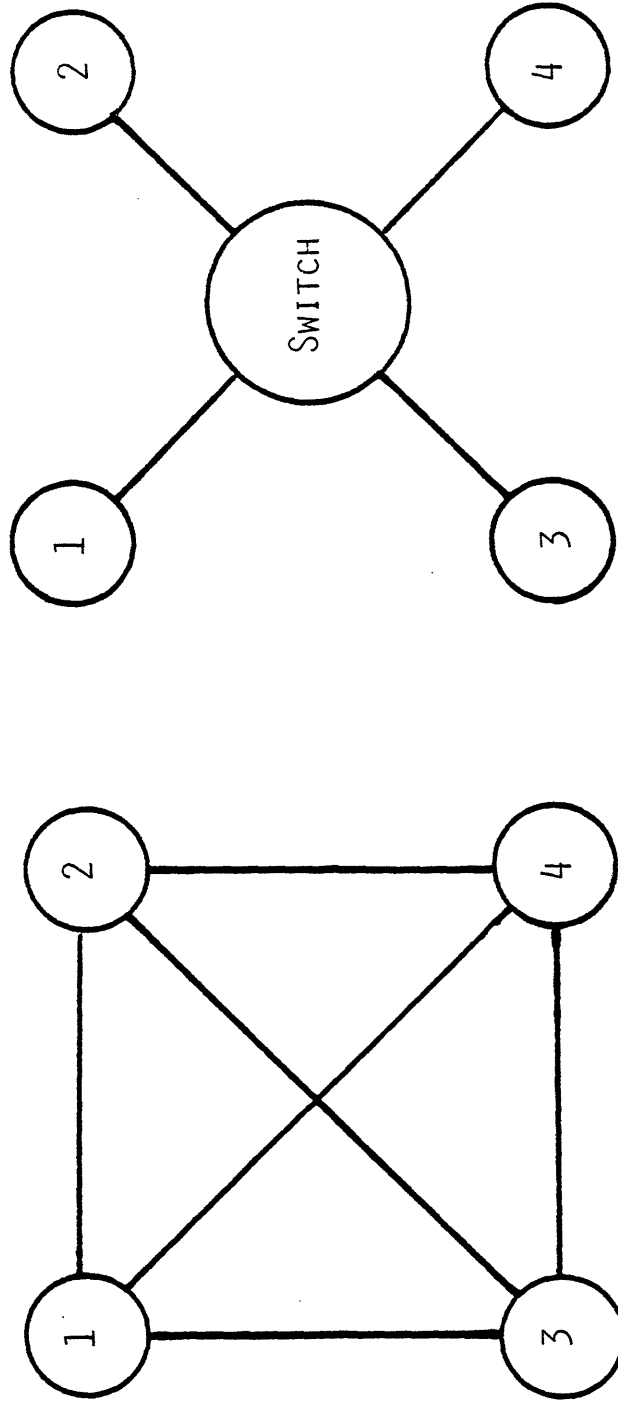
(a) MESH NETWORK

(b) HUB-AND-SPOKE NETWORK

Fig. III-4. Mesh and hub-and-spoke networks with four nodes.

currently available in several large cities, and as market demand increases, one should expect to see the service expanded.

Table III-2. Data-link requirements for mesh and hub-and-spoke networks

No. of Nodes	Required data links	
	Mesh Network	Hub-and-Spoke Network
2	1	2
3	3	3
4	6	4
5	10	5
6	15	6
7	21	7
8	28	8
9	36	9
10	45	10

CONCLUSIONS

Our study of communications for electronic interlibrary resource-sharing system has led us to several conclusions.

1. In intraorganizational situations where communication is entirely over private rights of way, coaxial cable that is installed and operated by the organization itself is highly attractive. A privately owned coaxial-cable system provides a simple, economical communication channel and minimizes over-all system complexity in that data compression is unnecessary in order to achieve high document throughput.

2. Microwave links are suitable for both intra-and inter-organizational situations where terminal points are within line-of-sight distances.

3. Beyond line of sight, conventional, 56 kilobit-per-sec digital data service is, on balance, the most effective class of services for inter-nodal communication. This data rate strikes a reasonable balance between document throughput capacity and transmission cost.

4. Finally, we conclude that, presently, communication-channel cost is the factor which imposes a practical limitation on the geographic extent of interlibrary electronic networks. This conclusion is based on the assumption that communications costs must be fully absorbed by network members; there is no subsidy available nor can document transmission piggy-back on communication links already in place for other purposes.

CHAPTER IV

DATA COMPRESSION

INTRODUCTION

We have investigated conditions under which data compression is desirable in an interlibrary resource-sharing system, and under such conditions the amount of compression that would be required in order to improve system response time. In principle it is advisable to employ data compression to shorten the transmission time of a document through any transmission link whose data rate is less than the data rate that is possible in the rest of the system -- the element of the system having the lowest data rate determines the data rate for the entire system.

Since the data-handling capacity of coaxial cable and the microwave transceivers greatly exceed the speeds with which document scanners and printers can be expected to operate, data compression is unnecessary when these data links are employed. We therefore confined our examination of data-compression requirements to wire links and we concentrated on 56 kilobit-per-sec lines, the most likely type that would be employed in an operational setting.

ANALYSIS OF REQUIREMENTS

The time required to transmit a page of scanned text involves three interrelated components: scanner time, operator time and data processing and transmission time. Relationships among these components are shown, in generalized form, in Fig. IV-1 for a pair of consecutive pages in a bound volume. It is assumed in this figure that scanning of the second page must be delayed until all data gathered from the preceding page have been compressed and transmitted. The reason is that the data-storage device is assumed to be capable of storing only one page of data and the device has read-and-write capabilities, but cannot do both simultaneously. Hence the storage device must be emptied before scanning can be resumed.

Observe in Fig. IV-1 that scanner time is composed of three parts: T_a is the time required for the scanner drive motor to accelerate to full speed; T_s is the time consumed in scanning the document and stopping the drive motor; and T_r is the time the device needs to get back to its original position. Figure IV-1 also indicates that page-positioning occurs immediately upon completion of page scanning and, in general, will involve two different time intervals. It can be expected that the time T_{p2} to turn a page and to set up a new page will be longer than T_{p1}, the time to set up a new page when no page-turning is required.

A further property of the over-all scanning process is implied in Fig. IV-1. For a given scanner design, the time variable that is under control of the designer is the time required to transmit the data, this time being dependent upon the amount of data compression built into the system. Figure IV-1 indicates, with respect to the first page, that the operator completes a new-page setup before data transmission has been completed, whereas second-page transmission has been completed before the operator is able to turn to a new page and position it for scanning. Hence, in this example, the time required to complete all operations on the first page is governed by the data-compression and transmission time, whereas for the second page it is operator setup time that controls the total elapsed time. Note, finally, that as the data-compression ratio is increased and the transmission time is correspondingly shortened, there is a value of data-compression ratio above which nothing is gained -- over-all page-scanning will then be governed solely by operator-setup and scanner times.

Time relationships can be expressed in terms of the symbols indicated in the various boxes drawn in Fig. IV-1. Let TT_1 be the time required to scan and transmit the first page. Then,

$$TT_1 = T_a + T_s + MAX(T_r, T_{p1}, T_{xm})$$

where $MAX(T_r, T_{p1}, T_{xm})$ means: Use T_r, T_{p1} or T_{xm} whichever is largest.
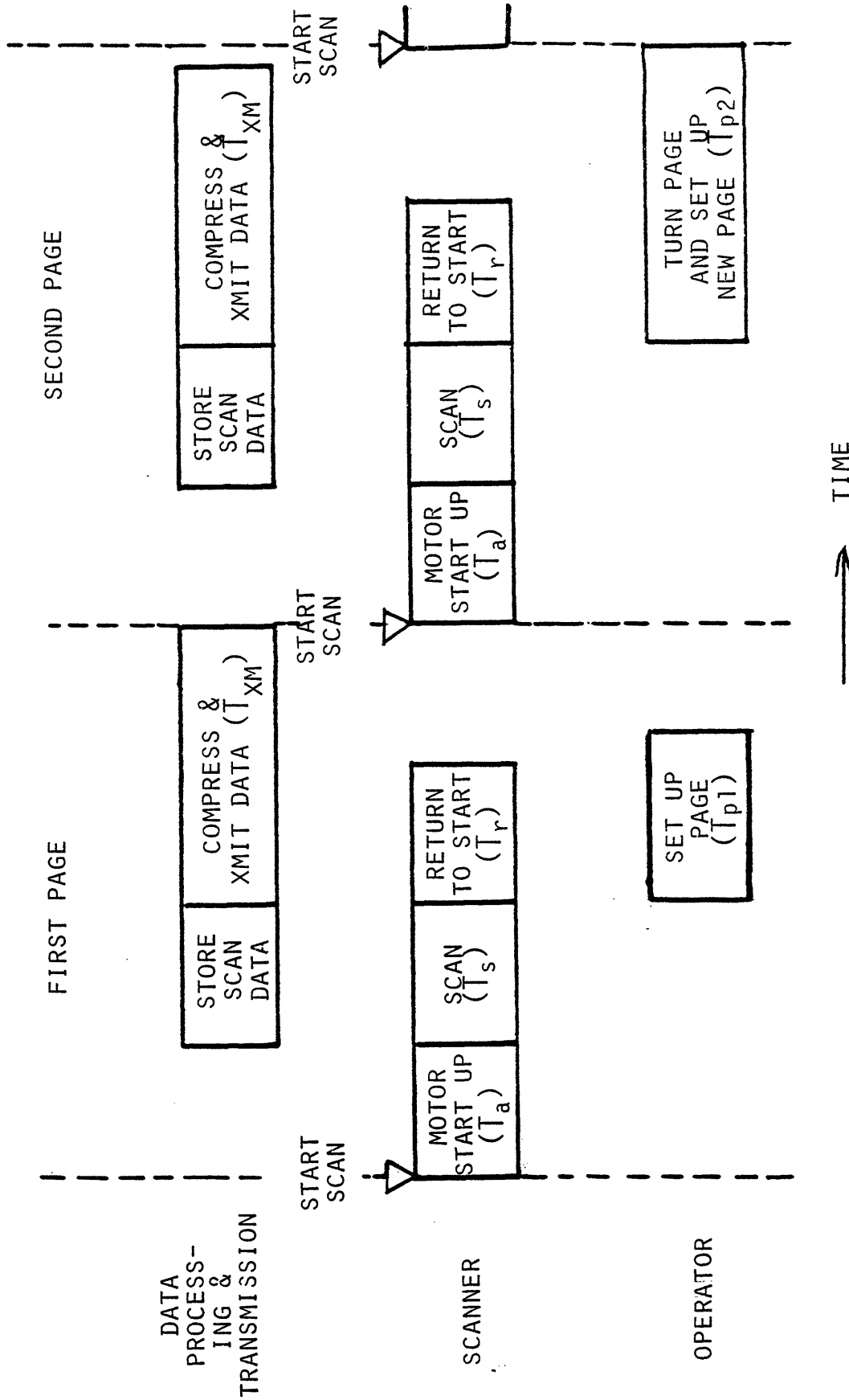
47

Fig. IV-1. Time relationships among operator, scanner, data processing and transmission over a two-scan cycle.

Similarly, the time TT_2 required to scan and transmit the second page is:

$$TT_2 = T_a + T_s + MAX(T_r, T_{p2}, T_{xm})$$

and the time TT_t to transmit the page-pair is:

$$TT_t = 2(T_a + T_s) + MAX(T_r, T_{p1}, T_{xm}) + MAX(T_r, T_{p2}, T_{xm}) \qquad \text{(IV-1)}$$

If the time required to perform data compression is negligible compared to data-transmission time, T_{xm} is given by the expression:

$$T_{xm} = \frac{B}{D_r \times C} \qquad \text{(IV-2)}$$

where B is the number of bits generated per page, D_r is the data rate of the communication link and C is the data-compression ratio.

For the experimental document scanner described in Chapter II, we are now able to show the influence of data-compression ratio C on TT_t, the time required to transmit a pair of consecutive pages over a 56 kilobit-per-sec data channel.

For the experimental scanner described in Chapter II the following data apply:

T_a = 2.5

T_s = 6.0   (9/10 sec. is added here to the 5.1-sec scan time to
          account for scanner stopping time)

T_r = 6.0

T_{p1} = 5.3

T_{p2} = 11.5

B = 3.5 x 10^6 bits

D_r = 56 kilobits per sec

T_{xm} = 3.5 x 10^6/(56 x 10^3 x C)

All times are in seconds. Substitution of these values into Eqs. IV-1 yields:

$$TT_t = 17 + MAX[6.0, 5.3, (3.5 \times 10^6)/(56 \times 10^3 \times C)]$$
$$+ MAX[6.0, 11.5, (3.5 \times 10^6)/(56 \times 10^3 \times C)] \qquad \text{(IV-3)}$$

Evaluation of this relationship for various values of C yields the curve of TT_t versus C in Fig. IV-2 for single-page storage capability.

The first break at C = 5.44 occurs because, at this compression ratio, page transmission time is exactly equal to the time required to turn a page and set up a new page. The second break occurs at C = 10.42 where page transmission time equals the scanner's return-to-start time. Higher C's have no further influence because the process becomes operator/scanner limited rather than data-transmission limited.

An improvement in response time can be achieved with a fixed amount of data compression if two pages of information can be stored and the storage devices are able to act independently of each other. In effect, scanning of a page can then proceed while data from the preceding page are still being transmitted. However, the preceding-page transmission should have finished at the end of the second scan interval so that the compression circuitry and transmission line will be free to accept new data. Time relationships are shown in Fig. IV-3. A general expression for the page-pair transmission time TT_t can be derived from the figure.

$$TT_t = MAX\{T_{xm}, [(T_a + T_s) + MAX(T_{p1}, T_r)]\} + MAX\{T_{xm}, [T_a + T_s + MAX(T_{p2}, T_r)]\}$$

$$(IV-4)$$

For the experimental scanner, this expression becomes:

$$TT_t = MAX\{3.5 \times 10^6/(56 \times 10^3 \times C), [(2.5 + 6.0) + MAX(5.3, 6.0)]\}$$
$$+ MAX\{3.5 \times 10^6/(56 \times 10^3 \times C), [(2.5 + 6.0) + MAX(11.5, 6.0)]\}$$

This equation reduces to:

$$TT_t = MAX(62.5/C, 14.5) + MAX(62.5/C, 20) \text{ seconds} \qquad (IV-5)$$

Page-pair transmission time TT_t is plotted in Fig. IV-2 as a function of C for this two-page-storage configuration.

Several observations can be made about the curves of Fig. IV-2. In a two-page storage system, no improvement in response time is possible for data-compression ratios above C = 4.3, since other parts of the system
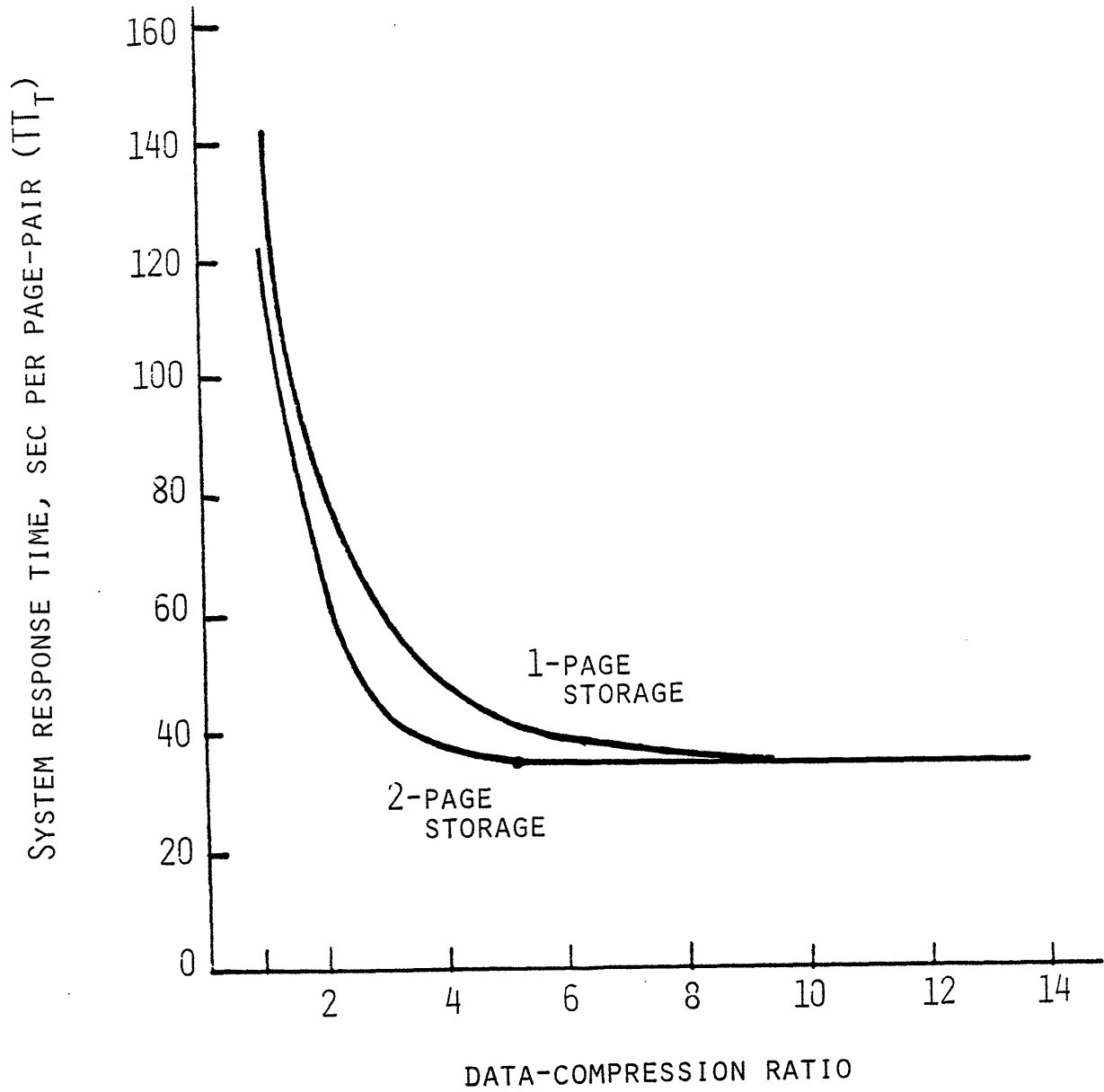
Fig. IV-2. The influence of data-compression ratio on the response time of the experimental scanner/printer system.
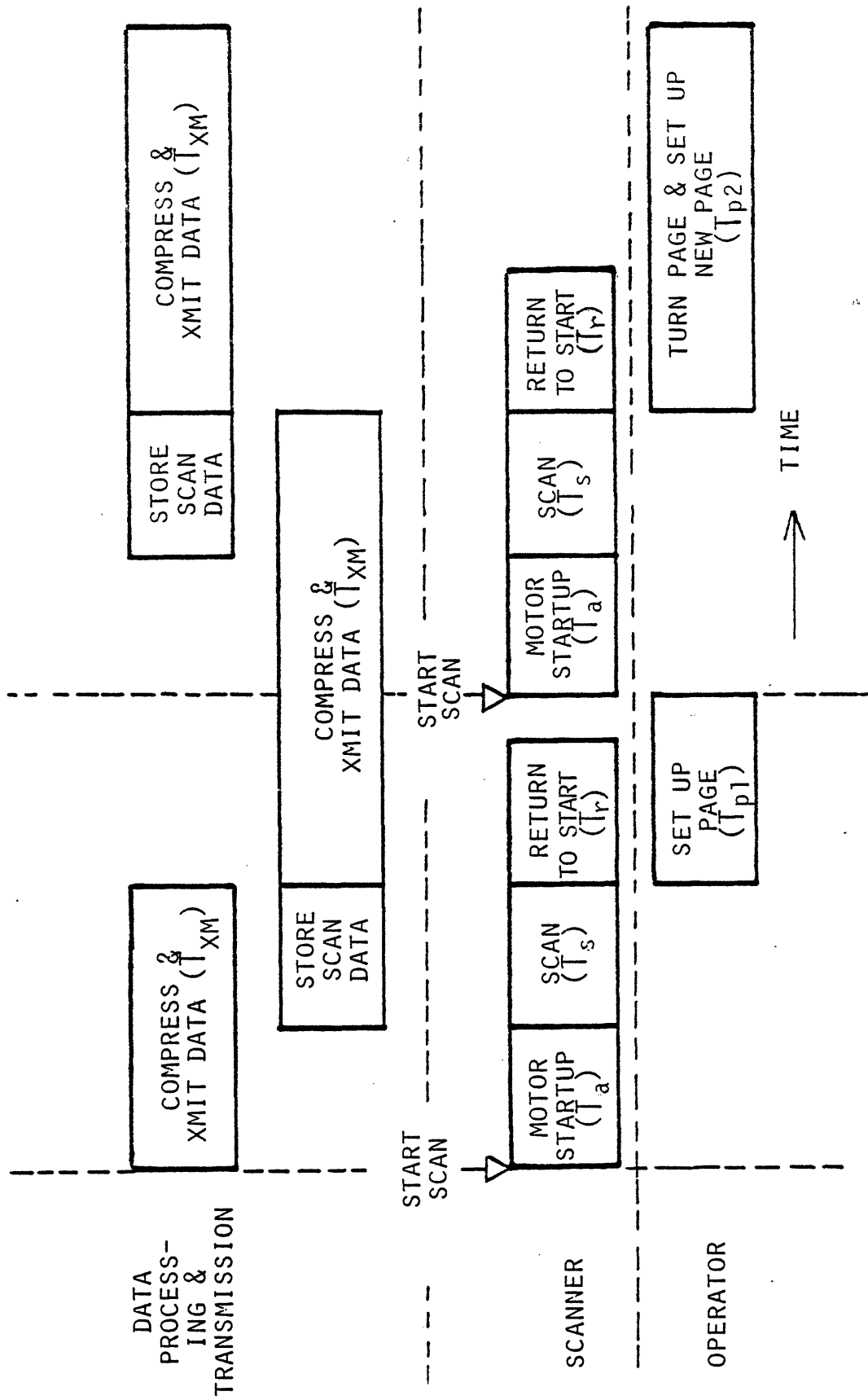
Fig. IV-3. Time relationships when two independent storage devices are available.

govern throughputs for larger values of C. On the other hand, response times for both the single-and double-page storage configurations improve rapidly as C is increased from one to five. Whether or not it is worth picking up the small extra response time that results from using two, rather than one, storage device when C = 4.3 depends on economic and human factors as well as throughput demands placed on the system; with a single-page-storage capability the throughput is 104 pages per hour in contrast to 78 pages per hour for two pages of storage. Finally, it should be noted that some improvement in system performance can be achieved by devising a book cradle that shortens the page-turning and re-positioning time. Once this operation is removed as the controlling time element, the scanner-printer response time must be shortened in order to gain further improvement in over-all response time.

Resolution versus data-compression ratio. The number of uncompressed bits of information required to encode a page increases as the square of the linear improvement in resolution. The bits per page that are generated when linear resolution is increased from 200 pixels per in. to 300 pixels per in., for example, is increased by a factor of 2.25. Hence, as resolution is improved, one must either increase the data-compression ratio by the square of the linear-resolution increase or accept reduced throughput at a constant compression ratio.

Choice of a data-compression scheme. Many techniques have been devised for compressing binary data into compact form. Basically, they all take advantage of the fact that most data consist of variable-length strings of 0's and 1's. Through choice of a set of codes that designate strings of various lengths and "color"* fewer bits are generally required, on the average, to codify the original bit string than there are bits in the original string. The amount of data compression that is achievable with a particular scheme is, in general, dependent upon the "busyness" of the material being compressed; a full page of text, for example, will be compressed less than a simple one-page line drawing.

Another characteristic of data compression is that the higher the compression ratio that is demanded, the more sophisticated will be the

---

*"color" means that a ZERO corresponds to white and a ONE corresponds to black, or vice versa.

algorithm, or coding scheme, required to achieve the compression. More circuitry will therefore be needed, and more data-processing time is likely to be consumed during compression and decompression of the raw data.

Run-length coding is commonly used as a compression scheme. In its simplest form it consists of a fixed list of binary-code words that are used to represent various runs of picture elements of the same polarity (color) along a line of an image. Choice of code words is crucial and several methods have been proposed [9, 10, 11, 12]. In general one can expect a compression ratio from 3.5 to 9 through use of one variation or another or run-length coding.

Other types of encoding that are employed include: prediction coding-- run-length encoding in two dimensions [13-16]; and edge-difference encoding--a two-dimensional scheme that works on black/white and white/black transitions [17]. Highly sophisticated algorithms that achieve compression ratios in the range of 10 to 20 are also possible.

From the results presented in this chapter, it is apparent that a simple run-length-coding scheme that yields an average compression ratio of the order of 4 or 5 is quite adequate for system characteristics akin to those of our experimental system. Should future systems with higher resolutions be required with no lessening in document throughput, more complex coding, obtained either by embellishments of run-length coding or by other encoding methods, would be required.

CHAPTER V

THE INDEX AND SOFTWARE

INTRODUCTION

The principal elements of an online electronic ordering system are the combined index of serials held within the network and the several computer programs that provide access to the index, enable requests to be placed for full text, and govern the periodic updating of the index. Separately, computer software is also required for controlling the electronic document-delivery process, including any data compression and decompression that may be present. The issues that must be resolved before implementation of these elements can proceed are discussed in this chapter. Several are rather fundamental and involve not only the structure of the data base itself but also the way in which the human operator will interact with it.

THE COMBINED INDEX

Consider, first, the various levels of computerization that are possible for a combined serials index. The lowest level would be no computerization at all. Note that it is possible to operate an online document-ordering system without an online file of serials holdings. In this mode the index would be stored in hard-copy form--either on paper or microform. A drawback of this approach is the propensity toward infrequent updating and the corresponding degree of obsolescense of the printed index.

The next higher order of computerization is to computer-store entries by code numbers only and to use an online printed list of titles (serial names) to identify code numbers. This scheme conserves computer-file space and should serve to resolve any mistakes or uncertainty in the precise title name before the computer is approached.

The third level is a fully computerized index with access gained exclusively at the computer terminal. The challenge here is to design an

index-file structure that will yield accurate information about the availability of the serial being sought on the basis of the information given by the end user to the terminal operator. A given document request could result in a computer response that the journal is: (a) in the network, or (b) not in the network. However, in case (b) the reason could be either that it is truly not among the combined holdings or that the form of the request failed to match the official name of the document as stored in the data base. Clearly, care needs to be exercised in file design to minimize such a mismatch. One possibility for an operator aid is to organize a lookup file consisting of all individial words used in all titles held in the network. Then, incomplete information in the end-user's document request could, with the aid of appropriate software, still lead quickly to correct title identification. The usefulness of this kind of operator aid is highest for operators with minimal bibliographic skills.

Several ways of expressing document requests should be considered: by full titles, abbreviated titles and coded forms. If some form of computer assistance is provided, human-factors are probably not very critical-- the computer would see to it that the correct form is transmitted to the file, and feedback from computer to operator would enable requests to be verified for accuracy.

Index-file organization. The basic issues with respect to index-file design concern the fields of information contained in each entry, or record, and how they are encoded. As a minimum, a single record should include the first N fields listed in Table V-1.

The serial title is the official name of the journal and its CODEN form is a 5-alpha character plus a 6th alpha check-character code. Alternatively, the International Standard Serial Number can be used instead of the CODEN. The ISSN is an 8-decimal-digit code written in the form dddd-dddc where d is a decimal digit and c is either a decimal digit or the character x. The CODEN or ISSN name is useful in placing document requests since it minimizes keystrokes and, hence, lessens possibility for typing errors.

The holdings of each library and branch within the library system should be expressed in a form that is cryptic enough to conserve storage space but still readily decipherable by the terminal operator. For example, one might encode a holding that starts with volume j and continues to date as vj*, where * denotes "to date". Holdings should probably be itemized by dates as well as volume and number, in which case a code such as "=" might be used as an exclusion or "except for"; for example, vj*=May, 1975 means that a holding starts with volume j and continues to date except for the May, 1975 issue. Library names and their branches can also be codified.

Table V-1. Fields of information in a single record

Field No.	Name
0	Record Number
1	Title
2	CODEN (or ISSN)
3	Library (name); Branch (name); holdings
4	Library (name) Branch (name); holdings
N	Name of Nth Library in the network and branches thereof that hold the title, plus the time span of their holdings
N + 1	Publisher's name and address
N + 2	Current subscription price
N + 3	Frequency of issue

One can envision that the record fields 0 through N could be expanded as an aid to library administration. Field N+1 might include the publisher name and address; field N+2 could be the current subscription price; N+3 the frequency of issue. Whether such information should be

included in the network file or be tabulated separately in each library's computerized business management system would have to be decided by the network members.

Index size and growth rate. Index size can be expected to vary significantly from network to network; hence, no precise generalizations can be made. Size will be some function of the number of network nodes, the degree of similarity of their individual collections, the time span covered by the combined index, the scope of coverage in the data base (all serials or subscription serials only, for example) and the number of fields in each entry.

For a given set of network members, the combined index size can be determined from serials-holdings information provided by the individual network members or from published data; Chemical Abstracts, for example, publishes a list of serials holdings, by libraries. One can get an idea of the order-of-magnitude number of serials titles that a typical research-oriented library might hold by examining two situations. At M.I.T., for example, subscription serials in 1981 numbered approximately 20,000 and the number of words per entry contained in the microfiche-stored list is estimated at 17.4. The Boston Library Consortium Microfiche Master Serials List has close to 70,000 entries (1979) and averages about 25 words per entry.* An order-of-magnitude figure for networks with attributes similar to those of the Boston Consortium might therefore be 50,000 -- it is unlikely the figure would be either 5,000 or 500,000. Such a number can be used as a rough guide to the amount of computer storage that might be required for a combined index for a few number of research-oriented libraries. Assuming a 50,000-entry index and 20 words per entry, we arrive at a storage requirement of 6.0 megabytes.

The expected growth rate for a network index is an interesting figure on which to speculate. We know that libraries are currently under severe budgetary pressures and are refraining as much as possible from expanding their list of serials holdings. Indeed many are seeking opportunities to eliminate subscriptions. Furthermore, if the interlibrary-network concept proves successful, further emphasis might be placed on the reduction of

*The Boston Library Consortium is a mix of 14 university and public libraries, three of which have multiple locations.

individual and multiple subscriptions. Nevertheless, each year new data would be added about serials that are retained. On balance, therefore, in terms of total storage requirements, one might envision a rather static-size data base over the first few years of network operation. On the other hand, our analysis of the Boston Library Consortium and the M.I.T. situations shows that entries grew at approximately a seven percent per annum rate over the past few years.

Once again, we stress that the preceding figures should not be applied as generalizations for all networks. They are intended to give the reader only a first insight to possible serials-index magnitudes and plausible procedures for estimating them.

COMPUTER PROGRAMS

Ordering system. The document-ordering system requires programs for updating the data base and for carrying out the ordering process. Prompt index updating is essential for high-quality service within the network, and frequent updates will be necessary in order to keep pace with the periodic arrival of serials at network nodes and with their appearance as references in online bibliographic systems. Monthly updates are reasonable with respect to currency of the data base and costs for doing the updating. To avoid discontinuities in service and to minimize cost, the shredding of the existing data base and the merging new information into it would probably be carried out during nighttime hours on a batch-processing basis.

Programs for executing orders for documents are governed by the ordering procedures finally settled upon. The process begins upon receipt of a request by the computer-terminal operator, who will verify the accuracy of the requested journal title either through use of offline references or by computer aids. The user's document request must include sufficient information to enable the operator to decide where, in the network, the particular article being requested can be expected to reside. As a minimum, this information will include serial title, volume number, number of the issues within the volume and page numbers; alternatively, or in addition, month and date of publication plus page numbers may be included. Note

that the title of the article being sought is unnecessary if page numbers are known, but even so, it may be useful to have the title for verification purposes.

The terminal operator's next step in the ordering process is to locate the node, or nodes, in the network that subscribe to the serial title. This can be accomplished by searching the data base by title and calling to the CRT screen all or some of the fields of information encoded in the record. Although display of the title field is not necessary, its presence on the screen may be helpful to verify the accuracy of the search. If the document's CODEN name is indexed, displaying it immediately may avoid need to keyboard a request for it later.

Having verified that one or more member libraries subscribe to the serial, the next step is to place the order. The computer-mailbox concept is appropriate for this purpose. Several design decisions must be made with respect to order placement to take care of the case where multiple nodes subscribe to the same serial. The issues are: does the terminal operator decide where the request is directed or will this matter be handled automatically? In either case, a further decision is needed on the algorithm used to construct the queue of libraries that will be tapped for the article. Several options are possible. One is to rank-order libraries so as to distribute the load as evenly as possible at the moment a new request is made; this requires knowledge of current activity in the network. Alternatively, the goal may be to average the load in the network over an extended period. Another possibility is to develop the queue on a minimum-transmission-distance principle. And still another approach, built largely on personal judgment, is to develop the queue manually in accordance with past experience of the ability of network members to fill requests and/or their promptness in doing so.

Another feature of the ordering-system software hinges on how one chooses to handle situations where the first requestee fails to supply the document either because of inability to do so or because of no action at all. One can allow a specific time limit for a response, after which the computer will automatically direct the order at the second node in the

queue, or one can make this a manual operation in which the terminal operator decides when and where to go with an unfilled request. As an aside, it should be noted that a cultural matter is involved here; the whole electronic-network concept will work well provided there is a strong commitment to it at each node. This implies, among other things, that full advantage of the online computer-mailbox principle is taken as a means of communicating the status of pending requests. Prompt, orderly work habits in the computerized environment must be established at each node to make the system work.

Another issue that must be addressed is the form in which the terminal operator places orders into the system. This relates to the user-command language that is agreed upon. The key points are that the order should be expressed in a way that is unmistakably clear at the far end, and, as an aid to accuracy as well as to streamlining the task, the number of key-strokes required to place the order should be kept as low as possible. A serial title can be requested by its index number in the data base, or by complete title or in CODEN form; the particular article can also be pin-pointed by page number within the particular issue; the issue, in turn, can be identified by volume and volume number or by date of issue.

The statistical-gathering ability of a computerized ordering system is among its strongest assets. Properly designed, this feature can provide insights into the movements of serials information that heretofore have been very time-consuming and uneconomical to gather. Among the data that are easily tagged in the process of filling document requests are these: The number of requests placed for a particular serial over a specified time period, the frequency of requests for a given volume, issue and article in the serial; the traffic flow among network nodes; percentage of requests filled (or unfilled) and profiles of unfilled requests--not held by network members, subscribed to but not available, and so forth. Programs for compiling such information will be part of the document-ordering software system; their form will depend on the measurement points decided upon.

<u>Document-delivery software.</u> Included in this software are programs that control scanner-printer operations in real time and execute the data-compression computations if data compression is used. The programs reside in the special-purpose digital controller configured to handle the control functions. Once a scanner and printer are operating satisfactorily under the supervision of the digital controller, there will likely be no need to revise the software. Hence, the programs can be imbedded permanently in a programmable read-only memory and made part of the controller.

CHAPTER VI

ECONOMIC STUDIES

INTRODUCTION

Our economic-feasibility studies of electronic interlibrary resource-sharing systems included an estimation of capital and annual operating costs of such systems and an analysis of possible savings that might result from their use. Since proven systems are not purchasable "off-the-shelf" from vendors, it was necessary to devise cost models and make estimations based upon the models. These models can serve as guides for cost-estimating similar system configurations.

CAPITAL EQUIPMENT

We confined our cost estimates of capital equipment to items described in earlier chapters and configured as shown in Fig. I-2. This approach is consonant with our over-all goal to investigate networks that would be financially attractive to the private-library community. A list of items needed at one node of an online ordering system and a separate electronic document-delivery system is presented in Table VI-1. The cost of items that are currently purchasable are quite accurate; for nonpurchasable items, amounts are what we estimate a commercial supplier might charge, assuming the supplier had no development costs or had previously written them off. In the case of the bound-document scanner, our experience with an experimental scanner was valuable in arriving at a representative figure.

We chose to include software packages as capital items, since they are "one-shot, up-front items". Two types of software are involved: One relates to the online-ordering system, the other to the document-delivery system. Since our software studies extended only to a delineation of issues rather than to detailed designs, cost estimate must be based on judgment and experiences with other software systems that appear, at least, to be of comparable functional complexity and magnitude.

Since cost estimation of future vendor-generated software packages

is by no means straightforward, special attention is given to this subject. Contributing to pricing policy are: the magnitude of the development effort, policy with respect to recoupment of development cost, market potential, the readiness of the market, expected life span of the packages, software-maintenance policy, expected return on investment and tax rates. Obviously, the assumptions made about all these points will influence final estimates.

TABLE VI-1. Capital equipment at one node.
A four-node network is assumed

Item	Available from Vendors	Approximate or Estimated Cost
For online ordering:		
1 CRT Terminal	Yes	$ 900.
1 1200 bit/sec modem	Yes	900.
Software package (see text)	No	6,000.
For document delivery:		
1 Bound-document solid-state scanner	No	4,000.
1 Electrostatic printer	Yes	10,000.
1 Microprocessor/controller	Yes	1,000.
1 Storage Disk and Controller	Yes	5,000.
Controller software (see text)	No	3,600.
Power supplies	Yes	1,000.
Equipment-interface electronics (see text)	No	3,600.
Total		$ 36,000.

Online-ordering software. As a means of arriving at a specific figure for online-ordering software, we made the following assumptions:

Two man-years of professional programming are required to develop the packages to a set of specifications.

A man-year of effort equates to $80,000., this amount to include

salaries and benefits of the professional programmer and all technical and administrative support staff, materials, services, and indirect costs chargeable to the programmer's activity.

An extra ten percent is added to the $80,000 figure to make the software compatible with multiple machine environments.

A market potential of 10 sets of packages is assumed before obsolescence caused by upgrading of the packages and/or the appearance of competitive packages.

All development costs are charged against the first four packages sold.

An after-tax return on investment of 40 percent will be realizable on the last six packages, only.

Net tax rate of 35 percent.

Total per-package charge, for the final six, equals 1.25 times the before-tax return on investment assignable to the package, the additional amount representing marketing, distribution and administrative costs. (Software-maintenance charges are assumed to be an operating cost item.)

Calculation of a set of software packages on this basis yields $22,500 per package on the last six sets <u>after development</u>, and $44,000 per package on the first four. However, since only one set of software is required for the whole network the amount allocated to each node depends on network size. In order to arrive at a specific capital-cost figure, we have assumed a four-node network and have rounded off the $22,500/4 = $5,600 figure to $6,000. This amount is entered in Table VI-1, third line.

<u>Document-delivery software.</u> This software comprises the programs that allow the microprocessor to control all document-delivery operations in real time at each node, and it directs the data-compression and decompression circuitry at each node. Again, since the exact software requirements will depend on the particular system organization and form of data compression chosen, cost estimates, only, can be given.

Based on our experience with our Laboratory document-delivery system and on analysis of data-compression requirements, we estimate that one-man year of program-development effort would be required. Using the same assumptions as before (except that no allowance need be made for different machine environments), we arrive at a per-package cost of $8,200. for each of the final six packages. The only question remaining, then, is how to allocate the $8,200 among each node of a multinode network, keeping in mind that an identical set of software packages will be needed at each node and that the programs will likely be contained in programmable read-only-memory chips (PROMs) at each site. Again, we faced a judgmental matter and we settled on the following formula: Charged to the entire network will be the full cost of the first package set plus 25 percent of full cost for each succeeding package set in the network. Hence, for a four-node network the total document-delivery software would be $8200 plus 3x$2050, or $14,350; averaged over the four nodes, this figure becomes $3,600 per node, rounded off. This amount has been entered in Table VI-1, line 8.

Finally, we need a figure for the electronic-interface equipment for the document-delivery system. This equipment will undoubtedly consist of circuit boards of solid-state logic elements for adapting the printer to the particular data format employed, for interfacing the scanner and micro-processor and making the local buffer-storage/controller devices compatible with the rest of the system. We estimate the per-node cost of the inter-face electronics to be similar to that of the document-delivery software and have assigned $3,600 to this item (line 10).

Thus we arrive at a per-node capital-equipment cost after development write-offs of $36,000. If one chooses, one can go further by assuming a specific cost-recovery period and an interest rate, and then calculating the per-annum charge for principal and interest over the recovery period. For example, if the interest rate is 12 per cent per annum and the recovery period is four years, the per-annum capital cost in each of the four years is: $36,000xF, or $36,000x.0.32923 = $11,850., F being the cost-recovery factor*

---

*The capital-recover factor F, which is the factor by which the Amount is multiplied to obtain the per-annum payment, is $F = i(1 + i)^n / [(1 + 1^n -1]$ where i is the interest rate and n is the number of years for capital recovery.

OPERATING COSTS

Operating costs are divided among staff salaries and wages, computer and communications fees, and system-maintenance charges.

Salaries and wages. This item is based on the assumption that at each node one paraprofessional handles the ordering system and a second paraprofessional operates the document-scanner/printer system, both on a full-time basis. They would be assisted on a half-time basis by a student employee or the equivalent. An annual salary of $24,000 is estimated for each paraprofessional; this amount includes employee benefits and indirect costs. Seven thousand dollars per year are allowed for the student-employee equivalent, including indirect costs. These items are listed in Table VI-2.

Table VI-2. Representative first-year operating costs per node of a four-node network.

Item	Estimated Cost	Totals
Salaries and Wages:		
2 Paraprofessionals (full-time)	$ 48,000	
1 Half-time employee	7,000	$55,000
Computer Fees:		
Connect time	6,500	
Storage	1,500	8,000
Communications Fees:		
High-Speed lines (see text)	24,100	
Low-Speed line	3,560	27,660
Maintenance Service		5,000
		$95,660

Computer fees. Computer fees include two components: storage costs for the index of documents held by the network and computer connect time, which, in turn, is composed of index-updating time, and the time consumed in placing orders for documents. If one assumes 30 hours of ordering time per week at $4.00 per computer-connect hour, $6,240 per annum must be

budgeted at each node for this purpose. To this amount we have added $260 per node for updating the index periodically during the year. [Note that total index-update cost is divided among the node numbers]. These two items are posted in Table VI-2. in the amount of $6,500.

Storage costs are principally a function of the size of the network's integrated index and the amount of information that is included in each entry. A small amount of storage is also needed for storing the programs for updating the index and for interactive online ordering. Based upon the discussion of storage requirements given in Chapter V and an estimated storage charge of $1.00 per thousand bytes per annum, we arrive at a first-year figure of $6,000 for storage cost for an integrated index for a four-node network, or $1,500 per node. This figure will increase steadily, of course, as the data base expands.

We turn, now, to another component of operating cost -- communications fees. These are composed of charges for the low-speed lines for connecting terminals to the online-ordering system and the high-speed lines for document transmission. In order to arrive at specific cost figures, we must make an assumption about the average length of these lines. Let us take 10 miles as the average length. Then, a representative per-month lease cost for a low-speed line, obtainable from published tariffs, is $297. per month ($3560. per year). Equation III-4 can be used to arrive at high-speed-line figures. It is $1338.50 per month ($16,062 per year) for a 10-mi line. However, the total required number of high-speed lines depends on the number of nodes in the network, the number being six for a four-node network in order that every node can be connected to every other node. Hence, the per-node annual-lease cost is $16,062 x 1.5, or $24,100, rounded off.

The final operating-cost item is for annual system maintenance. Systems purchased on a "turn-key" basis would likely be serviced by the vendor for a flat annual fee. A $5,000 per-annum maintenance fee is a reasonable figure for maintenance that includes service for equipment and computer programs.

INTERPRETATION OF COST FIGURES

The sum of the first-year, per-node operating cost and the annualized per-node capital cost amounts to $107,500 for a four-node interlibrary resource-sharing network configured along lines of Fig. I-2. Whether or not this expenditure is justifiable depends on several factors: The annual cost now being incurred by libraries to do "borrowing and lending" of serials in the traditional manner; projections of future costs for the conventional and machine approaches; savings that might result in electronic networks because network members are able to avoid replications of certain serials at each node; and the value of shortened response times to the end-user community to document requests.

We can cite rather firm cost figures for the serials borrowing-and-lending experience at our own Institute. There are indications that these figures are representative of those for other universities, although we gathered no other hard data. At M.I.T., the annual cost of interlibrary serials transactions is estimated at $110,000 for 1980-81 and is increasing at an annual rate closely tied to the inflation rate. If the average rate of inflation over the next ten years is 7 percent, for example, the tenth year cost will be $216,400. The current $110,000 figure is obtained as follows: The average cost of a transaction, (weighted with respect to borrowing and lending) is currently $8.15 and 13,450 transactions are completed during the year. One must decide, of course, whether or not these figures are generally representative of those incurred at major research-oriented libraries. We think they are.

A direct comparison should not be made between the $107,500 per-annum cost for the electronic network and M.I.T.'s photocopy-in-lieu-of-lending per annum cost of $110,000. The electronic-network figure is based on four nodes spaced a maximum of ten miles apart, whereas the photocopying figure encompasses transactions with many other libraries at widely dispersed locations. One interesting comparison can be made, however. The $110,000 per-annum photocopying figure is based on 13,450 transactions. If one assumes the electronic network upon which the $107,500 amount is based is busy 8 hours a day during each 5-day work week of the year,

the network has a potential annual throughput capability of approximately 53,000 documents. Hence, if the electronic network can be kept busy most of the work-day, the per-unit transaction cost will be much less than that of a photocopy transaction.

Figure VI-1 is a ten-year projection of the electronic network costs. The over-all cost curve is based on the assumption that the electronic equipment will function satisfactorily for ten years and network members will choose not to install replacement equipment during the ten-year interval. Observe that in the digital mode capital equipment accounts for a minor percentage of total annual per-node costs. Labor costs are substantial, but they can be reduced if end users, themselves, are allowed to do their own ordering at the CRT terminals. It is conceivable that introduction of this self-service feature will be feasible after the user and library communities have gained experience and confidence in the electronic-resource-sharing concept, but at the outset we suggest that a trained intermediary is preferable.

Clearly, there is great potential for savings in the cost of library operations if resource sharing through use of electronic networks proves so effective that network members can rely on each other to fill a major number of requests for information contained in serials. When this happens libraries can plan their annual subscriptions to serials on a network basis, rather than as individual libraries. The cost to libraries of owning periodical publications has been analyzed and reported. The Palmour study [18] indicates that paid annual subscriptions in the year of the study (1976 - 77) varied from $53.53 to $70.65, with the higher figure assignable to scientific and engineering journals. The cited study also assigns a storage cost of fifty cents per annual volume and approximately $25.00 per annum as an annual recurring cost. The latter cost is for annual review and selection of titles, subscription-renewal operations, check-in, shelving, binding, and so forth. For comparison, we have hard data for our own Institute; at M.I.T. the 1980-81 budget for 9,000 journal subscriptions was $850,000, and this figure is expected to increase at a rate of 15 percent per annum. Should interlibrary electronic document-delivery networks prove highly satisfying to the user and library communities, reducing the
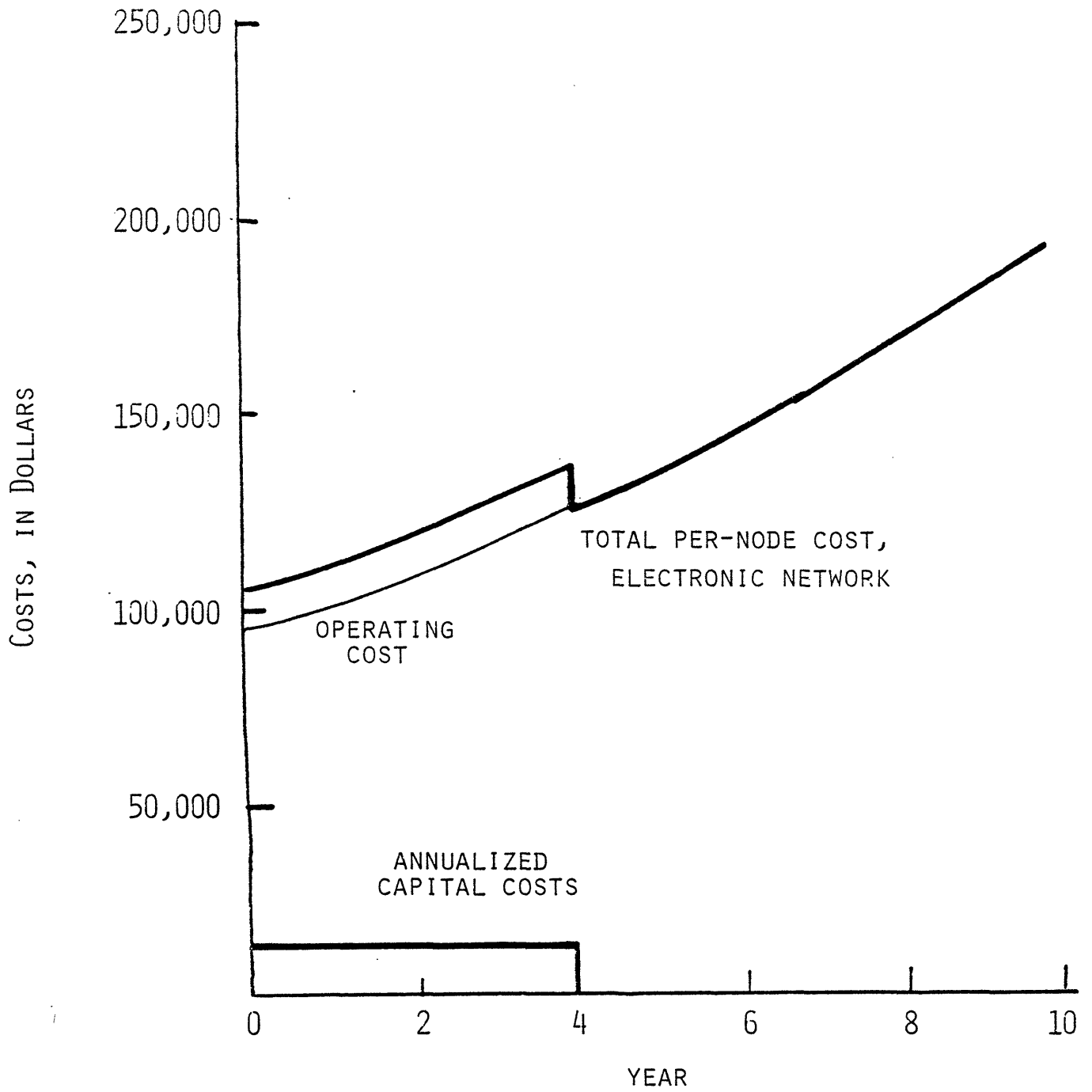
Fig. VI-1. Interlibrary document-delivery cost figures versus years from 1981-82. A seven percent annual inflation rate is assumed.

total number of serials holdings within the network become a real possibility and corresponding savings could be substantial -- perhaps enough eventually to offset the annual operating cost of the whole network.

CHAPTER VII

SUMMARY

Our investigations have shown that electronic networks of libraries assembled for the purpose of transferring information contained in serials are technologically feasible. Such networks are also economically attractive. A major cost factor resides in the communication links employed between network nodes. Hence, close attention should be paid to this element when networks are being synthesized.

It was clear from the outset uncertainties about the soundness of the concept stemmed from the unavailability of low-cost bound-document scanners. We therefore experimented with scanners to the extent that it was necessary to convince ourselves that they pose no inherent bottleneck. We satisfied ourselves there is not, but there is room for much more research on those devices. Such research should be aimed at improvements in resolution quality, refinements in the method used to support documents, automatic circuitry for setting voltage-threshold level in the analog-to-digital conversion part of the scanner and in alternative optical/mechanical configurations that would yield compact scanners.

We believe a resolution goal for bound-document scanners should be of the order of 300 lines per inch. In light of steady advancements being made in solid-state sensors such as charged-coupled and similar devices, attainment of this goal should be possible in the near future, say, in two or so years from how. It is less certain that a low-cost digital printing device that matches this resolution will be available, however, although there is evidence that a low-cost laser printer with a 300-line-per-inch resolution capability is nearing the final stages of development.

Our research has led us to the conclusion that the design of an electronic library network should be approached on a total system basis. To take full advantage of the available technology, they should perform both a document-ordering and a doument-delivery function. Since many tradeoffs are possible, each network should be treated uniquely, and specifications should

take into account factors such as anticipated volume of traffic within the network, distances between network nodes, number of nodes and the amount and kinds of materials to be handled. Each of these items will influence the detailed design of the network.

The choice of communication links for the document-delivery part of the network is crucial. Communication impacts upon traffic capacity, system complexity and operating cost. It will be the factor that will largely determine the geographic extent of the network, especially where minimization of cost is a prime consideration.

Electronic networks of libraries are suitable for intra-organizational as well as inter-organizational connections. In fact, it could well turn out that their first usage will be to electrically couple the various branches of an intra-organizational library system -- the library complex of a university, for example, or the branches of a municipal library system. We envision that such networks will provide wholly new opportunities for libraries in their dealing with serials through their ability to upgrade the quality of end-user service and to make fine-grain measurements of document usage.

Finally, we point to barriers that must be overcome in order for electronic networks to come into existence within the library community. One is that the library community must be convinced, before making an investment, that such networks can be made to work reliably and that the quality of end-user service they render will not only be uninhibited but actually enchanced. Another is that the industrial sector is unlikely to invest its own resources to make complete systems available on a "turn-key" basis unless a foreseeable market exists. Either subsidized demonstration experiments will be needed to break the deadlock, or high-risk capital must come into play.

A LOOK TO THE FUTURE

Impending advanced digital technologies such as very large-scale integrated circuitry (VLSI) and mass, archival, optical-disk storage devices will eventually offer opportunities for drastic changes in the

way we publish, store and retrieve information. Direct digital-electronic publishing and storage of new knowledge will become possible as soon as these technologies are perfected. Whether knowledge already recorded in printed form will be converted to the digital domain, or whether we shall have to contend with a dual-mode environment until printed documents fade into history and disuse is a moot question. In a two-mode environment one can conceive of digital mass storage as an add-on feature to the document-delivery concept discussed in this report. Refer to Fig. VII-1. Here an option is provided to deliver materials either by scanning printed material or retrieving them from a mass digital store. The complexity of such a system depends on the degree of compatibility that is built into the combined system. The techniques employed in each mode for handling graphics, data-compression and decompression and data transmission may well determine whether an integrated system will be possible or separate delivery systems will be necessary.
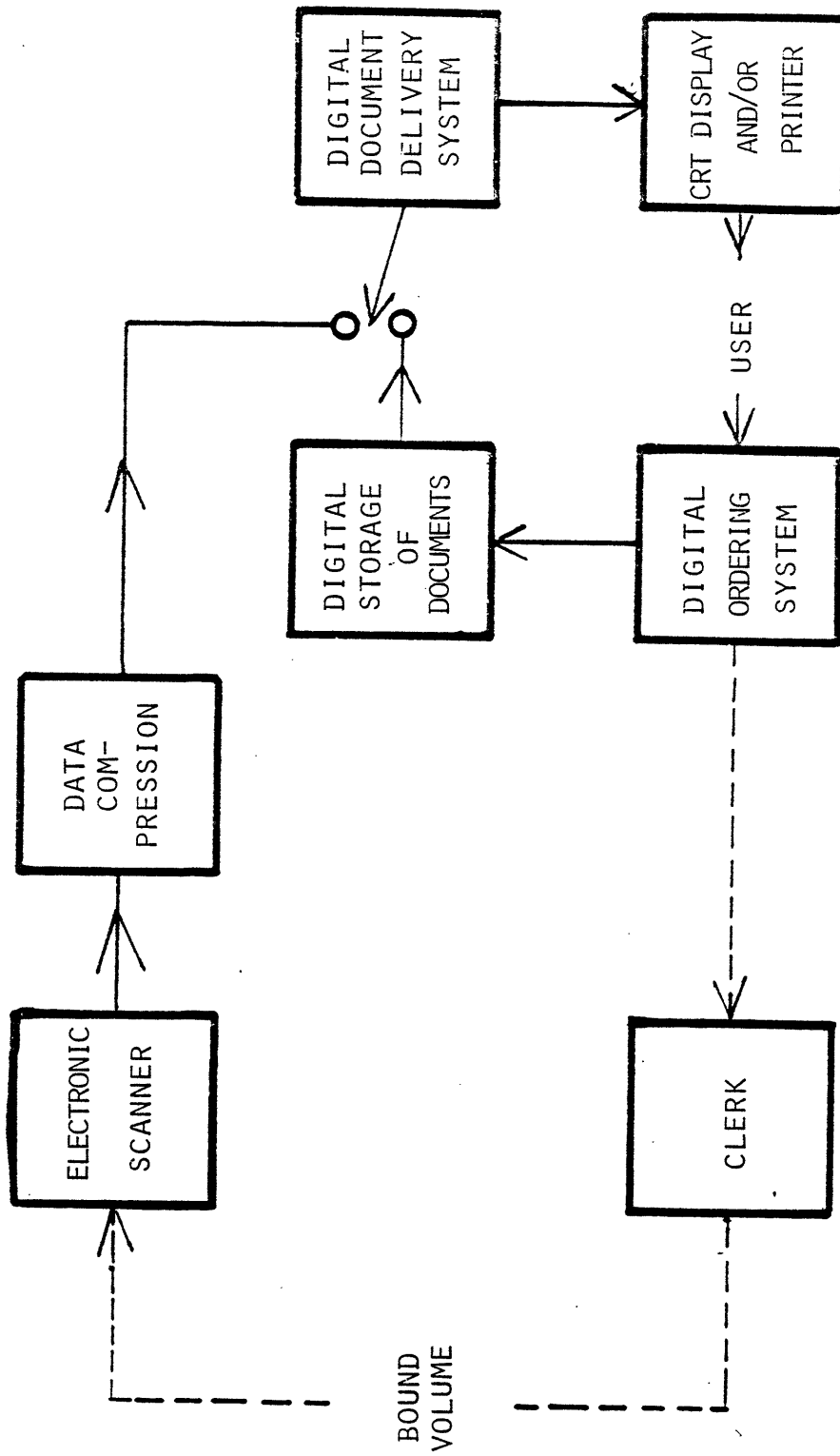
Fig. VII-1. A possible configuration for a future electronic ordering and delivery system.

BIBLIOGRAPHY

1. Library Photocopying in the United States. National Commission on Libraries and Information Science, October, 1977. U. S. Government Printing Office, Washington, D. C. 20402. Stock No. 052-003-00443-7.

2. Trevvett, Melissa D.. Characteristics of Interlibrary Loan Requests at the Library of Congress. College and Research Libraries, January, 1979.

3. Reekes, Pamela, Interlibrary Loan Activity at M.I.T.: Characteristics and Costs. A report submitted to Simmons College, School of Library Science, May 3, 1976.

4. Ferriero, David S., Interlibrary Borrowing for the M.I.T. Libraries, December, 1978. An internal memorandum.

5. Knudson, Donald R., Digital Encoding of Newspaper Graphics, Massachusetts Institute of Technology, Report ESL-R-616, August, 1975.

6. Hardy, A. C. and Perron, F. H., The Principles of Optics, McGraw-Hill, New York, 1932.

7. Smith, W. J., Modern Optical Engineering, McGraw-Hill, New York, 1966.

8. GTE TELENET public network rate schedule pricing summary effective June 1, 1981, Telenet Communications Corporation, Vienna, Va., 22180.

9. H. Meyr, H. G. Rosdolsky, and T.S. Huang, "Optimum Run Length Codes," IEEE Trans. Commun., vol. COM-22, No. 6, pp. 826-835

10. D. A. Huffman, "A Method for the Construction of Minimum Redundancy Codes," Proc. IRE, vol. 40, No. 10, pp. 1098-1101, September, 1952.

11. H. G. Musmann and D. Preuss, "Comparison of Redundancy Reducing Codes for Facsimile Transmission of Documents," IEEE Trans. Commun., vol. COM-25, No. 11, pp. 1425-1433, November, 1977.

12. J. B. Connell, "A Huffman-Shannon-Fano Code," Proc. IEEE, pp. 1046-1047, July, 1973.

13. A. N. Netravali, F. W. Mounts, and E. G. Bowen, "Ordering Techniques for Coding of Two-Tone Facsimile Pictures," B.S.T.J., 55, No. 10 (December, 1976), pp. 1539-1552.

14. A. N. Netravali, F. W. Mounts, and J. D. Beyer, "Techniques for Coding Dithered Two-Level Pictures," B.S.T.M., 56, No. 5 (May-June 1977), pp. 809-819.

15. F. W. Mounts, A. N. Netravali, and K. A. Walsh, "Some Extensions of the Ordering Techniques for Compression of Two-Level Facsimile Pictures," B.S.T.J., 57, No. 8 (October, 1978), pp. 3057-3067.

16. F. W. Mounts, E. G. Bowen, and A. N. Netravali, "An Ordering Scheme for Facsimile Coding," B.S.T.J., 58, No. 9 (November, 1979), pp. 2113-2128.

17. T. Yamada, "Edge-Difference Coding--A New, Efficient Redundancy Reduction Technique for Facsimile Signals," IEEE Trans. Commun., vol. COM-27, No. 8, pp. 1210-1217, August, 1979.

18. Palmour, Vernon E., et al, "Costs of Owning, Borrowing and Disposing of Periodical Publications." Public Research Institute, a Division of the Center for Naval Analyses, October, 1977.