On Line Data Processing

Proceedings of Two Sessions on
On Line Data
Processing Applications
To Be Presented at the
Winter General Meeting
New York, N.Y.
January 27 - February 1, 1963

Sponsored by the AIEE Subcommittee
on Data Processing Applications
of the
AIEE Computing Devices Committee

January 1963

I·E·E·E

Published by
THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, Inc.
Box A, Lenox Hill Station, New York 21, N.Y.

Printed in the United States of America

PRICE: $2.00 Members
$2.75 Nonmembers

FOREWORD

The Data Processing Applications Subcommittee of the Computing Devices Committee is pleased to present in this volume copies of papers to be presented at two sessions on On Line Data Processing at the Winter General Meeting, New York, N.Y., January 27-February 1, 1963.

On Line Data Processing refers to those applications in which data is received over a multiplicity of input links which are asynchronous with respect to processor. This collection of papers covers different aspects of the on-line field which is becoming increasingly important and in which there is considerable current interest.

These papers were stimulated by discussions at a Workshop on this subject held in Palo Alto, Calif., in May 1962.

Thanks are due to all who responded to the Call for Papers for these sessions and to the authors of the papers contained herein.

Thanks are due also to Dr. D. R. Helman of ITT Federal Laboratories, Nutley, N.J., for his contribution in coordinating the solicitation and review of papers and organizing the two sessions in which these papers will be presented.

C. A. Strom, Subcommittee Chairman
Data Processing Applications Subcommittee
Computing Devices Committee

Task Group on Papers

D. R. Helman, Chairman

PREFACE

In order to fully utilize the capability of present high-speed processors, it is becoming more and more necessary to automate data collection and distribution. Without some method of getting large quantities of data in and out of a processor, the problem set up time is apt to exceed the problem solving. In answer to this problem, there has evolved a large class of processors which are loosely referred to as on-line or real-time processors. These are generally characterized by a multiplicity of input and output devices which may be local or remote in nature. Since input data pertaining to many problems may be arriving simultaneously, the need for a multi-program processor is also a possibility that must be faced.

Session I will deal with various considerations which must be faced in dealing with a real time system. Such subjects as economics, memory needs, programming, and real time displays will be covered.

Session II will discuss several actual and proposed uses for real time processors. The topic covered will be store and forward switching, the concept of computer utilities, a data acquisition system, and an on-line bank processor.

TABLE OF CONTENTS

ASSESSING COMPUTING SYSTEMS

BY

LOUIS FEIN

Computers, Computing Systems, Assessing, Productivity, Cost, Effectiveness, Worth, Data-Processing, Figure of Merit, On-Line, Real-Time.

ASSESSING COMPUTING SYSTEMS

Louis Fein

SUMMARY. Four figures of merit--productivity, cost, effectiveness, and worth--useful to manufacturers for quantitatively assessing their data-processing systems, and to users for comparing and contrasting competing systems, are defined; design parameters affecting each figure of merit are given. Procedures are suggested for calculating the values of the figures of merit from the estimated values of certain system parameters on which they depend. Of the four figures of merit, worth is the most comprehensive.

INTRODUCTION

Productivity, Cost, Effectiveness, and Worth are figures of merit of computing systems that the author has introduced and used to assess control computers, gigacycle computers, and data-processors and to influence the design of such systems. 1,2. In this paper, these figures of merit are refined and procedures are given for calculating their values from estimated values of certain system parameters on which they depend. Since on-line real-time systems constitute one class of computing system, these figures of merit are suitable for them.

ON-LINE REAL-TIME SYSTEMS

An on-line data processing system consists of one or more central data processors with a multiplicity of remote but directly connected input and output terminals. Each equipment subsystem of the composite system is continuously and crucially involved in the operation. Since batching of inputs is unacceptable, no buffers are required. If an on-line system is to operate in real time then its reponse time must be acceptable to the system user. Thus, on-line and real-time are descriptors of physically integral data processing sub-systems constituting a system with a generally short (and acceptable) response time. It will be seen, however, that what will be said about the figures of merit of programmable data-processing systems in general is equally applicable to real-time or to non-real time systems in particular. The only significant parameters that must be taken into account for an "on-line" system but may be ignored for an "off-line" system are those associated with the communication links between the remotely located input and output terminals and the central processor(s). Whereas, in practice, the value of these communication parameters may seriously affect some properties of an on-line system, in principle, they present nothing new.

FIGURES OF MERIT

A figure of merit of a system is a numerical measure of a significant property of that system. Users and designers have used a variety of figures of merit to assess-and to compare and contrast--programmable data-processing systems; e.g. memory size, arithmetic speed, reliability, and price. But whereas these are important parameters of a system, they are usually only contributory to the appraisal of a composite system.

Four composite properties, deemed significant by users or manufacturers or both, have been selected for definition; a procedure will be given for obtaining their numeric measures; system parameters influencing each of the four figures of merit will also be identified.

Productivity

A popular measure of the attractiveness of a system is the amount of work it does of a certain kind in a given interval of time. The amount of work may be measured in terms of numbers of certain kinds of products produced or of operations performed, or in terms of the complexity of a product, or of an operation. A data-processing system used for payroll produces a number of paychecks and stubs of a certain complexity of format in a few hours; electronic banking machines produce a number of bank statements in a given period of time; inventory systems cope with an inventory while handling transactions of varying complexity at a certain rate; on-line hotel accounting systems produce complete customer invoices for hotels of a given number of rooms; computers operate on words of a given size with certain average arithmetic speeds; they solve equations of a certain type in a given interval of time. Apparently, there are at least as many suitable measures of work as there are classes of applications and types of operations.

The productivity parameters of a system are those characteristics of the system whose value or type influence the productivity, i.e., the amount of work done by a certain program in a given time on a particular class of tasks. In general, the productivity is some function of these productivity parameters; furthermore, many of the productivity parameters are mutually dependent. Representative productivity parameters of programmable on-line data-processing systems include speed of execution of individual instructions, size of computer word, capacity and access time of internal and peripheral stores, organization of the sub-systems and of the parts of each sub-system, efficiency of programs, channel capacity of the communication sub-systems, and reliability. Reliability itself has various measures affecting the down-time and hence the elapsed time to perform a task. These measures include error-detection time, error-diagnosis time, and error-repair time as well as the rework time--the time taken to rerun the program from the point of error. Program efficiency depends at least on system organization, storage size, and the capability of the program-writer, be it a compiler or be he a human being.

Because the productivity is rarely expressible as a known explicit function of the productivity parameters, and because the interdependence of the parameters is equally difficult to make explicit, the productivity of a system can rarely be calculated; it must usually be determined experimentally. To do so, a task must be selected, a program written by a compiler program or by a human programmer, and the program run on the system whose productivity is being ascertained, or on a simulated system. The productivity, expressed in appropriate units of work and time, will be a measure of the rate at which work is done on that task by a specific program.

To compare competing systems by productivity, the same tasks should be programmed and performed on all systems; furthermore, if the programs for competing systems were compiled from the same source program and were written in the same programming lanquage, then the productivity of each system will reflect the excellence of the compiler; if the programs were written by human programmers, then it will reflect their competence. Indeed, to increase the productivity of a system on a certain class of tasks, the manufacturer may consider the modification of others of the productivity parameters, e.g., memory size and access time, or system organization, or reliability. But, it is worth repeating and emphasizing that it is usually impossible to calculate the effect of changes in values or types of productivity parameters on productivity; the effect must usually be determined experimentally or by simulation.

Cost

A popular, but misleading, figure of merit, of a programmable data-processing system is its purchase price or rental. It is misleading because the total annual expenses, measurable in dollars, incurred by the user of a typical data-processing equipment system is from three to five times the annual rental (or price). Rental and price are only two of the several cost parameters, (i.e., only two of the accounts in the chart of accounts used in typical cost accounting). Here is a typical list of such accounts, which reflect the expenses incurred by a user in order to install and operate a programmable data-processing system: Feasibility study, Site Preparation, Installation, Conversion, Parallel Test; Price or Rental, Taxes, Insurance, Depreciation; Administration, Supplies, Housing Equipment, Spare Parts, Equipment Maintenance, Operation, Power, Cooling; Programming Expenses--Training, Programming, Coding, Compiling, Assembling, Testing, Maintaining, Documenting.

The annual cost of a system can easily be estimated from estimates of the values of these cost parameters. Thus each user can compare competing machines by annual cost to him. Since values of some cost parameters are unique to particular users, e.g., Programming and Taxes, the ranking by annual cost of several systems by one user may differ from the ranking of the same systems by another user.

The designer can thus try to influence and reduce only those costs that are common to all the users of his equipment system. Price or rental is one such item of expense; increased reliability, should result in decreased maintenance expense and improved system organization and order code may lead to reduced programming expense.

Effectiveness

Effectiveness--the hourly or weekly, or monthly, or annual production (work done) divided by the hourly, weekly, or monthly, or annual cost to the user, is a popular figure of merit for comparing and contrasting systems. The units of production may be the same as the units of work in productivity, e. g., number of checks, number of equations of a certain type, number of transactions, etc. The unit of cost is the dollar. Because of the user-dependent productivity parameters and user-dependent cost parameters, different users may assign different values of effectiveness to the same equipment system; for the same reasons, users might differ in their ranking by effectiveness, a set of competing equipment systems. Clearly, a manufacturer might be hard put to it to design his equipment system so that it would be at the top of the effectiveness list for every user.

Worth

What it is worth to a user to have and operate a system of a certain productivity, or cost, or effectiveness is almost always unique to that user. A user may obtain both tangible, measurable gains and intangible gains in other parts of his establishment that he knows are there, but the magnitude of which, he is unable to predict because he does not know the functional relationships. For example, suppose that an on-line programmable data-processing system would displace a certain number of persons. The tangible worth to the user may easily be measured in dollars, and a decision to install such a system may depend on whether the worth, measured say in annual savings, exceeds the annual cost. Intangible benefits may include gains obtained because the system provides such things as timely reports on budget performance, schedule performance, quality performance, product cost analysis, labor forecast, finance forecast, equipment utilization and labor distribution charges.

Worth-Cost

The most comprehensive criterion for ranking programmable data-processing systems is the worth-to-cost ratio (or difference) where presumably those systems with a worth-to-cost ratio less than or equal to one, or whose worth minus cost is less than or equal to zero, are unacceptable. Furthermore, the user who estimates the worth of the intangibles and adds this estimate to the tangible worth (or subtracts it if it is of negative value) has the optimum criterion for making a rational quantitative decision for selecting programmable data-processing systems.

CONCLUSION

Memory size and access time, reliability, arithmetic speed, instruction repertoire, channel capacity, price and other such properties of data processing equipment systems are misleading if used as a sole criterion for ranking these systems. Even more composite parameters such as productivity, cost, and effectiveness may be misleading as figures of merit. The worth-to-cost ratio (or difference), with intangible and tangible value included in the worth, is the most comprehensive and valid figure of merit for the user of a programmable data-processing system.

REFERENCES

1. "A Figure of Merit for Evaluating A Control Computer System," Louis Fein. Automatic Control, Vol. 12, May 1960, pp. 39-41.

2. "Suitable Applications for Computers Built of Gigacycle Storage and Logic Components," Louis Fein. AIEE Winter General Meeting, January 31, 1962.

MEMORY CONSIDERATIONS FOR AN ON-LINE PROCESSOR

BY

E. E. BARRETT

Computers, Memory, On-Line, Processor, Real-Time, Formats, Codes, Per Line Batching, Multi-Programming, Batching, Central Storage, Overflow Storage, Program Storage, Temporary Data Storage, Permanent Storage

MEMORY CONSIDERATIONS FOR AN ON-LINE PROCESSOR

E. E. Barrett
ITT Federal Laboratories
Nutley, N.J.

Summary

The advent of on-line or real time systems in which incoming data flow cannot be regulated by the processing device greatly increases the amount of storage space required for in-process data. Further complications arise because of such problems as different formats, codes, and processing requirements for various groups of inputs and outputs.

Several methods of providing sufficient data storage are discussed. These will include per line batching, multi-programming, and batching in a central storage. A quantitative comparison of these methods is made and the possible trade-offs between program time and hardware is discussed.

The use of overflow and permanent storage devices will also be discussed. Possible methods of reducing or eliminating the need for this storage will be covered.

A possible arrangement of a processor to provide the necessary functions in a reliable, economical manner will be briefly discussed in conclusion.

Introduction

In order to avoid confusion, it would be well to first define just what an on-line system is. Although these systems may have an infinite variety of special requirements, they will all have the following characteristics:

1. A multiplicity of input lines which provide independent data for processing. These lines must normally be served on demand and cannot be controlled by the central device.

2. A multiplicity of output lines which return processed results to the sub-stations or to display devices. These outputs are normally under control of the processor.

3. A need to provide the necessary processing and results in real-time. This time interval may range from seconds or less in a control system, to minutes or even hours in a routine bookkeeping application.

4. Processing routines which may range from a simple canned response for a fixed number of queries, to an elaborate data correlation routine for driving a real-time display.

In order to perform these tasks, a variety of storage functions must be provided for. Each of these will now be discussed in turn. The types of storage to be discussed are:

1. Program Storage
2. Temporary Data Storage
3. Permanent Data Storage
4. Overflow Data Storage

Program Storage

The program storage required for an on-line processor includes the normal program sub-routines and reference tables. However, it is also necessary to store many main programs instead of one or two. In an extreme case it may be necessary to have a separate program for each input and output device connected to the system. All of these programs may be active simultaneously on an interleaved basis. This requires that they be maintained in a quick access store such as a core memory. The operational program storage requirements may range from 3,000 to 20,000 words depending on the program complexity and the ingenuity of the programmer.

There is also a need to provide data input and data output routines which pick up or transmit data to and from the lines or line buffers. Such routines are usually simple if all data is in a common code and format. If various codes or formats must be provided for, the storage for these routines may become considerably larger. The use of indexing increments and the ability to handle partial words (i.e., characters) can greatly reduce the program storage requirements.

If an on-line system is to operate with a low down time, it will be necessary to provide an on-line maintenance program which provides self checking to quickly detect malfunctions. A set of off-line maintenance or diagnostic programs to pin-point the trouble spot can then be used to minimize the time to repair. To reduce the active storage requirements, these programs may be stored in an inactive file (tape, disc, or drum) and transferred to active storage only when needed. The on-line maintenance program can be transferred to active storage a block at a time and interleaved with the operational programs being run.

Temporary Data Storage

The amount of temporary data storage which is required must be tailored to the individual system. It is a function of data rates, block length, line utilization, and the amount of delay which can be tolerated in the system. The delay which can be tolerated also is the primary consideration in determining the type of storage media to be used.

The ideal storage media should allow for fast, random access, should be modularly expandable from 10,000 characters to 200,000 or more characters, and should be inexpensive. No such media exists; and we must therefore determine the best one for our system on a trade-off basis. A multiplicity of storage media must be considered. Some examples are:

1. Flip-Flops

   These are usually furnished as a per line buffer and because of the cost are limited to a few characters. The cost per character will range from $200 to $500.

2. Core Memory

   These may be furnished as a common storage unit or in smaller units on a per line basis. Depending on the size the cost will range from $1.00 to $4.00 per character. Sizes may range from 4,000 to 400,000 characters.

3. Magnetic Drum

   This would normally be used as a common storage unit with capacities of from 40,000 to 800,000 characters per drum. Prices range from $0.15 to $0.40 per character. Access times will generally range from 25 to 200 milliseconds.

4. Disc File

   This too would normally be used as a common storage unit with capacities of from 1 million to 10 million characters per disc. The cost will range from one cent to 5 cents per character. Access times will range from 50 to 200 milliseconds.

5. Magnetic Tape

   For temporary storage purposes this is normally used on a per line basis for high-speed inputs and outputs. It is expensive, ranging from $15,000 to $30,000. The storage capacity is about 1 million characters and the access time may range into minutes.

6. Paper Tape

   This too would be furnished on a per line basis and would be used for slow speed inputs and outputs. The price will range from $5,000 to $10,000 per line. The storage capacity will be about 100,000 characters and the access time immaterial since bin storage is normally used.

In most present day systems the use of per line paper and magnetic tape storage devices has been discontinued. We are then left with a system which may use flip-flops, cores, drums, disc files, and centralized magnetic tape units.

Overflow Storage

This type of storage is used as a back-up to the normal temporary data storage which is provided. It is needed during peak load periods when the input traffic exceeds the processing or the output capabilities of the system. It is also needed for badly skewed traffic distributions when a few outputs are badly overloaded.

Overflow storage should provide semi-random access so that specific items can be easily retrieved. It should provide one or two hours worth of storage which may amount to from 2 million to 4 million characters. It, of course, should be economical, and it must be cheaper than the normal data storage media.

The following media, in increasing order of cost, may be considered.

1. Magnetic Tape

   This is the cheapest form of overflow storage, particularly if tapes are used for permanent storage. Unfortunately, the poor access does not allow the ability to quickly retrieve a specific item. It is usable only for a brute force overflow approach.

2. Disc File

   This is probably the best choice for overflow storage. It is relatively cheap storage, has more than adequate volume and good access time. In order to reduce the number of accesses per unit time it may be necessary to utilize relatively large blocks--300 characters or more.

3. Drum

   The use of a drum for overflow allows the fastest access time of any of the media considered. It is also by far the most expensive. In order to provide the needed volume, several drums will be required. The advantage of using drums is that small blocks may be accumulated in the normal store for quick transfer to the drum. The size of the normal data store can therefore be reduced.

4. Elimination of Overflow Storage

   If a method to smooth out the high traffic peaks can be found, the need for overflow storage can be eliminated entirely. An easy solution is to refuse routine items from the originating subscribers or stations during these peaks. Normally each originator will have a local storage media. This storage may be a hard copy of a message, a perforated tape, a deck of cards, or a magnetic tape reel. In those few cases where storage is not available or where the originating item is of high priority, service will, of course, be granted. The refusal of such a routine item will not add any additional delay in handling the item since it would only be placed in overflow storage when picked up. The use of the existing subscriber storage media is therefore a cheap, practical method of solving the overflow storage problem.

Permanent Storage

Permanent storage of all information items passing through the center is usually required in an on-line system. In some systems (i.e., communications) it is required by law and a file must be maintained for six months. In other cases it is required to retransmit items which are later found to be in error and must be maintained in an active state for up to 48 hours.

The media used for this storage function must be capable of holding large volume. A single day's traffic may consist of 30 million characters or more. It must have a low cost per character in order to be economically feasible. In those systems which require dead storage files, it must be transportable and capable of being shelf-stored for long periods.

The most widely used media for this type of storage are listed below.

1. Hard Copy

   Printers may be obtained which will accept anything from 10 to 1000 characters per second. They range in price from $1000 to $30,000. This method will work for limited storage periods with manual search used to retrieve items. It is not recommended for long term storage unless it is backed up by a micro-film process.

2. Punched Cards

   The speed and price of this media is similar to hard copy. Retrieval of a specific item can be performed by semi-automatic means. Long term storage will be difficult due to the bulk unless a micro-film back-up is used. One disadvantage of cards is that, unless the items are short, each item will require more than one card.

3. Paper Tape

   Again the speed and price of this media is comparable to that of hard copy. Semi-automatic retrieval can be performed and there is no limitation on item length. The only problem is bulk storage since no size reduction can be obtained.

4. Magnetic Tape

   Magnetic tape units may be obtained with speeds from 15,000 to 100,000 characters per second. Prices will range upwards from $15,000. Retrieval of an item is simple and can be done automatically on those tapes still in position. Long term storage is reasonably bulky but not so bad as paper tape.

   A good choice for a combined overflow and permanent storage media would be a magnetic tape system. The cost is minimal, reliability is reasonably good, and the access time satisfactory if first in - first out overflow is not required.

Temporary Storage - Per Line

The use of a temporary storage buffer for each input and output line reduces the input-output load on the central data processor. The amount of storage required depends on the capability of the processor chosen and upon the processing needs of the system.

In order to get some quantitative figures of how the number of instructions and the costs vary with the length of the buffer storage, it is necessary to make some assumptions about the type of items handled by the system. It will be assumed that all items are 75 words or 300 characters in length. It will also be assumed that the data processor contains some sort of automatic interrupt or multi-sequence hardware so that it will require only one instruction to perform a data transfer.

1. Item Length Store Per Line

   It is entirely possible to store an entire item on a per line basis. This has been done by using an individual core memory per line or by time sharing a core memory over a group of lines. It is also necessary to use some sort of black box to convert the incoming bit stream to a parallel character or word for use in the memory and to perform the reverse function for outputs. If a simple word at a time transfer is used between the per line storage and the processor, it will require 75 instructions to transfer an item. If an automatic block transfer is implemented this number can be cut in half. The storage costs, if a shared memory is used, will range between $3,000 and $4,000 a line. This includes the cost of the interface black box.

2. Word Length Store Per Line

   The use of a single word buffer per input and output line reduces the cost somewhat over the item buffer method without greatly increasing the number of instructions required. The per input line serial to parallel converter and its opposite number per output line are still required. Since it will require one instruction to input each word, a total of 75 instructions are required per item. Storage costs including the converters will run about $2,500. Due to the small amount of storage per line, a flip-flop register is probably the best choice.

3. Character Length Store

   If only a single character is to be stored for each input and output, the use of an input serial to parallel converter and an output parallel to serial converter plus associated timing is all that is required. Flip-flop storage is normally used due to the small amount of storage. The average cost per line will be about $1,500, but it will require 300 instructions to input an item.

Bit Assembly

If the processor has a very high internal speed it is possible to completely dispense with per line storage and assemble the entire item into a central storage device. Only a level converter per input line, and a flip-flop plus a level converter per output line are required. If synchronous data transmission is used the equivalent of four instructions per character are required. This method then requires 1,200 instructions to transfer an item. If start-stop data transmission is used it will be necessary to repeatedly sample a bit in order to establish synchronization. This requires the time equivalent to 45 instructions per character or 13,500 instructions per item. The chief advantage of such an implementation is that the cost per line will only be about $100.

The choice on which way to go depends on the capability of the data processor used. If a slow or heavily loaded processor is utilized in the system, it will probably be necessary to provide at least a word of storage and perhaps an entire block. If an extremely fast processor is available, the use of bit assembly offers a considerable reduction in cost.

Temporary Storage - Central

The use of a word organized core memory as the centrally located temporary store will be assumed. This type of memory is the only device with the needed access time and storage volume which is economically feasible.

It is possible to permanently assign a block of storage within this memory to each input and each output line. Because of the possibility of a wait for processing, it is usually necessary to provide slightly more than a single block per line. The complications which are added to the control circuitry because of this partial block are so great, that it is usually easier to provide two full blocks per line. If an incoming item is longer than an assigned block, it must be assembled, block by block, either in an overflow storage device or in another portion of the core memory. When the item is completely assembled, it is processed and then transferred to the appropriate output blocks for outward transmission. With this approach, at least four storage blocks per duplexed line are required. In addition some overflow storage is required for long items.

If blocks are assigned by the processor on the basis of need, a substantial savings in overall memory size can be made. The number of storage blocks required will be a function of the length of time an item must be held in memory. This storage time (ts) is a function of:

Input Holding Time (ti)
Output Holding Time (to)
Output Queue Time (tq)
Processing Time (tp)
Wait for Processing Time (tw)

If one assumes infinite sources a simple calculation can be made for output queue time and processor waiting time. Actually, if more than 25 inputs are involved, only a small error is introduced by assuming infinite sources.

If an exponential distribution of item lengths is assumed, then the output queue time (tq) is given by the following expression:

$$tq = to \cdot \mu o/1 - \mu o$$

Where μo is the output line usage factor.

If a constant distribution of processing times is assumed then the wait for processing time (tw) is given by the following expression:

$$tw = to/2 \cdot \mu p/1 - \mu p$$

Where μp is the processor usage factor.

For present day processors both the processing and the wait for processing time are small compared to the input and output holding times. The main factor in determining the storage requirements is the output queue time which varies with the output line usage. The following table shows how (tq) varies with (μo):

μo	tq
0.2	0.25 to
0.4	0.67 to
0.6	1.5 to
0.8	4.0 to
0.9	9.0 to
0.95	18.0 to

For acceptable delays in any real time system, μo should be less than 0.8 and preferably less than 0.6.

In order to get some numerical values for the amount of storage, two sample calculations will be made. One will consider a system of teletype speed lines and the other a system of data speed lines.

Teletype Example

Assume

1. 300 character blocks
2. Line rate of 10 characters per second
3. μo = 0.6
4. Processing = 2000 instructions
5. Instructions = 20 μsec long
6. μp = 0.95

Then

ti = to = 30 sec
tq = 45 sec
tp = 40 millisec
tw = 360 millisec

The storage time (ts) then equals 105.4 sec.

The number of characters of storage required (Cs) is given by:

$$Cs = ts \cdot \mu o \cdot (\text{line rate})$$

For this example Cs is equal to 635 characters per duplexed line or 2.12 blocks per line. This compares favorably with the 4 blocks per duplexed line required using the permanently allocated method.

Data Example

The same assumptions are made except that a line rate of 300 characters per second is assumed.

The calculations then lead to the following:

ti = to = 1 sec
tq = 1.5 sec
tp = 40 millisec
tw = 360 millisec

The storage time (ts) then equals 3.9 seconds, and the characters of storage required (Cs) equals 700 characters. This is the equivalent of 2.33 storage blocks per duplexed line which again compares favorably to the four blocks per line obtained using permanently allocated blocks.

The method of assigning blocks offers several other advantages. The item remains in the same storage block during the input, processing, and output functions. No internal data transfers are required. If an item is longer than a block length, then several blocks may be strung together to hold this item. Overflow is never required unless the utilization of the output lines is greater than the 0.6 figure assumed.

Conclusions

If a per line storage device is used and this is coupled with permanently assigned central storage plus overflow, the data flow of Figure 1 results. Inherent in this system are 2 serial data transfers and 6 parallel data transfers. Each transfer requires time, equipment, and adds to the possibility of error.

This can be contrasted to the system shown in Figure 2 in which only 2 serial transfers are normally required with the addition of 2 parallel transfers during peak load periods. It is to be hoped that future systems will follow this trend and utilize the high internal speeds currently available to reduce interface hardware and storage requirements. For example, a system utilizing bit assembly and subscriber overflow storage would require only a main frame core memory plus some permanent storage memory. This system would certainly be minimal in terms of hardware and cost and would still perform the required functions of an on-line system.

CONVENTIONAL DATA FLOW

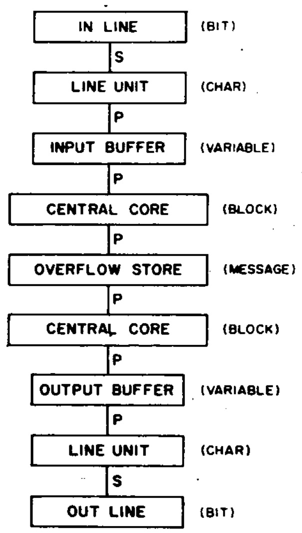

Figure 1.

IDEAL DATA FLOW

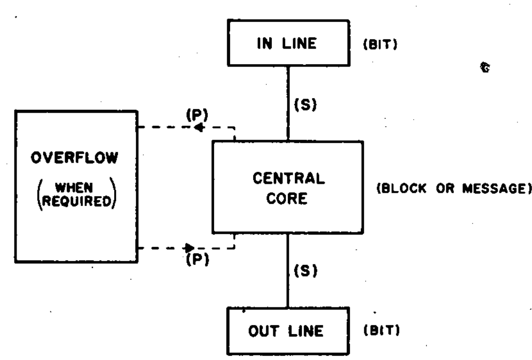

Figure 2.

ON THE DEVELOPMENT OF A SUPERVISORY SEQUENCING ROUTINE

BY

MALCOLM H. GOTTERER

Computers, Supervisory System, Sequencing Routine, Scheduling, Linear Programming, Transportation Method, Model, Lateness, Priorities, Customer Behavior

ON THE DEVELOPMENT OF A SUPERVISORY SEQUENCING ROUTINE

Malcolm H. Gotterer

SUMMARY

This paper describes a self-scheduling routine which can be incorporated into the supervisory system of a computer. The problem of scheduling a number of tasks on an overloaded computer can be formulated as a linear programming problem and then solved by means of the transportation method. Priorities for both jobs and customers, as well as a non-linear loss function for late completion, form the basis of the model.

INTRODUCTION

The modern large scale computer system requires substantial investments of not only funds but also of manpower. As such their effective use is of major concern to both those who are responsible for the operation of the computer as well as those who depend on it to provide them service. The final efficiency of the installation must, therefore, be judged by the dual standards of service to uses as well as the operating effectiveness of the computer itself.

One of the problems faced in the operation of a computer center is that of scheduling individual jobs on the machine. The problem is obvious in the case where the total demands for service exceeds the available processing time. However, it should not be ignored in cases where the servicing time exceeds total demands because the timing of these demands may result in overcapacity in some periods and overloads in others. In either case it is essential to develop and adopt some set of decision rules for the orderly sequencing of jobs on an overloaded computer to provide for its most effective use. In this paper a model for accomplishing this objective is developed and the method by which it can be made an integral part of the computer system is discussed. 1

THE SEQUENCING PROBLEM

The sequencing problem discussed in this paper is that of processing a relatively large number of jobs each of different, and unpredictable, length and originating from different sources. This type of problem is frequently encountered in the commercial computer service bureau, the university computer center, or a centralized computer within a firm, see Fig. 1. This demand pattern precludes the possibility of using a "prediction" or "one-ahead" method of scheduling. Prediction type scheduling can best be used when the workload is stable and can be forecasted. The one-ahead procedures are effective when the load is partially predictable. But when there is no basis for establishing in advance the nature of customer needs at any given time in the future sequencing must be in terms of actual orders, or demand scheduling.

One of the simplest methods of demand scheduling is that of first-come-first-served. Under this rule, each job is placed at the end of the line of all jobs awaiting processing. Then, as the preceding jobs are completed, it moves up in the queue until it is processed and returned to the customer. This procedure can be effective if the center is dealing with a set of customers who are of equal importance to the center and to the parent organization. A further requirement is that the jobs requiring processing be of equal importance.

It is frequently found that the conditions for the success of the first-come-first-served system of scheduling are not met. Some customers enjoy higher status because of their organizational position, the nature of the work they are doing, or some other reason. Similarly, not all jobs are of equal importance either. Combining these two

factors may lead to a situation in which a person with high status has a job which can be delayed while a person with low status presents a job that requires early processing. Under these conditions a logical modification to the first-come-first-served rule is to assign priorities to customers or projects, and then to establish two or more queues, each representing a different level of priority. Jobs in the highest priority queue are then completed before those in the queue with the next lower priority are started, each queue in turn being exhausted before beginning on the next. This method of scheduling can, however, create some major problems and these in turn may lead to considerable dissatisfaction on the part of the users.

One source of such dissatisfaction arises when a job with a high priority takes a long time, perhaps two weeks, to be processed, while a number of low priority jobs, requiring only a few minutes of computer time to complete, must wait. Another source of difficulty arises from the actual timing of job arrivals. Take, for example, a job with low priority has waited some time for all higher priority projects to be completed. Just as it is about to be processed itself, suppose that yet another high priority demand is made. Under the modified first-come-first-served decision rules, the new task is processed immediately and the waiting job is delayed still further. It must, therefore, be recognized that jobs frequently have a time value as well as customer and job priorities. Therefore, it may be desirable for the newer high priority jobs not to receive processing immediately. Instead it may be to the benefit of the organization if some, or all, of the older low priority jobs are processed first. Since the first-come-first-served scheduling procedure with priority queues does not take the time value of a demand into account, we must conclude that it is not satisfactory for scheduling a computer.

It is possible to further elaborate the decision rules of a first-come-first-served scheduling procedure with priority queues to take into account the length of time the job has been waiting. But as such elaboration takes place the system becomes more difficult to administer. Therefore, another approach to the sequencing problem must be sought. The general model described below seeks to minimize the total loss, or "cost," to the computer center of failure to complete all jobs by the time requested by the customers. This model can be incorporated into the executive, or monitor, system on the computer itself and thus provide a fully automatic self-scheduling system.

THE EFFECT OF LATENESS

Whenever a job is not completed by the time requested by the customer, we can assume that the customer will become dissatisfied. Short periods of delay should result in little discontent but as the delay increases, we can expect the dissatisfaction to increase rapidly. This dissatisfaction can then be approximated by a quadratic loss function, Fig. 2, having the form:

$$L_i^k = (a_i - d_i)^{m^k} \; ; \; a_i \geq d_i \tag{1}$$

where L_i^k = the loss for completing job i, for customer k, in time period a,

a_i = actual time of completion of job i,

d_i = deadline, or time requested by the customer for job i,

m^k = loss exponent for customer k

In some instances the early completion of a job may result in increase satisfaction to the customer. This then is the equivalent of a negative loss, or gain in good will to the computer center. Any such gain will, however, be bounded since the customer has himself specified the deadline for the completion of the job. This negative loss function is then:

$$L_i = -(d_i - a_i)^{m^k} \; ; \text{ where } d_i > a_i \tag{2}$$

The objective of the scheduling system would then be:

$$\min \sum_{i,k}^{n} L_i^k \tag{3}$$

PRIORITIES

Priorities provide a means by which it is possible to "weight" the relative importance of customers and jobs. As pointed out earlier, scheduling by priority classifications alone will not be satisfactory. Nevertheless, priorities do play an important role within the scheduling problem, and should be incorporated into the model.

Customer Priorities

It is generally true that not all customers are of equal importance to the organization. Certain customers may be assigned a higher priority than others by considering the status of the individual, organizationally or professionally. It then becomes possible to array the importance of the customers such that:

$$c^1 \geq c^2 \geq c^3 \geq \ldots\ldots \geq c^k \geq \ldots\ldots \geq c^n \tag{4}$$

where

c^k = an index of the relative importance of customer k

The relative values assigned to c are called the <u>customer priority</u>. Since some people may expect to receive preferential treatment, and the organization may want them to be treated in this manner, the effect of customer priorities is to increase the loss for the late delivery of a job. The loss function in Equation (1) is modified:

$$L_i^k = (a_i - d_i)^{m^k} + c^k \tag{5}$$

Some customer may be of such low status, or may be working in an area which is of such little importance to the organization, that it might be desirable to provide service to them only after all others have been served. As shown earlier, however, delaying such jobs until all others have been processed may produce too much dissatisfaction. This can be reflected in the model by assigning a negative customer priority, $-c^k$, to such persons. This has the effect of delaying the completion of these jobs.

Job Priorities

It can also be argued that not every job of each customer is of equal importance. Some jobs will be more important than others, independent of the priority status of the customer. Thus a high priority customer can have both important (high priority) and unimportant (low priority) jobs; similarly for a customer of low priority. The jobs of

any one customer can then be ranked as follows:

$$j_1^k \geq j_2^k \geq j_3^k \geq \ldots\ldots \geq j_i^k \geq \ldots\ldots \geq j_n^k \tag{6}$$

where

j_i^k = the relative importance of job i to customer k

Job priorities can be incorporated into the model by adjustments in the deadline for completing (d_i). A high priority job can be assigned a deadline (d_i') which is earlier than the time requested by the customer himself, which would have been d_i. The loss function then becomes:

$$L_i^k = (a_i - d_i')^{m^k} + c^k \tag{7}$$

It should be noted that, since $d_i' < d_i$, then $(a_i - d_i')^{m^k} > (a_i - d_i)^{m^k}$ or $L_i^{k'} > L_i^k$. For very unimportant jobs we may assign d_i' such that for any a_i, there exists a d_i' such that $d_i' \to a_i$ and therefore $L_i^k \to 0$. This is then the equivalent of a "stand'by" job, i.e., one which is to run only if the computer would otherwise be idle.

CUSTOMER BEHAVIOR

Certain rules and regulations are necessary in the operation of a computer center. These frequently result from the nature of the executive routines that control the system as well as from economic and managerial considerations. Deviations from these established norms may result in lost computer time or increased costs of operation.

Computer customers should therefore be encouraged to abide by all exisiting regulations, and should be penalized for any failure to do so. One method of doing this is to adjust the loss exponent, m, of the general loss function. If the value of m is decreased for continued violation of the established rules, the loss resulting from late completion of a customer's job will be reduced. Likewise, as m increases, indicating compliance with the regulations the loss from not completing a job by the time it was requested will be increased. Expressed in terms of the model, this means that if $m_i^{k'} < m_i^k$ then $(a_i - d_i)^{m^{k'}} < (a_i - d_i)^{m^k}$, resulting in $L_i^{k'} < L_i^k$.

The specific factors influencing the evaluation of the loss exponent m for any customer will vary considerably from one computer center to another. The customer's behavior relative to the norm will have to be measured and standards developed. In general form this will be:

$$m = f(r_1, r_2, \ldots, r_i, \ldots, r_n) \tag{8}$$

where

r_i = a measure of the customer's deviation from the norm on rule i.

Some of the factors that might typically be included in the calculation of m are:

1. Poor estimation of running time.

2. Errors in programming.

3. The use of inefficient input or output mechanisms.

4. Taking excessive dumps.

An integral part of any operating system will necessarily be a means to calculate the loss exponent, m. It should be noted that in the case of early completion of a job, i.e. $d_i < a_i$, the loss exponent must be less than zero so as to provide the bounded condition on the negative loss discussed earlier. Therefore there will be two values for the loss exponent m and m′, m > 0 for late deliveries and m′ < 0 for those instances in which the job is completed before the time specified by the customer.

In the calculation of the loss exponents the frequency and timing of the customer's deviation from the center's norms will be important. The user who deviates once, but then complies with the rules, should not be given as low a value of m as the person who continuously engages in poor work habits. Therefore a device for weighting recent experiences more heavily or for calculating trends in experiences of the customer will be desirable. This will permit the loss exponents to change as the behavior of the customer changes. One means of obtaining this weighting is, of course, by use of an exponential weighting formulation such as:

$$V_i^n = L\, R_i^n + (1-L)\,(V_i^{n-1}) \tag{9}$$

where:

V_i^n = customers deviation factor for rule i in period n

L = weighting factor

R_i^n = customers actual observed deviation from rule i in period N

The final behavior factor for each customer can then be a weighted average of each individual V_i. Or:

$$m = \sum_i b_i V_i \tag{10}$$

where:

$$\sum_i b_i = 1 \tag{11}$$

THE GENERAL MODEL

This scheduling model can be formulated as a transportation problem, somewhat along the lines suggested by Danzig, (3) and can then be solved by any one of the several established algorithms (2) (4). The objective of the final schedule is to minimize the total cost of lateness. The array would be:

Job Number	Time Period 1	2	....	j	....	n	Total time Periods Required
1	X_{11}	X_{12}		X_{1j}		X_{1n}	t_1
2	X_{21}	X_{22}		X_{2j}		X_{2n}	t_2
. . .							
i	X_{i1}	X_{i2}		X_{ij}		X_{in}	t_i
. . .							
m	X_{m1}	X_{m2}		X_{mj}		X_{mn}	t_m
	1	1		1		1	

$$\sum_j X_{ij} = t_i \tag{11}$$

$$\sum_i X_{ij} \leq 1 \tag{12}$$

$$n = \sum_i t_i \tag{13}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} L_{ij} = \text{minimum} \tag{14}$$

where:

X_{ij} = time units assigned to job i in period j.

t_i = time units required for job i.

L_{ij} = loss associated with job i being processed at time j.

n = number of time units being considered.

It will be noted that the entries in the array will be either 0 or 1. A one indicating that that particular job has been scheduled to be run at that time, while a zero indicates it has not been so scheduled. The effect of Equation (11) is to insure that the proper number of time units has been assigned to each job. Equation (12) insures that only one job has been scheduled for any time unit. All time on the computer is allocated to specific jobs (although there may be idleness) by means of Equation (13). Finally, the objective is represented by Equation (14).

The cost, or loss, tableau is:

Job Number	time					
	1	2	.....	j	....	m
1	L_{11}	L_{12}	. . .	L_{1j}	. . . .	L_{1n}
2	L_{21}	L_{22}		L_{2j}		L_{2n}
⋮	⋮	⋮		⋮		⋮
i	L_{i1}	L_{i2}	. .	L_{ij}	. . . . .	L_{in}
⋮	⋮	⋮		⋮		⋮
m	L_{m1}	L_{m2}		L_{mj}		L_{mn}

Since the transportation model will assign time for the performance of each job, a loss function must be assigned for each period of processing. This means then that the total loss for a job will be dependent, not on the final time of delivery, but rather on the time intervals during which it is processed. We then have to adjust the loss function so that

$$L_i = \sum_j L_{ij} \qquad (15)$$

$$L = \begin{cases} (p_i - d_i)^{m^k} + c \; ; \text{ if } p_i > d_i \\ -(d_i - p_i)^{m^{k'}} + c_i ; \text{ if } d_i > p_i \end{cases} \qquad (16)$$

AUTOMATING THE MODEL

An essential feature of the model developed above is that it can be automated, 5, and made an integral part of the computer supervisory system. The function of the supervisor is to provide for the progress of problems through the computer with the minimum possible delay. Therefore the supervisor would be made up of the scheduling subsection, an accounting subsection, as well as the overall control routines, 6.

This discussion will deal exclusively with the application of the model to multi-computer installations, Fig. 3, 4. This type of computer center is characterized by one, or more, smaller computers working in conjunction with a larger unit. The smaller processors serving as input and output devices for the larger computer as well as carrying out other computations as directed either in conjunction with the other computer or independently. The I/O computer will be referred to as the scheduler and the larger main processer being scheduled will be called the computer.

The sequencing routine, Fig. 5, uses a combination of the scheduler and computer to accomplish its objectives. When a customer presents a job it is read into the scheduler. The customer's file, stored permanently in the scheduler, is then interrogated to determine the customer and job priority as well as the current value of the loss exponent. Part of the job input data would be, of course, the expected running time, deadline, and other such information required by the scheduler. The loss to be incurred for late delivery can then be calculated by time periods. The schedule matrix is then adjusted by adding the new job as a row and as many new columns as are necessary to provide enough time periods to process all the jobs currently waiting for service. Using the transportation algorithm the optimal schedule is obtained. The sequence in which the jobs are to be processed together with estimates of when each will be started and completed can now be printed out for use by the administrators of the computer center.

The next task of the scheduler is to prepare the jobs for use by the computer. If magnetic tape units are used this can be accomplished by either one of two methods. First, a new input tape can be prepared with the jobs in proper sequence. Second, a table can be developed listing the jobs in sequence together with their storage location, at the same time prepositioning the tape so that it is ready to supply the computer with the proper file when requested. The second method is, of course, superior because it eliminates the need to rewrite input tapes. But even when using the second method it may be necessary to move from one end of a tape to the other end to locate the proper file, a time consuming task. If a random-access device is available this problem is greatly simplified. Once again a table containing the jobs in order of processing and their location is required. But the time consuming task of rewriting or pre-positioning tapes can be eliminated. Instead it is possible, of course, to go directly to the location of the next job in the random-access unit. For this reason a random-access device provides faster service to the computer and reduces the possibility of delaying the activities of the computer.

Obviously this scheduling routine requires processing time which is of value even on a smaller computer. The need for this expenditure has been noted by Newell, "The wastage from a half-second human tending a microsecond machine is forcing the development of supervisory routines and interrupt systems that will eventually eliminate the operator." 7. We feel, therefore, that the expenditure of this time must be made in order to assure the overall effectiveness of the utilization of the system. A more pertinent question is to what extent will the scheduling routine itself delay the operation of the computer.

The problem of conflict between the scheduler and the computer is of significance for two reasons. First, the purpose of the scheduler is to facilitate the operation of this computer. If it instead interferes there is reason to doubt its effectiveness. Second, the value of time on a modern large scale computer is extremely high when considered from either the point of view of direct cost of operation or opportunity costs. Therefore interference must be minimuzed. Fortunately, the problem can be eliminated by including suitable controls in the scheduler.

Ordinarily it is expected that the scheduler will complete its operation while the computer is processing a job, Case A, Fig. 6. If the computer completes its current job before the scheduler has arranged the new input then the next job as originally scheduled is read in by the computer and a new schedule is calculated, Case B, Fig. 6. It is only in Case C, Fig. 6, that a conflict is encountered. In this case the computer will attempt to read in the next job while the scheduler is preparing the input tape or table. By preserving the original schedule Case C can be converted to a modified form of Case B. When the computer calls for the next job it would get it, as before, from the original schedule. The scheduler would then stop preparing the new schedule and recalculate it as in Case B. This is Case C′, Fig. 6. Therefore the problem of interference between the scheduler and the computer can be eliminated.

CONCLUSION

The scheduling system suggested in this paper would appear to offer a number of advantages in the actual operation of a computer center. Since the scheduling would be done automatically anyone dissatisfied with the service they have been receiving can attribute it to either the lack of adequate facilities, their personal priorities or their adherence to the center's rules. Therefore any discontent on this score would become an issue to be settled by someone much higher than scheduling clerks or computer operating personnel. As was seen, however, even those with high priorities will not necessarily get their jobs processed immediately; instead there will be a more equitable use of the computer among those who need its services. The system will require the investment of computer time but it is felt that the result will be to provide a more efficient total operation of the computer center. Further development of this model and operating system will eventually result in the elimination of computer operating personnel.

ACKNOWLEDGEMENT

This project was sponsored by the Computer Center at the University of California, Berkeley. The author is indebted to Dr. Austin Hoggatt, Director, for his helpful comments and encouragement.

REFERENCES

1. Automatic Machine Scheduling, J. C. Hammerton. Computers and Automation, vol. 10, May 1961, pp. 17-22.

2. The Stepping Stone Method of Explaining Linear Programming Calculations in Transportation Problems, A. Charnes, W. W. Cooper. Management Science, vol. 1, October 1954, pp. 49-69.

3. Recent Advances in Linear Programming, G. Danzig. Management Science, vol. 2, January 1956, pp. 131-144.

4. Application of the Simplex Method to a Transportation Problem, G. Danzig. Activity Analysis of Production and Allocation (book, L. C. Koopmans, editor. John Wiley and Sons, Inc., New York, N.Y., 1951, pp. 359-73.

5. A High Speed Computer Technique for the Transportation Problem, J. B. Dennis. Journal of the Association for Computing Machinery, vol. 5, April 1958, pp. 212-221.

6. The Atlas Supervisor, T. Kilbourn, R. B. Payne, D. J. Howarth. Computers Key to Total Systems Control, Proceedings of the Eastern Joint Computer Conference, December 1961, Vol. 20 (book). The Macmillan Company, New York, N.Y., 1961, pp. 279-94.

7. On Programming a Highly Parallel Machine to be An Intelligent Technician, A. Newell. Proceedings of the Western Joint Computer Conference, May 3-5, 1960, Vol. 17 (book). Western Joint Computer Conference, San Francisco, California, 1960, pp. 267-282.

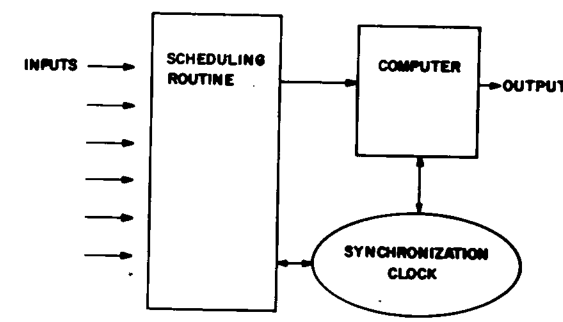

Fig. 1. The Scheduling Problem

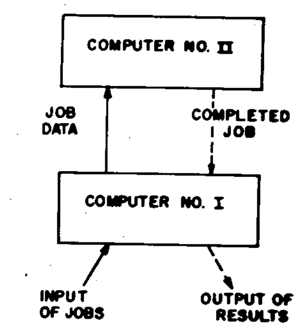

Fig. 3. Multi-Computer Configuration

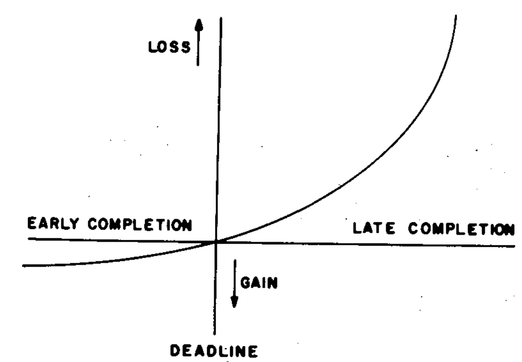

Fig. 2. Typical Loss Function

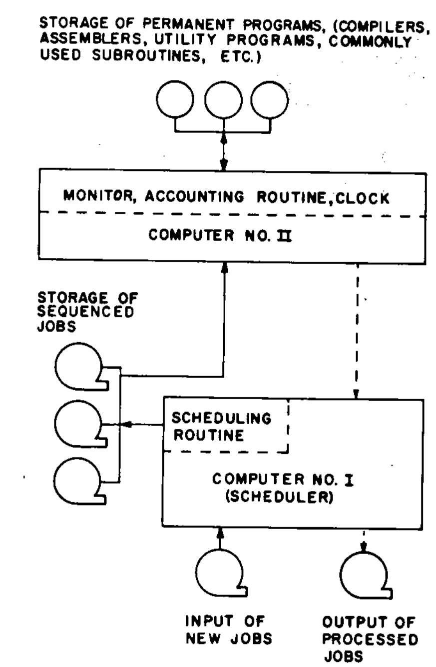

Fig. 4. Job Flow in Multi-Computer System

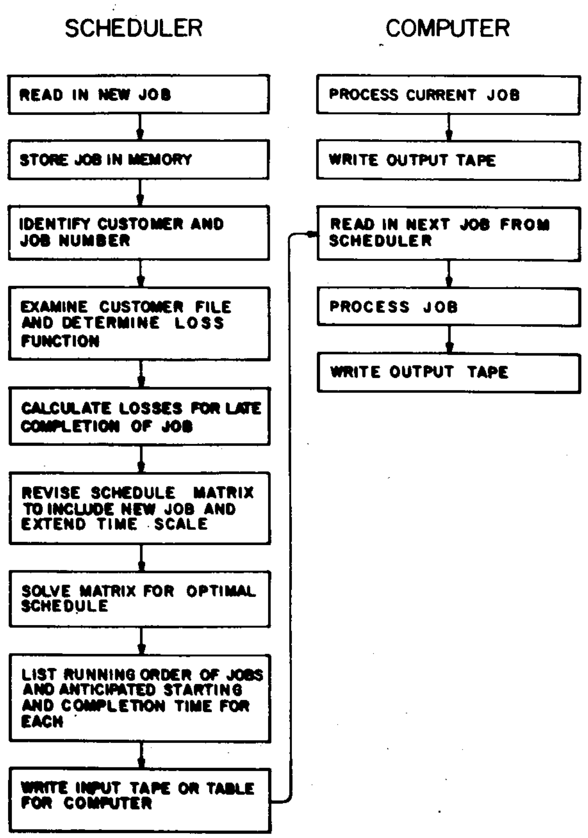

Fig. 5. Flow Diagram of the Sequencing Routine

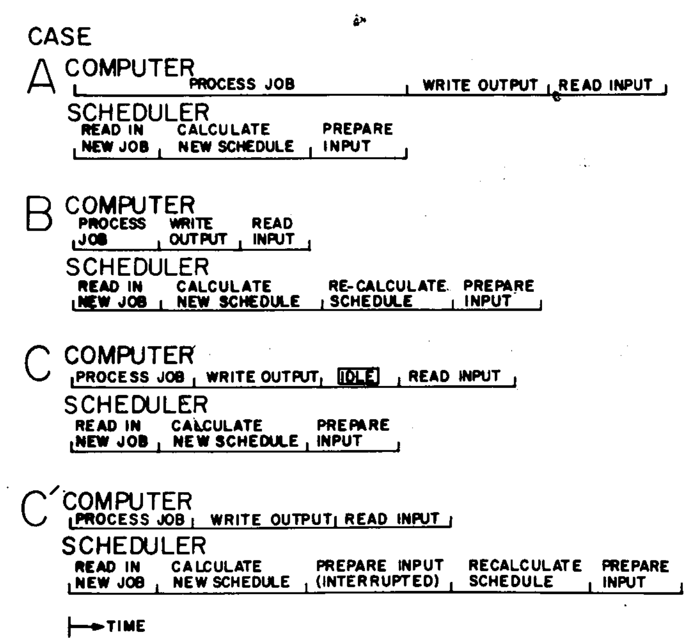

Fig. 6. Gantt Chart of Relations between the Scheduler and the Computer

COMPUTER POWER AS A PUBLIC UTILITY - AN EVALUATION

BY

R. S. STEIN

Summary

Computer power utilities have been discussed for at least five years as the next logical step in electronic data processing. Recent developments in computer hardware and communications facilities indicate that implementation of the idea may now be feasible. In this paper, specific aspects of computer power utilities are evaluated. These aspects include communications requirements, engineering, pricing, applications, and sources of supply. An analogy is drawn between the possible development of computer power utilities and the history of electric power utilities.

Computers, Computer Power, Public Utility, Data Processing, Communications, Engineering, Pricing, Sources of Supply, Time-Sharing, Computer Service Bureaus, Programming

COMPUTER POWER AS A PUBLIC UTILITY - AN EVALUATION

R. S. Stein
Associate Member AIEE
UNIVAC Division, Sperry Rand Corp.
St. Paul, Minn.

For several years, a recurring question has been asked by thoughtful men in the computer industry: "Can electronic data processing be made available to mass consumers as a public utility?" Who first conceived this idea is not recorded; it may have come to several minds at much the same time for the concept is reasonable and timely in the light of evolutionary trends in other public services. In particular, an analogy has been seen between the computer industry and the electric power industry. Electric power was first supplied to a small number of customers in an isolated, localized area, and the cost was necessarily high. Only after years of development and expansion did the service of supplying the generation and distribution of electric power to customers become a public utility, with costs lowered to a level that mass consumers could pay. Does electronic data processing possess the same potentialities to the extent of becoming a "computer power" utility? Developments in data processing techniques and in both communications and switching facilities would appear to make an affirmative answer possible at the present time.

The phrase "computer power" is a new one, but it is apt. Computing systems have a power that is uniquely their own. They have the power to exercise control, originate new transactions, organize conceptions, regulate principles, make decisions, and store and retrieve all manner of information. But can computer power become a public utility?

A short time ago, the public utility approach to widespread computer usage could be discussed only theoretically because of the state of hardware development and the high cost of handling large quantities of data. Now the situation has changed. In hardware, the requirements for computers with both real-time capabilities and large mass storage capacities are being met. Low cost media for transmitting data have become available, and methods are being developed for the efficient use of high-volume low-cost communications facilities that have versatile, on-line, remote, peripheral devices needed for the customer's use. Automatic switching of communications facilities have made this on-line use possible. Low-cost, high-speed transmission channels can now bring the benefits of electronic

data processing techniques to every individual, business or agency.

Alan O. Mann* has outlined points of similarity between existent public utilities and the proposed computer power utility. In his analysis, the following points are particularly significant:

1. Low rate of capital turnover.
2. Large proportion of relatively fixed costs.
3. Variable demands for service throughout the day, week and month.
4. Low utilization relative to rated capacity.
5. Need for reserve capacity.
6. Lower unit costs for larger equipments.
7. Large capital outlay that limits competition.
8. Similarity of early efforts in utility field to SPAN and service bureaus.
9. Ability to save through jointly produced services.
10. Business already large part rental.
11. Continuing trend toward simplifying closing the loop for service data acquisition.
12. Historical similarities in mergers and acquisitions.

Time-Sharing in Computing Systems and Public Utilities

A "utility", according to Webster's New Collegiate Dictionary, is a "power to satisfy human needs...the greatest good of the greatest number." By this definition, computing systems are already a form of "utility" within themselves. They share time among their units and their power is used to satisfy the greatest

*Mann, Alan O. "Computer Power: A Public Utility?" Computers and Automation, April, 1959, vol. 8 - No. 4, pp. 11-16.

good of the greatest number of units. Operating time is shared by the peripheral units used in a computing application: punched card, magnetic tape, printers, plotters, et al. Because some of the peripheral units are slower than others, buffers are included in the fast ones to store data until output from the slow units can catch up and get their data into the central processor. Furthermore, sequential signals in the central processor can initiate simultaneous operations in various peripheral units. Because a program is set up at the outset, the processing of data and all operations of the central processor and peripheral equipment are accomplished smoothly, swiftly, and accurately. The shared-time concept that is characteristic of today's computing systems is a forerunner of the shared-time concept for the greatest good of the greatest number of customers that would characterize a computer power public utility.

Immediate Problems

A major premise may be agreed upon: that it is possible to supply computer power to anyone who has need for 1) computations; 2) data manipulation such as sorting, merging, collating, or filing; 3) reports and their correlated visual aids such as graphs or charts; and 4) all manner of decision making in management. A minor premise is now possible that communication facilities exist to bring this computer power from a central processing station directly to the user whether he is in an office, factory, or home - providing, of course, that he has access to the high-speed communications lines of common carriers and the communications switching terminals that are necessary.

The logical conclusion from the above stated premises should be that all potentialities now exist for the inception of a computer power utility of wide usefulness. But does availability mean that the public is ready today to absorb computing power as a common utility, and that the computer industry is ready to change its entrenched practices? What are the factors to be considered in determining the support which such an enterprise could expect? What sectors of our economy would find this service most useful? How widespread would its acceptance be? What obstacles would have to be overcome? What are the advantages and disadvantages for the user? What are the benefits accruing to computer manufacturers and computer service bureaus from provision of such a service? What new resources would be required? What organized groups would engage in this business? These are questions that must be answered before plans can be made for so radical a change in industrial practice.

Computer Service Bureaus

One way to analyze the need and possible acceptance for a computer power utility would be to look at the growth of computer service bureaus. The expansion of computing centers has been spectacular; their present annual sales volume is estimated at over $50 million. Centers run by computer manufacturers and by groups that rent or buy their equipment have grown in about equal proportion. They have been most successful in offering data processing and programming services to a large number of customers, from both small and large firms. As reported by a news agency*, "the jobs being handled range from the dullest routine...keeping track of inventory, issuing paychecks - to bizarre problem-solving situations such as finding the best formula for blending perfume or analyzing Rorschach ink blots for psychologists."

The success of computer service bureaus has led to the formation of cooperative computing services within specialized groups, such as banks and stock brokerage houses. For example, the First National Bank of Chicago has set up an electronic data processing system for its correspondent banks, and the Mellon Bank of Pittsburgh plans to provide electronic data processing services for the 400 banks in its area.** Reports indicate that more and more small banks are pooling their resources to provide cooperative service bureaus*** both to serve their own needs and the accounting needs of customers for such applications as payroll preparation. Stock brokerage houses in increasing numbers are also linking together their branches by electronic data processing services. For example, several San Francisco stock brokerage firms are linked to a network of computers and are thereby able to obtain stock prices and other securities data almost immediately after the transactions occur in Wall Street. New York and Philadelphia brokerage houses have set up a network that now links or is planned to link more than 12 major cities over permanently hooked-up dataphone lines which carry data to stations from a master computing system. This latter network is so designed as to make possible an eventual network that will

---

*Wall Street Journal, Sept. 5, 1961, vol. XLI No. 226, "Computers for Hire."

**Data Processing Digest, March, 1962, p. 15.

***Data Processing Digest, April, 1962, "So You Want to Automate," by John J. Feldman.

embrace every financial center in the United States and Canada, with a possible extension to European capitals.*

The previously-cited reports and an increasing number of similar reports indicate an evolutionary trend toward combining and consolidating data processing facilities. This trend might take one of several alternative paths toward culmination but one end might well be regional computer power utilities. The means are at hand. At least one computer data processing service firm that handles general rather than specialized accounts is already probing the use of microwave or telephone links to connect an international electronic data processing service organization with centers in the United States and European capitals.**

The growth of service centers is noteworthy but so is the direction in which they are growing and the type of business they are seeking. Initially computer service bureaus provided one-shot, customized service for large users. Repeat business grew, and as programs were developed and center personnel acquired reputations in certain areas, customers brought in problems which could take advantage of past programs, thus reducing program preparation charges. Now, many centers are aggressively seeking to sell standard packages of computer-based service. Selling standard packages to many small users is a natural extension of present computer center operations, but the practice does not and cannot go far enough to cover all potential users of computer power.

Although present-day computing centers are a notable step forward in the evolution of computer power, the increasing data processing demands of the public and the high mortality rate among the non-manufacturing centers appear to indicate that computer bureaus are not the final answer to public needs. The use of centers is limited because costs, although lower than those for individual computer installations in a single company, are still higher than the vast majority of small firms can pay. Moreover, service bureaus presently lack adequate communication facilities between the center and the user; they cannot offer in a single center the variety of equipment needed to handle the widest possible spectrum of applications; and each center or

*EDP Weekly, June 1, 1962, pp. 6-7.

**"How An Industry Grows Up," an account of developmental aspects of C-E-I-R (Corporation for Economic and Industrial Research), Electronic News, July 24, 1961.

group of centers must have heavy initial capitalization to carry it while it awaits accounts.

Implementation of the computer utility concept would solve many of the problems now unsolvable in service bureaus. Furthermore, it would make computer power readily available to all potential users.

The Small Computer

Several manufacturing firms have been attempting to widen the computer market by offering small, low-priced machines. At present, no one has been able to produce a general-purpose digital computer that can be rented much under $750 a month. Even if the computer industry could offer a machine for $500 a month, the cost would still be beyond the financial means of the large majority of small business firms and far beyond the means of the average individual who now must spend tedious hours in the manipulation of data or in seeking answers to his problems. And even if it were financially acceptable, a small computer is not capable of providing adequate service for the many applications that all individuals or small business men and scientific researchers might require. The small customer, too, can rarely afford the cost of programming operations peculiar to his individual business, and must perforce limit his computer usage to standardized applications.

Implementation of the Computer Utility Concept

Under the computer utility concept, even the smallest requirements for data processing may be met, and program costs would be shared by many. The customer would have direct access to large-scale systems and to his own data storage without intervening paper work or tabulating. How would the utility operate? A communications device would be placed in the customer's home or office and tied to the computing system via common carrier facilities such as telegraph or telephone lines. The communication devices would have storage facilities so that communicating with the computing system might take place either at preset intervals or at the user's convenience. Communication at preset intervals would minimize line costs and computer interrupts in the lowest class of service. Use at convenience or for real-time applications would be offered in the higher classes of service.

Thus, through a versatile, remote input-output device, the customer would have access to whatever amount of computer power and peripheral devices his particular application might require -

all at a cost considerably lower than could be expected for any comparable, installed computer system.

Privacy of Data

In early stages of development for a computer power utility, each localized inquiry-response device would carry its own identification code so that requests for information could be verified in accordance with customer specifications. Thus data could be segregated to different remote devices at a customer's location as well as between customers. This approach is similar to the methods used to separate departmental information within a single-company installation. It is also similar to the methods used by banks which band together to process their data but segregate the transactions of the individual banks.

Applications

The selection of applications admittedly presents a problem. Those applications which are easiest to standardize would be most beneficial from the standpoint of the computer utility. This is especially true during the initial stages of the utility's development because many users must be added without incurring high setup and programming costs. Many subscribers would desire to have those applications handled which require real-time processing or could benefit from the joining of remote points. Are the goals of a computer power utility and user incompatible? Probably not. Standard package programs have already been developed in various application or industry areas as diversified as simulation, scheduling, and report writing.

Billing, customer, and revenue accounting are among the most widespread applications because they are necessary facets of all business that is not cash and carry, whether the primary function of the business is trade, manufacturing, or service. For example, consider a drug manufacturer with a single production facility selling directly to wholesale and retail outlets from 50 sales offices and supplying their needs from five warehouses throughout the country. Using the computer utility service as a nerve center, with an inquiry-response device at each of the various locations and at the home office, this user can streamline the bulk of his order-shipping-billing operations to offer the customer shipping within hours instead of days in addition to having all the required paper work rapidly produced at the point where it is needed.

A rejoinder may be made that so large a company as the one just cited can use a real-time computer system on its own premises.

It could. But how about the same type of concern with two offices and a single warehouse? Small firms face many of the same data processing and communication problems as do the large firms but without the chance to automate and solve them adequately. By utilizing centralized computer power the little businesses will not be penalized simply because they are little. They could be provided a complete "package service" with a program for all required calculations from entrance of sales order to billing, inventory control, and sales analysis with the addition of up-to-the-minute customer records. Far more small companies exist than large ones and even the large companies were once smaller. Furthermore, more individuals than firms exist that may need data processing or information retrieval.

It is true that computer manufacturers are attempting to meet the needs of the small businessman. Presently available are recording and reporting devices for collecting and transmitting information, but they are all off-line devices and the term "off-line" implies delay. They fall short of a total system concept that is necessary for effectively running a business. Inventory and other records are only available after a delay - which ties the system to cycle processing.

Under the computer power system both input and output could be on line. Men would not be needed to check inventory in the period between computer runs. There would be no delay in ordering and processing of orders because no paper would have to be transmitted.

No breakthroughs would be required in technology to implement this program. Even the cost of telephone-line equipment is presently low enough to facilitate general usage.

Programming and Service

Obviously a new set of programs cannot be developed for each customer when there are millions of customers. Standard packages must be designed and receive customer acceptance; they must be carefully conceived to meet nearly all possible customer needs and must be precisely programmed for efficient computer usage. Because all desirable applications could not be packaged and implemented at once, another question arises: Where is the line to be drawn between diversity of packages available and the size of the geographical area covered by a computer utility operations station? This question can be answered only after computer power is available to the public, and the needs of users are evaluated. Certainly, programmers will have to devise new techniques for producing software systems of greater diversity than envisioned today.

Certain types of applications would be most amenable to handling on the basis of a large geographical area. These applications are generally information storage and retrieval, such as medical diagnostic aid, legal search, or technical information storage and retrieval for a specific industry or centralized information bureau like a credit bureau. It is possible that each major center could specialize in one of these areas in addition to servicing customers with the previously mentioned applications packages. Then each of the centers would act as a switching center in handling requests for the main stations in their local areas. The number and location of the centers would depend upon the availability and cost of the communication line service.

The specialized service of the information retrieval type should not be overlooked. It is a high level of service which can be performed for customers especially when the information furnished was not provided initially by the client. Perhaps the only higher level would be providing needed information to users when they had not even requested it. Naturally as each successive level is reached, a higher level of customer service is achieved for which the user would be expected to pay a higher charge.

Before leaving this subject it would be well to mention one particular area of the information storage and retrieval type believed to possess tremendous potential service to humanity. This is the field of medical diagnostic assistance which is now being undertaken by several computer groups. Although this work has been mainly experimental, practical programs are resulting. The value of information storage and retrieval operations is particularly enhanced and made more economically feasible by acquiring a large subscriber list to spread the costs of accumulating information and of the extensive mass storage facilities required.

Manufacturing Requirements

The production rate of computers and standard peripheral equipment would necessarily have to be increased, at first slowly as regional computer power stations opened, then rapidly as they expanded, and finally slowing down to provide replacements and more efficient equipment. Special peripheral devices would require modification at intervals and new peripheral units would have to be developed; manufacturing would change periodically because of the modifications.

Computer power utilities would not necessarily require at the outset any different types of central processors than exist

today. But quantity manufacture would have to be stepped up, with attendant problems of retooling and plant expansion. Communications devices that connect to common carrier lines already exist in the designs of several manufacturers and only need development and mass production.

Engineering Considerations

As is true in the case of manufacturing, the effect on computer engineering would be slight. The necessary equipment is already designed and ready for development. The computer utility program can help underwrite the development costs of equipment it would use by spreading these costs over a greater number of units.

As already indicated, the development of new, more powerful systems also becomes more likely because of the need to use the most economical source of computer power just as an electric utility must obtain the most efficient power-generating plants. Another engineering consideration is the necessity that future large-scale systems be designed for direct hookup to large central processors.

Customer Engineering

Extensive service should not be necessary on peripheral equipment at a remote customer site. The equipment would be designed in a modular fashion so that the serviceman could easily replace a module.

Pricing Considerations

The pricing structure for computer utility power must place a premium on real-time or immediate-demand service so that it may be balanced by more moderately priced semi-real time service and even less demanding and cheaper, extended-time limit or overnight service.

Pricing policy for computer service should probably be on the basis of a minimum charge which varies according to class of service desired plus additional charges for services above those covered by the minimum charge. This policy represents a pricing approach closely akin to that of electric utilities, as it obviously should, so that the computer usage curve can be leveled out and the maximum computer power requirements reduced. Extended-time and off-peak service rates can be economically justified.

No attempt has been made to define the units of usage for measurement purposes. Usage time, by itself, is not the whole answer because this will vary according to type of computer system which is employed and will also be affected by the extent of the system (i.e., 10 tapes for sorting vs. 6 tapes, etc.) used at a given time. However, usage time may be the most feasible method, if the utility concept is to be followed in its fullest sense. In addition to a central computer usage-time charge, it would be necessary to add certain extras such as portions of on-line mass storage units reserved for the sole use of a customer (disc storage for real-time inventory control, etc.).

The following table lists some of the more common packaged applications that might be offered by a computer power utility. Admittedly, the applications are not all-inclusive for any of the listed customers; they are merely suggestive of many that may be offered.

Customer	Applications
Aircraft & Aerospace firms and agencies	Research and development computations
Airlines	Reservations systems, accounting, payroll
Banks	Accounting, payroll, balance verification
Brokers	Stock & security prices on immediate demand, billing, accounts receivable
Chemists & other scientists	Scientific computations, information retrieval
Civil Engineers	Computations for traverse, cut-and-fill, etc.
Department stores	Billing, accounts receivable, inventory, payroll, transportation routings
Doctors	Accounting, billing, diagnostic aid, payroll
Hospitals	Accounting, diagnostic aid, analysis of patients' records
Insurance companies	Accounting, information storage and retrieval
Lawyers	Information storage and retrieval
Management Consultants & all business management	Report generation
Message centers	Centralized services

Customer (Contd.)	Applications (Contd.)
Railroads	Accounting, payroll, centralized RR car information, replacement problems
Real estate brokers	Multiple listings, pricing
Savings and loan associations	Accounting, loan information, balance verification
Schools	Class scheduling, student records
Truckers	Accounting, loading, fueling, replacement problems, optimum load distribution
Warehousing	Order processing, inventory control, sales analysis

Competitive Considerations

A prime consideration is: Who can implement the computer utility concept? Would it become a monopoly of a single large computer manufacturer or would so ambitious a plan require the use of all present resources in equipment and personnel that are now engaged in computer manufacture and service? The history of the electrical industry may suggest the answer. Electric power is not supplied by one giant corporation that functions for the entire nation, but by a number of regional stations independently operated under local franchises. If this plan works for the electric power industry, why should it not work for the computer power industry?

Several large manufacturers in the computer industry already have equipment of the size and flexibility for use in computer power plants at costs to individual customers which would be less than operation of any individually operated computer. Besides computer manufacturers, computer service organizations of non-manufacturers are also capable of operating a computer power utility. However, a non-manufacturer has one handicap: he has to use the equipment made available by manufacturers and therefore has added cost of operations. Here would come a real problem in industrial change-over, but the evolution of any organized endeavor always meets such problems, and time combined with economics overcomes them eventually.

There should be a place in a regional scheme of computer power utilities for all who are within the electronic data processing industry today, whether manufacturer or non-manufacturer. The United States covers an area that is continental in extent. Within its borders are geographic features of every description, with population concentrations ranging from 2.6 per square mile in Nevada to 812.4 in Rhode Island. Computer power stations would certainly vary to the same extent that electric power stations vary today when Consolidated Edison with assets of over $2 1/2 billion services millions of customers and a small station in a mountainous region of the west with assets of less than $50,000 services scarcely more than a hundred customers. No national monopoly would or could be set up as all computer stations would necessarily operate over common carrier lines or over microwave networks owned by other industries.

Data Processing - Past, Present, and Future

Although the introduction of electronic data processing has had profound effects upon industry, science, and government the effects have been largely those of business expansion, speed of operations, and a shift in the occupations of workers. Computer memories have replaced ledgers and files; compute modules have replaced adding machines and manual calculations; electronic controls have replaced manual controls; electronic sorters and mergers and plotters have replaced mechanical equipment or hand labor. Business transactions have been tremendously speeded and increased, but they have not changed radically in nature. Large firms buy or rent their own electronic data processing equipment; a few firms may join together to use the same machines; a slightly larger number of firms take or send their data processing problems to service bureaus. But electronic data processing is still a largely localized matter and because of its cost is available to only a few customers. The mass consumer has no access to it at a price he can afford to pay.

The past, present, and possible future of the computer industry may be seen more clearly by comparison with the evolution of the electrical industry. The period around 1869 when Thomas Edison was completing development of the bipolar generator and the incandescent lamp may be likened to the early days of computer invention when central processors with mass memories, controls, and compute sections were first designed.

In 1882, Edison and a group of businessmen started the first United States central power plant, locating it in New York City to service 59 daring customers. After this first demonstration of the practicability of the generation of electricity and its delivery to consumers, electric power service was extended to other areas. It was still, however, confined to small portions

of the larger cities where small, isolated powerhouses were erected, local in scope of service, and independent of each other. Direct current could not be transported any great distance. The computer industry, in its equivalent stage of development, was also forced by the nature of its equipment and its method of services to confine usage to localized areas, accomplishing data processing on the immediate site of installations, or - in the case of service bureaus - in localized places of easy access to customers.

By 1900, pioneers in the electrical industry had increased the voltages at which electric power could be transported and lengthened the radius of transmission. The transfer of electric power in large quantities over distances of 200 mi. or more became an operating actuality, and by 1920 widespread and fundamental changes were necessary in the organization of services and in schemes for regional supply.

The computer industry stands in 1962 where the electric power industry stood in 1920. The transmission of electronic data processing over long distances has become an immediate possibility through the development of terminal communications devices and the means of using the common carrier communications channels and networks now available for leasing. The computer industry is ready for the same fundamental and widespread changes in organization and schemes for regional supply as occurred in the electric power industry. Costs must be lowered, and soon, if the industry is to meet the growing needs and demands of the public.

The major complaint of computer users is the cost of data processing systems. Today's costs run between $750 to $200,000 <u>monthly rental</u>. Computers have power, yes, but how many can afford to use that power?

The analogy between the electric power industry of the past and the computer industry of the past and present may be clearly shown. A continuing analogy may also be possible to predict the computer power industry of the future - should it become a public utility - with the electric power industry of the past and present.

The development of regional schemes for the supply of electricity to customers proceeded with such rapidity that by 1950 electricity was available in all urban communities in the United States and across Europe and in most rural areas in the United States. As the evolution progressed, advances and improvements were instigated in the technical aspects of production and distribution of electric energy and were made possible by economies inherent in the organization of the enterprises on a large scale.

Today, the larger of the electric power utilities represent enormous concentrations of both electric power and corporate wealth. Electric power is produced in great volumes at diverse and often remote sources, and is regulated and controlled to meet the constantly varying demands of consumers. In like manner, the larger of the computer power utility regional companies would service large areas of the population and small computer utilities with strictly limited capitalization would service areas of sparse population.

Should computer power utilities become a reality, they would, in all likelihood have to emulate the example of the electric power industry just before 1920, starting by the incorporation of small computer power-producing plants in urban areas. The successful delivery of computer power at points remote from the source of supply should soon make advantageous the abandonment of small stations and their replacement by larger units located at strategic points. Then there should occur - as with the electric power industry - the consolidation of many competitive urban undertakings into larger enterprises under regional jurisdictions, attended by a rapid growth in the size of the computer power-producing plants and by a marked increase in efficiency of operations and in number of subscribers. Computer power stations would likely remain regional in nature, as have electric power stations, but they would tie together many different power sources. Eventually this interconnection should result in the evolution of systems of great size that would have an effect upon the financial structure of the regional stations and a profound effect upon the life of the nation.

Admittedly, the foregoing statements reflect an optimistic point of view and present the more favorable side of the picture. Optimism is a stimulant, but no nation-wide system of public utilities ever got started without staggering problems. The Encyclopedia Britanica, in its account of the electric power industry, records that "Expansion of generation, transmission, and distribution was accompanied by attendant increase in the problems of engineering, finance, management, and public regulations."

Computer power utilities, organized on a regional basis, would undoubtedly meet equivalent problems in engineering, finance, and management. Competition to secure wider and wider areas of control would be inevitable. Federal and state regulations would undoubtedly be imposed upon the separate and group plant operators. But these problems are no more insoluble in the proposed computer power industry than they were in the electric power industry.

Uses for Computer Power

Skeptics may say that the foregoing analogy between the electric power industry and a proposed computer power industry is a misleading one. They may point out that everyone can use electric light to illuminate home, office, or plant, but not everyone can use computing power.

Skeptics may say, too, that electric power may be used for an endless variety of residential and industrial purposes, but only business men, industries, scientific researchers, and governmental agencies need high-speed, electronic computing power. The answer to this objection may be sought in the evolution of electrically powered devices for an endless variety of uses. When electric power became available to everyone, new uses for it were discovered, and the list of human needs, occupations, and activities that require only a cord plugged into an electric circuit continues to grow each passing year.

No one can predict today the uses that the public will find for computing power once it is made available to them at a price they can afford to pay. It may be difficult for the tradition-bound mind to envision, but the fact remains that computer power has wide potentialities for services to mankind. The public itself will discover the small uses that today would sound as much like science fiction as would have telling a man in the year 1900 that his daughter would plug in a cord to iron her clothes and his granddaughter would plug in a cord to run a typewriter. Public demand is a tremendous incentive to inventiveness!

Conclusions

The facts emerging from the foregoing evaluation are these:

*There is a definite trend toward offering computer service center facilities to customers with an ever shorter time lag between input and completion of task.

*Both large and small customers are being offered and are using computer based services.

*Customers are getting more service in a package form.

*Many similarities exist between growth of utilities and computer service center operation.

*A real-time computer utility operation with remote devices is now feasible.

*Such an operation could provide effective, economical computer power to large numbers of individuals and businesses in varied applications where no computer could otherwise be economically employed.

*Computer power as a common utility would, in all likelihood, find widely varied uses by the public and should be profitable both to computer manufacturers and utility entrepreneurs.

Serious planning for the implementation of this concept should be undertaken soon unless a better plan is found, for no industry can stand still. It must progress or retrogress, and the time to plan for the future is when current operations have reached a plateau. The time is right for a new and broad-based plan for supplying computing facilities to the public. All that is needed today is the vision and determination to start out on a new road. In view of the present situation in the computer industry, a strong possibility exists that it may have no choice.

Acknowledgments

This paper is an adaptation and extension of a report prepared by two co-workers of the author at the UNIVAC Division of Sperry Rand Corporation: Leo Higgins and Allen Trangle. The author also acknowledges indebtedness to other co-workers at UNIVAC who made contributions in the form of ideas, writing, and editing.

Bibliography

1. Encyclopedia Britannica, Vol. 8, "Electric Power."

2. Hurd, Cuthbert C., "Applications of Electronic Data Processing Machines," Mauchly, John W., "Data Processing," and Wooldridge, Dean E., "The Future of Automation," Automation in Business and Industry, Chap. 8, 11, and 19, John Wiley & Sons, N. Y., 1957.

3. Mann, Alan O., "A Publicly Regulated System of Management Control Series," Management Control Systems, Chap. 6 E, John Wiley & Sons, N. Y., 1960.

4. Mann, Alan O., "Computer Power: A Public Utility?" Computers and Automation, April, 1959, Vol. 8 - No. 4, pp 11-16.

5. McCarthy, John, "Time-Sharing Computer Systems," Management and the Computer of the Future, Martin Greenberger, Ed., Chap. 6, M.I.T. Press and John Wiley & Sons, 1962.

RESPONSIVE DATA HANDLING AT LOCKHEED

BY

NORMAN H. CARTER

Computers, Data Handling, Telephone Lines, Computer Interface, Random Access File, Error Detection, Correction Procedures

RESPONSIVE DATA HANDLING AT LOCKHEED

Norman H. Carter

SUMMARY

The data gathering project initiated at Lockheed Aircraft Corporation established criteria for equipment selection against which manufacturers' proposals could be evaluated. The equipment which was selected as being closest to the criteria has been field installed. The performance of the equipment is significant in that information can be transmitted over voice grade telephone lines through a computer interface to a random access file from which responses can be retransmitted to remote locations in any one of several different forms. Error detection and correction procedures are an integral part of the system. Successful completion of implementation of the data acquisition system will permit expansion into management display techniques.

PROBLEM

One of the major problems in business data processing today is that of organizing the gathering and dissemination of data so that the micro-second computing speeds available today are not nullified by old-fashioned, inefficient manual data collection and manual data dissemination systems.

While the computer has revolutionized data processing, data gathering methods are still back in the horse and buggy days. Men, key-driven machines, such as the teletype, mail and other slow feed-back is still used, and computers are still essentially tied to batch processing of data.

LOCKHEED ADA PROJECT

Because of this discrepancy, Lockheed in 1958 initiated project ADA, to define their requirements in Automatic Data Acquisition. A systems study was initiated to determine the broad data needs of Lockheed divisions in such applications areas as the collection and dissemination of shop control information, labor distribution, and the traditional payroll or time clocking systems. The systems needs, records layouts, and applicable volumes were reduced to a systems specification which further spelled out the physical locations involved. During the preparation of the specification, Lockheed contacted a number of other aerospace companies to determine their interest in the field of data acquisition and to exchange systems ideas. As a result of these contacts, much valuable information was received and the Lockheed specification showed great similarity to a generalized aerospace industry specification.

Copies of the finished specification were disseminated to all hardware manufacturing companies who showed any interest whatever in the field of manufacturing data gathering equipments. About 26 companies responded with proposals to produce hardware. At this early date only one company was able to promise delivery of prototype equipment and one company was ready to build a special purpose off-line recording device without transmission capability. All other proposals received were for "paper" equipment, proposals which could be reduced to hardware in times varying from twelve months to two years, and some of which would have required Lockheed to invest engineering and development funds.

FIRST EQUIPMENT TEST

Lockheed did, in fact, build two prototype models, with the manufacturer, of the off-line recording device which was used to test the systems criteria, and the reactions

of shop personnel, to equipment of this type. This device recorded data from two plastic cards, part of a punched card and a five bank keyboard on a wide paper strip locked in the machine. The information was not transmitted.

The test, covering about three months, was successful and three major things were discovered:

1) Transmission capability is mandatory in any data collection system. The problems of control of paper tapes and other records at remote locations is too great.

2) An individual receipt, for each operator, showing all recorded data from each transaction is unnecessary.

3) The concept of data gathering is acceptable to auditors, union members and officers and to customers if proper systems controls are designed into the system and if adequate training is performed.

After further equipment evaluation, Lockheed selected RCA as the vendor for its data gathering equipments. In the fall of 1960, the first twenty RCA EDGE input stations were installed in a cooperative test and evaluation program, which led in the fall of 1961 to orders for systems build-up in three major Lockheed divisions.

CRITERIA FOR EVALUATION OF DATA GATHERING SYSTEMS

The balance of the paper discusses the criteria which were established for evaluating data gathering systems, some discussion of experience to date with the off-line systems, a review of the on-line and real-time systems presently installed, and a glimpse into the future stages of Lockheed's management control system concept.

Of course, the system must be economically feasible. Dollar savings as well as intangible benefits must be present or the installation cannot be justified.

There are six equipment criteria against which Lockheed measures data gathering systems. These are as follows:

1) First, the system must be capable not only of sending information to a computer or computers, but must also be capable of receiving information back in forms such as page print-out, punched card, punched paper tape, and in the future, graphic visual display devices such as video tubes. This two-way capability is mandatory if the automatic data acquisition system is to be justified on economic grounds.

2) The second criterion is that the system must be capable of transmitting over distances from less than 1/4 mile to more than 500 miles, with the same basic equipment and without the addition of special sub-sets or other similar devices. This requirement is expressed to manufacturers by stating that the system must transmit over a two-wire cable of telephone voice grade or lower, not over multi-pair cables. This facilitates the use of an extensive private wire network which is operated between Lockheed plants in all parts of the continental United States.

3) Requirements further call for the provision of automatic error detection and message retransmission. It is felt that the operator should not have to make a determination as to whether the message has been correctly received at the data center. This determination must be made automatically

by the machine. The automatic error detection devices insure that after a given number of abortive attempts to send a message, the input machine will take itself out of service and at the same time lock all of the documents into the machine, and signal the operator to initiate an error procedure which is clearly defined and posted at each station.

4) A further criterion is that all parts of the input message must be inserted in the machine by the operator at one time prior to the beginning of transmission and that the machine must have such interlocks that the message will be properly composed. For instance, if a message calls for the insertion of a plastic badge and a punched card only, the machine will refuse to accept an entry into the variable keyboard of the machine. If the transaction code calls for a variable entry and plastic card only, then a punched card cannot be inserted, and if it is, it will be returned automatically to the operator and transmission cannot be initiated. These are referred to as inclusive interlocks. Another level of interlock requires that the correct punched card be inserted when a transaction requires a punched card. This is achieved by stating that for a given transaction code, the punched card inserted must have a specific punch in a specific column (say a two punch in column five) which the machine will sense. If this punch is not in the card, the machine will reject the entire transaction as being incorrect. This is referred to as an exclusive interlock.

5) A criterion which has been deleted from the requirements is that there should be a local print-out of all information transmitted from a given station. This was felt to be necessary to satisfy employees and line supervision that messages were being transmitted correctly. Experience in the field showed that this was not of concern to the users and could therefore, be deleted.

6) Although not a criterion in the sense that it affected electronic engineering design, it is mandatory that the equipment should be of such simplicity that all workers will be able to use it. It is not possible to justify the cost of machine operators in the factory or office areas for all of the systems which it is intended to implement on this equipment.

7) Modularity of design is a mandatory requirement to facilitate maintenance in the field.

IMPLEMENTATION OF FULL SCALE SYSTEMS

Some idea of the breadth of planning, systems development and implementation that is going on can be shown by the areas in which work is being performed. These applications areas include labor distribution, payroll, shop control, including job move, standard hours completed for performance reporting, machine loading and other similar functions, material control and planning, receiving and receiving inspection, and quality control and many others.

A typical complete network such as will be installed in the three major Lockheed divisions is illustrated in Figure 1. As can be seen, the system has input stations for sending information through a computer interface, either to a disk file for immediate action or processing, or to magnetic tape for storage and later batch processing. The inquiry station enables a remotely located operator to inquire for information through the interface to the disk file and to receive an answer through a printer, with control format; through a punched card output, either with or without the printer, or through punched paper tape.

The RCA EDGE input station is the basic input unit of this system and the RCA alpha-numeric inquiry station is the device for interrogation of the disk file. In one of the divisions, the computer interface is a dual RCA computer, with a Bryant disk file. In one of the other divisions, the interface may be a Beckman system, with wired program, feeding information to an NCR 315 computer with CRAMS.

It is mandatory to provide duality of computer interface as these systems may be required to receive information and return information to any of the remote points at any time during each 24 hour period. Therefore, the system cannot be completely shut down for preventive maintenance or for any other cause. The duplex computer interface must determine when it has a problem; must insure that all messages in transmission when the problem arose were retransmitted through the second computer. Although some degradation of speed may be acceptable, it is required that the employee using the input device should not be aware of the breakdown at the computing center.

Lockheed's data handling networks will be located as follows:

1) One at Sunnyvale, California, with a major sub-system located in Van Nuys, California;

2) One at Burbank, California, with a major sub-system at Palmdale, California; and,

3) One at Marietta, Georgia.

Two smaller systems will be implemented at Lockheed Aircraft Service plants at Ontario, California, and Idlewild International Airport, New York.

System build-up at the locations will reach a total of 735 input stations with appropriate support equipment and about 50 inquiry stations by June, 1963. The present number of input stations implemented in the off-line system is about 250 units.

Operating experience to date with the off-line systems has shown that the planning criteria which were used were slightly conservative although essentially correct. For instance, the ratio of input stations to central receiving units was established at approximately six to one. An actual operating level of seven to one appears to be adequate.

Experience is being gained rapidly as more than ten million characters a week are presently being transmitted over the networks.

ERROR DETECTION AND CORRECTION

Error rates are being reduced as equipment is shaken down and necessary field modifications are made. There are two rates of error that are significant in these systems. The first are those errors which are detected and corrected by the equipment. The second, those errors which are undetected by the data transmission system, but detected later in computer processing.

It is possible to automatically detect and correct errors using the retransmission feature built into the RCA equipment. The specification submitted by the equipment manufacturer for detected errors is that there will be fewer than 3% of retransmission occurrences when measured against the total volume of messages transmitted. The present rate of detected errors is considerably less than 1/2%.

The specification for undetected errors is 1/2% with present operating experience varying slightly at the different locations between .25% and .1%. It is expected that an undetected error rate of less than .05% will be reached after the equipment has been shaken down and necessary modifications have been made. It is felt that the present rates are reasonable for the state of the art of the equipment. As the computer programs used for processing this information are further refined, it is expected that some of the errors now charged to the equipment will turn out to be program errors, operator errors, or input documentation errors which we will then be able to detect.

The implementation of these data gathering networks provides an opportunity to achieve cost reductions, represented by fewer personnel in the functions concerned, as well as certain other advantages such as removing the human handling from these messages, thus eliminating transcription and keypunch errors, and insuring faster turn-around of information and permitting an approach to management by exception.

CONCLUSIONS

In summary, the Lockheed criteria for evaluation of equipments can be expressed as follows: examine proposed equipments in relation to the state of the art. Evaluate over-all reliability, and cost comparisons between equipments recommended. Measure the ability of a given equipment to be expanded as systems requirements expand.

The responsive data gathering networks discussed in this publication are in themselves only one step in the direction of a more comprehensive management information system. Once the ability to accurately gather and process information on a daily basis has been achieved, the next significant forward step can be taken, that of providing displays of critical information for management use in a control center. Furthermore, it will be possible to offer to management the ability to enter their proposed alternate courses of action directly into a computer stored model, then to observe the possible results of these courses of action displayed instantaneously. And, from the results displayed, management will make the necessary action decisions. This will close the loop from input through computing to display of results. From results, corrective action feed-back will occur to again control input and so to attain better production faster and at lower cost.

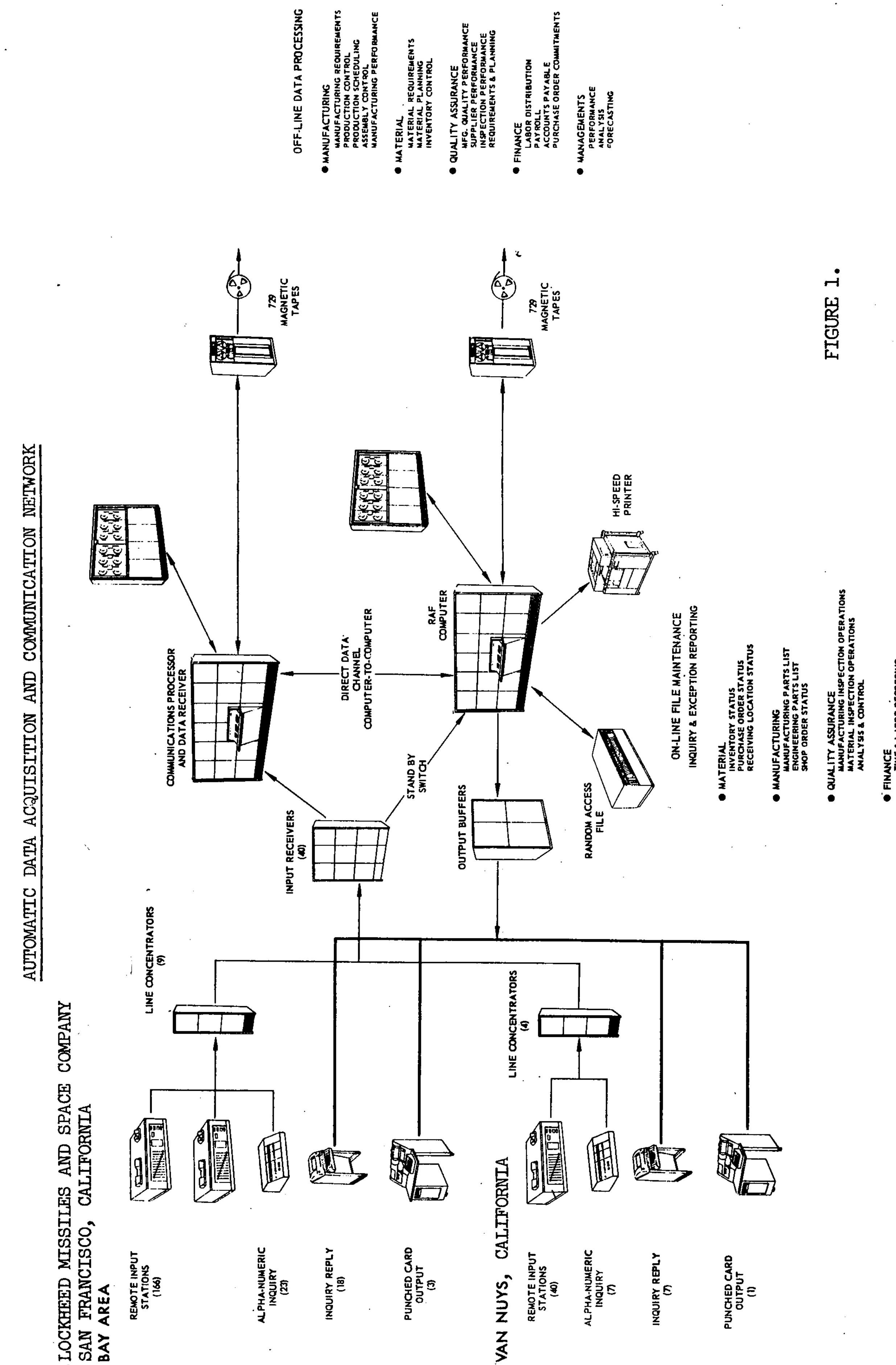

FIGURE 1.

TECHNICAL AND ECONOMIC JUSTIFICATION OF A STORE AND FORWARD SWITCHING CENTER

BY

DONALD J. BIRMINGHAM

Computers, Economic Justification, Switching Center, Store and Forward, Telegraph, Line Switching, Tariffs, Common Carriers, Way Circuits

TECHNICAL AND ECONOMIC JUSTIFICATION OF A STORE AND FORWARD SWITCHING CENTER

Donald J. Birmingham
Computer Department
General Electric Company

SUMMARY

The Store and Forward Switching Concept for telegraphic communications will not be replaced by the new line switching communication facility and tariff offerings of the common carriers. On the contrary, the Store and Forward Concept will assimilate these new services to provide a more flexible and capable system than that which was achieved in the past with the previous electro-mechanical systems. The economical justification will become more complex and the customer will be required to place a dollar value on added technical features provided by Store and Forward Message Switching.

INTRODUCTION

In the past, the justification for store and forward switching centers for telegraphic communications was based primarily on the cost savings achieved through line reduction. Along with the expanded requirements of data communications, which encompass both the old telegraphic requirements and the new computer on-line and off-line data requirements, there has developed a variety of new communications facilities and services; e.g., TELPAK, WATS, WADS, TELEX, and others. The introduction of these new services has not eliminated, but rather has expanded, the requirements for the store and forward system.

TECHNICAL DEFINITIONS AND COMPARISONS

The Store and Forward Switching Center has many definitions. As used in this paper, it is defined as: "A complex of equipment capable of receiving, switching, transmitting, and storing messages."

There are three basic system approaches which may be taken to a message network. These approaches are:

LINE SWITCHING - Terminal devices are interconnected to like terminal devices through a dialed-up connection. Coordination is performed by manual operator procedures.

STORE AND FORWARD/LINE SWITCHING - Terminal devices are connected to a store and forward switching center through a dialed-up connection. Messages are accepted by the switching center. As the switching center accepts the message, if an output trunk is available, an attempt is made to connect with the destination. If a busy condition exists, the message is temporarily stored. In certain cases, low priority traffic may be interrupted by high priority messages.

STORE AND FORWARD/WAY CIRCUITS - Like terminal devices are connected to a store and forward switching center through way circuits. This approach operates functionally in the same manner as present day conventional switching equipment.

Table I presents a list of technical features inherent in each of the three basic approaches. The technical justification of which approach to use is made by listing the technical requirements and matching them to the technical features offered. In many cases, compromises are made on technical requirements based on economic considerations.

The straight line switching approach is the simplest in complexity but offers few more features than those contained in the end device terminals. It allows operation with a minimum of operator training and has no system maintenance. To obtain a specific grade of service, both the line and terminal device utilization must be held to a minimum. This type of operation is best suited to the small volume users.

The solid state store and forward approach combined with line switching offers the greatest variety of technical features. This hybrid approach is designed to combine the best features and make up for the short-comings of the conventional store and forward/way circuit system and the line switching approach. The capability of code and format conversion allows the system to operate with mixed terminal devices.

It is also possible for intermixed operation where like terminal devices may bypass the store and forward center to communicate directly. The greatest system drawback is in the requirement for double transmission, one to the center and another to the end device. In an integrated data processing - communication system this may be greatly reduced if it is assumed that in future systems, over 50 percent of messages will be terminated and originated by a computer. This computer, in most cases, will be directly connected to the store and forward system as one integrated complex.

The solid state store and forward system when compared with the conventional store and forward system with way circuits provides a number of technical advantages because of its flexibility in operation with end devices. However, it offers less than the store and forward system with line switching, and in many cases it is advisable to provide duality in major functions to reduce system vulnerability.

ECONOMICS OF THE SMALL SIZE STORE AND FORWARD DATA NETWORK

The common carrier tariff services WADS and TELEX, are the chief competition for the private store and forward system. Tables II and III present a summary of message costs based on two general system models. The costs for the different types of systems are based as follows:

WADS - The WADS line-switched system is based on Model 33 and 35 Teletype units operating at 100 wpm with WADS tariffs.

STORE AND FORWARD/LINE SWITCHING - The store and forward switching center with line switching is based on a twenty line (10 input and 10 output) switching center with WADS input and output lines operating at 100 wpm with Model 33 and 35 Teletype units.

STORE AND FORWARD/WAY CIRCUITS - The store and forward switching center with way circuits is based on ten full duplex way circuits operating at 100 wpm with conventional type way circuits.

STORE AND FORWARD (CONVENTIONAL) - The conventional store and forward system is based on an average conventional system.

The cost totals tabulated in Tables II and III show that the difference between a straight line-switched system and a store and forward center with line switching is in the order of 10 to 15 percent. The justification of the store and forward switching center primarily would be made on a technical basis of requirements which could not be met by the straight line-switched system.

The conventional store and forward system with way-circuit control will normally have a higher cost than store and forward with line switching. However, there always will be special cases where the way-circuit method will provide a lower cost, but in general the technical advantages of the hybrid system may outweigh the cost savings.

THE INTEGRATED PRIVATE VOICE DATA NETWORK

The integrated Voice-Data Network requirement is normally generated by the company which has a private phone network consisting of a backbone of TELPAK-tariffed lines and which desires to add the data requirements "piggyback" on the voice system.

Consider the case of the Company "X" which has a nationwide leased-line network as shown in Fig. 1. This network supports a private voice system as shown in Fig 2. Table IV presents the combined data and voice requirements of Company "X". The geographic breakdown by percentage of volume is included.

The data requirements could be handled in one of three ways:

LINE-SWITCHED SYSTEM - Sufficient line and switching equipment could be added to handle the data in exactly the same manner as the voice network.

STORE AND FORWARD/LINE SWITCHING - A combination of line switching and store and forward could be made to provide sharing of low volume circuits by both data and voice and to achieve high efficiency on trunk circuits. This hybrid approach would attempt to maintain the best characteristics of both systems.

STORE AND FORWARD/WAY CIRCUITS - A non line-switched system could be used which has way circuits and direct-connected circuits to high volume points.

The line requirements and costs for the voice system and each of the three possible alternatives are presented in Table V. The Store and Forward/Line Switching approach is broken down into two systems, one with unrestrained input and output lines (lines which are not designed to handle more than 100 per cent traffic during busy hours) and restrained output lines (lines which are designed for queueing to provide a balance between cost and delay). The store and forward equipment in the hybrid approach does not have to be supplied in duplicate as the voice system itself can supply an alternate path for priority traffic during a down period.

The cost summary for each of the approaches in presented in Table VI. Two methods were used to compute the TELPAK line cost. One is based on allocation of only the additional cost of expanding the existing TELPAKs and the other on an appropriated cost where the data system pays for a percentage of each TELPAK link based upon the percentage of total lines used.

The cost figures for the store and forward system are based upon three switching centers located in San Francisco, St. Louis and Philadelphia. The cost of switching centers in the store and forward system with direct circuits and way circuits is based on complete duality.

The total cost figures for all systems are very close (within the realm of interpretation). The main point is that the availability of lines at wholesale costs does not rule out the use of store and forward systems, but it allows the system designer the freedom to generate hybrid approaches which provide more functions than the conventional store and forward approach.

CONCLUSIONS

The conventional store and forward teletype switching center, consisting of electro-mechanical switching or manual torn-tape switching for teletype, has become outdated. In order to compete with the new common carrier tariffs for line facility and outstation, the store and forward switching system must interconnect with line switching facilities, provide speed, code and format conversions, and interface with the expanding number of centralized on-line computing facilities.

The advent of TELPAK tariffs for the bulk voice and data communications user has lowered the price of line facilities to a minimum. However, the use of solid state computer-type equipment allows the store and forward switching center to keep pace and provides for the economic justification of store and forward equipment along with a network of TELPAK line facilities.

Table I

SWITCHING SYSTEM

TECHNICAL FEATURES

	LINE SWITCH	SOLID STATE STORE & FORWARD W/LINE SWITCH.	SOLID STATE STORE & FORWARD W/WAY CIRCUITS	CONVENTIONAL STORE & FORWARD W/WAY CIRCUITS
Message Switching		Yes	Yes	Yes
Line Switching	Yes	Yes		
Shared Facility (W/Voice)	Yes	Yes		
Line Facility Utilization	Low	Medium	High	High
Mixed Terminal Devices (Incompatible)		Yes	Limited	
Speed Conversion		Yes	Yes	Limited
Code Conversion		Yes	Yes	
Multiple Address Capability		Yes	Yes	Yes
Direct Computer Input/Output		Yes	Yes	
Journal Accounting		Yes	Yes	Limited
Unattended Operation	Yes	Limited	Limited	
Error Detection and Correction		Yes	Yes	
Operating Routines	Flexible	Fixed	Fixed	Fixed
System Maintenance	Non-Required			Non-Required
System Vulnerability	Low	Medium	High	Low-Medium

Table II

COST PER MESSAGE (50,000)*

	WADS	ST&F/LS	SOLID STATE ST&F/WC	CONVENT. ST&F/WC
Communication Service	.45	.35		
Switching		.10	.10**	.15
Lines			.20	.20
End Device			.10	.10
Operator & Program		.02	.02	.05
Installation & Facilities		.01	.01	.02
	.45	.48	.43	.52

* Based on location in Northern Illinois with 50,00 messages per month and 60 outstations locations, 50% of messages for input and output to a computer and 25% multiple address messages.

** Add .07 and .05 for duality.

Table III

COST PER MESSAGE (100,000)*

	WADS	ST&F/LS	SOLID STATE ST&F/WC	CONVENT. ST&F/WC
Communication Service	.30	.25		
Switching		.07	.07**	.10
Lines			.15	.15
End Device			.10	.10
Operator & Program		.02	.02	.03
Installation & Facilities		.01	.01	.01
	.30	.35	.35	.39

* Based on location in Northern Illinois with 100,000 messages per month and 60 outstations locations, 50% of messages for input and output to a computer and 25% multiple address messages.

** Add .07 and .05 for duality.

Table IV

COMPANY "X" TELEPHONIC & DATA COMMUNICATIONS REQUIREMENTS

VOLUME		
TELEPHONIC TRAFFIC	300,000	Long Distance Telephone - 8 minutes each
RECORD TRAFFIC		
Administrative Teletype	100,000	Messages - Average 80 Teletype words
On-Line Business Data	250,000	Messages - Average 250 characters
Deferred Business Data	5,000	Messages - Average 2000 characters
GEOGRAPHIC DISTRIBUTION		
EAST	40%	
Philadelphia		15
New York		10
Boston		5
Cleveland		5
Washington		5
CENTRAL	35%	
St. Louis		20
Chicago		5
Atlanta		5
Dallas		5
WEST	25%	
San Francisco		10
Los Angeles		10
Portland		5

Table V

LINE REQUIREMENTS AND COST

TELEPAK GROUP	TELPAK TYPE	MILEAGE	COST	LINES REQ	SP LINES	ADDITIONAL LINES REQ.		ADDED LINE COST	UNRESTRAINED LINES REQ.	UNRESTRAINED ADDED LINE CST.	RESTRAINED LINES REQ.	RESTRAINED ADDED LINE CST	
San Francisco-Portland	A	550	8,250	8	4	5	A-B	2,750	5	2,750	4	--	1
San Francisco-Los Angeles	B	330	6,600	18	6	10	B-C	1,650	10	1,650	6		1
San Francisco-Salt Lake City	B	620	12,400	18	6	15	B-C	3,000	2		2		2
Salt Lake City-Kansas City	B	950	19,000	18	6	15	B-C	4,950	2		2		2
Kansas City-Dallas	A	550	8,250	8	4	5	A-B	2,750	5	2,750	4		1
Atlanta-Charlotte	A	250	3,750	8	4	2			2		2		1
Atlanta-St. Louis	B	600	12,000	18	6	5			5		5		1
Chicago-Cleveland	B	340	6,800	18	6	23	B-C	1,700	2		2		2
St. Louis-Chicago	C	250	6,250	40	20	28	C-D	3,750	7		7		3
Cleveland-Philadelphia	C	360	9,000	40	20	28	C-D	5,400	7		7		3
Philadelphia-New York	C	90	2,250	40	20	15			15		15		2
Kansas City-St. Louis	C	210	5,250	40	20	20			7		7		3
New York-Boston	B	180	3,600	18	66	5			5		5		1
Philadelphia-Washington	B	140	2,800	18	6	5			5		5		1
Washington-Richmond	A	100	1,500	8	4	2			2		2		1
			$107,700					$25,950		$7,150		0,000	

Table VI

COST SUMMARY MONTHLY

	LINE SWITCHING	UNRESTRAINED	RESTRAINED	STORE AND FORWARD
ADDITIONAL COST METHOD				
Line Switching Cost	20,000	8,000	8,000	-
Store and Forward Cost	-	25,000	25,000	40,000
TELPAK Line Cost	25,950	7,150	-	-
Non-TELPAK Line Cost	-	-	-	10,000
Terminal Device	30,000	25,000	25,000	30,000
Switching Center Oper. & Prog.	-	9,000	9,000	9,000
Facility Cost	-	3,000	3,000	3,000
Amort. Installation Cost	2,000	2,000	2,000	2,000
	$77,950	$79,150	$72,000	$94,000
APPROPRIATED COST METHOD				
Line Switching Cost	27,200	10,000	10,000	
Store and Forward Cost	-	25,000	25,000	50,000
TELPAK Line Cost	46,000	27,000	25,000	10,000
Non-TELPAK-Line Cost	-	-	-	10,000
Terminal Device	30,000	25,000	25,000	30,000
Switching Center Oper. & Prog.	-	9,000	9,000	9,000
Facility Cost	-	3,000	33,000	3,000
Amort. Installation Cost	2,000	2,000	2,000	2,000
	$105,200	$101,000	$99,000	$104,000

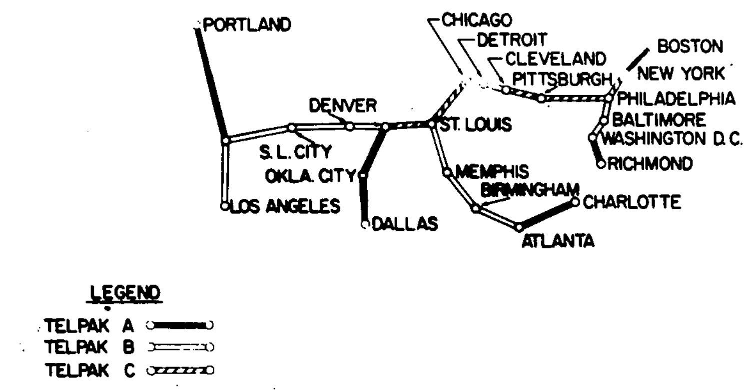

Fig. 1. Company "X" Telpak Routes.

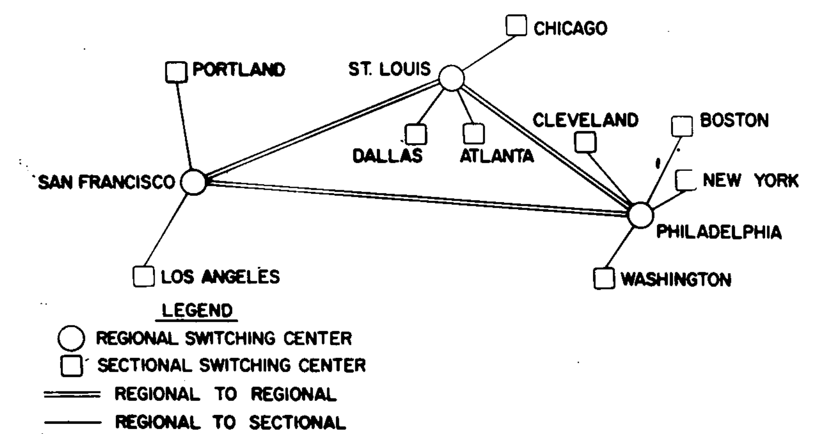

Fig. 2. Company "X" Private Phone Network Switching Plan.

THE HOWARD SAVINGS INSTITUTION -
A CASE HISTORY ON REAL-TIME DATA PROCESSING

BY

J. J. FELDMAN

Computers, Accounting System, Banking, Data Processing,
Real-Time, Savings, Mortgage, On-Line, Off-Line,
Telefile System

THE HOWARD SAVINGS INSTITUTION - A CASE HISTORY ON REAL-TIME DATA PROCESSING

J. J. Feldman

Summary

Nine years ago, in 1953, The Howard Savings Institution of Newark, New Jersey, launched a research study into automation. Today, Howard Savings Institution customers are enjoying the benefits of the fastest, most accurate savings accounting system in the world. This paper describes the customer benefits, the electronic data processing system that makes them possible, and the advantages and disadvantages of real-time data processing for the savings bank industry.

RESEARCH STUDY

The Howard Savings Institution began studying the feasibility of automation for savings and mortgage operations in 1953. In 1954, having concluded that the benefits of automation would be in the best interests of both the bank and its more than 300,000 customers, a search was started for a manufacturer who would build an electronic data processing system to meet all of the needs and requirements. The selection phase of the long range research program ended in 1956, when a contract was signed with The Teleregister Corporation of Stamford, Connecticut, for the first on-line, real-time electronic data processing system in the world for a savings and mortgage application.

From the beginning The Howard Savings Institution knew its needs and requirements. The search for an electronic data processing system to meet these objectives revealed many things. First, available off-the-shelf equipment was completely inadequate. Second, the problem was similar to that of the military system development with specified performance without specified hardware types. Third, in order to preserve the fine image of customer service The Howard Savings Institution earned during its 105 years as a public service institution, acceptable guarantees of system reliability and accuracy had to be assured. A final important revelation resulting from system studies was that there was no advantage gained by going to a purely off-line system; "balance" between input, output, storage devices and main frame equipment was far more important than inherent speeds and capabilities of individual components.

On-Line vs. Off-Line

During the selection process, cost analyses revealed that there would be no payoff in converting savings and mortgage applications to off-line equipment. Greater economies could be achieved through the use of the more mundane punched card accounting systems. Therefore, to provide justification, customer service had to be bettered appreciably. This could take place only at the customer window (teller window if one is standing on the "other" side of the counter), a location where off-line processing has little if any tangible effect. Teleregister's on-line, real-time teller posting machine, was the solution to the problem. Fig. 1 illustrates this machine.

Equipment "Balance"

As the equipment field was surveyed, it was observed that speed and capability range varied radically between similar components. At first, interest was held only in the apparent fastest or best of each equipment type. Soon, however, came a realization that cost was almost directly proportional to inherent component characteristics and that there was little sense to pay for equipment twenty-four hours a day when the equipment would be employed for a much shorter period of time. Therefore, the system selected, though certain components are the earliest versions of their types, provides the essential "balance" sought. In other words, it is geared to needs, not to niceties.

Customer Service

The Howard Savings Institution is a Mutual Savings Bank. Their only product is service. Therefore, it was necessary to guard carefully the fine reputation as a provider of customer service while entering into the field as an electronic data processing system user. To insure continuity of service, sound system uptime and reliability guarantees were requested and received as part of a long term, eight year lease contract.

THE TELEFILE* SYSTEM

The data processing system, referred to as a Telefile system, contains dual Telefile computers, random-access memory drums, magnetic tape handlers, automatic typewriters, in-line punched card equipment, a communications network and on-line, real-time teller machines called Teller-Registers.® The relationship of each of the sub-systems to the computers is maintained by a priority control system which permits multiplexed operation and overlapping of computer internal operations with sub-system operation.

Central Processor

Each Telefile computer is a large scale, general purpose, stored program, fully transistorized machine. These are capable of operation in either dual or duplex modes. In the dual mode, both processors operate simultaneously on the same data, permitting a cross check of each operation. In the duplex mode, each processor operates independently, but with access to the common pool of sub-systems. These processors, therefore, also serve as control and switching centers for the sub-systems. Consoles associated with each processor and the overall system console display the internal operations as an aid to the operating personnel and maintenance technicians and are used in the test and analysis of programs. The randomly accessible, digit addressable core memory of each processor has a capacity of 10,000 decimal digits and a cycle time of sixteen microseconds. This relatively small core memory is used programmatically for sub-routines and data required by the immediate program. Other sub-routines and data are contained in the mass storage sub-systems and are available on demand.

Magnetic Drum Sub-System

The magnetic drum sub-system consists of drum units, drum control units and a drum central unit. Each drum unit consists of a vertically mounted drum and read/write circuitry housed in a cabinet. The drums revolve at a speed of 1800 revolutions per minute, providing an average access time to any piece of information of 17 milliseconds. The drum system control unit and the drum central unit provide the necessary logic and solid state switching circuits.

Magnetic drums are used in this system for mass storage of on-line available account balance information and program storage. Each magnetic drum has a capacity of 1,050,000 decimal digits. Single or multiple digit accessibility is provided on a random basis. Parallel processing of other information may be carried out in the central processor during access time. Presently, the system employs five of these drums: three as balance record drums, one as a transaction record storage drum, and one for processing and fall back.

* Trademark, The Teleregister Corporation

® Trademark Registered

Magnetic Tape Sub-System

The magnetic tape sub-system consists of tape handlers with their associated electronics connected through a control unit. Both automatic and manual control of tape handlers is provided. Depending on tape length, a single magnetic tape provides storage capacity for up to 4,500,000 decimal digits of information. When running, the tape passes the read/write head at a speed of approximately 60 inches per second. This corresponds to a read/write rate of 9,000 digits per second. Information is contained in tape blocks, each consisting of a 300 decimal digit storage capacity, $1\frac{1}{2}$ inches long. The inter-record gap is $\frac{1}{2}$ inch. Four of these magnetic tape handlers are currently in use.

Automatic Typewriter Sub-System

The automatic typewriter sub-system provides communication between the Telefile and the operator. The sub-system is made up of two racks of solid state logic, two relay panels, and two automatic typewriters. Each typewriter with its sub-system is directly associated with one of the central processors.

Punched Card Sub-System

The punched card sub-system permits one IBM 514 Reproducing Punch and one IBM 403 Tabulator to be directly connected to the Telefile System by a multiconductor cable. The cable can be disconnected for conventional utilization of these machines whenever necessary. The tabulator is used for direct output of printed reports. The reproducing punch is used for either direct input or output of punched cards. Though this equipment was not provided by Teleregister, it was modified by them to facilitate operating modes under the direct control of the Telefile.

Communications Sub-System

The data processing center is connected radially by communications lines to each office. At the processor, each communication line terminates in a unit called a message register. This unit stores incoming messages temporarily until the processor is prepared to process the call, at which time the stored program transfers the message register contents into the core memory. Conversely, the answer or response to the call is transferred from the core memory into the message register where it is stored for transmission by the communications equipment. The average processing time is 137 milliseconds per call. Fig. 2 is a simplified diagram illustrating the relationship of one office and its communication line with the central processor site.

Three types of data communication lines are employed in this system. High-speed (Dataphone), slow-speed (Teletype) and multiconductor cable (local). The choice between high-and low-speed was a matter of transaction volume created by each office. As a result, two offices justify the high-speed circuits while four others use slow-speed transmission. One office, due to its proximity to the Telefile (it is in the same building), employs local lines. All telephone communications are in dual, i.e. one primary and one alternate route network.

At the offices, communication circuits terminate in a transceiver. This unit is capable of service for three Teller-Registers in slow-speed circuits, up to eight in high-speed circuits and up to twelve in local circuits. Again, as in the selection of line types, the activity role is the primary controlling factor in the quantity of Teller-Registers which are to be served by a single transceiver. The functions of a transceiver include connection of a bidding Teller-Register, control of transmission, storage of returned data, error detection and control of the printer in the Teller-Register.

The Teller-Register is a counter top input/output device consisting of a keyboard and printer housed in a single unit. This device was designed specifically for on-line teller operation rather than an adaptation of a conventional window machine. The design includes the following features: interlocked keyboard construction, check and money order writing facilities, provision for two tellers to use one machine with separate audit functions and key controlled security for each, provision for special supervisory input under key control, key locked journal tapes, and parity checking on all data. The printer is a modular and removable unit contained in a separate compartment within the Teller-Register. This also was designed specially for its purpose, featuring echo-checking techniques whereby the final positioning of the printing mechanism associated with the printing of the balance column data is checked against the stored answer in the transceiver. This printer is capable of printing on two surfaces simultaneously, the journal tape and the passbook, or forms contained in the carrier. The passbook carrier provides for the handling and alignment of passbooks, checks, vouchers and other forms. Presently, the system contains 32 Teller-Registers.

The communications sub-system also has a special device known as the Executive Monitoring Console associated with it. One of these devices is located in the office of Howard Board Chairman, William L. Maude. Upon depression of a button, this device displays total deposits or withdrawals accumulated through the day up to that moment.

THE TELEFILE OPERATION

Presently, approximately one quarter of a million savings accounts and 30,000 mortgages are being processed in the system. The savings accounts consist of approximately 150,000 regular, 60,000 school, 25,000 club and 7,000 payday savings accounts. Teller-Registers are located in all seven of the bank's offices. Their distribution is shown in Table 1.

On-Line

All normal banking transactions are accomplished in this system through the Teller-Register. By the single process of entering an account number on the Teller-Register, the teller commands the system to verify the account number and print the current balance on the Teller-Register journal tape. The equipment, not the teller, retrieves the current balance automatically. This is an important factor since it avoids erroneous old balance entry. Deposits, withdrawals, interest posting and unposted "no-book" transactions then can be processed immediately. Holds and other special conditions are signalled to the teller automatically and must be "processed" by the teller before further banking transactions can be made. The passbook is updated automatically, teller journal tapes are printed in the Teller-Register and each transaction is recorded, with teller identity, in the central magnetic records. Proof totals in five categories for each teller, stored in the Telefile, are available for print-out on the journal tape of each Teller-Register. Teller-Register keyboard layout is illustrated in Fig. 3 and the transaction types and controls, with operational procedures, are listed in Table 2.

Off-Line

The system computes anticipated interest and updates detailed account records at the end of each day. The interest calculation is performed automatically during non-banking hours. Thus, the tedious task of manually calculating anticipated interest is completely eliminated as a daily routine. As part of this processing, trial balances and special reports can be obtained in any degree of detail desired.

Mortgage processing takes place on the "second" processor, off-line, during banking hours. This includes billing, remittance processing, escrow analyses and other related mortgage functions. If the on-line processor should fail, control is transferred to the second processor in less than 30 seconds with no loss or distortion of information.

SYSTEM RELIABILITY

The measurement basis for system reliability is during periods of attended maintenance, 42 hours per week. Experience during the first six-month measurement period from March 13, 1962, to September 13, 1962, was as follows:

Component	% Availability
Processor 1 Complex	99.91
Processor 2 Comples	99.93
Drum #00	93.81
#01	99.17
#02	100.00
#03	99.70
#04	99.26

NOTE: Drum #00 was physically damaged by scoring of the drum surface. Over 90% of the period outage time for this drum was for installation and testing of the replacement drum.

Component	% Availability
Magnetic Tape Handler #00	97.76
#01	100.00
#02	94.56
#03	95.79
Message Register #1	99.84
#2	99.72
#3	100.00
#4	100.00
#5	100.00
#6	99.96
#7	99.99
#8	100.00
#9	99.99
#10	100.00
403 Printer	100.00
514 Punch	100.00

NOTE: This does not include outages caused by conventional features of the punched card equipment.

Component	% Availability
Processor 1 Automatic Typewriter	98.89
Processor 2 Automatic Typewriter	99.99
Teller-Registers	At least one was available per transceiver location 100% of the time.

Component	% Availability
Transceiver #1	99.65
#2	99.71
#3	99.51
#4	99.98
#5	99.32
#6	99.99
#7	100.00
#8	100.00
#9	99.65

During this period, over two million teller calls were processed. Of this total one may have been processed in error by the central processor. There is no certainty, due to its nature, but it has been charged arbitrarily against the central system.

Several printer errors have been detected, such as misprint on account number and failure to print transaction amount. These errors are not harmful, however, since the balance is echo-checked and is always correct. The printing errors are superficial errors only, as the internal processing is correct. If they are detected immediately, a reverse entry (correction transaction) negates the transaction and it is re-executed. If the error is detected at a later date, the degree of seriousness is determined and action is taken accordingly, i.e., write for the passbook, manually correct the journal tape on account number error (account number does not print in the passbook).

EMERGENCY OPERATING PROCEDURES

A worse case situation would be the complete loss of the entire Telefile System for an indefinite period of time. If such a catastrophe were to occur, the customer passbook would be processed by pen and ink, accepting the passbook balance as the correct balance. At the same time, the master magnetic tape files would be delivered to other Telefile System users. Arrangements have already been made to have these users reproduce these tape files into either punched cards or hardcopy listings if such a disaster should occur. This would provide sufficient information and machineable media to insure the continuity of bank operation.

Between a perfect operation and a complete disaster there are too many emergency configurations to enumerate. Such configurations can consist of the loss of one Teller-Register to the loss of a complete sub-system. Manual and automatic procedures have been developed to cope with virtually all configurations. As previously stated, should the on-line processor fail, control can be transferred to the second processor in less than 30 seconds with no loss or distortion of information. If prime power fails, an emergency power generator will re-supply power in 8 seconds. If a primary communication line fails, an alternate route line is switched in. Up to one drum, one tape handler, one message-register and several Teller-Registers can be out of commission prior to even considering the loss of an emergency nature. In summation, there is full cognizance of the things that can and will happen. Operating procedures and programs have been readied in advance.

SYSTEM BENEFITS

The benefits The Howard Savings Institution is realizing as a result of this revolutionary new system are already greater than anticipated. Some of the most important of these benefits are as follows:

1. Customers can bank at any teller window in any Howard Office with equal speed and facility.

2. The back office teller work is reduced over 90%.

3. Transaction processing time is reduced 30%.

4. Tellers are now able to handle two customers in the time previously required to handle one.

5. Controls are strengthened.

6. Operating costs are fixed.

7. Prime banking floor space is made available for more productive use.

8. Tedious bookkeeping and clerical jobs are eliminated.

9. Controlled management by exception is now possible.

10. Operating efficiency and flexibility previously unattainable are now a reality.

These are but a few of the significant benefits. There are many more. To bring them into focus, during the first half of this year tangible displaceable cost savings realized are over 50% of the total annual system operating costs. By the end of the first 18 months of operation it is anticipated that the entire cost on an annual basis will be recovered. These cost recoveries, coupled with both projected savings and the numerous intangible benefits, clearly prove the wisdom of the Howard's top management decision to automate.

THE FUTURE OF ON-LINE SAVINGS ACCOUNTING

At this writing, there are only three banks in the country processing savings accounts on an on-line, real-time basis. They are the Howard Savings Institution, the Society for Savings in Hartford, Connecticut, and the Union Dime Savings Bank in New York City--all Telefile users. Approximately ten banking institutions (nine savings, one commercial) have contracted for, but have not yet installed similar systems marketed by other manufacturers. The benefits of employing such systems are great. So too, are the costs. Therefore, the future of such computer installations in the savings industry will be on a cooperative basis.

Cooperative automation may take many forms. Inroads are being made for systems from the prime-user-operated data center to cooperative data center, where many banks underwrite the cost of the total system. As a matter of fact, one can predict that many commercial banks will justify more sophisticated data processing systems on the basis that they can automate savings accounting at the window. Probably, less than one hundred banks in the United States can afford an on-line, real-time savings accounting system individually. Yet, in the next five years, hundreds of banks will implement such an operation. This will come through the use of cooperative automation.

CONCLUSIONS

Banks unable to take advantage of automation will have to pass increased operating costs along to their customers in the form of higher service charges, higher interest rates, lower dividends or reduced earnings. These banks may find themselves at a serious competitive disadvantage with the larger more automated banks that perform their

services with greater efficiency.

1. On-line, real-time savings accounting operations are now a reality.

2. The benefits are significant but cannot be measured in terms of tangibles alone. Emphasis must be placed on the intangibles of faster and better customer service, more flexible operations and more timely reports.

3. Though the costs are high, the cooperative approach can enable many banks to share in these benefits.

Table I. Bank-to-Bank Communications

Office	Distance from Center	Type of Communication	No. of Teller-Registers
Main Office	1/4 mile	Data-Phone	13
Downtown	Same Building	Directly connected Multi-conductor cables	5
Bloomfield Branch	3 miles	Teletype	3
Springfield Branch	4 miles	Teletype	3
Vailsburg Branch	6 miles	Data-Phone	4
South Orange Branch	8 miles	Teletype	2
Wessex Branch (Caldwell, N.J.)	12 miles	Teletype	2

Table II. List of Teller-Register Transactions and Controls

TRANSACTION	CONDITION BAR	TRANSACTION BAR	RELEASE BAR
QUERY			
Query of Account Status		Account Number (16)	Operate (21)
PASSBOOK TRANSACTIONS			
Deposit		Deposit (15)	Operate (21)
Withdrawal		Withdrawal (14)	Operate (21)
Miscellaneous Deposit		Misc. Deposit (13)	Operate (21)
Miscellaneous Withdrawal		Misc. Withdr. (12)	Operate (21)
Dup. Deposit Slip Posting	No-Book (2)	Deposit (15)	Print Only (20)
Dup. Withdr. Slip Posting	No-Book (2)	Withdrawal (14)	Print Only (20)
Balance Print Only		Balance (10)	Print Only (20)
NO-BOOK TRANSACTIONS			
Deposit	No-Book (2)	Deposit (15)	Operate (21)
Withdrawal	No-Book (2)	Withdrawal (14)	Operate (21)
CORRECTIONS			
For Items Above			Correct (19)
INTEREST POSTING			
Interest Only	Interest (1)		Operate (21)
Previous to Deposit	Interest (1)	Deposit (15)	Operate (21)
Previous to Withdrawal	Interest (1)	Withdrawal (14)	Operate (21)
Prev. to a Misc. Deposit	Interest (1)	Misc. Deposit (13)	Operate (21)
Prev. to a Misc. Withdr.	Interest (1)	Misc. Withdr. (12)	Operate (21)
CHECKS & MONEY ORDERS			
List Print Only		List (11)	Print Only (20)
SUBTOTALS			
Deposit		Deposit (15)	Proof (18)
Withdrawal		Withdrawal (14)	Proof (18)
Miscellaneous		Misc. Deposit (13)	Proof (18)
Miscellaneous Withdrawal		Misc. Withdr. (12)	Proof (18)
List		List (11)	Proof (18)
SUPERVISORY CONTROL			
Entry of Teller Number		Teller A/B (24)	Operate (21)
Grand Total		Stop/Caution - Grand Total (23)	Proof (18)
Entry of Stop Condition*	Stop (4)	Deposit (15)	Operate (21)
Removal of Stop Condition*	Stop (4)	Withdrawal (14)	Operate (21)

* Supervisor must set Stop/Caution - Grand Total Switch (23)

Fig. 1. The Teller-Register, an important part of The Howard Savings Institution's on-line banking system, brough about considerable improvement in customer service. This teller posting machine was designed specifically for on-line, real-time data processing as opposed to an adaptation of a conventional window machine.

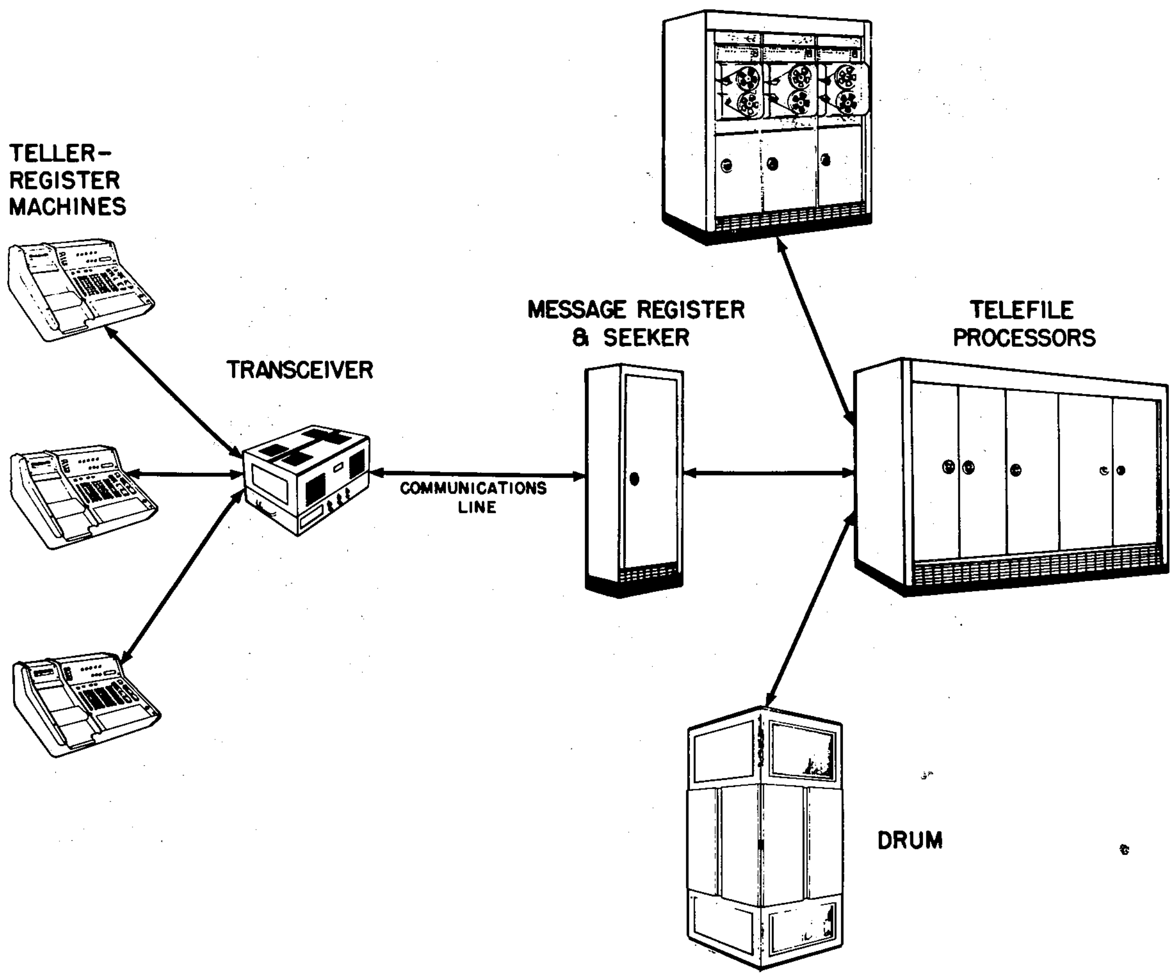

Fig. 2. Simplified Block Diagram of Telefile System

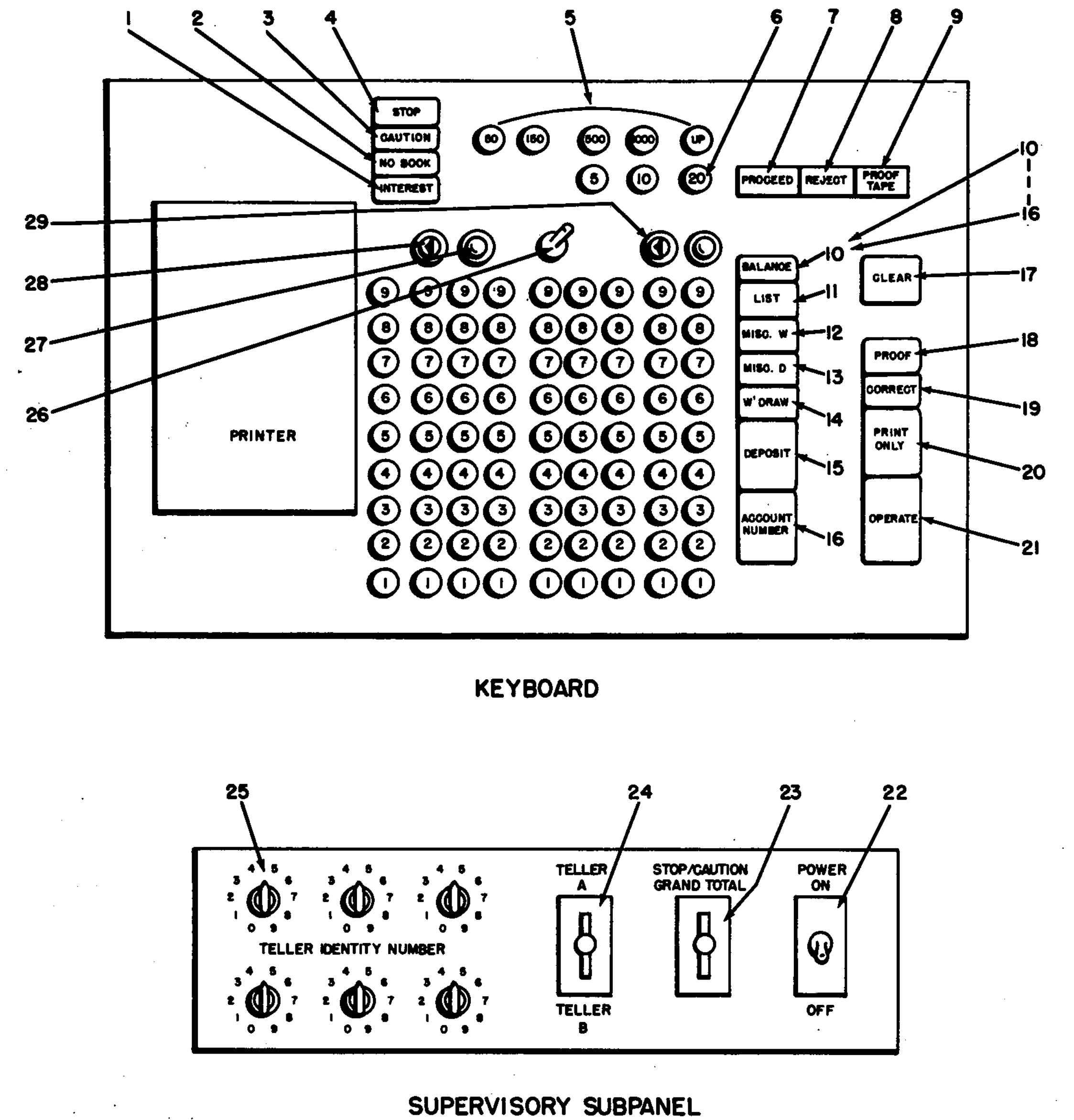

Fig. 3. Teller-Register Keyboard Layout

THESE ARE THE CONTROLS:

CONDITION BARS	
Interest	(1)
No-Book	(2)
Caution	(3)
Stop	(4)
CHECK CONDITION CONTROLS	
Check-Range Key/Lamps	(5)
Check-Days Buttons	(6)
OPERATION CONTROL LAMPS	
Proceed	(7)
Reject	(8)
Proof Tape	(9)
TRANSACTION BARS	
Balance	(10)
List	(11)
Misc. W.	(12)
Misc. D.	(13)
Withdrawal	(14)
Deposit	(15)
Account No.	(16)
RELEASE BARS	
Clear	(17)
Proof	(18)
Correct	(19)
Print-Only	(20)
Operate	(21)
SUBPANEL CONTROLS	
Power On-Off	(22)
Stop/Caution Grand Total	(23)
Teller A/Teller B	(24)
Numeric Switches	(25)
Upper Row for Teller A;	
Lower Row for Teller B.	
TELLER IDENTIFICATION CONTROLS	
Teller Position Switch	(26)
Lock Indicator Lamps	
for Teller A	(27)
for Teller B	not numbered
Front Panel Locks	
for Teller A	(28)
for Teller B	(29)