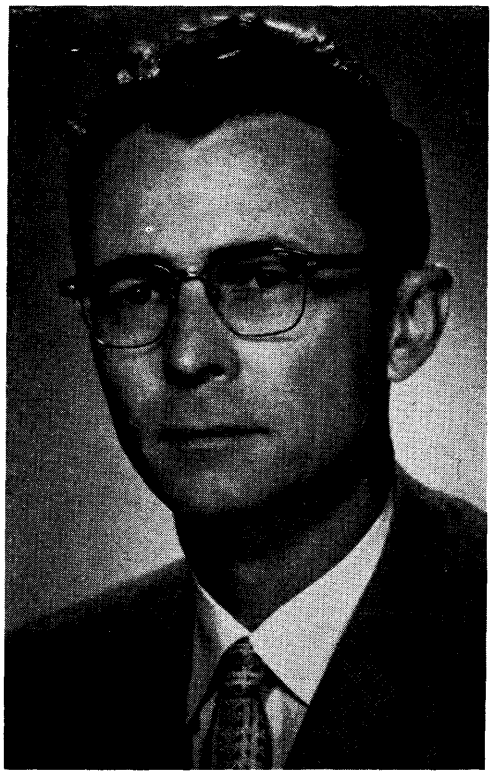
General Electric in the Field of Business Data Processing



HARRISON TELLIER

Mr. Tellier has had more than twenty years experience in the gathering and anlysis of data. He is presently Manager - Data Processing Operation at General Electric's Hanford Atomic Products Operation, where data processing and computational services are performed by a single installation for the six General Electric components which operate the United States Atomic Energy Commission Hanford Works. He is a member of the Inland Empire Chapter NMAA at Spokane.

General Electric is unique in that it is one of the major users of data processing equipment in the world and also a manufacturer of data processing components and systems. General Electric's faith in the future of such equipment is evidenced by R. J. Cordiner's Testimony on Automation before the Sub-committee on Economic Stabilization, Joint Congressional Committee on the Economic Report on October 26, 1955.

Mr. Cordiner, while speaking specifically on experience with computers in General Electric, stated as follows:

"If there is one idea which I would like to leave with you today, it is that the computer represents an important extension of the powers of the mind of man. When the history of our age is written, I think it will record three profoundly important technological developments.

"Nuclear energy, which tremendously increases the amount of energy available to do the world's work.

"Automation, which greatly increases man's ability to use tools and computers, which multiply man's ability to do mental work.

"Some of our engineers believe that of these three, the computer will bring the greatest benefit to man."

In 1956 a new Computer Department was formed in General Electric. The home office of this department is in Phoenix, Arizona.

The fact that such a department has been just recently formed does not mean that this is a new field to General Electric. In the past, various components of the Company have created an impressive series of digital and analog devices. These products have arisen out of specific customer needs, and are well known in the fields where they have application. The General Electric computer activity is perhaps less well known to business and industry in general because of this concentration on the more specialized and difficult problems, and on many projects still classified as secret by the armed forces.

However, recently the Computer Department has entered into two contracts which are well known in the business world. The first is with the Bank of America to manufacture a computer called the Electronic Data Processor.

The second contract is with the National Cash Register Company to design and manufacture the electronic portion of a data processing system known as the NCR-304. The NCR-304 will be a transistorized data processing system, consisting of a central electronic computer, magnetic core memory, magnetic tape storage units, storage medium converters and various high-speed input and output units.

The basic aim of the Computer Department is to build equipment which is custom-designed to the fulfillment of over-all customer needs, rather than forcing the customer to compromise his own operation and warp his procedures to meet the limitations of the data processing equipment he can purchase.

As previously mentioned, General Electric is one of the major users of data processing equipment in the world. Just the fact that a company is large does not necessarily make its data processing operations interesting to other users of such equipment. It is the nature of General Electric's business and form of organization that make such operations interesting.

General Electric is one of the most diversified companies in the world, with some 350 distinct product lines and about 3 million catalog items. This wide range of products places the Company in direct competition with about 600 other principal companies. Obviously such variety of products presents a wide variety of data processing problems.

Of even greater interest is General Electric's form of

organization. The Company is decentralized, both in structure and geographically. The Company is organized into over 100 product departments. These product departments represent 159 plants in 123 communities.

Under our form of decentralization, management decisions take place at the local level. Although guidance and counsel are provided to all decentralized operations by the Company-wide functional services, the identification and solution of local problems are made within each individual organization. Thus, the problem of paying some two hundred thousand employees is not solved by a few men in New York, but is a problem to be solved by over 100 Managers of Finance throughout the Company.

I have mentioned our product lines and form of organization to illustrate that although we are a large company, our data processing problems are comparable to hundreds or perhaps thousands of American businesses.

In the area of developments made by our Company in the use of equipment, I would like to place particular emphasis on the use of magnetic tape systems such as the IBM 700 series machines and the UNIVAC. I do not wish to imply that the perhaps less glamorous punched-card installation does not have its place, or that it's on its way out. However, I believe that true integration of data processing must start with these larger systems.

At the present time, there are 11 large-scale magnetic tape processing systems in use within General Electric:

One is a Sperry-Rand UNIVAC
Two are IBM 702's
One is an IBM 705 and
Seven are IBM 704's

In many of these cases, the departments operating these machines are performing services for other General Electric departments on a service bureau basis. The Computer Department at Phoenix is able to provide services to anyone. It was recently announced that the Computer Department has been awarded a contract by the Army to operate a computer center at the Army Ballistics Missile Agency at Huntsville, Alabama.

In the IBM family of machines, the 702 and 705's were designed for business data processing, while the 701 and 704's were designed for engineering computation. However, there is an increasing trend to use these machines for both types of work. Our 702 at Hanford was installed to perform engineering calculations as well as business systems.

The reverse situation is that the Flight Propulsion Laboratory Department's 704 at West Lynn is being used by the Medium Steam Turbine, Generator and Gear Department, for payroll preparation in addition to the normal engineering calculations. Similarly the recently installed IBM 704 at Heavy Military Electronic Equipment Department will be used for both types of work.

As time goes on, I believe we will see more use of data processing systems for both commercial and engineering work. Often the workload of either type is not sufficient for economic justification of a large-scale machine, but pooling of the workload will justify a large-scale processor.

The distinction between data processing and computing, in my opinion, is rapidly disappearing for two reasons:

1. More and more mathematical techniques are being brought to play in areas which used to be considered as simple accounting operations. Any properly staffed data processing organization should include at least one mathematician.
2. What once were considered purely mathematical problems are becoming a part of data processing systems where large volumes of actual data are used rather than theoretical figures or averages.

As my topic today is Business Data Processing, I would like to call to your attention that General Electric was America's first commercial user of a large-scale electronic data processor.

Appliance Park was organized into five separate product departments. In 1953 many existing operations were to be moved into the new Appliance Park. This seemed an ideal time to install automation in office work before employing a large number of clerical workers who would have to be transferred to other jobs if automation did come later.

The Remington-Rand (now Sperry-Rand) UNIVAC system was available in time to be phased in with the planned expansion at Appliance Park. A feasibility study indicated that the UNIVAC system had the capacity to handle the planned workload and its rental cost could be justified on the basis of clerical savings alone.

In 1953, General Electric Management elected to assume a calculated risk to become America's first commercial user of a large-scale electronic data processing system.

In 1954, the UNIVAC was installed and during that year the payrolls for the five product departments were converted to the new system. At the present time, the UNIVAC is being used for 23 different applications.

Within General Electric there are many others who hold a position similar to mine. Naturally there will be many differences of opinion as to methods of operation, techniques and approach to the problems of data processing. As I do not feel equipped to speak for all of us as a group, my report on data processing must be based to a large degree on our experiences at Hanford.

General Electric, under a contract with the U. S. Atomic Energy Commission, operates the Commission's Hanford Works. The operation at Hanford is known as the Hanford Atomic Products Operation and is made up of six separate departments or operations. One of these operations is known as Relations and Utilities. Within the Financial Operation of Relations and Utilities is the Data Processing Operation.

This data processing operation at Hanford for six years has been a service organization, offering computational and data processing services to all functions of the plant. Such services are performed just like a service bureau—every dollar spent for salaries, equipment rental, supplies and services purchased from other service organizations must be billed to our customers. Our operation has many of the problems of an individual business.

In addition to maintaining a relatively uniform level of billing, we must also provide a reasonable level of service which from time to time has required us to replace equipment with new and more powerful machines. In 1955 we replaced the Card Program Calculators and related punched card machines with one of the first IBM 702 Electronic Data Processing Machines. As our 702 has been loaded for some time, we are considering its replacement by an IBM 709.

During the two-year period the 702 has been in operation at Hanford, close to 200 programs have been written for the routine processing of scientific and commercial data, and some 400 special requests for non-routine data processing and computation have been filled. It is estimated that some one-half million program steps have been written by our staff during this period.

Except for maintenance, our machine is in constant use around the clock Monday through Friday, and only twice during the past 23 weeks has the machine been turned off at the end of Friday. For 7 consecutive weeks, the machine has run constantly. Scheduled and unscheduled maintenance averages about 9% of the total time.

The cost of our services to our customers is about a quarter of what corresponding services would cost at a commercial service bureau. Costs are low because of the high rate of utilization which takes advantage of the reduced rates for aditional shift utilization.

Our 702 workload is divided roughly 20% for engineering calculations and 80% data processing. The following are a few of the routine data processing applications currently in operation on the 702:

- Weekly payroll and salary distribution
- Monthly payroll and salary distribution
- Work order cost
- Stores inventory control and material distribution
- Services cost distribution
- Property accounting
- Classified document accountability
- Vehicle analyses
- Personnel exposure control
- Radiation monitoring instrument calibration
- Reactor operating calculations

As each individual data processing application should reflect the peculiarities of the business, there would be little interest in any explanation of how these systems work. What I believe will be of interest to you as machine accountants would be some description of the data processing techniques which have been developed at Hanford. The following are techniques which I would like to review in the sequence of their development.

The IBM 702 system at Hanford was the first such system installed to perform engineering calculation as well as perform the normal commercial applications such as payroll, etc. It was recognized in 1955 that the problem of conversion from punched card machines to the 702 could be accomplished only if an automatic programming system were available. Such a system was developed at Hanford known as SCRIPT: The word SCRIPT stands for "Scientific and Commercial Subroutine Interpreter and Program Translator." SCRIPT is a symbolic programming system which permits the programmer to use symbolic operation codes and addresses, and to automatically incorporate in his program long sequences of machine instructions from single handwritten instructions. SCRIPT was the first programming system of its kind in the country. It has been immensely successful by reducing the programming time on some applications by as much as 80%.

Almost every major application requires a re-arrangement of the data at some point in the processing, either for reporting purposes, or for matching the data with data in another file. About 20% of our 702 processing time is devoted to sorting tape records.

Soon after installation of the 702 in 1955, it was recognized that to write a special program at every point in processing where the data is to be put in a different sequence would be wastful of programming effort. Sorting on an electronic data processor is a very complex logical process.

Late in 1955, a Generalized Sort Routine was developed that would sort any file, regardless of the data it contains. To use this generalized sort routine, all the programmer has to do is supply certain parameters which specify key elements in the sorting process. The parameters have the effect of particularizing the generalized routine to the specific application. Once the parameters have been supplied, the generalized routine may be inserted at any point in the processing where re-arrangement of data is required. The parameters which define a particular sort are the length of the records to be sorted and the size and position in the record of the control fields which establishes the new sequence. That is all that is required to "program" a sort pass.

SCRIPT, which was the first automatic programming system developed at Hanford, was a vast step forward. However, as with many firsts, it had certain limitations—the major limitation was the fact that the programmer had to think too much in terms of the characteristics of the 702 system.

In 1956, a new system to be known as OMNICODE was being developed to succeed SCRIPT. This was to be a system by which the programmer could define the problem in English letters and names, rather than symbols. However, when the installation of an IBM 650 Magnetic Drum Data Processor was projected in

1956 to supplement the 702, efforts were re-directed to develop OMNICODE to produce 650 programs. OMNICODE has been in use since September 1956 and has been used with considerable success by engineers and scientists on a self-service basis.

The use of OMNICODE is rather unique in that cards punched from OMNICODE written programs are processed on the 702 to prepare 650 instruction cards.

Even with the SCRIPT programming technique, we found the following unfavorable facts to be true about using a large-scale electronic data processor:

1—Manpower required to install new applications was measured in man-years.

2—All programming other than the sorting of data by use of the generalized sort routine was "hand-tailored" to meet the demands of the application.

So much time was invested in hand-tailored programs that it was too expensive to create "models" of proposed systems for test and demonstration.

3—Programs for specific applications were inflexible after original definition. Changes took too long to make. It was difficult to prepare one-time, non-routine analyses from the tape records.

4—Too much emphasis had to be placed on communications with the electronic machine at the expense of concentration on the customer's problem.

5—Last and perhaps the most difficult problem is what to do when the existing electronic data processor is completely loaded as our machine is. We feel that the installation of a second machine of the same type is too expensive and that it is more economical from an operational standpoint to install a larger system.

The number of active routine 702 passes is a fluctuating number, but as of last week there were 235 of them. To reprogram each of these passes for a new machine would be an impossible task while still attempting to maintain a constantly growing program.

These unfavorable facts which I have mentioned have been a source of concern to us in data processing as well as management.

The success of the generalized sort routine lies in the fact that it is many times easier and faster to supply parameters than it is to write special sort programs. Because of the success of generalized sorting, it is only natural to expect that the same benefits can be derived by identifying and generalizing other broad data processing operations of this type.

In evaluating what we were actually doing on the 702, we determined that there are four basic functions in data processing:

1—File Maintenance, which includes creation and and updating of unit records

2—Calculations

3—Sorting

4—Report preparation

Sorting had already been generalized. Calculation is the operation that distinguishes one data processing system from another and at the moment does not appear to be subject to generalization. Actual calculations (add, subtract, multiply and divide) at Hanford represent less than 20% of operating time.

Early in 1957, projects were begun to create generalized routines for the two remaining functions—file maintenance and report generation. The Generalized File Maintenance Routine was completed in June and the Generalized Report Generator Routine was completed in May. We are now equipped with generalized routines for three of the four functional areas of electronic data processing.

Before description of the Generalized File Maintenance Routine, three terms must be defined. A reel of tape upon which is recorded information about each person on the payroll will be used as an example. Such persons are identified by employee number and are recorded on the tape in sequence of man number.

This reel of tape is referred to as a "tape data file." The collection of information about a specific employee is a "unit record" and a specific piece of information on each employee, such as birthdate, is a "field."

The Generalized File Routine will create, revise and update any tape data file. Fields in all unit records are established, altered or eliminated by feeding a Field Format Change card into the processing. Unit records are established or eliminated and particular fields of particular unit records changed by feeding a "Change Notice" card into the processing.

During processing, the 702 generates the particular machine instruction pattern required to make the defined field changes to each unit record in the file and also to update specific fields of specific unit records.

As part of processing, the 702 prepares a change report listing the before and after status of every field that has been affected.

Under conventional programming for file maintenance, the addition of new fields or changes to field lengths is a most difficult task. Equipped with a Generalized File Maintenance Routine, it is possible to allow information in data files to grow and it is not necessary for complete definition of record format before the application goes into operation. All of you who have worked with punched card equipment have experienced the unhappy situation when it is discovered that a card field was not large enough or an additional field had to be added. On electronic machines the situation is not much different, but we have found that Generalized File Maintenance solves this problem.

One of the great selling points of mechanized data processing is the ability to prepare special reports as required, that is, on an exception basis rather than routinely in case they are needed. This concept is relatively simple, using punched card machines. In most cases, a collator pass will select the class of unit records desired and some existing printer panel will list the selected unit records.

Unfortunately with magnetic tape records, the comparable operation of selection of records and listing such selected records must be programmed. Until we had developed the Generalized Report Generator, we

would have, on the average, two analysts doing nothing but programming special reports from data tapes. Even worse than the manpower requirement was the necessary elapsed time required to produce such special reports.

With the Generalized Report Generator, the task of programming any report from a tape data file is reduced to filling out a form. The top part of the form is just like a conventional punched card printer layout sheet. Page and columnar headings are printed in the appropriate type wheel grids. Type wheel printing positions for listed detail are indicated by x's.

The bottom part of the form is a simple designation of what unit records are to be selected from the tape file, where each desired piece of data is located in the magnetic tape record and where the data is to print on the report.

From each such output definition form a packet a cards is punched. Packets for desired reports, the Generalized Report Generator Routine program and the magnetic tape source file are fed into the electronic processor.

As the first step in the processing, the report generator routine reads the output definition cards and generates the particular instruction pattern required to prepare each report. The generated program then takes over, selecting out of each record in the file those fields required and placing them in an auxiliary record in a form suitable for printing. When the file has been completely processed, the auxiliary records are printed on a line printer.

The important feature of the Report Generator is its ability to produce any reasonable number of independent reports from a given file on a single machine run. During processing, the Report Generator program generates a separate, independent program for each packet of cards representing a single report. As each record in the file is read, the generated programs are executed, one after another, so that the various reports are generated essentially simultaneously. Certain additional features of the Generalized Report Generator are worth noting:

1—It is not necessary to list information out of every record in a file; it is possible to list information out of selected records only.
2—Totals may be printed of designated fields and provision is made for taking up to 15 levels of totals, including final totals.
3—Pages may be numbered.
4—Record counts taken.
5—Elimination of zero-printing.
6—Group indication.
7—And last but not least, the generation of records for use in subsequent processing.

In general, what we have done is simulate on a general purpose machine three special purpose machines. In the case of the Report Generator we have created a huge collator and printer without the problems of panels to wire.

Other users of large-scale processors have been developing techniques similar to certain of our generalized routines. However, to our knowledge we are the first to have the three routines of sort, file maintenance and report generation in actual operation.

Now that these routines are available, I would like to review what they have already done to our operation at Hanford:

1—There has been a substantial reduction in the time and cost of installing data processing applications.
2—It is possible now to develop integrated systems without firm definition of all input and output before the systems are placed in operation. Such systems are highly flexible.
3—Special analyses may be prepared quickly and at low cost.
4—The conversion of existing operation may be made to a higher level of data processor without complete re-programming of all present processor passes individually and at the same time take full advantage of the characteristics of the new processor.

Payroll preparation can be taken as an excellent example, as this so often is the first application placed on a new processor. Our original 702 payroll, developed in 1955, uses 22 separate handwritten programs containing a total of 50,000 instructions and required *six man-years* to develop. We are in the process of developing an integrated personnel accounting system for which the payroll preparation portion is estimated will require at most 3000 handwritten instructions from *six man-months.*

This improved ratio of man-month to man-year as the unit of measure is the result of having to hand program *only* the calculation passes. The organization of a revised weekly payroll for Hanford shows one major calculation pass and three minor calculation passes. There are nine generalized passes, six of which use the Report Generator, two File Maintenance and one sort. This is a total of 13 passes as contrasted to 22 and only four of them hand-programmed.

In conclusion, I would like to say that after 27 months of using a large-scale processor, we still feel that such machines will be the answer to the paperwork and communications problems that have held back progress. The most important point is that at last we see the way clear to divorce for the first time the two problems of definition of the data processing application and the machine operations. As machine accountants we have always felt that the proper definition of the system is really the major problem. But as long as the characteristics of the data processing equipment have had to be considered in the system's definition, it has been the equipment that has been blamed for poor systems. Every effort must be made to provide an atmosphere in which management can define communications system with complete disregard for peculiarities of the data processing equipment used. I believe we have overcome this problem and at last will be able to give management what they gambled for when they approved installation of electronic data processing.