Technical Information

For Release:
Tuesday,
March 5, 1968

International Business Machines Corporation
Data Processing Division
112 East Post Road
White Plains, New York 10601
Bert Reisman, Manager of Information

E. F. Rodgers
IBM Corporation
914 WHite Plains 9-1900

COMPUTER-AUTOMATED LABORATORY EQUIPMENT USED IN IBM/VARIAN DEMONSTRATION

The automated laboratory demonstrated at the 19th Annual Conference on Analytical Chemistry and Applied Spectroscopy, March 4 - 8 in Cleveland, is the result of a joint study by International Business Machines Corporation and Varian Associates, Palo Alto, California. The laboratory is designed to help researchers conduct their experiments faster and more accurately than ever before by linking their instruments directly to a digital computer and time-sharing its processing power. Included in the demonstration is an IBM 1800 data acquisition and control system and an array of analytical instruments used in chemical research. A description of the equipment follows:

IBM 1800 DATA ACQUISITION AND CONTROL SYSTEM

The IBM 1800 data acquisition and control system can regulate traffic signals, monitor an assembly line, control a steel-making process and handle a host of other manufacturing and research jobs.

The system can gather information from measuring devices at rates up to 24, 000 signals a second and analyze it rapidly. It can generate the commands required to control production or test equipment. The system also processes routine computing jobs and produces management reports.

The IBM 1800 can be tailored to meet the specific needs of the user. It can be used independently for process control and data acquisition or, coupled with IBM System/360 Models 25, 30, 40, 44, or 50, to serve as part of a more powerful computing complex.

- more -

Ten different combinations of speed and memory size are available, together with a wide selection of peripheral devices. These devices range from magnetic disk drives with removable disk packs to printers, plotters and punched-card equipment.

Two different central processing units are available, one of which can accommodate magnetic tape units. Both processing units are equipped with two- or four-microsecond main-memories which come in five storage capacities ranging from 4, 000 to 32, 000 words. Each word contains two checking bits and 16 data bits -- the equivalent of four decimal digits or two letters.

VARIAN M-66 MASS SPECTROMETER

The Varian M-66 mass spectrometer is a double-focusing cycloidal mass spectrometer designed to handle a wide range of materials and to produce accurate data immediately in an easy-to-use form.

The M-66 can be used in structural organic chemistry, or in more fundamental work such as determining ionization and appearance potentials, and making quantitative measurements.

VARIAN A-60 NMR SPECTROMETER

The Varian A-60 nuclear magnetic resonance (NMR) spectrometer is used for qualitative and quantitative analysis of chemical compounds, analysis of mixtures, isotope identification, and biomedical systems studies.

VARIAN AEROGRAPH GAS CHROMATOGRAPH

Gas chromatographs provide chemical analysis, trace analysis, and separation of multi-component mixtures. This latter capability makes the gas chromatograph an ideal companion to larger analytical instruments. In this application, it is used as a "preparative instrument, " supplying the effluents later used for concentrated study.

The two units being shown are the Model 202, a thermal conductivity gas chromatograph, and the Model 1200, a hydrogen flame ionization gas chromatograph.

# # #

DPD 340 B-C2
3/5/68

Text of a talk to be presented by C. H. Sederholm at the Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy, March 4-8, 1968

IBM

A COMPUTER SYSTEM FOR AUTOMATION OF THE ANALYTICAL LABORATORY

C. H. Sederholm
P. J. Friedl
T. R. Lusebrink

As was demonstrated at last year's conference, a large number of instrument manufacturers and a reasonably large number of instrument users have already taken the first steps toward automation. They have automated individual instruments with remarkably good results. A few of the more advanced laboratories have already started to automate entire laboratories. Surely widespread laboratory automation appears to be just around the corner.

In general, instrument automation has taken place by devoting and interfacing a small computer to an individual instrument. When one considers automation of an entire laboratory containing many such instruments, the question arises as to the advisability of acquiring many small computers, one attached to each instrument, or whether it would be preferable to have a single computer system which serves the entire laboratory.

There are three reasons for attacking the problem of laboratory automation in terms of successive instrument automation; i.e., using one small stand-alone computer per instrument. The most important consideration is one of isolation. Each individual user does not want to concern himself with any problems but his own. He does not want malfunctions of other instruments or their interfaces to jeopardize his interface to the computer system.

The programming considerations associated with his own instrument are sufficiently complex that he doesn't want to worry about other users' programming problems too. He is rightfully afraid of being forced to factor parts of his programming requirements into general purpose programs which serve the entire laboratory (programming by committee).

To a somewhat lesser extent, availability of the computer system encourages one to favor multiple computers, one per instrument. Waiting for a few seconds, or at most one or two minutes for computer facilities to become available would not be intolerable, but the prospects of having to schedule one's experiments and schedule the use of the computer facilities long in advance is very unpalatable to most users. In general, users prefer fewer facilities which are routinely available to larger facilities available by appointment only. Hence, the desire for availability tends to favor multiple instrument automation over laboratory automation.

Finally, the question of cost should be considered. A computer system which is capable of automating an entire laboratory will, of necessity, be more expensive, even in its most stripped-down version, than a computer system capable of automating a single laboratory instrument. Therefore, when one is taking the first step toward laboratory automation; that is, the automation of a single instrument, there is a very strong tendency to automate that instrument using one small computer. "Worry about laboratory automation after the first few instruments have been automated." However, if one considers automation of the entire laboratory as an ultimate goal

from the beginning, many of the considerations below imply that a single, shared, laboratory computer would give more performance per dollar.

The automation of a laboratory by the use of a single, shared computing facility has several outstanding advantages over multiple automated instruments. The purpose of automation is not only to acquire data but, also, to control the instrument during the data cquisition step, to process the data which has been acquired, to standardize it, to compare it against known parameters (e.g., to compare an unknown spectrum against a file of spectra of known compounds for identification), and finally to present the results in a form which is usable to the experimenter. The data acquisition and control steps can very often be adequately performed by a small computer. However, the data reduction steps, the comparison with data files, and the presentation of results in usable form, often require much larger compute capabilities. One solution to this is to record the raw data which has been acquired with a small stand-alone computer on a recording medium such as magnetic tape or punched paper tape. This data may then be processed on a large computer when time is available. However, if the raw data is at all voluminous, and magnetic tape must be used, the cost of the magnetic tape drive relative to the cost of the small computer can become very high. Hence, going this route, one is encouraged to minimize the quantity of data taken, often resulting in less precise results. Likewise, the turn-around time on very large computer facilities still is much longer than the individual investigator would like to wait between epochs of his experiment. That is, if he could have the data from the last epoch back

immediately, he could start planning the next experiment and executing it.

The larger shared laboratory computer then has three impelling reasons for its use in automation of laboratories. It can have sufficient core and sufficient processing capabilities to do reasonably large-scale data reduction and file look-up tasks at a modest expense per instrument attached to it. By having one central computer facility, one may take advantage of the fact that most analytical and spectrographic instruments have a low duty cycle; that is, much of the time is spent in sample preparation or with the instrument completely idle. Using a centralized computer facility, the computer need not be idle while an individual instrument is idle, since the computer can be used for other jobs. Finally, by using a larger centralized laboratory computer, the input/output devices may be shared by all the users of the system. Hence, for a given cost per user, each user may have the use of more sophisticated input/output and peripheral storage devices. E.g., it is possible to have a large disk file which can hold hundreds of thousands of words. This is most useful in the analysis of spectroscopic data. It is possible to have a line printer which increases the amount of printed output which a user can request and receive in a reasonable period of time. It is possible to have a card reader such that reasonably massive input data can be prepared in a reasonable format for input into the computer system. If one attached these input/output devices to each of the individual computers in a multiple instrument automation situation, the total expense for the input/output devices alone would be far greater than the cost of the laboratory automation facility.

The centralized laboratory automation facility seems to have several

important advantages. If the problems of unavailability and non-isolation of individual users could be successfully overcome using the centralized computer facility, it would seem to be a far more reasonable solution to automation of a laboratory. We believed that these problems could be overcome, and hence set out to construct a system designed to service multiple, isolated, asynchronous users, each of which had demands for closed-loop control, large-scale data reduction and file look-up programs, as well as smaller data acquisition and control programs. The specifications of this system were aimed at automation of a laboratory containing a group of analytical and/or spectrographic instruments. This system has been completed to a point, and has been demonstrated in the exhibition hall during this week.

The system features complete program independence and complete hardware independence of each instrument from all others. An application program for one instrument need in no way to take into consideration other programs running in the system simultaneously with it. An application program need not be modified as a result of a change in the total instrument mix attached to the computer system. Each instrument has its own individual data path directly to the core of the computer via a pair of data channels, so that there need be no sharing of the interfaces between various instruments.

This system uses an IBM 1800 computer. It is possible to operate this system on a computer with 16K words of core. However, it is only economically practical if the laboratory to be automated is sufficiently large to support a 24K or 32K word machine.

A new I/O device has been designed and built, specifically to support this laboratory automation system. This is a digital multiplexer channel for the 1800 which provides up to 32 discrete data paths between the laboratory and the core of the computer. This device allows data to be acquired from, or presented to, the individual instruments in a demand/response mode with a minimum of computer overhead. The reduction in computer overhead increases the allowable total data acquisition rate from all instruments by more than an order of magnitude. The demand/response mode of data acquisition is of great value in that it allows the instrument to say when data is available rather than the computer saying that data should be presented now. An example of the usefulness of demand/response data acquisition is in acquiring data from an infrared spectrometer which has automatic scan suppression. Since the scan rate is a function of the first derivative of the absorption, data acquired at equal intervals of time would not be at equal increments in wave length or wave number. However, with a demand/response interface, the instrument could be run with scan suppression, anddemands to take data could be made by the instrument at equal wave length or wave number increments.

The present operating system, known as the 1800 Laboratory System, requires approximately 13K of core which, with additional work, can be reduced. The rest of core (known as variable core) is broken up to 512-word blocks, called pages, which are assigned on a dynamic basis.

When a user wishes to sign on the system, he types a command at the 1800 console typewriter. This command causes the system to load a master program associated with that instrument into variable core and put that program into execution. Depending upon the size of a user's program, it may require one or multiple pages of variable core. If the program requires more than one page of core, the system loads the program into any available pages. The pages used need not be adjacent to each other.

This master program activates various buttons and indicator lights at the instrument interface. By pushing one of several buttons, the user may ask for one of several application programs associated with his instrument. When the master program recognizes which program or routine the user wishes to run, the master program asks the system to load the appropriate program into variable core and set it into execution. The master program then exists, releasing the core that it occupied back to the variable core pool. The specific application program may acquire parameters for the run by interrogating user-operated thumb switches at the instrument interface. Then the application program continues with its execution. At any time, an application program presently in execution may be aborted from the instrument interface. If the application program is very lengthy, the system provides facilities for chaining smaller segments together. The system also provides

facilities for disk-resident subroutines which only occupy core when they are called.

Many system subroutines are provided, and these are also loaded into pages of variable core when needed. If no user program presently in execution is using any of the subroutines on a given page, that page is returned to the pool of variable core. Because of this dynamic allocation of core, it is possible to run the system to execute a program requiring 15K of core, and a few minutes later, with no changes to the system, have the same computer executing 15 different programs simultaneously, each of which occupies 1K of core; i.e., a given instrument is not restricted to a fixed partition of core!

Presently, users' programs may be executed with two different priorities -- a foreground priority used for real-time data acquisition and control, and a middle-ground priority used for non-real-time calculations. All users' programs operating with foreground priority have the same priority. The system provides each program, in turn, with a 5 millisecond time slice for starting or stopping the data acquisition process and/or "massaging" current data and/or feedback control of instruments. Time slicing is also done for middle-ground programs, only the time slices here have 100 milliseconds. Only when there is no foreground program with processing to be done do the middle-ground programs receive time slices. If a foreground program finishes acquiring data and, hence, has some foreground processing to do (e.g., for instrument control) any active middle-ground time slice is interrupted to process the foreground program.

Once a data acquisition activity has been initiated in a foreground program, the new digital multiplexer channel provides for the activity at instrument rates. This activity procedes independently from and concurrently with the executions of time sliced programs.

Systems capability for handling background batch processing job streams are present. However, facilities for such processing have not been implemented yet.

Data stored on the disk file is referenced in terms of logical tapes. That is, each application program may be written as if there were available to it a large number of magnetic tape drives. Hence, an application program may ask the system to present it with 50 words of data off of logical tape SPEC starting with the 31,352nd word on the tape.

As far as the application programs are concerned, the entire system is oriented toward performing a variety of tasks, most of these tasks being associated with input/output. In general, it takes 12 words of a user's program to specify a task for the system. These tasks are accepted by the system and performed as soon as possible. Multiple tasks may be queued for a single input/output device; e.g., the line printer may have a current task in execution while 5 other tasks are waiting for the line printer to be free. A given application program may have several tasks outstanding simultaneously. For instance, a given application program may instruct the system to (1) acquire a block of data from a given instrument, (2) read a card in the card reader, (3) print a line of the line printer, (4) light a light at the

user interface, and (5) write a block of data on a logical tape. These tasks would be given to the system sequentially, but all five tasks would be set into execution before the first one was completed.

A partial list of tasks the system is capable of performing will demonstrate the power of the system:

1. Read a block of 50 words from logical tape DATA starting at word 5273.
2. Read the thumb switches at the NMR spectrometers.
3. Read 30,000 words from the gas chromotography columns at a data rate of 50 points per second and write these onto a logical tape called GCDAT.
4. Load and set into execution a smoothing program (SMUTH) in the foreground mode. (A common data area between the calling program and the called program is provided).
5. Release all core and system subroutines that this program presently occupies.
6. Read all the cards in the card reader and put them onto the logical tape FILE.
7. Print the logical tape OUT on the line printer.
8. Dump core presently occupied by program SMUTH.
9. Plot logical tape PLOT on the plotter.
10. When push button X is depressed, transfer control to entry point ENT in this program.

Having defined and constructed this system for use in the automation of an entire laboratory, we wished to investigate its actual usefulness. As

a result, we engaged in a joint study venture with Varian Associates of Palo Alto, California. Varian interfaced three of their analytical instruments to this system. These instruments included an M66 mass spectrometer, an A60 NMR, and a pair of aerograph gas chromatography columns. Multiple application programs for each of these instruments were jointly specified and written by Varian and IBM. Varian has had previous experience with the automation of individual instruments. Many of the programs previously written for stand-alone systems were modified and augmented to run on this system. For the A60 NMR, the following programs have been written: data acquisition and control, data smoothing, data catting, data presentation, and field homogeneity adjustment. The following programs have been written for the mass spectrometer: data acquisition and control, peak identification, and compound identification based upon the position of the five most intense peaks in a mass spectrum. For gas chromotography, a data acquisition and control program has been written. Peaks are detected, their areas are resolved and calculated, and retention time and areas are reported.

Once the system was defined and operational, the coding of the application programs went extremely rapidly. For the programs so far written, there has been no indication of unfavorable interaction between programs concurrently running on the system. Several of the programs which have been written, in particular for the mass spectrometer, could not have been written for a small stand-alone computer with no peripheral storage devices.

In conclusion, we believe that we have been able to produce a system for use in laboratory automation which provides each individual user the isolation, the availability, the real-time control responsiveness, and the price per instrument associated with multiple computers, one per instrument. In addition, this system provides sophisticated input/output devices, disk files, and a large amount of support for the I/O devices and the files, which would not be available on a small computer associated with a single instrument. Each user, then, has the impression that he has a large computer attached to his instrument and completely at his disposal.

#####