Technical Developments and Future Progress of Digital Computers*

J. P. Eckert**

I am very glad to be here in Osaka today to talk with you about a subject in which I have been involved for 20 years.

People frequently ask how Dr. Mauchly and myself at the University of Pennsylvania came to be interested in building a digital computer. There were three reasons why we became interested in building a computer.

The first reason and most obvious reason was the war and the fact that the Government needed numerous calculations for aiming guns. Second, the thing that personally made the computer attractive to me was my interest in building things with electronics. I grew up a few blocks away from the laboratory of Philo Farnsworth who built the first all electronic non-mechanical television equipment. I felt that electronics deserved a more important use than simply entertaining people and for communication purposes.

The third factor was that we had two machines; one at the University of Pennsylvania, one at the Aberdeen Ballistic Research Laboratory which were copies of Dr. Bush's mechanical differential analyzer to work with and to inspire us.

Many firing tables for American guns were useless in Africa. In addition, automatic devices for aiming guns had been invented which required new calculations, and similarly, devices for aiming guns in airplanes had been developed and required new calculations. All these requirements together meant more computers would be needed.

Two programs were started to try to speed up our calculating ability. One of these programs involved the training of over 300 people to use desk calculators. The second program was to speed up and make more accurate the differential analyzer at the University of Pennsylvania and Aberdeen. We succeeded in speeding up the original differential analyzers by about ten times and were able to make them about ten times more accurate.

At this point we had improved the accuracy of the mechanical integrator to better than one hundredth of one percent. Further improvement in speed and accuracy seemed possible but looked like a dead end in terms of flexibility and increasing costs. Therefore, we decided to turn to some better way of doing calculations.

We had by this time incorporated several hundreds of vacuum tubes in the mechanical differential analyzers and since they performed so well, we thought of adding more vacuum tubes to a differential analyzer system as a next step. We originally thought simply of making a better integrator electronically. We decided next to eliminate the gears and motors also. We thought of building an integrator which simply added or integrated long string of pulses.

Actually we never did this. At this point we had, I suppose, developed the idea of the digital differential analyzer. We never used this name at that time. The original idea of using long train of pulses to represent numbers is inefficient. To represent a million would require a string of one million pulses to pass through the circuits. To overcome the inefficiency of this type of coding information we decided to binary code the information as it passed from one unit to another.

Finally we realized it would be better to make

* An Invited Talk Presented at the Meeting of JIPS Held in Osaka on May 8, 1964.

** Vice President of UNIVAC Division, Sperry Rand Corporation, New York.

an arithmetic calculator and forget about the differential analyzer in our future designs. Dr. Mauchly was an authority on numerical calculations. He had taught statistics in college to physics students and others interested in weather prediction. I was a student of electrical engineering and electronics. I also had some radar experience and other electronic experience in building gating circuits. We chose to build a general purpose device so that problems other than those related to guns could be studied.

We finally received a contract to start this work on my birthday in 1943. The Ordnance Department wished to have a general purpose computing device and gave us a contract to build a general purpose digital machine.

There was one more factor involved in the project. Building a machine in this way took 19,000 vacuum tubes, 70,000 resistors and over 100 kilowatts of power. I was only 24 years old and I did not know any better than to attempt such a massive undertaking. As chief engineer of this project I insisted that all the engineers use a wide margin of safety, the details of which we all agreed upon. I also insisted that the operation of all circuits would be calculated by algebra and arithmetic to insure that we have these margins of safety. The ENIAC, as we called the first machine, would not have been successful if we had not proceeded in this careful way. I was fortunate in this respect in that my professors from whom I had learned electronics, had taught me to design in this careful fashion. We all had a healthy respect for mathematics and the calculated design when we started out.

I might talk a little bit about the speed of digital computers. The ordinary desk calculator is about 10 times as fast as a human being. The MARK-1 at Harvard, whose design preceded the electronic ENIAC, was about 200 times faster than a human being. I am speaking of speed on bare arithmetic, not in overall problem solving ability. The ENIAC was about 200,000 times as fast as the human being who has no help, not even a Soroban.

Since then electronic machine speed has jumped by another factor of 100. The question of speed is an interesting one, however, because unless some new way of organizing the operation of a digital computer is developed, it does not look likely that we will ever be able to build the machine much faster than about 1,000 times as fast as we have today. In scientific computing it does not help to build a machine a little faster. The progress must be on a logarithmic scale. If this is true, and since we have developed devices which are 10 million times better than a human being, and if we can only go 1000 times better, then we have to conclude that unless something new occurs we have come 70% of the distance that we can ever come. It seems ikely that much future progress is going to have to come from improved organization in our machines and improved ways of using these machines rather than just improved speed of the machines.

We might discuss what speed progress seems possible for the coming period and what progress has been made in the last 10 years. In the last 10 years scientific computers have improved about 100 to 1 speed, but they have gone up in price 5 to 1, so that the performance per dollar has gone up perhaps 20 to 1. It is also true that for a given performance, prices have reduced of the order 20 to 1, perhaps over a somewhat longer period than 10 years. In other words, a small computer like the UNIVAC 1004 today is certainly about as powerful as the original ENIAC and it is more than 20 times less expensive to build than the original ENIAC.

In the next 10 years scientific computers will go up in performance another 100 to 1, and probably not increase in price this time as much as 5 to 1. I think that the cost of performing data processing in the next 10 years would go down by a large factor, perhaps not as great as 20 to 1 however. I should make it clear that I am comparing with the performance of existing machines, not machines that have simply been announced for the future.

In the last 10 years a great deal of time has been consumed in learning how to use the transistor. We have learned how to use transistors and magnetic cores not once but three or four times, because every time we designed a new machine around a transistor, a new kind of transistor was invented and we had the job to do over and over again. Recently, however, the transistor has shown some sings of stabilizing in its characteristics and methods of construction and it is now possible to concentrate engineering effort on optimising the use of the transistor rather than having to adapt to a new type every few years.

It is now possible to obtain in the United States 800 megacycle silicon transistors in quantities for a price of perhaps 60 cents. It is also possible to obtain silicon epitaxial diodes, which match speed requirements of the fast transistor at a price which is about 4 or 5 times less than the transistor. We have developed circuits using these parts which can provide logical decisions in about 1.5 ns. To this 1.5 ns we must add perhaps 0.5 ns or 1 ns (for a total of 2 to 2.5 ns) to allow for wiring delays and other factors. The machine we have been building up to now involved circuits with delays of 40 to 50 ns for a logical decision. This means we are now in possession of circuits which are 20 times as fast as in present days' machines and 10 times as fast as the fastest machines which have been built.

I am sure that the reaction to a remark like this is how are we going to have memory to match the speed of the circuits? We have been working for about two or three years on a new kind of wire plated memory, sometimes called a "cylindrical thin film memory."

Let me tell you what we have done in wire memory so far. We have delivered to the United States Government a small, comparatively slow wire memory which contains 1,000 bits and consumes only 1/10 watt. A larger memory for another Government Agency will contain 100,000 bits and will consume only 3/10 watt. we are working on a wired memory which will contain about 600,000 bits. It will operate in 1/4 μs for reading, and slightly longer than this for writing, which is, of course, less frequent. This memory will take several watts to operate. It is not aimed at low power consumption; it is aimed at low construction cost and high operating speed.

These memories use new unique selection systems and we believe that in computers we will ultimately achieve in the smaller size memories a 1/10 μs cycle time, and in very large memory sizes about a 0.5 μs cycle time. We think we can achieve these speeds with a cost per bit several times lower than can be achieved with a ferrite core memory.

Now if we have these high speeds of operation for the memory and for the computing circuits then, except for a scientific computer, what can we do with this speed? Even a scientific computer has to have input and output abilities to match this speed. We have a device called the FASTRAND* at the present time. This device consists of two large drums about 2 feet in diameter and about 5.5 feet long which together store approximately one billion pulses. A movable bar or boom mounted between the two drums carries 64 flying heads. For ordinary data processing it is sufficient to read the output of one of these heads at a time because each head puts out 1.1 million bits per second. However, for scientific purposes we might well read all 64 at once, and this will give as a read-out or read-in rate for the drum about 70 million bits per second. By using some fixed heads in addition to movable heads it is possible to synchronize this information into the computer memory without unreasonable cost or excessive equipment. It is further possible to at least double the drum

* Trademark—Sperry Rand Corporation

speed and pulse density so that an effective transfer rate of several hundreds of megacycles is possible in the future. Such pulse rates are over 100 times that available from the best tape units today and thus input-output balance can be maintained in the future.

But how can such high speeds be used in data processing? The most obvious use is to allow the inefficient use of one machine to simulate another and by this means to give the user the flexibility he would like in selecting the best or most available machine for his use. This use of high speeds, while attractive, may have short term value to some users however, since they may reprogram anyway after a while. Another obvious use is to allow the data processing user to become more scientific in his applications and use more complicated methods, and thus solve more difficult problems or solve his existing problems better. The most exciting use of speed, however, is to try and reduce the cost of data processing by time-sharing.

I say this because attempts to simplify a computer to a small number of orders and functions reduces productivity at such a high rate as compared with cost, that a point of negligible savings, even with low cost components, is soon reached if speed improvements are absorbed in this way. It seems better to build fast machines in the efficient size range and then find a means of time sharing such machines over telephone wires to many users. Many examples of this exist. The airline reservation systems are examples of machines shared by many users. Here, however, there is more than just the economics of speed. There is also a common interest.

I am speaking generally of sharing machines with many users, even though they may have only a small common interest, or may in fact have no common interest. They even may be competitors. Several efforts of this type have been started in the United States. One example is project MAC at M.I.T. Another example is a similar project at Stanford. There is a small commercial group in the city of Boston seeking to share a common machine among a group of hospitals. Of course, these people have a common interest. I would like to make it clear that the work at M.I.T. and Stanford is aimed at sharing a scientific computer system between the various departments of the University. While this is certainly an important area, the larger area of economic interest lies in sharing a computer for data processing in much the same manner between the various parts of a given company and more generally between different companies.

I expect the plated wire memory to be the least expensive memory we have in the speed range now covered by ferrite cores.

As we head into the use of integrated circuits, which are still too slow and too expensive to be generally attractive we will probably reach a point where circuit speed is essentially independent of cost. Even with present day separate circuit elements we approach this situation. Thus you can see that from either the point of view of logical design or circuit design, the only fruitful way of cutting cost that can have big effects in the next 10 years is time sharing.

Time sharing a computer has no real technical difficulties. You may need a moderate size, multiple head drum on which you can dump almost the entire memory contents except for some input and output routines and associated data buffer areas in the memory. If there were, for example, 50 people using one computer, the memory capacity of the drum might well be 50 times the internal capacity of the high speed memory, not counting the input-output area. Supposing we assumed we have a computer with half a million bits in its memory, not counting the input-output area. Let us assume that information is transferred through 64 channels at a rate of 2.5 mc per channel. This means that information is transferred at a 160 mc rate to the drum. Thus 3 ms is required to completely dump the memory, another 3 ms is required to fill the memory for the next user. This means that a total of 6 ms is required to switch from one user to another. Under these conditions a cycle of 2 seconds to cycle through 50 users would result in an average waiting time for any given user of about 1 second. This would still allow 85% of the time available for computing. Only 50×6 ms or 0.3 seconds out of 2 seconds would be wasted making transfers.

A high speed computer like this chopped into 50 pieces may cost as little as a few hundred dollars a month to each user for the computer itself. To this must be added printer, input keyboards and readers and telephone line costs. The printer would be the largest cost at the present time unless the user can use a simple teletype. The costs of the telephone line will be negligible in a large city like New York or Tokyo but can become significant in the large areas of the United States. Input devices can range from very small to very large costs.

I do not wish to minimize the organization problem that is involved in a new idea like this. In the United States we have sometimes called this idea "a computer utility." You may wish to call it "a real time-time shared service center."

I believe that considering a very small user, this principle, without other improvements, can reduce the cost of computing 2 to 1. For larger users the factor can be much higher.

There are some special things about time sharing. Punched cards and magnetic tapes have a much different place in time sharing than they do now. Punched cards may still have a place for providing input information at the user's location but the main storage of information for the user must be provided for by mass storage equipment at the central location. Mass storage equipment is just reaching a point, in its economics and speed, which justifies it for time sharing use. If you do not have economical mass storage, then 50 users might need perhaps 6 tape units apiece on the average. We have then to visualize a room with 300 tape units with many people running back and forth with reels of tape in this room. This is too horrible a situation to talk about any more.

The mass storage should be in the same place as the computer and not at the many user locations for two reasons. First, a large mass storage, due to the sharing of much of the electronics, casework, drive and other items can provide storage at lower cost than each user having his own small mass storage unit. Second, the main input-output traffic in a data processing system is not between the keyboards, readers and printers, but between the computer and the tape units or other forms of mass storage used in the system. Consequently, the telephone line costs will only be small if the heavy input-output burden from mass storage is confined to local wiring at the time shared center. A few tape units—but certainly not 300 may be needed at a time shared center to handle long term dumps and to provide periodic dumps to back up the regular mass storage in case of failure.

Up to now, there has not been great pressure to reduce the costs of high speed printers because the cost of the computer, the cost of tape units and readers tended to mask the relatively high printer costs. But once we change to time shared centralized computers, the printers may well become the most expensive items in the system. Fortunately, printers can be further cost reduced and work on this is underway.

All customers in a time shared system would not necessarily have the same size printer, in fact, very small customers might have no more than teletype units for their input and output. All customers would also not receive the same share of computer time. In fact, peak loads could be handled in the same way that power loads are shared among electric power station users.

There is even an input and output device for the very, very small user. Soon in the United States the telephone company will install large numbers of push button telephones. With push buttons information can be keyed in much more rapidly than with a dial telephone. Certainly this type of telephone can be used as an input of digital information to a time shared computer.

Some time ago, for airline use, we made a device called UNITEL* with a simple voice drum. With this, a computer can talk back to users on the telephone. With this drum and push buttons even a very small user can have economic access to a very large flexible computer system. He can even code this problem in COBOL, ALGOL, FORTRAN or some other programming system because the memory of the central time shared computer will allow this. Up to now, a small computer is prevented from using an extensive programming system because of its limited memory size.

In a way, we have a time shared computer in Toronto. It is a regular UNIVAC 1107 which is tied to 1000 traffic lights in the city of Toronto. It is also tied to several thousand magnetic sensing units in the streets of Toronto. It controls the traffic load in Toronto through these inputs and outputs. This device is expected to decrease the traffic problems by 20% in the city of Toronto. It would cost 10 or 20 times as much to get an equivalent improvement by widening streets or building additional streets in Toronto. The Toronto traffic control system is not entirely used for traffic control. There is time on the computer to perform some calculations about road maintenance problems, some calculations about traffic tickets and for other problems associated with the running of a large city. All this is time sharing.

Thank you very much. It has been very nice to be here and talk with you this afternoon.

---

* Trademark—Sperry Rand Corporation