Some comments on connecting a small on-line computer to a large data processing machine

JOHN O. AIKIN and RICHARD B. MILLWARD
Brown University, Providence, Rhode Island 02912

Many psychologists with access to a small on-line computer for control of experiments also have access to a large central computing facility. In this paper, we propose that the interconnection of these two machines can enable the research psychologist to exploit both more effectively at reasonable costs. The paper is divided into three parts: first, a discussion of some properties of large and small computers and the implications of these properties; second, a discussion of some benefits of interconnection; and third, a discussion of certain considerations relevant to interconnection.

PROPERTIES OF DATA PROCESSING AND PROCESS-CONTROL MACHINES

Computers may be divided into two categories: *data processing machines* (DPMs) and *process control machines* (PCMs). DPMs are designed primarily for the manipulation of symbols, especially numbers. This design purpose is reflected in six major characteristics of DPMs: (1) DPMs have long word lengths and large core sizes to facilitate high-accuracy computations on large amounts of data; (2) DPMs have powerful hardware for fast arithmetic processing; (3) DPMs have fast bulk storage (disk and/or tape) to handle large volumes of data; (4) DPMs have fast I/O such as card equipment, high-speed printers, scopes, and plotters for easy input and output of masses of data; (5) DPMs are programmed by the user in high-level languages (e.g., FORTRAN, PL1) oriented toward data manipulation; (6) DPMs are often delivered with documented and user-tested software, including extensive arithmetic subroutines, complete data-handling programs (e.g., analysis of variance), and operating systems for time sharing or multiple-task batch processing.

The characteristics of the DPM define a machine with great power for symbol manipulation but one that is too expensive to be dedicated to a single user and usually too sluggish to allow real-time control applications when also being time-shared. Thus, DPMs are usually found in central installations serving a number of users with diverse purposes and requirements whose only common need is the manipulation of symbols.

PCMs, however, are designed primarily for the manipulation of devices rather than symbols. PCMs take as input real-time sequences of external events (switch closures, voltage level changes, etc.) and produce as output other real-time sequences of external events (relay actuations, signal lamp illuminations, etc.). The task of the PCM is thus to respond to and to control a sequence of external events in real time. This design purpose is reflected in two major characteristics: (1) PCMs have a reduced symbol manipulation capability manifested in short word lengths, small amounts of core, limited instruction sets not particularly suited for arithmetic processing, and slow, limited data I/O devices; and (2) PCMs have expanded capabilities for real-time control characterized by speed, flexible I/O instructions for programming responses to external events, and hardware which is easy to interface to special-purpose external devices.

The characteristics of the PCM define a machine having only limited capacity for symbol manipulation but ideally suited for on-line control and sufficiently inexpensive to be justifiably dedicated to a single real-time user for his special purposes. Thus, PCMs are usually found in decentralized installations serving, at most, a few users with very similar real-time control needs.

The experimental psychologist will likely select the PCM as the logical way to automate the collection of data and the DPM as the logical way to analyze data. These two processes are not independent: a PCM can generate astonishing amounts of data requiring sophisticated analyses possible only on a DPM. Hence, efficient transfer of data from the PCM to the DPM will be highly desirable and frequently indispensable.

Data transfer between computers is traditionally accomplished by hand entry into the DPM, by paper media such as cards or punch tape, or by compatible magnetic tape. Hand entry is suitable only for small amounts of data and is subject to transcription errors. Paper media often cannot be used for PCM-DPM transfer because PCMs tend to be equipped for punch paper tape but not cards, while DPMs tend to be equipped for cards but not paper tape. Even where cards or punch tape can be used, non-reusable paper media can be a significant operating expense and the handling of easily torn paper tape or easily misordered cards can be a serious inconvenience. Compatible magnetic tape is probably the best transfer medium from a technical viewpoint, but the cost (between $7,000 and $10,000) of a suitable tape drive for the PCM is often prohibitive and can really only be justified if the amount of data to be regularly transferred is very large.

BENEFITS OF INTERCONNECTION

An on-line connection between the PCM and DPM is an attractive alternative to the above methods of data transfer. A dial-up connection over standard telephone lines costing between $1,500 and $2,500 allows for data transmission rates from a minimum of 10 char/sec to upwards of 150 char/sec. These rates are adequate for many real-time data collection situations. Even much higher rates are possible, although costs increase rapidly with rate.

Perhaps the greatest advantage of interconnection is that in addition to solving the data-transfer problem economically, interconnection of a PCM to a DPM offers an enormous increase in the range of experimental procedures possible in the automated laboratory. This is because direct interconnection, unlike the other three methods of communication, can be interactive in real time. Both the DPM and the PCM can have programs running simultaneously, and these programs can interact, passing information back and forth.

Each program can handle those tasks in the experimental procedure for which its respective machine is best suited. Thus, the PCM can handle the recording and control of external events, and the DPM can perform the symbol manipulations which are so difficult on a small machine. The special-purpose routines which are required to deal with the external devices in the laboratory in real time can be written for the PCM to take advantage of its use as a dedicated machine. The more general flow of experimental contingencies from trial to trial can be written in the more powerful symbol-manipulation languages found in the time-sharing environment of the DPM.

Using this sort of division of labor, it becomes possible to implement experimental procedures which would otherwise be extremely difficult or impossible because of the limited programming languages, restricted core size, and arithmetic capability common on PCMs.

RELEVANT CONSIDERATIONS

The actual process of establishing a PCM-DPM interconnection involves considerations of hardware, software, and system reliability. With respect to hardware, the particular peripherals required for the PCM and DPM will depend on the particular machines. In general, the simplest and least expensive approach is to make the PCM resemble a remote terminal of the DPM, using a dial-up system which includes an asynchronous serial line interface and modems. For PDP-8 series PCMs, this involves only off-the-shelf equipment, and relatively complete information about what is required can be found in the *DEC Communications Equipment Handbook* available from Digital Equipment Corp., Maynard, Massachusetts. The existence of requisite hardware at the DPM end is a precondition for any consideration of interconnection because this equipment will generally be too expensive to be supplied by the user of the PCM.

Software to service the interconnection is very straightforward for PDP-8 series PCMs: The interconnection is programmed exactly like a teletypewriter. Information is sent to the DPM character by character and received from the DPM in the same fashion. Input programming should take advantage of the interrupt facility to avoid possible loss of characters due to overwriting by the DPM before the PCM has fetched the previous character.

Software for the DPM will depend on the particular time-sharing system in use. In most cases, information will be transmitted and received simply by using the equivalent of FORTRAN "read" and "write" statements referencing the PCM terminal as the device. It is highly desirable, however, that the overall DPM time-sharing system have certain characteristics. Chief among these is some variant of a "virtual machine" concept (cf. IBM 360 Cambridge Monitor System or DEC PDP-10 software system). In such a system, each time-sharing user appears to have the entire DPM at his disposal and powerful macros are available for using the DPM. Furthermore, the system is designed to protect user files against system "crashes," to respond quickly to action at the terminal, and to minimize connect charges.

Some conditions based on DPM system reliability must be attached to the advisability of choosing to interconnect a laboratory's PCM with a central computing facility's DPM. The DPM installation should be expected to remain stable both in hardware and software for the foreseeable future. The DPM time-sharing system must be available for a sufficient number of hours each day at the appropriate times. Cooperation of systems-wise DPM personnel is necessary both in establishing and using the interconnected system. Under these conditions, a hybrid PCM-DPM system can be a powerful laboratory tool of reasonably low cost.

AVALA, a small on-line system with off-line communication with a large computer

JOSEPH L. YOUNG
State University of New York at Stony Brook, Stony Brook, New York 11790

When I first heard about this conference, I knew I wanted to attend it because I use an on-line computer in my own research and was interested in the insights of others regarding this important area of instrumentation and experimental technique. I was less certain that I wanted to give a presentation. After all, I had not designed a computer, I had not discovered a new hardware wrinkle, and I had not made any major software innovations. All I had was an idea, which had been implemented in our on-line system, of how communication between a small on-line system and a large central campus-wide facility could be effected and the conviction that such communication could eliminate many problems commonly griped about by users of small on-line systems. Since I did not know if my idea was novel enough to present, I turned to Professor Tepas, the conference director, for guidance; he felt it was, and so I am here today. Thus, if you get great insights from my talk, thank me, but if you are bored, blame him.

Briefly, the idea is as follows: Include in the on-line system one or more magnetic-tape drives compatible with the central campus facility (at Stony Brook, an IBM 360/67). This meant, for us, 9-track 800-bpi drives with interrecord gaps within IBM tolerances.

The total system configuration can be described quite simply. It was manufactured to our specification by Infotec, Inc., of Westbury, New York, and we call it AVALA, Automated Verbal Associative Learning Apparatus. The core computer is a PDP-8/L with 4K memory and standard Teletype reader-punch. Interfaced to it is a tape-drive controller with two tape drives, completely IBM compatible, but limited to tape reels of 7 in. or smaller; communication between the tape drives and the 8/L is via the data break and bypasses the accumulator. In another room, connected by cables to the main console, is the subject station. It consists of a stimulus display and a response keyboard. The stimulus display is a unit with two banks of eight Burroughs NIXIE tubes each, one bank atop the other; each NIXIE has 14 individually addressable filaments which can be lighted, so that all alphanumeric characters, many special characters, and all manner of stick figures can be presented. There are two response keyboards that can be used. The first is a Teletype Model 33 keyboard, which is simply a keyboard that does not make a hard copy of what is typed on it but only transmits it. The second is a response panel with four buttons. Our PDP-8/L is connected by cable to another one with 8K memory and a high-speed reader-punch, enabling us to use those facilities if necessary.

What does this arrangement buy us? Briefly, it gives us the full power of the IBM 360/67 to create sequences of stimulus materials, which are read off the tape on one of the drives, and to analyze the data, which are collected on a tape on the other drive. These capabilities are important to me. I run verbal learning experiments in which temporal and procedural variables are varied within the same S's trials. The algorithms I use to set the sequence of trials for a simple S are quite complex, using much of the available core storage on the 360/67; moreover, they are not linear in that the content of a later trial may well be determined before that of an earlier trial. Furthermore, the analyses I run are often quite complex; I typically run grid searches for the set of parameters of a model giving the best fit to a set of data, a time-consuming process even on the 360/67. Thus, I must be able to communicate effectively with the large computer or else beef up the core