COMPUTER CONSIDERATIONS

PREPARED FOR THE

10th NUTRIENT DATA BANK CONFERENCE

by
Peter Zilahy Ingerman
Systems Consultant
40 Needlepoint Lane
Willingboro NJ 08046-1997
(609) 871-7474

This paper should be viewed as a check-list of concerns when considering the acquiring of a micro-computer, rather than a list of recommendations.

Computer Selection Considerations

There are three possible choices with respect to a new on-site computer system that is based on microcomputers:

- Stand-alone microcomputers
- Networked microcomputers with a file-server as the central unit
- Terminals or microcomputers connected to a multiprocessing central unit

Stand-alone microcomputers

A stand-alone microcomputer is, as its name implies, an entirely self-contained unit. It consists of a display, a keyboard, the computing unit itself, some internal memory (RAM and ROM), and some secondary storage (floppy disk and, possibly, a hard disk).

A single stand-alone microcomputer may be sufficient for your purposes. However, it is important to keep in mind that if there is only one computer available, and it goes down for any reason, you are out of business. It is therefore reasonable to consider either having more than one microcomputer, or having arranged for access to another microcomputer on an emergency basis, if it is essential that you be operational.

The major advantage to this approach is that, if care is taken in the purchase of these computers and their accessories, so that every capability is available on at least two computers, then each computer has a "secondary" computer to serve as back-up in case the "primary" computer fails for some reason.

The major disadvantage of this approach is that the only way of transporting data from one computer to another without reentering it is by physically transporting a floppy disk from one machine to another. Unless great care is taken to control this environment, it will be difficult to determine which copy of a particular file is the current one, and to be sure that all current files are backed up. (See below, however, under Telecommunications.)

An additional disadvantage is that the size of a reasonable stand-alone microcomputer is (comparatively speaking) limited, and hence it becomes difficult to handle large quantities of data: either the data is on a number of floppy disks, which must be continually interchanged, or the data is on a hard disk, which can become inaccessible if the hard disk or the microcomputer breaks down.

Networked microcomputers

This configuration requires a number of independent microcomputers, as described in the preceeding section, connected ("networked") to a "file-server" through a Local Area Network (LAN). The file-server is also a computer, but is not used as a terminal. Rather, it acts as the custodian for copies of programs and data that will be required for use on the individual computers; when something is required, a message is sent via the network, the required file is moved from the file-server to the requesting microcomputer, and work then proceeds. When the work is completed, the updated copy of the data (as appropriate) can be sent back to the file-server for storage.

The major advantage to this approach is that the problem of manual control of the current copy of data is eliminated, since there is only one current copy at any point in time. Similarly, there need be only one current copy of any program that is to be used in the network.

In addition, the terminals, since they are complete microcomputers, can continue to function even if the file-server breaks down; if careful backup has been performed from the file-server, the temporary loss of the file-server means a reduction in efficiency, not a total inability to work.

There are several disadvantages to this approach.

There is no single neworking standard, so that your organization may be at risk if the vendor selected is no longer available for support.

Certain proprietary software packages, because of the protection schemes used internally to them, may not operate in a networked environment.

There is a potential problem with multiple users updating the same data base. This can be solved by careful programming, together with a suitable operating system, but it is frequently not addressed, and can lead to badly garbled data.

The individual terminals, since they are microcomputers in their own right, are more expensive than the "dumb" terminals that can be used in the multiprocessing environment described next.

Multiprocessing

This approach presumes that the individual terminals have little or no computational ability and internal storage (that is, they are

explicitly terminals, rather than microcomputers), and are all connected to a single computer that performs all of the computations required for all of the tasks that are currently in process, as well as providing the capabilities of a file-server with respect to data and program files. This is, of course, the typical environment with a "large main frame"

The major advantage of this approach is that the incremental cost to add a work station is very much smaller ($700 to $1500 depending on the bells and whistles), compared with the cost of a full microcomputer ($1500 to $8000, depending on the bells and whistles).

The major disadvantage of this approach is that if the central computer goes down for any reason, the entire system goes down. Even if an entire back-up system is available, data that is stored internally to the central computer (e.g. on a hard disk) may not be available without considerable effort and time.

This approach also has the disadvantage that the central processing unit must be rather larger than that required for the file server (and hence more expensive). At the same time, it is uneconomical to get a central processer that is capable of providing all of the computational power of a stand-alone microcomputer for each work station. Hence, when the work stations are being heavily used, there is an increasing likelyhood that performance will be degraded. This will be particularly true if compute-intensive tasks (e.g. statistical work, reformatting while word-processing) are all being done at once.

Hybrid Systems

If your environment is one that includes a "large main frame" computer, most of the discussion above does not apply directly to you. However, both microcomputers and terminals can be attached to a main frame as terminals. If microcomputers are attached to your main frame, the situation is similar to that discussed under "Networked microcomputers", and all of the considerations discussed elsewhere in this paper with respect to microcomputer ancillary devices and peripherals apply.

There is an additional disadvantage that arises in this environment. If microcomputers are used as terminals, there will be a temptation to have "personal" copies of data and programs, that have been downloaded from the main frame, or developed individually on microcomputers. If this situation is not controlled with great care, it will be difficult to determine which copy of a particular file is the current one, and to be sure that all current files are backed up.

Telecommunications

Most microcomputers presently on the market can be adapted for communications on ordinary telephone lines, by using either a built-in or external modem (q.v.). It is probably not reasonable to purchase a computer that does not have this capability.

However, it is reasonable to consider the implications of having telecommunications capabilities, particularly if the microcomputer is left unattended, operating, and connected to the telephone lines. In this situation it is possible that unauthorized users may access the microcomputer. In the best case, nothing will happen. In the worst case, data and programs can be erased

Hence, in a telecommunicating environment, there are additional questions with respect to security, privacy, and integrity that must be addressed in a manner appropriate for your particular installation.

Software and Data Considerations

The concerns with respect to software (that is, the programs that will process data to produce the desired information) and data (the base from which the programs produce the desired information) are essentially the same:

- is the software, or data, purchased "as is", with no representation as to its usefulness or correctness, or is there some kind of guarantee?
- if there are problems (bugs in software, errors in data), will they be able to be fixed, and, if so, by whom?
- will some kind of service contract have to be purchased, to assure that updates will be received as available, or will each update have to be purchased separately, as a new product?
- will the author of the program or compiler of the data, whether a company or an individual, be available for consultation?
- if the author of the program or compiler of the data, whether a company or an individual, chooses to abandon support (or not to provide support) will the original source coding for the software, or the maintenance tools for the data, be available to present owners of the software or data?
- is the documentation for the software or data sufficiently complete with respect to the intended user?
- what is the cost of the software, or data? If that cost does not include maintenance and updates, what is the cost of maintenance and updates?

These same considerations apply whether the author of the software or the compiler of the data is an outside vendor or a part of the same organization as the user.

Peripheral Selection Considerations

Floating Point Hardware

Most of the computers that are presently available are inherently capable only of arithmetic on integers When real numbers (e.g. 1.25)

are used, it is necessary to provide a set of programs that perform the arithmetic on these numbers.

If your environment is one where significant numerical processing will be done, an additional device that should be considered is a "floating point processor". This is a piece of equipment that can plug into the computer, and that increases the speed of arithmetic on real numbers by approximately a factor of 10.

Modems

Modems are devices used when two computers communicate with each other over telephone lines. It is necessary to have a modem at each end of the communication link, and the two modems must be compatible with each other.

In addition to the modem itself (which is a piece of hardware), it will be necessary to have communciations software at each end of the communication link, and the software (as well as the hardware) must also be mutually compatible.

Printers

There are four fundamentally different types of printers on the market: dot matrix and fully-formed are the two types of contact printers, and ink-jet and laser are the two types of non-contact printers.

The dot-matrix printers are the most flexible of the contact printers. In the price range of $1000 to $2000 it is possible to get a dot matrix printer that can print in a draft mode (at upwards of 400 characters per second) and in a near-letter-quality ("NLQ") mode (at about one-fourth the draft speed) as well as having extensive graphics capabilities. In addition, it is possible to have multiple fonts, and to change fonts on a letter-by-letter basis. Dot-matrix printers that print in colors are available.

Fully-formed printers (true letter quality) print at about the same speed as dot-matrix printers in NLQ mode. However, their graphics capabilities are usually more limited, and, since font-changes involve changing a print element, it is not practical to make numerous font changes. Also, there is some software available that can print a wide report "sideways" on a dot matrix printer; this capability is simply not available on a fully-formed printer.

Non-contact printers are significantly faster and quieter than contact printers, since there is no requirement that there be the mechanical impact of an element (or a wire) against the paper

Ink-jet printers are reasonable when they are in use more often than not (so that the ink does not dry and clog); they are also particularly useful when printing on irregularly shaped objects (many brands of soda have a bottling date "sprayed" on the side of the cap). One major advantage of an ink-jet printer is that it can print in colors.

Laser printers are quite fast (8-12 pages per minute), have most of the flexiblity of the dot-matrix printers, and are expensive ($2500 and up).

Stand-alone Printer Buffers

A stand-alone printer buffer is a device that appears to be a printer to a computer, but appears to be a computer to a printer. It has its own internal memory (typically, from 32K characters to 256K characters), and can receive data from a computer as fast as the computer can generate it. The data is then stored in the internal memory, and is retransmitted to the printer as fast as the printer can accept it.

This device has the effect of freeing the computer from having to slow down to printer speed, and, therefore, permits the use of the computer while a long document is printing, rather than requiring the operator to wait

Floppy Disks

Floppy disks are the storage media used in floppy disk drives. They are considered to be "expendable" items, in the sense that they are thrown out when they are no longer functional.

It is a poor economy to use cheap floppy disks in an environment where reliability is important. If the only purpose for a particular floppy disk is to be written on once and then, shortly thereafter, to be read once (as when a floppy disk is used to send data or a program through the mails), then bulk-packed, off-brand disks (at a cost of about $1.00 each) may be sufficient. High-quality disks, on the other hand, can be read and written many times without deterioration, and are stable (when stored in reasonable environments) over long periods of time. (High-quality disks can cost $2.00-$4.00, depending on size of diskette and the number purchased at any one time.)

Floppy disks must be stored in an environment that does not exceed 125 degrees F, nor 80% relative humidity (see Media Safes).

Floppy Disk Drives

A floppy disk drive uses a single floppy disk as the recording medium. Floppy disks (and their associated drives) come in 8-inch, 5.25-inch, and a variety of 3-inch sizes. There are a number of different ways of organizing the data when it is recorded on a floppy disk, and the fact that the physical diskette is interchangable with a machine other than the one on which it was recorded is no assurance that the data will be readable on that machine. Many programs are available, however, that can be used to transcribe the data from one format to another.

The organization of the data on a floppy disk is determined by the operating system of the computer on which the disk was written, as well as the computer itself. Hence, if more than one operating system

is available on a given computer, it is possible that disks written by using one operating system will not be readable by the other operating system. (An "operating system" is a program that provides the working environment for other programs. CP/M, MS-DOS, and UNIX are examples of operating systems for microcomputers, and CMS and TOS are examples of operating systems for IBM main frames.)

The advantage of a floppy disk is that it is easily removed from the computer and stored in a safe place. (It can also be used to transport files from one computer to another, similar, computer.) Also, it is easy to make a copy of a floppy disk (assuming that a computer with two disk drives is available), and a disk can be completely duplicated in less than 120 seconds.

The disadvantage of a floppy disk is that it is very sensitive to the way in which it is handled, and can be easily damaged by being scratched or having its surface contaminated (e.g. with smoke particles). Also, if floppy disks are the only "bulk" storage in use on a system, there is a tendency for copies of programs and data to proliferate, with the consequent difficulties of maintaining control.

Hard Disk Drives

A hard disk drive is the logical equivalent of a number of floppy disks, packaged with a drive mechanism. The smaller drives use the "Winchester" technology, in which the drive and the media are sealed in a box, so that the only interfaces are electrical. Drives of this type are available in the 10 megabyte to 50 megabyte range, at a cost of about $100/megabyte.

A hard disk has the advantage of compact storage, without requiring the handling of individual volumes (e.g. floppy disks).

A major disadvantage, however, is that it is necessary to back up the data on a hard disk regularly, so that in case of disaster it will be possible to reconstruct the hard disk.

Streaming Tape Drives

A streaming tape drive is a special-purpose tape drive that can be used to back up a hard disk more rapidly than simply transcribing the contents of the hard disk to floppy disks. (Note that it takes on the order of 25 5.25" floppy disks to hold the complete contents of a 10 megabyte hard disk.)

Certain computers that have hard disks have built-in streaming tape drives. This is a consideration if the purchase of a computer with a hard drive is contemplated.

Reel-to-Reel Tape Drives

A reel-to-reel tape drive can also be used for backup of a hard disk.

This device has, however, the additional capability of being able to

read tapes that are written by other computers, and being able to write tapes that other computers can read. Such a capability might allow your organization to utilize data bases that are presently available only on tape. (It should be noted that certain data bases are available from the Government on both magnetic tape and floppy diskettes.)

(Data can, of course, be sent over telephone lines --- but this is a slow process at best, and can take hours for large data bases.)

There is clearly no necessity for the $10,000 investment that is required for a reel-to-reel drive at the initial stages of computerization. Its long-term desirablily, however, should be considered carefully at the initial stages, and, if there is any expectation that reel-to-reel tape capability will be required, only those computers that have the proven ability to have such a drive attached should be considered.

Color and Graphics Displays

Graphics displays permit the display of pictures (e.g. bar graphs, pie charts) as well as letters and numbers. Graphics displays are now common, and it would be unreasonable to select a computer that did not provide that capability.

Color displays, while more expensive than monochrome displays, can be significantly more effective in displaying data. There are camera attachments that permit 35mm slides to be made directly from the computer display

Plotters

Certain dot-matrix printers can also be used for graphic displays; if the dot-matrix printer is a color printer, the graphics can be printed in color. (There are programs that will "dump" the screen of an Apple Computer to a suitable color dot-matrix printer with a single key-stroke.)

Plotters, that are designed to put graphic displays on paper, usually in multiple colors, may be an alternative to color dot-matrix printers. They are usually best suited to the displaying of colored lines, rather than masses of color, and the inexpensive models do not have the flexibility of a color dot-matrix printer. On the other hand, they can be easier to use, since a line in any direction can be specified by giving its end-points and its color, and the plotter will do the arithmetic. (Note that some color printers also have this capability.)

Ancillary Considerations
--------- -------------

Furniture

It may be desirable to purchase specially designed tables for the

work stations. If ergonomic stands are not included in the terminals or microcomputers that are acquired for the work stations, then they can be purchased independently, as part of the tables.

The chairs that are to be used at the work stations should be designed for such use; this implies a somewhat greater degree of adjustability than may otherwise be required.

Stands for printers will also be convenient, as well as acoustic hoods if the printers are not isolated.

Static mats, to protect the computer and media from static electricity, are advised for each work station.

Diskette files, for the orderly storage of floppy disks at each work station, will be desirable.

Suitable lighting should be assured for each work station.

Cables

Cables will be required to interconnect pieces of equipment such as printers and microcomputers. Some of the interconnection cables may be provided as part of a "package" price, while others may have to be purchased separately.

Media Safes

An ordinary "fire safe" is not satisfactory for the storing of magnetic media. Magnetic media must not exceed a temperature of 125 degrees F, nor a relative humidity of 80%; exceeding these limits will do permanent damage to the disks, and will make them unreadable. A media safe is designed to maintain the internal environment below these limits in case of a fire.

Enough media safes will be required to store all of the required on-site back-up copies of magnetic media. It may also be reasonable to provide capacity for the storage of all on-site working disks for nights and weekends, although if backing up is done carefully and regularly, it may not be strictly necessary.

It may, however, be reasonable to have a media safe off-site so that a secondary back-up copy of all data will exist.

Media safes are available for floppy disks (list price approximately $1300) and for magnetic tapes (list price approximately $2000).

Supplies Storage

Adequate space must be provided for the storage of supplies (e.g. disks, paper) prior to their use. Paper, in particular, is heavy, and should be stored near the printers (or, alternatively, paper dollies are available, and should be acquired).

Power Supplies

The electrical power supply that is conventionally available is usually not adequate, for a computer system, in terms of reliability and "noise" on the line.

The first type of protection for computer components to be considered is a surge/spike protector. This device tends to eliminate sudden changes in voltage caused by lightning and by the starting and stopping of other equipment (typically, those containing motors). A top-of-the-line surge/spike protector costs on the order of $300.

Next is a line conditioner. A line conditioner tends to remove high-frequency noise (such as that caused by fluorescent lighting) from the line, and also tends to protect against "brown-outs" by producing a near-constant output voltage even when there are fairly wide variations in the input voltage. A reasonable rule of thumb for price for line conditioners is $200 plus $400 per KVA.

Finally comes an uninterruptible power supply (UPS). This device contains batteries that provide power to the computer for some number of minutes (8 to 15, typically) when the supply power fails entirely. This allows for the orderly shut-down of the computer system in the event of a power outage, without losing data. A reasonable rule of thumb for price for a UPS is $100 plus $1000 per KVA.

(A KVA, which stands for kilovolt-ampere, is approximately a kilowatt.)

(There are larger UPS available --- including some with diesel generators --- that can maintain power to the computers for 24 hours or more. They are probably not reasonable in the kinds of environments contemplated here.)

A power line monitor should be rented, and the power monitored for at least a week (longer, if possible) to determine the exact nature of any problems with the power as supplied. This, together with the kinds of usage that will be made of the computer system, will determine the exact kinds of power conditioning equipment that will be required

Personnel Considerations

Day-to-Day Usage

All personnel who will be using the computer system on a day-to-day basis should be trained to perform such routine maintenance as the cleaning of display screens on terminals and the changing of paper and ribbons on printers.

If this is an environment using stand-alone microcomputers, then, in addition, at least two persons should be trained as "key operators", in the sense of being familiar, e.g., with option switch settings, clearing paper jams, replacing fuses, and the like.

Programming

It is probably not cost-effective to have a staff programmer, at least for the short term. Rather, the standard report programs should be contracted out, and the users of the system should become sufficiently familiar with the software tools available to be able to prepare one-time reports; if it becomes desirable to make these reports "pretty", it will be reasonable to contract to have that done.

Ultimately, if the dollar cost of contract programming becomes high enough, it will become reasonable to hire a staff programmer.

Maintenance

One of the considerations in the selection of both hardware and software is the availability of maintenance.

While it is possible to get deep discounts on hardware, they are usually only available from firms that provide little or no maintenance support. Hence, for an organization that is interested in using computers, rather than repairing them, it is reasonable to shop for price --- but only in the context of the price including hardware maintenance.

Most dealers, and some independent organizations, will provide after-warranty service contracts. Unless your organization contemplates having an in-house person with the technical competence to perform hardware repair, a service contract will be essential. The type of service contracts vary from service on a bring-in basis to round-the-clock service on location, with a guaranteed response time. (Obviously, the latter is significantly more expensive than the former.)

If your organization has sufficient equipment so that every piece of equipment is duplicated, then there is, by definition, a degree of in-house back-up for all of the hardware, and it may not be strictly necessary to purchase the most extensive possible service contract.

With software, on the other hand, the maintenance typically involves the acquiring of upgraded versions of software that has already been purchased and installed. A single person at your organization should be designated as "software coordinator", and should be responsible for upgrading the in-house software as appropriate. Note that this implies careful control of all copies of each software product, so that all copies can be upgraded at once.

(On the other hand, whenever an update is obtained, it is important that it be checked out before being put into the general environment. Unfortunately, much of the available software is not carefully debugged before it is distributed, and there is no sensible reason to upgrade a piece of software that is working to one that has not been shaken down.)