Networking VS Central Database

R. L. Grundy
Librarian
The University of Texas at Austin
Marine Science Institute
Port Aransas, Texas 78373-1267

ABSTRACT

As technology advances, marine science libraries are finding that networking is becoming a desirable possibility. It will allow us to consult other libraries directly, rather then pay to build and maintain large central databases -- then pay again to use these same databases and it will allow us to concentrate on our own library and its users and still serve the general marine science community.

INTRODUCTION

First note that there are libraries who have long since moved into the world of automation. This paper is not intended to suggest that they jump into another system; however, it would really be useful if their data bases could be available to outside users while the rest of us try to catch up. This paper will discuss the basic requirements for sharing information, particularly the unique information which is located at each of our respective libraries. The basic idea is to build machine readable data bases locally which can be shared collectively rather than spend our time and energy feeding into some central data base which charges us to input the data and then charges us again to access it.

There are several requirements for such an ambitious project and they are: 1) static memory which can be accessed both locally and by phone lines, 2) the ability to communicate/network, 3) cost and maintenance of the hardware and software, 4) record cost, and 5) multi-user capability.

DISCUSSION

MEMORY. The reason for building a central database, in the past, is simply because only the mainframe had the ability to address the memory needed for a data base; however, this is no longer the case. Microcomputer architecture is capable of addressing a large amount of memory, making it possible for us to catalog our own collections and providing ready access to these files. Compaq Computers have just introduced their new microcomputer using Intel's new 80386 microchip which can address 4 gigabytes of physical memory and supports 64-terabyte (246 bytes) virtual memory capacity, with a sustained performance of 3 to 4 million instructions per second; previously available only on mainframes. It is a full 32-bit machine. "The 80386 combines pipelined architecture, segmentation, and paging in a multitasking environment", according to Paul Wells1. The use of this architecture in a distributed system would give us the speed and

memory access needed to build a local data base and allow public access. One possible configuration is a distributed system (fig. 1).

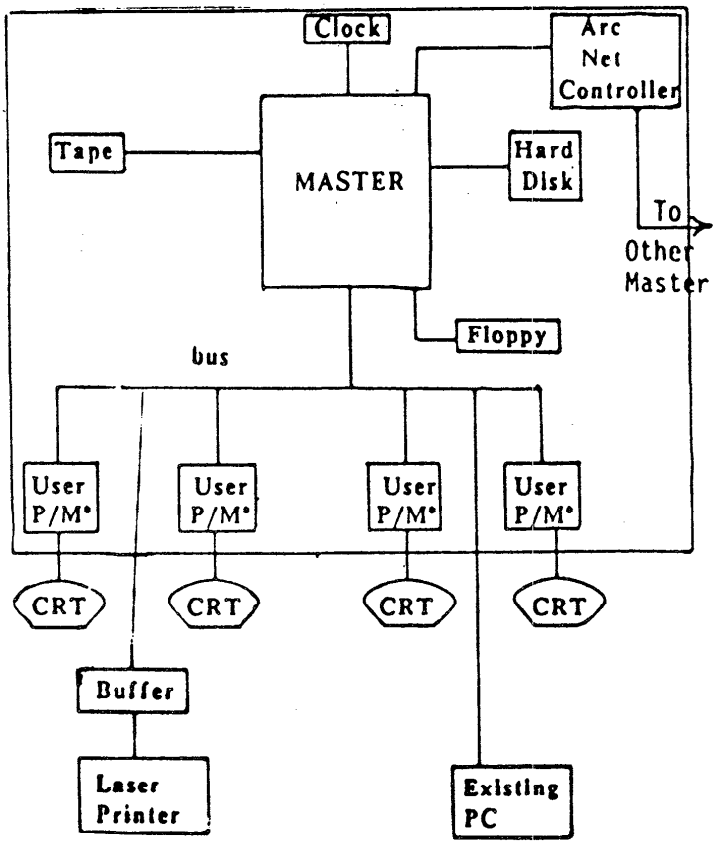

Figure 1 *P/M Processor & Memory

COMMUNICATION/NETWORKING CAPABILITIES. Available on the market at this time is software which will allow you to have dial-into "read-only" data bases. These programs set up user identification and then access to the data base. I am sure that you are acquainted with the public bulletin boards. The application of such a system would be to give users "on-line" access to library data bases and eliminate the need to go into the library before one knows that the material needed is there and theoretically it would "free-up" the public services person. Ideally it would save the user time as well as save the library valuable professional time that can be spent on other task.

EQUIPMENT/OPERATION/MAINTENANCE COST. The difference between a microcomputer-based system and a mainframe is astronomical. The Compaq 80386 microcomputer is priced in the neighborhood of $8,000, and if recent marketing history continues this price will be dropping. A mainframe with similar capabilities would cost in the neighborhood of $150,000. Both would need additional equipment (printers, software, etc.).

Microcomputers were built for individuals to use and to operate interactively. They were planned for small businesses, individuals, libraries, or schools; however, the 80286 is faster than

the mainframes of ten years ago and address much more memory. Also remember mainframes are designed to have a full-time "operator". Thus above the initial cost of equipment, one must hire a full-time operator to do the day by day chores required to run a mainframe. Mainframes must also have special controlled environments, which in itself is expensive.

The microcomputers are designed to plug-in and run; as opposed to the mainframe which require a maintenance contract. These contracts often cost more per-year than the microcomputer systems do initially.

There is, however, considerably more responsibility placed on the buyer of the microcomputer system. One needs to learn about the hardware, the software and know what exactly is needed, then take on, somewhat, the responsibility for the system design. This is true only to the extent that the real needs must be identified. There will be some maintenance on any system; however, microcomputer architecture has been tested and has been found to be very reliable.

RECORD COST (INPUT & MAINTENANCE/UPDATES). In looking for comparative cost between these two, there was found no reliable comparison. It appears that it would be like comparing oranges and apples; however, input record cost should be comparable. Records can be input by purchasing them from a supplier or from the keyboard.

The comparison of maintenance/updates of records between the microcomputer and mainframe was much more difficult. Looking at possibly weekly updates on a mainframe with the data base and the update being preserved separately (in the event of a crash) then reconstructing of the data base from the backup and updates is already known and standardly used in mainframe programing technology. Safeguarding the data base is standard practice in programing; however, it is costly. There are usually several (possibly the last six) updates kept, dropping the oldest updated data base each time. With a microcomputer, similar safegaurds must be used; the major difference is the size of the memory required. Comparison is made between the cost of maintaining copies of, let us say ASFIS as compared to the cost of maintaining backups of our own local data base. Stream taping is normally used with microcomputers. Even if the cost of backups per record were the same, maintaining backups for several million records would logically be more than maintaining backups for several thousand records.

MULTI-USER. Both systems are or can be multi-user. The speed or response time is both software and hardware dependent. There is no reason why a reasonable number of users could not be placed on either system. In the microcomputer system, by using a shell over DOS and a program such as Quick Connect2 in a 386 environment, up to ten users would not even be visible to one another (there would be no delay in response time) and up to 30 users with less than 2 second response time. Mainframes can have as many as 100 or more users with little or no delay. When we speak of "users" we are not limiting terminal connections we are speaking of "users at a given a time".

Again this approach allows the patrons of the local collection to access the collection from their office and allows outside users the same privilege.

To make full use of this technology and share resources, we should adopt standards in vocabulary use and in record format (such as MARC format); however, to search another collection (say for ILL purposes) that would not be necessary. We presently search several data bases which use differing search schemes and different vocabulary standards and retrieve what we want. One advantage of adopting a standard would be the sharing of records, we could copy free or purchase from another data base the record needed for our collection without rekeying.

SUMMARY

Hard/Soft-ware.

	MICROCOMPUTER	MAINFRAME
Memory	Virtual	Virtual
Communication	per use cost*	per use cost*
Equipment Cost (full system)	$35,000 (est)**	$200,000 (est)**
Operation cost normal personnel	normal personnel	designated operator.
Maintenance of Equipment	no annual contract (this does not mean that there is no cost)	annual contracts on both hardware and software.
Record Input Cost	comparable	comparable
Multi-user	yes	yes.

* Large central databases must charge for every use as apposed to cost only outside your own database.
** Verbal estimates from dealers.

General.

As information specialist, we seek to supply our local users with the information they require and at the same time organize our collection so that future users will have access without relying upon individuals. We have long-sinse passed the era where the "librarian" must keep all the information in his/her head; however, as you well know a great deal of that is still demanded of us.

With the present information explosion, we must rely upon fast access electronic equipment to help us help our users. The same problem that faced our predecessors still face us today and that is to give quick access to the information stored in our collections.

You are aware that there are still many problems which must be approached before such an idea could develop, only one being the compatibility of different brands of equipment to communicate. Probably the major problem would be security of the data base.

However, we are the organizers, the retrievers, the system designers of today. With the technology now available, we can begin to make our collections readibly available, at the touch of a keyboard.

LITERATURE CITED

1. Wells, Paul. 1986. Intel's 80386 Architecture. Byte. 10(11):89-106.
2. Computer Shopper. October 1987. 7(10):260.