TRENDS IN HARDWARE FOR GEOGRAPHIC INFORMATION SYSTEMS

Jack Dangermond
Scott Morehouse
Environmental Systems Research Institute
380 New York Street
Redlands, CA 92373

ABSTRACT

This paper presents a description of and comments on various trends in the hardware available for Geographic Information Systems (GIS's). After a brief introduction the paper deals with fast geoprocessing hardware, parallel processing, memory, workstations, networks, and hardware for specialized processing functions. Finally there are some comments on the UNIX operating system and on various peripheral devices including the need for new kinds of specialized workstations for GIS use.

INTRODUCTION

This paper describes and discusses some of the trends which we see in the hardware used in Geographic Information Systems (GIS's). We think such a paper may be useful because of the major impact that hardware developments have on geographic information system architecture, on GIS software, on costs and efficiency and on the whole nature of the user interface with GIS's. Hardware is certainly one of the three major drivers of developments in GIS's and probably the one in which change continues to occur most rapidly.

FAST GEOPROCESSING

There is now some interest in such tasks as processing geographic data for all of the North American continent or weather information for the whole world in near real time. Over the next 10 years CRAY-type machines with the speed and performance necessary for such tasks are going to become available at affordable prices, and1,000 MIPS CPU's within large organizations will become common.

Providing one or two network stations at people's desks, just to satisfy their need for interactive display and editing, doesn't solve the problem of data analysis for very, very large geographic data sets. In five years 20 to 30 MIPS workstation processors are likely. There is already discussion of 100 MIPS PC's in the early 1990's for about the same price paid for a PC today. At that time100 to 1,000 MIPS processors that are the file servers are also likely, and in the price range of today's minicomputers .

The implication for software developers is that they need to begin right now to develop the algorithms for that hardware environment; they need to concentrate on doing it correctly and not worry so much on whether it takes 5 minutes or 1 minute to actually perform a particular operation. They need to consider that performance with large databases is the real issue, not speed. Of course algorithms should be written economically and efficiently, but the spending of hundreds of hours to save minutes of time on a particular process is going to be of less interest given the state of the art of the hardware and the direction that hardware development is going.

PARALLEL PROCESSING

Even with 1,000 MIPS machines there will still be limitations on the performance of individual computing units based on the amount of circuitry that can be packed into a small space, the speed of light and the need to dissipate heat from the circuitry.

Because of these limitations, further dramatic breakthroughs in the performance of serial processing computers are not certain and order of magnitude increases in speed and computing power may have to be based on the use of parallel processing in which the computational task is divided among many processors. Even given a single task like generating a map or computing a polygon overlay, algorithms which inherently support parallel processing will benefit. Different computer manufacturers are now developing prototype multiple parallel processors and are beginning to provide support for parallel processing operating systems.

This development means that the software will have to change to take advantage of the parallel processing. There are two ways to handle true parallel geoprocessing. One is that the operating system breaks the problem into pieces, does all the optimization, does the sorting in one box, does each of the other needed operations in other separate boxes. The other approach is to program specialized parallel processors that take advantage of the fact that geography can inherently be spatially subdivided: for example, each of a hundred processors takes one-hundredth of the problem and does the polygon overlay on just that piece. This approach will require that we reorganize the way GIS software runs. The tiling structure in ESRI's map library system might be a suitable beginning point for such an approach: each of the map tiles would be dealt with by a single processor. Geoprocessing can also be viewed as a kind of pipelining through the various processing stages. If a process involves piping through eight stages these are now done sequentially; but if these stages could be piped to five different computers where the partial results of one were then fed to the others and acted as partial input for the next stage, then considerable time savings could be achieved. This may be the only way to get dramatic, orders of magnitude performance improvement.

MEMORY

The major trend in memory development seems to be toward lower and lower costs for ever larger memories. Indeed the trend might be said to be toward "zero cost" hardware solutions in memory and in mass storage devices. Nevertheless, at the present, memory costs remain significant.

A few years ago it appeared that laser read disks might offer a major breakthrough in this area. The present use of the laser read compact disks for storage, which now makes it possible to distribute very large geographic data bases to large sets of users by mailing each user a compact disk, is one present form of that technology. We can expect further developments in this area as the "writing" technology involved becomes more affordable to match the present low costs of the reading technology.

WORKSTATIONS

We are now seeing the gradual elimination of terminals in GIS hardware. They are being replaced by desk top computers. This means that for the price paid several years ago for a high resolution terminal, a high resolution screen together with compute power on the desk top can now be obtained.

By workstations, we mean interactive graphic workstations with very tight connections to the CPU so that integrated graphics and CPU power are available as a single tool, transparent to the user. One of the characteristics of the workstation is a very high graphic performance. Workstations add the ability to build much more sophisticated user interfaces than are possible with terminals by having the compute power and the high performance graphics on the desk top. Interactive graphic user interfaces, such as the MacIntosh or the Sun window/icon-based environments are possible. This will soon become a standard; people will expect this type of interface, which just isn't possible with terminals connected to central machines.

Workstations are of two kinds. One kind is the workstation that exists as part of a family of CPU's, like the VAX station and the high performance Sun or Apollo types. The other kind is the PC which is a workstation on a network. The latter often doesn't provide some of the networking and communication possibilities of the former; nevertheless, it's a very powerful standalone workstation that has some of the same fundamental components in it. Those workstations are becoming very fast. Upgrades to the PC/AT type systems (with something like the Intel 80386 chip running at 4 MIPS with a 32 bit word length) will make them equivalent to the MICROVAX in speed (and perhaps faster) for substantially lower costs. Moreover they can be linked with fast communication devices into larger networks.

What will be the impact of this kind of workstation? It will mean that processing, both analytic and graphic, can be done on desktop CPU's. So workstations and network technology will begin to have to work together in the GIS solution. One of the nodes on these networks still has to be a larger CPU for centralized processing and database management. The so-called file server device, which is a larger box shared by a number of workstations, has to be able to have information efficiently extracted from it, put onto local disks, and used locally, perhaps for updating, perhaps for analysis, perhaps for cartography, perhaps for reinsertion into a central library. Software solutions for organizing spatial libraries for rapid entry and extraction will work hand in hand with the hardware architecture of that central database. In our own software we have chosen the notion of spatial subdivisions in the form of tiles; these will allow pieces of the database to be pulled out into workstation environments for interactive work.

NETWORKS

In recent years the dominant hardware system architecture for GIS's has been the "multiuser host" system architecture; in this architecture one central processor supports a number of graphics and alphanumeric terminals. The single CPU provides computation, file and system management and drives the various shared peripheral devices (printers, plotters, etc.).

The network architecture is a rapidly emerging alternative to this multiuser host architecture. In this model there is a multiuser network rather than a single host. The network functions to integrate user workstations, compute servers, file servers, and shared peripherals. In such networks computing function is moved to single user workstations or to batch oriented compute servers. In its most extreme form each user has a dedicated workstation; less extreme networks have a combination of single user and multiuser nodes. The success of this approach is based on very fast networks making sharing of data practical and on inexpensive, high performance, single user workstations making the solution cost effective in comparison with the host oriented architecture. If such high performance workstations are available, the network is highly suited to the computation and graphics intensive applications of geoprocessing.

Hardware manufacturers are beginning to offer network architecture in their system configurations, meaning that their operating systems, their databases, and their communications, support the linking together of CPU's, and operate transparently to the end user. This means that different nodes within organizations have the ability to share and use data and software in common on a local area network. Wide area networks are also possible, making use of one of the various microwave transmission networks which are now being rapidly developing in the United States for the transmission of data over long distances. Hardware is thus being designed to support certain database concepts and this is very important for GIS. Instead of being just a set of tools, the hardware is becoming a fabric within which databases can easily be woven.

Also, because the networks usually use smaller CPU's as nodes, some of the degradation problems that we have had with one database being shared by many users are starting to go away; the hardware platform and architecture supplied by single vendors is becoming able to support a true database environment for geoprocessing.

A problem in using geographic databases in a network is that there are usually a lot of data to which the user needs access and there is also a lot of computational power needed on the desk top. PC's can give you the computational power on the desk top but you cannot share the data. A minicomputer can give you shared access to the data, but then you are competing for the same CPU with other people. The promise of networking is that it gives you the shared access and also gives you the performance at your desk top that you need for complex processing tasks and graphics.

We are just beginning to see the emergence of network architecture for CPU's by vendors such as Digital, Prime, Data General, Sun, Apollo and IBM. There are communication tools like Ethernet, that allow us to move among different vendors using TCP/IP, but we don't really see that successfully implemented at this time. Yet this is a trend that will eventually lead to multiple vendor hardware devices being connected in single networks. This will lead to both the inter-system and the intra-system database sharing.

One likely result of this trend is that there will be reduction in the power of the data processing management structures now in place; data processing managers and the whole computer center mentality will receive less emphasis and may eventually disappear. Instead transactional structures will be set up for updating, maintaining, managing and providing user access to the data.

HARDWARE FOR SPECIALIZED PROCESSING FUNCTIONS

One other trend is the development of hardware for specialized functions, especially relevant as movement toward networks occurs. Right now, in the network architectures, there are compute servers and file servers that do certain specialized functions for the network. There will be more development of other types of servers; for example, we already see sort servers and search servers such as the TRW Fast Data Finder and the Excel Sorting Engine. Other companies will be developing hardware assists for performance. Intergraph has done this a lot with their scanning graphics processors, and there will be an increasing trend toward moving geoprocessing into hardware to get increased performance.

In time, polygon overlay processing by intersecting tools may be built into a chip and become an integral part of geoprocessing machines. But there will have to be a much larger potential market for such chips before this occurs.

COMMENTS ON THE UNIX OPERATING SYSTEM

Despite some indications a few years ago that it might take over the operations system world, UNIX has not done so. Manufacturer's proprietary operating systems are still very strong and will continue to be supported indefinitely. Digital, with the introduction of their VAX cluster, local area cluster and the introduction of all their workstations that run VMS, is clearly committed to keeping VMS its operating system, or one of its two operating systems. IBM, with the introduction of minicomputers that run MBS and CMS, also have decided that they will continue to develop and support their proprietary operating system. What this means is that there will continue to be a diversity of operating system software and hardware and the differences between manufacturers will continue, thus making it more difficult to develop truly affordable software.

It will also continue to be difficult to connect different hardware devices in one big network. UNIX is not really an answer for making it all come together. Nevertheless, more and more boxes and workstations are committed to having UNIX as their only operating system.

Can the same kind of performance be obtained with UNIX operating systems by people like Sun or Apollo or others, that can be had from VMS or Primos operating systems? Probably these larger companies, because of their larger installed base, their larger proprietary interest, will devise new features in their operating systems that UNIX will not be able to keep up with. There will be attractive, end user-oriented types of solutions that UNIX just doesn't offer. The result will be two classes of machines. The inexpensive ones, based on operating systems like UNIX, will be fairly functional but won't do some of the really effective operations that proprietary operating systems will allow and offer.

One of the challenges for people who want to develop software that is machine independent will be to be able to develop very fast import/export procedures for traveling across machine environments with their data and also developing hardware which can run with standard compilers within different operating systems, with the operating system independent. One of the main things that we have done in ARC/INFO is the AML system which has freed us from some of the machine dependency on the command language processors which are resident in the operating system of each manufacturer; AML allows us to not only move across different vendor machines, but allows us to build macros which can move across them more effectively and pick up some of the key features from each of the different operating systems.

COMMENTS ON TRENDS IN PERIPHERAL DEVICES

The big trend in peripherals is the emergence of low cost reliable raster output devices. There are really two technologies now that are replacing the pen plotting type technology that we have used in the past. One is the electrostatic output, which is available in black and white and color. The other is the laser printer technology. The raster output device has great promise for increasing the quality of cartographic products because with the raster devices and the hardware that has been developed for these devices, half-tone patterns and high quality text fonts, high quality line symbols, and high quality point symbology are all possible and can be put directly onto the output media rather than having to be stroked in by pen. The other impact with the raster devices is that they can serve as input to pre-press page composition and graphic systems. Right now, there are only very expensive color, high resolution pre-press systems like Scitex's and inexpensive black and white small format systems like the Pagemaker on the MacIntosh. In time these two will converge and raster output from cartographic systems will go directly into high performance, color, pre-press systems.

A year or two ago scanners were being proposed as the new solution for input of cartographic data. While scanning has found its niche, it certainly hasn't taken over in cartography. Probably the main reason is that in a lot of data situations in cartography there is revision going on at the same time that manuscripts are being captured. The scanning technology still is not intelligent enough to deal with symbolized lines and many of the complicated text recognition problems.

NEW SPECIALIZED WORKSTATIONS

One of the devices that is badly needed for data entry is an instrument that allows data to be corrected as part of the data entry process. It will probably be a large, flat, display which will allow multiple planes of graphic memory to be used to display different layers of maps and images. It will allow compilation and adjustment to occur interactively among and between these layers relative to reference maps, base maps or images. What is needed next is hardware which can integrate display technology for pictures and vector screens and also serve as a platform for data entry and for overlay of manual maps and images. ESRI has performed manual map integration for about 15 years, using light tables and highly skilled craftsmen that know the relationships among various geographic data coverage types. They often make scientifically based decisions which result in cartographic adjustment to the data. With the introduction of this type of compilation tools, some kind of an interactive light table display mechanism with digitizer capabilities would be very powerful. We have also talked about using this instrument like an electronic sandbox to do thematic displays and suitability modeling in support of land use planning.

A host of new specialized workstations will also emerge. The analytic photogrammetric workstation is going to emerge into more and more of an integrated sort of device. The first ones were introduced this last year and they are still getting the bugs out, but some provide graphic super-imposition of database on the photogrammetric workstation.

Historically the whole orientation for workstation technology has been data capture, data editing, and some limited types of data use. There will be a trend in the next decade toward user workstations-- zoning workstations, land use planning workstations, water resources simulation workstations-- that actually are equipped with the kinds of output devices and analytic tools that allow users to do more than simply enter, display or edit data.

Human factors analysis has to enter into designing new kinds of workstations. One can envision the emergence of some kind of facilities, even entire rooms, that serve specific functions. Strategic planning rooms, emergency preparedness rooms, dispatch rooms, satellite tracking rooms, war rooms all have been equipped in the past with very crude ties to actual databases. Soon geoprocessing databases will be pipelined right into these rooms. Imagine full-wall CRT's which are touch-sensitive, panning and zooming on such a CRT, talking with it and having various tools to work directly with the data base. Rooms where users do geoprocessing will change. Users will not be sitting at a little CRT. Instead of just single persons, groups of people will be able to interact with the data; there will be informed conversations about situations; there will be more participatory planning.

CONCLUSIONS

The trend of rapid developments in hardware is continuing and these developments will have important effects in the next few years on the way in which geoprocessing is done and the way in which geographic information systems are structured.