DIGITAL MAPPING IN A DISTRIBUTED SYSTEM ENVIRONMENT

Dr. M. Mosaad Allam

Topographic Survey Division
Surveys and Mapping Branch
Department of Energy, Mines and Resources
615 Booth St., Ottawa
Canada K1A 0E9

ABSTRACT

Over the last two decades there has been a gradual advancement of digital technology within most mapping organizations, especially in the fields of photogrammetric computer mapping and computer-assisted cartography, the efficiency of which is greatly influenced by the processes of data acquisition, the encoding of various information for storage, and the manipulation of data within the system.

The current trend in system development (hardware, software and operation) is directed towards the configuration of sub-systems in a distributed computer network. The development of distributed systems is a natural outgrowth of the requirements to increase our computational abilities and to optimize the sub-systems resources. Even though the individual components operate autonomously, the common master data base is the unifying element. The configuration of the total system depends mainly on the user requirements, the availability of the system to various users, the source of digitization (aerial photography, existing maps, etc.), the structure of the common topographic/cartographic data base, the interaction between the various sub-system components and the form of output.

The functional requirements of distributed digital mapping systems are described. The various methods that can be used to achieve the interconnection between the various sub-systems, the communication mechanism and the characteristics of distributed data bases are also given.

INTRODUCTION

The most compelling benefit of digital mapping in a distributed system environment is the ability to share information. Distribution means not only one program can visit many different processors, but, conversely, that one processor and its associated data bases can be accessed by many different sub-systems. Time-sharing is not a new idea. Time-sharing was developed out of the need to obtain raw computing power at reasonable cost, and data sharing was initially viewed as a secondary benefit.

Today, the economics have changed considerably to the extent that raw computing power is available at low cost through microprocessors and minicomputers thereby reducing the need to share a processor for computational power alone. While this suggests that there may be a shift in emphasis from time-shared computers to stand-alone dedicated systems, the recent trends are directed towards the multicomputer configurations, mainly because of the need to share the data bases.

The increase in adapting digital technology in mapping organizations and the success achieved in automating the various topographic/ cartographic processes, coupled with the desire to optimize sub-systems resources are focussing the attention of system designers, managers and users to the need for building digital topographic data bases. In large mapping organizations, this can be achieved by automating the compilation of data base elements in the topographic map production processes. This includes the automation of the following functions: data acquisition, reduction, indexing, storage/retrieval, computation, analysis, editing, data base maintenance/management, and generation of output in digital or analog forms.

To achieve these objectives, it is necessary to design and develop a plan from a systems approach which integrates the following various functional activities into a unified production system:

- computer-aided photogrammetric map compilation,
- digital compilation using automated photogrammetric instruments,
- graphic digitization (manual, semi-automatic and/or fully automatic digitizing devices),
- automated cartography,
- editing and/or digital map revision,
- generation and manipulation of local data bases by various sub-systems, according to their functional activity,
- translation (restructuring) of local data base elements into a master data base,
- graphic output and digital output sub-systems, and
- etc.

The analysis of various production sub-systems (activities) clearly indicates that a distributed system environment is the most practical and efficient methodology to utilize in the system design.

FUNCTIONAL REQUIREMENTS OF DISTRIBUTED DIGITAL MAPPING SYSTEM

The functional requirements for a digital mapping system in a distributed environment can be summarized as follows:

1. Data Acquisition

The source material used for the digital compilation of topographic/ cartographic products are received in both analog or digital form. In the analog form, we find aerial photography and graphic maps, while digital sources include scanned aerial photographs or digital imagery acquired by airborne solid-state linear array cameras. The essence then is to capture the data digitally in order to provide the necessary input information for data base manipulation procedures.

1.1 Digitization from Analog Aerial Photographs

Computer-aided photogrammetric map compilation using analog aerial photography as source is achieved by the interfacing of photogrammetric plotters to digital computers. Varius system configurations have been successfully developed in numerous mapping organizations. As a prime example, we find the multiple-station digital photogrammetric compilation systems developed in the Topographical Survey Division, Surveys and Mapping Branch, Ottawa, Canada (Allam, 1979).

1.2 Digitization of Existing Graphics

Graphic sources are existing maps, photogrammetric manuscripts, orthophoto maps, etc., and the process of digitization is regarded as a conversion of data records from analog to digital form. This is achieved by:

a) manually retracing the graphic features on a digitizer table with a cursor or stylus,
b) employing a semi-automatic line-following device, or
c) processing the entire graphic automatically with a raster scanning digitizer.

In the first two methods, the output is a series of X and Y coordinates in the digitizer coordinate system. The third method produces a digital output in a raster format, which for data purposes requires conversion from a raster to a vector format.

Systems for digitizing existing graphics consist of multiple work stations (terminals) for data entry, display, editing, etc., interfaced to a digital computer.

1.3 Compilation from Digital Imagery

Digital stereo imagery may be acquired by scanning aerial photographs on automated photogrammetric instruments or in real-time by means of airborne linear array cameras. The compilation of topographic detail requires an automatic correlation (matching) of the conjugate digital stereo-imagery, while the compilation of planimetric detail requires a solution of the complex problems of digital image analysis and pattern recognition. This mode of compilation is heavily dependent on processing using digital computers.

2. Data Storage and Retrieval

Depending on the mode of data acquisition, the data is stored in a retrieval format in the sub-systems local data base. To build a master digital data base, the data is restructured according to a standardized digital data base, also in a retrievable format. The stored data must include positional coordinates and sufficient description information to make it useful across a spectrum of digital mapping products. In addition, they are stored at a resolution capable of supporting the majority of the products.

3. Data Manipulation

When a specific product is requested, the data contained in the master data base (or local data base) within the geographic area of interest is extracted and is used as input to the digitial manipulation process. This function requires several user dependent operations, and includes:

- selection of specific feature types,
- display of digital data,
- cartographic editing and enhancement,
- data compression,
- projection adjustment and transformations,

- mathematical processing/restructuring (e.g. generation of digital elevation models - DEM's),
- symbolization,
- statistical analysis,
- report generation, etc.

The output of the process may be a digital file structured according to a standardized digital exchange format or a digital file that can be used to drive a finishing device.

4. Graphic Output

This function entails the use of the digital data desired to produce the necessary map/chart overlays using automatic plotting/scribing devices. This function also includes a quality control task and other manual processes as required.

In addition to these four basic tasks a digital mapping system in a distributed computer network will support other functions such as:

- resource sharing
- system management and control
- batch processing and program development
- multiple user defined data base inquiry operations

DEVELOPMENT OF DISTRIBUTED DIGITAL MAPPING SYSTEM

A distributed system is achieved by interconnecting the computers of the digital mapping sub-systems in a manner that provides optimum interchange of digital data between the various sub-systems, in addition to the inter- change of data between the various terminals, graphic work stations and peripherals within each sub-system. In essence, the development of a distributed digital mapping system may be regarded as a computer network including hosts (central) computers, nodes (minicomputers), terminals (work stations, digitizers, video terminal, graphic editing stations, etc.), and transmission links.

A host in this context is a central computer whose function is separate from that of switching data. A node is a computer that primarily is only a switch. Some designs permit a computer to be both node and host. Terminals are interface between the user and computer network. Transmission links join the hosts, nodes and terminals together to form a network. A path is a series of end-to-end links that establishes a route across part of the network. The links and nodes along with the essential control, make up the communication sub-net, or the digital mapping data network.

Classification Schemes of Distributed Systems

Network topology, as a mean for categorizing digital data communication networks evolved from graph theory (Deo, 1974). Terms borrowed from graph theory include the previously defined host, node, link and path. This terminology forms the basis for topological classification of data networks as follows:

1. Centralized Network

It is essentially a star configuration (links radiating from a single node), is the simplest arrangement, as shown in Figure 1. This topological scheme requires a link to be dedicated between the central computer (node) and each terminal. This type of configuration is normally found in digital mapping sub-systems dedicated for one activity, e.g. computer-aided photogrammetric compilation. The reliability of the centralized network depends on the central computer (node). Its failure suspends all activity in the network, whereas individual link failures affect only a single device per link.

2. Centralized Network with Concentrators/Multiplexers

Where several terminals in a geographically dispersed system are close to one another, they are often connected to concentrators or multiplexers, which in turn are connected to the central computer as shown in Figure 2. The devices (normally microprocessors) obtain more efficient link utilization at the expense of an occasional delay in response time. A multiplexer is used where the information transfer rate of all simultaneously active terminals never exceeds the information transfer capacity of the link to the central node; but where the potential input capacity exceeds link capacity a concentrator is necessary because of its storage capacity. Concentrators can also merge several low-speed links into one high-speed link. However, concentrators and multiplexers cannot save a network when the central node fails.

3. Decentralized Networks

As shown in Figure 3, a decentralized network is an expanded centralized configuration with the concentrators/multiplexers replaced with minicomputers or microcomputers (nodes), having more memory and storage capacity. The added reliability of decentralized processing power comes at the expense of additional computers (nodes) and corresponding connecting links which permit some paths to be duplicated. Within reason, the reliability of the network can be increased almost indifinitely by duplicating paths and adding links and nodes.

4. Distributed Networks

As shown in Figure 4, a distributed network is similar to the decentralized configuration, but with each node connected to at least two other nodes in mesh form. A single ring is the simplest form, with each node connected exactly to the two other nodes. Each node can switch systematically between links according to a routing algorithm, to maximize the capacity of the network.

Distributed networks could be designed with a homogeneous structure, that is identical hardware at each node or hierarchical. In hierarchical configuration, there exists several sub-nets connected to one or more nodes. This hierarchy will allow for the use of various types of computers with varying power and function at each level and the backbone processor of such a system configuration is designated for the master data base. This type of hierarchical form makes it difficult to differentiate between the decentralized network structure and the distributed network and presently most designers can no longer distinguish between the two classes.

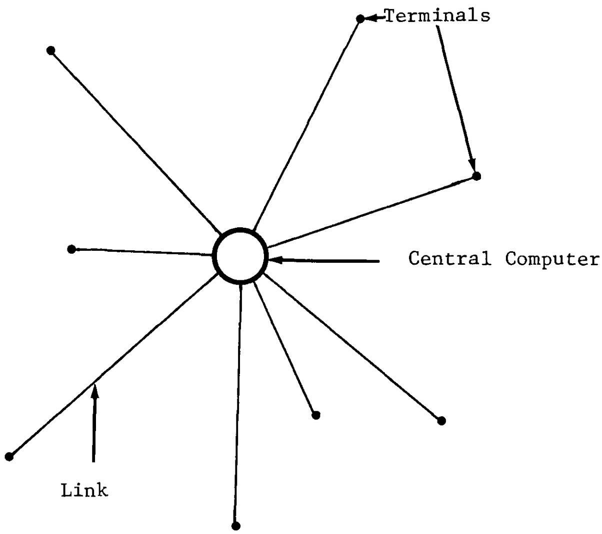

Figure 1 Centralized Network

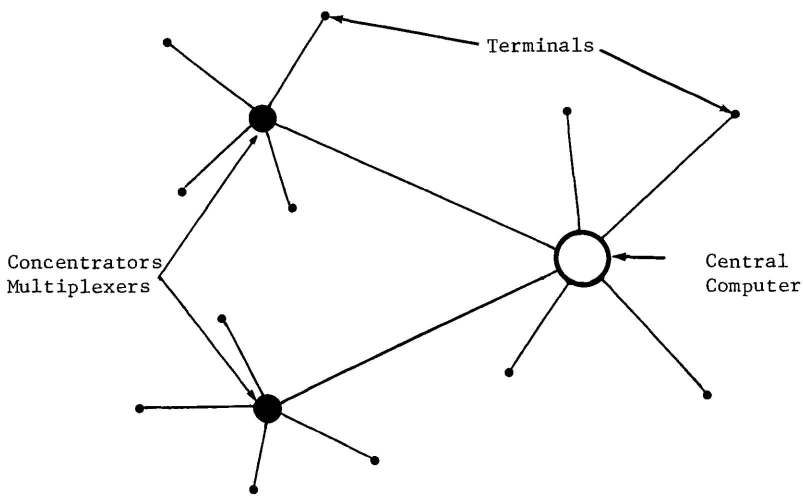

Figure 2 Centralized Network with Concentrators / Multiplexers

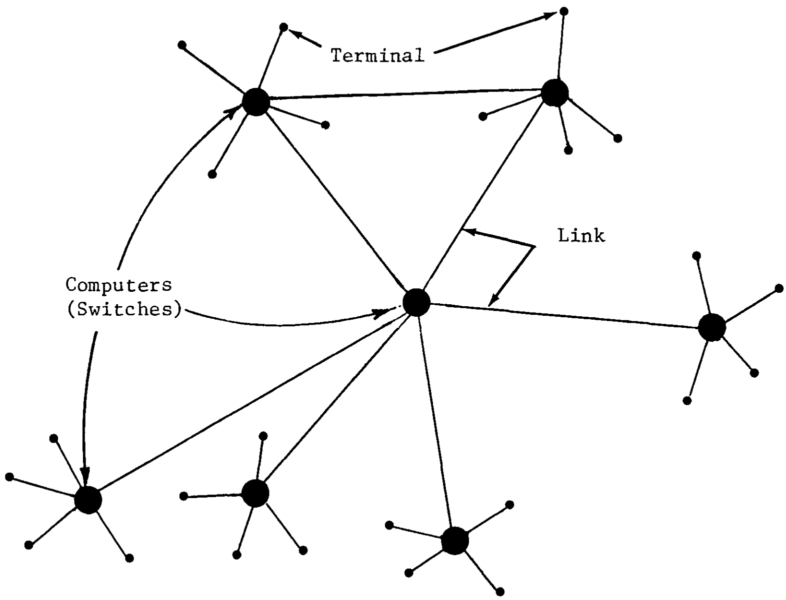

Figure 3 Decentralized Network

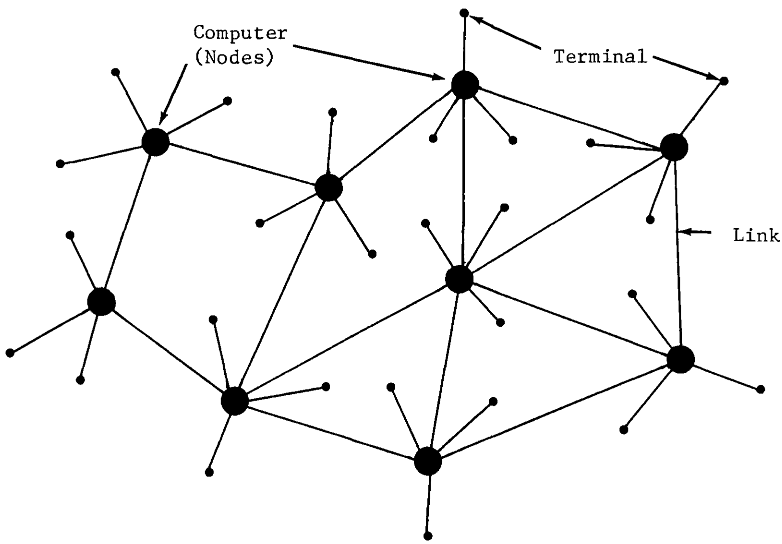

Figure 4 Distributed Network

COMMUNICATION AND NETWORKING

As shown in Figure 5, it would be a relatively straightforward task to design and build a communication mechanism whereby two programs, residing in the same computer, can send information or messages to each other. This is what is commonly known as interprocess communication facilities in operating systems.

In a distributed system the communication between programs reside in separate processors and hence a distributed communication mechanism must be created. Two halves of a single mechanism or service must communicate and synchronize to perform a unified function. This communication is used not only to pass information between the systems but also to pass control information between the communication mechanism programs them- selves. For that, a communication protocol is needed and is the only connection between the two distributed parts as shown in Figure 6.

Software is needed to enable multiple digital computers to be interconnected. This software is normally implemented as a set of layered network protocols, each of which is designed to fulfill specific functions whithin the network.

Software for each digital mapping sub-system activity, whether a data acquisition sub-system(s), data manipulation sub-system, data base sub-system or output sub-system, will be implemented as application software. This software must be designed and developed to achieve a synchronization of the system activity across multiple processors with the objective of sharing local data bases within each sub-system and the master data base containing the central data files.

DISTRIBUTED DATA BASES

The successful implementation of many distributed systems requires solutions to problems of data management in a distributed environment. These problems include data collection, concurrency control and update processing, failure recovery, and query processing.

In a distributed digital mapping system, data collection is fulfilled by the data acquisition sub-systems, and creation of local data bases. To collect the data in a central (master) data base, translation programs are necessary to reformat the files according to a standardized data base structure.

The coordination of concurrent updates initiated by multiple users must be done in a way that preserves the consistency of the data base. Updating should be made according to a data base manager program which acts on updates based on their sequence and/or time of requests.

Distributed system should have a failure recovery mechanism allowing the system to operate in a resilient manner and preserving the integrity of the data base. Ideally the system should continue functioning, although perhaps in a degraded mode.

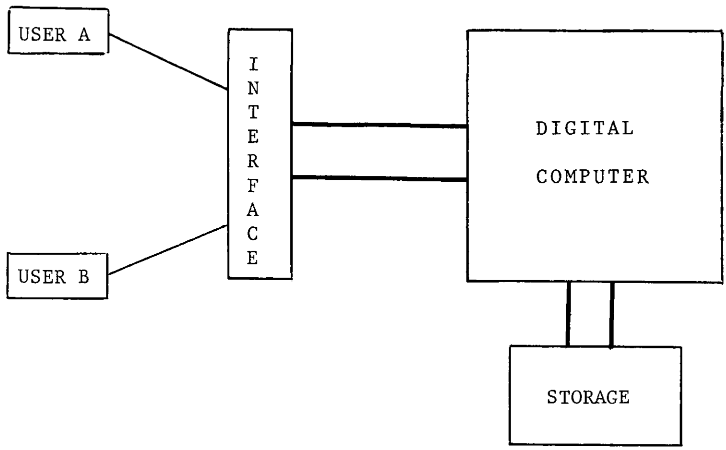

Figure 5 Local or Centralized Communication Mechanism

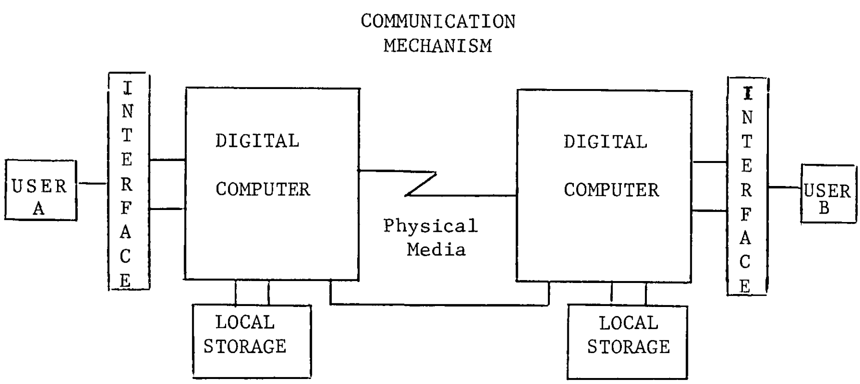

Figure 6 Distributed Communication Mechanism

Locating stored data in a distributed system will be simplified if the data is translated and transferred from the local data bases, in each digital mapping sub-system, to the master data base. If this is not the case, the distributed system program must determine where the data required by a query is stored. This problem can be solved by using a distributed data base management system, which allows the data to be sorted at multiple locations in the network and yet accessed as a single unified data base.

In a digital mapping environment, the approach of constructing a central data base from local data bases generated by various sub-systems will be preferred, due to the need for the exchange of digital map data among various organization/users, not directly linked to the distributed digital mapping system.

CONCLUSION

Distributed data processing systems are a new class of organizations and operations that exhibit a high degree of distribution in all dimensions, as well as a high degree of cooperative autonomy in their overall operation and interaction. By interconnecting the digital computers controlling various digital mapping sub-systems in a distributed environment, large mapping organizations will benefit from the improved performance, improved system availability, easier adjustment to a growing workload, extended control over computer resources, and the advantage of master digital data base. The dramatic drop in the cost of computers versus the increased cost of software and computer personnel makes distributed systems inevitable.

REFERENCES

1. Allam, M.M., 1979, Multiple Station Interactive Map Compilation with Distributive Computer Network, Proceeding Auto Carto IV, Reston, Virginia, November 1979, PP. 379-404.

2. Deo, N., 1974, Graph Theory with Applications to Engineering and Computer Science, Englewood Cliffs, N.J., Prentice-Hall, 1974.