GEOGRAPHIC INFORMATION SYSTEM
CONCEPTUAL DESIGN
FOR
WEST VIRGINIA

Submitted to:

Mr. Thomas E. Holder, Project Coordinator
West Virginia Development Office
West Virginia State Capitol
Building 6, Room B-553
Charleston, West Virginia 25305-0311
(304) 558-4010

Submitted by:

Mr. Peter L. Croswell, Executive Consultant
Tom Herrick, Design Analyst
PlanGraphics, Inc.
202 West Main Street, Suite 200
Frankfort, Kentucky 40601
(502) 223-1501

February 17, 1993

The preparation of this report was financed in part by an Appalachian Regional Commission grant to the West Virgnia Development Office (Project Number WV-11005-92-I-302-0310; Contract Number 92-76).

TABLE OF CONTENTS

Title **Page**

TABLE OF CONTENTS (continued)

Appendices

List of Figures

TABLE OF CONTENTS (continued)

List of Tables **Page**

SECTION 1
INTRODUCTION

1.1 PURPOSE

This report presents a design concept for a statewide GIS in West Virginia. Participants in this planning project may use this concept as a basis for strategic planning and detailed designs. The concept is a description of the fundamental components required as a foundation for West Virginia's GIS. PlanGraphics developed a concept for three primary components of the planned GIS, and they are:

- GIS database model
- Computer hardware configuration and communication networks
- Organizational structure.

For each of the major components, PlanGraphics considered the land-related activities and the potential GIS applications of the varied group of participants in this study. These considerations were documented in the previous study report, West *Virginia GIS Needs Assessment.*

1.2 BASIC PRINCIPLES AND GOALS GUIDING THE GIS DESIGN

1.2.1 Overall Goals

The basic principle guiding PlanGraphics' recommendations for this GIS concept is to coordinate the management and sharing of geographic information. The concept is developed with a goal of reducing redundant collection and storage of information and improving the accuracy and reliability of that data among the study participants. Overall goals that guide the GIS concept include the following:

- Establish a structure to coordinate GIS development efforts, in order to minimize duplicative expenditure of labor and funds.

- Increase the productivity of management and staff in the exercise of their land-related activities.

- Improve the quality and availability of geographic information available to support decision-making and management.

Major objectives that must be reached to meet these goals include:

- Establishing and enforcing data standards to facilitate use of information that may be managed by many different organizations

- Facilitating access to data and GIS functionality by multiple users
- Developing databases that support GIS applications with the greatest utility for multiple organizations
- Pooling financial, staff, and technical resources to build the GIS.

1.2.2 Major Factors Affecting West Virginia's GIS Project

There are several major factors that affected preparation of this GIS concept. PlanGraphics considered the following pervasive factors in all aspects of this project:

- The participants in this study are a very diverse collection of organizations, including federal, state, regional, and local government; universities; large publicly regulated utility firms; and small private firms, many of whom share some commonality in their need for geographic information, and yet are diverse in the scope of their activities. Opportunities for these organizations to share information should be considered when evaluating approaches to developing a statewide GIS. The West Virginia Geographic Information Systems Coordinating Committee (WVGISCC) could play a role in facilitating information exchange, although the future activity and authority of this body cannot be predicted at this time.
- The participants in this study are also spatially diverse. The largest single collection of participant organizations are state government agencies located in Charleston; the second largest contingent is located in Morgantown. Many of the other participants are distributed throughout the state. The wide spatial distribution of the participants adds dimension and complexity to issues of sharing and managing geographic data.
- Several participating organizations have already made significant use of computer mapping, GIS, and image processing software for certain projects. Experience with such public domain packages as GRASS and MOSS, as well as commercial software like AutoCAD, Intergraph, MIPS, ERDAS, SPANS, GeoSQL, and ARC/INFO, provides a base of knowledge about GIS. The diversity of experience with multiple systems is an obvious benefit to the overall project, but, is also a source of added complexity for sharing data and programs.
- The West Virginia Geographic Information Systems Coordinating Committee has established an ad hoc forum through which organizations, both experienced and inexperienced with GIS, can communicate. The nature of this organization encourages sharing of experience and ideas without the restrictions normally inherent with a formal committee. The activities of the WVGISCC are of significant benefit to the State of West Virginia in the development and implementation of GIS technology. The ad hoc forum has initiated and forged relationships between organizations that, otherwise, may not have occurred. When considering the activities and responsibilities of new, more formal,

organizations necessary to manage GIS development, the role of the WVGISCC is taken into account.

1.3 PHASED DEVELOPMENT

While the development of West Virginia GIS will be a continual process, four development phases are defined which identify key milestones in the process from Phase 1 (Detailed Design/Organizational Development) through Phase 4 (Mature Operations). All participants should adopt an attitude that the GIS is never “completed” or fully mature. In the long-term, there will be user needs that can be addressed by the GIS and technological changes which cannot be fully anticipated at this time. The intent of the system design, therefore, is to create a structure that allows the system to grow easily, to respond to new applications, and to take the best advantage of improvements in the technology. GIS development phases are explained below. These phases will form the basis for the GIS Plan.

In defining these phases, it is acknowledged that some state agencies, University groups, and private companies that have participated in this GIS planning process have been conducting GIS activities for several years. This existing experience with the technology provides a sound basis to proceed with a statewide coordinated approach to GIS. We recommend that current GIS activities continue, while a firm organizational foundation is being established for a long-term, multi-organizational GIS program.

1.3.1 Phase 1: Detailed Design/Initial Development

This phase will begin in mid-1993 and will last approximately 18 months. Its purpose is to lay a strong technical and institutional foundation for the GIS and detailed development to be initiated in Phase 2. This includes an organizational structure addressing state government agencies as well as other governmental and non-governmental organizations that are critical participants in the statewide GIS. A GIS pilot project will be planned and completed, and GIS activities currently ongoing in State Departments will continue within the framework of a statewide plan. While there will be a focus on state agencies during this phase, significant “outreach” efforts will occur with non-state organizations encouraging long-term participation and cost-sharing for the GIS. Where applicable, formal agreements will be established with non-state government groups for cost sharing and data exchange. Appropriate staff and GIS development responsibilities will be assigned, and a long-term organizational structure will be put in place. Technical activities will include development of database standards, procurement specifications and procedures for GIS hardware and software, initial database development, and implementation of several key applications.

1.3.2 Phase 2: Continued Development and Early Operation

At the start of this phase, which will last approximately two years, the organizational structure established in Phase 1 will be firmly in place, and initial applications relying on available portions of the GIS database will be in operation. Operating procedures for accessing and exchanging GIS data will be put in place, and GIS capabilities will be expanded to all major state agencies with high-priority GIS applications. Continued outreach to non-state government organizations will occur, and agreements not concluded in Phase 1 will be established. A critical objective of this phase will be completion of the GIS database layers needed to support high-priority applications and the development of key applications.

1.3.3 Phase 3: Expanded Operation

This phase will last approximately two years. During this period, all high-priority and many lower-priority applications will be developed and placed in operation, and development work will continue with lower priority applications. All database development work for GIS layers needed to support high and medium priority applications will be completed, and portions of the database needed for future applications will be developed. During this phase, there will be significant expansion of GIS capabilities to many organizations in the state, and batch data transfer over data communications lines will be in operation.

1.3.4 Phase 4: Mature Operations

Phase 4 is open-ended and describes a period of continued GIS expansion during which additional applications and users are added to the system. A key aspect of this phase is the adoption of higher-speed communications networks and more mature system and data standards that will permit more flexible distribution and exchange of GIS data. This phase is described to depict the long-term direction of the system—the specific nature of which will become more clear later in the development cycle. No specific configuration information and cost estimates are presented for this phase.

1.4 CONTENTS AND FORMAT OF REPORT

This report contains four additional sections that provide recommendations developed by PlanGraphics for West Virginia’s GIS project. The concept includes organizational and staffing concerns; a system configuration; and database development. A summary of the contents of each section of this report follows.

Section 2 deals with the GIS database and lays out a model for handling geographic data to satisfy application needs at a variety of scales and accuracy levels. This section includes a discussion of the content and general format of the database; a discussion of methods for developing the database; and cost estimates for development.

System configuration issues are addressed in Section 3. This section provides a description of the data communications model, links to outside systems, and a typical hardware configuration.

Section 4 provides background about critical issues for GIS management and staffing and lays out an organizational structure for the statewide GIS. It makes specific recommendations on staffing required to support management and coordination of the GIS.

Closing comments and a summary of conclusions are included in Section 5. In addition, this section provides a preview about issues that will be covered in the *GIS Development Plan,* which will follow this conceptual design.

Five appendices are also included as a companion document to this report. Each appendix provides basic tutorial information and a description of terms and concepts relating to computer hardware, software, data communications, and database issues which are intended to help readers more fully understand technical discussions in this report.

SECTION 2
DATABASE CONCEPT

2.1 INTRODUCTION

The GIS concept in this section is a description of the content and characteristics of the data needed to support development of West Virginia's initial GIS applications. An initial set of GIS applications was identified in the *West Virginia GIS Needs Assessment.* This database concept is intended for use as a guide to developing detailed database designs and specifications for data conversion.

For a general overview of the issues raised in this section, please reference GIS Database Concepts in Appendix A.

In this section, PlanGraphics presents our findings and recommendations regarding the GIS database concept. The overall concept was developed from our analysis of the following list of associated issues:

- Concept of GIS database tiers
- GIS application requirements
- Core GIS database recommendations
- Data conversion methodology, sources, and cost
- Relationship between GIS data and databases that reside on systems external to the GIS
- Data standards.

The system configuration concept, outlined in Section 3 of this report, will enable access to data from different automated systems, including different GISs. The database concept for West Virginia's GIS involves two general categories of data: 1) a core set of map layers and associated attributes that are routinely used by multiple agencies, and 2) external tabular databases maintained independently that can be linked with the map layers for a specific analysis.

The core GIS database contains all map features required by multiple users with a defined set of tabular attributes stored permanently in the core database. External databases can be linked with the core map layers as needed for analysis using the GIS. This view of the database is consistent with the tabular data linkage described in Appendix A. The linkage concept is that any database can be tied to a map location based on a unique geographic identifier.

2.2 DATABASE TIERS

2.2.1 Database Tier Concept

This project involves a variety of participants, each with different types of land-related activities. The focus of these activities ranges from parcel-specific site details to statewide aggregations of data. Perhaps ideally, one database with a singular level of detail and accuracy could meet the wide range of activities. However, this ideal is not yet practical in the current environment of data processing. The use of tiers is a practical approach for organizing the base of mapping and geographic data in tiers.

The concept of tiers, as applied to the GIS database, refers to developing separate groups of data at different levels of accuracy and detail. Different applications require different levels of accuracy and detail. For example, site analysis for a potential sewage treatment plant should be conducted using large-scale mapping with a high degree of accuracy; to prepare a statewide statistical map of current flow from treatment plants does not require the high degree of accuracy required for the site analysis. The difference between these applications, as related to a discussion of data tiers, is in data resolution. Practical data processing factors, such as data storage, database structure, and retrieval, must also be considered. A tiered database approach is designed to address the practical issues of GIS data storage and performance in a way that supplies users with the types of data, covering the appropriate areas, that are needed to support the different requirements of their land-related activities.

By structuring the GIS database into four tiers by level of detail, data will be organized efficiently to balance the trade-offs of data content, map accuracy, and geographic coverage that impact all land-related activities. Figure 2-1 indicates the relative area of coverage for a single base map file in each of the tiers. The tiers actually represent a continuum of accuracy from one level of detail to the next. The four tiers of the core base map data are described on the following pages and summarized in Table 2-1.

- *Local Data:* This tier consists of a series of map layers registered to a base map that has an absolute accuracy of approximately ±2.5 to ±10 feet. Data storage requirements for local data are significantly more than the other tiers. Because of the level of detail, the area of coverage for these tiers must be geographically subdivided into smaller areas to facilitate retrieval and processing. Many of the standard maps associated with the Local Data tier are produced at scales ranging from 1:1,200 (1" = 100') to 1:4,800 (1" = 400'). In some instances, base map scales may range from 1:600 (1" = 50') in densely urbanized areas to 1:12,000 (1" = 1,000') in areas that are quite rural.

  The map and tabular databases associated with this tier can be used to perform many site-specific or aggregated analyses, to identify or inquire about specific locations, and to create large-scale maps. Activities associated with this tier include tracking property ownership, tax assessment, case tracking, field operations, local infrastructure, and others. The primary users of databases that fall within the Local Tier category are local governments and utilities. State agencies that perform site-specific functions, such as hazardous waste contamination remediation, will often require data at the level of detail and accuracy associated with this tier. Databases in this tier will be primarily stored and maintained at the municipal or county government level.

- *Sub-regional Data*: Databases that are associated with the Sub-regional Tier will support land-related activities that typically require a larger view of an area than the local data tier supports. The Sub-regional data tier would provide map data appropriate for viewing an entire county, or perhaps multiple counties. The base map accuracy for this tier should be approximately ±40 feet. As in the local data tier, map data in the Sub-regional Tier will likely be divided into geographic sub-areas for storage and retrieval, although the extent of each map file will be larger than those used for the local data tier.

  For example, if your agency has data that are effectively shown on USGS 1:24,000 scale topographic maps, that data would be included in the Sub-regional data tier.

- *Regional Data:* Map data in this tier will be used to support land-related activities that typically require a view of larger regions. Map accuracy requirements for activities supported by this tier will be less than for the Sub-regional and Local Tiers. PlanGraphics recommends that the USGS 1:100,000 scale DLG files be used as the base map for the Regional Tier. These files are designed to maintain a horizontal accuracy tolerance of approximately 400 feet.

  Additional map data to be included in this tier may be input from sources at scales other than 1:100,000. For example, soils data recommended for this tier is available at an original input scale of 1:250,000. In some cases, it may be more appropriate to generalize data directly from the Sub-regional data tier. Because the regional databases contain less data, they are generally less costly to produce, easier to store and manage, and faster to process than Sub-regional databases, particularly when large areas are being mapped or analyzed. Map and tabular data in this tier will also likely be divided into sub-areas for storage and retrieval.

- *Statewide Data:* This tier is designed to meet the needs of land-related activities that require an overview of the entire State. This tier will depict political boundaries, the major transportation network, and other major locational reference features. PlanGraphics recommends that the base map for statewide data be the USGS 1:500,000 scale map of West Virginia.

  In large measure, map data at the Statewide Tier would be prepared as special maps for display or publication. Few permanent databases containing specialized thematic data are required at this level. Data at this level can be used for generalized thematic mapping or statewide presentations. To a large extent, this tier provides a large-area coverage of basic boundary and landmark features to serve as a foundation for special mapping projects.

2.2.2 Database Tier Recommendations

For each tier, PlanGraphics recommends particular factors for scale and accuracy. The recommended factors should be applied to the base map, which will be used in support of database development within that tier. Base maps of a known level of accuracy and detail will come to be associated with a particular tier by those involved with the planning and use of the GIS. PlanGraphics' recommendation for base maps for each of the tiers is shown in Table 2-1.

PlanGraphics has organized requirements for each of the GIS databases with one of the four tiers identified in this concept based on the scale and accuracy criteria. Our recommendations concerning databases will relate users' applications to a particular database tier.

Project participants should expect that details of specific scales and accuracy which are eventually adopted and developed will vary slightly from those stated here. PlanGraphics' recommendations should be viewed as a goal; in the interim, existing data may well be used that does not meet the ultimate recommendations.

Existing sources, funding levels, and other issues will eventually come together to refine details of the concept recommended in this report. However, we believe that the concept described hereunder will provide a model for the overall data requirements which will carry on throughout the GIS project.

2.3 GIS DATABASE RECOMMENDATIONS

2.3.1 Overview

In this subsection, we present our recommendations for the GIS database. These recommendations are primarily based on requirements associated with the GIS application categories identified in previous reports of this study. Tables 2-2 through 2-5 show the relationship between GIS application categories and data categories; each table represents a separate database tier. The relationships shown in these tables are the basis for the database concept.

Following the discussion of applications and data requirements, we present the overall GIS database recommendations. We have prepared four tables (Tables 2-6 through 2-9), each containing map features and core attributes associated with one of the four tiers.

The map data categories associated with GIS application categories are listed below:

- Survey Control
- Orthophoto/Planimetric
- Transportation Network
- Hydrography
- Topography
- Political Boundaries
- Administrative Boundaries
- Parcels
- Energy Transmission
- Utility Distribution/Collection
- Land Use/Land Cover
- Bedrock Geology
- Floodplains

- Well Locations
- Environmental Features
- Soil Classification Boundaries
- Incidents/Point Features
- Historical Archeological Sites
- Recreational Facilities
- Demography
- Meterological Data
- Biological Distribution.

2.3.2 Application Categories and Data Requirements

In the previous report, *West Virginia GIS Needs Assessment*, PlanGraphics identified categories of GIS applications that encompass the participants' land-related activities. It is on the basis of these application categories that PlanGraphics has prepared our recommendations for the database concept.

On the following pages are Tables 2-2 through 2-5. Those tables present the relationship between GIS application categories and map and tabular data associated with each database tier.

2.3.3 Data Category Descriptions

Each of the data categories identified in the previous tables are further described below. PlanGraphics presents our recommendations for the GIS database concept in tables and text below. Tables 2-6 through 2-9 contain PlanGraphics' recommendations for the graphic and core attribute data associated with each database tier. The tables are as follows:

- Table 2-6 - Local Tier Data Recommendations
- Table 2-7 - Sub-regional Tier Data Recommendations
- Table 2-8 - Regional Tier Data Recommendations
- Table 2-9 - Statewide Tier Data Recommendations.

It is not feasible to develop all tiers all at once, so, some priority should be established. PlanGraphics believes that the Regional Tier should be the initial target for development by West Virginia. Data to be included in the Regional Tier will support a relatively wide number of users early on in the project; and, the Regional Tier will be much less expensive and time-consuming than the Sub-regional and Local Tiers. The Statewide Tier is not recommended as the initial priority since it will support far fewer applications than the Regional Tier.

Our recommendation should not be interpreted to mean that all recommended data in the Regional Tier must be complete before going to the next tier. We are indicating a general priority for focusing the activities and resources from a statewide project perspective. We fully expect that it will be feasible to begin some databases within other tiers before the Regional Tier is completely populated.

While databases of the Regional Tier are being converted or acquired, management can begin to set up agreements for joint funding of the Sub-regional Tier. These agreements take time to establish in principle, and still more time to get into the funding cycle.

Following the tables is a description of each data category.

Table 2-9 cont'd

Survey Control and Grid

Survey control databases should include geodetic horizontal and vertical control points and reference grids. PlanGraphics recommends use of the NAD83 horizontal datum and NAVD88 vertical datum as the primary geodetic reference datums. Survey control databases should contain records of all monuments within a local jurisdiction. The state government should maintain records of all control points in the state that are established using approved methods to third order accuracy or better.

We also include the West Virginia State Plane Grid (WVSPG) under the category of survey control. Local Tier mapping and surveying is generally conducted in reference to the West Virginia State Plane Grid. Most coordinate values for state government mapping projects that fall into the category of Sub-Regional and Regional are in reference to the WVSP grid.

Survey control will play a larger role in the development and daily applications associated with the Local and Sub-Regional Tier databases than the Regional and Statewide Tier databases. The Local Tier survey control database should contain all control that exists within a city or county jurisdiction. This would include all geodetic control from first through third order.

The Sub-Regional Tier survey control database may contain all geodetic control from first to third order. The Regional Tier may contain only first and second order geodetic control. We do not recommend creating a database of survey control for the Statewide Tier; at this level, survey control could be plotted to display the general network layout.

Orthophoto/Planimetric

An orthophoto is a prepared aerial photo map that has been photogrammetrically altered to remove spatial distortions. A planimetric map shows surface features using lines and symbols. In the context of GIS database development, a photogrammetrically prepared orthophoto or planimetric map is used as a base to control the location of most other map information. This is particularly true for the Local and Sub-Regional Tiers where the importance of quantifying map accuracy and the ability to reliably analyze map overlays is more critical. Further, when base maps are plotted with other overlays, they provide a spatial reference that helps users identify the location of other features.

Some local communities have opted for planimetric mapping as their base map because of additional requirements for that data in digital form. A planimetric mapping project typically will include transportation and hydrographic features. For purposes of this database concept, transportation and hydrographic features are included in separate database layers because of their potential uses. Additional planimetric features that may be included under this layer, should planimetric mapping be selected, include building outlines, vegetation, retaining walls, cemeteries, substations, and other physical features.

PlanGraphics recommends that orthophotos be prepared to support development of Local Tier and Sub-Regional Tier databases. Orthophotos are much less expensive to produce than planimetric maps and provide a solid foundation for building local GIS databases. For the Local Tier, orthophoto, or planimetric, mapping should range in scale between 1:1,200 and 1:4,800. PlanGraphics recommends that orthophotos for the Sub-regional Tier be prepared at a scale of 1:24,000.

Sub-Regional Tier orthophoto mapping may be developed to fit the same coverage as USGS 1:24,000 scale quadrangles. The orthophotos provide many users with a richer base of information than a planimetric line and symbol map (USGS quad). The USGS and Soil Conservation Service have programs for developing orthophoto coverage, and they may be willing partners to share costs for such a project.

Another alternative for Sub-Regional Tier mapping may be to enter into an agreement with a private firm that has digitized USGS 1:24,000 scale quadrangles for their own purposes. CNG Corporation is one firm that has converted a significant number of quadrangle maps to digital form, and may be interested in establishing an agreement to supply that data, provided it meets other users' needs.

PlanGraphics does not recommend orthophoto mapping for the Regional or Statewide Tiers.

Transportation Features

Transportation features include roads, railroads, scenic trails, airports, waterways, and public transit routes. PlanGraphics recommends that each database tier include transportation features as vector data even if the base map selected is an orthophoto. Transportation features, and particularly roads, are used for many applications either as a general base map or as the primary data to support an application.

Transportation features shown in the Local and Sub-Regional Tiers should be organized in a topological structure. This structure will enable users to trace through the network to perform many queries and analytical applications. In the Local Tier, road segments should be related to address ranges in order to support address-matching queries and computer-aided dispatching. For the Local Tier, the apparent road centerline can be digitized directly from orthophoto base mapping. Sub-Regional and Regional transportation features may be digitized from existing USGS quadrangle maps or from highway maps.

For the Regional and Statewide tiers, PlanGraphics recommends that transportation features include those selected by USGS for presentation on their 1:100,000 and 1:500,000 scale maps. These features should be updated with information available from the West Virginia Geological and Economic Survey and from the West Virginia Department of Transportation.

Surface Hydrography

Surface hydrographic features include rivers, streams, lakes, and reservoirs. Hydrographic features may also be captured by direct stereocompilation (on the photogrammetric stereoplotter), by trace digitizing from orthophoto base mapping, or from existing topographic maps. PlanGraphics recommends that the hydrographic features be organized in a topological structure for the Local and Sub-Regional Tier databases. Topologically structured features will enable users to conduct network tracing and other query functions on the hydrographic database.

For the Statewide, and perhaps Regional Tier, databases, the hydrographic features need not be topologically structured. The 1:250,000 and 1:500,000 scale USGS maps are appropriate sources for surface hydrography for the Regional and Statewide Tiers, respectively.

Local tier hydrographic data would be acquired through means of a county or municipal photogrammetric mapping project (e.g., City of Beckley, City of Morgantown, Jefferson County).

Topography

Topography includes map features that define elevation and the characteristics of landforms. Topographic map features include contour lines, spot elevations, shaded relief, and digital elevation models (DEMs). Topographic data may be stored in three basic formats: a) as points with X, Y, and Z values, b) as triangular irregular networks (TINs), or as, c) grid cells. All three formats have advantages. As part of a Local Tier database, elevation data stored as points are ready for use in most computer-aided engineering systems. The TIN format in a Local Tier, or perhaps Sub-Regional Tier, database enables users to create highly accurate interpolated contours or slope mapping with relative quickness. The grid cell format is generally more appropriate for the smaller scale, less detailed Regional and Statewide Tiers.

For many Local Tier users, topographic data may be a “luxury” item due to the cost, particularly so in rural areas. Topographic data for a Local Tier database may include a DEM and contours, with the contours at intervals between two and ten feet. Two foot contour intervals are mostly associated with mapping at a scale of 1:1,200; ten foot intervals are associated with the 1:4,800 scale. If users of the Local Tier databases have need for topographic data, PlanGraphics recommends that the users acquire a DEM in addition to the more traditional contours. The DEM can directly support analyses for preliminary earthworks projects as well as watershed and storm drainage.

A DEM could be produced as part of a Sub-Regional Tier digital orthophoto production effort and would provide the best procedure for maintaining accuracy during data capture. A less accurate alternative is the interpolation of DEM elevations from USGS 7.5 minute topographic quadrangles. The latter process is being used by the USGS to create their DEM files. Topographic data along hydrographic features in this tier may be used to support hydrologic modeling.

The Division of Environmental Protection already has acquired USGS 1:250,000 scale DEM files. PlanGraphics recommends that these files be used as the topographic data for the Regional Tier database.

There is generally little requirement for topographic data at the Statewide Tier level, except for general portrayal. For the Statewide Tier, if users need to portray topographic data for the entire state, PlanGraphics recommends that contours shown on the 1:500,000 scale state map by USGS be digitized.

Political Boundaries

Political boundaries include legally established lines of incorporation or charter, such as cities, counties, or state.

Political boundaries change; some, in areas of intense development, change on a relatively frequent basis. State government activities that require up-to-date county and municipal boundaries should use the data prepared by local governments. A procedure to acquire political boundary updates, either in digital or hard copy form, should be established with the local jurisdictions.

Political boundaries for the Local and Sub-Regional Tier should include all city, town, and county boundaries. These boundaries should be derived from the legal description for the Local Tier. For the Regional and Statewide Tiers, political boundaries may include only counties and the larger cities.

Administrative Boundaries

Administrative districts define areas of jurisdiction, service areas, or statistical units. These districts include census tracts, regional planning council areas, utility rate areas, appraisal districts, soil conservation districts, and others.

Districts represented in the Sub-Regional and Regional Tier will generally be those which are established by state organizations. Sub-Regional tier administrative boundaries would be prepared using map sources showing the districts. Entry to the GIS should be accomplished after plotting the boundaries from existing maps to the Sub-regional base map.

Some administrative districts are established by state law, but are managed by local boards. Local boards typically have a need to be able to relate district boundaries to parcel boundaries. This application would take place using data from the Local Tier. However, state and federal agencies have need for current and accurate portrayals of the district boundaries. In this and similar instances, PlanGraphics recommends that a procedure be established for the local boards to supply the state and/or federal agency with boundary updates on a timely basis.

Administrative boundaries in the Regional Tier database could best be derived from the Sub-regional Tier.

The Statewide Tier may include U.S. Congressional, U.S. Senatorial, and West Virginia state legislature districts.

Parcels

Parcels are the limits of property ownership, and may include either or both surface ownership or subsurface mineral ownership. For purposes of simplifying ongoing discussions, road rights-of-way and legal easements are included under the general map data heading of parcels.

Comprehensive property mapping is appropriate at the Local Tier. This level of effort is outside the scope of the statewide GIS planning project. However, over time, detailed parcel mapping may be available to state agencies as part of a coordinated effort for standards and data sharing between the county assessors and the state. The remaining tiers are at too small of a scale to support a detailed, comprehensive property mapping base. However, large parcels of interest to state or federal agencies may be included in the Sub-regional and Regional Tiers. These may be tracts of land exceeding 20 acres, as a minimum size, that may be mapped using the smaller scale base maps associated with the aforementioned tiers.

The West Virginia Department of Tax and Revenue may wish to work with the county assessors to develop a comprehensive statewide standard for parcel mapping in the Local Tier. These standards would address issues such as content, map accuracy, and data format. Other states that have taken this approach include Kansas, North Carolina, and Oklahoma. Efforts to establish parcel mapping standards for the Local Tier should not affect timing for the overall statewide GIS project.

Energy Transmission Features

Physical features related to power/energy transmission, including electric generation plants, substations, electric transmission lines, gas transmission lines, oil pipelines, etc. Transmission differs from distribution in that energy is transmitted between major population centers where it is distributed to the consumer.

The USGS 7.5 minute quadrangles include above and below ground energy transmission facilities. It appears that the quad maps are an appropriate source for energy transmission facilities at the level of a Sub-Regional Tier. The Sub-Regional Tier orthophoto base map can be used to identify transmission lines that have been constructed since the last USGS quad update.

Utility Distribution/Collection

Utility features include the infrastructure of sewer, water, electric, gas, and other distribution and collection facilities.

The components of utility infrastructure systems can be mapped appropriately at Local Tier scales. Utility features can be plotted using the orthophoto as the base for their location. In general, a utility infrastructure database is used by the provider to more effectively manage the utility. Additional uses of these databases include development planning and background data for locating underground utilities in the field. Details of the data to be included in Local Tier utility databases are largely dependent on local conditions and informational needs.

The only utility distribution or collection features appropriate for mapping in the Sub-Regional Tier are the larger facilities such as substations, pumping stations, and treatment plants. At this tier, other features are too detailed to be effectively mapped. Users that need utility infrastructure features should refer to the Local Tier.

Due to their small scales, PlanGraphics does not recommend that the Regional and Statewide Tiers include utility distribution or collection data.

Land Cover/Land Use

Land cover refers to the physical surface characteristics of the land (e.g., vegetative type, water or wetland cover, bare ground). Land use refers to the human activity on the surface (e.g., timber harvesting, urban use, agricultural use). In the Local Tier, databases may also include zoning under this data category. These two geographical feature types are closely related but are logically separate, particularly when mapped at a detailed, large-scale level.

At the local government level, land use is generally more important than land cover. Land use boundaries will generally follow parcel boundaries in urban and suburban areas. However, in rural areas, one parcel may encompass several uses. Orthophoto base maps and parcel maps can be used to plot the location of land use boundaries.

In addition to land use, zoning boundaries and classification may become important for local governments. Like land use boundaries, zoning boundaries are generally coincident with parcel boundaries. PlanGraphics recommends that, where zoning ordinances exist, zoning boundaries also be plotted using the orthophoto base maps and parcel maps in the Local Tier. Zoning boundaries should be updated in the Local Tier database as soon as the action is complete.

Land use and land cover mapping may be accomplished in the Sub-Regional Tier using USGS 7.5 minute quadrangles or orthophoto quadrangles as the base map. Data for land use analysis may be derived from SPOT Image satellite coverage or from Landsat imagery. Since there is little, if any, land use/land cover mapping being created at the Local Tier, the Sub-Regional Tier would likely contain the most detailed information of this type. A project to prepare land use/land cover mapping at 1:24,000 scale would be a large and expensive undertaking. PlanGraphics recommends that land use/land cover mapping be acquired for the Regional Tier first. On occasion, specific projects include gathering and analysis of land use and land cover data. The land use/land cover mapping that results from these projects should be stored as part of the GIS database and can be the basis of land cover/land use in the Sub-Regional Tier.

The Regional Tier may include the LandUse/Land Cover data prepared by the U.S. Geological Survey. These maps are available in digital form (GIRAS data) at a scale of 1:250,000 which may be purchased and imported to the Regional Tier database. Since these products are somewhat generalized and out of date, a better approach may be to re-map land cover using a combination of satellite imagery and aerial photography. Such a project could use the 1:250,000 scale USGS data as a foundation with satellite data and other sources to update and augment that data. Another possible source to aid in land cover mapping are the wetland maps resulting from the National Wetland Inventory (NWI) being conducted by the U.S. Fish and Wildlife Service (USFWS).

PlanGraphics does not recommend that a database be prepared for land use or land cover for the Statewide Tier. Statewide representation of these factors could be derived as needed from the Regional Tier.

Bedrock Geology

Bedrock geology data will primarily be a series of boundaries indicating rock type and geologic classification, including unconsolidated stream or talus deposits. The data may also include spot features such as outcrops and linear features such as fault lines.

PlanGraphics does not recommend that bedrock geology be mapped specially for the Local Tier database. The effort of collecting data at this level of detail would simply be too costly for the benefit that might be derived. Local Tier data users would best use geological data prepared for the Sub-Regional Tier.

The geological quadrangles at 1:24,000 scale are the best sources for data at the Sub-Regional Tier. However, this mapping is not yet complete. Most of West Virginia is covered by geological mapping at 1:62,500 scale. A program was recently approved by Congress for national geological mapping which could accelerate geologic mapping in West Virginia at the 1:24,000 scale. PlanGraphics recommends that where geologic quadrangles at 1:24,000 exist, these data be input to the Sub-regional Tier bedrock geology database. Owing to the fact that these geological boundaries are mapped using the USGS quadrangles as the base map, we expect that no further cartographic adjustment will be required.

Geologic mapping has also been prepared by the West Virginia Geological and Economic Survey at a scale of 1:250,000. PlanGraphics recommends that geological mapping at 1:250,000 scale be included in databases for the Regional and Statewide Tiers, respectively.

Floodplains

Floodplains are areas identified as prone to flooding and fall into different categories established or carried on by the Federal Emergency Management Agency (e.g., 100-year or 500-year floodplain boundaries and floodways). The 100-year floodplains shown on the Flood Insurance Rate Maps are classified as “A-zones” and are further divided to identify different values based on local conditions.

At the local government level, the FEMA maps are used to determine flood insurance eligibility and to identify areas of restricted development opportunity. In most cases, FEMA maps and the county parcel maps are not at the same scale, and they are also of varying, generally unquantifiable, levels of accuracy making analysis of 100-year floodplain boundaries and parcel boundaries a questionable activity. PlanGraphics recommends that floodplain boundaries from the FEMA maps be re-plotted onto the orthophoto base maps in the Local Tier. The cartographic accuracy of the floodplain boundaries could be further refined by interpolating the base flood elevations within the contours prepared for this tier. Such an effort of re-compiling the floodplain boundaries should be coordinated with the appropriate FEMA representative in order that the new maps can be adopted by the local jurisdiction and by FEMA as the new official floodplain maps.

For general planning purposes, the 100-year floodplain boundaries may be included in the Sub-Regional Tier. PlanGraphics recommends using data prepared for the Local Tier in developing the Sub-Regional Tier. Local tier boundaries would be shown in a more generalized manner appropriate for Sub-Regional Tier presentation scales.

PlanGraphics is not aware of a need to prepare a separate database of 100-year floodplain boundaries as a separate database for the Statewide Tier. The Regional Tier, however, would be suitable for portraying, as needed, the location of potential flooding problems for general emergency response planning.

Well Location

Well features include the locations of water, gas, and oil wells. Smaller scale presentation of fossil fuel well locations may show the limits of the oil or gas fields instead of individual sites.

Individual wells may be included in the Local and Sub-Regional Tier databases. At the orthophoto scale recommended for the Sub-Regional Tier, however, only large individual wells will be visible and where wells are close together, some cartographic adjustment may be necessary for clear portrayal. In the Sub-Regional Tier, individual wells are included mainly for inventory and permit tracking applications. PlanGraphics recommends that existing water and oil and gas well maps be used as the source for plotting their locations at the Sub-Regional Tier.

The Regional Tier can be most effectively used as a geographic index to the location of oil well and gas well fields. The oil and gas well field features at this level are simply boundaries showing physical groupings of well locations. These boundaries are not intended to shown the underground reservoir of oil and gas reserves. PlanGraphics recommends that boundaries of the oil and gas well fields be used for this tier. Only the largest oil well and gas well fields would be shown on a map of statewide coverage. PlanGraphics does not recommend that a separate well location database be prepared for the Statewide Tier.

Environmental Features

This is a rather broad category that includes areas of environmental significance encompassing features that identify sensitive natural areas and areas of toxic human activity. Some examples of area features for this database include wildlife habitats, research natural areas, coal mine subsidence, abandoned mine lands, and threatened and endangered species. Some of this data may be restricted from general distribution due to natural or economic sensitivity.

The Sub-Regional Tier will likely include the most environmental databases of any tier since the most detailed existing mapping is at a scale of 1:24,000. Most of the mapping data would be derived from existing hard copy maps, many of which use USGS 7.5 minute quadrangles as the base. Minor positional adjustments may be required to plot the location of features that can be seen on orthophotos, if the orthophoto base map is selected.

The Regional Tier may be prepared by aggregating features included in the Sub-Regional Tier. PlanGraphics does not recommend preparing a separate database of environmental features for use with the Statewide Tier.

Soils

Mapped by the USDA Soil Conservation Service, soil maps show classification boundaries and include a code to indicate the associated soil type. Soil mapping is available from the Soil Conservation Service at three scales: 1:24,000; 1:250,000; and 1:7,500,000. Mapping at the 1:24,000 scale has been completed for a portion of West Virginia, and work is in progress. The 1:250,000 mapping is complete and is available in digital form from the SCS.

PlanGraphics recommends that soil classification mapping for the Sub-Regional Tier be acquired as SCS completes their work. The 1:24,000 scale soil mapping is ongoing as part of the Soil Survey Geographic Database (SSURGO). SSURGO data will be correlated to USGS quadrangles and will fit within the intended scale and accuracy of the Sub-Regional Tier. Three counties in West Virginia are complete at the 1:24,000 scale.

Statewide soil association mapping has been prepared at a scale of 1:250,000 by SCS and is referred to as the State Soil Geographic Data Base (STATSGO). PlanGraphics believes that this data would be appropriate to include in the Regional Tier.

Soil data is also available for the entire state at a scale of 1:7,500,000 in the National Soil Geographic Data Base (NATSGO). This data is rather coarse for much detailed analysis, but may be helpful to some users that require only a general overview.

Incidents/Point Features

This is another rather broad category of map features of various types. Incidents and point features may include hazardous waste locations, pollutant point sources, water quality sampling points, traffic accident locations, or permit sites. Their commonality is that they are generally shown as point features on a map, such as a pin map.

Some data maintained as polygons in the Local Tier may be shown as point features in the Sub-Regional, Regional, and Statewide Tiers. Determining when to drop boundary features and represent the location with a point will vary depending on two factors: 1) the relative boundary size, and 2) use of the data. For instance, city parks will certainly be shown as polygons in the Local Tier, and perhaps at the Sub-Regional Tier if they cover a city block. But, these same features become point features at the presentation scales associated with the Regional Tier.

Based on the activities identified in the previous report, PlanGraphics recommends that Sub-Regional Tier point features include pollution discharge points, traffic accident locations, water quality sampling points, and hazardous waste locations. There is a large potential for including many other features as point data in this tier.

Historical/Archaeological

PlanGraphics recommends that designated historic, cultural, or archaeological sites be included in the GIS. These data may be represented at the local, and perhaps the Sub-Regional, Tier as site boundaries.

Historic, cultural, and archaeological sites are also recommended for inclusion in the Sub-Regional Tier. At Sub-Regional Tier mapping scales, many of these features will be shown as points.

Recreational Facilities

These facilities typically include parks or recreational sites, including public and private parks, playing fields, golf courses, trails, and camping areas.

Sub-Regional tier representation of recreational facilities will vary depending on whether the site is a small picnic area or a park of some size. At the scale of 1:24,000, features such as a single football field are about the smallest feature that can be shown to scale. Most recreational features in the Sub-Regional Tier will be shown by their boundaries. As for the Local Tier, these boundaries will generally coincide with the property lines.

Biological Distributions

The locations and distributions of species, both animal and plant, are referred to as biological distribution features.

PlanGraphics does not recommend creating databases specifically for biological distribution at the Local or Statewide Tiers. An aggregation of distribution boundaries from the Sub-Regional Tier may be included in the Regional Tier, however.

Demography

In terms of map features, demography describes the physical boundaries of statistical units used to measure characteristics of the population, including population density, income, or age. Demographic map features generally include census tracts and blocks, or may be county boundaries.

PlanGraphics recommends that the census tract, block group, and block boundaries be derived from the TIGER Line files. These boundaries will have to be adjusted to fit the Local Tier base mapping; the TIGER Line files are based on USGS 1:100,000 scale mapping, and census boundaries will not fit the larger scale, more accurate Local Tier base maps without adjustment.

The Local Tier demographic database can include the boundaries of the smallest statistical unit - the census block. PlanGraphics recommends that the Local Tier databases include all statistical unit boundaries. The Sub-Regional Tier, however, is of a scale that cannot effectively portray the census block and block group. PlanGraphics recommends that for the Sub-Regional Tier census tracts be the smallest statistical unit boundary included in the database. Demographic reporting at the Regional and Statewide Tiers would be limited to county boundaries (see Political Boundaries) due to map scales.

Meteorological Information

The location of weather phenomena and average conditions are shown on maps, including temperature contours (isotherms), rainfall contours (isohyets), and other weather information. PlanGraphics does not recommend creation of a meterological database for the Local Tier.

PlanGraphics recommends using precipitation data from maps prepared by the USDA Soil Conservation Service as the source for input to the Sub-Regional or Regional Tier database. Meteorological data prepared for the Regional Tier can be used in aggregated form for inclusion in the Statewide Tier. We do not recommend preparing a database of meterological data for the Local Tier.

Dam Locations

PlanGraphics recommends that a database of dams be included as part of the Sub-Regional Tier. Local governments may also need for a separate database of dams which would be included in the Local Tier. We do not recommend creation and maintenance of separate databases of dam features and attributes for the Regional and Statewide Tiers.

Dam features in the Sub-Regional Tier will be shown as line symbols. Most dams included in this tier will be shown on the USGS 7.5 minute quadrangle.

PlanGraphics recommends that the USGS 7.5 minute quadrangle be used as the source for dam features. The quad map can be used to identify the feature location, and the orthophoto base map can be used as the source from which to trace the feature.

2.4 DATABASE SOURCES, CONVERSION METHODS, AND COST ISSUES

PlanGraphics presents our recommendations for data sources, conversion approaches, and our opinion of probable cost for each database layer in Tables 2-10 through 2-13. Each table covers a separate database tier:

- Table 2-10: Local Tier
- Table 2-11: Sub-Regional Tier
- Table 2-12: Regional Tier
- Table 2-13: Statewide Tier.

Background information regarding sources, methods, and cost issues are presented in the text following the tables.

2.4.1 Database Sources

For those data that are to become part of a comprehensive GIS, there will be a variety of sources. The primary sources of data for the comprehensive GIS include existing GISs, new aerial photography and field reconnaissance, and existing hard copy maps.

Perhaps the most expedient source of data is from existing GISs. Where users determine that data currently residing on one of the existing GISs should be placed on a common GIS, and is of sufficient content, accuracy, and currency, this data can be re-formatted and translated to the new system. Mostly, where needed, mapping and attribute databases from existing GISs will be moved totally to the new system. There may be some isolated cases where certain data may be maintained on both the existing GIS and on a new GIS. In general, this approach is not recommended. However, on occasion, applications may have been built on the existing GIS that either are too costly to replicate on a new system, or the new system does not support development of the applications. And still, other users need access to the data in a manner that supports both quick access and the ability to carry out new applications with the same data. Where both of these situations exist, it may be cost-effective to maintain redundant databases. Special update procedures must be developed to ensure that both copies of the database are kept current in accordance with the user standards for that data.

Different hardware and software platforms have been used to create the existing automated data sets by various participants. To share data between different systems, data must be translated. Many programs are available to perform data translations. In fact, many GIS software packages can be acquired with programs that translate data to and from several popular formats that are widely used in the mapping industry. These popular formats have been used as intermediate exchange formats for moving data between systems. Some of these include Intergraph's SIF, ESRI' ARCExport, MOSS, ERDAS' DIG, USGS' DLG, Census Bureau DIME and TIGER, and AutoCAD's DXF. Some vendors have developed custom exchange programs that bypass an intermediate format and translate data directly between two specific data formats (e.g., AutoCAD to Intergraph). Programs are now available that can translate map data now being used by USGS to formats used by most of the popular commercial and public domain GIS software.

New data collection efforts such as aerial photography and field reconnaissance are additional sources for GIS data. In particular, new aerial photography will be needed to support photogrammetric base mapping. Aerial photography is a very cost-effective data source for physical features. Large areas can be viewed and analyzed quickly in an office setting. Beyond photogrammetric base mapping, aerial photography can be used as a source to augment soils classification, to analyze land use and land cover, and for many other features. Field reconnaissance, long a tedious job combining on-site technical analysis and surveying, is being revolutionized with the advent of low-cost GPS receivers. GPS data collectors can be programmed by a user to accept information about the feature, species, or event being monitored in the field - and its coordinates. The location of threatened and endangered species habitats, timber stands, abandoned mines, water diversion structures, state-owned land boundaries, pavement quality, and many other features can be collected by a single person with relative ease. The data collected in the field can be "downloaded" to the GIS to create or update a particular layer.

Existing hard copy maps will be a major source of data for the GIS. Existing maps perhaps are the most troublesome source - some are not up to date; the accuracy may not be known. When creating a database from "scratch," the user knows how it was prepared. Often, existing maps were originally created prior to the recollection of current staff, and, without documentation on the process. However, all of these problems considered, using existing maps is generally better than starting from scratch. Some work will be required to determine accuracy and currency of the data. Additional work may be required to "scrub" the map of inaccuracies and to update it with a review of its source materials. In most cases, the analysis and scrubbing work will be worth the effort rather than trying to create the mapping anew.

An important part of developing the GIS database will be the use of digital data already existing as sources in digital map or tabular form. Many of these sources will be important for building the core and custodial GIS databases, while others will remain as external data sources accessed for specific GIS analysis. Data from existing GISs, as well as the many dBase and CAD files, may be incorporated directly into a common statewide GIS, either to replicate the functions on the common GIS, or as periodic input for file updating. The prerequisite issue for incorporating data from a single, separate database to the GIS is whether the effort to replicate existing application programs using the GIS software is offset by the benefit of increased access and potential to conduct analyses more efficiently.

2.4.2 Conversion Methods

Graphic information is entered into a GIS database using a variety of techniques that are selected depending on format, content, and condition of source data; capabilities of the

GIS software; user requirements for data accuracy; and funding availability. The most common methods include trace digitizing from a hard copy map, direct input using aerial photography sources through photogrammetric stereoplotters, coordinate geometry entry from survey descriptions, or optical scanning.

Source materials are available at a specific scale and level of accuracy. The graphic data stored in a GIS will maintain the same level of accuracy but may be plotted at nearly any scale. A scale is assigned only when a map is displayed or plotted. Most GISs have flexible routines for selecting any scale for map plotting. However, while a map can easily be enlarged, its accuracy and resolution are no greater than that of the original source data. This is an important consideration, since users will invariably desire to plot maps at scales that exceed the base resolution. For example, if an infrastructure map is converted from source maps at a base map scale of 1:2,400 and then replotted at 1:600, the resolution (or relative accuracy) of the features remains as they were on the 1:2,400 source document. Conversely, themes derived from 1:600 source maps retain their projected resolution when reproduced at 1:2,400.

Users should also be mindful that data translated between various systems can be subject to degradation of its inherent accuracy.

As a general practice, PlanGraphics recommends that most GIS database development be performed by private contractors specializing in conversion of particular types of data (e.g., photogrammetric, parcel, utility). Our recommendation is based on the fact that most users are already quite busy, and not set up for digital cartographic production. User agencies will be responsible for preparation of source materials, and for quality control review of data delivered by the contractor(s). Some data layers may in fact be created “in-house” if the amount of data to be converted is not large.

A very important issue to consider as part of GIS database development concerns the timing of GIS software selection and database conversion. A myth that has been circulating for some years states that conversion may begin without knowledge of which GIS software will ultimately be used, since data translators exist for nearly all systems. One element of truth in this myth is that many inter-system data translators do exist. However, it is critical that the “target” GIS software be known prior to beginning conversion. Data structures between systems are widely disparate; these differences stem from various schemes used to establish topology within the graphic database by different vendors. The data structure used in the selected GIS software must be understood and used by the conversion contractor in order that data translated from “data capture” software is compatible with the target system.

PlanGraphics has prepared recommendations for the approach to data conversion for each of the databases listed in Tables 2-10 through 2-13.

2.4.3 Database Cost Issues

Costs are presented in ranges that PlanGraphics has found to be typical for similar types of mapping and digital data conversion projects. The costs assume a need for a contracted data conversion project, except, where PlanGraphics recommends acquiring existing databases from federal agencies. Some of the recommended databases may have been converted to digital form and are in use by local, state, or federal agencies, universities, and private firms.

No costs have been estimated for those layers that are identified as candidates for in-house conversion or from aggregation from another tier. In many cases, in-house conversion is the most appropriate approach, and staff time will need to be allocated for this work. One option for the state is to use the West Virginia Geological and Economic Survey as a contractor for projects where their staff have appropriate technical background. In addition, the state may pursue hiring temporary or university intern staff for map conversion projects. With the recommendation to establish a Technical Support Center at West Virginia University (see Section 4), it is assumed that this group will play a major role in planning and carrying out database development.

There are a number of items, in the above mentioned tables, for which an estimation of costs is unfeasible under the focus of this study. Local Tier data, whose users are largely beyond the scope of this study, has the least cost information (for example, the number of sewer line segments within each municipality and county). In a number of cases, the background information and analysis were outside the scope of this study. Where applicable, we have indicated so in the cost tables. Much of the data for the Regional and Statewide Tiers may be derived by aggregating data from other tiers or in-house digitizing. Costs associated with programmatic aggregation and in-house efforts are also not estimated.

Costs are presented in terms of their unit of measure. For example, the range of probable costs for orthophoto base mapping in the Local Tier is per square mile. PlanGraphics' recommendations for the Local Tier orthophoto base mapping identified three scales; the variation in map scale has a significant impact on the price per square mile. Hence, we have further indicated, under the Unit of Measure for this item, these special cost considerations.

The following costs are summarized in Tables 2-10 through 2-13 for conversion of the Sub-regional, Regional, and State Tiers. Cooperative Cost-sharing with the federal present opportunities to reduce state outlays for this database development. These opportunities will be discussed in the *GIS Development Plan.*

Sub-regional	$1,500,000	to	$2,700,000
Regional	$100,000	to	$252,000
Statewide	$0		

2.5 GIS DATA AND EXTERNAL SYSTEMS

2.5.1 Existing Tabular Databases

There are two general approaches to making use of existing tabular data with the GIS. One approach is to incorporate the tabular database into the GIS as the attribute data set to a particular set of map features. Another approach is to establish common attribute data items in the GIS and external files in order to programatically relate other attribute data. We explain both of these approaches more thoroughly below.

Strictly from the point of view of GIS data processing, the best approach for using existing tabular data that are related to map features is to translate, format, and load that data into the GIS as part of the database. This will eliminate the need to link the GIS to other data processing systems to access and process attribute data as part of a particular GIS application. Generally speaking, this approach is less cumbersome and easier to accomplish. However, in many cases, these tabular databases were created to support existing applications for which software was purchased or developed. If these databases are made part of the GIS, either redundant data will be maintained (which is not a good objective) or new software programs must be written and programmed to carry out existing applications (which are costly and may not work as well as the original program). A review of the current non-GIS application programs that are used to process existing tabular files should be made to determine whether or not it is realistic and cost-effective to incorporate them into the GIS.

The other approach of linking the GIS to other systems may result in somewhat more complex requirements for data communications and programming, however, it may be less disruptive to the overall operation. If existing databases are left in place to support ongoing non-GIS applications, these files may be accessed for use in a GIS application program. A common attribute data element is necessary to link a non-GIS record to a particular feature within the GIS.

For example, in the GIS, a parcel feature carries with it a parcel identification number in its attribute database. The non-GIS parcel ownership file maintained by the assessor contains the same parcel identification number. This common data item can be used to identify a particular non-GIS parcel record to extract information that is not maintained in the GIS. This approach requires a limited set of redundant data items between the GIS and other files, but eliminates the need for large duplicative databases.

2.5.2 Document Management

Another potential use of a GIS is to assist in the management and retrieval of non-GIS files such as plans, drawings, photographs, log books, or field sketches. In many cases, there is no need to include details of, for example, park facilities in the GIS database, but, GIS users need to access the detailed data. These records can be maintained in a separate document management system and linked to the GIS.

Many organizations are becoming involved with document imaging systems to manage their various records. Where these records relate to land, such as subdivision plats, timber cutting permits, or survey record books, the GIS can be used as a geographic index. These applications will also require that the GIS and the document management system carry the same logical record identifier in order to link the systems.

With proper planning, it may well be possible to use one workstation or terminal to access both the GIS and document management system. A system user that knows the area of interest and the type of records needed could access, for instance, the image of all water well permits within a particular sub-region. In many cases, users don't know the specific document file number and must search in indices for some geographic reference. Linkage of data and software between the GIS and document management system could be used to significantly reduce the time required to retrieve these records.

2.6 DATA STANDARDS

Perhaps the most important issue in regard to the GIS database is that of data standards. GIS users must have confidence in the data used for analyses and decision-making, or the resource and investment will diminish in value, or perhaps be lost. When developing procedures for ongoing database maintenance, proposed methodologies must ensure adherence to mutually compatible data standards.

Below, PlanGraphics describes several topic areas relating to data standards. The topics include:

- User design standards
- Meta-data standards
- Data maintenance and updating standards.

2.6.1 User Design Standards

There are a host of standards issues that influence the design and implementation of geographic information systems. In any automation project, the creation of a database should be preceded by a design phase that describes the content and format of the database and products generated from the system. These design decisions impact attribute data schemas and coding rules; map accuracy criteria; quality control; and map design criteria, all of which are largely independent of a particular vendor's software.

Schemas for storing GIS attribute data should adhere to existing standards within an organization as a whole and, if applicable, in a broader community of users. The schema specifies field lengths, data element formats (e.g., integer, character, etc.), and other characteristics of the database. One example of GIS data for which standard schemas are appropriate is site address information. Standard address formats define street name, address number, street type, prefix, and suffix as individual data fields with specified field lengths and coding scheme fields. Use of a consistent schema can greatly reduce the problems encountered in exchanging attribute data between systems. Users should take sufficient time to review existing schemas and comply with existing standards.

Alpha or numeric coding and classification schemes for data are used for many types of data that are stored on computer systems. Classifying data into parent and subclasses and coding these classes appropriately makes it easy to retrieve and analyze data at different levels of detail. For example, land use or land cover classification schemes normally use some hierarchical classes each of which have alpha or numeric codes assigned (e.g., code of “100” for urban with a subclass code of “110” for residential). A large amount of the data in GISs is stored in coded form, and it is easy to see why consistent use of coding and classification can facilitate exchange of data among systems. Unfortunately, it is all too often the case that database design is done in a vacuum, and standard coding schemes are not followed. Users should adhere to accepted coding and classification standards within their own organization as well as standards established by government and professional groups, at national and international levels.

Selecting suitable accuracy levels for the compilation and automation of map layers in GIS is a perplexing issue for most users. The objectives of map accuracy standards are to explicitly define and measure levels of accuracy, not to dictate what accuracy is appropriate for particular users. Accuracy standards address both horizontal and vertical placement of map features and describe maximum errors of displacement relative to their actual position. Since 1947, the National Map Accuracy Standards originally published by the U.S. Bureau of the Budget has been the accepted standard for defining the accuracy of photogrammetrically prepared base maps. This standard has traditionally been used in the preparation of base maps for GISs. Recently, a new accuracy standard for large-scale line maps was developed and proposed by a Committee of the American Society of Photogrammetry and Remote Sensing (ASPRS) which addresses some limitations of the National Map Accuracy Standards and is more suited to automated mapping. While this standard has not yet been officially approved by a federal government agency, it is likely that it will become a standard in the near future.

With the diverse sources from which GIS databases are built, it is extremely important to maintain information about the integrity and quality of data in the GIS. These factors determine how the data can be appropriately used in specific GIS applications. “Integrity” and “quality” encompass many characteristics of a GIS database such as map accuracy, completeness, and data currency. Several approaches may be used to track this information, including the creation of data dictionaries that maintain information about data “layers” or the inclusion of codes assigned as attributes for individual map features to characterize and track quality and integrity. This “data about the data” is referred to as meta-data. No standards have been proposed to guide the rating of GIS database quality and how quality indicators should be maintained. PlanGraphics recommends that the state government GIS community begin to address this concern, because it is so vital to flexible exchange of data. It also has great implications on application development through the use of newer technologies such as object-oriented database architectures, expert systems, and artificial intelligence.

Data standards should be based on current and anticipated user requirements. PlanGraphics' presented our initial recommendations for data accuracy and update timing are presented in this report section. Additional work will be required on the part of the various project participants to further refine our basic recommendations.

PlanGraphics recommends that a standing task force be established to coordinate development and enforcement of data standards.

The process of determining standards may require some negotiation when non-custodian users have more stringent needs than the agency responsible for a particular data type.

2.6.2 Meta-Data Standards

Meta-data is data about the data. Meta-data is used to inform the user about particular characteristics of a database. Meta-databases may hold information about the content; source and lineage; geographic area of coverage; quality and accuracy; custodian responsible for maintenance; and other important characteristics of the GIS database. The meta-databse can be considered an automated data dictionary from which users, or potential users, can review the characteristics of data and make decisions on how it may be used. At the federal government level, the Federal Geographic Data Committee is developing a standard meta-data format that West Virginia may wish to consider as a prototype for their own format.

One of the most important uses of meta-data is for users to determine limitations on the use of information. In addition to the separate meta-database, PlanGraphics recommends that each map and geodata database contain certain elements of meta-data to indicate its characteristics in order to expedite the process of determining their usefulness for a particular application. Data that is particularly useful for one application may provide inappropriate results when used for another application.

2.6.3 Data Maintenance

After initial creation or digital conversion of map features and their associated attributes, the databases must be continuously updated and maintained through operational activities. In most cases, the agency responsible for creating each database component will also be responsible for maintenance of that component. However, the effort and resources required to perform this indispensable task often require coordination between more than one agency. For example, updates to the survey control layer will be provided by local and state agencies involved in field surveying operations. The prompt update of the land cover/land use layer will be very important to several state government agencies.

The benefits of a cooperative, coordinated GIS effort for each agency places the burden of prompt and accurate database maintenance on custodial agencies. A coordinated effort also means that some agencies may have to maintain their databases more frequently and accurately than is required for their internal needs, so that their databases remain valuable to other participating agencies. Nothing will extinguish enthusiasm and funding for a coordinated GIS effort more readily than a lack of confidence in the accuracy and currency of another agency's database, and the feeling that another agency is not carrying its share of the load.

PlanGraphics recommends that a series of formal procedures be developed and adopted for database updating. These procedures will identify the agency responsible for updating, how often the data are updated, data sources, and the technical update process. These procedures should also include a provision on how to notify the "custodian" agency of errors or inconsistencies in their database to establish a quality control loop between data users and creators. In defining maintenance responsibilities, statutory responsibilities of certain agencies for data maintenance should be taken into account.

For the purpose of organizing a development and maintenance strategy, these layers can be grouped into two categories: a) **Common** GIS layers which include those that will be used routinely by agencies outside the custodian agency, and b) **Agency-specific** layers which will not routinely be shared and are the specific responsibility of a particular agency. We have listed the recommended basic data layers below and have identified each layer with a "C" for common or an "S" for single agency use. The division between what is labeled as C or S is based on anticipated general usage of an entire layer, not individual data items that may reside within the layer. Our groupings are also not intended to be restrictive, but simply to indicate our vision of the general trend of data use.

- Survey Control Points and Grid (C)
- Orthophoto/Planimetric (C)
- Transportation Network (C)
- Hydrography (C, S)
- Topography (C)
- Political Boundaries (C)
- Administrative Boundaries (C)
- Parcels (S)
- Energy Transmission (S)
- Utility Distribution/Collection (S)
- Land Use/Land Cover (C)
- Bedrock Geology (C)
- Floodplains (C)
- Well Locations (C, S)
- Environmental Features (C)
- Soil Classification Boundaries (C)
- Incidents/Point Features (S)
- Historical/Archeological Sites (S)
- Recreational Facilities (S)
- Demography (S)
- Meteorological Data (S)
- Biological Distribution (S).

This reference to “common” vs. “single agency” use refers to the core set of map features and attributes on the particular layer. It will be normal practice for individual agencies to build additional tabular attribute sets and enhance existing GIS layers with additional features to meet their needs.

The main purpose of this designation of common or agency-specific is to establish a basis to assign roles for custodianship. Custodianship implies a responsibility for update and for maintaining meta-data to facilitate access to GIS layers. It is assumed that all common layers will be available for access by all GIS participants.

TABLE 2-1
DATABASE TIERS

Database Tier Name	Geographical Extent to which Tier is Typically Applied	Base Map and Recommended Scale	Map Layers Typically Associated with Tier
Local	Small: Parcel or site-specific, up to an entire county. Each base map file typically covers an area ranging from 1/4 to 4 square miles.	Possibly photogrammetric base map (e.g., orthophoto). Accurate parcel-level maps compiled from plats and survey data. Depending on density of development, scales should range from 1" = 100' to 1" = 400'.	Survey control Parcel Utility infrastructure Zoning Voting precinct
Sub-regional	Medium: Entire county to multiple counties. A base map file may cover an area from 25 to 50 square miles.	USGS topographic quad sheets or reduced local base maps. Recommended scale is 1:24,000.	Survey control, planimetric, district boundaries, energy transmission, land cover, environmental, large tracts of land (parcels), and others which require a reasonably accurate small-scale base map to create and maintain.
Regional	Large: Large regions of the state. A base map file may cover 1,000 to 3,000 square miles.	1:100,000 scale USGS topographic map files.	Same layers as the regional database; tabular data removed and aggregated into larger files. Some layers may be generalized graphically.
Statewide	Very Large: Whole state is shown on a single map. Data portrayed at the statewide tier may be at a scale for wall display or perhaps for inclusion in a bound report.	An overview map that may show county boundaries, primary highways, primary rivers, and a few features for general location reference. The 1:500,000 scale USGS state map is recommended for the statewide base map.	Primarily used for the portrayal of generalized information. Few thematic databases would be created for ongoing maintenance in this tier.

TABLE 2-6
LOCAL TIER DATA RECOMMENDATIONS

Layer	Data Type	Map Features	Primary Attributes
Survey Control	Point	Horizontal Geodetic Control Monuments	ID number Geographic coords. State Plane coords. Responsible agency Order of accuracy
	Point	Vertical Geodetic Control Monuments	ID number Elevation State Plane coords. Responsible agency Order of accuracy
	Line	State Plane Grid	Grid values
Orthophoto/ Planimetric	Raster image, Point, Line, Polygon	Orthophoto image, Building outlines, Retaining walls, Substations, Other physical features	Map ID
Transportation	Line	Road Centerline	Road name Highway number Functional classification
	Line	Railroad	Railroad name
	Polygon	Airport	Airport name
	Line	Waterways	Waterway name
	Line	Scenic trails	Trail name
	Line	Public transit routes	Route number
Surface Hydrography	Line Line	River Stream	River name Stream name
	Line	Lock and dams	Lock and dam number
	Polygon	Lake	Lake name
	Polygon	Pond	Pond name
	Point	Water intake structure	Structure ID Owner's name Permit number
	Polygon	Reservoir	Reservoir name
	Point	River mile marker	Marker mileage

TABLE 2-6 (continued)
LOCAL TIER DATA RECOMMENDATIONS

Layer	Data Type	Map Features	Primary Attributes
Topography	Line	Intermediate contour	Elevation value
	Line	Intermediate depression contour	Elevation value
	Line	Index contour	Elevation value
	Line	Index depression contour	Elevation value
	Point	Spot elevation	Elevation value
	Point	DEM mass point	Elevation value
	Line	DEM breakline	Elevation value
Political Boundaries	Polygon	State boundary	State name
	Polygon	County boundary	County name
	Polygon	Municipal boundary	Municipality name
Administrative Boundaries	Polygon	Senatorial districts	District ID
	Polygon	Congressional districts	District ID
	Polygon	Voting districts	District ID
	Polygon	Voting precincts	Precinct ID
	Polygon	Census tract	Tract ID
	Polygon	Census block group	Block group ID
	Polygon	Census block	Block ID
	Polygon	Highway maintenance districts	District number
	Polygon	Utility rate areas	Utility name Rate area ID
	Polygon	Appraisal districts	District ID
	Polygon	Soil conservation districts	District name
	Polygon	Regional planning council areas	Council name
Parcels	Polygon	Parcel boundary	Parcel ID Address Deed book and page Owner's name
	Polygon	Highway ROW	Parcel ID Deed book and page Owner's name

TABLE 2-6 (continued)
LOCAL TIER DATA RECOMMENDATIONS

Layer	Data Type	Map Features	Primary Attributes
Parcels (continued)	Polygon	Railway ROW	Parcel ID Deed book and page Owner's name
	Polygon	Utility easement	Parcel ID(s) Deed book and page User's name
Energy Transmission	*Not Recommended for this Tier*	*Features to be Included in Utility Distribution/Collection*	
Utility Distribution/ Collection	Line/Network	Overhead electric distribution lines	Owner's name Wire segment ID KVA value
	Line/Network	Underground electric distribution lines	Owner's name Wire segment ID KVA value
	Point	Electric poles and control facilities	Pole ID #s, type, normal status
	Polygon	Substation	Owner's name Address
	Line/Network	Sanitary sewer lines	Sewer segment ID
	Point	Sanitary sewer manhole	Manhole ID
	Polygon	Treatment plant	Plant name
	Point	Pump; misc. equipment	Pump name or ID
	Line/Network	Storm sewer pipeline	Sewer segment ID
	Point	Storm sewer manhole	Manhole ID
	Line/Network	Storm ditch	Ditch name or ID
	Line/Network	Water distribution pipeline	Pipeline segment ID
Land Use/Land Cover	Polygon Polygon	Land use category boundary Land cover category boundary	Land use category Land cover category
	Polygon	Zoning boundary	Zoning classification Case number
Subsurface Hydrogeology	*Not Recommended for this Tier*	*See Sub-Regional Tier*	
Bedrock Geology	*Not Recommended for this Tier*	*See Sub-Regional Tier*	

TABLE 2-6 (continued)
LOCAL TIER DATA RECOMMENDATIONS

Layer	Data Type	Map Features	Primary Attributes
Floodplains	Polygon	100-year flood zone category	Flood zone category
	Polygon	500-year flood zone category	Flood zone category
	Polygon	Floodway	
Well Locations	Point	Water well	Well ID Permit number Owner's name
	Point	Oil well	Well ID Permit number Owner's name
	Point	Gas well	Well ID Permit number Owner's name
Environmental Features	*Not Recommended for this Tier*	*See Sub-Regional Tier*	
Soil Series	*Not Recommended for this Tier*	*See Sub-Regional Tier*	
Incidents/Point Features	Point	Hazardous waste locations	Permit ID Owner's name Address Type of waste
	Point	Water quality sampling points	Sampling point ID
	Point	Pollutant point sources	Type of pollutant Address
	Point	Traffic accident locations	Highway number
Historical Archeological Sites	Polygon	Historic site boundary	Name of site Address Owner's name
	Polygon	Archeological site boundary	Name of site Address Owner's name
	Polygon	Cultural site boundary	Name of site Address Owner's name

TABLE 2-6 (continued)
LOCAL TIER DATA RECOMMENDATIONS

Layer	Data Type	Map Features	Primary Attributes
Recreational Facilities	Polygon	Park boundary	Park name Owner's name Address
	Polygon	Golf courses	Course name Owner's name Address
	Polygon	Playing field	Field name Type of field Owner's name Address
	Line	Trails	Trail name Owner's name Type of trail
	Polygon	Camping areas	Campground name Address Owner's name Facilities
Demography	Polygon	County boundary	County name
	Polygon	Municipal boundary	Municipality name
	Polygon	Census tract	Tract ID
	Polygon	Census block group	Block group ID
	Polygon	Census block	Block ID
Meterological Data	*Not Recommened for this Tier*		
Biological Distribution	*Not Recommened for this Tier*		
Dam Locations	Line	Dam feature	Dam ID Owner's name Permit number Last inspected date

TABLE 2-7
SUB-REGIONAL TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Survey Control	Point	Horizontal Geodetic Control Monuments	ID number Geographic coords. State Plane coords. Responsible agency Order of accuracy
	Point	Vertical Geodetic Control Monuments	ID number Elevation State Plane coords. Responsible agency Order of accuracy
	Line	State Plane Grid	Grid values
Orthophoto/ Planimetric	Raster image	Orthophoto	Map ID
Transportation	Line	Road Centerline	Road name Highway number Functional classification
	Line	Railroad	Railroad name
	Polygon	Airport	Airport name
	Line	Waterways	Waterway name
	Line	Scenic trails	Trail name
	Line	Public transit routes	Route number
Surface Hydrography	Line	River	River name
	Line	Stream	Stream name
	Line	Lock and dams	Lock and dam number
	Polygon	Lake	Lake name
	Polygon	Pond	Pond name
	Point	Water intake structure	Structure ID Owner's name Permit number
	Polygon	Reservoir	Reservoir name
	Point	River mile marker	Marker mileage
Topography	Line	Intermediate contour	Elevation value
	Line	Intermediate depression contour	Elevation value
	Line	Index contour	Elevation value

TABLE 2-7 (continued)
SUB-REGIONAL TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Topography (continued)	Line	Index depression contour	Elevation value
	Point	Spot elevation	Elevation value
	Point	DEM mass point	Elevation value
	Line	DEM breakline	Elevation value
Political Boundaries	Polygon	State boundary	State name
	Polygon	County boundary	County name
	Polygon	Municipal boundary	Municipality name
Administrative Boundaries	Polygon	Senatorial districts	District ID
	Polygon	Congressional districts	District ID
	Polygon	Voting districts	District ID
	Polygon	Voting precincts	Precinct ID
	Polygon	Census tract	Tract ID
	Polygon	Census block group	Block group ID
	Polygon	Census block	Block ID
	Polygon	Highway maintenance districts	District number
	Polygon	Soil conservation districts	District name
	Polygon	Regional planning council areas	Council name
Parcels	Polygon	State-owned parcel boundary (large)	Parcel ID Address Deed book and page Agency's name
	Point	State-owned parcel (small)	Parcel ID Address Deed book and page Agency's name
Energy Transmission	Line	Electric transmission lines	Owner's name KVA value
	Polygon	Major substation	Owner's name Address
	Line	Gas transmission pipeline	Owner's name
	Polygon	Gas storage area	Owner's name

TABLE 2-7 (continued)
SUB-REGIONAL TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Energy Transmission (continued)	Line	Oil wellhead collectors	Owner's name
	Polygon	Oil collection storage area	Owner's name
Utility Distribution/ Collection	Polygon	Substation	Owner's name Address
	Polygon	Treatment plant	Plant name
	Line	Major storm ditch	Ditch name or ID
Land Use/Land Cover	Polygon	Land use category boundary	Land use category Land use permit
	Polygon	Land cover category boundary	Land cover category
Subsurface Hydrogeology	Polygon	Aquifer boundary	Aquifer name
Bedrock Geology	Polygon	Geologic rock units	Geologic rock unit name
	Point	Geologic point feature	Feature type, measured quantity
	Line	Fault line	Fault name
Floodplains	Polygon	100-year flood zone boundary	Flood prone stream name
Well Locations	Point	Water well	Well ID Permit number Owner's name
	Point	Oil well	Well ID Permit number Owner's name
	Point	Gas well	Well ID Permit number Owner's name
Environmental Features	Polygon	Significant wildlife habitat boundary	Wildlife species
Soil Classification Boundaries	Polygon	Soil classification boundaries	Soil classification Productivity category
Incidents/Point Features	Point	Hazardous waste locations	Permit ID Owner's name Address Type of waste
	Point	Water quality sampling points	Sampling point ID

TABLE 2-7 (continued)
SUB-REGIONAL TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Incidents/Point Features (continued)	Point	Pollutant point sources	Type of pollutant Address
	Point	Traffic accident locations	Highway number
Historical Archeological Sites	Polygon/Point	Historic site boundary (point if small)	Name of site Address Owner's name
	Polygon/Point	Archeological site boundary (point if small)	Name of site Address Owner's name
	Polygon/Point	Cultural site boundary (point if small)	Name of site Address Owner's name
Recreational Facilities	Polygon/Point	Park boundary (point if small)	Park name Owner's name Address
	Polygon/Point	Golf courses (point if small)	Course name Owner's name Address
	Polygon/Point	Playing field (point if small)	Field name Type of field Owner's name Address
	Line	Trails	Trail name Owner's name Type of trail
	Polygon/Point	Camping areas (point if small)	Campground name Address Owner's name Facilities
	Polygon	Sportsman access sites	Site name
	Polygon	Wildlife management areas	Area name
Demography	Polygon	County boundary	County name
	Polygon	Municipal boundary	Municipality name
	Polygon	Census tract	Tract ID
Meterological Data	Line	Temperature contours (isotherms)	Contour temperature value
	Line	Rainfall contours (isohyets)	Contour rainfall value

TABLE 2-7 (continued)
SUB-REGIONAL TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Biological Distribution	Polygon/Point	Animal species boundary (point if small)	Specie name
	Polygon/Point	Plant species boundary (point if small)	Specie name
Dam Locations	Line	Dam feature	Dam ID Owner's name Permit number Last inspected date

TABLE 2-8
REGIONAL TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Survey Control	Point	Primary Horizontal Geodetic Control Monuments	ID number Geographic coords. State Plane coords. Responsible agency Order of accuracy
	Point	Primary Vertical Geodetic Control Monuments	ID number Elevation State Plane coords. Responsible agency Order of accuracy
	Line	State Plane Grid	Grid values
Orthophoto/ Planimetric	*Not Recomended for this Tier*		
Transportation	Line	Road centerlines of 1st thru 3rd class highways	Highway number
	Line	Railroad	Railroad name
	Point	Airport	Airport name
	Line	Waterways	Waterway name
Surface Hydrography	Line	River	River name
	Line	Major stream	Stream name
	Polygon	Major lake	Lake name
	Polygon	Major reservoir	Reservoir name
Topography	Line	Intermediate contour	Elevation value
	Line	Intermediate depression contour	Elevation value
	Line	Index contour	Elevation value
	Line	Index depression contour	Elevation value
Political Boundaries	Polygon	State boundary	State name
	Polygon	County boundary	County name
	Polygon	Municipal boundary	Municipality name
Administrative Boundaries	Polygon	Senatorial districts	District ID
	Polygon	Congressional districts	District ID
	Polygon	Soil conservation districts	District name

TABLE 2-8 (continued)
REGIONAL TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Administrative Boundaries (continued)	Polygon	Regional planning council areas	Council name
Parcels	Polygon	Major state-owned parcel boundary	Owner agency
Energy Transmission	Line	Major electric transmission lines	Owner's name KVA value
	Line	Major gas transmission pipeline	Owner's name
Utility Distribution/ Collection	*Not Recomended for this Tier*		
Land Use/Land Cover	Polygon	Land use category boundary	Land use category
	Polygon	Land cover category boundary	Land cover category
Subsurface Hydrogeology	Polygon	Major aquifer boundary	Aquifer name
Bedrock Geology	Polygon	Geological rock unit	Rock unit name
	Line	Major fault line	Fault name
Floodplains	Polygon	Flood prone area boundary	
Well Locations	Polygon	Oil field boundary	Field name
	Polygon	Gas field boundary	Field name
Environmental Features	Polygon	Wildlife habitat boundary	Wildlife species Responsible agency
Soil Classification Boundaries	Polygon	Soil series	Soil series codes Soil characteristics
Incidents/Point Features	Point	Hazardous waste locations	Permit ID Owner's name Address Type of waste
	Point	Permit sites boundary (parcel or area)	Permit ID Permittee Address Type of permit
	Point	Water quality sampling points	Sampling point ID
	Point	Pollutant point sources	Type of pollutant Address
	Point	Traffic accident locations	Highway number

TABLE 2-8 (continued)
REGIONAL TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Historical/ Archeological Sites	Point	Historic site	Name of site and ID#
	Point	Archeological site	Name of site and ID#
	Point	Cultural site	Name of site and ID#
Recreational Facilities	Point	Park boundary	Park name Owner's name Address
	Point	Golf courses	Course name Owner's name
	Point	Major camping areas	Campground name Owner's name
	Polygon	Sportsman access sites	Site name
	Polygon	Wildlife management areas	Area name
Demography	Polygon	County boundary	County name
	Polygon	Municipal boundary	Municipality name
	Polygon	Census tract	Tract ID
	Polygon	Census block group	Block group ID
	Polygon	Census block	Block ID
Meterological Data	Line	Temperature contours (isotherms)	Contour temperature value
	Line	Rainfall contours (isohyets)	Contour rainfall value
Biological Distribution	Polygon	Major animal species boundary	Specie name
	Polygon	Major plant species boundary	Specie name
Dam Locations	Point	Major dam	Dam name

TABLE 2-9
STATEWIDE TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Survey Control	*Not Recomended for this Tier*		
Orthophoto/ Planimetric	*Not Recomended for this Tier*		
Transportation	Line	Road centerlines of major highways	Highway number
	Line	Railroad	Railroad name
	Point	Airport	Airport name
	Line	Waterways	Waterway name
Surface Hydrography	Line	River	River name
	Line	Major stream	Stream name
	Polygon	Major lake	Lake name
Topography	Line	Intermediate contour	Elevation value
	Line	Index contour	Elevation value
Political Boundaries	Polygon	State boundary	State name
	Polygon	County boundary	County name
Administrative Boundaries	Polygon	Senatorial districts	District ID
	Polygon	Congressional districts	District ID
	Polygon	Regional planning council areas	Council name
Parcels	*Not Recomended for this Tier*		
Energy Transmission	Line	Major electric transmission lines	Owner's name KVA value
	Line	Major gas transmission pipeline	Owner's name
Utility Distribution/ Collection	*Not Recomended for this Tier*		
Land Use/ Land Cover	*Not Recomended for this Tier*		
Subsurface Hydrogeology	*Not Recomended for this Tier*		
Bedrock Geology	Polygon	Major rock units	Major rock unit name

TABLE 2-9 (continued)
STATEWIDE TIER DATA REQUIREMENTS

Layer	Data Type	Map Features	Primary Attributes
Floodplains	*Not Recomended for this Tier*		
Well Locations	*Not Recomended for this Tier*		
Environmental Features	*Not Recomended for this Tier*		
Soil Classification Boundaries	*Not Recomended for this Tier*		
Incidents/Point Features	*Not Recomended for this Tier*		
Historical/ Archeological Sites	*Not Recomended for this Tier*		
Recreational Facilities	*Not Recomended for this Tier*		
Demography	*Not Recomended for this Tier*		
Meterological Data	*Not Recomended for this Tier*		
Biological Distribution	*Not Recomended for this Tier*		
Dam Locations	*Not Recomended for this Tier*		

TABLE 2-10
LOCAL TIER
DATABASE SOURCES, CONVERSION METHODS, AND PROBABLE COST

Layer	Data Source	Conversion Method	Range of Probable Cost[1]	Unit of Measure
Survey Control	NGS, WVDOT, and Local Survey Control Files	File Import and Translation; Coordinate entry	UNKNOWN	Per point - entry Per point - survey
Orthophoto Mapping	Aerial photography, Survey control	Photogrammetry - Scanning	$3,000 - $5,000 $1,000 - $1,500 $250 - $500	Per sq.mi. 1" = 100' Per sq.mi. 1" = 200' Per sq.mi. 1" = 400'
Transportation	Aerial photography, Orthophoto	Photogrammetric compilation, Trace digitizing	UNKNOWN	N/A
Surface Hydrography	Aerial photography, Orthophoto	Photogrammetric compilation, Trace digitizing	UNKNOWN	N/A
Topography	Aerial photography, Survey control	Photogrammetric compilation	$2,000-$3,000 $1,000-$1,500 $500-$750	Per sq.mi. 2.5 foot Per sq.mi. 5 foot Per sq.mi. 10 foot
Political Boundaries	Survey descriptions, Existing maps	Keyboard entry, Trace digitizing	UNKNOWN	Per polygon
Administrative Boundaries	Survey descriptions, Existing maps	Keyboard entry, Trace digitizing	UNKNOWN	Per polygon
Parcels	Survey descriptions, Existing maps	Keyboard entry, Trace digitizing	$12 - $25 $5 - $7	Per parcel - COGO Per parcel - Digitize
Energy Transmission	Orthophoto base maps	Trace digitizing	UNKNOWN	Per orthophoto
Utility Distribution/ Collection	Orthophotos, existing utility maps	Trace digitizing	UNKNOWN	Per feature/segment
Land Use/Land Cover	Existing land use/land cover mapping	Trace digitize	UNKNOWN	Per polygon
Subsurface Hydrogeology	*Not recommended for this Tier*	*See Sub-Regional Tier*		
Bedrock Geology	*Not recommended for this Tier*	*See Sub-Regional Tier*		
Floodplains	Orthophotos, topography, FEMA maps	Trace digitizing	UNKNOWN	Per map
Well Locations	Orthophotos, Existing well maps	Trace digitizing	UNKNOWN	Per well
Environmental Features	Orthophotos, Existing environmental maps	Trace digitizing	UNKNOWN	Per polygon
Soil Classification Boundaries	*Not recommended for thisTier*	*See Sub-Regional Tier*		
Incidents/Point Features	Orthophotos, Existing maps, Tabular files	Trace digitizing, address matching	UNKNOWN	N/A
Historical Archeological Sites	Orthophotos, Existing site maps	Trace digitizing	UNKNOWN	Per site
Recreational Facilities	Orthophotos, Existing facility maps	Trace digitizing	UNKNOWN	Per facility
Demography	Orthophotos, TIGER census maps	Trace digitizing, Automated adjustment	UNKNOWN	Per map
Meterological Data	*Not recommended for this Tier*	N/A	N/A	N/A
Biological Distribution	*Not recommended for this Tier*	N/A	N/A	N/A
Dam Locations	Orthophotos, Dam records	Trace digitizing	UNKNOWN	Per feature

[1]In general, no attempt has been made to project costs for Local Tier layers since these costs will vary considerably from location to location.

[2]Options for conversion of land parcels include coordinate geometry (COGO) entry form survey records and digitizing from existing maps. The specific approach depends on the condition of source materials, desired accuracy, and other considerations.

TABLE 2-11
SUB-REGIONAL TIER
DATABASE SOURCES, CONVERSION METHODS AND COST

Layer	Data Source	Conversion Method	Range of Probable Cost	Unit of Measure
Survey Control	NGS and WVDOT Files for Control Points	File Import and Translation	N/A	
Orthophoto/Planimetric Mapping	Aerial photography, Survey control	Photogrammetry; Optical Orthophoto Production, then Scanning	$700,000 to $1,000,000 Cooperative project with USGS and/or SCS possible to reduce costs	For Statewide Coverage of 7.5 Minute Orthophoto Quadrangles
Transportation	TIGER Line Files and Orthophoto Quads	Adjust Roads to Orthophoto Quads In-house	N/A	N/A
Surface Hydrography	TIGER Line Files and Orthophoto Quads	Adjust Hydrography to Orthophoto Quads In-house	N/A	N/A
Topography	Aerial Photography, Survey Control	Part of Orthophoto Photogrammetric Project	$1,000,000 to $1,500,000 if produced as part of orthophoto project	For Statewide Coverage
Political Boundaries	Local Tier	Aggregate from Local Tier	N/A	N/A
Administrative Boundaries	Local Tier, Existing Admin. Boundary Maps	File Import and Translation, In-house Trace Digitizing	N/A	N/A
Parcels	State-owned Parcel Files - Existing Plats	In-house Trace Digitizing from plats	N/A	N/A
Energy Transmission	USGS quads	In-house Trace Digitizing	N/A	N/A
Utility Distribution/ Collection	*Not recommended for this Tier*			
Land Use/Land Cover	Satellite Imagery and High Altitude Aerial Photography	Automated Image Classification and Manual Interpretation	*Undetermined Due to WideVariables*	
Subsurface Hydrogeology	Existing hydrogeological mapping	In-house Trace Digitizing	N/A	N/A
Bedrock Geology	Existing USGS geologic quadrangles at 1:24,000 scale	In-house Trace Digitizing	N/A	10% of State is Available at 1:24,000 Scale
Floodplains	Local Tier	In-house Aggregation from Local Tier	N/A	N/A
Well Locations	Permit Applications; Complaints; Aerial photography	In-house Data Input	N/A	N/A
Environmental Features	Heritage Database; Miscellaneous existing maps; Field reconnaissance	Import from Heritage Database; Board digitize; Input from data collector	N/A	N/A
Soil Classification Boundaries	SSURGO Database from SCS	Purchase from SCS when Files are Complete	$27,500 - When Available	Statewide Coverage of SSURGO Database
Incidents/Point Features	Miscellaneous external permit and complaint files; field reconnaissance	In-house Tabular input; Download from data collector	NA	N/A
Historical Archeological Sites	Existing maps	In-house Trace Digitize	N/A	N/A
Recreational Facilities	Orthophoto, Existing facility maps	In-house Trace Digitize	N/A	N/A
Demography	TIGER files	File Import	N/A	N/A
Meterological Data	Existing precipitation maps	Trace digitize In-house	N/A	N/A
Biological Distribution	Existing files	Trace digitize In-house	N/A	N/A
Dam Locations	Existing maps	Trace Digitize In-house	N/A	N/A

Likely Cost Range for Sub-regional Tier **$1,500,000** **to** **$2,700,000**

[1]No cost has been entered if layer is likely candidate for in-house conversion.

[2]Cost to West Virginia will be dependent on the level of cost sharing available from the federal government.

[3]Cost to West Virginia will be dependent on possible cost sharing with the SCS.

TABLE 2-12
REGIONAL TIER
DATABASE SOURCES, CONVERSION METHODS AND COST

Layer	Data Source	Conversion Method	Range of Probable Cost	Unit of Measure
Survey Control	Sub-regional Tier	Aggregate from Sub-regional Tier	N/A	N/A
Orthophoto/Planimetric Mapping	*Not recommended for this Tier*	N/A	N/A	N/A
Transportation	USGS 1:100,000 scale DLGs	Translation to selected GIS format	Less than $1,000 for statewide coverage	DLGs providing coverage for all of West Virginia
Surface Hydrography	USGS 1:100,000 scale DLGs	Translation to selected GIS format	See Transportation	See Transportation
Topography	USGS 1:250,000 DEMs	Translation to selected GIS format	N/A - Already Possess Files	N/A
Political Boundaries	USGS 1:100,000 scale DLGs	Translation to selected GIS format	See Transportation	See Transportation
Administrative Boundaries	Sub-regional Tier	Aggregation from Sub-regional Tier	N/A	N/A
Parcels	Sub-regional Tier	Aggregation from Sub-regional Tier	N/A	N/A
Energy Transmission	USGS 1:100,000 DLGs (where data available); Sub-regional Tier	Translation to selected GIS format, and/or aggregation from Sub-regional Tier	N/A	N/A
Utility Distribution/ Collection	*Not recommended for this tier*	N/A	N/A	N/A
Land Use/Land Cover	Satellite imagery and high altitude aerial photography augmented by USGS GIRAS data and USFWS NWI maps	Automated image classification and manual interpretation	$100,000 to $250,000[1]	N/A
Subsurface Hydrogeology	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Bedrock Geology	WVGES 1:62,500 and 1:250,000 scale Geological Mapping	Board Digitize In-house	N/A	N/A
Floodplains	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Well Locations	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Environmental Features	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Soil Classification Boundaries	SCS Soil Association mapping at 1:250,000-STATSGO	Purchase from SCS	$500 for STATSGO Database	N/A
Incidents/Point Features	Miscellaneous field collection efforts; permit files; incident reports	Miscellaneous In-house Programmatic Efforts	N/A	N/A
Historical Archeological Sites	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Recreational Facilities	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Demography	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Meteorological Data	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Biological Distribution	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A
Dam Locations	Sub-regional Tier	In-house Aggregation from Sub-regional Tier	N/A	N/A

Total Cost Range for Regional Tier **$100,000 to $252,500**

[1]Cost depends on classification detail decided upon, amount of in-house labor used, and potential use of low-cost staffing from university. Cost included LANDSAT TM scenes and labor costs for classfication work.

TABLE 2-13
STATEWIDE TIER
DATABASE SOURCES, CONVERSION METHODS AND COST

Layer	Data Source	Conversion Method	Range of Probable Cost	Unit of Measure
Survey Control	*Not recommended for this Tier*	N/A	N/A	N/A
Orthophoto/Planimetric Mapping	*Not recommended for this Tier*	N/A	N/A	N/A
Transportation	Regional Tier; or USGS 1:500,000 State Map	Aggregate from Regional Tier; or Trace Digitize USGS Map	N/A for Aggregation or In-house Digitizing	N/A
Surface Hydrography	Regional Tier; or USGS 1:500,000 State Map	Aggregate from Regional Tier; or Trace Digitize USGS Map	N/A for Aggregation or In-house Digitizing	N/A
Topography	Regional Tier; or USGS 1:500,000 State Map	Aggregate from Regional Tier; or Trace Digitize USGS Map	N/A for Aggregation or In-house Digitizing	N/A
Political Boundaries	Regional Tier; or USGS 1:500,000 State Map	Aggregate from Regional Tier; or Trace Digitize USGS Map	N/A for Aggregation or In-house Digitizing	N/A
Administrative Boundaries	Regional Tier	Aggregate from Regional Tier	N/A	N/A
Parcels	*Not recommended for this Tier*	N/A	N/A	N/A
Energy Transmission	Regional Tier	Aggregate from Regional Tier	N/A	N/A
Utility Distribution/ Collection	*Not recommended for this Tier*	N/A	N/A	N/A
Land Use/Land Cover	*Not recommended for this Tier*	N/A	N/A	N/A
Subsurface Hydrogeology	*Not recommended for this Tier*	N/A	N/A	N/A
Surficial Geology	*Not recommended for this Tier*	N/A	N/A	N/A
Floodplains	*Not recommended for this Tier*	N/A	N/A	N/A
Well Locations	*Not recommended for this Tier*	N/A	N/A	N/A
Environmental Features	*Not recommended for this Tier*	N/A	N/A	N/A
Soil Classification Boundaries	*Not recommended for this Tier*	N/A	N/A	N/A
Incidents/Point Features	*Not recommended for this Tier*	N/A	N/A	N/A
Historical Archeological Sites	*Not recommended for this Tier*	N/A	N/A	N/A
Recreational Facilities	*Not recommended for this Tier*	N/A	N/A	N/A

TABLE 2-13 (continued)
STATEWIDE TIER
DATABASE SOURCES, CONVERSION METHODS AND COST

Layer	Data Source	Conversion Method	Range of Probable Cost	Unit of Measure
Demography	*Not recommended for this Tier*	N/A	N/A	N/A
Meterological Data	*Not recommended for this Tier*	N/A	N/A	N/A
Biological Distribution	*Not recommended for this Tier*	N/A	N/A	N/A
Dam Locations	*Not recommended for this Tier*	N/A	N/A	N/A
Total Cost Statewide Tier			**$0**	

SECTION 3
SYSTEM CONFIGURATION CONCEPT

3.1 OVERVIEW OF SECTION

This section of the report presents PlanGraphics' concept for a statewide GIS hardware configuration and data communications. The configuration concept covers the following.

- Data communications
- Linking the GIS to other systems
- Typical hardware configuration.

3.2 SYSTEM CONFIGURATION ASSUMPTIONS

PlanGraphics has made the following assumptions during its analysis to formulate recommendations regarding the system configuration:

- The GIS must allow for communication and data exchange among sites which are geographically dispersed around the state. There will be no single, central GIS that will be used by all potential GIS participants. Rather, the system will allow for a decentralized approach to GIS.

- Existing use of GIS by state government departments, universities, and other organizations, will influence, in part, decisions on hardware and software used by participants in West Virginia's GIS.

- A decision should be made to identify a small number of commercial and/or public domain GIS software packages for use by participants. At the same time, procedures should be instituted to ensure the ability to exchange data with other automated mapping and GIS software packages.

- Systems should grow incrementally through the series of development phases, with measured expansion of hardware and software as GIS databases and applications are developed.

- System acquisition and development should be guided by appropriate standards for hardware, software, and data communications, in order to encourage long-term system integration and flexibility. To the greatest extent possible, GIS standards should be developed in concert with broader information system and database standards being defined through IS&C.

- There are many existing non-GIS tabular databases and files residing on mainframe and microcomputer systems which contain geographically-related data useful for GIS applications. There must be procedures established for efficient exchange of data with these systems.

- UNIX workstations and servers will be the principal hardware platform for GIS, although DOS microcomputers will be used as stand-alone automated mapping workstations or networked to GIS servers for direct query and display. It will be assumed that existing microcomputers will be used, where possible, to support appropriate GIS applications.
- The technology itself is changing, and computer hardware and software vendors are making rapid advances in product performance and degree of "openness" allowing for better integration with other systems. The GIS should be developed in such a way to provide the greatest flexibility in growth and upgrade.

3.3 COMPUTER HARDWARE, SOFTWARE, AND DATA COMMUNICATIONS ISSUES

General concepts and specifications guiding this conceptual design are explained in Appendices B, C, and D. The reader should examine these appendices before evaluating the specific recommendations presented in this section. The recommendations presented in the next sub-section recognize that GIS configurations have evolved from traditional centralized approaches using non-intelligent workstations connected to host minicomputer processing units to a more distributed processing environment based on powerful UNIX workstations and servers. This implies that users have a significant amount of local processing power and mass storage of data at their sites. Microcomputers running DOS-based software may be used as part of an overall GIS configuration for limited stand-alone processing or as query workstations connected to a network or directly to a workstation or server.

The principal hardware devices that will be used in the GIS are described in detail in Appendix B. Table 3-1 summarizes the important characteristics and role of these devices in the GIS. These specifications provide a basic characterization of hardware devices that will be included in the GIS, but it is not meant as a set of detailed specifications for procurement. Technical and environmental factors must be taken into account for particular users and specific sites when a procurement is being made.

A general overview of a mature GIS configuration is presented in Figure 3-1. The reliance on UNIX servers and workstations supporting users connected to a local area network and/or acting in a stand-alone mode with remote connection to an existing network is illustrated. This configuration concept will guide system development for the West Virginia GIS. This recommended model for a GIS network does not imply any specific local area network protocols as "standard." Hardware and software vendors would be required to specify the details of local area network protocols when purchase specifications are made.

Since the State has accepted IBM's Token Ring protocols as a standard for local area networks, it should be noted that not all GIS software vendors and UNIX hardware platforms support Token Ring. Currently, most use the Ethernet standard and employ higher level standard protocols, Network File System (NFS) and TCP/IP, to provide communication between UNIX workstations and servers from multiple vendors. For integration with existing Token Ring networks, it is possible to use special hardware and software to establish bridges with the GIS network to support data exchange and real-time communication with microcomputers.

Table 3-1

Table 3-1 cont'd

Table 3-1 cont'd

Figure 3-1

3.4 REMOTE DATA COMMUNICATIONS

Organizations that will eventually require access to a statewide GIS are located throughout West Virginia. While there are clusters of potential GIS users - the cities of Charleston and Morgantown contain the largest concentrations of study participants - a great many of the organizations are geographically disbursed around the state. Physical distance between GIS user sites has significant impact on data communications characteristics, timing associated with adding users, and methods of data management.

In addition to the geographic distribution of West Virginia GIS participants, several other factors must be considered in determining an appropriate configuration for a statewide system. These factors include distribution of responsibilities; development of local, sub-regional, regional, and statewide databases; and data sharing with existing systems.

Communication links to support exchange of data between remote sites will use media exchange, direct batch file transfer, or interactive access techniques as explained in Appendix D. We do not believe that high-speed interactive communications should be established between all remote sites. In many cases, GIS applications can be supported adequately using less expensive and less technically sophisticated media transfer (using an exchange of tape or disks) or lower speed on-line batch file transfer.

PlanGraphics envisions that West Virginia's GIS configuration will grow in a gradual, phased manner. Part of this growth encompasses a combination of additional users and a greater desire to share data between them. We have prepared descriptions and accompanying figures to portray three stages of remote GIS communications in West Virginia. These stages are essentially "snapshots" of long-term system development. Considering the large number of variables that will determine the actual system characteristics, these three stages should be viewed as a general indication of the growth path and not specific, detailed recommendations.

In cases where on-line communications is necessary for interactive or batch file transfer, we recommend making use of existing wide-area communications facilities to the greatest extent possible. Several options exist which should be further examined in a detailed system design:

- Use of the State's IS&C Communication Network
- Leasing of dedicated lines from C&P Telephone
- Use of communications facilities managed by WVNET.

The IS&C network is designed specifically to link data terminals and peripheral devices used by State government offices to the IBM mainframe located in Charleston. This network connects remote mainframe terminals and peripheral devices (generally at speeds of up to 19.2 KB/sec.) and supports IBM's System Network Architecture (SNA) protocols. The existing high level of traffic (at least during normal business hours) and protocol dependencies of this network limit its use to support GIS users on a routine basis.

C&P Telephone maintains a fiber optic network linking many areas of the state with high-speed communications that can support both voice and data communications. At this time, a large number of cities in all portions of the state are linked by fiber optic lines, and additional links will continue to be added through the next four years. These fiber optic lines will, in most cases, support speeds needed for large batch file transfers and considerable interactive GIS access between remote sites.

WVNET administers the state supported network relying on a variety of remote communication links to support data communications to state offices and other sites at virtually all critical sites around the state.

Considering the cost, speed, and communications management concerns impacting remote communications, it is recommended that the GIS use the communications facilities administered by WVNET as the primary means for remote communications for the GIS. It may be necessary, in some cases, for GIS users to lease point-to-point lines directly from C&P to support remote communications between specific sites if such links are not directly suppported as part of WVNET-administered facilities.

PlanGraphics has prepared a description of three stages of remote data communications development in West Virginia. These descriptions are accompanied with figures to provide an indication of a logical growth path. We have not identified specific organizations involved with each stage; the variables are numerous and under the control of separate organizations.

3.5 LINKING GIS WITH OTHER COMPUTER SYSTEMS

As mentioned previously, there are existing tabular databases containing tabular GIS data that may be important for GIS applications. These include databases located on mainframe or minicomputers as well as smaller departmental-specific databases on microcomputers. The GIS database is designed to hold map features and a defined set of “core” tabular attributes associated with those map features. It is not intended that the GIS will store and independently maintain a full range of attribute data that may reside on external database systems. It is therefore necessary that techniques and procedures be established to exchange data between the GIS and these external systems.

The ideal situation is for the GIS to directly access the data files on the external system at high speed and interactively. This would eliminate a need to go through a step of batch file transfer and redundant storage of data on two or more systems. In a practical environment, there are limitations to high-speed “transactional” communication between the GIS and external systems. These limitations have to do with incompatible communication protocols; lack of standardized application interfaces between the GIS and external system; incompatibilities in the database design; access security restrictions;

and speed limitations of the physical network. It will be advantageous to integrate existing DOS-based microcomputers with the GIS. Issues and approaches for GS integration with mainframe and microcomputers systems are discussed below.

3.5.1 Information Exchange with Mainframe Systems

The IBM mainframe systems use proprietary operating systems and communication protocols which limit the ability to link them directly to GIS networks for sharing tabular data. In planning an approach for information exchange between the IBM mainframes and the GIS, different networking concerns must be examined:

- **Physical network** factors involving the type of cabling used to connect the IBM mainframe with the GIS network and the impacts of distance on the type and speed of communications that are implemented.

- **Network protocols** defining the structure of information packets which are sent along the network. The packet contains the actual message or data along with control information defining addresses for the packet and error checking information. Direct communication between systems requires that a protocol conversion occur whenever information is passed between systems.

- **High-level database and application communication** impacting the way specific data elements are extracted from one system to be sent to another. Databases on the IS&C IBM mainframe system are currently maintained under several software environments (e.g., VSAM files with access via Cobol and Fortran application programs, SAS database management system, and DB2 database management system). Whether or not a specific GIS application requires the exchange of a small or large volume of data in an interactive or batch mode, a processing step for extraction of the data must occur. For databases supporting the SQL standard (like DB2), SQL calls can be embedded in data requests. For VSAM files or other database formats, customized programs may have to be written on the IBM to extract the data. Some applications may also require some level of reformatting of tabular data after extraction.

There are many software and hardware products (from IBM and third-party vendors) on the market today that are designed to facilitate information exchange between IBM mainframes and other systems, including UNIX-based systems on which the GIS will be built. Multi-protocol routers and gateways include both hardware and software components to link systems and perform protocol conversions for terminal emulation, batch file transfer, and interactive communications.

IBM is in the process of defining its System Application Architecture (SAA) which is a general blueprint for increasing the capacity and flexibility for integration of mainframes with other systems of IBM manufacture (e.g., PS/2s, AS/400s, RS/6000) and computers from other vendors. IBM is also a member of the Open Software Foundation (OSF), a multi-vendor consortium with goals to develop a portable operating system and a suite of network communication and management tools to operate transparently on platforms of multiple vendors. Again, product development from the OSF is still at an early stage, and there is no indication at present regarding how IBM's participation in the OSF will impact communication with its mainframe systems.

In the foreseeable future, information exchange between the GIS and the IBM mainframe must rely on approaches using hardware and software products from a variety of vendors with custom programming for data extraction and reformatting. In the early phases of GIS development, it is recommended that IBM information transfer use technically unsophisticated approaches for media exchange, batch file transfer, and perhaps terminal emulation which would allow a GIS UNIX workstation to "look like" an IBM 327x terminal to access IBM databases. Later phases of GIS development could include testing and implementation of more sophisticated interactive data exchange approaches.

While it is possible to establish interactive access between the GIS and an external system IBM mainframe or minicomputer in some cases, it will often be sufficient to exchange data through less sophisticated (and less expensive) means. Such a data exchange can use batch file transfer or media exchange procedures as described in Appendix D. A program written on the external system would be used to extract data from its databases and formatted into a flat file for transport to the GIS. After some re-formatting, if necessary, the data is loaded into the tabular database management system of the GIS. It is recommended that batch file transfer or media exchange be used exclusively during Phase 1 and Phase 2 of the GIS. After Phase 2, more sophisticated data exchange with external systems may be examined and implemented in selected cases.

3.5.2 Integration of Microcomputer Systems

Microcomputers operating in a DOS environment may serve one of the following roles in the GIS network:

- As an independent node supporting microcomputer databases with tabular data for uploading to the GIS

- As an independent workstation for the capture of graphic data (through digitizing, scanning, using DOS-based software) to prepare files for uploading to the GIS

- As a query terminal to directly access a GIS database server to query and display graphic and tabular data. This approach would require a special software or hardware board in the microcomputer to emulate a graphic query terminal or
  X-terminal.

Most GIS software vendors provide support for microcomputers in one or more of the scenarios above, but the specific approach and flexibility will vary widely from one vendor to the next. Specific requirements that should be considered in preparation of detailed system specifications regarding support for microcomputers include:

- A standard graphic user interface for GIS data capture which has a consistent "look and feel" for both the microcomputer and the UNIX workstation

- Support for a common database management system at the microcomputer and UNIX workstation

- Binary exchange of data to eliminate the need for data translation in batch file transfers between the microcomputer and the DNR network

- Support for direct graphic query access by the microcomputer to the GIS network in a non-intelligent graphic terminal mode (through emulation) and as an intelligent X-terminal mode in which some processing is occurring locally at the microcomputer

- Bridges to existing microcomputer networks to enable the capabilities above.

Currently, the state (IS&C) supports IBM's Token Ring as a standard local area network protocol.

3.6 RECOMMENDED PHASED CONFIGURATIONS

3.6.1 Recommended Configuration-Decentralized Network

This recommended decentralized approach uses a combination of stand-alone UNIX workstations and local area networks at distributed sites in Charleston, Morgantown, and other locations around the state to support user organizations. A main site should be designated for location of most of the common GIS data layers required by multiple GIS users and to act as a service center for users needing copies of digital data or hard copy products to be generated from the GIS. Individual user sites should have local processing and mass storage capabilities through UNIX workstations. In some cases, microcomputers will be used for stand-alone processing or as query workstations accessing a UNIX server or workstation.

It is recommended that the main site be established at West Virginia University in Morgantown with technical support provided by designated University staff with support from the West Virginia Geological and Economic Survey. Initially this site will function as a technical center supporting database development, application development, and technical support. It will evolve later into a service center providing GIS data and hard copy products to users. GIS use by state agencies will continue and expand as the GIS implementation progresses. High-speed digital communication links will be established between state offices in Charleston and the Morgantown site when the GIS database reached a point where data exchange requires such a link. Relationships and technical links between the main network at other non-state government sites (cities, counties, and Regional Planning and Development Councils) will evolve over Phases 1, 2, and 3.

3.6.1.1 Phase 1 GIS Configuration

Figure 3-2 illustrates the Phase 1 configuration. As described in Section 1, Phase 1 will concentrate on establishing the organizational framework for the GIS and on initial database and application development. The system will not be highly “operational” during most of this phase except for applications already in place or those that are currently being developed by organizations that now have GIS technology. PlanGraphics does not recommend major procurement of hardware and software for users during this phase. Communications between remote sites will not be sophisticated during this phase because the status of the database and maturity of applications at user sites will not demand high-speed or frequent exchange of data. A dedicated line between Charleston and Morgantown will use a 19.2 Kbit/sec or 56Kbit/sec line for data exchange, and other data exchange will be supported by media transfer using cartridge tapes or disk.

During this Phase, State government and WVU representatives will work closely with federal agencies and private companies in database development. GIS development by local government agencies and Regional Planning and Development Councils will be encouraged and supported and data will be exchanged with these organizations through media transfer.

3.6.1.2 Phase 2 Configuration

In this Phase, GIS use has been expanded within state government agencies and additional local governments, and Regional Planning and Development Councils have instituted GIS programs. The communication network has increased in sophistication at local sites and in linking remote sites together. Figure 3-3 shows that local area networks have been installed in Charleston and Morgantown and a high-speed communication line has been established to link these sites. Other remote users still rely largely on media transfer to exchange data, although where appropriate, remote users may be linked by direct lines for batch file transfer.

It is anticipated that a significant amount of additional hardware and software will be acquired during this Phase to support an expanded set of users and applications supported by a completed Tier 3 and 4 database and parts of the Tier 2 database in development during this Phase. The role of the Technical Support Center at WVU will increase to fulfill a demanding need for technical assistance and application development. Additional technical staff will also be needed in Charleston to provide on-site support to systems in operation there.

Figure 3-2

Figure 3-3

3.6.1.3 Phase 3 Configuration Overview

Phase 3, depicted in Figure 3-4, is characterized by a well-integrated system with a completed Tier 2, 3, and 4 database. GIS use has expanded considerably in state government agencies in Charleston, Morgantown, and other state offices. Significant hardware/software procurements have been made to support expanded use, and existing systems are upgraded as necessary. Individual local area networks operating at Morgantown and Charleston are linked by high-speed bridges, and high-speed communication lines are established with key remote sites at cities, counties, and Regional Planning and Development Councils. Regional Planning and Development Councils have taken on the role as the principal GIS service centers for small communities within their region. During this Phase, high-speed links are also put in place to exchange data with external systems such as the state's mainframe in Charleston.

3.6.1.4 Phase 4 Configuration

No specific configuration is provided for Phase 4. It represents long-term operation of the system characterized by additional expansion of GIS use, particularly at the local level, and greater integration of systems with higher-speed communication lines lining remote sites around the state.

Figure 3-4

3.7 RECOMMENDED HARDWARE PROCUREMENTS AND COSTS

To keep pace with system expansion over Phases 1, 2, and 3, hardware and software acquisitions must occur to provide a base for database development, application development and testing, and actual execution of the applications by system users. It is premature to make specific recommendations for acquisitions for each organization participating in the GIS. To support long-term budgeting, PlanGraphics presents aggregated projections (for state government agencies) of hardware/software for each of these three Phases. These recommendations, presented in Table 3-2, assume that Phase 1 GIS development will make use, to a large extent, of existing hardware and software now operated by state agencies.

Table 3-2

3.7.1 Hardware and Software Selection

This recommended configuration implies that there will be a standardization of a small number of GIS software packages to satisfy user needs for automated mapping, geographic analysis, and image processing. It is assumed that this will include designated commercial and public domain packages with procedures for flexible exchange of data between them.. In addition, a decision to support only a small number of software packages eases the problems of technical management and creates a better environment for training and development of applications.

We recommend that, as part of Phase 1 or even before the initiation of this Phase, specific packages be identified as being “pre-qualified” for use by state government agencies. Several key factors should be taken into account in this decision:

- The range of capabilities to support mapping, geographic analysis, and image processing
- Existing use and expertise by West Virginia state agencies
- Existing use and maturity of applications by other state governments
- Existing use by Universities in the State and availability of University staff for training, application development, and technical support
- History of product development and user support by the GIS vendor
- Level of use by non-state government organizations (public and private) that will be participating in state GIS programs and exchanging data with state government agencies.

A software package should be selected only after a detailed examination of all these factors. Based on this study and experience with other state governments, PlanGraphics feels that there is reasonable justification to include, as software candidates, ARC/INFO and associated packages from the Environmental Systems Research Institute (ESRI); GRASS, a public domain package from the U.S. Army Corps of Engineers; and specialized packages from the Intergraph Corporation. In addition, it will be advantageous to specify at least one DOS-based package for mapping and GIS on microcomputers. Atlas*GIS by Strategic Mapping and GeoSQL from the Generation 5 company are two possible candidates both of which have been used by organizations in the state.

Even with the identification of several pre-qualified packages, there will be other GIS packages in use, and it will be necessary, on occasion, to exchange GIS data with these other packages. For instance, CNG Transmission uses Synercom software and maintains base map data for a large part of the state. In some cases, state agencies themselves may be justified in deviating from the pre-qualified list of software if it can be shown that another package is necessary to perform specific functions not present in the pre-qualified packages. This underscores the need to consider data exchange issues in the software selection and application development process.

SECTION 4
ORGANIZATION ISSUES AND RECOMMENDATIONS

4.1 BACKGROUND

This section provides a general overview on issues and concerns critical to a successful GIS implementation. It also provides specific recommendations by PlanGraphics for an organizational structure that best addresses the needs of organizations in West Virginia. It is acknowledged that the major goal in creation of an organizational structure for GIS is the integration of functions and information to avoid redundancy and inconsistency in the wy information is managed and maintained. Any of a variety of institutional relationships could meet this objective. In the end, it will come down to the commitment and dedication of technical staff and management personnel to see that system goals are achieved.

4.1.1 Institutional Challenges

Listing the known institutional challenges can serve to further define the possible duties and roles to be assigned to different GIS organizational components and thereby provide guidance in the selection of the overall organizational alternative best suited to West Virginia. This subsection examines the internal and external factors or forces which will have a major impact on building a successful multi-organizational GIS.

The following are identified as critical success factors in GIS implementation that are impacted by the organizational structure:

- Communication between state government divisions and departments must increase and become more efficient.

- Building and maintaining the GIS database is dependent upon a close relationship with federal agencies and private companies.

- The state must encourage and support GIS development among local government agencies and Regional Planning and Development Councils, and the GIS should provide mechanisms for information exchange with these groups.

- Geographic information must be consolidated and shared, reducing redundancy of data storage, maintenance, and retrieval efforts.

- Accuracy and currency of geographic information must increase.

- The GIS must continually evolve to embrace existing and future industry technology.

- The system must be responsive to official policies, regulations, and legislation in developing applications and providing products and services to the general public and to other public or private organizations.

- A long-term financing strategy must be developed to take into account development and operational costs. This involves allocation of funds by the state as well as cooperative funding agreements with outside parties and creative ways to generate revenues through sale of GIS products and services.

4.1.2 Organizational Components

Experience in GIS operations over the past 15 years has indicated that four basic organizational components are normally critical to success. These components are a policy and decision-making unit, management unit, steering committee, and technical sub-committees. These components will operate as users, information providers, and system managers. The following is a brief description of each.

4.1.2.1 GIS Policy Body

Many organizations which are developing multi-organizational GISs create formal policy bodies. These groups consist of senior representatives of the multiple departments or divisions with programmatic applications served by the GIS. The group exercises oversight and control of the system pursuant to chartered procedural rules and the approved implementation plan. A policy body typically develops or approves broad policies and procedures that guide the day-to-day activities of staff and users. The policy body should assume a high-level policy role by making decisions on major design, development, staffing, and procurement issues, and directing the activities of a GIS Coordinator. The policy body also is responsible for dealing with and resolving problems or issues of coordination between multiple divisions or departments, and allocation of staff time to the GIS effort. The following major policy responsibilities are important for guiding and oversight of GIS development and operation:

- Providing oversight on the GIS management and coordination functions

- Approving data standards and other standards as defined in GIS planning and design documents

- Establishing priorities for application development, with appropriate emphasis on the importance of specific programs administered by different divisions

- Approving a funding plan and cost-allocation approach for participants

- Setting policies for access to the GIS, participation of outside organizations, and like-kind exchange or sales of GIS products and services.

PlanGraphics recommends that the GIS Steering Committee take a lead role soon in formation of this Policy Body. We also recommend that consideration be given to broadening the scope of this Body to include policy issues relating to all information systems issues in state government--not GIS alone. The rationale for such recommendation is that GIS should be one part of a larger information system strategy. Issues regarding standards, procurement policies, the state's data communication infrastructure, etc., impact all information system areas which should be considered together.

4.1.2.2 GIS Management

GIS management should be centralized within a single management unit under CLER. The critical roles for staff assigned to GIS management positions are implementing and managing policies and procedures approved by the policy body. GIS management includes coordination activities among users and technical system administration activities. A GIS coordinator or manager would be permanently assigned and be in direct communication with the policy body and agency coordinators.

GIS management responsibilities include non-technical administration and coordination functions as well as technical administration activities. Some specific GIS management responsibilities (regardless of the specific management structure) include:

- Preparation of detailed management and action plans and system specifications
- Leading the preparation of a detailed database design and data standards with input from user divisions
- Systems operation and administration
- Leading the preparation of standards and specifications to guide hardware and software procurements
- Management of contracts for database development and other activities affecting multiple divisions
- Routine database administration and oversight on compliance with standards
- Liaison with outside organizations for use and access to the GIS
- Technical support and education for GIS users.

4.1.2.3 GIS Steering Committee

The Steering Committee would consist of designated coordinators from key agencies participating in the GIS. The individuals on the Steering Committee are required to serve as liaisons between the users in their particular agency, their representative on the Policy Body, and the GIS Management Unit. The primary role of the Committee is to keep abreast of developments within the Management Unit, as they relate to respective agencies' GIS applications, and communicate this to GIS users within their agency. Agency coordinators on the Steering Committee also have a role in communicating and educating their agency's users about GIS standards and procedures, with a responsibility to see that they are adhered to when a division is involved in application development projects. The Steering Committee should communicate with the Management Unit and other agencies to avoid duplication of effort and ensure adherence to technical standards.

4.1.2.4 Technical Sub-Committees

From time to time during planning, development, and operation phases, temporary sub-committees of the Steering Committee may be formed to investigate or research specific issues, technical approaches, products, etc., and to provide recommendations to the GIS Management Unit. These sub-committees would be assigned specific tasks with deadlines and would include designated members of the user group and other appropriate agency representatives who are not formal members of the Steering Committee.

4.2 RECOMMENDED ORGANIZATIONAL STRUCTURE

Based on an evaluation of management approaches, PlanGraphics recommends that West Virginia develop an organizational structure that has the following key elements.

4.2.1 Recommended Organizational Components

4.2.1.1 GIS Policy Body Recommendations

We recommend establishing a state government Policy Body with representatives from all state government divisions that are key participants in the GIS. In addition to representatives from key user agencies serving on the existing Steering Committee, this Policy Body should have representatives from the Governor's Office, and Information Systems and Communications (IS&C). In keeping with the high-level function of this Body, members should be selected from senior management levels and should have the authority to make decisions on issues impacting statewide policy, funding and expenditures, inter-agency coordination, and relationships with organizations outside state government.

4.2.1.2 GIS Management Recommendations

GIS coordination and system administration should be centralized to a sufficient degree to support the work of individual divisions and ensure compatibility and consistency in GIS development. To address multi-agency coordination and management, we recommend that the state establish a full-time GIS Coordinator position. This position, during Phase 1 will be administratively assigned to an existing state government office (See 4.2.2) without the formation of a new formal GIS Management Unit. Before the end of Phase 1, the state should establish a formal GIS Management Unit with necessary staff to handle the on-going operation of an expanded GIS. At the beginning of Phase 2, this staffing of the Management Unit should be expanded to include a full-time System Administrator (see discussion below in Subsection 4.3) and an administrative support staff person. Recommendations of the organization placement of the Management Unit are presented in Subsection 4.2.1.5.

In addition to this GIS coordination position evolving to a formal management unit in Charleston, we recommend that a Technical Support Center be established at West Virginia University. The purpose of this center is to provide support, in coordination with state government personnel, in technical areas of system development and operation, including the following:

- Preparation of detailed database designs
- Database development for specified GIS layers with an oversight and quality control role on GIS layers development by other parties (e.g., federal government agencies, private contractors)
- Maintaining and providing access to key common GIS layers
- Working with users to design, develop, and test applications
- Training and technical assistance in the use of GIS software
- Assistance in the promotion and expansion of GIS use in local governments and Regional Planning and Development Councils
- Acting as a service bureau to provide GIS products and services to government and non-government "customers."

The justification for such a center at WVU is that it capitalizes on the existing strong GIS expertise and facilities located there. To a large extent, this center would make use of existing facilities and resources (hardware, software, staff, building space), although some monetary support would be needed to offset additional costs to the University in filling this role. The specific duties and responsibilities of this Technical Support Center would need to be negotiated and documented in a formal agreement. The Center would operate with a designated director who would report and be responsible to the State GIS coordinator in Charleston. This Technical Support Director would not be a member of the GIS Policy Body but would sit on the Steering Committee. This Director would be active in the GIS Users Group (see below) and would participate (along with other University staff) in activities of Technical Sub-Committees.

4.2.1.3 Steering Committee Recommendations

PlanGraphics recommend that the state establish a formal GIS Steering Committee that would be based on the informal Steering Committee now established. It would include Agency Coordinators from state government agencies now on the committee along with representatives from state universities. It would formalize the function of the current Steering Committee. This Committee would have a rotating Chair. The state's GIS Coordinator would attend all meetings of the Steering Committee. This Committee would deal with important issues of a technical nature or problems and concerns involving inter-Departmental coordination. Decisions would be reached and acted upon at this level where appropriate. When necessary, this Committee would seek approval or decisions from the Policy Body and would provide the Policy Body with the background information and recommendations upon which to make decisions.

4.2.1.4 Technical Sub-Committee Recommendations

On an as-needed basis, the GIS Coordinator or the Steering Committee may establish a time-limited technical sub-Committee to investigate any technical issues and provide results and recommendations for action. Generally, these sub-groups will include selected members of the Steering Committee and GIS User Group, but they may include any state government staff, University staff, or other appropriate representatives of other government agencies or private companies.

4.2.1.5 Additional Recommended Groups

In addition to these formal bodies, we recommend that the following additional bodies be organized:

- External Organization Policy Advisory Group

  The involvement of non-state government organizations is essential for the long-term success of the GIS program. The Policy Body described above is limited to state government representatives, but it is important to have high-level input from other organizations that may affect decisions reached by this group. An External-Organization Policy Advisory Group should be established to meet this need. It would include senior representatives from all interested private companies and government agencies which have a demonstrated interest in the state's GIS. It would be the responsibility of the GIS Coordinator to seek input, by mail or phone, from these representatives and, where appropriate, to convene meetings of this group. On occasion, members of this group would be invited to participate in meetings of the state's Policy Body to provide opinions on issues being evaluated.

- GIS User Group

  The purpose of this User Group is to provide a forum for GIS users in the state to share ideas, technical accomplishments and solutions, technology news, and to share information that helps maintain a collective atmosphere for GIS development and operation. This User Group would evolve from the current West Virginia GIS Coordinating Committee (WVGISCC), and all members of the WVGISCC would become part of this group. Membership would be open to representatives from state government agencies, Universities, local and federal government agencies, and private companies who are current users of GIS or who have an interest in using GIS technology. The GIS Coordinator should participate in meetings and activities of this group. It should include at least one member from each state agency active in GIS, but multiple representatives from any single state agency may be included. The User Group should be chaired by a person agreed to by group consensus who would have the responsibility for keeping records, informing all members of the Group's activities, and organizing meetings. The chair position should rotate, probably on an annual basis.

4.2.2 Administrative Placement of GIS Management

An important decision in making this organizational structure work properly is the administrative placement of the GIS Coordinator and GIS Management Unit discussed above. In fact, there are multiple options that could work effectively assuming that responsibilities and relationships were clearly defined and understood by all parties. Some of the factors that should guide a decision on where to house state GIS management are discussed below:

- Administrative complexity

  Assignment of a new staff position followed by formal creation of a new management unit requires formal action through executive order or potentially through legislative action. It is important to decide on an approach that is not so administratively or legally complex that the process is delayed for too long a period of time.

- Real or perceived agency biases or conflicts

  The GIS Coordinator and Management Unit must be in a position that is free from unbalanced influence of a particular Department's or Division's mission and is not perceived to have biases of a particular agency that would limit working with others.

- Sufficient authority and administrative channels to respond to all state agencies

  The Coordinator position and Management Unit should be placed in an administrative location that takes advantage of administrative channels and lines of authority that permit effective communication with all state agencies. In addition, there must be sufficient authority associated with that placement to make binding decisions with assurance that actions agreed to by the Policy Body and Steering Committee are carried out.

- Proper status and authority to work with and negotiate agreements with outside parties (e.g., federal agencies, private companies)

  As already discussed, work with outside organizations will be essential. Most of the direct communication and negotiation with outside organizations will be the responsibility of the GIS Coordinator. Therefore, this individual must have an administrative status that permits effective working relationships with these outside parties.

Several options have been considered for administrative assignment of the GIS Coordinator and establishment of the GIS Management Unit:

- New, separate agency
- University group
- Information Systems and Communications
- Department of Commerce, Labor, and Environmental Resources
- Governor's Office.

PlanGraphics recommends that the GIS Management Unit be housed within CLER. We believe that this placement will provide the best overall solution for West Virginia.

4.2.3 Summary of Organizational Relationships and Structure

Figure 4-1 is a diagram depicting this structure. The “Strategic Plan” which will follow this report provides more specific recommendations on the timing for development of this organizational structure.

Figure 4-1

4.3 STAFFING ISSUES AND TECHNICAL SUPPORT IN STATE GOVERNMENT DEPARTMENTS

The previous sub-section provided recommendations about staff to support central GIS coordination and system administration functions. PlanGraphics recommends establishment of a state GIS Management Unit and a Technical Support Center to provide a central resource coordination. These organizations are intended to provide support services to the user agencies.

Coordination must also occur within the user agencies. There must be staff who understand how to interact appropriately with the state-wide GIS. To ensure proper communication and coordination among divisions, state agencies with significant GIS use will need to assign staff to GIS development and maintenance functions.

Principal roles that must be filled within user departments and divisions are briefly described below. Descriptions of positions that could be used to fill these roles are presented in Appendix F.

- <u>GIS Agency Coordinators:</u> These positions function as a Department's or Division's representative to the Steering Committee and will require approximately a 10 to 30 percent full-time equivalent (FTE) depending on the department or Division, the level of delegation of tasks to other personnel, and the stage of GIS development or operation.

- <u>Database Administration:</u> This function involves detailed GIS database design for the Division (complying with Departmental standards); establishing procedures for and overseeing database development; setting up and overseeing ongoing data maintenance; establishing procedures and overseeing database documentation and the creation of data dictionaries or meta-databases; and work with other Departments and agencies to evaluate GIS data that may be incorporated in the system.

- <u>System Administration:</u> In coordination with the Department GIS system administration staff, this function includes overseeing the acquisition, installation, and testing of hardware and software; establishing communication networks; setting up user accounts and security; and providing technical support to Departmental users on GIS hardware/software questions.

- <u>GIS Application Development/Programming:</u> This function would be performed by individuals with in-depth knowledge about the software package(s) being used for GIS as well as the structure and design of GIS databases. They would work with users of the Department (in coordination with other Departments and the State GIS Coordinator) to design applications and carry out programming, customizing, documentation, and training of users.

- GIS Database Capture: While it is recommended that outside contractors be used to develop high-volume layers of the GIS database, it will be appropriate in some cases, for Divisions to assign internal staff for data capture which may include digitizing, keyboard entry, re-formatting of existing digital data from other sources, and other steps in GIS database development. These staff people must be trained on the rudimentary skills, and user-customized entry routes to carry out this work.

Depending on the specific division and stage of development or operation, any one of these positions may be full or part-time. In early phases or where GIS use is not intense, these functions may be combined in one or more individuals. In some cases, one department or division may “share” staff time with another to satisfy internal requirements.

In several cases, most notably in CLER, staff time has been formally or informally assigned to provide these internal GIS support roles and other information system support requirements. PlanGraphics recommends that participating departments assign full-time equivalent staff to address technical administration and development activities. In the summary below, a burdened salary figure is included reflecting the recommended staff levels. Where possible, these functions should be assigned to existing staff and therefore not all of the estimated salary costs summarized here represent "new" funds for staff. A more detailed evaluation of staff requirements in departments must accompany detailed operational planning at the department and division level. These projected costs do not include salaries of the GIS Coordinator or other staff of the GIS Management Unit.

Fiscal Year	System Admin.	Database Admin.	Application Development	Data Capture	Total FTE	Projected Cost
93-94	1.0	1.0	1	2	5	$170,000
94-95	1.5	1.0	3	3	8	$280,000
95-96	1.5	1.5	3	4	10	$340,000
96-97	1.5	1.5	3	4	10	$340,000
97-98	1.5	1.5	3	4	10	$340,000

These additional staff costs represent approximately ten positions to support GIS activities. Internal decisions need to be made as to whether any or all of these positions are re-assigned from existing positions or whether they represent new hires. In all cases, a key concern is training on specific aspects of the GIS hardware and software important for these positions. WVU staff associated with the recommended Technical Support Center can supply some of this training. This should be augmented by training provided by the vendors themselves and in-house training programs conducted by State government personnel.

GEOGRAPHIC INFORMATION SYSTEM
CONCEPTUAL DESIGN
FOR
WEST VIRGINIA
APPENDICES A, B, C, D, E, and F

Submitted to:

Mr. Thomas E. Holder, Project Coordinator
West Virginia Development Office
West Virginia State Capitol
Building 6, Room B-553
Charleston, West Virginia 25305-0311
(304) 558-4010

Submitted by:

Mr. Peter L. Croswell, Executive Consultant
Tom Herrick, Design Analyst
PlanGraphics, Inc.
202 West Main Street, Suite 200
Frankfort, Kentucky 40601
(502) 223-1501

February 17, 1993

APPENDIX A
GIS DATABASE CONCEPTS AND DATA CONVERSION ISSUES

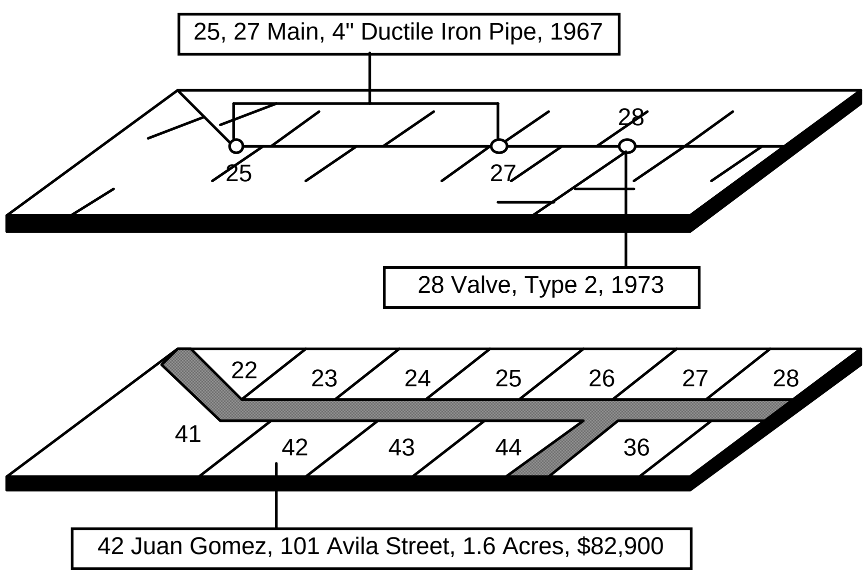
A.1 BASIC CHARACTERISTICS OF GIS DATABASES

Three major concepts are important in understanding the structure of GIS databases and the way they can be used in specific GIS applications. While different software vendors use a variety of approaches for structuring GIS data, the following concepts are fundamental in understanding most GIS databases:

- Linkage between map features and tabular attributes
- Organization of geographic information in layers
- Continuous spatial coverage.

A fundamental characteristic of any GIS is the linkage between map features (graphic) and tabular attributes (nongraphic) that hold information about the map features (see Figure A-1).

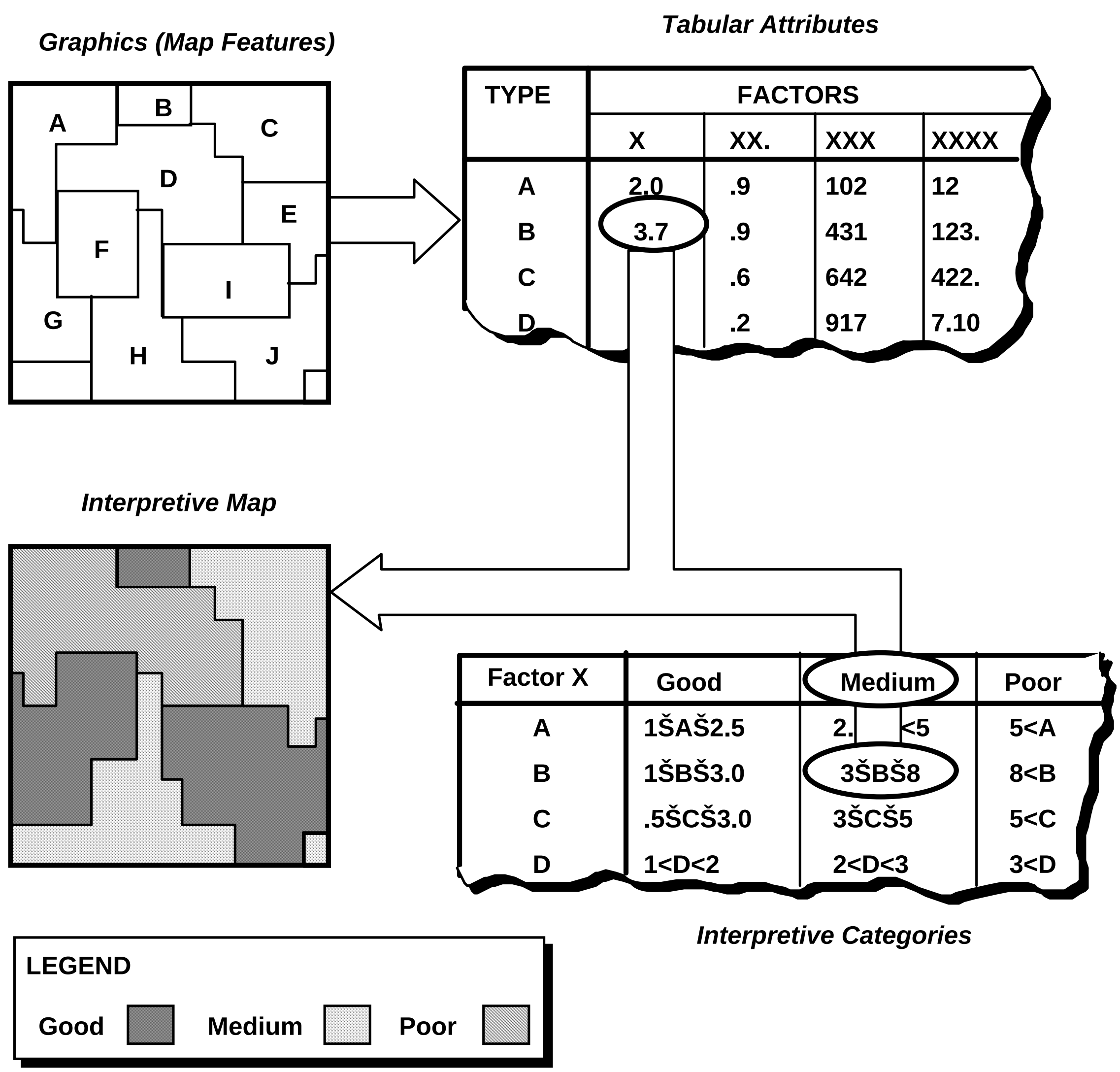
FIGURE A-1
MAP FEATURE ATTRIBUTES AND GEOGRAPHIC REFERENCE



The map features represent actual surface structure features, boundaries, or incidents' positions which are defined by X, Y, and Z coordinates. Examples of map features that may be represented digitally in a GIS database are stream lines, gas wells, parcel boundaries, or avalanche incidents. GISs offer such powerful capabilities for spatial analysis because these map features are linked with tabular attributes that may contain virtually any information about the map features that can be defined and entered by users.

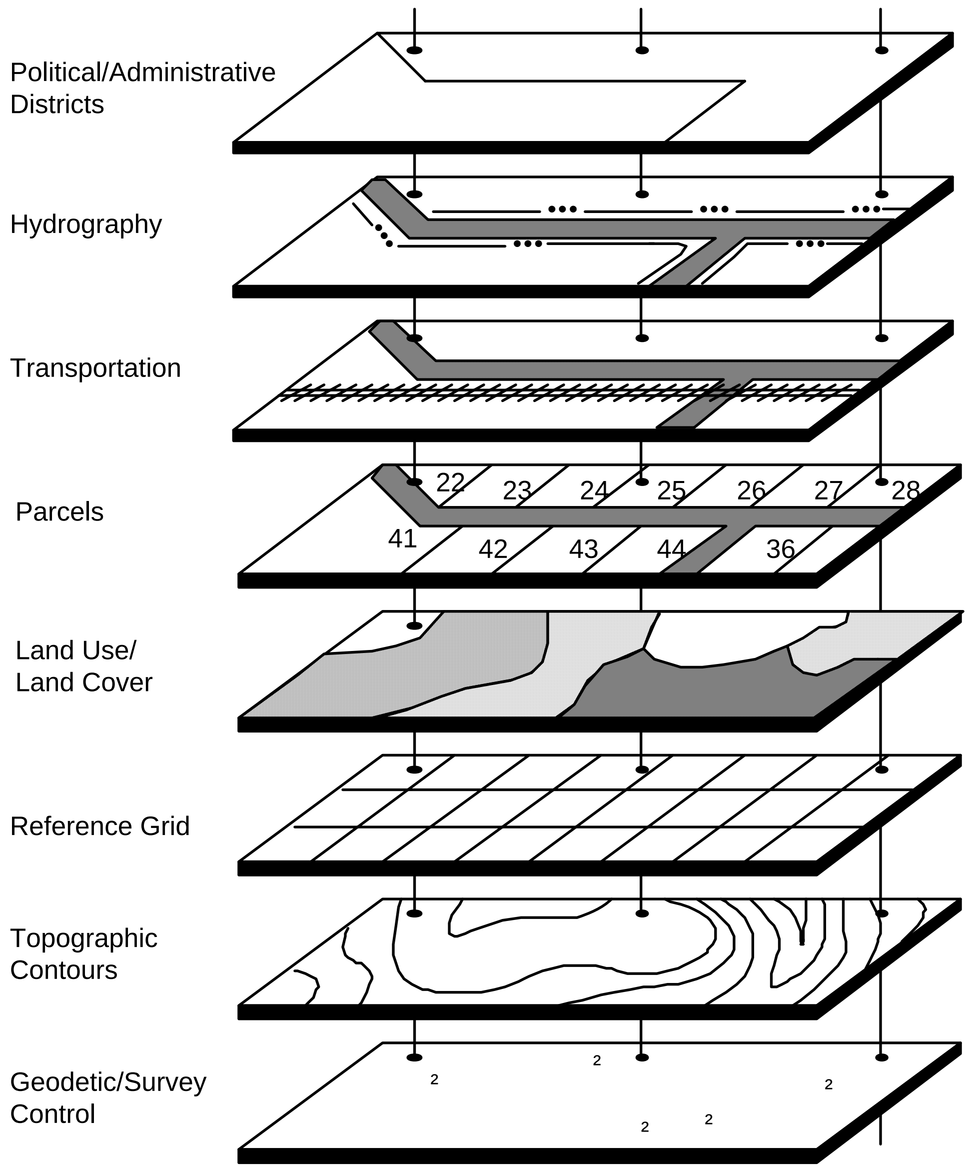
GIS software developed within this GIS database architecture provides powerful tools for users to perform mapping, geographic queries, and complex spatial analysis. Figure A-2 illustrates the use of the map feature/attribute data linkage to generate a custom thematic map.

FIGURE A-2
GRAPHIC/TABULAR DATA RELATIONSHIP



It is helpful to conceive of the overall GIS database in the form of "layers." This layering concept is based on the analogy of a manual map overlay approach in which multiple map sheets of transparent drafting material are registered together, with each sheet containing specific map information. GIS databases then can be conceived as a series of digital map sheets which are geographically registered to one another (see Figure A-3). This series of map layers begins with the creation of control points which define specific X, Y (and sometimes Z) locations that are precisely located on the Earth's surface and a base map which includes a defined set of map features or boundaries to which other layers can be referenced. As shown in Figure A-3, a variety of other map layers can be defined that address specific application needs of GIS users.

FIGURE A-3
MAP DATABASE LAYERING CONCEPT



The term "layer" should not be interpreted in a restrictive sense to describe specific physical or logical formats implemented by software vendors to store GIS data. In fact, most GIS software vendors in recent years have been using GIS storage approaches in which map features can be independently manipulated and are not tied to a specific layer in a highly structured or permanent sense. This raises the concept of "feature-based" or "object-based" techniques for storing GIS data. Despite this trend away from strict layer storage formats, the term will be used in a conceptual or cognitive manner to help organize and present GIS data requirements.

To be most effective, the GIS should allow storage of map features and attributes in a spatially continuous format over a specific geographic area (e.g., city, county, conservation district, state). Such an approach frees the user from having to access and analyze information on a map sheet-by-map sheet basis. Information for any geographic area, large or small, can be extracted, queried, and analyzed. Carried to its full extent, a continuous spatial database has no "seams" or boundaries that represent physical or logical file breaks in the storage of GIS data.

A.2 TYPES AND FORMATS OF GIS DATA

Previously, two general categories of GIS data were described - map features (graphic data) and tabular attributes (nongraphic data). GIS data storage techniques actually format graphic and nongraphic data in a variety of different formats to facilitate processing and retrieval. This sub-section briefly explains major concepts about GIS data structures as a basis for examining the requirements for an organization's GIS database.

A.2.1 Types of Map Features

Map features can be classified in the following categories, illustrated in Figure A-4, that define specific characteristics which govern how the features are stored:

- Point Features

  These are features or geographically-defined occurrences whose locations can be represented by a single x, y or x, y, z location. As stored in the GIS, points have no linear or area dimensions but simply define the location of a physical feature (e.g., control point monument, gas well) or an occurrence (e.g., avalanche incident or minerals exploration permit).

- Line Features

  Lines represent features that have a linear extent in the GIS database but no area dimensions. Centerlines of roads, stream threads, and hiking trails are examples of line features.

- Area Features

  Area features, sometimes called "polygons," have a defined two-dimensional extent and are delimited by boundary lines that encompass the area. Typical area features are property parcels, watersheds, and conservation districts. GIS data structures that store area features do not simply store the lines that represent the closed boundary; they reference that boundary to the interior area and to adjacent area features to make possible overlay operations and many types of area analysis.

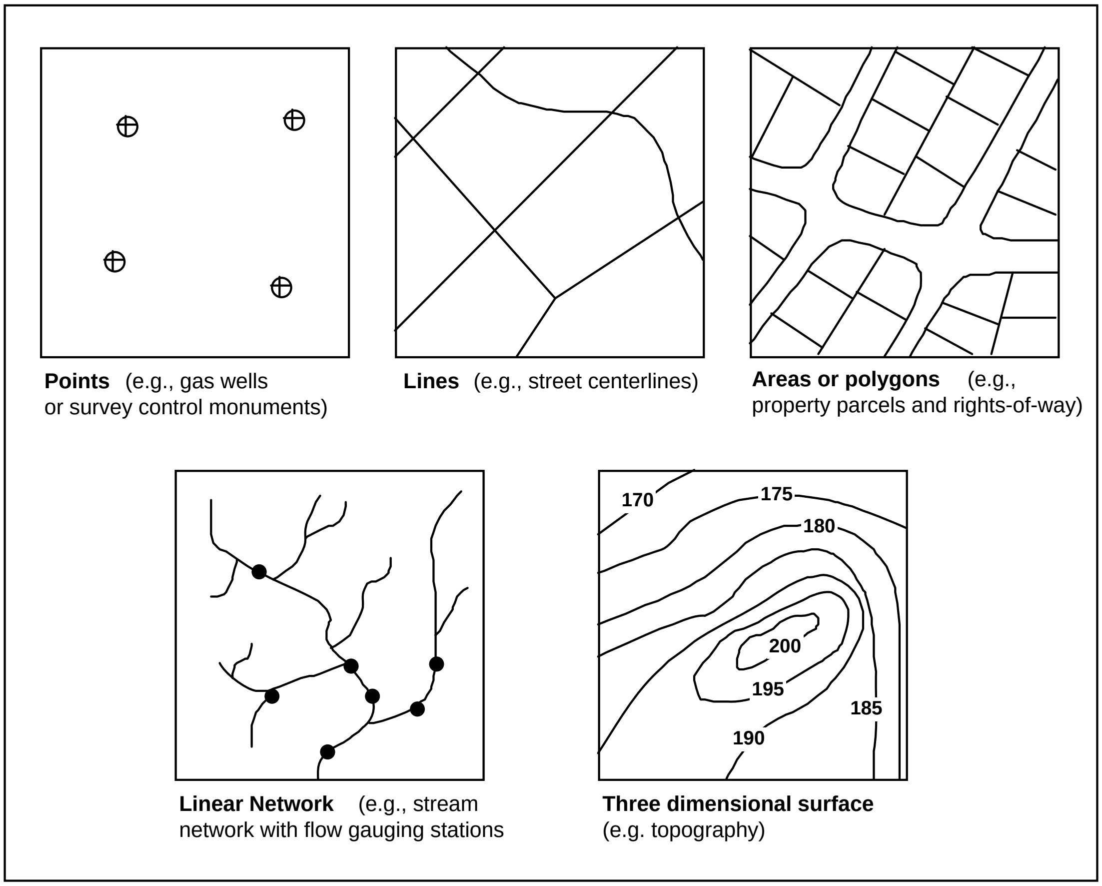
- Linear Networks

  While map features in a GIS may often be stored discretely as points, lines, and polygons, points and lines may be combined to form "networks." A network is an interconnected system of lines (and often point features) which define a pathway for flow. For example, a stream system is a network defined by linear features (e.g., streams, intermittent streams) and point features (e.g., gauging stations, dams, sampling points). The linear features define a flow path for water, and the point features are locations of changes where flow is being affected or monitored. The network storage format can be considered a "model" of a linear system that provides a basis for very useful analysis.

- Three-Dimensional Surfaces

  Some geographic information is best suited for representation in three dimensional form covering an area. An example is a topographic or terrain surface which represents the variations in elevation of the Earth's surface. Various techniques may be employed to store this data in a GIS. This concept can be applied to other spatially continuous data as well. For instance, population density or other statistical variables could be mapped as a "third dimension" to support critical demographic analysis.

FIGURE A-4
TYPES OF MAP FEATURES STORED IN A GIS DATABASE

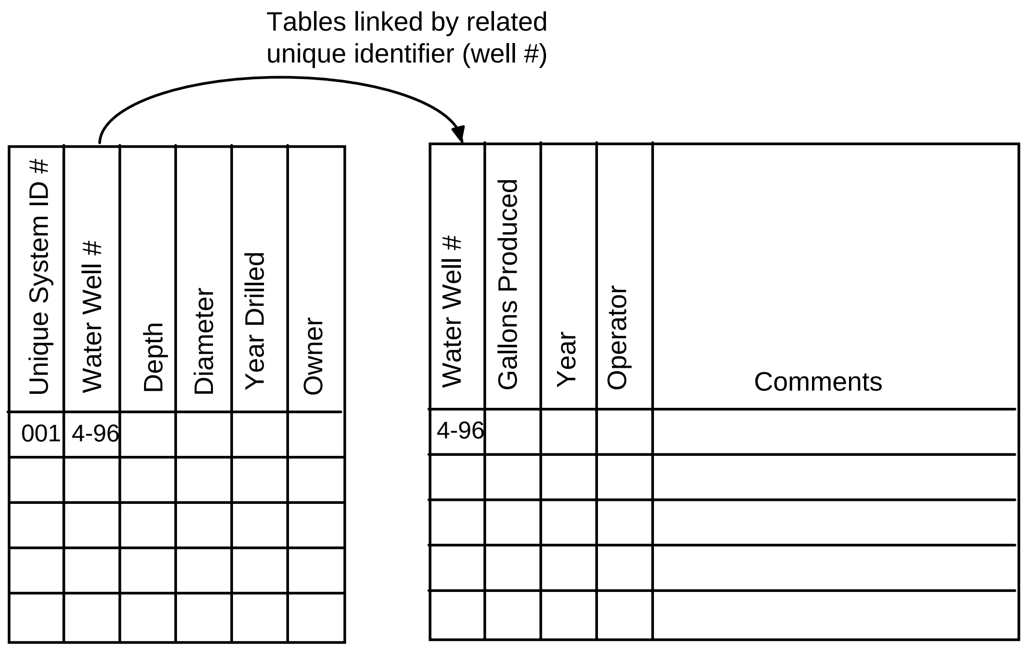


A.2.2. Tabular Attribute Data

GIS software vendors have employed a variety of techniques to store tabular attribute data and to link these data with map features. All approaches, however, use the concept of a database management system (DBMS). A DBMS allows a user to describe the particular contents of a database and the formats of data elements (e.g., integer, decimal, date, character). While some GIS software vendors have developed their own proprietary DBMSs, the trend over the past five years has been to use the capabilities of a third-party DBMS package (e.g., ORACLE, DB2, ADABAS, INGRES) and create and program a functional linkage with the map feature (graphic) database. In fact, many vendors are now beginning to support multiple third-party DBMS packages to give users greater flexibility (and to open a larger market for their GIS software).

The relational DBMS model for storing attributes is, by far, the most popular approach in the GIS software industry. The relational model is based on the storage of attributes as two-dimensional tables. As illustrated in Figure A-5, multiple tables can be linked or related based on a common and unique identifying attribute. Columns represent what traditionally have been referred to as data "elements," and rows represent data "records."

FIGURE A-5
RELATIONAL DATABASE TABLES

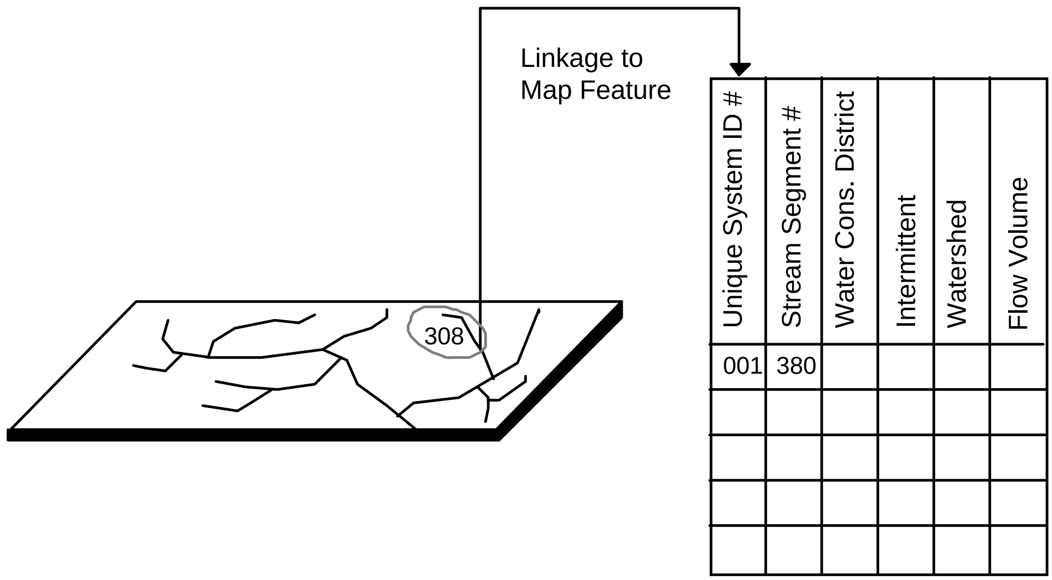


Wells Table

Related Table for Production History

As in any tabular database, a GIS requires records or rows that can be uniquely defined. In relational databases, there are one or more key fields (columns) in each table that identify records (rows) as being unique. In Figure A-5, "water well #" is a key field that uniquely identifies records. Map features in GISs are "tagged" with a unique number (usually an internally generated sequential number) that is used to link the feature to a database as depicted in Figure A-6. The capabilities of GISs to manipulate both map features and tabular data allow virtually any automated tabular files to be linked for GIS analysis as long as there are unique identifying fields that can be matched to map features. This capability allows GISs to make effective use of previously automated databases without requiring redundant data entry or update.

FIGURE A-6
LINKAGE OF TABULAR ATTRIBUTES TO MAP FEATURE



A.2.3 Raster and Vector Data

GIS graphic data are stored in either vector or raster form. GIS data structures adhering to a "vector" format store the position of map features as sequences of x, y (and sometimes z) coordinates. A vector format represents the location and shape of features and boundaries precisely. The precision is limited only by the accuracy and scale of the map compilation process; the resolution of input devices; and the skill of the operator inputting data.

In contrast, the "raster" or "grid-based" format (shown in Figure A-7) generalizes map features as cells of a grid matrix. The fineness of the grid or, in other words, the size of the cells in the grid matrix, will determine how accurately original map features are represented. There are advantages to the raster format for storing and processing some types of data in GIS. Two principal cases where the raster format is very useful are explained below:

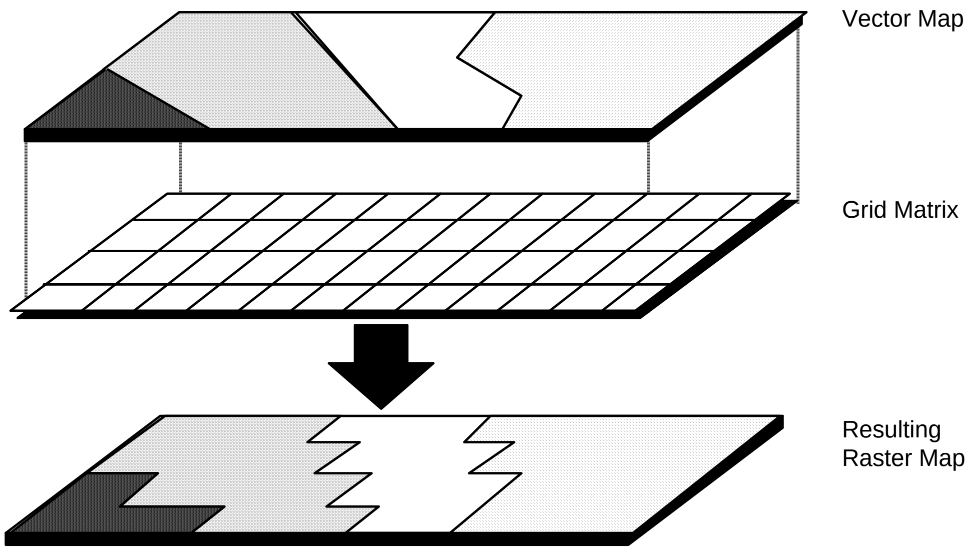
- Map display and plotting

  Most graphic screen display devices and many hard copy plotting devices present data in raster form which, in most cases, has been converted from an original vector format. This conversion, which occurs rapidly at display or plot time, creates raster formatted data with a very small cell or pixel size and, thereby, represents the original vector data quite accurately. Many high resolution graphic monitors display images at a resolution of greater than 80 pixels per inch. Raster plotting devices such as electrostatic plotters can typically generate hard copy maps at a resolution of 400 pixels per inch.

- Storage and manipulation of images

  Some data that may be important for a GIS does not require "intelligence" at the map feature level. This is appropriate for data that is simply an image such as an aerial photograph or orthophotograph. This type of product is a "map" in the sense that it represents surface features for a specific geographic area, but it is different from vector GIS databases described earlier, because individual map features are not uniquely defined or linked with attribute databases. This type of image can be encoded in raster form to create a digital image that can become part of the GIS database. Many GIS software vendors have developed capabilities to overlay digital images of this type with digital layers in vector format.

FIGURE A-7
VECTOR-RASTER RELATIONSHIP

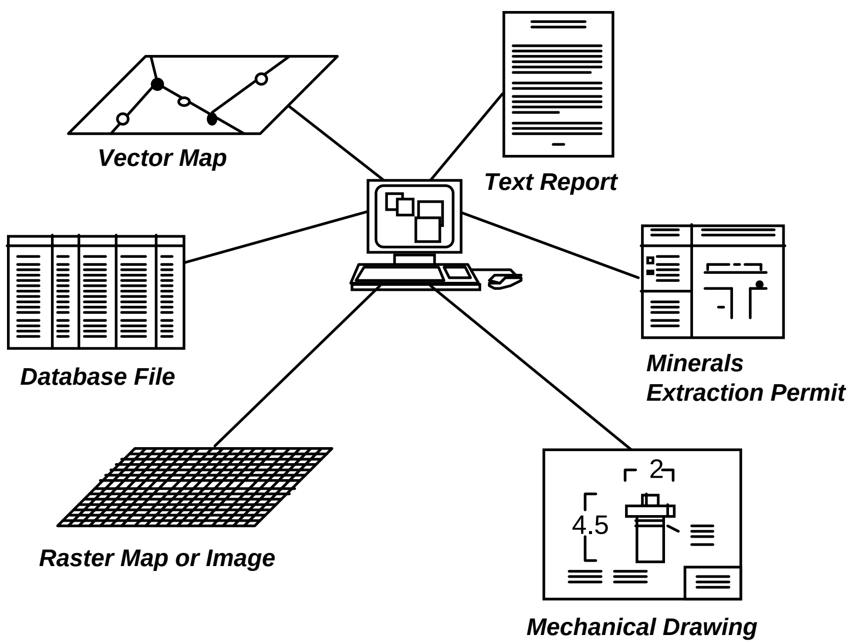


While most map data that would potentially be stored in the GIS database would be stored in vector format with associated tabular attribute data, some data would likely be maintained in raster format.

A.3. HYPERMEDIA AND GIS

Some GIS applications may effectively apply some of the concepts of "hypermedia access," which is a recently coined term in the data processing community. Hypermedia describes an information system environment which provides access to geographic information that exists in a variety of forms. In GISs, hypermedia tools can provide a way to integrate GIS data of different types and formats in a manner that is transparent to the user. Much of the discussion so far in this section has concerned the GIS database in terms of map features in vector form linked to tabular databases. In examining the user environment, however, it is apparent that there exist other types of geographically-referenced data that do not easily fit this vector map-tabular database model. Some examples of other data types include engineering or mechanical drawings, architectural plans, aerial photographs or orthophotographs, minerals extraction permits, field sketches, site photographs, or textual documents (see Figure A-8). In many cases, it is infeasible or impossible to "force" these into a vector map format. They can be considered as documents that could be geo-referenced to a location or feature on a vector map, but remain as separate automated documents that can be retrieved quickly.

FIGURE A-8
HYPERMEDIA ACCESS



Optical scanning provides one of the best approaches to automate these documents. Hard copy documents can be scanned and stored in raster format and accessed by geographic coordinate or other unique identifier on the vector map. For instance, an engineer could access a vector map, point to a specific mine site outline, and then display a detailed schematic of the underground mine layout which is stored as a separate document.

GIS software vendors are just beginning to apply hypermedia tools in their packages so the technology is still quite immature. The future, however, holds some exciting promises for the use of hypermedia concepts in an integrated GIS environment.

A.4 VIEWS OF THE GIS DATABASE

In a GIS, data display is independent from the actual storage of the GIS data. For example, different users may want to generate network maps, but each user may wish to see the map in a slightly different form. In other words, for each user's "view," there may be different criteria set for scale, colors, point or line symbols, annotation, etc. Flexible GIS databases do not store these display parameters as permanent parts of the GIS database, but provide users with tools to customize displays at the time of display. This customizing, in the case of map displays, is based on tabular attributes linked to map features. Symbols and colors, for instance, can be assigned to correspond to particular tabular attributes or combinations of attributes. Also, text annotation may be tailored by accessing particular data elements from the tabular database.

A.5 SCALE, ACCURACY, AND DATA INTEGRITY

A number of factors, discussed below, contribute to the overall quality of the GIS database. Each of these factors should be examined in development procedures for creating and maintaining the GIS database.

A.5.1 Scale and Spatial Accuracy

Spatial accuracy, in the traditional sense of inherent scalar accuracy in placement of all features mapped, is established by the base map used in any GIS. The base map is a collection of features generally used for locational reference. In the context of GIS database development, a photogrammetrically prepared orthophoto or planimetric map is used as a base to control the location of most other map information. This is particularly true where the importance of quantifying map accuracy and the ability to reliably analyze map overlays are more critical. Further, when base maps are plotted with other overlays, they provide a spatial reference that helps users identify the location of other features.

A.5.2 Completeness

Completeness relates to both map features and corresponding tabular attributes. In the case of a map "layer" of currently producing gas wells, two questions relating to completeness can be posed:

1. Are all water wells that existed at the time of mapping depicted on the map?

2. Are any features that are not actually water wells erroneously depicted on the map as water wells?

These questions illustrate the issue of "errors of omission" and "errors of commission" in mapping. With tabular databases, the concept of completeness is self-explanatory. In creating tabular databases, it may be permissible not to fill all database fields for all records, however, certain fields may be critical to the integrity of the database.

A.5.3 Coding and Classification Integrity

Integrity also relates to correctness of coding or classification. In the case of a land use map, we may ask whether all parcels coded as "high density residential" actually fall into this category. Procedures for creating and updating GIS databases are subject to coding inconsistencies that may result from errors in initial interpretation, map compilation, or data entry. Systematic procedures for database creation and quality control can be applied to reduce these types of errors to a tolerable level.

A.5.4 Currency

For many components of a GIS database, currency is a critical issue that directly influences the applications for which the data may be used. Some database components change frequently and demand routine update if the data is to be applied effectively. It is a useful approach to keep records as part of the GIS database on the dates and times of updates. In some cases, it may be advisable to include an attribute field for individual map features indicating status or date of change. Currency is an issue that must be addressed as procedures are devised for update of "layers" of the GIS database.

A.6 DATABASE DEVELOPMENT

A.6.1 Overview

This subsection describes approaches to building the GIS database. It focuses on procedures for initial preparation and conversion of source documents in their original hard copy form (e.g., paper maps, aerial photographs). It also discusses procedures for analysis or reformatting of digital data to conform with the specific GIS database design and software being used.

A.6.2 Source Material Preparation

The source documents used in the creation of a GIS database will be in the form of maps, card files, automated databases, aerial photography, and in some cases satellite imagery. There is usually some preparation of source documents required before they are usable for conversion. Source preparation may range from organizing and assembling maps, to a detailed review and revision of their content. Each data layer source material needs to be evaluated to determine the amount of preparation, or “scrubbing,” and whether the process should be undertaken in-house or by a contractor.

An example of a more complex scrubbing process involves information which has to be gathered from more than one source. Document scrubbing can be a time-consuming process; however, it allows resolution of inconsistencies, provides a source which can more rapidly be digitized, and ultimately, results in better quality information in the GIS database. Besides redrafting, document preparation tasks may include photographic enlargement, reduction and removal of extraneous marks and notes, or clarifying handwritten notes. Field verification may be required in some cases to verify map features and locations. Depending on their condition, source materials may be used as input into the geographic information system without redrafting.

A.6.3 Data Conversion Approaches

Photogrammetric Compilation

The primary data used in the process of photogrammetric compilation are aerial photography and the results of the analytical triangulation. Generally, the process involves using specialized equipment (a stereoplotter) to project overlapping aerial photos so that a viewer can see a three-dimensional picture of the terrain, known as a photogrammetric model. In the photogrammetric model, the location of survey control points can be identified, and the mathematical relationship between the control points and all other features visible on the photos can be reestablished. Using the three-dimensional view and the accurate locational relationships between features, the photogrammetrist is able to compile a map by digitizing the locations of roads, buildings, and other features on the stereoplotter. The information digitized from the photographs can then be checked and edited, annotation can be added, and the result plotted as a planimetric base map.

Other than the geodetic control, the planimetric base map is the most positionally accurate map in the system, and is used as the positional base for other data layers.

In a GIS, map features are measured, stored, and referenced in x, y or x, y, z east-west coordinates. These measurements are made on the basis of a map projection. Many projections exist, which can be divided into two types: those that recognize and depict the curvature of the Earth, and the plane coordinate projections that depict a flat surface. The plane systems are simpler and can be used without significant distortion in relatively small areas of the Earth, small enough to prevent the Earth's curvature from causing a significant error. As a result, the State Plane Coordinate System is widely used for GIS.

The State Plane Coordinate System was originally established for each state by the Coast and Geodetic Survey, now known as the National Geodetic Survey (NGS), and is referenced by monumented geodetic control points. In preparing base maps, the relationship between the control points and the reference grid is established to provide a common reference for all information subsequently mapped that uses the base maps as control.

Trace Digitizing

A digitizing workstation with a digitizing tablet and cursor are typically used to trace digitize. Both the tablet and cursor are connected to a computer that controls their functions. Digitizing involves tracing features on a source map, taped to the digitizing tablet, with a precise cross hair in the digitizing cursor and instructing the computer to accept the location and type of feature. Instructions are often coded in button-type switches located on the digitizing cursor. The person performing the digitizing may separate features into map layers, or attach an attribute to identify the feature.

Digitizing is used in a variety of data conversion tasks, including:

- Trace digitizing map graphics from source documents
- Editing trace digitized graphics
- Editing graphics entered using COGO (see discussion below)
- Editing vectorized graphic output from scanners.

Since digitizing is both labor and machine-intensive, it is among the most expensive of all data conversion processes, but the predominant method used for conversion of existing hard copy maps.

Field Surveys Coordinate Geometry (COGO)

COGO is a technique for entering boundary information to a GIS database by key entering distances and bearings surveyed from a known starting point. When distances and bearings define the position of boundary lines based on a coordinate grid, such as the State Plane Coordinate System, GIS software can use this information to create a graphic representation of the lines. This technique is most commonly used for entering real property boundaries, although some detailed locations for features in the field may be collected in automated form with positions “downloaded” to the GIS from field computers.

Map Scanning

Optical scanning systems automatically capture map features, text, and symbols as individual cells, or pixels, and produce an automated product in raster format as described earlier in this section. Scanned graphics are said to be in raster format. Most scanning systems provide software to convert raster data to a vector format differentiating point, line, and area features. However, scanning software has not advanced to the state that it can reliably differentiate scanned graphics from scanned text and symbols. Also, most scanners are still very sensitive to variations in line types, widths, and stray marks or creases on maps, and post-scanning clean-up and editing is normally a very time-consuming process. Establishing correct relationships between graphics and text has also not reached a level suitable for most municipal mapping applications.

Scanning is useful and cost-effective in GIS data conversion for two types of source documents. One application of scanning is to convert data that will remain in raster format, such as detailed engineering drawings that are only required to be displayed or plotted. Scanning has also been used successfully to convert maps consisting primarily of sharply delineated line sets with little or no text features. The line information is transformed from raster to vector format for use on a GIS.

Document Scanning

Smaller-format scanners can also be used to create raster files of documents such as permit forms, service cards, site photographs, etc. These documents can be indexed in a relational database by number, type, date, etc., and queried and displayed by users. They can also be referenced to map locations or features in the GIS graphic database. As in the case of map scanning, scanned documents usually require some clean-up and image enhancement after scanning.

Heads-up Digitizing

Heads-up digitizing capabilities work with the scanning and editing procedures described above to provide a semi-automated environment to efficiently convert hard copy maps to vector format suitable for the GIS. After a map has been scanned, the raster image, with the scanned line work and annotation, is displayed on a workstation monitor. The operator uses a mouse to invoke batch routines and/or interactively edit and clean the raster image to remove stray marks or line gaps picked up in the scanning process. The operator may then use heads-up digitizing tools to perform vector conversion and to enter annotation and attributes. Tools may allow the user to select individual raster features for vector conversion, invoke automatic line-following and thinning vector conversion, direct keying of attribute data, and other tools to speed the process of vector conversion.

Tabular Data Entry

Some of the tabular attribute data that will be part of the core database exist on maps as annotation and paper files. Information from these sources will be required for GIS applications and will have to be converted to digital form through keyboard entry.

This kind of data entry is commonplace and relatively easy to accomplish. Furthermore, GIS software can facilitate this process by providing the capability for creating data entry screens which can provide quality control on information being entered. This control involves checking entries against lists of acceptable entries or ranges of acceptable values. Use of automated quality control needs to be evaluated thoroughly for the types of data appropriate for this type of editing. It may be necessary to manually check some data elements, comparing the source document and an edit printout to ensure data quality.

Use of Existing Digital Data

Existing automated data may be available from existing tabular files maintained by state agencies or outside sources. In some cases, the federal government or other organizations may have automated map data that can fill some of the GIS database. It is possible to access multiple sources of tabular data residing on external platforms. Tabular data files that now reside on mainframes, or PCs, can be extracted and exchanged with the GIS in a batch file transfer or, in some cases, direct transactional data exchange in real-time may be feasible.

Image Enhancement and Classification

Digital raster images from satellites or, in some cases, aircraft are often used for map creation. This is particularly true in the case of land cover mapping using such satellite data sources as LANDSAT or SPOT. These raster images can be purchased for specific geographic areas and dates and used with raster GIS software. In their initial form, they consist of raster files with brightness values assigned to each pixel for one or more spectral bands. Several processing steps may be followed in using these images for mapping work:

- Geographic rectification to match these images to a coordinate grid

- Image enhancement to manipulate contrast and other visual parameters to portray greater detail

- Automated classification to analyze the spectral values of all pixels and assign pixels to specific categories representing geographic variables (e.g., land cover categories).

APPENDIX B
GIS HARDWARE CONCEPTS AND CAPABILITIES

B.1. INTRODUCTION

Computer hardware that supports geographic analysis and mapping is comparable to the hardware used for business and scientific applications. The term "computer hardware" refers to any tangible or physical device used as part of a computer system. GIS computer hardware may be categorized into the following two groups: 1) processing units,
and 2) peripheral devices.

The **processing unit** in a computer system directs and supervises all of its functions. The processing unit receives data and commands, performs requested operations, and generates a result or product that is delivered to the user through a peripheral device. Functions of a processing unit include:

- Controlling user access to the system and monitoring security
- Compiling high-level language into machine-readable form
- Controlling and monitoring communications with peripheral devices
- Performing system accounting and diagnostics
- Executing all user program commands in an interactive or batch mode.

A **peripheral device** is any hardware component that is physically separate from the processing unit and is used to enter and manipulate data; to store data; to generate screen displays or hard copy products; or to support special functions such as data communications. Peripheral devices commonly used in geographic information systems include workstations for query, analysis, and data capture; alphanumeric terminals; plotters; and other devices used for entry, analysis, and generation of products from a tabular or graphic database. In this discussion, peripheral devices are organized into the following categories:

- Mass Storage Devices
- Workstations
- Hard Copy Output Devices
- Communications Hardware.

The sections that follow provide an explanation of typical hardware devices used in GIS. Table B-1 provides a summary of basic characteristics, price ranges, and vendors who manufacture these devices.

INSERT TABLE B-1

B.2. PROCESSING UNITS

The processing unit in a computer network is the device that controls the execution of instructions and manages communication between devices on a network. Traditionally, most computer systems have used a centralized processing approach, with a "host" processing unit, to control all functions of the system. Centralized processing arrangements are still used today, although in recent years, the trend has been toward decentralization. In decentralized processing systems, some processing tasks, such as communications, plot generation, and program execution, are distributed from a main host processing unit, and handled by separate processing units. It is common now for high-speed networks to be configured for GIS applications which may contain multiple-linked processing units to distribute computing power and data to different user locations.

Processing units have traditionally been categorized as "mainframe," "mini," or "micro." Developments in computer architectures and performance over the past five years have resulted in many types of processing units that don't fall neatly into these categories, and it is valid to think of processing units today in terms of a continuum from small desktop computers to large mainframes and super-computers (see Figure B-1). While it is difficult to precisely categorize types of processing units available, it is useful to identify the following classes of computers that typically serve as platforms for mapping and geographic information systems:

- Microcomputers: These devices, often referred to as "personal computers," have limited processing capabilities and data storage capacity as compared to other computers. They normally support one user with a limited number of peripheral devices and may also play the role of a server on a network of microcomputers. Maximum main memory is usually 16 megabytes, and mass storage on disk drives does not usually exceed 150 megabytes when used as single-user devices. Main memory and mass storage capacity are higher when these devices are configured as network servers. Most of the computers in this class use the DOS operating system, but other operating systems for microcomputers also exist (e.g., Macintosh).

- Supermicrocomputers: The operating systems used on these computers (most often a UNIX derivation) support multi-tasking operations, very compute-intensive tasks, and high resolution graphics. They support a maximum main memory, often exceeding 64 megabytes, and may have mass storage capacities of over five gigabytes. These processing units support high-capacity processing and graphics. Intelligent workstations operating in a stand-alone mode or as part of a network for GIS applications are based on these types of processing platforms. They are often used extensively as data servers in a client-server network providing to access a GIS database by workstations and peripheral devices.

- Minicomputers: These are computers which can support multiple users (up to about 100) in a traditional centralized configuration or as part of a network of other minicomputers, supermicrocomputers, and other devices. They use operating systems with features that are designed for efficient multi-user access and security. Main memory capacities exceed 64 megabytes, and maximum mass storage is usually about two gigabytes. During the 1970s and for much of the 1980s, most GISs were implemented on minicomputer platforms, but today, most vendors have moved their packages to microcomputers or super-microcomputer workstations or servers.

- Mainframes: There is no absolute distinction between a minicomputer and a mainframe, but it is generally understood that mainframes have processing power and mass storage capacities that significantly exceed that of minicomputers. They support more users (500+) and have special features and operating systems designed for heavy "transaction processing" which implies intensive on-line database operations.

FIGURE B-1
PROCESSING UNIT CHARACTERIZATION

The processing unit or units in a network allow for the connection of peripheral devices that users interact with directly to enter, edit, analyze, display, and generate hard copies of data from the geographic information system. In a distributed network, the processing unit may act as a server to store and transmit data to peripheral devices, workstations, or other processing units on the network. Workstations and peripheral devices on a high-speed network may also use power of processing units to increase the efficiency of computing tasks. Under this scenario, the network processing unit is sometimes called a **compute server** as opposed to a **data server** whose primary function is to manage and transmit data. A server on a network may also be dedicated to performing certain communication functions such as controlling access to peripheral devices, or supporting bridges or gateways to other networks.

Table B-2 summarizes the roles that these processing unit platforms play in different types of computer configurations.

TABLE B-2
PROCESSING UNIT TYPES AND POSSIBLE ROLES

Unit Type	Single User Workstation	Network Workstation	Network Data Server	Compute Server	Central Host Processor
Microcomputer	x	x	x		
Super-microcomputer	x	x	x	x	
Minicomputer			x	x	x
Mainframe computer			x		x

B.3. MASS STORAGE DEVICES

The main memory capacity of a computer processing unit is limited and usually insufficient to store the large amounts of data and programs to which users need access. Mass storage devices are used to store data and programs needed by the processing unit. Mass storage devices are categorized as "off-line" or "direct access." The most popular device used for off-line storage is the **tape drive**, which is used to archive data sets on magnetic tape for later access by system users. Tape drives are considered an off-line technique because a tape physically must be mounted before data can be accessed. Tape drives are useful for the archival storage and transport of large volumes of data or software. In contrast, direct access storage devices (DASDs) allow users to store and retrieve data directly without loading physical media. The most popular direct access storage device, the **disk drive**, uses magnetic or optical technology to store information. Disk drives are connected to the processing unit through high-speed channels so that data can be retrieved directly by system users.

B.3.1. Disk Drives

Disk drives are connected to the processing unit through high speed channels and store all software and data for direct access. Disk drives are used in computer systems of all types from microcomputers up to large mainframes. In most cases, disk drives use magnetic media, however, optical drives employing a laser read/write mechanism are becoming more popular, particularly for environments requiring read-only access to large volumes of data. Microcomputer systems are often configured with floppy drives allowing the storage of data on a disk that can be removed and transported. Microcomputer systems may also have "hard disk drives" with large storage capacity for active access by users. Multiple disk drives are connected to processing units through very high-speed communication lines allowing efficient interaction with the main memory of the processing unit. For mini and mainframe computers or intelligent workstations, mass storage units may be "rack mountable" in pre-designed slots in the processing unit cabinet, or they may be external units with their own cabinet. Rack mountable units have a capacity ranging from approximately 300 megabytes to over 800 megabytes, while external units may exceed 3 gigabytes in capacity.

B.3.2. Tape Drives

Tape drives are available in various styles to meet the particular needs of users. All of them magnetically store fairly high volumes of data on flexible tapes. They are not intended to be used for quick on-line access but are more appropriate for periodic data backups, archiving data, and for inexpensive shipment of data between different sites. In a minicomputer or mainframe system, large reel-to-reel units which use removable reels of 1/2-inch width, 9-track tape are common. Smaller systems (small minicomputers, microcomputers, super-microcomputer workstations) often use cartridge tape units. These are self-contained units usually built into the cabinetry of the processing unit or workstation. There are a number of different sizes and formats of cartridge tape units with capacities up to approximately 3 gigabytes.

B.4. WORKSTATIONS

B.4.1. Full-performance Workstation

The full-performance GIS workstation is a combination of components used for the entry, edit, display, and analysis of graphic and nongraphic data. These workstations may be based on microcomputer platforms or on super-microcomputers employing the UNIX operating system. These workstations are usually equipped with significant local mass storage and local processing capability functioning either in a stand-alone mode or as part of a larger network. All capabilities of the GIS software (data entry, edit, query, and analysis) are accessible at these workstations. The main components in the full-performance workstation include one or two graphic monitors with color display, usually a digitizing tablet (at least 36" x 24"), cursor, and keyboard. Digitizing stations make use of efficient screen menus that are accessible by the cursor to issue commands, perform any analysis or modeling, and enter or edit complex graphic features. These full-performance workstations normally support all the functions of the graphic edit/query or graphic query workstations and alphanumeric terminals described below.

B.4.2. Graphic Edit/Query Workstations

Graphic edit/query workstations are similar in function and specifications to the full-performance workstation except that they are not used for digitizing functions and, depending on the specific software package being used, not all GIS analysis capabilities may be accessible. They are used to perform graphic query and display functions and can be used for editing of the graphic or attribute databases of the GIS, including the use of coordinate geometry (COGO) and precision entry procedures for graphic data capture. Many routine map update applications not requiring large digitizing tablets can make use of these devices. Depending on the vendor, the edit/query workstation may include hardware components similar to the full-performance workstation except that a large-format digitizing tablet is not included, and main memory and/or mass storage may be less than the full-performance workstation. A mouse or small digitizing tablet is normally included to manipulate the graphics and enter commands through screen menus.

As in the case of the full-performance workstations, edit/query workstations may use microcomputer or super-microcomputer platforms. In addition to the functions described above, the graphic edit/query workstation has the same capability of the query workstation and alphanumeric terminal.

B.4.3. Graphic Query Workstations

Graphic query workstations are used primarily for standard graphic and tabular query and display applications. They are normally based on inexpensive graphic terminals or microcomputers that operate in a terminal emulation mode with a host processing unit. They are configured with a mouse for the entry of commands through screen menus. Many GIS applications involving geographic queries will benefit from the display of graphic information from these devices. The graphic query workstation has all the capabilities of the alphanumeric terminal. In this category of graphic query workstation, a new type of device known as an **X-terminal** has recently been introduced by hardware vendors. These devices have local processing power but limited or no mass storage. They are designed to run on a client-server network to access and display information from the network's data server. Much of the processing required to display the information is carried out locally by the X-terminal as opposed to non-intelligent graphic terminals that rely on the server for processing power.

B.4.4. Alphanumeric Terminal

Alphanumeric terminals will be used to carry out tabular queries and other functions which do not require graphic manipulation or display. They are relatively inexpensive devices that consist of a keyboard and text monitor. They can be used for database query, report generation, macro development and programming operations, data entry, checking on system status, and other nongraphic operations. Alphanumeric terminals can be acquired with a variety of features, but generally will consist of a full character set keyboard (sometimes with special function keys) and an 80-column display.

B.5. HARD COPY OUTPUT DEVICES

This category of peripheral devices comprises those used to generate hard copies of tabular or graphic products. Maps and other products can be produced in large-sheet and page-sized formats. Devices such as **pen plotters** generate hard copy products directly from data in vector form, while most other devices produce hard copy graphics using a raster format technology. Devices best suited for printing text include high-speed line printers and low-end dot matrix printers.

The role of hard copy output devices in GIS is determined by the requirements of specific applications and, ultimately, the product's user community. Features and characteristics such as size, speed, resolution, color capability, and hard copy media should be considered. The most common categories of hard copy output devices in computer mapping systems are discussed in this section. With the wide variety of devices and plotting technologies available for hard copy output, it is important to weigh selection decisions on the key factors described below:

- <u>Print or graphic capability:</u> Devices are designed to produce hard copies of graphic and/or text products. Some devices that are primarily built for text have capabilities for low quality graphics as well. Text or graphic production needs should take into account the functional nature of the device. A device that is appropriate for printing reports or letters may be inadequate for printing more elaborate material.

- <u>Color or black & white:</u> Certain devices are designed specifically for color or monochrome (black & white) production. Those that are designed for color production use various types of printing approaches (e.g., colored printer ribbons, thermal transfer, ink-jet, electrostatic, laser) and have varying limitations for number of colors, size of output, resolution and speed. A color printer may be an unnecessary extravagance for printing letters, but may be essential for printing maps or charts.

- <u>Resolution:</u> This is a measure of graphic quality and is normally described in terms of "lines" or "dots" per inch. A specification of 150 or fewer dots per inch (typical of some dot matrix impact printers and ink-jet plotters) results in a graphic product with a very noticeable raster or stair-step appearance. A 200 to 300 dot per inch resolution (represented by ink-jet and some thermal plotters and laser printers) results in a fairly good graphic product with the stair-step appearance noticeable only with very close inspection. At resolution above 300 dots per inch (typical of some laser and electrostatic plotters), lines appear very smooth unless examined very closely. Generally, as the resolution decreases, the available range of colors decreases as well. The intended use of the printed product should determine the necessary resolution. If a product is intended for reproduction by offset printing, for example, a very high resolution may be required, while xerographic reproduction requires a lower level of resolution.

- Output size: Printers and plotters are available which provide output at a large range of sizes from "A" size (8.5" x 11") to "EE" size (40" x 46") and beyond. Most systems requiring text and low-resolution graphics can be satisfied with printing capabilities at a maximum "B" size (11" x 17"), but mapping systems and GISs generally require high quality graphic plotters for hard copy output of at least "E" size (34" x 44").

- Production speed: Production speed for text and graphic output varies widely and is dependent on the production technique and output size. Users' production speed requirements will vary among applications, but it is very important to choose a technology with appropriate production speed since this can be a main bottleneck in generation of products.

- Media type supported and media cost: Users may need the flexibility to generate hard copy products on a variety of materials—particularly in the case of mapping systems and GISs which may need to produce maps on opaque or transparent media (mylar). Also, some plotting devices require specially treated material, while other devices may allow the user to use inexpensive plain paper to print out products at lower cost.

Various plotting and printing technologies are compared in Table B-3, and price ranges are presented in Table B-1.

Insert Table B-3

B.5.1. Desktop Impact Printers

These are inexpensive devices used primarily for printing text documents, but may also be used for low-resolution graphics. Most use dot-matrix techniques in which a print head with small pins strikes a ribbon to form characters and graphics on a continuous feed paper roll. Print speeds may range from approximately 150 to 300 characters per second.

B.5.2. Small-format Laser Printers

Laser printing is a relatively new technology used to produce both hard copy text and graphics. Laser printers produce images one page at a time by using a laser to apply charges to a rotating drum. The drum is then exposed to a dry toner, which causes toner particles to adhere to charged portions of the drum. Laser printers are used to generate high-volume, high-quality text reports and graphics. Low-volume laser printers that can print page-size (8.5" x 11") documents at a rate of five to eight pages per minute are used extensively in word processing and electronic publishing applications. This technology is fast compared to other screen copy devices, and offers relatively high resolution (300 dots per inch is common). Page-size laser printers are best suited for black & white copies. Although several manufacturers recently have offered color models, the technology is costly; and it has not yet been accepted in the geographic information industry. Improved color capabilities, however, likely will make these devices more popular in the near future.

B.5.3. High-speed Line Printers

For high-volume printing of reports, tables, and other text documents, the line printer is favored because of its speed and relatively low cost. Line printers use various techniques to print documents line by line. They generally use two kinds of printing techniques that provide either fully formed characters or dot matrix characters. Printing speeds for line printers vary from 400 lines per minute to more than 2,000 lines per minute. Printers in the range of 600 to 1,200 lines per minute are typical in geographic information systems that require text output.

B.5.4. Large-format Pen Plotters

Pen plotters are capable of plotting a vector format file from a host processing unit using multiple line weights and colors. Most plotters sold today are capable of using plotting media on a role for continuous plotting with multiple pens for plotting line work and annotation. Pen plotters currently on the market vary in size of plotting area, ability to use cut sheets versus rolls, plotting speed, and a number of other special features. Models now available from various vendors combine features for ease of use and maintenance with the speed required for production work. Pen plotters produce high quality products, but are not well-suited to extremely high volume production, since they are slow compared to electrostatic plotters. Most pen plotters on the market suitable for GIS configurations have 4 to 8 pens which are controlled by the plotting software and commands imbedded in the plot file. These large-format plotters can generate output at up to 44" in width, at pen speeds of 24 to 30 inches per second, on various plotting media, including paper, vellum, or mylar.

B.5.5. Electrostatic Plotters

Electrostatic plotters connected to a host processing unit through a parallel or serial port will produce hard copy plots from a GIS database. In recent years, electrostatic plotters have become the workhorses for automated mapping systems for the high volume production of plots. The electrostatic plotter produces a plot in raster form in which the image is composed of a matrix of dots produced line by line as the plotting media on a drum feed moves through a liquid toner reservoir. Most electrostatic plotters are equipped with on-board rasterizers which convert an original vector map file from the system database into the raster format necessary for plotting. The raster format results in an image with somewhat lower quality line work than a pen plotter. Major advancements and price reductions over the past two years, however, have made high resolution units available at a reasonable cost. Common resolutions for electrostatic plotters are 200 and 400 dots per inch. The 400 dot per inch units are acceptable for most finished plotting work.

Electrostatic plotters are available in many sizes and in color or black and white. Most units accept multiple plotting media, including opaque paper, translucent bond, and mylar. The plotters are available in a variety of sizes but, in GIS configurations, those with a 24" or 36" width are most common. Both color and black and white units are available. Most color models employ a process requiring four passes of the plotting media through the toner reservoirs. Color electrostatic plotters have the capability to generate a wide range of colors for plots that are suitable for presentation purposes.

B.5.6. Large-format Laser Printers

Large-format laser printers now have a resolution of 400 dots per inch. Laser printers have not yet been used extensively in geographic information systems, but their versatility, high-quality graphics, and expected price decreases make it likely they will see increasing use. They can be used for raster plotting of maps in black & white and provide somewhat better quality than electrostatic plotters but, at this time, they are more expensive than electrostatic models and do not offer color output.

B.5.7. Screen Copy Devices

These devices connect to a graphic workstation and are used to produce a hard copy of the image displayed on the graphic monitor. To produce the plot, an operator enters a simple command from the keyboard. These devices usually use electrostatic, ink jet, or thermal technology to produce a page-size or slightly larger hard copy image. Recently, devices have come on the market using newer color output technologies, including color laser and dye sublimation. Inexpensive impact dot matrix printers or laser printers may be used as well, although each of these devices has major limitations on production of color output. Models are available from a variety of vendors that produce color or black & white copies. Most GIS vendors can support multiple models of screen copy devices. Resolution generally ranges from 150 dots per inch to 300 dots per inch. Depending on the device and vendor, both black and white or color products can be generated. These devices provide a quick way of producing a text or graphic image which may be suitable for final presentation purposes or as a proof plot during digitizing and editing.

B.5.8. Optical Scanners

Optical scanners digitally encode information from hard copy maps and documents by optically scanning an image consisting of line work, text, and symbols. The scanner senses variations in reflected light from the surface of the document. Any marks on the hard copy, including stains or wrinkles, that exceed a given threshold of contrast between the mark and the background media will be encoded as "information." Scanners on the market today that accept large-format maps and drawings use one of three types of approaches: flatbed, drum, or pass-through.

The digital file that results from the scanning process is in raster (grid) format, and does not generally differentiate between different types of features, annotation, or symbols appearing on the original hard copy map. The raster image, a non-intelligent picture, cannot identify individual features uniquely or link them to nongraphic files. Most scanning systems come with special software that converts these raster images to a vector format with line segments individually identified. In this format, nongraphic attributes can be linked in a post-scan editing session.

Small-format optical scanners suited to page-size documents are becoming important parts of data management systems. Most GIS software packages have the capability to index raster documents (such as field sketches, minerals exploration permits) and retrieve them based on geographic queries. These desktop scanners are available in a wide variety of models with different levels of capabilities for discriminating graytones in monochrome images or colors.

B.6. COMMUNICATION HARDWARE

Hardware devices in this category play various roles in facilitating data communications in a GIS network. The devices described below serve to connect devices in local or remote networks and efficiently manage data flow among processing units and peripheral devices in a GIS configuration.

B.6.1. Modems

Modem is the accepted acronym for "MOdulator-DEModulator," a device used to convert analog signals from a communication carrier to digital signals that a computer can interpret. The analog signal typical of phone communications transmits voice data as an electrical signal that continuously varies in frequency and amplitude. These analog communication media require modems to encode and decode the signals. In geographic information systems, modems are most commonly used to connect computers and peripheral devices through voice-grade phone lines.

Two categories of analog phone service are typically used to support computer communications -- dial-up lines and dedicated phone lines. A dial-up modem may be "direct connect" in which the phone jack is connected directly to the modem, or it may be used as an acoustic coupler in which the phone receiver is placed in a receptacle to send and receive audio signals that are converted by the modem to electrical signals. Dedicated phone lines set up point-to-point service between two locations, often using specially conditioned lines that have lower error rates than dial-up facilities.

Modems using analog phone lines can be categorized as:

- Low-speed modems, which include, primarily, dial-up modems that can transmit at speeds of up to 1,200 bits per second

- Medium-speed modems, which include dial-up or dedicated line modems transmitting data at speeds from 1,200 to 4,800 bits per second

- High-speed modems, which include dial-up or dedicated-line modems transmitting at speeds from 4,800 to 19,200 bits per second

- Limited-distance modems (LDMs), which are designed for transmitting data on dedicated circuits over short distances (usually within two miles, depending on line quality and speed). They are relatively inexpensive and can transmit data at rates of up to 1 million bits per second. Generally, as distance increases, the maximum speed will decrease.

B.6.2. Multiplexors

Multiplexors maximize the efficiency of a communication line by allowing multiple users to communicate on one high-speed link. A multiplexor splits a high-speed line into multiple channels supporting multiple devices. Multiplexors employed on remote lines can alleviate the need to set up multiple remote lines to support devices. This approach eliminates the multiple lease and modem costs that otherwise may be incurred. The multiplexor is placed between the computer hardware component (processor or peripheral device) and the modem, in a remote link.

Multiplexors can accommodate communication lines ranging in speed from 9,600 bits per second up to several million bits per second. Typically, maximum capacities of up to 56,000 bits are found in those used to communicate via phone lines. Multiplexors typically support from four to 24 separate channels. Special multiplexing equipment is used to support high-speed digital communication facilities such as DS-1 or DS-3 (T-1 or T-3) lines.

B.6.3. Digital Service Units

Digital service units can be considered analogous to modems because they are used to connect computer devices using an outside communication carrier. However, these units connect devices across digital lines where modems are not needed to convert analog signals. In many metropolitan areas, local phone companies, government organizations, or private companies have instituted digital service to support both voice and data transmission. Digital service units are used to allow communication between sites that are connected by digital lines such as digital dataphone service (DDS) or T-1 and T-3 lines that are often provided by local phone companies.

B.6.4. Device Servers

Device servers are hardware components used to connect multiple non-intelligent devices, such as host-dependent terminals, query workstations, plotters, and printers, to a local area network. Device servers are intelligent devices containing computer processors dedicated to supervising communications between the devices and the network. They translate signals from the devices using a protocol that is understood by the local area network and route messages from a processing unit on the network to the appropriate device. Any device that cannot perform communication processing functions on its own must be connected to a local area network through a device server.

TABLE B-1
CHARACTERISTICS AND PRICES OF COMMON GIS COMPUTER HARDWARE DEVICES

Hardware Device	Typical Characteristics*	Selected Manufacturers	Typical Price Range
Processing Units			
Mainframe Processing Unit	• High-end central processing unit for large organization • Supports proprietary operating systems (almost always) • Users supported: 50 to 500 • Main memory: 64 MB to 1 GB • Mass storage: 10 GB to 500 GB	Amdahl; Hitachi; Data Systems; Digital Equipment Corporation; IBM; Unisys	$1 to $8 million (price depends on model and options)
Large Minicomputer Processing Unit	• Used as mid-sized central processing unit or network server for one or more departments • Supports proprietary or UNIX operating systems • Users supported: 20 to 200 • Main memory: 96 MB to 512 MB • Mass storage: 2 GB to 50 GB	Data General; Digital Equipment Corporation; Hewlett-Packard; IBM	$50,000 to $300,000 (with minimal memory and mass storage)
Small Minicomputer Processing Unit	• Used as low-end central processing unit or network server in office environment for single department • Supports proprietary or UNIX operating systems • Users supported: 8 to 25 • Main memory: 16 MB to 96 MB • Mass storage: 1 GB to 10 GB	Data General; Digital Equipment Corporation; Hewlett-Packard; IBM	$18,000 to $50,000 (with minimal memory and mass storage)
Super-Microcomputer Server (Low-end)	• Used database server or compute server • Usually support versions of UNIX operating system • Users supported: up to approximately 25 users and peripheral devices when used as network server • Main memory: 16 MB to 64 MB • Mass storage: 500 MB to 3 GB	Data General; Digital Equipment Corporation; Hewlett-Packard; IBM; Intergraph; Silicon Graphics; Sun Microsystems	$8,000 to $30,000 (with about 32 MB main memory and 500 MB mass storage)

*MB - megabytes; GB - gigabytes

TABLE B-1 (continued)
CHARACTERISTICS AND PRICES OF COMMON GIS COMPUTER HARDWARE DEVICES

Hardware Device	Typical Characteristics*	Selected Manufacturers	Typical Price Range
Processing Units (continued)			
Super-Microcomputer Server (High-end)	• Used as network database server or compute server • Usually support versions of UNIX operating system • Users supported: up to approximately 50 users and peripheral devices when used as network server • Main memory: 16 MB to 256 MB • Mass storage: 500 MB to 30 GB	Data General; Digital Equipment Corporation; Hewlett-Packard; IBM; Intergraph; Silicon Graphics; Sun Microsystems	$40,000 to $90,000 (with about 64 MB main memory and 2 GB mass storage)
Microcomputer	• Single user workstation or PC network file server • User supported: 1 user and peripheral devices • Mainly DOS operating system environment using Intel 386, or 486, processor, but includes most Apple Macintosh models • Main memory: 2 MB to 16 MB • Mass storage: 40 MB to 250 MB[1]	Apple; AST; Dell; Compaq; IBM; Zenith	$3,000 to $18,000 (with math co-processor, 4 MB main memory, 80 MB hard drive
Workstations			
Full-performance Workstation (Low-end)	• Uses low-end super-microcomputer as processing platform • Graphics workstation with onboard processor and mass storage (500 MB to 2.5 GB) • Includes large-format digitizing tablet (36" x 48") • Single high-resolution, color monitor (e.g., 19", 1,280 x 1,024) • Used for data entry from source documents and for all geographic analysis and queries	Data General; Digital Equipment Corporation; Hewlett-Packard; IBM; Intergraph; Silicon Graphics; Sun Microsystems	$14,000 to $32,000 (with about 16 MB main memory, 500 MB mass storage and large-format digitizer)

*MB - megabytes; GB - gigabytes

TABLE B-1 (continued)
CHARACTERISTICS AND PRICES OF COMMON GIS COMPUTER HARDWARE DEVICES

Hardware Device	Typical Characteristics*	Selected Manufacturers	Typical Price Range

[1]Up to 1 GB or more when configured as network server with auxiliary storage drives.

*MB - megabytes; GB - gigabytes

TABLE B-1 (continued)
CHARACTERISTICS AND PRICES OF COMMON GIS COMPUTER HARDWARE DEVICES

Hardware Device	Typical Characteristics*	Selected Manufacturers	Typical Price Range
Workstations (continued)			
Full-performance Workstation (High-end)	• Uses low-end super-microcomputer as processing platform • Graphics workstation with onboard processor and mass storage (600 MB to 30 GB) • Includes large-format digitizing tablet (36" x 48") • Single high-resolution, color monitor (i.e., 19", 1,280 x 1,024) • Used for data entry from source documents and for all geographic analysis and queries	Data General; Digital Equipment Corporation; Hewlett-Packard; IBM; Intergraph; Silicon Graphics; Sun Microsystems	$35,000 to $90,000 (with about 32 MB main memory, 1.5 GB mass storage and large-format digitizer)
Edit/Query Workstation (Low-end)	• Uses low-end super-microcomputer as processing platform • Graphics workstation with onboard processor and mass storage (250 MB to 2.5 GB) • Includes small-format digitizing tablet • Single high-resolution, color monitor (i.e., 19", 1,280 x 1,024) • Used for graphic database update, database queries	Data General; Digital Equipment Corporation; Hewlett-Packard; IBM; Intergraph; Silicon Graphics; Sun Microsystems	$6,000 to $25,000 (with about 16 MB main memory, 500 MB mass storage)
Edit/Query Workstation (High-end)	• Uses high-end super-microcomputer as processing platform • Graphics workstation with onboard processor and mass storage (600 MB to 30 GB) • Includes small-format digitizing tablet • Single high-resolution, color monitor (i.e., 19", 1,280 x 1,024) • Used for graphic database update, database queries	Data General; Digital Equipment Corporation; Hewlett-Packard; IBM; Intergraph; Silicon Graphics; Sun Microsystems	$30,000 to $90,000 (with about 32 MB main memory, 1.5 MB mass storage)

*MB - megabytes; GB - gigabytes

TABLE B-1 (continued)
CHARACTERISTICS AND PRICES OF COMMON GIS COMPUTER HARDWARE DEVICES

Hardware Device	Typical Characteristics*	Selected Manufacturers	Typical Price Range
Workstations (continued)			
Query Station	• Graphics terminal dependent on host computer or server on a network • May have mouse or small bit pad for input • Could be PC running emulation software • Medium-resolution color monitor (640 x 480) • Used for graphic query and displays	Digital Equipment Corporation; Hewlett-Packard; IBM; Textronix; Wyse	$1,000 to $3,000
X-Terminal Query Workstation	• Query device in client-server network • Local intelligence for network access in window environment • Mouse as input device • 1 to 8 MB local memory • High-resolution color graphic display with 15" to 19" monitor	Hitachi; Data Systems; Digital Equipment Corporation; Hewlett-Packard; Tektronix; Wyse	$1,500 to $5,000
Peripheral Devices			
Desktop Text Printers	• Alphanumeric printing device for desktop or office environment • Dot matrix or ink jet printing technology • Print speed: 80 to 300 characters per second	C. Itoh; Digital Equipment Corporation; Epson; IBM; NEC; Okidata	$500 to $2,500
Small-Format Laser Printer	• Desktop printer using laser printing technology • Can support alphanumeric and graphic output • Print speed: 8 to 12 pages per minute • Most models just monochrome	Apple; Digital Equipment Corporation; Canon; Hewlett-Packard; IBM; NEC; Okidata: QMS; Seiko	$1,500 to $3,500
High-Speed Line Printer	• Alphanumeric print device for high volume printing • Use line, band, or matrix techniques • Large cabinet, floor standing device, computer room environment, office environment • Print speed: 400 to 2,000 lines per minute	C. Itoh; Data Products; NCR; QMS	$8,000 to $15,000

*MB - megabytes; GB - gigabytes

TABLE B-1 (continued)
CHARACTERISTICS AND PRICES OF COMMON GIS COMPUTER HARDWARE DEVICES

Hardware Device	Typical Characteristics*	Selected Manufacturers	Typical Price Range
Peripheral Devices (continued)			
Large-Format Pen Plotter	• Graphic output device for 36" to 44" wide drawings • Different models accept cut-sheet or roll plotting media • Typically has 4 to 8 pens • Plots at 24 to 30 inches per second • Connects directly to host processor or intelligent workstation	Calcomp; Hewlett-Packard	$7,000 to $11,000
Electrostatic Plotter a) Large-format (36" to 44") b) Medium-format (24")	• Graphic output device for A through E-size drawings • Uses electrostatic plotting technology - color or black & white • Host-dependent device, but may have local memory and processor for rasterizing vector drawings • Resolution: 200 to 400 dots per inch • Connects directly to host processor • Connects to LAN by device server	Calcomp; Xerox	$20,000 to $24,000* (black & white, 24") $25,000 to $34,000* (black & white, 36") $50,000 to $60,000* (color, 24") $55,000 to $70,000* (color, 36") *Price includes on-board rasterizer.
Large-Format Laser Printer	• Graphic output in black & white • 300 to 400 dots per inch • 36" width	Calcomp; Xerox	$70,000 to $80,000
Small-Format Screen Copy Device	• Graphic output device for duplicating images displayed on graphic workstation • Uses ink jet, thermal, or electrostatic printing technology • Produces color output up to 11" x 17" • Resolution: 150 to 300 dots per inch • Connects directly to graphic workstation	Calcomp; Hewlett-Packard; QMS; Seiko; Tektronix; Xerox; Optigraphics	$3,000 to $7,000

*MB - megabytes; GB - gigabytes

TABLE B-1 (continued)
CHARACTERISTICS AND PRICES OF COMMON GIS COMPUTER HARDWARE DEVICES

Hardware Device	Typical Characteristics*	Selected Manufacturers	Typical Price Range
Peripheral Devices (continued)			
Large-Format Optical Scanner	• Drum or pass-through scanner • Accepts up to 36" wide media • Variable resolution 300 to 1,000 dots per inch • Variable gray-tones up to 256 • Half-tone or gray-tone	Optigraphics	$12,000 to $40,000
Small-Format Optical Scanner	• Accepts 8 1/2" x 11" media • 200 to 400 dots per inch • 24 to 48 pages per minute • Half-tone or gray-tone	Bell and Howell; Eastman Kodak; Ricoh	$4,000 to $8,000

*MB - megabytes; GB - gigabytes

TABLE B-3
COMPARISONS OF PRINTING/PLOTTING TECHNOLOGY SPECIFICATIONS

Printing/ Plotting Technology	Specifications: Print/ Graphic Capability	Color Capability	Graphic Resolution (dots per inch)	Maximum Output Size	Production Speed	Media	Media Cost/Copy
Impact dot matrix	Text; low resolution graphics	Low (dependent on multiple ribbons)	Low to High (150 to 300)	24" width	Medium	Plain paper	Low
Laser	Both	Low (but improving)	High (300 to 400)	36" width	High	Plain paper	Low
Ink jet	Both	Medium (6 primary colors plus black)	Medium (150 to 200)	11" width	Low	Paper/ Transparencies*	Low to Moderate
Thermal wax transfer	Both	High; 7 to 16 million hues and shades (depends on model)	Medium (200 to 300)	24" width	Medium	Special paper/ transparencies	High
Thermal phase change (dye sublimation)	Both	High; up to 16 million shades	Very high (300**)	8 1/2” x 11”	Medium	Specially coated paper	Very High
Electrostatic	Both	High; 16 million hues and shades	High (400 to 500)	44" width	High	Special paper/ special mylar	High
Pen plotter	Both	Medium (dependent on number of pens)	Very high	44" width	Very low (depends on plot complexity)	Paper, mylar requires punched holes for plotter tracks	Medium

*Many models can now use plain paper.

**Excellent dithering capability gives very apparent resolution by blending colors.

APPENDIX C
GIS SOFTWARE CONCEPTS AND CAPABILITIES

C.1 INTRODUCTION

In modern computer systems, software can be depicted in a model composed of layers, as shown in Figure C-1. This layered model shows the system hardware in the center surrounded by three layers of software:

- Operating system
- Special system utilities and support programs
- Application software.

FIGURE C-1
SOFTWARE LAYERS

The outer layers rely on the inner software layers to perform a particular task for a user. The computer system provides the capability for each software layer to interact with the other layers. Although this model vastly simplifies the actual collection and relationship between software components, it helps clarify the interaction of software within a computer system.

C.1.1 Operating Systems

Operating systems include the programs that supervise and direct the fundamental operations of a computer system. The operating system provides the interface between the user, the application programs, and the computer system hardware. Although the concept of an operating system continues to evolve in the computer industry, and there is no universally accepted set of functions that an operating system should perform, major operating system functions that are characteristic of all types of computers, from microcomputers to large mainframes, include the following:

- **Memory Management** monitors the use of the system's main memory and communication with mass storage devices.

- **Access Security** allows for specific limitations to be placed on the type of functions that users may perform and the data which they have to access.

- **Communications Management** coordinates the function and operation of peripheral devices, including workstations, printers, and plotters.

- **Command Processing** provides a system language for users to invoke operating system commands

  Computer systems offer “command languages,” special sets of instructions through which users can execute operating system functions. Command processing programs are used to interpret these commands to perform a particular task. An example of a command language is the “job control language” (JCL) used to submit batch programs on mainframes and minicomputers.

In the past five years, there has been strong interest in the use of universal or **portable operating systems** (as opposed to vendor-specific operating systems) that can be used on different processing units, thereby freeing users and application programs from the limits of a specific computer system. Carried to its full extent, the concept of portability means that any application program, for example, a computer mapping package, could be loaded on another processing unit regardless of the make or model and could operate without program modification. This concept has not been realized fully in the computer industry, but there is a strong trend toward the development of portable operating systems.

One example of a generally accepted operating system “standard” is DOS used on a majority of microcomputers. This operating system, developed by IBM and Microsoft Corporation, has become a de facto standard in the microcomputer industry, and there are many desktop automated mapping and GIS packages that run in a DOS environment.

The UNIX operating system, originally developed at the University of California at Berkeley and at AT&T Bell Laboratories, has seen a dramatic increase in use in recent years. Originally developed for scientific applications on large processors, it is becoming increasingly popular on multi-purpose computers. Many computer system manufacturers have developed software for UNIX platforms and sometimes offer UNIX as an alternative to their proprietary operating system. Its attractiveness as a “portable” operating system undoubtedly will continue to fuel its rise in popularity.

C.1.2 System Utility Software

There is no precise definition of “system utility” software. Programs that are considered special system utilities in one system may fall under the operating system or application software category in other systems. Many utility programs are closely associated with the operating system and are provided by the computer vendor as part of the operating system package. Other utility or support programs are delivered with application packages, or are acquired from third-party vendors for installation on the system. Major utilities and support software used frequently in geographic information systems are described below.

- **Language Compilers** which translate the source code of an original programming language (e.g., Fortran, COBOL, BASIC, or C) to “machine code” which can be executed by the computer.

- **Device Drivers** which are programs that provide a communication interface to support a specific peripheral device (e.g., pen plotters, electrostatic plotters, and other graphic output devices).

- **Disk Backup Utilities** that provide support programs for the efficient tape backup of data and software on a system disk drive.

- **Subroutine Libraries** which are program modules written in any of a variety of programming languages, such as Fortran, and C, that can be accessed by users to help perform special tasks.

- **Special Communication Software** to support complex communication tasks like local area and wide area network management; gateways for communication between networks; and special software to help monitor network traffic and resolve network problems.

C.2 GIS APPLICATION SOFTWARE

Automated mapping and GIS application software is provided in the form of software “packages,” each consisting of multiple programs that are integrated to supply particular capabilities for mapping, management, and analysis of geographic data. Application software developed for geographic information systems can be conceptualized in two parts, as depicted in Figure C-2:

- A “core” package of basic mapping and data management capabilities

- Separate applications that are integrated with the core package to perform a specific mapping or geographic analysis operation and generate products.

FIGURE C-2
APPLICATION SOFTWARE CONCEPT

The specific types of capabilities provided as part of the core package or as special applications will vary among software vendors. In a general sense, however, the core package will include functions that fall into one or more of the following categories:

- Graphics processing
- Database management
- Basic tools and capabilities for cartographic operations and geographic analysis.

Packages offered by software vendors as full "geographic information systems" (GIS) normally include significant functionality in both graphic and nongraphic data management with strong links between the map (graphic) database and nongraphic attributes.

C.2.1 Graphics Processing Capabilities

Graphics processing capabilities in automated mapping and GIS packages include functions that allow the user to enter or edit map features and associated annotation, and to generate screen displays or hard copy maps. Graphic entry capabilities in GIS allow users to input map features and store these as x,y (and sometimes z) coordinates based on a geographic reference grid. Text annotation and feature identifiers also can be entered to define a map feature uniquely and provide a basis for associating the feature with tabular attributes stored in a database. GIS and automated mapping packages offer special capabilities for the editing of graphic features, including a variety of delete and modification functions. These functions allow an operator to update or modify graphics and annotation in an interactive environment. Also, automated mapping and GIS packages often have capabilities for line generalization, transformation of map projections and coordinate grids, and other cartographic functions.

Part of the graphics processing component of automated mapping and geographic information systems are capabilities to customize the display and hard copy production of maps on graphic monitors or plotters. Software capabilities of automated mapping systems or GISs allow an operator to control the appearance and format of the display or plot by setting parameters like line patterns and weights, symbol types, annotation text fonts, etc.

C.2.2 Tabular Database Management

The power of a GIS is its ability to query and analyze both graphic and tabular attribute data together. This concept of graphic/tabular data linkage provides powerful capabilities for mapping and geographic analysis. GIS software packages include capabilities to store and retrieve nongraphic attribute data associated with map features. For instance, a user may want to store information about road characteristics associated with particular road segments in a GIS database (e.g., road name, surface type, speed limits, etc.). Vendors of GISs store tabular attributes and link them with their associated map features to support map display and analysis. These GIS software packages use one of two approaches to manage the nongraphic database associated with the map features:

- Proprietary database management software
- Commercially available third-party database management software.

Regardless of the approach, nongraphic database software used in GIS has the following common components:

- A **Data Definition Language** (DDL) describes the characteristics or “schema” for files that will contain tabular attributes. The schema holds information such as the name of data elements; the size of element fields in bytes or columns; the data element format (e.g., alpha, integer, binary); and other data required by the software to process the attribute data.

- **Data Entry** facilities create a file defined by the data definition language and allow users to enter data according to system prompts. Most advanced database management systems provide the ability to design special screen formats for more efficient entry of nongraphic data. These entry format design programs also allow for logical error checking during the entry operation to flag or prevent all entry that violates established validity checks.

- **A Data Manipulation Language** (DML) queries the GIS database. Data query languages generate searches using commands that operate on elements of the database. Ideally, the GIS package should provide a query capability that operates in a unified manner on the graphic and tabular data to generate reports, screen displays, or hard copy maps in response to the user's commands.

In addition to the basic characteristics described above, some database management systems include other features that increase their flexibility. These include capabilities for complex report design, using high-level programming languages, sometimes called fourth-generation languages (4GL), that can be used to develop applications.

C.2.3 Basic Cartographic and Geographic Analysis Capabilities

Most GIS packages come with a set of capabilities for performing routine mapping and geographic analysis. These utilities can be considered a "toolbox" of programs that are invoked by simple commands that operate on graphic and nongraphic database elements. These tools combine aspects of the graphics processing and tabular database management environments discussed previously to give users the ability to build complex applications. These basic capabilities represent a range of tools that are considered important in a fully integrated geographic information system. They may be grouped into the following categories:

- Data Entry and Editing
- Data Query, Manipulation, and Analysis
- Data Display and Output
- Data Format Translation and Conversion.

The particular set of basic cartographic and analysis utility programs included as part of the core software package will, of course, vary among commercial software vendors. They include such standard cartographic functions as map overlay, map windowing, area/distance calculation, buffer analysis, and other basic mapping and analysis operations. Several critical GIS functions are described below with examples for the types of applications they might support.

C.3 THE GIS TOOL BOX CONCEPT

The tool box concept is based on the assumption that users need a flexible software environment with which they can design and customize applications (see Section 1) that suit their particular needs. GIS applications all make use of certain basic functions for mapping, data query, and spatial analyses that operate upon components for the GIS database (map features and their tabular attributes) to generate products. The GIS tool box approach, conversely, provides users with the basic geographic functions and efficient programming tools to develop custom applications to address their specific requirements. Some of the fundamental GIS software tools are described below.

C.3.1 Standard Map Update and Production

This tool includes a number of functions for entering, maintaining, displaying, and plotting standard map sets that must be updated on a regular basis. Standard map sets may depict land parcels, utility networks, administrative boundaries, or other map information that changes on a regular basis as land development or redevelopment occurs. It is common for most GIS packages today to incorporate an extensive set of functions for interactive digitizing and edit, tabular data entry and edit, map design, and plotting to generate hard copy products.

C.3.2 Geographic Query and Report Generation

Geographic query and reporting tools include a range of functions that are fundamental to the database management component of the GIS. Query capabilities let users ask geographically oriented questions and receive meaningful responses. Typically, users will interact with the GIS at a workstation to answer two fundamentally different types of questions. The first involves simply pointing to a map feature to display pertinent attribute data. This type of query could be used to quickly determine the owner and characteristics of a specific land parcel or the material and diameter of a specific water pipe segment. The other major type of query establishes search criteria and asks the GIS to identify all map features that satisfy that criteria. For example, a utility engineer could ask the GIS to find and identify all gas mains of a certain pressure rating that were originally installed before 1962, and receive input in the form of a map, graphic display, or report. Both of these types of queries can be applied to all database components of the GIS making use of the strong linkage between map features and tabular attributes to generate map displays and textual reports.

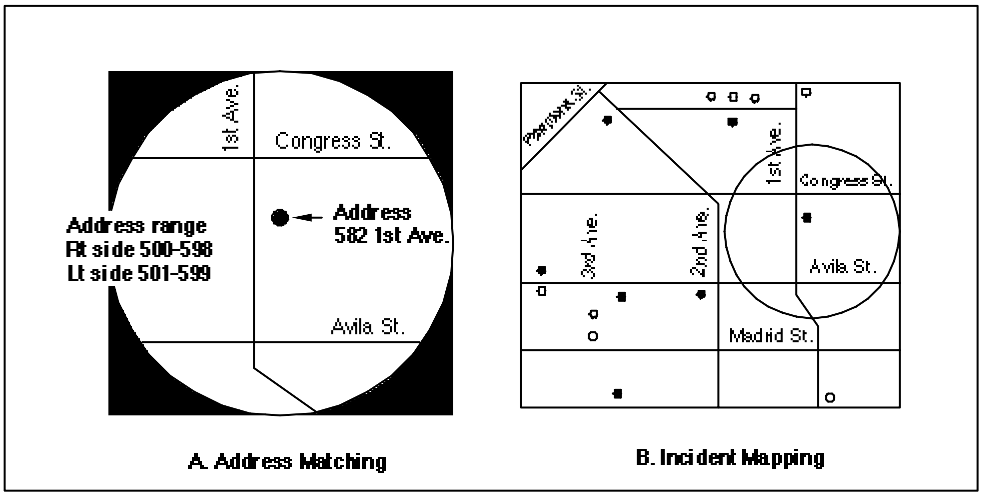
C.3.3 Distance and Area Calculation

These capabilities are usually included as part of the core GIS software for the calculation of areas and distances. These functions may operate interactively, providing immediate calculations from user input at a graphics workstation or in a batch process for an existing set of map features in the GIS database. A simple example of a batch process would be the calculation of areas of land parcels and the generation of a textual table listing those area figures. Users have the flexibility to output distance and area figures in any desired unit of measure.

C.3.4 Address Matching and Incident Mapping

Because of the large amount of information in an urban environment that is geographically defined by a street address, many vendors of GIS software provide tools to map these address-related features or occurrences relative to a street centerline map. The digital centerline map itself portrays the path of all public streets with address ranges assigned to each street segment normally determined by a block. The address information for the street centerline is stored as tabular attributes and indicates the range of address numbers (from address, to address) for each side of the street. The address matching capability reads specific addresses from a tabular database (e.g., building permit file, emergency response file) and automatically places a point along the street segment by interpolating a location relative to the address range of the street segment. Figure C-3 illustrates the concept of address matching and incident mapping. Incident mapping tools can allow the user to define specific point symbols or colors determined by attributes of the address file. For instance, different symbols could be created for type or status of building permits.

FIGURE C-3
ADDRESS MATCHING AND INCIDENT MAPPING

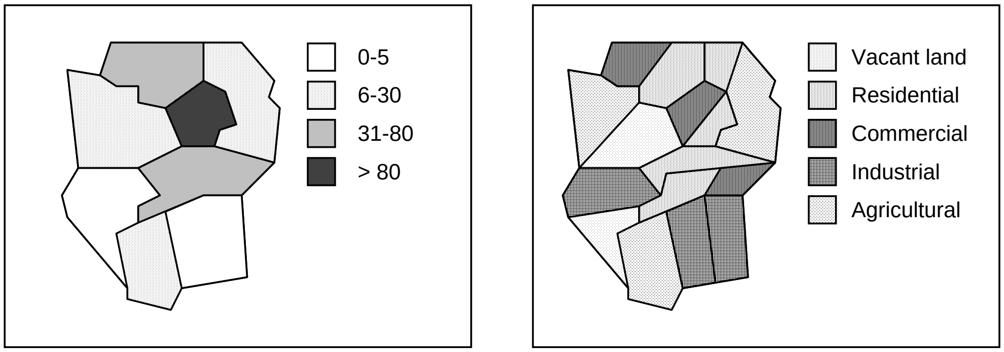


C.3.5 Thematic Mapping

Thematic mapping is the creation of custom maps designed to portray particular topics or themes. This broad definition encompasses many types of maps but, in all cases, a thematic map symbolizes or shades map features based on tabular data linked to those features. Thematic mapping techniques may be applied to polygon features as shown in Figure C-4. Figure C-4a represents a type of thematic map known as a “choropleth map,” which organizes quantitative data (in this case, population density) into specific class ranges and then assigns a shading pattern to geographic areas based on their value. In
Figure C-4b, qualitative data representing land use is being mapped in the same way. In each case, polygons are being shaded based on attribute data values stored in a tabular database. Thematic maps can also be produced for linear map features as well. For

example, the GIS could be used to depict traffic volumes for a road network or pressure ratings for segments in a gas or water system.

FIGURE C-4
THEMATIC MAPPING

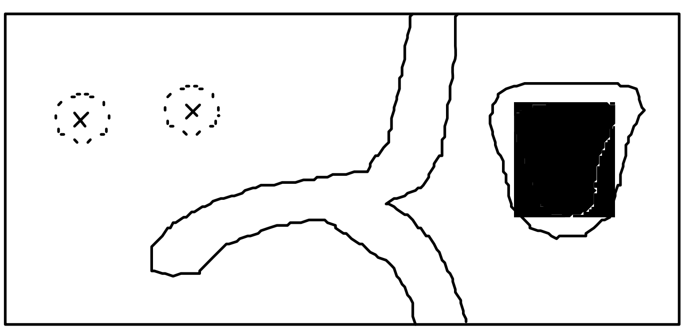


(a) Population density thematic map (b) Land use thematic map

C.3.6 Proximity/Buffer Analysis

These tools are used to search or to generate polygons and select map data within a specific distance from a point, line, or polygon feature as illustrated in Figure C-5. This capability is used in many urban applications that require ownership searches; environmental impact assessments; and other similar applications. This capability is very useful for local governments or regional planning agencies that routinely evaluate land use changes and must locate and notify all land owners within a specified distance of a parcel.

FIGURE C-5
BUFFER GENERATION



C.3.7 Polygon Overlay

Polygon overlay tools vertically combine two or more polygon layers of a GIS database to produce a resulting layer with new polygon boundaries and associated tabular attribute data (see Figure C-6). Polygon overlay tools of many GIS packages allow users to perform modeling to produce the resulting layer based on “rules” established for the combination of the initial layers being overlaid. This polygon overlay modeling concept is being applied to more efficiently evaluate restrictions on land use development projects such as large housing or industrial projects or in evaluating optimal sites for parks or landfills. Users can overlay parcels with other layers depicting current land use zoning designations, flood hazards, environmentally sensitive areas, utility lines, and other layers to review acceptability for development.

FIGURE C-6
POLYGON OVERLAY

C.3.8 Spatial Aggregation

For many mapping and geographic analysis applications, it is useful to generalize detailed spatial information and present it in a form which is easier to interpret and more suitable for evaluating broad conditions or trends over specific geographic areas. This is the purpose of spatial aggregation tools in GIS packages illustrated in Figure C-7. These tools generally use polygon overlay capabilities to tabulate frequencies or averages of map features or incidents over specific polygons which serve as reporting districts. Products generated from an aggregation process may be maps or hard copy reports summarizing statistics. Aggregation tools may be used to summarize land use data collected by parcel over specific planning districts; to tabulate numbers of facilities (poles, transformers, etc.) within maintenance areas; or to evaluate the frequency of crimes within police reporting districts.

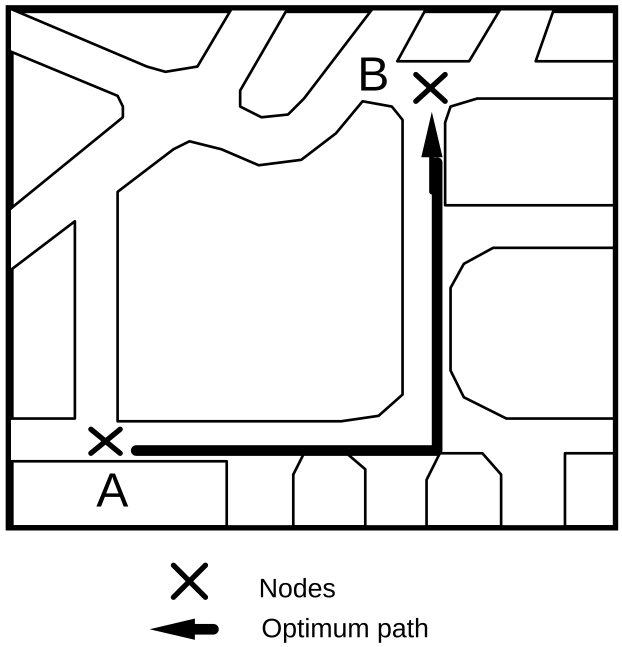
FIGURE C-7
SPATIAL AGGREGATION

C.3.9 Network Analysis

Network analysis tools evaluate the path, flow, or pressure in a linear network. Spatial networks that can be analyzed may include road systems; gas, water, or electric distribution networks; sewer collection systems; or drainage networks. Figure C-8 illustrates an "optimal path" analysis using network analysis tools to find the quickest path between two points based on street distance and attributes assigned to street segments. GISs that provide efficient network analysis tools normally use topologic data structures with which true network models are defined. These topologically structured databases explicitly define the spatial connectivity between all segments of the network and point features (nodes) that represent connection points between segments or important point features on the network.

For example, network analysis tools could be used to evaluate pressures and flow in water mains from a network model that included the water main segments, valves, pump stations, and service connection points.

FIGURE C-8
NETWORK ANALYSIS



C.3.10 District Delineation

The purpose of district delineation is to define the optimal or most efficient boundaries for providing a municipal service or administering a program. Defining districts in a manual environment is complex because it requires simultaneous evaluation and comparison of many geographic variables such as population, economic conditions, transportation facilities, and many physical site conditions. GIS tools can be applied to simplify what can be a very time-consuming task requiring multiple interactions to re-orient district boundaries. In the United States, property tax assessment offices are beginning to use GIS capabilities to define “tax neighborhoods” that exhibit a homogeneous set of site conditions and can be used to help establish fair market values for taxation purposes. Districting functions can also be used to help define optimal areas for allocation of emergency services (e.g., police and fire station areas), administration of public assistance programs, and other municipal service functions.

C.3.11 Terrain Analysis

Some GIS packages incorporate capabilities for mapping and analyzing three-dimensional services. Most often, these tools are applied to a physical topographic surface, however, they could be used with statistical “surfaces” to portray demographic statistics that can be modeled in the form of a continuous surface. Terrain analysis uses data layers called digital elevation models (DEMs) that, in their raw form, consist of an array of points each of which has x, y geographic coordinates and a z coordinate for elevation. With the proper terrain analysis tools, these DEMs can be used to produce traditional contour maps, three dimensional displays and, for more complex studies, to evaluate slope, erosion, and viewsheds, and to make engineering calculations.

A more inclusive set of GIS functions is presented in Table C-1.

TABLE C-1
BASIC GIS FUNCTIONS

Data Entry and Editing

Interactive Digitizing/Graphic Editing

Parametric Feature Entry and Edit

Heads-up Digitizing

Coordinate Geometry (COGO) and Precision Entry

Image/Document Scanning

Annotation Entry and Placement

Data Query, Manipulation, and Analysis

Map/Control Point Adjustment (rubber sheeting)

Map Projection Transformation

Coordinate System Transformation

Line Smoothing/Generalization

Area/Perimeter/Distance Calculation

Vector Map Overlay

Vector/Raster Overlay

Grid (Raster) Overlay

Polygon Overlay

Graphic Data Query

Tabular Data Query/Analysis

Data Query, Manipulation, and Analysis (continued)

Spatial Aggregation

Statistical Analysis

Buffer Generation

Thematic Mapping

Address Matching/Incident Mapping

Network Analysis

Terrain and Three-Dimensional Analysis

Area Districting

Data Display and Output

Tabular Display

Graphic Display

Hard Copy Report Production

Hard Copy Map Production

Graph/Chart Production

Data Format Translation and Conversion

Vector Map Data Translation

Raster/Vector Conversion

C.4. APPLICATION DEVELOPMENT

C.4.1. Application Development Capabilities

As discussed earlier, most GIS applications will generally not come "pre-packaged" with the GIS software. It is most often a requirement, therefore, to carry out some level of application development and customization. The "toolbox" concept of a GIS implies that there must be efficient procedures available for users to build applications that are customized to meet their needs. Application development in a GIS environment has the following objectives:

- To simplify a complex set of basic GIS commands into simple keyboard or menu commands
- To make user interfaces more efficient to reduce time and complexity in running applications
- To implement special applications that require complex analysis and processing
- To facilitate the translation and exchange of data between different platforms
- To ensure consistency, integrity, and efficiency in the way data is maintained and products are formatted.

Application development procedures, therefore, help to customize and tailor the user environment to maximize efficiency and consistency. In selecting a GIS package, a key criterion should be the range and flexibility of application development tools provided. Major categories in the application development process are described below.

C.4.1.1 Customizing Data Entry Procedures and User Interface

Building of Macro Programs

Macro programs are logical sequences of system commands and high-level programming statements that collectively perform complex functions through selection of menu commands or with a limited number of keystrokes.

Menu Design

Most GIS software packages now allow the user to access complex commands, macro programs, and digitizing commands via tablet or screen menus and windows. Menus are used by pointing to the screen or tablet using the cursor, mouse, or joystick. Special applications will often require application developers to design and program custom menus.

Data Entry Form Design

Most map or database update applications require the entry of nongraphic attribute data using a formatted screen. These forms will control cursor movement to optimize the speed of the data entry process and may also have some "intelligence" in the form of programmed logical error checks to flag incorrect or suspect entries.

Data Translation and Exchange

The use of GIS data on multiple platforms will require special software to translate and exchange data using programs provided by the GIS vendor. There are a number of "standard" formats for the interchange of digital maps (e.g., IGES, ISIF, DXF) in North America, but no universally accepted exchange format for GIS data has yet emerged. Over the last three years, however, technical work groups in the U.S. have been working on the development of exchange formats that will eventually become standard in the industry.

C.4.1.2 Customizing Output and Product Generation

Report Form or Screen Display Design

Many applications will require formatted displays of tabular data from attribute files, either on a workstation monitor or in a hard-copy form. Database management capabilities offered by vendors should allow users to design and implement these forms for the generation of text outputs. For example, popular database management packages integrated with GIS core capabilities such as INFO, ORACLE, and DB2 provide tools for form design.

Development of Map Plotting and Display Criteria

All applications that will generate graphic output as screen displays or as hard copy maps must have specific display and map design criteria established. This includes conventions for line and point symbology, the location of the map title and legend, and the map text font. Many of the point and line symbols required will be available in the GIS package, in the form of line type and symbol tables available with the GDS software. Other special symbols may need to be created.

C.4.1.3 Special Application Programming

Procedural Programming

For some applications, special programs using what traditionally have been called "third generation" languages may be coded by in-house staff or acquired from an outside source. These programs may be used to perform a variety of operations, and may be coded in COBOL, Fortran, C, BASIC, or some other procedural programming language.

High-level Programming

Some applications may be developed by in-house staff or from an outside source using high-level programs based on fourth generation languages. These languages, most of which are proprietary products offered by particular vendors, use powerful, non-procedural commands to perform complex operations.

C.4.1.4 CASE Tools

Over the past five years, there has been considerable maturing of software tools falling under the category of CASE which stands for "computer aided software engineering" or "computer aided systems engineering." There is no universally accepted definition of CASE, but the software packages that exist combine a number of capabilities to aid the user in such tasks as database design, generating application programs, and preparing documentation. CASE packages have some common attributes:

- Use of graphic user interfaces to enter commands and manipulate data
- Tools for automatic flow-charting and entity-relationship diagramming
- Fourth generation languages to generate application programs
- Automatic documentation tools for application programs.

There is currently no wide use of CASE tools for application development in GIS, although a number of GIS software vendors have indicated interest in building CASE tools appropriate for GIS. Several third-party software vendors have developed rudimentary CASE packages to work with specific GIS packages.

C.5. GIS SOFTWARE VENDORS

C.5.1 Available Packages and Capabilities

At the current time, there are over 25 vendors who offer packages for CAD and GIS running on microcomputers, super-microcomputers, minicomputers, and mainframes. A contact listing of these vendors is provided in Table C-2. Some of the basic capabilities of these packages are compared in Table C-2. Users evaluating software packages should be aware that commercial packages are continually being upgraded and enhanced by vendors. It is, therefore, important to evaluate packages based on current capabilities or near-future software releases to make a valid comparison between commercial offerings. Table C-2 provides a starting point for users in this evaluation process.

Insert Table C-2

APPENDIX D
DATA COMMUNICATION CONCEPTS AND SYSTEM CONFIGURATIONS

D.1 CONCEPTS

A computer network may include devices that occupy a single room or are distributed across a nation. A single-user microcomputer GIS may consist of the microcomputer, disk drive, color graphics monitor, a pen plotter, and a printer in close proximity, all connected by various types of cables. More complex systems may extend cabling to connect multiple devices and perhaps multiple processing units throughout a building or group of buildings. Remote communication technologies using the telephone system, microwave transmission, or other communication techniques' satellites are able to connect devices over very large areas.

To provide a basis for understanding GIS configurations, this appendix describes major concepts on which communication networks operate.

D.2 ANALOG AND DIGITAL COMMUNICATION

Two basic modes of electronic communication are used to transmit data in computer networks. The analog method transmits a signal that may vary continuously in frequency or intensity within a given interval. Digital transmission, on the other hand, transmits data as discrete impulses in two states that represent the binary digits 0 and 1. Both digital and analog communication techniques are used to transmit data in computer networks. Analog communications are used primarily where distances preclude a direct cable connection between computer devices. Since computers require digital signals to store data and execute instructions, modems must be used to convert between analog and digital signals in cases where analog techniques, such as the telephone system or microwave, are employed to carry data.

D.3 SPEED AND DISTANCE PARAMETERS

All computer networks have limitations on the speed at which data can be transferred and the distance that may separate devices. The speed and distance limitations are, in part, determined by the type of transmission media used and by environmental conditions that may affect the strength and quality of a signal. RS-232 cables, for instance, may be used reliably to send and receive data at distances only up to about 50 meters.

Speed, as it is applied to data communications, is normally expressed as "bits per second" (bps). This measure indicates the volume of data that can be transferred within a given time period. A measure of speed often used in discussions of analog transmission is "baud" or one "signal unit" per second. Baud has often been incorrectly used synonymously with "bits per second." With some analog transmission techniques, the baud rate may be equal to bits per second, but special modulation techniques are used frequently that deliver a bits-per-second rate higher than the baud rate.

Different devices and geographic data management applications have varying speed and distance requirements. For instance, a pen plotter, since it plots a map relatively slowly, has modest speed demands that may be met by connection to a host computer through an RS-232 cable or even a dial-up phone line using a 1,200 bits per second modem. Connections to support interactive data access between multiple computer processing units, on the other hand, might require the higher speeds of a local area network. It is necessary to evaluate, closely, the specific needs of the system to choose the optimal means of communication.

D.4 LOCAL COMMUNICATIONS

Local communication implies a direct cable connection between devices in a computer network which allows data to be transmitted in digital form. Local communications may support devices within the same room, an office building, or a collection of buildings within close proximity. Generally, it becomes less technically and/or economically feasible to use local communication techniques when devices are separated by greater distances.

A variety of cabling types is used to support local data communications, each of which has its own speed and distance limitations. Table D-1 summarizes common types of cabling used in local computer networks and general speed and distance guidelines. In addition to direct physical cabling, wireless LANs using radio transmission, laser, or infrared beams are gaining in popularity.

Two scenarios for local communications must be considered:

- Direct connection to host processing unit

- High-speed local area networks.

TABLE D-1
COMPARISON OF SPEED AND DISTANCE LIMITATIONS OF CABLING SCHEMES FOR DIGITAL TRANSMISSION[1]

Cable Type	Maximum Transmission Speed[2]	Maximum Length between Devices[3]
Twisted pair copper wire	4 million bps[4]	500 meters
RS-232 cable	19,200 bps	50 meters
Standard coaxial cable	50 million bps[5]	500 meters
Thin-wire coaxial cable	20 million bps	300 meters
Multi-mode fiber optic cable	1 gigabyte per second	10 kilometers

[1]Figures are dependent on site conditions and the specific communication hardware and software used.
[2]Speed in bits per second (bps). Speed depends partly upon the total length of cable. These speeds may be exceeded when the network extends over short distances.
[3]Maximum expected distances without repeaters.
[4]Capacity can be extended to over 10 million bps within limited distances.
[5]Higher speeds up to approximately 100 million bps can be achieved over short distances (less than 100 feet).

In the first scenario, non-intelligent terminals or peripheral devices (e.g., pen plotters) are connected to a central host processing unit (or to an intermediate communication controller) for direct communications. In this scenario, the non-intelligent device relies on the central host for all or most processing and data storage. The user receives output from transactions with the host in the form of screen displays or hard copy reports or maps. The line speed (between the host and peripheral device) needed to support this type of communication will vary depending on the type of peripheral device and application. Alphanumeric terminals may be very effective at 1200 bps or even lower speeds, while graphic terminals usually require about 9600 bps or more to operate with adequate response time.

The local area network (LAN) is a special type of network designed to support multiple devices within close proximity on a relatively high speed network. A local area network may include multiple processing units and peripheral devices of various types (workstations, servers, microcomputers, printers, plotters). LAN applications are dependent upon the movement of high volumes of data and therefore support raw data transmission speeds of 4 million bps to 16 million bps. Standard protocols and products are now being offered to support LAN communication at 100 million bps. The Fiber Distributed Data Interface (FDDI) is a communications protocol approved by the American National Standards Institute (ANSI) and the International Standards Organization (ISO) to support high-speed local area networks or connection of distributed LANs at this 100 megabit per second speed. While it was designed for use with fiber optic cables over limited geographic regions, coaxial cable and twisted pair wiring can also be used to support FDDI over short distances inside buildings.

D.5 REMOTE COMMUNICATIONS

Remote communication techniques are used when technical, environmental, or cost factors make the use of local communication connections technically and/or economically infeasible. Depending on the particular system, remote devices on multiple LANs could be in the same building or distributed throughout a large region. Remote communications often use the facilities of a third-party communication carrier, such as the telephone company.

D.5.1 Dial-up Telephone Lines

Dial-up voice grade lines may be used to carry data via analog signals. Therefore, modems are required to support the connection of computer devices. Due to the frequency of the telephone channel, dial-up telephone lines are limited with respect to the amount of data they can carry. Theoretically, dial-up lines cannot support a transmission rate of greater than 30,000 bits per second. In reality, the noise impairments on the line put a practical upper limit on the data transmission rate that is between 10,000 and 15,000 bps.

D.5.2 Analog Dedicated Leased Lines

A more effective wide area network solution to a dial-up circuit is to establish a fixed transmission path. This type of path is known as a leased or private line and is based on a point-to-point telephone line that terminates at the two sites being linked. Since the path is fixed, electrical adjustments known as conditioning can be made to the line to correct noise and other anomalies thereby creating a higher quality connection. With the negative effects of line noise minimized, transmission speeds and data reliability can improve to provide communication at speeds up to 19,200 bps and sometimes greater. Analog leased lines also require a modem on each of the links.

D.5.3 Digital Transmission Options

Despite line conditioning, analog leased lines are also susceptible to a level of noise impairment and signal attenuation. Digital leased line facilities transmit digital pulses, which are not as susceptible to line impairments, so channel performance is improved. Digital line services are categorized into speed classes as follows:

Class	Speed
Sub-rate Speeds	Generally 1,200; 2,400; 4,800; 9,600; and 19,200 bps
DS0	56,000 or 64,000 bps
DS1 (T1)	1.544 million bps
DS1C	3.152 million bps
DS2	6.312 million bps
DS3 (T3)	44.736 million bps

DS4	274 million bps

These digital transmission services do not require modems for communication between devices on a network, but communication devices known as channel service units (CSUs) and digital service units (DSUs) are required for connection of devices.

D.5.4 Other Communication Options

Many other media are available for remote data transmission, each of which has special advantages and limitations in specific situations. Some of the more popular remote communications facilities used to transmit data include microwave systems, broadband cable networks (Cable TV), radio transceivers, infrared systems, and satellite systems. Table D-2 compares the characteristics of remote data communication media.

TABLE D-2
COMPARISON OF REMOTE COMMUNICATION ENVIRONMENTS[1]

Method	Normal Transmission Speeds[2]	Distance Limitation between Devices
Dial-up phone line	300 bps to 9,600 bps	Theoretically unlimited where telephone service is in place
Dedicated analog phone line	300 bps to 19,200 bps	Theoretically unlimited where telephone service is in place
Digital phone service	1,200 bps to more than 44 million bps employing multiple channels	Normally within metropolitan areas or specified service regions
Broadband analog	Up to about 50 million bps	Up to about 50 km transmission on coaxial cable
Microwave	-	Up to about 50 km depending upon signal frequency[3]
Radio	Up to 9,600 bps	Theoretically up to about 75 km[4]
Laser or Infrared beams	Up to 19,200 bps	Up to about 1.5 km[3]

[1]Figures are very dependent upon environmental conditions, and the specific communications hardware and software used.

[2]Speed in bits per second (bps).

[3]Dependent on atmospheric conditions and requires direct line-of-site path between transmitter and receiver.

[4]Normally used within a 15 km range.

D.5.5 Public Network Value-added Options

Telephone companies typically offer a suite of services to facilitate communications for subscribers. These options provide support for routing of communications among multiple sites, special error detection/correction, and integrated communications. Some of these special offerings are discussed below:

- X.25 Packet Switching: This switching technology uses internationally standardized protocols to send messages from source locations to multiple destinations which have been assigned addresses. Messages are sent inside a data packet which also contains information bits defining a unique address along with other control information for error detection and correction. The phone company maintains the directory of addresses and performs all needed error correction. The X.25 systems route packets dynamically through the most efficient route. X.25 networks can support speeds of up to about 19,200 bps.

- Fractional and Multi-drop Digital Lines: High-speed digital lines (DS0 and greater) traditionally have been offered only in standards increments and on a point-to-point basis. Some phone companies, in specific regions, have begun to offer more flexible services such as Fractional Lines in which a user may lease a portion of the bandwidth on a digital line or establish multiple termination points of service.

- Frame Relay: In North America, public carriers are beginning to offer Frame Relay services which is a technology similar to X.25, but provides higher speed transmission. Frame Relay does not rely on the public network to perform error checking and therefore reduces the overhead of this task. Like X.25, Frame Relay services keeps a directory of addresses and efficiently routes messages over the best route. Frame Relay can be used to transfer data over large distances at speeds of up to DS1 (T-1) (about 1.5 million bits per second). Frame Relay networks adhere to standard protocols defined by the IEEE.

- Switched Multi-megabit Data Service (SMDS): SMDS is a new offering of some phone companies to data networking within defined areas of service (usually restricted metropolitan areas). It is based on the use of fiber optic backbones extending through the service area, and it can support DS1 to DS3 speeds (from about 1.5 million bits per second to 45 million bits per second).

- Cellular Telephone: In most areas of the country, phone companies are setting up cellular phone networks to support voice, and potentially data, communications from mobile and portable telephones. These telephones receive and transmit radio signals. The cellular system establishes a series of geographically spaced cellular telephone switching centers (or "cells") which house radio transmitters and receivers. These centers are the links between the radio signals to and from the mobile or portable phones and the public telephone network. While cellular phone networks are not normally used at this time for carrying data, the possibility exists to use these systems for data communications.

D.6 MODES OF INFORMATION EXCHANGE

Most complex computer networks connect multiple devices capable of performing local processing. The existence of multiple processing units, whether locally or remotely connected, often creates a requirement for the exchange of data between these processing units. For instance, in a typical local area network, multiple servers, microcomputers, or workstations may be sharing common data but carrying out processing tasks independently. In more complex environments, it may be necessary to exchange information between dissimilar systems such as a case in which data must be downloaded from a mainframe to a workstation or microcomputer. There are different methods, which vary in complexity, that can be employed to enable this information exchange. The particular approach used depends on the needs of the application (specifically response time) and the technical complexities in implementing it.

Typically, information may be exchanged between computer processing units or networks in one of the following modes:

- Media Transfer: Data may be exchanged through the transfer of magnetic tapes or other storage media. This method is the least sophisticated linkage and requires that the source and target systems are compatible with respect to tape or storage media formats. This method results in the duplicated storage of the information on the source and target systems. It is appropriate where some time delay in the exchange will not negatively impact the applications using the data. It has the advantage of being inexpensive and technically simple with the ability to transfer large volumes of data.

- Batch File Transfer: Systems may be linked directly to facilitate batch file transfers using any of the local or remote communication facilities described above. In this scenario, a communication network with appropriate software to support file transfers is created between the two systems. This data transfer method will also create a duplicate copy of data files on both the source and target systems.

- Interactive Data Communications: Systems may be linked directly using high-speed communications to facilitate interactive communication. In this scenario, programs or processes operate on different systems interactively in a peer-to-peer mode. It is important to note that database records rather than entire files are being transferred as they are required by the programs being executed. The GIS process is distributed between the GIS itself and another peer system. This data transfer is the most sophisticated and requires the greatest level of application development, complex network management software, and high-speed communication lines typical of local area networks.

APPENDIX E
STANDARDS AND OPEN SYSTEM COMPUTING

E.1 INTRODUCTION

The current trend in computer system standards is toward "open architecture" and the definition of standards that run on multiple platforms. These standards are hardware and operating system-independent and are used to integrate computing systems. With an open architecture approach, "strategic standards" will be defined for specific system layers. These standards would be platform-independent and will facilitate communication between different system architectures without identical layer implementations on each system.

One of the key tasks during GIS implementation will be the definition and adoption of information system standards by organizations. This section provides a tutorial on the background, types of standards, status of and directions in standards development, and organizations active in developing standards that impact GIS.

E.2 BACKGROUND

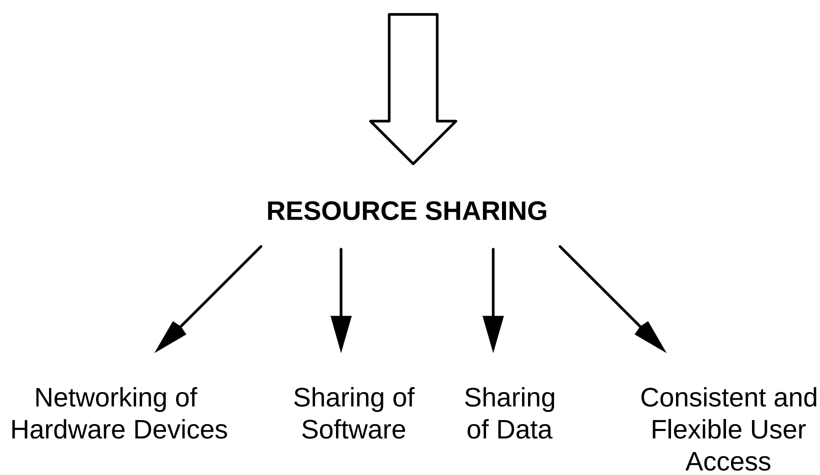
The underlying justification for the adoption of standards in computer systems is that information has value to users, and benefits are derived from the ability to share and exchange information globally in an organization or between multiple organizations. This implies the elimination of duplicative and inconsistent practices in data management with an open computing environment allowing easy access to information by those who need it.

Historically, information systems have been developed as independent stand-alone systems each with its own set of applications and database to support the specific needs of that system's applications. As existing systems evolved and additional systems were developed, considerable duplication of data has evolved because it was easier to duplicate the information in each database than to provide access to other system applications. This is particularly true in cases where institutional boundaries inhibit the exchange of information and where technical barriers resulting from incompatible computer systems complicate the process of data exchange. In the past five years, however, most large organizations have become focused on the need to share information and to develop more "open" institutional and technical environments to allow this to happen. The fundamental element necessary to enable an open computing system environment is a realistic set of information system standards.

"Standards," as the term is applied in data processing, covers many different topic areas that relate to hardware, software, data formats, and procedures through which users access computer systems. This section will explore the topic of computing standards and will illustrate how they impact GIS.

The central theme pushing the adoption of standards is "resource sharing" (see Figure E-1). The interest in distributed computing has demanded products which allow connection between devices, exchange of data, flexible access to software programs, and consistent procedures for user interaction. The importance of standards is most apparent in cases where hardware and software products from different vendors must be linked, or where independently created databases must be exchanged. Despite impressive technological advances in the last ten years that address these concerns, solutions for communicating among disparate systems and databases have been accomplished without overall guidelines using special hardware and customized software for a particular implementation.

FIGURE E-1
COMPUTING STANDARDS



E.3 TAXONOMY OF STANDARDS

Table E-1 presents a list of standards categories organized into the following major classes and further explained below:

- Hardware and Physical Connection Standards
- Network Communication Standards
- Software Standards
- Data Format Standards
- Data Presentation/User Access Standards
- User Design Standards.

Insert Table E-1

E.3.1 Hardware and Physical Connection Standards

This class of standards addresses important concerns relative to the physical connection and cabling of hardware devices. Accepted standards for cabling types, electrical interfaces, and cable connectors are now widely adhered to in the computer industry. The popular RS-232 standard for asynchronous communications is one example of a widely adopted standard. Other standards are in place for communication along other types of cabling (twisted wire, coaxial, fiber optic) and remote communication such as microwave or radio.

The computer hardware industry, with support from professional associations such as the Institute of Electrical and Electronic Engineers (IEEE) and the Electronics Industry Association (EIA), has been prompt in establishing physical connection standards. As new communication media becomes available (e.g., fiber optic cable), standards for its efficient use are quickly being accepted. Also included in this category are physical format standards for mass storage of data on tape and disk. Physical formats for magnetic disk storage media are fairly well established. Of particular concern to GIS, however, is the lack of standards for optical disk systems. Optical storage systems will play an increasingly important role in handling the large volumes of data that will characterize GISs.

E.3.2 Communication and Network Management Standards

Communication between two or more devices simply involves the transfer of binary digits (bits). For devices to make sense of these bits, however, they must be "packaged" in a specific format. This is the role of a particular communication "protocol"--to describe how the bits are arranged so that they can convey all information necessary for effective communication. Protocols describe the format of transmission to accomplish the following tasks:

- Transfer messages from one device or node to one or more destination nodes.
- Control routing of messages between multiple networks.
- Provide information for error checking and connection.

Specific protocols are now used to transfer information among devices in local networks, using a variety of cabling types and wide area networks (e.g., digital or analog telephone circuits, radio, microwave). National and international standards organizations have made significant progress in drafting standard protocols, but data processing environments are still largely characterized by many independent proprietary software programs that operate on specific vendor platforms.

The most ambitious example of standardization in communication protocols is represented by the Open Systems Interconnect (OSI) model defined by the International Standards Organization (ISO). The OSI model does not in itself define specific protocols. It is structured as a series of "layers" for communication between devices (see Table E-2). Specific protocols addressing the requirements of this series of layers are evaluated and accepted by the ISO if they comply with the tenets of the model. Upper layers of the OSI model are not yet fully defined, but the progress made by the ISO in the past five years has been impressive. Continued demands by users for open systems will motivate the computer industry to further support standards efforts and offer products that comply with accepted standards.

TABLE E-2
LAYERS OF THE OPEN SYSTEMS INTERCONNECTION MODEL

7.	Application Layer	Provides access of specific applications to lower level OSI layers.
6.	Presentation Layer	Provides applications with the means to interpret information exchanged on a network; provides specific data formatting rules.
5.	Session Layer	Provides coordination of display and presentation functions between multiple nodes on a network.
4.	Transport Layer	Establishes protocols for transparent flow of data once a networking has been established, and includes high-level error checking procedures.
3.	Network Layer	Controls high-level network routing; maintains address information and error reporting to Layer 4.
2.	Data Link Layer	Establishes protocols for proper transmission of messages between two network nodes; detects errors that occur at Layer 1.
1.	Physical Layer	Defines electrical interfaces, cabling, and physical connection standards.

There are not current standards fully accepted for all layers of the OSI model, and this has inhibited the development of software and hardware products to fully comply with the model. While development continues, products are in existence which offer partial compliance with standards accepted under OSI. A de-facto network protocol standard addressing most of Levels 2 through 4 of the OSI model is the TCP/IP protocol. TCP/IP was developed by the U.S. Defense Department to provide a means to network computers from multiple vendors. The purchase power of the Defense establishment influenced most major vendors to develop products which supported this protocol. TCP/IP is now popular in many large government and private organizations where multi-vendor networks are in place. It is expected that, in the future, more complete and robust sets of products that comply more fully with the OSI model will replace TCP/IP.

Most GISs and an increasing number of general purpose data processing systems are based on distributed networks using processing units called "servers" often storing a central database and workstations (also know as "clients") capable of local processing which access the server's database to perform applications. This "client-server" network model is dependent upon specific communication protocols and programming tools that support network operations. Many client-server networks in operation today use the X-Window standard which establishes rules for communication, network protocols, user interfaces, and programming tools to support the network.

In complex networks supporting large numbers of workstations or personal computers, there is an important need to manage the network, including such functions as problem diagnosis and correction; setting up addresses for system nodes; establishing access security; and other high-level management tasks. There is a class of software on the market today referred to as "network operating systems." These software packages work with the operating systems of the individual computer platforms to provide tools to support these network management tasks. Packages such as Novell's NETWARE, Banyan's VINES, and Microsoft's LAN Manager are examples of network operating system packages. There are certain network management protocol standards now accepted by standards bodies which network operating system vendors use in their products. These standards include the Simple Network Management Protocol (SNMP), originally developed for TCP/IP networks, and the Common Management Information Protocol (CMIP), an ISO compliant standard.

E.3.3 Software Standards

Until very recently, the core software or "operating system" that directs all fundamental operations of a computer system was unique to particular makes and models of processing units. Because operating system instructions have traditionally been specifically tied to a particular hardware architecture, they have traditionally run only on a specific family of computers. Until recently, the concept of "portable" operating systems that run on different models of computers by different vendors was not a practical alternative and, in fact, the computer industry is still dominated by system-specific operating systems. A notable exception to the dominance of proprietary operating systems is DOS for personal computers which has become a de facto industry standard for desktop systems.

Movement toward operating system portability on large systems is being led by vendors and professional groups promoting UNIX. UNIX has made a tremendous impact on the GIS community for users of intelligent workstations. UNIX is well entrenched as the most popular operating system for workstations and is just beginning to be implemented on larger processing units. Complicating the adoption of UNIX is the fact that many vendors have developed their own versions of UNIX which are not fully compatible and therefore limit the portability of application software among processing units from multiple vendors. Two major industry consortiums (Open Software Foundation and UNIX Systems and UNIX International) are addressing this problem by developing "standard" versions of UNIX and network management software and are encouraging the development of application software running on this operating system. In addition, the U.S. National Institute of Standards and Technology (NIST) has issued specifications for a portable operating system interface called POSIX with which many hardware and software vendors comply. For some computing environments, the POSIX-based system is a reasonable choice for a portable environment.

Over the next five years, the popularity and functionality of UNIX will undoubtedly increase and will offer a high degree of portability. Proprietary operating systems offered by specific vendors will not disappear but, for GIS installations, UNIX will very likely become the operating system of choice.

Of great concern in standardization of software is the role of high-level programming tools that users employ to build and customize applications on a given GIS package. As discussed in Section 3, the GIS package can be thought of as a "toolbox" providing major functions for graphic processing, database transactions, and geographic analysis. Most GIS software packages provide macro languages of varying levels of functionality, and some provide more powerful "fourth generation" languages for the customization of applications. Also, fourth generation languages are often included with third-party database management systems (e.g., ORACLE, INFORMIX, DB2), which many GIS vendors have integrated with their packages for attribute handling. Currently, the high-level application programming environment is characterized by proprietary languages and procedures with minimal consistency or standardization regarding functionality, commands, syntax, and user interfaces. Over the past year, there have been some trends toward standardization of application development tools, but this is an area that still warrants major development and the adoption of consistent and portable programming tools.

Interest in standard languages for user commands and database queries has been evident for at least the past five years in the GIS community. The standard query language (SQL) developed and promoted by IBM has become a de facto standard by many GIS and database management software developers. The benefit of a standard query language is that users are given a consistent vehicle to interact with different software packages running on different systems.

While SQL has become a de-facto standard for tabular database query and reporting, it has not yet been extended for full use by GISs. A number of groups, with vendor involvement, are in the process of developing a standard "geographic query language" (GQL) which adds spatial operations to the suite of tabular-based commands in SQL. Within the next three or four years, it is likely that a commercially viable "GQL" will be available.

E.3.4 Data Presentation and User Access

This category includes standards that impact the presentation and display of data from a GIS and the vehicles through which users gain access to the software and database. GIS software packages allow the user to generate products in the form of screen displays and hard copy reports, maps, or other graphics. The GIS database describes data content and relationships, not the appearance of products generated from an extraction or analysis of the GIS database. Therefore, tools are provided to control such display parameters as size, scale, symbology for map features, color, and annotation.

GIS software packages normally provide "libraries" from which users can select and manipulate display parameters. Specific display and plotting devices vary, and interfaces are needed to translate these user-defined display parameters into specific commands that drive the output devices (e.g., graphic monitors, pen plotters, electrostatic plotters). In the computer graphics industry today, there are several program libraries that let programmers easily translate user display parameters in "primitives" that can be sent to a device (e.g., pen plotter) for output. By predefining a set of primitives (lines, arcs, geometric shapes), the programmer is freed from having to develop specific programs to drive display and plotting devices. Adherence by GIS vendors to a display interface standard is important, because it ensures optional flexibility in customizing displays and supports a wide range of output devices.

A very popular topic in data processing today is the graphical user interface (GUI). It provides programmers with a "tool kit" of program modules to access the operating system and the application software. In addition to providing a flexible development environment for programmers, GUIs provide an efficient way for users to invoke commands and access portions of the database. Rather than typing in command statements, users access graphic windows with a mouse or other pointing device and run applications using a "point and click" technique. The adoption of an industry-wide standard GUI could allow users to quickly learn and become proficient with a particular GIS package regardless of the specific software vendor or hardware platform.

Several industry "standard" GUIs are popular today. The two most often cited are MOTIF, developed by the OSF consortium, and Open Look, developed by the UNIX Systems Laboratory. Other vendor proprietary GUIs are in use, however, including Presentation Manager supported by IBM for the OS/2 operating system, WINDOWS developed by Microsoft Corporation for personal computers, and the Macintosh GUI supported by Apple.

E.3.5 Data Structures

This class of standards covers the logical structure of data stored in a computer system. The term "database architecture" is sometimes used to describe particular approaches to the storage of data elements and the logical relationships between the data elements. For the storage of nongraphic, attribute data, three main architectures (hierarchical, network, and relational) or models are now popular. In each case, there are specific rules governing how these databases are structured and the functionality of particular implementations of these models by vendors. In the GIS industry, as in the broader data processing community, in general, the relational database model is now preferred for most database applications. Many GIS software vendors offer links between their relational database management packages from third party vendors.

Spatial data for storage of map features can be broadly categorized as vector (storing actual x, y coordinate strings) and raster (grid format). Specific formats for storing vector data are almost as numerous as vendors offering GIS or automated mapping software. The trend toward standards should not dictate specific file formats for implementation by GIS vendors, because this would inhibit the development of more efficient formats. However, it is important that basic guidelines are followed to allow flexible interchange of data for use with different software.

The need for exchange of graphic data has, to date, been addressed through two different strategies:

1. Development of specific programs to translate formats directly from one vendor's structure to another.

2. Translation of a specific vendor format to an accepted intermediate exchange format followed by translation to another vendor's structure.

The second approach ultimately holds the best long-term solution for flexible exchange of data between multiple hardware platforms but, to date, no fully acceptable intermediate exchange format has been devised, and unofficial interface formats are in use. Some of these unofficial standards include the interactive graphics exchange standard (IGES), ISIF (developed by Intergraph Corporation for the exchange of vector graphic data), and DXF (developed by Auto Desk, Inc., for the exchange of microcomputer vector graphics). The U.S. federal government also provides data for distribution in several de-facto formats, including DLG (for USGS topographic data) and TIGER (for Census Bureau boundary and demographic data).

Most existing exchange formats for vector graphics work well with non-topologically structured data but have limitations in their ability to exchange more complex topologically structured data. Topologically structured data explicitly defines the spatial connectivity and adjacency relationships of map features and, therefore, data translation procedures must maintain these relationships. Also, most existing exchange formats were designed to handle engineering graphics which contain many parametrically defined graphic entities (arcs, spline curves, ellipses) which do not always translate directly into graphic formats supported by many GIS software packages.

A major project initiated by the U.S. Geological Survey (USGS) and carried out by the National Committee for Digital Cartographic Data Standards (NCDCDS)[1] proposed an exchange format appropriate for digital cartographic data now referred to as the Spatial Data Transfer Standard (SDTS). SDTS was accepted as a federal government standard by the National Institute of Standards and Technology (NIST) in 1992. SDTS described standards rules and an overall structure for storing topological vector data, spatial attributes, and characteristics about the database. The actual use of SDTS as a technical file format standard is dependent on the development of "profiles" which apply SDTS to specific database scenarios. The federal government is testing a profile for storage and transfer of vector map data handled by federal agencies, but no profile has yet been developed specifically for local or state governments or utility governments. SDTS provides a very sound foundation for a standard GIS data exchange process, but work must still be done to apply SDTS to user environments outside the federal government.

[1]The NCDCDS was reorganized in 1991 into the Federal Geographic Data Committee (FGDC).

E.3.6 User Design Standards

E.3.6.1 Database Design

There are a host of standards issues that should influence the design and implementation of GIS. In any automation project, the creation of a database should be preceded by a design phase that describes the content and format of the database and products generated from the system. These design decisions impact attribute data schemes and coding rules, map accuracy criteria, quality control, and map design criteria, all of which are largely independent of a particular vendor's software.

Schemes for storing GIS attribute data should adhere to existing standards within an organization as a whole and, if applicable, in a broader community of users. The schema specifies field lengths, data element formats (e.g., integer, character, etc.), and other characteristics of the database. One example of GIS data for which standard schemes are appropriate is site address information. Standard address formats define street name, address number, street type, prefix, and suffix as individual data fields with specified field lengths and coding scheme fields. Use of a consistent schema can greatly reduce the problems encountered in exchanging attribute data between systems. Users should take sufficient time to review existing schemes and comply with existing standards.

Alpha or numeric coding and classification schemes for data are used for many types of data that are stored on computer systems. Classifying data into parent and subclasses and coding these classes appropriately make it easy to retrieve and analyze data at different levels of detail. For example, land use or land cover classification schemes normally use some hierarchical classes, each of which has alpha or numeric codes assigned (e.g., code of "100" for urban with a subclass code of "110" for residential). A large amount of the data in GISs is stored in coded form, and it is easy to see why consistent use of coding and classification can facilitate exchange of data among systems. Unfortunately, it is all too often the case that database design is done in a vacuum, and standard coding schemes are not followed. Users should adhere to accepted coding and classification standards within their own organization as well as standards established by government and professional groups, at national and international levels.

Selecting suitable accuracy levels for the compilation and automation of map layers in GIS is a perplexing issue for most users. The objectives of map accuracy standards are to explicitly define and measure levels of accuracy, not to dictate what accuracy is appropriate for particular users. Accuracy standards address both horizontal and vertical placement of map features and describe maximum errors of displacement relative to their actual position. Since 1947, the National Map Accuracy Standards originally published by the U.S. Bureau of the Budget has been the accepted standard for defining the accuracy of topographic and other types of base maps. This standard has traditionally been used in the preparation of automated base maps for GISs. Recently, a new accuracy standard for large-scale line maps was developed and proposed by a Committee of the American Society of Photogrammetry and Remote Sensing (ASPRS) which addresses some limitations of the National Map Accuracy Standards and is more suited to automated mapping. While this standard has not yet been officially approved by a federal government agency, it is likely that it will become a standard in the near future.

E.3.6.2 Meta-data

With the diverse sources from which GIS databases are built, it is extremely important to maintain information about the integrity, quality, and lineage (history of use and changes) of data in the GIS. These factors determine how the data can be appropriately used in specific GIS applications. "Integrity" and "quality" encompass many characteristics of a GIS database such as map accuracy, completeness, and data currency. Several approaches may be used to track this information, including the creation of data dictionaries that maintain information about data "layers" or the inclusion of codes assigned as attributes for individual map features to characterize and track quality and integrity.

In recent years, there has been considerable activity in the development of meta-databases--tabular databases that hold information about the GIS to provide users with a view of the data to make decisions on how to use it, before accessing the GIS data itself. Meta-databases usually contain the following components:

- Description or listing of the data features and attribute content of the GIS
- Geographic coverage information
- Data sources and methods of collection
- Quality (including completeness, currency, accuracy, reliability)
- Lineage data tracking the history of changes or transformations of the data.

A number of GIS users have built their own meta-databases, but several standards formats are in development. Coordinated efforts by the GIS Committee of the ASTM (formally the American Society of Testing and Materials) and the U.S. Federal Geographic Data Committee (FGDC) are developing a standard meta-data format. At the same time, vendors are beginning to understand the importance of meta-data and are beginning to incorporate software tools to automatically create and update meta-databases as GIS databases are developed and changed. Over the next three years, GIS users can expect to have standard formats available for meta-data, but it is expected that specific implementations of these will usually require some modification to suit the needs of particular users.

E.4 STANDARDS ORGANIZATIONS

A large number of organizations are now actively involved in developing and promoting computing standards. Government organizations establish and enforce standards for more consistent communication among agencies and international bodies, and with multi-national representation, seek to encourage communication standards globally. Several independent professional associations and industry consortiums are also heavily involved in the standards movement.

Organizations that are active in the development, approval, and promotion of standards may be classified into the following categories:

- Government Organizations: Government agencies in the U.S., Canada, and other countries with specific responsibility for approval of information system standards (and other types of standards) for use by their constituencies. Example: National Institute of Standards and Technology (NIST).

- Independent Standards Bodies: Non-aligned standards bodies with representation from government agencies, private companies, vendors, professional organizations with formal processes for standards consensus and development, approval, and testing. Example: American National Standards Institute (ANSI).

- Vendor Consortia and Trade Associations: Formal or informal associations of information system product or service vendors with missions for joint definition, development, and promotion of standard-based products for their customer base. Example: Open Software Foundation (OSF).

- Professional Organizations: Professional organizations or associations with missions involving education, interaction between its members, and advocacy on information system standards issues. Example: Urban and Regional Information Systems Association (URISA).

- User Groups and User Consortia: Formal user groups established around a particular vendor product or groups of user organizations established to promote and influence the development and adoption of open system products by vendors. Examples: Digital Equipment Computer Users Society (DECUS); Corporation for Open Systems (COS).

There are both formal and informal relationships between the many organizations in these categories. The confusing web of organizations apparent now is, in part, the result of the relative immaturity of standards development. The field is quite volatile, and activities are influenced by a number of forces, including:

- Requirements for large user organizations to find better ways to share information and to cut costs

- Government organizations needing to institute formal mandates to direct information system procurement and development

- Market forces and interest by vendors in maintaining and increasing their own customer base.

In the United States, the approval and adoption of standards at the federal level is led by the National Institute of Technology and Standards (NIST), formerly the National Bureau of Standards. This agency is responsible for the testing and approval of standards of all types and has directed considerable effort in recent years to data processing standards. NIST oversees the continued development of the Federal Information Processing Standards (FIPS) which include specific standards for data processing. Many of the FIPS standards include operational standards accepted by the ISO, IEEE, and other standards approval organizations. One major component of FIPS is the Government Open Systems Interconnection Profile (GOSIP) which is a layered model responding to OSI model compliant standards. It was developed as a mandate for computer acquisitions by federal government agencies. In the near future, all federal computer acquisitions must adhere to GOSIP.

For years, professional organizations operating in the U.S. and internationally have played an important role in developing and promoting standards important to the data processing community. Two independent standards bodies in the U.S., the American National Standards Institute (ANSI) and the Institute of Electrical and Electronic Engineers (IEEE), have the active participation of government agencies, industry, and academic institutions. Computing standards accepted by these groups generally become universally accepted by the industry and users. Each of these organizations maintains several subcommittees formed to investigate and test various standards. ANSI is the U.S. representative to the International Organization for Standardization (ISO) and contributes to the development and testing of proposed standards which comply with the OSI model.

Trade organizations and special user consortia play an important role in the development and acceptance of computing standards. The last five years, in particular, has seen a significant jump in the participation in standards development and promotion. This is largely because of the maturing of standards development efforts and the increased user demand for products that comply with standards. Computer hardware and software vendors have accepted the fact that, to remain competitive, they must develop effective strategies to meet this demand while protecting their existing proprietary markets. Groups like the Corporation for Open Systems (COS) have been established by large user organizations specifically to test and promote products based on open system standards.

E.5 THE STATUS OF STANDARDS AND THE ROLE OF THE USER

As indicated in the discussion of standards categories above, there is much activity in the adoption of specific standards and the development of compliant products by hardware and software vendors. Despite impressive progress that has been made, we are quite far from the open system computing utopia of complete interoperablility of hardware and software products, transparent user interfaces, and portable software. Table E-3 lists several key items of concern in open system computing and the status in the development of specific standards and products that address those concerns.

Users can take an active role to encourage and accelerate the standards process, first by becoming educated about standards issues that will impact them and their organizations, and then by using formal avenues to incorporate standards into system design and development activities. Education can be facilitated through participation in professional organizations, special training seminars on specific standards issues, and regular reading of key trade publications.

The most direct way to influence the computer industry is through procurement specifications that require compliance with standards. However, users must be knowledgeable about the current status of product offerings to make realistic demands on vendors and to avoid making requirements that are not feasible or very expensive to meet. With the current immaturity of standards definition and industry products, it is often less expensive to develop systems based on a specific vendor's proprietary products that do not adhere with standards. The benefit of standards, of course, is the ability to connect and support products from multiple vendors resulting in a long-term payoff, both functionally and financially.

While users can have some influence on the direction of product development by hardware and software vendors, they can have the greatest positive impact by concentrating on standards in the "User Design" category described above concentrating on the format, coding conventions, and feature definitions for the database. In cases where standards exist within an organization or a wider professional community, these should be used in the database design process. Users from different parts of the organization should be involved in the database design process to ensure that it is complete and establishes the most effective foundation for smooth sharing of information. This must be followed by developing a set of rules and procedural standards for data access and updating to guarantee that standards are complied with on an ongoing basis.

Insert Table E-3

TABLE E-1

TAXONOMY OF COMPUTING STANDARDS WITH SELECTED EXAMPLES

Hardware and Physical Connection Standards	
Cabling and couplers	Cabling types (e.g., twisted wire, coaxial, fiber optic), physical couplers, and connectors.
Electric interfaces	Voltage and frequency standards.
Storage media format	Physical format for data storage on tapes, magnetic disks, and optical storage media.
Network Communication Standards	
Local network protocols	Protocols supporting communication on network connecting devices in close proximity over direct cabling schemes.
Wide area network protocols	Protocols supporting communication among widely spaced devices using remote communication media (e.g., X.25 packet switching).
Integrated data communications	Digital communication standards to support data, voice, and other communication channels on integrated network (e.g., ISDN).
Distributed network management	Tools to provide effective communication on networks with distributed processing and/or data.
Software Standards	
Operating systems	Portable operating systems allowing more flexible networking and applications which run without modification on multiple platforms (e.g., DOS, UNIX).
Programming languages	3rd generation language standards (e.g., Fortran, C).
Query languages	Standard query languages allow consistent interface to different databases (e.g., SQL)
High-level application development languages	4th generation application development software with tools for customizing, screen and report generation, and analysis.
Data Format Standards	
Database architecture	Basic structure for storage of spatial and attribute data, and the relationship between data elements (e.g., hierarchical, relational).
Spatial data structures	Structure for storing map features and other graphic entities (in raster and vector form) and the relationship between them.
Exchange formats	Intermediate file formats that allow for translation and exchange of data from one format to another (IGES, ISIF, DXF), and standard exchange formats such as SDTS and its derivative profiles.
Object-based formats	Data structures for storage of objects, relationships between objects, and the rules or operators associated with them.
Presentation Access Standards	
Display and plotting interfaces	Standard interfaces and libraries describing primitives for communication and output on display and plotting devices (Graphic Kernal System, Plot 10).
Graphical user interfaces	Standard window environments for interactive user access (X-Windows, Motif, Presentation Manager, Open Look).
User Design Standards	
Data schemas	Attribute data file parameters such as field length, format, and other characteristics of data elements.
Data coding	Classification and coding schemes for data elements providing consistent reference by multiple users (e.g., standard land use or zoning codes).
Map design	Consistent selections of feature content by layer, placement of annotation; and use of symbols, line types, shading, and color.
Map accuracy	Standards for evaluating and assigning horizontal and vertical accuracy to maps.
Data integrity and quality rating	Guidelines for rating and storing information on map integrity and quality (accuracy, source material, currency, completeness).

TABLE E-3
STATUS OF IMPORTANT COMPUTING STANDARDS ISSUES

Theme	Key Issue	Current Status
Application Portability	Ability to move high-level applications (using macro or 4GL code) to other platforms and software	Most high-level application development languages are vendor-proprietary; some portability exists for SQL-based routines.
Common Graphic User Interface	Graphic user interfaces with common "look and feel" across different hardware and software platforms	Several "standard" open GUIs (Motif, Openlook); with continued use of vendor-specific proprietary menu systems such as IBM Presentation Manager and Microsoft Windows.
Common Query Dialogue	Common command/query dialogue for entire GIS (graphic data and attributes)	Some strong commitment by vendors to SQL-based systems for tabular data; many proprietary approaches and variations still in use by GIS vendors; no standard GIS query language is in use, but some initial developments are in progress to develop SQL extensions for a full GIS environment.
Database Exchange	Transfer of GIS databases from one platform to another without need to do translation or data restructuring	RDBMS vendors offering good multi-platform exchange for tabular data and advances in distributed databases; most GIS software running on multiple platforms requires batch data translation using specially designed programs available from vendors; some transfer of data in binary form to speed process; SDTS offers foundation for standard intermediate GIS data exchange format.
Flexible LAN	More flexible options for configuring local area networks	Standards now exist for cabling and connection schemes (e.g., Ethernet - ISO 8802/3 and Token Ring ISO 8802/5); wiring hubs; ability exists to use multiple types of cable (fiber, coax, twisted pair); wireless LAN standards emerging.
Full System Distribution	Distribution of databases and processing power on a network to gain overall performance and flexibility	Most GIS networks still require data downloads for local processing by workstations; vendors are developing more robust tools to track use of central GIS database; strong commitment now to client-server networks based on X/windows; standards being developed for fully distributed systems for tabular data by RDBMS vendors.
GIS Meta-data	Standard formats for storing descriptive, quality, and lineage data about the GIS database and software tools to maintain the meta-database	Meta-data component of SDTS provides basics, but no fully accepted standard format exists; current efforts in progress by ASTM and the FGDC to approve a meta-data format standard; some GIS vendors and users beginning to develop tools to support direct update of the meta-database.
High-speed Multi-platform Networks	Mixing of different hardware platforms and operating systems from multiple vendors on a high-speed network for interactive operations	GIS networks still need common operating system for efficient interactive environment; industry-standard network protocols exist (TCP/IP, NFS) for file transfer, terminal emulation; more sophisticated offerings for multi-protocol bridges/routers; vendors of network operating systems (e.g., Novell, Banyan, Microsoft) beginning to offer capabilities to tie together networks with multi-vendor hardware and operating systems.
Higher-speed Networks	Higher speed local area networks and more flexible/less expensive wide-area communications	FDDI and FDDI-2 are emerging standards for higher-speed in local networks; Asynchronous Transfer Mode (ATM) is an emerging high-speed local protocol; Frame Relay and switched multi-megabit data service (SMDS) standards; more flexibility and standards compliance by communication carriers and manufacturers of modems, multiplexors, digital service units.
Network Management	Easy and robust tools to manage and monitor computer networks (multi-vendor)	SNMP and CMIP are standards for defining network "objects" and monitoring/diagnostic statistics; recent commitment by vendors to standards compliant software to ease problems in configuring, monitoring, diagnosing problems, and optimizing network performance.
Software Portability	Running vendor application packages on different platforms without need to revise and recompile code	Current strong commitment to UNIX versions and use of C as programming language. UNIX versions are different, and application software contains specific machine dependencies inhibiting portability; POSIX and OSF, DME offer path to more portability.

Explanation of Acronyms

ASTM	Independent standards organization formerly known as the American Society of Testing and Materials.
CMIP **Common Management Information Protocol**	An internationally accepted protocol for network management.
GUI **Graphic User Interface**	A menu-driven system allowing users to point at graphic selections on a screen using a mouse to issue commands or access data.
ISO **International Standards Organization**	International body responsible for approval of standards. Responsible for development of the OSI model.
NFS **Network File System**	An industry-standard local area network communication protocol (originally developed by SUN computer) now supported by most computer vendors.
OSF **Open Software Foundation**	A consortium of computer vendors (IBM and DEC included) developing open system software products, including a standard UNIX operating system and open tools for distributed database management and network management.
POSIX **Portable Operating System Interface**	An operating system standard for UNIX platforms allowing the development of portable applications to be developed using the C programming language.
RDBMS	Relational database management system.
SDTS **Spatial Data Transfer Standard**	National Institute of Standards and Technology (NIST approved standard giving framework for exchange of GIS data.)
SNMP **Simple Network Management Protocol**	An industry-standard protocol for monitoring and controlling network devices.
SQL	An industry-standard command dialogue for data query and retrieval.
TCP/IP **Transmission Control/Internet Protocol**	An industry-standard local area network communication protocol supported by most computer vendors.

APPENDIX F
SAMPLE GIS POSITION DESCRIPTIONS

Contents: This Appendix contains some sample descriptions for GIS positions that fulfill technical and management roles described in the *West Virginia GIS Conceptual Design* report. These positions are described in general terms, and it will be necessary to modify them to meet specific needs of state Departments and Divisions. Also, salary ranges represent national averages, and some adjustments will be needed here to make these consistent with other salary grades in state government.

GIS Coordinator or Manager

Job Responsibilities and Duties:

This is a highly responsible administrative position with work in planning, organizing, and directing GIS activities for multiple user organizations.

Work involves responsibility for providing technical expertise and supervision for the day-to-day implementation and operation of the GIS within the context of programs and policy directives. Work includes coordinating GIS activities, managing GIS vendor contracts, planning and organizing system development activities, and other GIS project management activities. Supervision is exercised over a staff of subordinate GIS staff positions. Work is performed independently within established policies, procedures, and guidelines and is reviewed through reports, conferences, and system performance.

Specifically, the GIS Coordinator:

- Manages resources of GIS; confers and coordinates with user organizations and GIS personnel to determine user needs.
- Oversees all GIS operations; supervises directly, or through GIS staff, all activity associated with implementation, operation, and enhancement of the GIS.
- Coordinates contracts with GIS hardware, software, data conversion, and other providers of GIS products and services.
- Organizes and coordinates partnerships with other government organizations and the private sector for GIS development and operation.
- Organizes application and other development activities.
- Prepares annual budget for GIS operations, capital expenditures, and professional services.
- Coordinates design and enforcement of GIS standards and operating procedures.
- Develops GIS project schedules and work programs to support statewide GIS operations.
- Performs related work as required.

Qualifications:

The GIS Coordinator must have any combination of education and experience equivalent to graduation from an accredited four-year college or university with major course work in geography, computer science, planning, engineering, or related fields, and thorough experience with geographic information system design, implementation, and management, including considerable experience in both administrative and project supervision. In addition, applicants must have at least four years experience in a GIS technical or management role for a large organization. Applicants for this position must demonstrate sound skills in administrative and technical management, budget responsibility, and technical writing. In addition, applicants must have excellent communication skills and the ability and interest to work directly with individuals and groups, and find innovative and efficient solutions to administrative and technical problems. More specifically, the GIS Manager must have the following knowledge, skills, and abilities:

- Thorough knowledge of automated mapping and spatial information processing methods and techniques
- Thorough knowledge of the capabilities of automated mapping and geographic information processing systems
- Knowledge of program budgeting and contract management
- Ability to plan, organize, and manage the programs and activities of a GIS personnel team in GIS design, operation, and implementation
- Ability to translate technical concepts and terminology in terms understandable to elected officials and other department heads
- Ability to make decisions based on factual data, and to evaluate progress or success of computerized projects and systems
- Ability to establish and maintain effective relationships with user departments, administrative officials, and employees.

Salary Range (includes 25% multiplier on base salary): $55,000 to $80,000

GIS Database Administrator

Job Responsibilities and Duties:

This is a highly responsible administrative and technical position with work in managing, supervising, and assisting in GIS database creation and maintenance activities.

Work involves design, development, and management of GIS data resources within the context of programs and policy directives. Work includes coordination of GIS database activities, assistance in managing GIS data conversion contracts, establishment and enforcement of database standards, participation in planning and organizing system development activities, and other GIS database related activities. Supervision is exercised over a staff of subordinate GIS staff positions. Work is performed independently within established policies, procedures, and guidelines and is reviewed through reports, conferences, and system performance. Specifically, the GIS Database Administrator:

- Manages database resources of the GIS; confers and coordinates with the GIS Manager, user organizations, and GIS personnel to determine user needs
- Oversees all GIS operations; supervises directly or through GIS staff all activity associated with implementation, operation, and enhancement of the GIS
- Assists in coordinating contracts with GIS data conversion and related service vendors
- Responsible for development and enforcement of data standards impacting coding schemes, database designs, and data update procedures
- Oversees data quality control procedures and loading of data to the system; coordinates activities in data transfer and translation
- Establishes and enforces procedures for data access security and access rights
- Produces or reviews preliminary and detailed designs for new or modified database elements
- Performs or assists in investigating and recommending sources and approaches for automating GIS database elements
- Provides documentation and training for system users
- Performs related work as required.

Qualifications:

The GIS Database Administrator must have any combination of education and experience equivalent to graduation from an accredited four-year college or university with major course work in geography, computer science, planning, engineering, or a related field, and considerable experience with geographic information system design, implementation and management, including experiences in GIS database-related activities. At lease two year's of work experience in government and familiarity with information management, mapping, and geoprocessing operations of government agencies is required. More specifically, the GIS Database Administrator must have the following knowledge, skills, and abilities:

- Thorough knowledge of automated and spatial information processing methods and techniques
- Thorough knowledge of the capabilities of automated mapping and geographic information processing systems
- Considerable knowledge of cartographic principles, automated mapping, GIS database design concepts, and database structure
- Considerable knowledge of data conversion, translation, and transfer methods and techniques
- Ability to coordinate with GIS staff and users to develop standards for system access, security, and data integrity
- Ability to work with and oversee efforts of private database development contractors
- Ability to translate technical concepts and terminology in terms understandable to management, department officials, and system users
- Ability to establish and maintain effective relationships with user departments, administrative officials, and employees.

Salary Range (includes 25% multiplier on base salary): $45,000 to $65,000

GIS Technical System Administrator

Job Responsibilities and Duties:

This is a highly responsible administrative and technical position with work in operating and managing GIS hardware and software resources.

Work involves design, operation, and management of GIS hardware and software resources within the context of programs and policy directives. Work includes coordination of GIS hardware and software activities; assistance in managing GIS hardware, software, and service vendor contracts; participation in planning and organizing system development activities; and other GIS hardware and software-related activities. Supervision is exercised over a staff of subordinate GIS staff positions. Work is performed independently within established policies, procedures, and guidelines and is reviewed through reports, conferences, and system performance. Specifically, the GIS Technical System Administrator:

- Manages hardware and software resources of GIS; confers and coordinates with GIS Manager, user organizations, and GIS personnel to determine user needs
- Oversees all GIS hardware and software operations; supervises directly or through GIS staff and groups all activity associated with implementation, operation, and enhancement of GIS hardware and software
- Assists in coordinating contracts with GIS hardware, software, and related service vendors
- Oversees installation, testing, and acceptance of hardware and software
- Responsible for adherence to applicable standards for computer hardware, software, and communication networks
- Coordinates maintenance and service contracts with hardware and software vendors
- Develops procedures for system backup and overseeing backup operations
- Monitors performance and conducts system performance evaluations
- Oversees and optimizes data communications and monitors network transaction flow
- Responds to user questions and resolves system operation problems
- Evaluates requirements for system upgrades and contributes to the selection and procurement of hardware, software, and related services

- Performs and assists in application design and development
- Performs related work as required.

Qualifications:

The GIS Technical Systems Administrator must have any combination of experience equivalent to graduation from an accredited four-year college or university with major course work in geography, computer science, planning, engineering, or a related field, and at least two years experience with computer system operations, geographic information system design, implementation and management, including experience in GIS hardware and software-related activities. Additional work experience in government and familiarity with computer system administration, data communications, mapping, and geoprocessing operations are desired. The GIS Technical Systems Administrator must have the following knowledge, skills, and abilities:

- Thorough knowledge of automated mapping and spatial information processing methods and techniques
- Thorough knowledge of the capabilities of automated mapping and geographic information processing systems
- Considerable knowledge of cartographic principles, automated mapping, GIS database design and structure
- Considerable knowledge of GIS hardware and software components, data communication, and network methods and techniques
- Ability to coordinate with GIS staff and users on system hardware and software
- Ability to translate technical concepts and terminology in terms understandable to management, department officials, and system users
- Ability to establish and maintain effective relationships with user departments, administrative officials, and employees.

Salary Range (includes 25% multiplier on base salary): $45,000 to $65,000

GIS Analyst/Programmer

Job Responsibilities and Duties:

This is a highly responsible technical position with work in application design, programming, and providing user support.

Work involves design and maintenance of data resources and operation of GIS hardware and software within the context of programs and policy directives. Work includes application development, user support, and other GIS related activities. Work is performed with considerable independence and is reviewed through reports, conferences, and system effectiveness. Specifically, the GIS Analyst/Programmer will:

- Analyze user organization functional requirements and develop GIS application designs
- Write GIS application programs
- Test GIS applications
- Prepare documentation of GIS application programs
- Prepare GIS application user manuals
- Train user staff in the operation of GIS applications
- Provide user support for GIS application programs
- Perform GIS hardware and software operations; perform activity associated with implementation, operation, and enhancement of the GIS hardware and software
- Assist in evaluating requirements for database enhancement and system upgrades
- Assist in designing and enforcement of GIS database standards for content
- Produce or review preliminary and detailed designs for new or modified database elements
- Investigate and recommend sources and approaches for automating GIS database elements
- Perform related work as required.

Qualifications:

The GIS Analyst/Programmer must have any combination of education and experience equivalent to graduation from an accredited four-year college or university with major course work in geography, computer science, planning, engineering, or a related field, or considerable experience with geographic information system design and implementation, including experience in GIS hardware and software-related activities. Work experience in government and familiarity with mapping and geoprocessing operations of government agencies are desired. The GIS Specialist must have the following knowledge, skills, and abilities:

- Extensive knowledge of GIS software used by West Virginia agencies
- Considerable knowledge of automated mapping and spatial information processing methods and techniques
- Considerable knowledge of the capabilities of automated mapping and geographic information processing systems
- Knowledge of cartographic design and drafting principles, automated mapping, and GIS database design
- Knowledge of GIS hardware and software components, data communication, and network methods and techniques
- Considerable skill in the design and programming of GIS applications
- Ability to think in logical terms and to communicate well with system users to translate application concepts into system applications
- Ability to translate technical concepts and terminology in terms understandable to system users
- Ability to establish and maintain effective relationships with user departments, administrative officials, and employees.

Salary Range (includes 25% multiplier on base salary): $35,000 to $50,000

GIS Technician

Job Responsibilities and Duties:

This is a highly technical position with work in performing GIS operations.

Work involves operations, including database creation and maintenance, executing applications, and other high level user functions. Work is performed with some independence and is supervised for efficiency and quality. Specifically, the GIS Technician will:

- Assist in design and implementation of applications; attend training and execute applications
- Perform GIS hardware and software operations; perform data quality control, produce plots and reports, and report system maintenance and performance problems
- Participate in installation, testing, and acceptance of hardware and software
- Provide input in evaluating requirements for database enhancement and system upgrades
- Assist in developing and following GIS database standards for content, format, design, and maintenance procedures
- Provide input in developing designs for new or modified database elements
- Assist in investigating and recommend sources and approaches for automating GIS database elements
- Perform related work as required.

Qualifications:

The GIS Technician must have any combination of education and experience equivalent to an associate degree in cartography, geography, engineering drafting, or automated cartography and training in use of geographic information system technology. Work experience in government and familiarity with mapping and geoprocessing operations of government agencies are desired. The GIS Technician must have the following knowledge, skills, and abilities:

- Knowledge of automated mapping and spatial processing methods and techniques
- Knowledge of the capabilities of automated mapping and geographic information processing systems

- Knowledge of cartographic drafting and design
- Ability to understand technical concepts and terminology, and ability to effectively use automated information processing methods and techniques
- Ability to establish and maintain effective relationships with user departments, administrative officials, and employees.

Salary Range (includes 25% multiplier on base salary): $20,000 to $40,000