MASTER ADDRESS FILE (MAF)/TIGER MODERNIZATION STUDY

FOR THE GEOGRAPHY DIVISION
UNITED STATES CENSUS BUREAU

JUNE 7, 2000

BOOZ·ALLEN & HAMILTON

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The Geography Division of the U.S. Census Bureau is responsible for developing geographic applications and executing the geographic and cartographic activities needed to support the Bureau in collecting and disseminating census data. As part of the 21st Century Master Address File /Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) Enhancements initiative, the Geography Division is proposing an effort to streamline the process for collecting and disseminating the geographic and cartographic data and modernizing the corresponding technology.

In the early 1980s, in an effort to streamline these activities, the Geography Division broke new ground by developing a valuable national resource to assist in conducting the census—the TIGER geographic data base. TIGER® includes geographic features, such as roads, railroads, water, boundaries, and other geographic information needed to support the programs of the U.S. Census Bureau. TIGER® is used extensively within the Bureau, and for many purposes other than the census by local/tribal governments and private industry.

The TIGER data base collects street address ranges, ZIP Codes, and feature names—it does not contain specific addresses. To provide accurate address data that could be maintained and updated, the Bureau developed the MAF in the 1990s. The MAF includes addresses that allow census forms to be mailed and, in some cases, provides descriptions of living quarters enabling census enumerators to deliver forms. Further, the MAF has a geocoding linkage to the postal information in TIGER®. Although TIGER® is part of the public domain, MAF is restricted from public use by Title 13 of the *U. S. Code*.

The use of MAF/TIGER greatly improved the accuracy of addresses, geographic features, and boundaries and improved the efficiency of Bureau operations. Today, however, the functional processes that MAF/TIGER were designed to support have expanded significantly and the technology available to support these processes has improved dramatically. Additionally, preparations for Census 2000 highlighted the need for this aging national resource to be updated to successfully execute the activities up to and including the 2010 decennial census.

As a first step to better understand the expanded functions, the high-level requirements, and unmet needs, Booz·Allen & Hamilton MAF/TIGER Modernization Study team members spoke with approximately 70 users and stakeholders. The team interviewed staff and managers across the Bureau and Geography Division and spoke with selected organizations external to the Bureau. The comments and opinions identified during these interviews are included in the body and appendices of this document.

During interviews with Geography Division managers, these requirements were mapped to the following four initiative objectives that the Geography Division proposed to supplement and enhance the Geographic Support base program:

- Correctly locate every street and other map feature in the TIGER data base, each MAF address, and implement an effective automated feature change detection methodology

- Implement a modern processing environment for the MAF/TIGER system that will support rapid application development, allow extensive use of commercially available software, and reduce staff training time
- Expand and encourage geographic partnership programs with state, local, and tribal governments that update the MAF/TIGER data base, including a “rolling LUCA” to maintain a current address list for use by all Bureau censuses and surveys and computer-based updates of governmental unit boundaries and other geographic areas

Implement a comprehensive plan for periodic MAF/TIGER evaluation and corrective activities that will guide planning for cost-effective future coverage and geocoding improvement operations.

Recent legislation and current efforts underscore the need to implement these objectives. For example, in 1994, Congress passed the Census Address List Improvement Act of 1994 (Public Law 103-430) mandating that the Bureau improve its partnerships with local and tribal governments by allowing them to review and update the census address list before the decennial census.

The address list review requirement alone leads to functional process issues, such as the mechanism for communicating and updating address file data, that must be addressed by appropriately selecting technologies that will meet the objectives of the Bureau. For example, improving the ability to provide the Bureau’s partners with timely data that is accessible and easy to work with would demonstrate the Bureau’s commitment to partnership. Such examples have led the Geography Division to embark on the 21st Century MAF/TIGER Enhancements initiative.

While multiple examples of functional and operational improvements will be addressed through the 21st Century MAF/TIGER Enhancements initiative, three key areas that will see significant improvement include—

- **Increased Use of the Internet**—To increase voluntary participation, the Geography Division proposes developing an Internet-based geographic and address update program. This program will enable participating local and tribal governments to incorporate the review and update of address, street, and boundary information into their normal business processes, improving timeliness and accuracy of the data.
- **Increased Accuracy of Address List and Map Feature Information**—To maintain this valuable and permanent asset, the Bureau proposes establishing a program to validate the current address list information and provide for its continuous maintenance until the 2010 decennial census. This program will provide a highly accurate and up-to-date resource that will be available to support other core activities of the U.S. Census Bureau that rely on address list and map feature information.

Improved Tools Supporting Field Operations—The 21st Century MAF/TIGER Enhancements initiative proposes improving the quality and efficiency of collecting and processing geographic data by taking advantage of commercially available state-of-the-art tools and techniques. Taking advantage of current technology moves data collection from a less accurate, less efficient paper-based system to a more precise digitized approach.

This modernization study is the first step in addressing the functional requirements and technologies needed to more accurately and efficiently conduct the decennial census. As the 21st Century MAF/TIGER Enhancements initiative continues, it is expected that the functional requirements will be further refined. Additionally, it is expected that applicable technologies will continue to mature, contributing to further improvements in the accuracy and efficiency of this national effort. This study addresses these issues as follows:

Chapter 1: Introduction and Background—Provides an introduction to the MAF/TIGER Modernization Study, including a high-level review of the challenges currently faced by the U.S. Census Bureau, a review of its overarching goals, and a review of the various geographic databases targeted by the initiative.

Chapter 2: Requirements and Objectives—Presents, at an aggregate level, the requirements that the MAF/TIGER Modernization Study team collected as a result of extensive interviews and facilitated discussions with both U.S. Census Bureau personnel and external stakeholders.

Chapter 3: Implementation Scenarios—Suggests a number of potential implementation scenarios for each of the four objectives of the U.S. Census Bureau's 21st Century MAF/TIGER Enhancements initiative.

Appendix A: Requirements for the MAF/TIGER Modernization Study—Provides the data that was the basis for the information provided in Chapter 2. Contains a detailed analysis of the more than 300 requirements that were extracted from interviews and facilitated sessions with both U.S. Census Bureau personnel and external stakeholders. The requirements are arrayed across 7 major requirement categories.

Appendix B: Mobile Computers and Global Positioning System—Provides information on the current capabilities and technology that support mobile (portable/hand held) computers and GPS applications.

Appendix C: Imagery Assessment—Presents a discussion of the current state of aerial and satellite imagery and its potential use in connection with improving the accuracy of the geographic locations of the natural features, manmade features, and existing structures within the MAF/TIGER data bases.

Appendix D: Current System Engineering Best Practices—Provides a review of current system engineering best practices. Included is a review of methodologies and practices employed to support the migration of large-scale, heterogeneous legacy applications to an integrated and homogeneous environment. Highlighted are those practices that improve programmer productivity, reduce application development time, and increase the quality of the systems produced.

Appendix E: Commercial Geospatial Software—Discusses commercially available software that potentially could provide the required tools and functionality to support the major geographic information system requirements of the U.S. Census Bureau.

Appendix F: Commercial Database Software—Presents key considerations and best practices for developers implementing large-scale database applications supporting geographically dispersed users. Included is a review of current database products and their application development tools sets.

Appendix G: Internet-Enabled Applications—Presents key considerations and best practices for developers implementing Internet-based applications. Discussed are issues such as security, capacity, and remote access.

Appendix H: Acronyms and Abbreviations

Note: *The discussion of specific vendors and products in this document is for the sole purpose of offering representative examples of products available for specific technology needs. This information is not intended as a promotion of any specific vendor or product. Additionally, the vendors and products discussed herein should not be considered exclusive from other potential vendors and products in the event of future procurement requirements.*

1. INTRODUCTION AND BACKGROUND

The mission of the U.S. Census Bureau is—

> *"to be the preeminent collector and provider of timely, relevant, and quality data about the people and economy of the United States."*

1.1 Introduction

The U.S. Census Bureau is the largest statistical agency of the U.S. Government. It conducts the constitutionally mandated Census of Population and Housing every 10 years to apportion seats in the House of Representatives, and it conducts 8 censuses related to economic activities and state and local government operations every 5 years. In addition, the Bureau conducts more than 100 demographic and economic surveys on a monthly, quarterly, and annual basis.

The tremendous array of information the Bureau collects describes the Nation's population, housing, businesses, government finances, foreign trade, and other vital characteristics. Analysis of this information forms the basis for decisions by a wide variety of entities including the Congress; the administration; private businesses; and state, local, and tribal planners.

The Bureau is composed of three core program areas: Decennial, Economic, and Demographic.

The Decennial program's principal activity is the decennial census, which is the Nation's oldest and most comprehensive source of population and housing information. The U.S. Government uses the results of the decennial census to apportion the seats in the House of Representatives. Decennial census data are also used in redistricting decisions and distribution of billions of dollars in federal funds.

The diverse activities of the Economic program produce information that is critical to the economic functioning of the Nation. Data such as housing starts, retail sales, wholesale trade, merchandise imports and exports, mining, and manufacturing are collected each month. These data support several leading indicators of national economic performance that influence decisions made by the Federal Reserve Board and Wall Street investors.

The Demographic program activities include the preparation of official population estimates and projections and the collection of survey data for official statistics on income and poverty, families, race and ethnic groups, and the Nation's housing characteristics. The collection and analysis of housing and socioeconomic data also supports the decennial census.

Reimbursable agreements with other federal agencies compose a large part of the Demographic program. For example, data from the Bureau's Consumer Expenditures Survey is an important element of the Consumer Price Index, which the Bureau of Labor Statistics calculates and publishes. The Current Population Survey gathers data to calculate employment and unemployment statistics. The Bureau's Survey of Income and Program Participation provides other government agencies with data to evaluate the effects of their programs on the population.

The statistical sampling techniques employed by the Bureau rely on relatively small samples of data that must be highly accurate to be considered valid. The credibility of the studies and surveys conducted by the Bureau relies on the accuracy of the address list and geographic database, which are critical elements of the Bureau's geographic support program.

The goal of geographic support is to provide the basic maps, address lists, address and geographic reference files, and associated processing systems needed to meet the geographic requirements of all Bureau programs.

To support its goal, the Bureau has developed an integrated computer-based Geographic Support System (GSS). The GSS needs to acquire large volumes of data from many external sources to establish and maintain a current and complete inventory of streets, roads, street names, address ranges, housing unit identifiers and addresses, and accurate geographic boundaries and their identifiers and other map information. This information must be updated periodically to meet the needs of the economic census, the current demographic statistics program, the intercensal demographic estimates program, and the future continuous measurement activities conducted by the Bureau.

Among the most important information technology resources used by the Bureau is the electronic address list of housing units located throughout the United States and the related island areas where it performs the censuses and surveys. This electronic address list is referred to as the Master Address File (MAF) and is one of the primary geographic support databases.

Another critical geographic database is the Topologically Integrated Geographic Encoding and Referencing (TIGER) system. The TIGER system is a companion system to the MAF that contains an inventory of geographic features and manmade structures. Key to this inventory is the geographic locations of all the features and structures and their linkage to specific addresses in the MAF. The geographic features and structures contained in TIGER® describe the street network for the United States and the current governmental and statistical boundaries and their identifiers used for so many of the Bureau's activities.

The accuracy and coverage of the current address and geographic databases is a significant issue for the three major programs of the Bureau. Although the accuracy and coverage of the existing inventory of residential structures is adequate for meeting Bureau performance levels, the inventory of business structures is barely adequate. The current state of the inventory of addresses is compromised by the absence of addresses for existing structures, the presence of addresses for structures that no longer exist, and most problematic, duplicate addresses for discrete structures.

The TIGER system is based on custom software developed by the Bureau during the 1980s. The TIGER system has essentially remained in its originally designed format and operating environment. Many of the original design and platform decisions were based on the state of information technology at that time and do not meet current technological expectations for data management and user accessibility.

The current TIGER system does not support the real-time concurrent updating of data. Bureau staff who are responsible for updating the information in the TIGER data base are required to "check out" the particular portion of the database, known as a "partition," that contains the geographic area requiring updating. During this time, the "checked out" geographic area is available for viewing only by other Bureau staff and does not include visibility into the updates being applied. The period of time that a partition can remain "checked out" can be extensive depending on the extent of the changes that need to be made.

The inability of the TIGER system to support real-time concurrent updating of geographic information causes significant operational issues concerning workload management, coordination of program activities, and accuracy of the TIGER products produced.

These problems cause the need for more resources and time for the execution of the field operations that support the three primary programs of the Bureau than would otherwise be necessary. Strategies to mitigate these issues have proven very expensive. With important matters such as federal funding and congressional redistricting depending on the accuracy and completeness of the geographic databases, the Bureau has been forced to execute costly field operations to ensure the integrity of the databases. Over the past decade, the Bureau has spent approximately $250 million on projects such as block canvassing and address listing to improve the accuracy of the address and geographic databases.

The principal address and geographic databases that support the Bureau's core programs, including MAF and TIGER®, are administered and maintained by the Geography Division within in the Decennial Census Directorate.

The mission of the Geography Division is—

> *"the development and accomplishment of all geographic and cartographic activities necessary to support the U.S. Census Bureau's data collection, processing, tabulation, and dissemination programs for the United States and its related island areas."*

The Geography Division is proposing a 21st Century MAF/TIGER Enhancements initiative to provide significant improvements in the operational processes that support the decennial census, quinquennial economic censuses, and numerous economic and demographic surveys that measure changing individual and household demographics and the economic condition of the Nation. These improvements will help the Bureau address the differential undercount of children, renters, and minorities. At the same time, they will improve the efficiency and effectiveness of enumeration operations and in the process save money and deliver results more quickly. The three key areas that will contribute to these performance improvements are—

- **Increased Use of the Internet**—To increase the voluntary participation of the 39,000 local and tribal governmental units in the United States in maintaining the Bureau's MAF and TIGER data base, an Internet-based MAF/TIGER update program will be developed. By establishing this Internet-based system, the local and tribal governments will be empowered to incorporate the review and update of address and related statistical information into their normal business process, improving the timeliness and accuracy of the data and streamlining the current manual process.

- **Increased Accuracy of Address List and Map Feature Information**—The Bureau uses the MAF information to deliver census questionnaires to every housing unit, check in returned questionnaires, and make follow-up visits to those housing units from which a questionnaire was not returned. Without a control address list, an accurate census is not possible. The current practice is to update the address list information used in the previous decennial census with current information. This is a costly and resource-intensive effort that requires an estimated $300 million each decade to complete. To maintain in a current and complete state this valuable and permanent asset, the Bureau will establish a program that will validate the current address list information and provide for its continuous maintenance until the 2010 decennial census. This program will provide a highly accurate and up-to-date resource that will be available to support other core activities that rely on address list information.
- **Improved Tools Supporting Field Operations**—The current field operations supporting the decennial census require nearly 20 million black and white map sheets and 13.1 million pages of address lists. Any identified updates to the information presented on these paper instruments must be manually documented and processed at a later time. The current state of information technology has rendered these practices inefficient and archaic. To improve the efficiency of field operations, the Bureau will initiate a program to develop the use of commercially available software and hardware, to create tools designed to streamline the capture of decennial census information and the geographic information required to support census execution.

1.2 Geographic Databases

The seven information systems that will be impacted by the 21st Century MAF/TIGER Enhancements initiative are—

Master Address File (MAF) System
Topological Integrated Geographic Encoding and Referencing (TIGER) System
Geographic Catalogue (GEOCAT) System
Geographic Programs Participant (GPP) System
Problem Referral (PRS) System
Production Control (PCS) System
Windowing Index (Windex) System.

The MAF and TIGER systems provide the core information and work products that support the Bureau's programs while the other systems, particularly the Production Control System (PCS), provide critical support functions for managing the execution of the programs.

1.2.1 Master Address File System

Following the 1990 decennial census, the Bureau undertook, for the first time, the development of a permanent housing unit address list, the MAF. Previously, address lists were prepared for each decennial census and then discarded. The Address Control File (ACF) from the 1990 census was a starting point for the MAF.

The MAF contains millions of records of housing unit mailing addresses, locational descriptions known to the Bureau, and nonresidential addresses such as those for businesses. The MAF lists every known residential address, whether occupied or vacant. The address list record includes the street address, mailing address, and geographic codes linking the record to U.S. Census Bureau geography. The MAF is for internal Bureau use only; the individual address data it contains are restricted from public use by Title 13 of the *U.S. Code*.

The MAF contains several types of addresses. City-style addresses include a house number and street name (e.g., 123 Main Street). They also may contain a unit designator for multiunit structures (e.g., Apartment B). They generally are a permanent identification for each structure and are part of an address system. Non-city-style addresses do not use a house number and street name description. They often are found in rural areas. Post Office box numbers are the most common form of non-city-style address; other forms include general delivery, rural route, and highway contract route addresses. These types of addresses rarely conform to any address system, and they are **not** a permanent identifier for a structure; they often are used by the occupant of the structure and can be taken or used by the occupant if (s)he moves to a different structure within the same postal delivery area. Also, they are not associated with any street name, which makes relating them to U.S. Census Bureau geography (geocoding) difficult. Residential addresses are associated with a dwelling or housing unit. Nonresidential addresses are for structures or units within a structure that are used for business or other nonresidential, purposes. Residential and nonresidential addresses may be city- or non-city-style addresses.

The MAF is updated regularly from numerous sources. For Census 2000, the primary source of updates was the U.S. Postal Service's (USPS) Delivery Sequence File (DSF). Other sources include field operations sponsored by the Bureau (e.g., Address Listing, Block Canvassing, Special Place Address Listing, List/Enumerate). In addition, local and tribal governments were asked to participate in the Local Update of Census Addresses (LUCA) Program and the new construction program to review each address in the MAF and provide missing addresses and missing streets/street names for the TIGER data base.

The Geography Division updates the MAF and produces address-related products from it. These products include address listings, address extracts, and housing unit counts. The MAF is geographically structured through its geocoding linkage to the street name/address range/ZIP Code/map spot location information in the TIGER data base.

1.2.2 Topologically Integrated Geographic Encoding and Referencing System

In the early 1980s, the Geography Division created the TIGER data base, initially for use in the 1990 decennial census, to manage all relevant geographic features, boundaries, and area identifiers needed to take the census. The database includes roads, railroads, water features, official boundaries, and other information for the entire United States and its related island areas. In addition to feature information, attributes of the features are collected (e.g., feature names, street address ranges, ZIP Codes). The database also contains codes for every geographic area, used for U.S. Census Bureau data tabulation, down to the census block level. Currently, TIGER® contains some housing unit location information for non-city-style addresses (digitized map spots) that are linked to associated records in the MAF.

The database and its rich set of attributes are used by every branch in the Geography Division and by many other divisions within the Bureau. Also, one extract of the TIGER data base, called the TIGER/Line® file, is distributed to the general public for use in numerous geographic and demographic applications. The TIGER/Line® files are credited with being a vital contributing factor in the development of the Geographic Information System (GIS) industry in the United States, and TIGER® is considered a national resource.

The Geography Division updates the TIGER data base with current geographic feature information and attributes, such as new boundary information, road information, and housing unit locations. Several programs to improve TIGER® currency are sponsored by the Geography Division, such as the MAF Geocoding Office Resolution (MAFGOR). In addition, the Geography Division uses the TIGER data base to produce maps for collection and dissemination of data, perform geocoding, provide geographic reference files, and prepare a variety of other products.

Although the TIGER data base has been modified over time, it remains in basically the same format and on basically the same platform as it was for the 1990 decennial census. The Geography Division's internal version of the TIGER data base is housed on a Compaq (digital alpha) computer in Bowie, Maryland. The information in the database is contained in approximately 3,200 distinct files called partitions. Each partition represents a segment of geographic information, usually based on a county. It is stored in a "homegrown" format and accessed only through Bureau-created software. Many branches in the Geography Division have large investments in in-house software that provides mission-critical functions using the TIGER data base. In addition, the internal Census TIGER processing is performed on a variety of systems (Compaq Open Virtual Memory System [VMS], Silicon Graphics IRIX, Windows).

1.2.3 Geographic Catalogue System

The Geographic Catalogue System (GEOCAT) is an ORACLE-based system that provides non-spatial reference information for every "named" geographic area. These "named" geographic areas can be either a legal entity such as a state, county, or city or a statistical entity, such as a census tract or designated urbanized area. The GEOCAT stores nonspatial information concerning these "named" geographic areas, such as functional status, land area, and population. In addition, the GEOCAT contains an historical thread, or audit trail, of all the boundary changes that have occurred over time to the legal geographic area.

The GEOCAT and TIGER geographic area lists are not integrated. To ensure consistency, the lists of legal and statistical geographic areas in the two systems are synchronized periodically through an electronic batch process.

1.2.4 Geographic Programs Participant System

The GPP System is an ORACLE-based system that supports the execution of Bureau programs by providing a tool to manage and coordinate the interaction between the Bureau and external participating organizations. The GPP establishes a relationship among the specific program, the geographic area targeted by the program, and the organization that is representing the geographic area that is participating in the program. Only legal geographic areas or tribal statistical areas are included in the GPP. The participating organization can be a local government or a pseudo-governmental organization, such as the Washington Regional Council of Governments.

The GPP contains information documenting the timeline and milestones of an organization's interaction with the Bureau and its participation in a specific program. Information such as when an organization was invited to participate, when the organization agreed to participate, when material was sent out, and when the updated materials were returned is captured in the GPP. The GPP is a key resource in the execution of critical Bureau programs such as the LUCA program, and the Boundary and Annexation Survey (BAS).

The GPP and GEOCAT geographic area lists are not integrated. The lists of legal geographic areas in the two systems are synchronized periodically through an electronic batch process to ensure consistency.

1.2.5 Production Control System

The PCS is a system that supports the execution of Bureau programs by providing a tool to manage and coordinate the maintenance of TIGER partitions. The PCS uses the same Bureau-developed database engine that supports both MAF and TIGER®.

The information stored in the TIGER system periodically requires review and adjustment. This review and adjustment is performed by various Bureau organizations throughout the country. Every significant Bureau program that involves maintaining the geographic information in the TIGER system has a specific set of activities that are tracked from initiation to completion within the PCS.

PCS provides an automated system that coordinates these activities, tracks the status of each TIGER partition, and analyzes the adjustments to ensure that data integrity is maintained. In addition, the PCS also coordinates the execution of other activities and processes that require the availability of the TIGER partitions and ensure that prerequisite activities have satisfactorily been completed before permitting a subsequent activity to be initiated.

The PCS and GEOCAT geographic area lists are not integrated. The lists of geographic areas in the two systems are synchronized periodically through an electronic batch process to ensure consistency.

The PCS receives date-related information from the GPP through an electronic batch process.

1.2.6 Problem Referral System

The Problem Referral System (PRS) is an ORACLE-based incident tracking system that serves as a repository for all incidents that are related to the address and geographic databases. Incidents of operational failure can be entered into the system automatically or entered manually by Bureau staff. The PRS provides Bureau staff with visibility into any potential system-related issues with the address and geographic databases that would affect the staff's work schedule. Once an incident has been resolved, the PRS is updated to reflect the current status of the incident and its resolution.

The PRS is tightly integrated with the PCS to support the automatic entry of operational failures in the PCS.

1.2.7 Windowing Index System

Windex is a TIGER DB-based spatial index of the TIGER partitions. The TIGER data that represent a geographic area can span multiple TIGER partitions. Windex contains the specific geospatial coordinates for determining the dimensional boundaries of geographic areas. This information indicates what TIGER partitions are necessary to achieve a complete representation of the targeted geographic area.

1.3 MAF/TIGER System Modernization

The MAF and TIGER systems have served the Geography Division and other Bureau customers well. However, modernization of the databases and the systems that support them will yield efficiency and accuracy that cannot be accomplished with the current processing system and database content. Therefore, as part of the 21st Century MAF/TIGER Enhancements initiative, the Geography Division is proposing a processing system modernization objective. This objective will allow the Bureau to expand the sources of data from which the database is built, forge stronger relationships with its partners, and provide higher quality products for internal and external customers. Although all programs that use geographic and address data will benefit from the initiative, some new programs depend directly on the success of the initiative and the ability of the Bureau to update the address database continually.

The Geography Division has developed budget estimates for its Geographic Support base program. The base program estimates allow periodic updating of the address list and geographic features that meet the minimal needs of mapping, reference file creation, and address listings. The base program funding will not support additional automation, improved accuracy of coordinates, or increased production needs. The base program has four components:

1. Activities that build/maintain the MAF of housing units
2. Continuing operations to link the MAF with the TIGER data base
3. Continuing partnerships with state, local, and tribal governments
4. Quality assurance for geographic support products.

To improve MAF and TIGER®, unfunded enhancements to the base program are included in the initiative. The proposed enhancements include the use of Global Positioning System (GPS) and aerial/satellite imagery technology to provide more precise location information that can correct the locations of all streets and other map features in the TIGER data base and also provide a correct location for every housing unit address in the MAF; use commercial software tools where none existed at the time the databases were created; encourage improved relationships with Bureau geographic program partners; and improve database quality in all areas.

The unfunded enhancements will enable the Geography Division to meet major goals. One major goal is improving the processing environment and using commercially available tools wherever feasible, economical, and practical. Another goal of the 21st Century MAF/TIGER Enhancements initiative is to improve accuracy of the database using aerial/satellite imagery or other sources to realign existing coordinates for features and existing housing units and place them in the TIGER data base. Improving the accuracy of the coordinates will streamline Bureau operations by allowing GPS technology to gather accurate housing unit location information in relation to an accurate base map, which will enable enumerators to relocate those structures in the field, when necessary, for follow-up operations. In addition, a more accurate base map will alleviate a publicly perceived problem with census geographic data and also facilitate digital data exchange with federal, tribal, state, local, and private partners. Sharing of geospatial data will improve relationships with partners and ultimately enhance the Bureau's geographic programs.

1.4 Objectives of the 21st Century MAF/TIGER Enhancements

During interviews with Geography Division management, the following objectives were identified and associated with the four initiative components:

1. Correctly locate every street and other map feature in the TIGER data base, each MAF address, and implement an effective automated feature change detection methodology—

   Integrate the address list, geographic update, and data collection activities
   Improve feature change detection
   Enhance the accuracy of geographic features
   Automate field data collection
   Add geographic coordinates for building structures.

2. Implement a modern processing environment for the MAF/TIGER system that will support rapid application development, allow extensive use of commercially available software, and reduce staff training time—

   Modernize the processing environment
   Identify new technology and commercial software useful to the system.

3. Expand and encourage geographic partnership programs with state, local, and tribal governments that update the MAF/TIGER data base, including a "rolling" LUCA to maintain a current address list for use by all Bureau censuses and surveys and computer-based updates of governmental unit boundaries and other geographic areas—

   Promote and expand partnerships with data providers
   Identify new address and geographic data sources.

4. Implement a comprehensive plan for periodic MAF/TIGER evaluation and corrective activities that will guide planning for cost-effective future coverage and geocoding improvement operations—

   Improve the quality of geographic products
   Implement a MAF/TIGER quality assurance program
   Institute a field operation to evaluate the quality of MAF/TIGER (e.g., maps, lists, questionnaires) with electronic products.

1.5 Standards and Guidelines

The 21st Century MAF/TIGER Enhancements initiative represents a significant investment in information technology for the U.S. Census Bureau. The Bureau will follow a management strategy that supports the information technology management improvement goals of the Clinger-Cohen Act of 1996, the Paperwork Reduction Act of 1995, Office of Management and Budget (OMB) Circular A-16, OMB Circular A-119, Executive Order (E.O.) 12906, and related guidance provided by the Department of Commerce.

Standards for geographic and spatial data have begun to emerge over the past few years. However, many of them are not yet mature or accepted by the industry, many are still in the discussion stage, and standards for many important topics have not yet begun to be developed. Regardless of the state of development of standards, it is in the interest of all federal government agencies, the U.S. Census Bureau in particular, to follow the policy of OMB Circular A-119 to comply with these emerging and voluntary standards. Failure to do so will compromise the ability of the Bureau to share data with other organizations in the future. The Geography Division should continue to monitor the development of geospatial data standards and participate in their development when appropriate. The primary organizations involved in the development of these standards are the Federal Geographic Data Committee (FGDC); the International Standards Organization, Technical Committee 211 (ISO/TC 211); and the American National Standards Institute (ANSI). In addition to the formal standards bodies, the Open GIS Consortium (OGC) is attempting to establish best practices for geospatial software developers to follow, which may become de facto standards in the future. Again, the Bureau should continue to monitor these activities and participate as appropriate.

1.6 Study Organization

The MAF/TIGER Modernization Study has been organized into three chapters and includes eight appendixes. A brief description of each section follows:

Chapter 1: Introduction and Background—Provides an introduction to the MAF/TIGER Modernization Study, including a high-level review of the challenges currently faced by the U.S. Census Bureau, a review of its overarching goals, and a review of the various geographic databases targeted by the initiative.

Chapter 2: Requirements and Objectives—Presents, at an aggregate level, the requirements that the MAF/TIGER Modernization Team collected as a result of extensive interviews and facilitated discussions with both U.S. Census Bureau personnel and external stakeholders.

Chapter 3: Implementation Scenarios—Suggests a number of potential implementation scenarios for each of the four objectives of the U.S. Census Bureau's 21st Century MAF/TIGER Enhancements initiative.

Appendix A: Requirements for the MAF/TIGER Modernization Study—Provides the data that was the basis for the information provided in Chapter 2. Contains a detailed analysis of the more than 300 requirements that were extracted from interviews and facilitated sessions with both U.S. Census Bureau personnel and external stakeholders. The requirements are arrayed across 7 major requirement categories.

Appendix B: Mobile Computers and Global Positioning System—Provides information on the current capabilities and technology that support mobile (portable/hand held) computers and GPS applications.

Appendix C: Imagery Assessment—Presents a discussion of the current state of aerial and satellite imagery and its potential use in connection with improving the accuracy of the geographic locations of the natural features, manmade features, and existing structures within the MAF/TIGER data bases.

Appendix D: Current System Engineering Best Practices—Provides a review of current system engineering best practices. Included is a review of methodologies and practices employed to support the migration of large-scale, heterogeneous legacy applications to an integrated and homogeneous environment. Highlighted are those practices that improve programmer productivity, reduce application development time, and increase the quality of the systems produced.

Appendix E: Commercial Geospatial Software—Discusses commercially available software that potentially could provide the required tools and functionality to support the major geographic information system requirements of the U.S. Census Bureau.

Appendix F: Commercial Database Software—Presents key considerations and best practices for developers implementing large-scale database applications supporting geographically dispersed users. Included is a review of current database products and their application development tools sets.

Appendix G: Internet-Enabled Applications—Presents key considerations and best practices for developers implementing Internet-based applications. Discussed are issues such as security, capacity, and remote access.

Appendix H: Acronyms and Abbreviations

2. REQUIREMENTS AND OBJECTIVES

The MAF/TIGER system is an aging national resource. It has been used to support not only the various censuses and household surveys managed by the U.S. Census Bureau, but also as the foundation of the burgeoning Geographic Information System (GIS) industry in the United States. Although the MAF/TIGER system has been able to meet, and often exceed, most of the Bureau's geographic requirements up through Census 2000, operational managers throughout the Bureau have expressed concerns that the Bureau cannot successfully conduct the 2010 census without significant enhancements to the system.

To document these concerns the Booz·Allen & Hamilton MAF/TIGER Modernization Study team conducted interviews and facilitated discussion groups with approximately 70 MAF/TIGER users and stakeholders (see Table 2-1). These users and stakeholders represented the Bureau's data collection and data analysis organizations; state, local, and tribal governments; and other federal agencies. The interviews and facilitated discussions were held between November 1999 and May 2000. Please refer to Appendix A for additional information concerning the organizations that participated in the interviews and facilitated discussions.

Table 2-1. Organizational Areas Interviewed

ORGANIZATIONS	NUMBER OF INTERVIEWS/ DISCUSSIONS
U.S. Census Bureau Customers	
Geography Program Managers	
Field Operations Managers	4
	14
External Users/Customers	

After all of the interviews and facilitated discussions were completed, the Booz·Allen &

captured to facilitate the extraction of general requirements. All comments and opinions concerning a particular MAF/TIGER subject area were grouped together and combined where

Technical, functional, administrative, and organizational areas were covered by the comments and opinions. The results of the interviews and facilitated discussions were then

a result, 354 general MAF/TIGER requirements were identified. It is important to note that these general MAF/TIGER requirements are not intended to represent detailed functional requirements

current MAF/TIGER system feel are not being satisfied.

Once the general requirements were extracted, the
Modernization Study team analyzed them. Based on this analysis, the team grouped the 354 general MAF/TIGER requirements into seven high-level categories. Table 2-2 provides a

in each high-level category, and the sources that identified the requirement.

Table 2-2. High-Level Requirement Categories

HIGH-LEVEL REQUIREMENT CATEGORY DESCRIPTION	NUMBER OF RELATED GENERAL REQUIREMENTS	SOURCE OF REQUIREMENTS				
		U.S. Census Bureau Customers				External Users/Customers
		Geography Division Program Managers	Field Operations Managers	Program Sponsors/Analysts	Census Documents	
Maintain an accurate and complete inventory of all residential, group quarters, and commercial building structures, along with the address and location for each, and the within-structure designations when a structure contains more than one unit.	44	✓	✓	✓	✓	✓
Maintain an accurate and complete inventory of legal, administrative, and statistical geographic areas, their attributes, and interrelationships.	29	✓	✓	✓	✓	✓
Maintain an accurate and complete inventory of road, rail, water, landmark, power transmission line, and other features mapped for Bureau purposes.	31	✓	✓	✓	✓	✓
Make use of geospatial data provided by federal, state, local, and tribal government partnerships, as well as private-sector sources, wherever possible.	48	✓	✓	✓	✓	✓
Provide geographic products and other support services to Bureau programs and census data users.	56	✓	✓	✓	✓	✓
Comply with applicable geographic data standards, guidelines, and directives and meet the Federal Geographic Data Committee (FGDC) objectives for the National Spatial Data Infrastructure (NSDI) initiative, where appropriate.	3	✓	✓	✓	✓	✓
Improve the accuracy and responsiveness of the geographic support system to meet the needs of Bureau program managers, geographic program participants, and U.S. Census Bureau data product users.	191	✓	✓	✓	✓	✓

The general MAF/TIGER requirements were further reviewed to determine which of the Bureau's 21st Century MAF/TIGER Enhancements initiative objectives satisfied the requirement. Table 2-3 identifies by high-level category, which of the Bureau's 21st Century MAF/TIGER Enhancements initiative objectives will contribute to addressing the requirements.

Table 2-3. Objectives Addressed By High-Level Requirements

HIGH-LEVEL REQUIREMENT CATEGORY DESCRIPTION	NUMBER OF RELATED GENERAL REQUIREMENTS	Base Program	21ST CENTURY MAF/TIGER ENHANCEMENTS: MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality
Maintain an accurate and complete inventory of all residential, group quarters, and commercial building structures, along with the address and location for each, and the within-structure designations when a structure contains more than one unit.	44	◑	✓		✓	✓
Maintain an accurate and complete inventory of legal, administrative, and statistical geographic areas, their attributes, and interrelationships.	29	◑	✓		✓	✓
Maintain an accurate and complete inventory of road, rail, water, landmark, power transmission line, and other features mapped for Bureau purposes.	31	◑	✓		✓	✓
Make use of geospatial data provided by federal, state, local, and tribal government partnerships, as well as private-sector sources, wherever possible.	48	◑	✓	✓	✓	✓
Provide geographic products and other support services to Bureau programs and census data users.	56	◑		✓	✓	
Comply with applicable geographic data standards, guidelines, and directives and meet the Federal Geographic Data Committee (FGDC) objectives for the National Spatial Data Infrastructure (NSDI) initiative, where appropriate.	3	◑				✓
Improve the accuracy and responsiveness of the geographic support system to meet the needs of Bureau program managers, geographic program participants, and U.S. Census Bureau data product users.	191	◑	✓	✓	✓	✓

The remainder of Chapter 2 provides additional details concerning each of the high-level requirement categories.

2.1 High-Level Requirement Categories

The seven high-level requirement categories are discussed below.

2.1.1 High-Level Requirement #1—Maintain An Inventory Of All Structure Addresses, Within-Structure Designations, And Location Descriptions

Maintain an accurate and complete inventory of all residential, group quarters, and commercial building structures, along with the geographic location for each, and the within-structure designations when a structure contains more than one unit.

The U.S. Census Bureau must have a database that contains a record for each location for which the Bureau should collect data. For demographic censuses and surveys, these places would include housing unit building structures and also a wide variety of other locations used as primary residences, referred to as special places and group quarters. These special places and group quarters include college dormitories, hospitals, prisons, monasteries, nunneries, half-way houses, homeless shelters. Less permanent habitations, such as common areas where homeless people tend to congregate (e.g., tents, cardboard boxes), are not included in the database but are visited in special operations during censuses and surveys. For economic censuses and surveys, the Bureau has determined that it will not use the MAF/TIGER database as its source for nonresidential addresses. There is, however, a strong need for the nonresidential addresses in demographic data collection. The status of a structure may frequently change between residential and nonresidential, and frequently a mix exists within a single structure. Leaving nonresidential addresses out of the MAF/TIGER database results in an incomplete list of addresses, which can lead to uncertainty about the accuracy of the list. The incomplete list can also make it very difficult for field workers and local/tribal governments to determine which units are nonresidential and therefore not to be included in the demographic census or survey. Also, a list of nonresidential structures is needed to support demographic data collection related to the workplace.

The major change proposed by the 21st Century MAF/TIGER Enhancements initiative is to base the residential and nonresidential inventory on building structures that are visible on the ground or in aerial/satellite imagery rather than lists of mailing addresses as has been done in the past. Mailing addresses are not a reliable source from which to determine a geographic location. Many mailing addresses cannot be geolocated (box numbers, rural routes), addressing schemes vary throughout the country, names of streets change or have alternates, house numbers are revised by local jurisdictions, and ZIP Codes are changed frequently by the U.S. Postal Service. Having a list composed of the coordinate location of each structure with the mailing address as an attribute will solve many of the problems of the past.

COMPONENT	BENEFITS AND IMPROVEMENTS
Maintain a mailing address for each structure.	Reduce duplicate addresses by basing the inventory on structures visible in imagery or on the ground rather than on mailing addresses. Reduce the need for massive field operations, such as address listing and block canvassing through the use of aerial or satellite imagery and the current MAF. Improve coverage by using imagery and field visits.
Maintain the geographic location of each structure by collecting a latitude/longitude coordinate.	Make structures in the field easier to find Tabulate data more accurately. Use GPS to locate structures.
Maintain the structure inventory by regular matching to other address list data files, such as USPS Delivery Sequence File, administrative records of the Social Security Service and other federal agencies, and files from local/tribal program participants.	Identify new and missing structures through comparisons with sources outside of the Bureau. Use the results of matching for targeting areas for research through field operations or imagery.
Maintain a link to a street segment for each structure.	Make structures in rural areas easier to find.

2.1.2 High-Level Requirement #2—Maintain an Inventory of Legal and Statistical Geographic Areas

Maintain an accurate and complete inventory of legal, administrative, and statistical geographic areas, their attributes and interrelationships.

The Bureau must tabulate and report the data it collects for a wide variety of geographic areas. These include legal governmental areas, such as states, counties, cities, American Indian reservations; statistical areas, such as census tracts, census blocks, urbanized areas; and administrative areas, such as school districts, voting districts; and even internal administrative areas for census field work, such as regions, local census offices, and individual assignment areas. The current TIGER and GEOCAT data bases manage more than 9,000,000 geographic areas within 32 types of geographic activities. Detailed current boundaries are maintained along with identifying codes, names where appropriate, various status codes, relationships to other geographic areas, relationships to physical features, population and housing counts, and land and water areas.

In addition to the data maintained in the past, the future MAF/TIGER data base should manage changes in geographic boundaries and their descriptions over time, provide “boundary flags” that control the rules for changing boundaries, and maintain nonspatial relationships between geographic areas, such as place/minor civil division (MCD) dependency, coextensivity, American Indian Reservation—trust land relationships.

COMPONENT	BENEFITS AND IMPROVEMENTS
Maintain accurate boundaries of all geographic areas including a record of their changes over time.	Eliminate the need for historic or “benchmark” copies of TIGER to preserve past information. Improve update operations by maintaining “boundary flags” to direct the changes of related geographic areas. Improve boundary accuracy by including line work from the Public Land Survey System (PLSS) in MAF/TIGER.
Maintain attribute information about each geographic area including names, identification codes, area type classification, governmental status.	Allow for varying geographic area definitions and relationships in different parts of the country instead of forcing all states into the same model.
Maintain nonspatial relationships between geographic areas.	Provide more accurate control of geographic coverage in dependency and coextensive relationships.
Eliminate need for separate “collection geography.”	Because building structures can be automatically allocated to any polygon based on the coordinate location of the structure, final tabulation census blocks can be generated by computer as soon as the final boundaries are available. Thus, there is no need for a separate set of blocks for collection purposes or massive “block split” operations.
Impose standards for updating the database and create definitions for feature types.	Eliminate the practice of accepting the most recently submitted change as the official version of the data or subjective interpretation of feature types.

2.1.3 High-Level Requirement #3—Maintain an Inventory of Road, Rail, Water, Landmark, Power Transmission Line, and Other Features

Maintain an accurate and complete inventory of road, rail, water, landmark, power transmission line, and other features mapped for U.S. Census Bureau purposes.

The current TIGER database maintains all street features along with their names, address ranges, and ZIP Codes where available. Water and rail features were derived from the United States Geological Survey (USGS) Digital Line Graphs in rural areas and have been maintained only as needed for Bureau purposes. Within metropolitan areas, water and rail features were derived from Geographic Base File/Dual Independent Map Encoding (GBF/DIME) files and updated as needed in response to Bureau operations. The overall positional accuracy of current TIGER features is in the 50–100 meter range.

In addition to the data above, the future system needs to contain a coordinate for every building structure, the lines from the Public Land Survey System (PLSS), more complete and more accurate hydrographic and rail features in urban areas; cadastral-level detail, such as street right-of-way and property lines in selected areas where needed to represent complex corporate boundaries and corridors; and all USPS geographic concepts, such as carrier routes and ZIP+4 Codes. Coordinates for all features must be corrected to within 4–8 meters of ground truth.

COMPONENT	BENEFITS AND IMPROVEMENTS
Maintain information about each feature, such as feature type classification, coordinate location, names and postal geography.	Expand geographic data from USPS and keep more current. Allow flexibility to represent unusual or alternative mailing addresses systems, such as Queens, Salt Lake City, Puerto Rico.
Develop and impose a set of standards for feature classification and attribute content.	Maintain the feature classification more simply and accurately, which will correct many problems with misleading symbolization on mapping products and reduce confusion in the field.
Include features from adjacent countries, such as Canada, Mexico, and Russia, near the borders of the United States.	Improve mapping products by providing continuity of features that cross international boundaries.
Provide adequate coverage of water features in former GBF/DIME file areas.	Improve accuracy of boundaries that must follow poorly defined water features in urban areas.
Perform regular geocoding of the mailing addresses in the structure inventory to ensure the accuracy and completeness of the ZIP Codes, street names, and address ranges associated with each street segment.	Improve the accuracy of the postal geography in MAF/TIGER.
Impose standards for updating and provide precise metadata.	Eliminate the practice of accepting the most recently submitted change as the official version of the data.

2.1.4 High-Level Requirement #4—Use Geospatial Data Provided by External Sources

Make use of geospatial data provided by federal, state, local, and tribal government partnerships, as well as private-sector sources, wherever possible.

Many state, local, and tribal governments along with other federal agencies have collected data that is potentially very useful to the Bureau. The current methodology of gaining access to those data relies primarily on paper forms and maps and simple computer text files. The Geography

Division has developed an automated process for exchanging some geographic data (the digital exchange process), but that system requires expert manual intervention and is limited in the types of data that can be retrieved.

Participant programs will be improved to allow Internet access to MAF/TIGER data and expanded electronic file submissions as well as continuing the traditional paper form and map methodology.

COMPONENT	BENEFITS AND IMPROVEMENTS
Expand the use of geospatial data holdings from state, local, tribal governments, and other sources.	Reduce costs and improve accuracy by taking advantage of local geographic and address list knowledge.
Expand the use of electronic data exchange between the Bureau and program participants.	More accurate feature coordinates will allow automated matching of features between partner databases and MAF/TIGER.
Provide the ability for program participants to directly review and update MAF/TIGER.	Reduce complaints from local/tribal partners about delay in feedback by providing direct access to MAF/TIGER on the Internet.

2.1.5 High-Level Requirement #5—Provide Geographic Products and Support Services

Provide geographic and address list products and other support services to U.S. Census Bureau programs and census data users.

To meet the needs of Bureau programs and data-user requirements, the Geography Division produces millions of paper maps, geocodes hundreds of millions of addresses, and creates a wide variety of data file extracts. Operations to provide these products consist of interactive update and database maintenance activities in the regional offices and the National Processing Center (NPC); cooperative partnership programs with state, local, and tribal governments; and massive batch computer processing supported by a highly automated work flow management system known as the Production Control System (PCS).

Future product and support services will need to be expanded to send and receive data files to portable computers being used in field operations, provide on-demand data viewing and downloading capabilities throughout the Bureau, and provide data products via the Internet.

COMPONENT	BENEFITS AND IMPROVEMENTS
Support operational and data user needs with mapping products.	Reduce the need for paper maps through the increased use of computer displays of geographic information.

COMPONENT	BENEFITS AND IMPROVEMENTS
Expand and improve geocoding services.	Provide a National Geocoding System that is capable of geolocating any mailing address through a mix of automated and manual techniques. Provide interactive, on-demand geocoding for single addresses.
Improve data file extracts.	Provide spatial and nonspatial data in formats that are commonly used with commercial software packages.
Provide ability to tabulate census data to any geographic area boundary.	With a coordinate location for building structures available, data tabulations can be done for any geographic area whose boundary is known. The data tabulation is performed on all housing units that fall within the boundary, and the relationship to census blocks is irrelevant.
Provide better MAF/TIGER viewing and browsing capabilities.	Provide on-demand viewing and data file downloads to data users inside the Bureau.

2.1.6 High-Level Requirement #6—Meet Requisite Standards

Comply with applicable geographic data standards, guidelines and directives, and meet the Federal Geographic Data Committee (FGDC) objectives for the National Spatial Data Infrastructure (NSDI) initiative, where appropriate.

As mentioned in Chapter 1, standards for geographic and spatial data have begun to emerge over the past few years. However, many of them are not yet mature or accepted by the industry, many are still in the discussion stage, and standards for many important topics have not yet begun to be developed. Regardless of the state of development of standards, it is in the interest of all federal government agencies, the U. S. Census Bureau in particular, to follow the policy of the Office of Management and Budget (OMB) Circular A-119 to comply with these emerging and voluntary standards. Failure to do so will compromise the ability of the Bureau to share data with other organizations in the future. The Geography Division should continue to monitor the development of geospatial data standards and participate in their development when appropriate. In addition to the formal standards bodies, the Open GIS Consortium (OGC) is attempting to establish best practices for geospatial software developers to follow, which may become de facto standards in the future. Again, the Bureau should continue to monitor these activities and participate as appropriate. It must be noted that there are mandatory directives such as E.O. 12906 with which the Bureau must comply.

COMPONENT	BENEFITS AND IMPROVEMENTS
FGDC sponsored and recommended standards.	Content Standard for Digital Geospatial Metadata. Spatial Data Transfer Standard. Cadastral Data Content Standard. Geospatial Positioning Accuracy Standard.
Other standards	The International Standards Organization Technical Committee 211 (ISO/TC 211) is in the process of developing standards for many major topics in geospatial data.

2.1.7 High-Level Requirement #7—Improve Accuracy and Responsiveness of Geographic Support

Improve the accuracy and responsiveness of the geographic support system to meet the needs of U.S. Census Bureau program managers, geographic program participants, and U.S. Census Bureau data product users.

While both the TIGER and MAF databases were major innovations in the past, both are in need of very significant improvements to meet the needs of the future and solve problems of the past. Modernizing the database technology for MAF/TIGER can solve many operational problems of the past, such as multiuser access and availability of TIGER partitions. Commercial software and rapid application development techniques will reduce the amount of time and labor required for maintenance and product delivery. Use of GPS and portable computers can reduce labor in the field as well as in large office operations. Improving the accuracy of coordinate data will allow expanded partnership programs and reduce errors in many census and survey activities.

COMPONENT	BENEFITS AND IMPROVEMENTS
Create an integrated geographic support database.	Integrating MAF, TIGER, GEOCAT, GPP, PCS, PRS, and Windex will eliminate redundant and inconsistent data and the associated problems.
Use modern database technology.	Provide a seamless database of the entire United States without partitions; thus, eliminating long waits for access to partitions. Provide multiuser access as needed in update operations. Allow for multiple versions of the data rather than creating benchmark versions of the entire database.

COMPONENT	BENEFITS AND IMPROVEMENTS
Modernize the application development process.	Allow commercial GIS Packages to replace the functionality of custom-developed MAF/TIGER applications, where appropriate. Use other commercial software such as database report tools, spreadsheets, where possible. Use a component architecture for the development of all software modules to speed application development and allow for gradual system evolution. Attract programmers who will be able to further their careers with cutting-edge software.
Provide for computer-assisted field operations.	Provide computer-assisted navigation to lead field representatives to their assignments. Use portable computer displays instead of paper maps in field operations to provide more complete information. Use portable computers in field operations to provide automated error detection. Use GPS to locate structures for follow-up operations and to locate new structures and features. Avoid costly office operations by transmitting updates directly from the field rather than shipping, keying, and digitizing paper products. Use GPS for quality assurance of private sector and partner geospatial data.
Use aerial and satellite imagery for data maintenance.	Use imagery to correct the location of features and locate building structures. Use the automated change detection capability of imagery to target areas of growth and change.
Expand and improve electronic digital exchange.	Provide the ability to do automated matching and data transfer with other geospatial databases based on geometry.

3. IMPLEMENTATION SCENARIOS

In response to the feedback received during the interviews and facilitated discussions described in Chapter 2, the Booz·Allen & Hamilton MAF/TIGER Modernization Study team has developed a number of suggested scenarios that would potentially satisfy the identified general requirements of MAF and TIGER®. These suggested scenarios are directly related to the four objectives of the U.S. Census Bureau's 21st Century MAF/TIGER Enhancements initiative. The proposed scenarios include the use of GPS and aerial/satellite imagery technology, the use of commercially available database systems and associated software tools, the increased use of the Internet and electronic data exchanges, and the strengthening of existing quality assurance programs for the Bureau's geographic products and systems. The suggested scenarios were developed with sensitivity to both cost and schedule. Specifically, the scenarios provide approaches using different degrees of technology and contractor support, which influence the potential cost and the scenario's comprehensiveness in satisfying the objective.

In support of the suggested scenarios, a potential implementation approach has also been developed by the Booz·Allen & Hamilton MAF/TIGER Modernization Study team. This chapter provides a high-level description of these scenarios and their respective implementation approaches.

3.1 Objective 1—MAF and TIGER Data

Objective 1 of the 21st Century MAF/TIGER Enhancements initiative is as follows:

> Correctly locate every street and other map feature in the TIGER data base, each MAF address, and implement an effective automated feature change detection methodology.

For the Bureau's geographic databases to fully satisfy census requirements, the geographic locations of features and structures must be accurate enough to effectively locate them, and the linkage between the MAF and TIGER systems must achieve 99 percent accuracy.

There are approximately 118 million housing units in the United States, and approximately 87 percent of these housing units are located on just 10 percent of the 3.8 million square miles that make up the United States. Clearly, a critical success factor in meeting the requirements for the Bureau's geographic databases is to focus its resources in these densely populated areas.

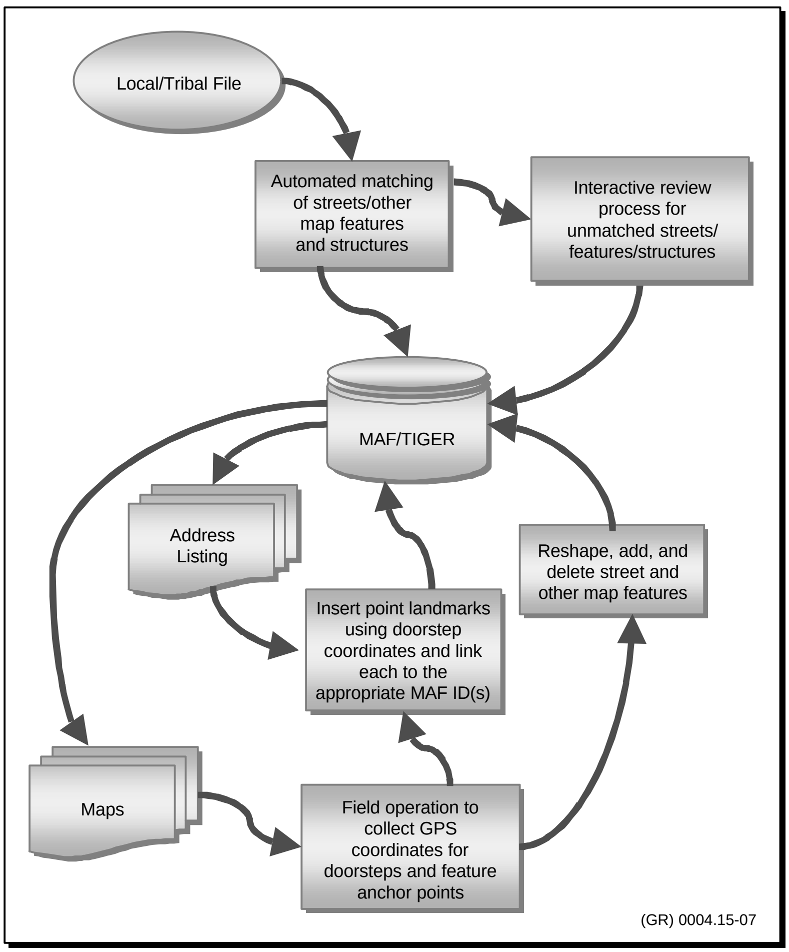
To support the Geographic Support base program, the Bureau would continue to update geographic features and structures periodically in the MAF and TIGER data bases with input from local and tribal governments and from information received as a result of the Bureau's field operations conducted in support of other programs.

In support of Objective 1, three scenarios were developed for validating and correcting the geo-locations of the geographic features and structures in the TIGER data base.

3.1.1 Scenario 1—Local Government Files and Field Operations

The first scenario proposes a number of activities that, when combined, will improve the accuracy of the location information in the TIGER data base. A key activity would be using local government files, where available, to more accurately depict the position of geographic features, update postal geography, and insert geographic locations of structures into the database. Another critical step in successfully implementing this scenario would be the Bureau conducting a specialized field operation to collect the geographic location of geographic features and of building structures in areas where local government files are not available. The GPS and mobile computer technology would be used in support of this field operation (see Appendix B—Mobile Computers and Global Positioning System). Figure 3-1 provides a graphical representation of the implementation of Scenario 1.

Figure 3-1. Implementation of Scenario 1 for Objective 1

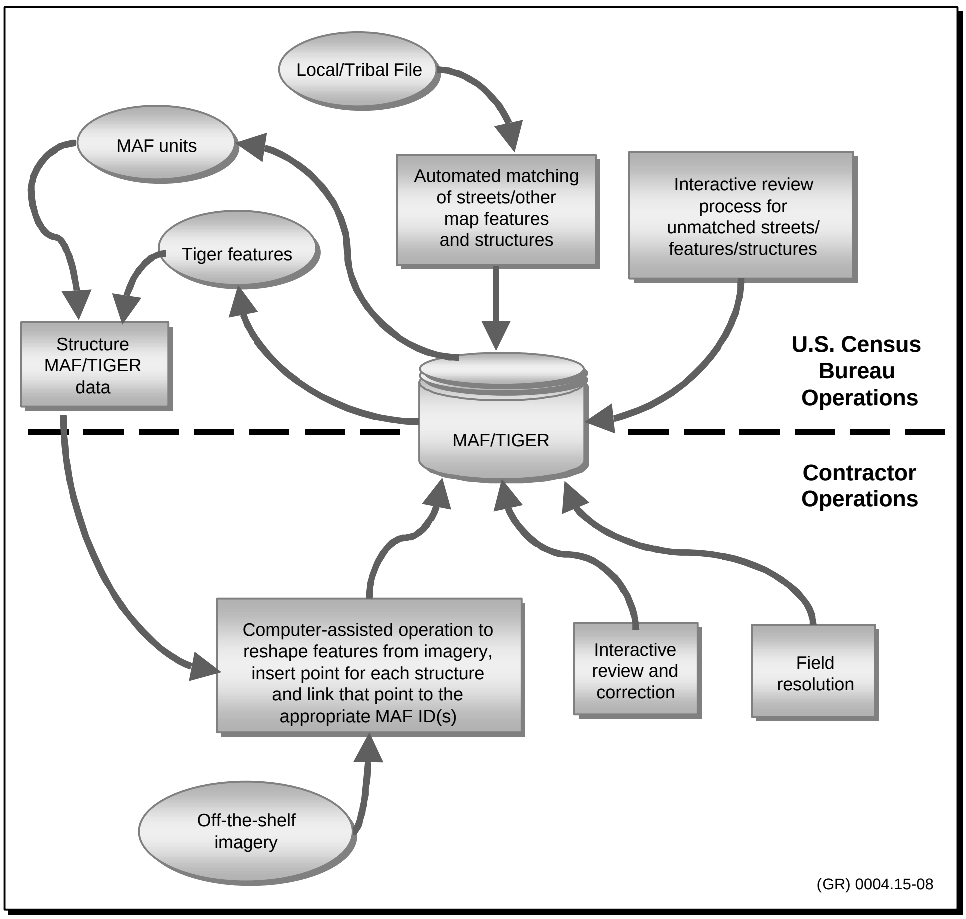


3.1.2 Scenario 2—Off-the-Shelf Precision Imagery

The second scenario would also use data obtained electronically from local governments but would introduce the use of high-precision imagery in correcting the locations of the geographic features and structures identified in the TIGER data base (see Appendix C—Imagery Assessment). Off-the-shelf high-precision imagery would be employed to provide images of high-growth areas in the United States while older high-precision imagery would be used to provide images of the more stable but densely populated areas of the United States. The high-precision imagery would be the basis for adjusting the locations of features in the TIGER data base and adding and deleting features. Contractor support would be required for extracting features and structures from the high-precision imagery and in reconciling the alignment of the extracted features and structures with the TIGER data base.

For areas for which the use of high-precision imagery would not be cost effective, such as "small settlements" in rural areas, less expensive alternatives would be employed. For small settlements, a field operation using GPS-equipped vans would be initiated to collect structure locations and anchor points. For rural areas, older, less precise imagery would be used to correct feature positions and locations of structures. Figure 3-2 provides a graphical representation of the implementation of Scenario 2.

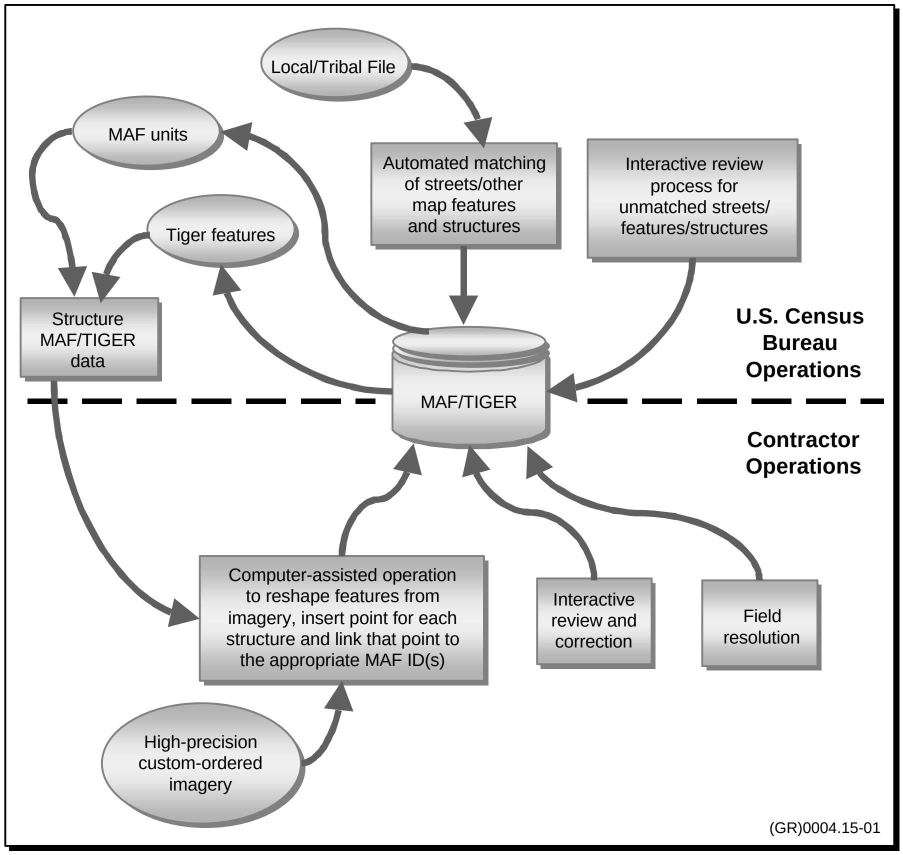
Figure 3-2. Implementation of Scenario 2 for Objective 1



3.1.3 Scenario 3—Custom-Ordered High-Precision Imagery

The third scenario for Objective 1 is virtually the same as Scenario 2, but this third scenario would use a combination of off-the-shelf high-precision imagery and custom-ordered high-precision imagery for adjusting the locations of features in the TIGER data base and adding and deleting features. Figure 3-3 provides a graphical representation of the implementation of Scenario 3.

Figure 3-3. Implementation of Scenario 3 for Objective 1



3.2 Objective 2—New Processing Environment

Objective 2 of the 21st Century MAF/TIGER Enhancements initiative is as follows:

> Implement a modern processing environment for the MAF/TIGER system that will support rapid application development, allow extensive use of commercially available software, and reduce staff training time.

As an organization's information technology assets increase, more attention is focused on the processes and technology used to build and maintain those assets. When systems are small and support only a small portion of an organization's mission, it is possible to consider redesigning and replacing those systems that no longer adequately satisfy the business requirements or have become technologically obsolete. Systems that significantly support the mission of an organization are substantial investments, and their replacement would introduce a significant level of risk to an organization's operation. Increasingly, software that supports the core business processes of an organization is being viewed as an asset that represents an investment that grows in value. Therefore, a major challenge facing organizations today is how to build software assets whose technical design and environment support functional modularity, seamless integration, and architectural flexibility (see Appendix D—Current System Engineering Best Practices).

As evidence of the increasing importance of software to organizations, attention is being focused on architecture and the systems that support the core business processes of an organization. Just as manufacturing industries have found they can leverage their asset base by restricting the amount of variation in their products, system developers are learning how to design and build software to serve multiple purposes by controlling and managing the amount of variation allowed in the software. They have learned that a core set of software assets can be leveraged to produce flexible inventories of software applications that can readily evolve to meet the changing requirements of an organization.

In the recent past, traditional software development practices have produced systems that were custom developed as one-of-a-kind products. Integration between these types of systems can be effectively implemented only if the underlying business requirements are clearly understood. When systems can be successfully integrated, an organization has an information technology asset that effectively supports the targeted core business process or organizational mission.

The core business processes supported by the Bureau's geographic databases require information technology assets that are both flexible and reliable. The current geographic databases used by the Bureau have adequately supported the core business processes but were designed and constructed as separate systems integrated through the use of links and pointers between related data sets. This design was established as a set of systems that require significant maintenance and locally acquired system knowledge.

The MAF, TIGER, and PCS systems use the TIGER data base that was developed locally by the Bureau. The TIGER data base, developed for the 1990 census, provided the Bureau with a database management system that exceeded all expectations. In accordance with the SPAD report, the Bureau chose to develop the TIGER data base because no commercial alternative was available that supported the Bureau's geospatial requirements. Since the initial development of the TIGER data base, the geospatial capabilities of commercially available relational database management systems (RDBMS) have increased substantially (see Appendix E—Commercial Geospatial Software). In addition, the performance of current hardware platforms can easily compensate for any performance degradation that may result from any system overhead related to a relational model implementation (see Appendix F—Commercial Database Software).

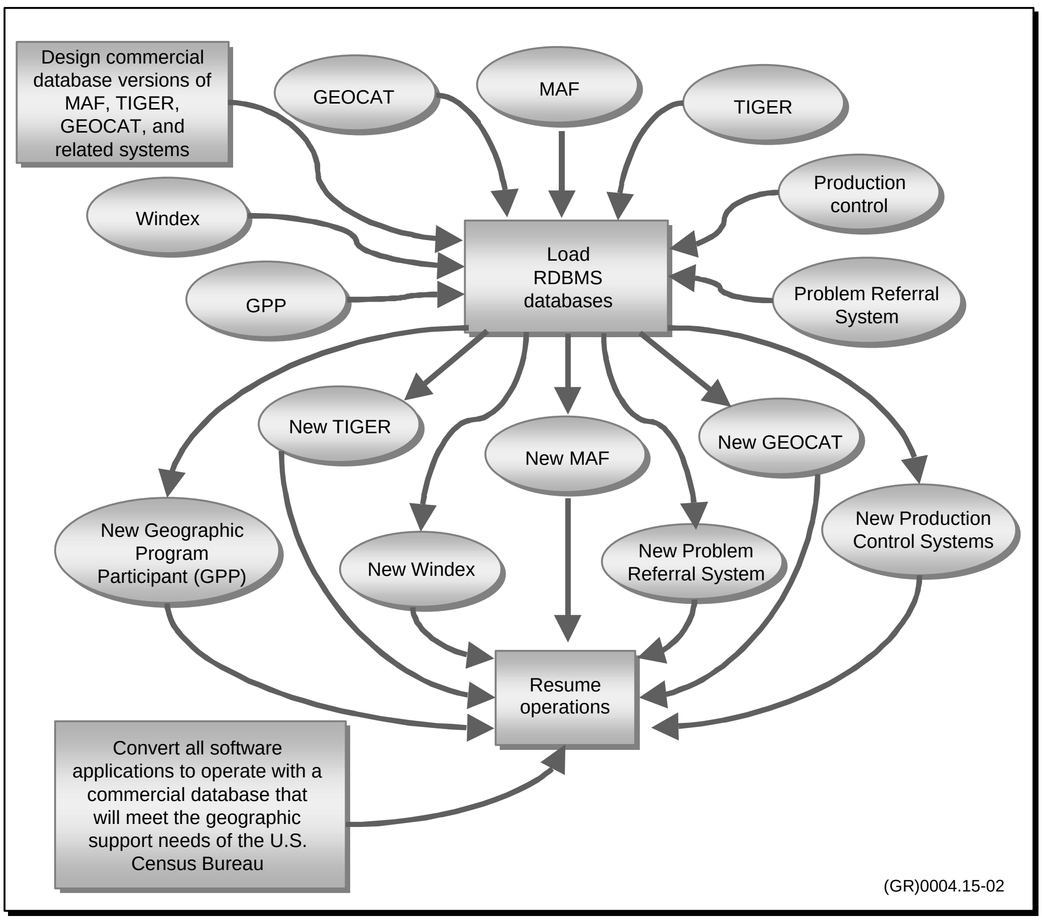
To support the Geographic Support base program, the Bureau would maintain the geographic databases and their associated application software as required to support the business processes of the Bureau.

To improve the overall effectiveness of the geographic databases in supporting the Bureau's core business processes, three scenarios have been developed. The scenarios are based on the migration of the TIGER data base to a commercially available RDBMS. The three scenarios differ in the extent to which the databases are redesigned and the application software associated with the various geographic databases is adjusted or replaced.

3.2.1 Scenario 1—Database Migration

In the first scenario, the MAF, TIGER, and PCS databases along with Windex and the PRS would be migrated to a commercially available RDBMS maintaining the current structure and data content of the two databases. The main effort required to successfully complete this scenario would be adjusting the respective system's application software to use the target RDBMS's application program interfaces (API) instead of the APIs that were previously used with the TIGER data base. The level of effort to complete this readjustment would be significant. Every application program would require individual inspection to determine where and how the program interacts with the TIGER data base and how it will interact with the targeted RDBMS. Once this analysis is completed, the application program would have to be modified, tested, and staged for implementation. The methodology to accomplish this scenario has characteristics that are similar to the effort required to prepare application software for the year 2000 transition. Figure 3-4 provides a graphical representation of the implementation of Scenario 1.

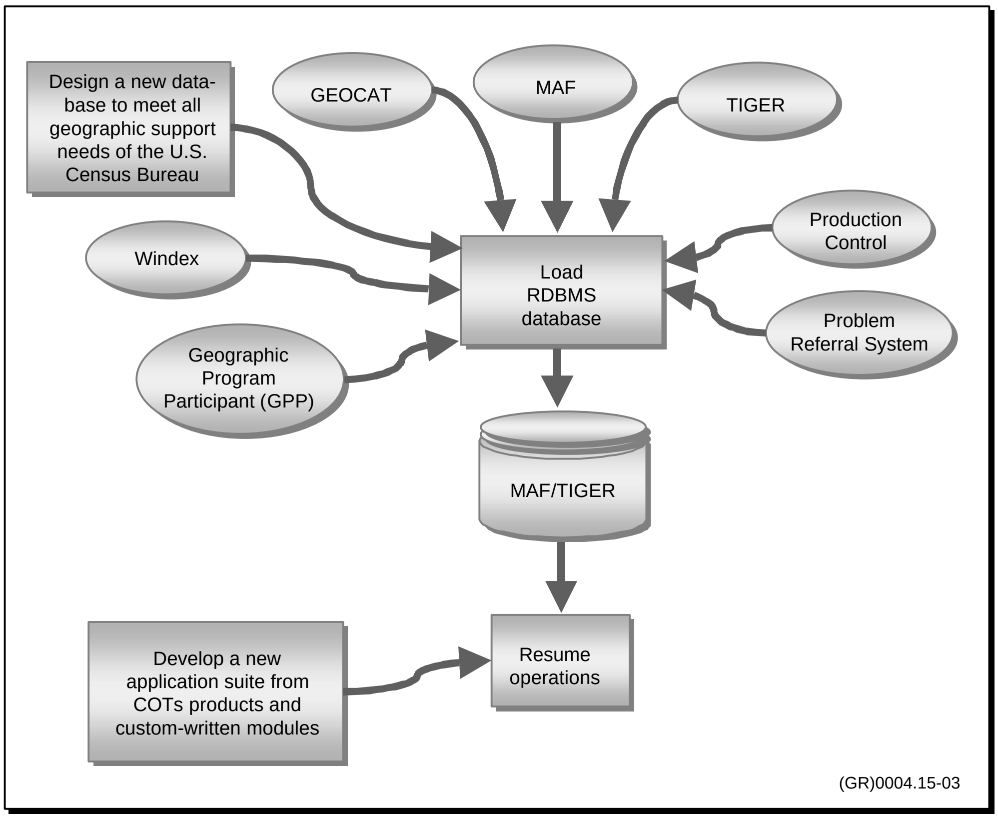
Figure 3-4. Implementation of Scenario 1 for Objective 2



3.2.2 Scenario 2—System Redesign

In the second scenario, the geographic databases would be migrated to a commercially available RDBMS, but the current data structure and content of the geographic databases would be redesigned. The current application software supporting the geographic databases would also be redesigned. It is expected that the software development effort would take advantage of the development tool suites associated with the selected RDBMS, such as Oracle's Developer 2000. Where practical, commercially available GIS would be incorporated into the system architecture design to provide system functionality rather than develop unique software to provide the functionality. Figure 3-5 provides a graphical representation of the implementation of Scenario 2.

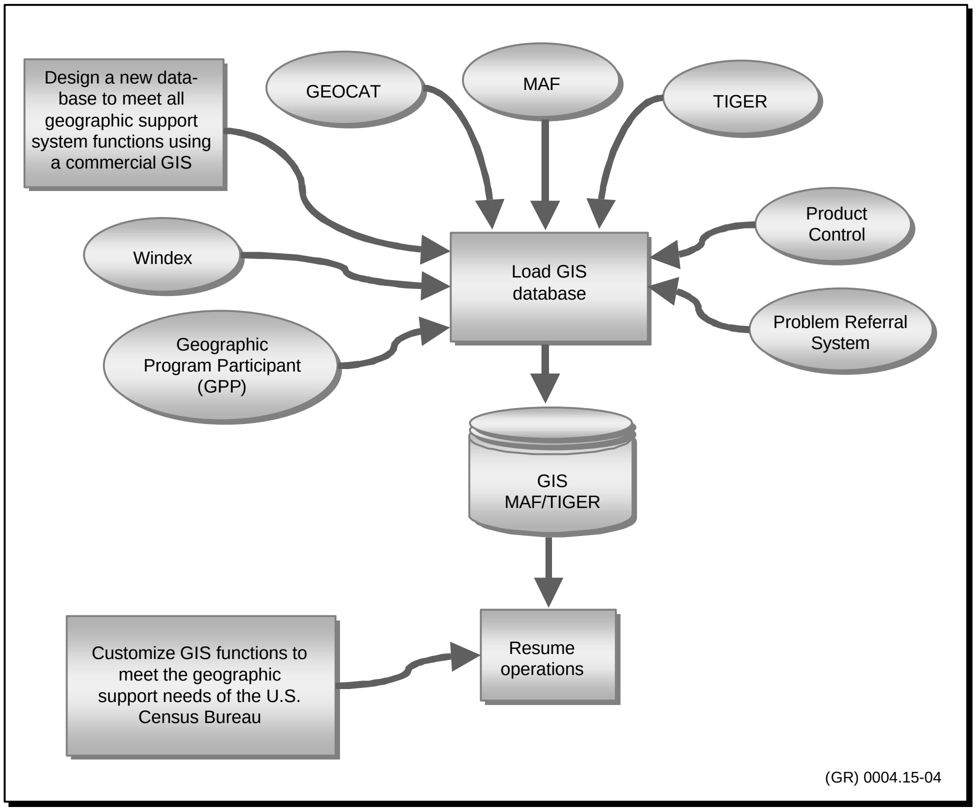
Figure 3-5. Implementation of Scenario 2 for Objective 2



3.2.3 Scenario 3—Commercially Available GIS

In the third scenario, the functionality provided by the geographic databases would be migrated to a commercially available GIS product or suite of products. The geographic databases would be migrated to the database management system supported by the GIS vendor. In addition, a thorough gap analysis of the current functionality that the geographic databases provide the Bureau and the standard functionality supported by the GIS product(s) would be required. Any identified deficiencies would require adjustments to the Bureau's business processes that currently rely on the functionality or would require that the functionality be added to the GIS product(s) through custom development. Figure 3-6 provides a graphical representation of the implementation of Scenario 3.

Figure 3-6. Implementation of Scenario 3 for Objective 2



To support execution of the various scenarios, a structured project management approach would be followed that incorporates, where appropriate, a rigorous application development methodology. The application development methodology would be based on an incremental development strategy that supports the deployment of discrete segments of functionality as they are developed.

3.3 Objective 3—Geographic Partnerships

Objective 3 of the 21st Century MAF/TIGER Enhancements initiative is as follows:

> Expand and encourage geographic partnership programs with Federal, state, local, and tribal governments that update the MAF/TIGER data base, including a "rolling" LUCA to maintain a current address list for use by all Bureau censuses and surveys and computer-based updates of governmental unit boundaries and other geographic areas.

One of the reasons the information technology environment is changing so rapidly is the emergence of integrating infrastructures. With improved integration, the Internet, particularly the World Wide Web (WWW) and electronic commerce have flourished. Where once information systems were segregated and difficult to access, they are now engineered to be accessed using the Web and interfacing software (see Appendix G—Internet-Enabled Applications).

The Internet affords a number of innovative ways to connect users both inside organizations and between organizations. Within organizations, intranets are being used not only to connect divisions and departments but also to facilitate the interaction and collaboration of work groups and project teams located in disparate geographical locations. Between organizations, extranets connect businesses with their trading partners and customers. The extranet is becoming a medium for exchanging information and providing a clear channel for timely communications.

The Bureau values the active participation of its geographic partners (i.e., state, local, tribal governments, and private sector) in maintaining the accuracy and currency of its geographic databases. The Bureau encourages local participation in its programs to the maximum extent feasible. The quality and meaningfulness of data describing the local communities are ensured when this partnership exists. This partnership has proven successful over the years but relies on processes and tools that have lagged behind changes in technology.

To continue supporting the Geographic Support base program, the Bureau will follow established procedures and processes in soliciting current geographic information from its geographic partners. These procedures and processes support programs such as the Boundary and Annexation Survey, the Local Update of Census Addresses, and the Census 2000 Redistricting Data Program. These processes include providing printed materials that are manually updated by the geographic partners and returned for processing by the Bureau through manual office operations. Electronic files are also used for updating the geographic databases when available.

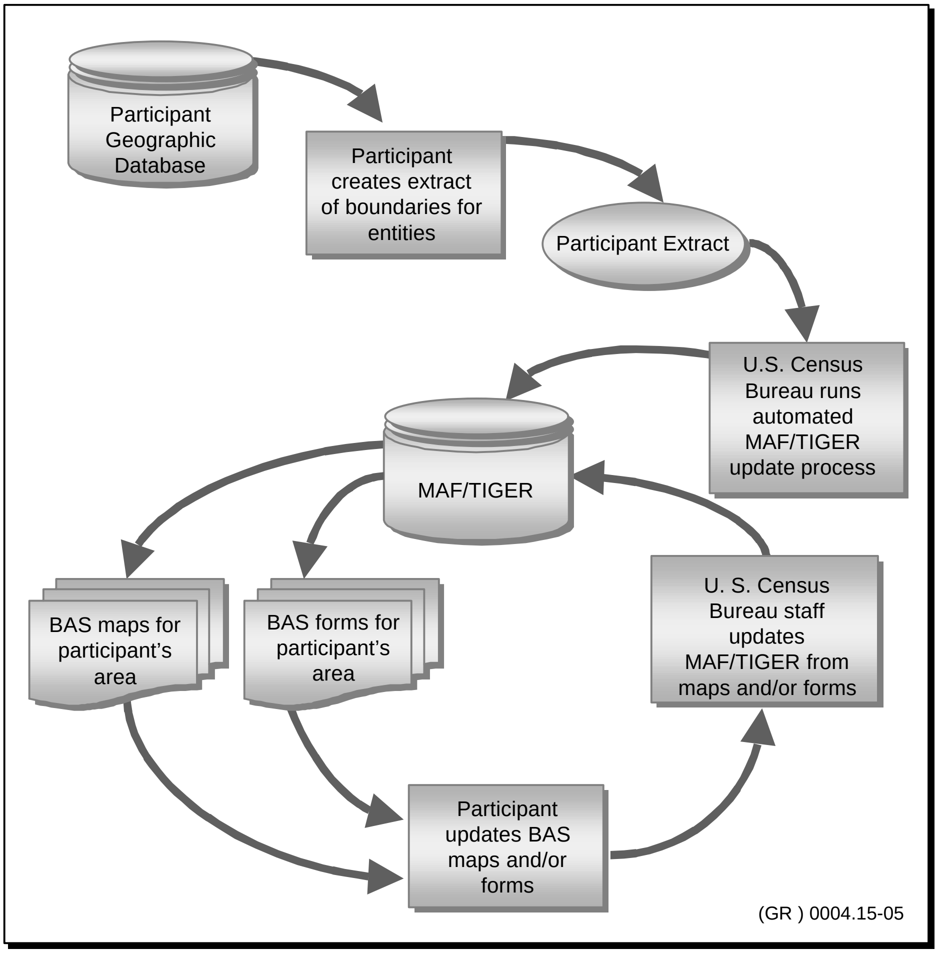
To improve the interaction between the Bureau and its geographic partners, and in the process increase their level of participation, three scenarios have been proposed.

3.3.1 Scenario 1—Annual Boundary and Annexation Survey

The first scenario calls for the development of an electronic data interchange program that supports the submission of boundary and other geographic changes in a computer file rather than on paper maps and forms. The BAS would determine the inventory of legally defined entities and the correct names, governmental descriptions, and legal boundaries of counties and equivalent areas; minor civil divisions; incorporated places; and Native American reservations. Specifically, the BAS would attempt to obtain current boundary information for local and tribal governments; information on the legal actions that had changed local and tribal governmental boundaries; the correct legal name and designation for an entity, such as a city, township, or reservation; updates to the streets and other information shown on Bureau maps; and correct address break information at the boundary of local or tribal governments.

The development of an interactive process that facilitates the automated update of the Bureau's geographic databases from information supplied by the Bureau's geographic partners is key to this electronic data interchange program. The process would accommodate extracts from the geographic partner's GIS systems as well as updates supplied manually. Figure 3-7 provides a graphical representation of the implementation of Scenario 1.

Figure 3-7. Implementation of Scenario 1 for Objective 3

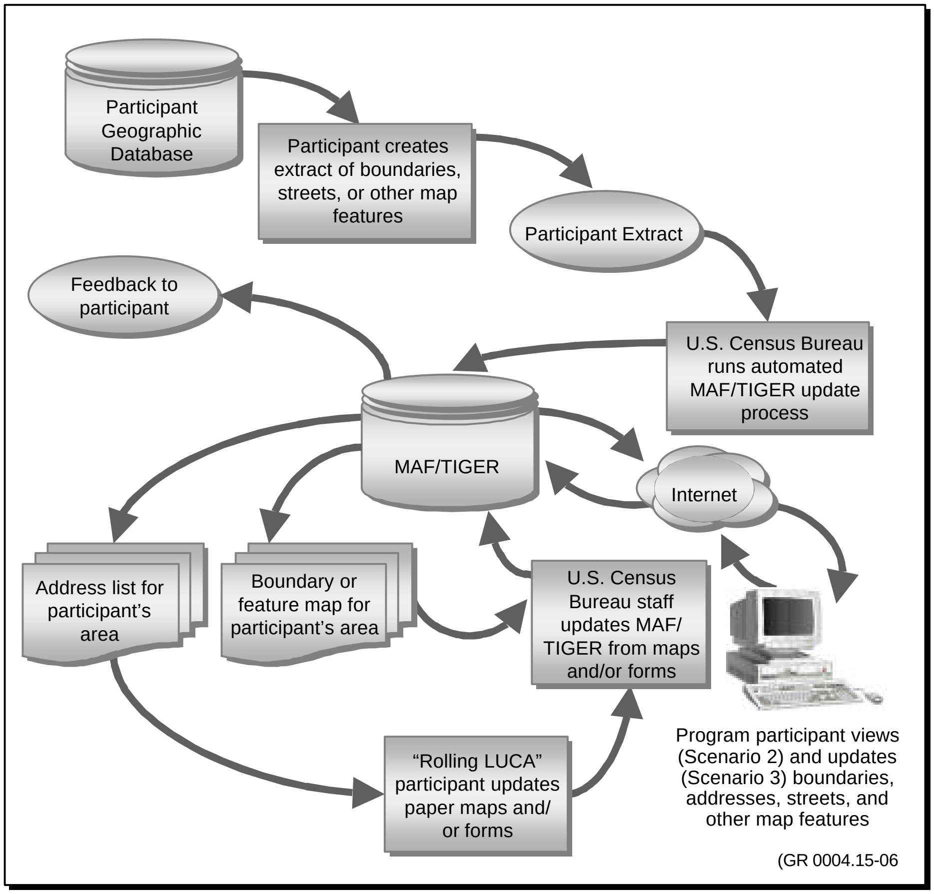


3.3.2 Scenarios 2 and 3—"Rolling" Local Update of Census Addresses Program

In the second and third scenarios, an electronic data exchange program would be developed to support a continuous, or "rolling," LUCA program. The LUCA program goal is to update the information concerning housing units, such as house number; street name and type, including any directional indicator (e.g., S Main St or Apple Blvd SW); unit designator for units in multiunit structures (e.g., 101 Main St, Apt A); and the five-digit ZIP Code. The LUCA program also solicits information to update the Bureau's maps, such as the location of any missing streets, corrections to street names, and to correct other information about streets, such as their location in relation to the boundary of the jurisdiction or displayed intersections.

Both of these scenarios require that a Web-based application be established to support the ability of the Bureau's geographic partners to participate in the LUCA program. The initial scenario (Scenario 2) would establish the capability for geographic partners to access the Bureau's information for review purposes. Any updates to the Bureau's information would be communicated to the Bureau. The companion scenario, Scenario 3, would provide the capability for the geographic partner to update LUCA-related data through a Web application. Figure 3-8 provides a graphical representation of the implementation of both Scenario 2 and Scenario 3.

Figure 3-8. Implementation of Scenario 2 & 3 for Objective 3



3.4 Objective 4—Improve Quality

Objective 4 of the 21st Century MAF/TIGER Enhancements initiative is as follows:

> Implement a comprehensive plan for periodic MAF/TIGER evaluation and corrective activities that will guide the planning for cost-effective future coverage and geocoding improvement operations.

An organization's success is measured by how successfully it meets its mission. To that end, the Bureau must first satisfy its customers' requirements. The geographic databases that support its numerous programs and activities are key to the Bureau's success. The systems that provide access to these geographic databases must be engineered to ensure the completeness and accuracy of the data and the products that are produced.

Current software engineering practices specifically incorporate quality assurance and quality management as core attributes throughout the system development process. As a result, the quality of the system's data and its products is intrinsic to the system's overall design and operation. Although it is preferable to design quality into systems in their initial development, it is also beneficial for existing systems to have quality controls retrofitted into their design to increase end-user confidence and satisfaction.

In support of the Geographic Support base program, the Bureau will follow established quality control procedures and processes in reviewing samples of each output product to verify conformance to specifications. In addition, formal quality standards will be developed and published to control the acceptance of data from the Bureau's geographic partners.

To strengthen the overall quality of the Bureau's geographic support products, three scenarios have been developed. The first two scenarios call for the programmatic improvement of the current systems that support the geographic databases in the area of quality assurance. The third scenario proposes a field evaluation program that would be designed to support the verification of the data in the geographic databases.

3.4.1 Scenario 1—Product Quality

The first scenario is intended to strengthen the current systems that support the geographic databases by expanding their existing capability to monitor the quality of their respective products, such as maps and address lists. This initiative would be accomplished by adjusting the current application software to provide for automated error detection. New software would be developed that would support the error tracking from detection through resolution and facilitate the appropriate level of communication to interested parties. In addition, this scenario would provide for the development of specialized data extracts to support other quality assurance programs initiated by the Bureau.

3.4.2 Scenario 2—Comprehensive Quality Program

The second scenario would expand the focus of Scenario 1 to all areas of the systems that support the geographic databases—not just the output functions. This would include the functions that support data entry and data manipulation.

3.4.3 Scenario 3—Field Verification

The third scenario proposes that a field evaluation program be instituted to verify the accuracy of the data content of the address-related data in the geographic databases. Under this proposal, the Bureau would sponsor the periodic verification of the information in the geographic databases. This operation would be completed through the physical verification of data stored in the various geographic databases for a representative sample of 500,000 housing units.

3.5 Summary

The successful implementation of the scenarios supporting the 21st Century MAF/TIGER Enhancements initiative will enable the Bureau to meet its identified goals. These goals are strengthening its existing partnership with state, local, and tribal governments by improving and simplifying the exchange of information; significantly increasing the accuracy and currency of the information contained in the Bureau's geographic support systems; and improving the quality of the products and tools that support the Decennial Census execution.

APPENDIX A
REQUIREMENTS FOR THE MAF/TIGER MODERNIZATION STUDY

A.1 Background

The Booz·Allen & Hamilton team gathered the following requirements from interviews, meetings, and supporting documentation for potential support to the Master Address File/Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) Enhancements initiative. Interviews were the primary source of information for the detailed requirements. The Booz·Allen & Hamilton team interviewed Geography Division management as well as staff in the Geography Division's branches and functional units; other U.S. Census Bureau divisions; and selected external U.S. Census Bureau customer, partnership, and advisory organizations to identify requirements for modernizing the MAF/TIGER data base. Meetings and interviews were held between November 1999 and May 2000. The results of all interviews were recorded in electronic format and were used in the development of functional requirements.

A.2 Interviews and Requirements Meetings

The complete list of Geography Division interviews and requirements meetings is as follows.

Geography Division Management Meetings

Chief, Geography Division
Assistant Division Chief, Geographic Operations Advisor
Assistant Division Chief, Geocartographic Services
Assistant Division Chief, Geographic Operations
Assistant Division Chief, Geographic Application Systems
Assistant Division Chief, Geoprocessing Systems
Chief, Geographic Planning and Budget Staff

Branch Interviews

Branch	Abbreviation	Branch	Abbreviation
Address List Review Branch	ALRB	MAF Operations Branch	MAFOB
Cartographic Operations Branch	COB	MAF Products Team	MPT
Computer Support Branch	CSB	Mapping Services Team	MST
Economic Programs Branch	EPB	Matching Systems Branch	MSB
Geographic Areas Branch	GAB	Production Operations Branch	POB
Geographic Areas Systems Branch	GASB	Products and Services Staff	PSS
Geographic Planning Branch	GPB	TIGER Mapping Branch	TMB
Geographic Products Quality Assurance Team	GPQA	TIGER Operations Branch	TOB
Geographic Technologies Team	GTT	TIGER Systems Branch	TSB
Geographic Update Systems Branch	GUSB	Update Operations Branch	UOB
Geospatial Research and Standards Staff	GRaSS		

Group Requirements Discussions

Meeting	Attendees
Meeting 1: Boundary/Area Delineation and Maintenance	GAB, GUSB, GASB, GPB, POB, TSB, UOB, GPQA, PSS
Meeting 2: Address Matching and Geocoding	MAFOB, MSB, MPT, ALRB, EPB, GUSB, POB, TOB, UOB, GPQA
Meeting 3: Expansion/Maintenance of Address Boundaries or Ranges	MAFOB, MSB, MPT, ALRB, EPB, GUSB, POB, GPQA
Meeting 4: Mapping for Both Internal and External Customers	TMB, MST, COB, GPB, POB, GPQA
Meeting 5: Reference File and Extract Generation	GAB, GASB, TSB, GPB, POB, GPQA, PSS, TOB
Meeting 6: Maintenance of Road/Railway/Hydrography Network and Attributes	UOB, GUSB, MAFOB, GPB, POB, GPQA, PSS, TOB

Interviews Within Census

American Community Survey (ACS)
Decennial Management Division (DMD)
Decennial Statistical Studies Division (DSSD)
Decennial Systems and Contracts Management Office (DSCMO)
Demographic Statistical Methods Division (DSMD)
Denver Regional Census Center (RCC)
Economic Programs Directorate
Field Division (FLD)
Housing and Household Economic Statistics Division (HHES)
National Processing Center (NPC)
Planning, Research and Evaluation Division (PRED)
Population Division (POP)
Redistricting Data Office
Regional Directors

Groups Contacted for Interviews Outside of Census

2000 Census Advisory Committee
Boundary and Annexation Survey (BAS) Respondent
Census Advisory Committee on the American Indian and Alaska Native Population
Census Advisory Committee on the Asian and Pacific Islander Populations
Census Advisory Committee on the Hispanic Population
Federal Transportation Officials
Mapping Sciences Committee, National Academy of Sciences
National Association of Counties (NACo)
National Center for Education Statistics (NCES)
National Conference of State Legislatures (NCSL)
National League of Cities
National State Data Center (SDC) /Business and Industry Data Center Program (BIDC)
National States Geographic Information Council (NSGIC)
State of Georgia Reapportionment Services
United States Postal Service (USPS)
Vermont for Geographic Information, Inc.

A.3 Requirements Organization

After all of the interviews were conducted, the team began organizing the responses to each question of the interview guide to facilitate the extraction of the requirements. All responses to a particular question were grouped into a single document and similar answers were combined, where appropriate.

The Booz·Allen & Hamilton team developed a detailed requirements matrix to organize the requirements. The detailed organization of the requirements matrix includes—

- **Requirement Number** —Indirectly associates each requirement within a category.
- **Level of Importance**—Indicates the level of importance of each requirement to the success of MAF/TIGER Modernization. Each individual requirement is coded with an E for essential, a D for desirable, or an O for optional.
- **Detailed Requirement (by Major Category)** —Provides a description of the requirement.
- **Source of Requirement** —The source for most requirements is interviews. The five sources of requirements are as follows:
  - *Geography Program Managers*. Labeled **GEO Program Managers**, this category is marked when a requirement's source is attributed to a meeting involving a Geography Division manager.
  - *Field Operations Managers.* Labeled **Field Operations Managers**, this category is marked when a requirement's source is attributed to feedback from the Field Division.
  - *Program Sponsors/Analysts.* Labeled **Program Sponsors/Analysts**, this category is marked when a requirement's source is attributed to a Geography Division interview.
  - *Census Documents.* Labeled **Census Documents**, this category is marked when a requirement's source is attributed to a document provided by the U.S. Census Bureau.
  - *External Users/Customers.* Labeled **External Users/Customers**, this category is marked when a requirement's source is attributed to a conversation/interview with a selected external U.S. Census Bureau-related organization.

- **Base Program—**Documents the extent to which the detailed requirement is being met by the current MAF/TIGER system. Labeled **Base Program**, this category uses the Harvey Ball methodology for estimating the extent to which the current system is meeting the requirement and is coded as follows:

  ○ Current system does not meet this requirement at all
  ◔ Current system meets 1–25 percent of this requirement
  ◑ Current system meets 26–50 percent of this requirement
  ◕ Current system meets 51–75 percent of this requirement
  ● Current system satisfies all the requirement

- **21st Century MAF/TIGER Enhancements**—At the highest level, the functional requirements are organized within the major elements of the Geography Division's budget materials. The four high-level categories, in addition to the existing base program, are italicized and explained below:

- *Activities that build/maintain the MAF of housing units.* Labeled **MAF and TIGER Data**, this category is checked when a requirement deals with correctly locating every street and other map feature in the TIGER data base, and each MAF address, and implementing an effective automated feature change detection methodology.
- *Continuing operations to integrate the MAF with TIGER data base.* Labeled **New Processing Environment**, this category is checked when a requirement deals with implementing a modern processing environment that will support rapid application development, allow extensive use of commercially available software, and reduce staff training time.
- *Continuing partnerships with state, local, and tribal governments.* Labeled **Geographic Partnerships**, this category is checked when a requirement deals with expanding and encouraging geographic partnership programs with state, local, and tribal governments that update the MAF/TIGER data base, including "rolling LUCA" to maintain a current address list for use by all Bureau censuses and surveys and computer-based updates of governmental unit boundaries and other geographic areas.
- *Quality assurance for geographic support products*. Labeled **Improve Quality**, this category is checked when a requirement deals with implementing a comprehensive plan for periodic MAF/TIGER evaluation and corrective activities that will guide planning for future cost-effective future coverage and geocoding improvement operations.

• **Comment/Issue—**Captures general comments and issues.

Once the requirements were extracted from all interviews, meetings, and supporting documentation, the Booz·Allen & Hamilton team grouped all requirements into seven high-level requirements. The high-level requirements are intended to capture the main thoughts of the underlying requirements at a summary level. The team created an additional category labeled "other observations" to include requirements that did not fit into one of the seven high-level requirements but were captured in the requirements gathering. The seven high-level requirements are as follows:

1. Maintain an accurate and complete inventory of all residential, group quarters, and commercial building structures, along with the geographic location for each, and the within-structure designations when a structure contains more than one unit.

2. Maintain an accurate and complete inventory of legal, administrative, and statistical geographic areas, their attributes, and interrelationships.

3. Maintain an accurate and complete inventory of road, rail, water, landmark, power transmission line, and other features mapped for U.S. Census Bureau purposes.

4. Make use of geospatial data provided by Federal, state, local, and tribal government partnerships, as well as private-sector sources, wherever possible.

5. Provide geographic and address list products and other support services to U.S. Census Bureau programs and census data users.

6 Comply with applicable geographic data standards, guidelines, and directives and meet the Federal Geographic Data Committee (FGDC) objectives for the National Spatial Data Infrastructure (NSDI) initiative, where possible.

7. Improve the accuracy and responsiveness of the geographic support system to meet the needs of U.S. Census Bureau program managers, geographic program participants, and U.S. Census Bureau data product users.

REQUIREMENTS

1.0 Maintain an Inventory of All Structure Addresses, Within Structure Designations, and Location Descriptions

Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	Source of Requirement: U.S. Census Bureau Customers: GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	21st Century MAF/TIGER Enhancements: MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
1.1	**E**	Automate field operations by use of handheld computers	•		•			◔		×		×	• Need to be able to digitize new streets and such from the field
1.2	**E**	Correlate five-digit Zip Code with only one default U.S. Post Office	•		•			◑	×	×			
1.3	**E**	Develop a type code for every type of structure	•					○	×	×			
1.4	**E**	Display individual housing units as point features	•					◑	×	×	×	×	• This would assist in adding up individual units geographically • Would help identify housing units within a given block
1.5	**E**	Eliminate duplicate addresses in the MAF	•		•			◑	×	×		×	
1.6	**E**	Ensure that information obtained by Local Update of Census Addresses portrays the same block numbers on the address list and on maps	•					◕	×	×	×	×	
1.7	**E**	Improve the MAF's identification of housing units in rural areas	•					◔	×			×	• Noted that there are 2 million structures that do not have addresses in the MAF (most of which are in rural areas)

1.0 Maintain an Inventory of All Structure Addresses, Within Structure Designations, and Location Descriptions													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
1.8	**E**	Include key geographic information, such as trailer parks, in the block level	●					◑	×	×	×		• Suggestion to give internal address and give them a descriptive address in addition to mailing address
1.9	**E**	Include ZIP Code information in the MAF	●					◕	×	×		×	
1.10	**E**	Maintain dual record if business and residence are located at the same place	●					○	×	×			
1.11	**E**	Maintain timestamps and history of address/structure updates	●					○	×	×		×	• Allows for reconstruction of structures and MAF inventory at some point in time
1.12	**E**	Match accurate X and Y coordinates to the address	●					◔	×		×	×	
1.13	**E**	Minimize field operations	●					◔	×	×	×		• Bureau could save millions of dollars a year by minimizing field operations • Imagery would support this
1.14	**E**	Provide a 100 percent comprehensive listing of structures in the United States in the MAF	●					◕	×	×	×	×	
1.15	**E**	Provide ability to move subdivisions to correct locations			●			◑	×	×			

1.0 Maintain an Inventory of All Structure Addresses, Within Structure Designations, and Location Descriptions													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
1.16	**E**	Provide better control over data entry	●					◑	×	×		×	• Mentioned that there is a need to hire qualified personnel to digitize changes
1.17	**E**	Provide better inventory of multistructure units	●					◑	×	×	×	×	• Basic street address (BSA) does not work well • Need to work with building managers
1.18	**E**	Provide coordinates for every address	●	●	●			◑	×	×		×	
1.19	**E**	Provide for improved positional accuracy	●	●	●	●	●	◔	×	×	×	×	• Provide to use detailed data exchange with partitions
1.20	**E**	Provide user friendly tallying capabilities on MAF address information	●					◔	×	×	×	×	• For example, total number of commercial units in Houston
1.21	**E**	Report addresses by type to determine workloads	●					◕					• Improve efficient field operations
1.22	**E**	Resolve addresses that cannot be matched to TIGER®	●		●			◕	×	×	×	×	
1.23	**E**	Resolve non-city-style address that appear in multiple locations in the database	●					○	×	×		×	
1.24	**E**	Standardize data entry	●					◑	×	×		×	• This affects comparability

1.0 Maintain an Inventory of All Structure Addresses, Within Structure Designations, and Location Descriptions														
			Source of Requirement						21st Century MAF/TIGER Enhancements					
			U.S. Census Bureau Customers											
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
1.25	**E**	Support addresses for Local Update of Census Addresses	•					◕	×		×		
1.26	**E**	Support addresses for the 2002 economic census	•					◔	×				
1.27	**E**	Target high-growth areas that should be priorities for the Community Address Update System and MAF updating operations				•		◕	×	×			
1.28	**E**	Validate address ranges through a quality assurance process before such information is entered into the database	•		•			◔	×	×	×	×	
1.29	**D**	Allow a distinction between address and occupant in the MAF	•					○	×	×	×		
1.30	**D**	Allow Z coordinates for address unit identifiers	•		•			○	×	×	×		• Multiple Zip Codes for single structures • Allocate work loads for large units
1.31	**D**	Calculate address ranges ad hoc	•					◕	×	×			• To minimize overhead cost of maintaining them
1.32	**D**	Develop methodology to determine if coordinates reference a mobile home or the pad where the mobile home sits	•				•	○	×	×		×	

1.0 Maintain an Inventory of All Structure Addresses, Within Structure Designations, and Location Descriptions													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
1.33	**D**	Employ Global Positioning System (GPS) data to assist with accuracy assessment	●					○	×	×	×	×	
1.34	**D**	Ensure MAF/TIGER remains current with ZIP Code changes	●					○	×		×	×	
1.35	**D**	Identify and record attribute information on structures when enumerators are in the field	●					◔	×	×		×	
1.36	**D**	Provide a single address list for both housing units and businesses	●					◑	×	×		×	
1.37	**D**	Provide ability to automatically enter the next available map spot number to the next map spot			●			○	×	×		×	• Would like to be able to enter 1 and 95 and have the computer evenly space all others that are between them
1.38	**D**	Provide functionality in MAF to link address to census tract	●					◕	×	×		×	
1.39	**D**	Use imagery to assess data quality	●					○	×	×		×	• Ongoing change detection
1.40	**D**	Use imagery to pinpoint new development	●		●			○	×			×	
1.41	**O**	Differentiate large farms and plantations from other map spots	●					○	×	×		×	
1.42	**O**	Maintain the nine-digit ZIP Code related to geography	●					◑		×			
1.43	**O**	Reduce dependence on the Delivery Sequence File (DSF)	●					○	×	×			

1.0 Maintain an Inventory of All Structure Addresses, Within Structure Designations, and Location Descriptions														
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
1.44	**O**	Remove individual addresses from under the Title 13 requirement	●					○	×	×		×	
1.45	**O**	Use the U.S. Postal abbreviations in the MAF					●	◕	×	×		×	

2.0 Maintain an Inventory of Legal and Statistical Geographic Areas													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
2.1	**E**	Allow for correct block definitions with ongoing block comparability based on standard criteria	●					◕	×		×	×	
2.2	**E**	Capture individual voting districts and precincts, and state legislative districts	●					◕	×	×	×	×	
2.3	**E**	Capture nonvisible voting districts and other administrative boundaries	●					○	×	×	×	×	
2.4	**E**	Create relationships between features and boundaries	●					◕	×	×	×	×	
2.5	**E**	Eliminate duplicate addresses in the MAF	●		●			◑	×	×		×	
2.6	**E**	Improve boundary, name, and code information for statistical areas	●		●			◑	×		×	×	
2.7	**E**	Maintain a time stamp for all boundary and feature updates	●					○	×	×			
2.8	**E**	Provide accurate border information for the US–Canada and US–Mexico border	●					◕	×	×	×	×	
2.9	**E**	Provide better control over data entry	●					◑	×	×		×	• Mentioned that there is a need to hire qualified personnel to digitize changes
2.10	**E**	Provide boundary files for individual housing units	●					○	×	×		×	• This would ignore block structure for tabulation

2.0 Maintain an Inventory of Legal and Statistical Geographic Areas													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
2.11	**E**	Provide economic geography information	●					○	×	×	×	×	• Potomac Mills and Tysons cited as examples
2.12	**E**	Provide for the integration of the MAF and TIGER data bases so that a boundary change will automatically update the corresponding geocoding	●					○	×	×	×	×	• This is the most essential requirement for accurate and correct geocoding and products (like LUCA)
2.13	**E**	Reevaluate Geographic Update System (GUS)	●					◑	×	×			• Enforce relational changes
2.14	**E**	Standardize data entry	●					◑	×	×		×	• This affects comparability
2.15	**E**	Support American Indian, Alaska Native, and Native Hawaiian geographic boundaries	●					◔	×	×	×	×	
2.16	**E**	Support automatic changes related to boundary changes	●					◔	×	×	×	×	• Change in county boundary means change in census tract, etc.
2.17	**E**	Support nonspatial relationships	●					○	×	×		×	
2.18	**E**	Support variety of methods of input for boundary changes, to include those from TIGER/Line®, Digital Exchange (DEX) files, and updates from school districts	●					◔	×	×	×	×	

2.0 Maintain an Inventory of Legal and Statistical Geographic Areas													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
2.19	**E**	Target high-growth areas that should be priorities for the Community Address Update System and MAF updating operations	●				●	◕	×	×			
2.20	**D**	Allow for immediate attribute update for new geographic entities	●					○	×			×	• Minimizes discrepancies between spatial and nonspatial databases and need for batch updates
2.21	**D**	Include county-size partitions	●					◕	×	×			
2.22	**D**	Include feature standardization	●					◑	×		×	×	• Need feature standardization for water features, blocks, military establishments, parks • Must hold up to scrutiny • Must contain adequate and clear definitions
2.23	**D**	Link legal change action data to areas	●					○	×	×	×		
2.24	**D**	Maintain a comprehensive boundary inventory of different types of special districts	●					◔	×	×	×		• Estimated at hundreds of thousands of special districts
2.25	**D**	Produce boundary maps quickly	●					◕	×	×			
2.26	**D**	Provide for the definition of census tracts before the census is conducted	●					◑	×	×	×	×	
2.27	**D**	Provide for the easy adjustment of census tracts during the census	●					○	×	×			

2.0 Maintain an Inventory of Legal and Statistical Geographic Areas													
			Requirement						MAF/TIGER				
			U.S. Census										
Requirement Number		Detailed Requirement	GEO Program	Field Operations	Program Sponsors/	Census Documents		Base Program		New Processing	Geographic	Improve Quality	
		display information for a						◔	×	×			
2.29	**D**	Provide functionality to level	●						×	×			
2.30	**D**	Provide information on geographic areas	●					◑	×	×	×		
2.31		Provide ongoing update area definitions						◕	×			×	
2.32	**D**	Require spatial Boundary and surveys	●					○	×	×		×	
		with delineating blocks	●					○	×	×	×	×	

3.0 Maintain an Inventory of Road, Rail, Water, Landmark, Power Transmission Line, and Other Features														
			Source of Requirement						21st Century MAF/TIGER Enhancements					
			U.S. Census Bureau Customers											
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
3.1	**E**	Allow streets to be entered into MAF/TIGER without a suffix					●	○	×	×			• Not all roads have a suffix and one should not be forced
3.2	**E**	Develop methodology for keeping database current with demolished structure information	●					○	×	×	×	×	
3.3	**E**	Develop type code and capture as a landmark for each structure	●					○	×	×			
3.4	**E**	Differentiate (on digitized maps) whether road line is center of highway or side of highway	●					○	×	×		×	
3.5	**E**	Employ GPS data to provide more accurate coordinates for TIGER features and MAF addresses	●		●			○	×	×		×	
3.6	**E**	Improve the positional accuracy of roads	●	●	●			◕	×	×	×	×	
3.7	**E**	Include Public Land Survey System (PLSS)	●					○	×	×	×	×	
3.8	**E**	Maintain a timestamp and history of all base feature updates	●					○	×	×		×	• Allows for reconstruction of base features at any point in time
3.9	**E**	Maintain primary and alternate road naming	●		●			◕	×	×	×	×	
3.10	**E**	Provide better control over data entry	●					◑	×	×		×	• Mentioned that there is a need to hire qualified personnel to digitize changes

3.0 Maintain an Inventory of Road, Rail, Water, Landmark, Power Transmission Line, and Other Features														
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
3.11	**E**	Provide economic geography information	•					○	×	×	×	×	• Potomac Mills and Tysons cited as examples
3.12	**E**	Provide way of bringing up an image/map to assist with feature alignment			•			◑	×	×		×	
3.13	**E**	GUS	•					◑	×	×			• Enforce relational changes
3.14	**E**	Standardize data entry	•					◑	×	×		×	• This affects comparability
3.15	**E**	Support nonspatial relationships	•					○	×	×		×	
3.16	**E**	Target high-growth areas that should be priorities for the Community Address Update System and MAF updating operations				•		◕	×	×			
3.17	**D**	Allow for large regional landmarks	•					○				×	
3.18	**D**	Define minimum water criteria	•					◔	×	×		×	• Small bodies of water can cause large problems—can become their own blocks
3.19	**D**	Develop methodology to determine if coordinates reference a mobile home or the pad where the mobile home sits	•				•	○	×	×		×	
3.20	**D**	Differentiate on maps between regular roads, overpasses, and underpasses	•					○	×	×		×	• Particularly important for underground roads, like in Boston

3.0 Maintain an Inventory of Road, Rail, Water, Landmark, Power Transmission Line, and Other Features													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
3.21	**D**	Expand TIGER® to include roads less than 200 feet					●	○	×	×		×	
3.22	**D**	Include complete road information					●	◔	×	×		×	
3.23	**D**	Provide current railroad information	●					◑	×	×		×	
3.24	**D**	Provide documentation on shopping center files, Yellow Pages, postal services, "Zip plus 4" files, and city/state files	●					○	×	×	×		
3.25	**D**	Provide information for special places and other land-use areas	●					◔	×			×	
3.26	**D**	Provide sufficient water definition criteria	●					◑	×			×	• Noted that there are too many water definition criteria; many unclear
3.27	**D**	Support multiple feature types discretely that occupy same line or area	●					◔	×				
3.28	**O**	Include detailed cadastal information	●					○	×		×		
3.29	**O**	Minimize number of alternate names for a given location		●	●			○	×	×	×	×	• One source cites as many as 37 for a given entity
3.30	**O**	Provide current bikeway information	●					◔	×	×			
3.31	**O**	Provide current Bureau of Land Management (BLM) forest area information	●					◔	×	×			
3.32	**O**	Provide current park areas information	●					◔	×	×			
3.33	**O**	Provide current pipeline information	●					○	×	×			
3.34	**O**	Provide current powerline information	●					○	×	×			

3.0 Maintain an Inventory of Road, Rail, Water, Landmark, Power Transmission Line, and Other Features													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
3.35	**O**	Provide current trailways information	●					○	×	×			
3.36	**O**	Provide pictures to help clerical coders find locations and intersections	●					○	×			×	• This would assist with quality assurance from extraction results
3.37	**O**	Provide ridge line information	●					○	×	×			
3.38	**O**	Use the U.S. Postal abbreviations in the MAF					●	◕	×	×		×	

4.0 Use Geospatial Data Provided by External Sources													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
4.1	**E**	Extract necessary data for the geocoding system				●		◕		×			
4.2	**E**	Geocode by matching to an individual address instead of matching to an address range				●		○		×			• May geocode by a carrier route or box number, a USPS PO box, or a city-style address that was not used to create an address range
4.3	**E**	Grant permission for locals to update TIGER®	●					○	×	×	×	×	
4.4	**E**	Have ability to cut out portions of TIGER® and replace with better comparable local/tribal files	●					○	×	×	×	×	
4.5	**E**	Improve coordinates to ensure better match with detailed local files	●					◔	×		×	×	• Essential for digital data exchange
4.6	**E**	Improve relationships between the Geography Division and local organizations	●		●			◔		×	×	×	• Local Update of Census Addresses worked in New York City because of excellent communication between local organizations
4.7	**E**	Include a "Group Quarters facilities, Building, or Apartment Name Match" module				●		○	×	×			
4.8	**E**	Include a "MAF Match Geocoding" module				●		◕		×			

4.0 Use Geospatial Data Provided by External Sources													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
4.9	**E**	Include an "Address Geocoding using a Mailed Questionnaire and Area Map" module				●		○		×			
4.10	**E**	Include an "estimation for project size, timing, and cost" module				●		○		×		×	
4.11	**E**	Include an "intake" module				●		○		×		×	
4.12	**E**	Include control flags				●		◕		×			• Control flags will determine the processes each address will go through to obtain the target set of geocodes
4.13	**E**	Include in estimation module estimates for coding rates and contribution of each geocoding option				●		○		×		×	• Samples of addresses to be provided by sponsor
4.14	**E**	Include in estimation module information on completion schedule, cost, and coding				●		○		×		×	• Information to be provided by sponsor
4.15	**E**	Incorporate local file metadata into TIGER®	●					○	×	×	×	×	
4.16	**E**	Integrate all point, line, and area TIGER data for reshaping to better coordinates	●					◑	×	×	×	×	
4.17	**E**	Integrate the U.S. Postal Service files completely with databases	●					◕	×	×	×	×	
4.18	**E**	Provide functionality to copy delivered file onto off-line storage				●		◕		×			
4.19	**E**	Provide functionality to copy delivered file onto the National Geocoding System disk				●		○		×			

4.0 Use Geospatial Data Provided by External Sources													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
4.20	**E**	Provide functionality to have the MAF queried and manipulated directly via the Internet by U.S. Census Bureau and external organizations	●	●	●		●	○	×	×	×	×	
4.21	**E**	Provide functionality to receive address information to be geocoded from the sponsor				●		◔		×		×	
4.22	**E**	Provide functionality to reference the PLSS	●					○	×	×	×	×	
4.23	**E**	Provide security measures to lock out unauthorized users			●			○		×		×	• Only allow uses access to data who must use data
4.24	**E**	Support the reference of information from all sources of input				●		○		×			• Sources include computerized, computerized-assisted look-up, and paper
4.25	**E**	Support the use of U.S. Census Bureau addresses for government contact data	●					◑	×	×	×		
4.26	**E**	Support the use of local address ranges and the unsolicited address information from local and tribal governments	●					◔	×		×		
4.27	**E**	Support the use of spatial databases from local and tribal governments	●					◔	×	×	×	×	

4.0 Use Geospatial Data Provided by External Sources													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
4.28	**E**	Use Community Address Update System (CAUS) field staff to locate and map spot ungeocoded addresses				●		○				×	• Staff uses local sources • For travel efficiency, the uncoded addresses must be grouped into compact areas
4.29	**E**	Use computer-assisted telephone interview (CATI) geocoding for sponsor-provided telephone numbers or addresses with a telephone number provided by a commercial source				●		◔	×	×			• This approach works faster and cheaper than a field operation • OMB approval needed
4.30	**E**	Use the special place names and addresses assembled in the 2000 Census				●		◐	×	×			
4.31	**D**	Assign a processing number to project and record				●		○		×			• This ID will identify both the sponsor and individual record/address
4.32	**D**	Develop strategy for updates from local and tribal entities that usually do not update databases	●		●			◔	×	×	×	×	
4.33	**D**	Ensure that local entities receiving software have appropriate hardware to run system			●		●	○		×	×		

4.0 Use Geospatial Data Provided by External Sources														
			Source of Requirement						21st Century MAF/TIGER Enhancements					
			U.S. Census Bureau Customers											
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
4.34	**D**	Expand Geography Division focus to better include external customers	●					◔	×	×	×		• Internal operational constraints conflict with public use and perception of product
4.35	**D**	Expand system ability to geocode to major special places				●		◑	×	×		×	• Such special places would be carried in TIGER® as Key Geographic Locations (KGL)
4.36	**D**	Improve working relationship with the U.S. Postal Service; consider working at local Post Office level			●			○	×	×		×	
4.37	**D**	Include a “CATI Geocoding” module				●		○		×			
4.38	**D**	Include a “Match to References Sources” module				●		○		×			
4.39	**D**	Include a “Regional Office and CAUS Staff Field Geocoding” module				●		○		×			
4.40	**D**	Include an “Address Range Geocoding” module				●		◕		×			
4.41	**D**	Provide capability to merge United States Geological Survey data into system	●					◑	×	×		×	
4.42	**D**	Provide functionality so that respondents can spot their addresses on a map sent to them and have the information geocoded into the system	●			●		○		×	×	×	• Expect back 50 percent of maps mailed out

4.0 Use Geospatial Data Provided by External Sources													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
4.43	**D**	Provide functionality to trigger a process to geocode new addressees in the DSF				•		◕		×			
4.44	**D**	Provide Sequenced Query Language (SQL) like ability to query data	•		•			◔	×	×	×		• Noted that querying TIGER® is not intuitive
4.45	**D**	Reduce minor data content changes	•					○	×	×		×	
4.46	**D**	Standardize the process for entering data into the system	•					◕	×	×		×	
4.47	**D**	Support extracts from ZIP Code tab areas	•					◔	×	×	×		
4.48	**D**	Support the process of having a field geocoding clerk identify unknown areas on a map				•		◕				×	
4.49	**D**	Use the highest and lowest potential address as well as address parity				•		◕		×			• To assign a set of geocodes to the address
4.50	**O**	Calculate road distances between housing units			•			○	×	×	×		• Need to be able to determine driving distances
4.51	**O**	Create interdisciplinary division within the U.S. Census Bureau to review and accept anything provided from local organizations					•	○	×	×	×	×	
4.52	**O**	Include a "MAFGOR and TMU Uncoded Feature Resolution" module				•		◕		×			

5.0 Provide Geographic Products and Support Services

Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	Source of Requirement: U.S. Census Bureau Customers: GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	21st Century MAF/TIGER Enhancements: MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
5.1	**E**	Allow the Geographic Catalogue to support current and previous geography data	●					◔		×		×	
5.2	**E**	Create detailed, specialized maps using display parameters interactively	●					◔		×			
									×			×	
5.4	**E**	Develop software to edit and maintain geographic data files	●					◑		×			
		for the geocoding				●				×			
5.6	**E**	Geocode 100 the United States	●					◕		×			
5.7	**E**	Geocode by matching to an individual address instead of matching to an address range				●		○					

Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	U.S. Census Bureau Customers										
			GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
5.8	**E**	Include ''MAF Match Geocoding" module				●		◕		×			
5.9	**E**	Include an "Address Geocoding Using a Mailed Questionnaire and Area Map" module				●		○		×			
5.10	**E**	Include in estimation module information on completion schedule, cost, and coding				●		○		×		×	• Information to be provided by sponsor
5.11	**E**	Integrate the U.S. Postal Service files completely with databases	●					◕	×	×	×	×	
5.12	**E**	Provide a better way to organize and classify data	●					◔	×	×	×	×	
5.13	**E**	Provide documentation on Master Address File Update File (MAFUF)	●					◔	×	×		×	•
5.14	**E**	Provide documentation on the Master Address File Geocoding Office Resolution	●					◔	×	×		×	
5.15	**E**	Provide functionality to copy delivered file onto off-line storage				●		◕		×			
5.16	**E**	Provide functionality to copy delivered file onto the National Geocoding System disk				●		○		×			
5.17	**E**	Provide functionality to receive address information to be geocoded from the sponsor				●		◔		×		×	
5.18	**E**	Provide method of distributing Local Update of Census Addresses participant addresses to appropriate geographic entity	●					●					

5.0 Provide Geographic Products and Support Services													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
5.19	**E**	Provide same geocoding in both TIGER and MAF data bases	●					◑	×	×		×	• MAF has geocodes but not the same as TIGER®. MAF has census geocode (block) and a TIGER geocode (could be different) until the TIGER® is updated
5.20	**E**	Support ad hoc MAF queries for new surveys and sample expansions				●		◕	×	×	×		
5.21	**E**	Support address breaks and address range extracts	●		●			◕	×	×			• In accordance with Federal regulation (Title 13), the Bureau can divulge only address ranges, not individual addresses
5.22	**E**	Support density assessment maps	●					◑	×	×			
5.23	**E**	Support the automated geocoding of Puerto Rico	●					◔	×	×	×	×	
5.24	**E**	Support the community address update system with DSMD	●					○	×	×			
5.25	**E**	Support TIGER/Line® products in the geographic database	●					◕	×	×			

5.0 Provide Geographic Products and Support Services													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
5.26	**E**	Synchronize the MAF and TIGER identification numbers to link geocoding	●		●			◔	×	×		×	
5.27	**E**	Use CATI geocoding for sponsor-provided telephone numbers or addresses with a telephone number provided by a commercial source				●		◔	×	×			• This approach works faster and cheaper than a field operation • OMB approval needed
5.28	**E**	Use CAUS field staff to locate and map spot ungeocoded address				●		○				×	• Staff uses local sources • For travel efficiency, the uncoded addresses must be grouped into compact areas
5.29	**D**	Develop generic TIGER input/output (I/O) routines	●					○	×	×		×	
5.30	**D**	Develop in-house, Web-based tool to download MAF data	●					○	×	×	×	×	
5.31	**D**	Document differences between versions of TIGER/Line®	●					○	×	×		×	• Metrics may include miles of paved roads
5.32	**D**	Employ GPS data to provide more accurate geocoding	●					○	×	×	×	×	
5.33	**D**	Expand system ability to geocode to major special places				●		◑	×	×			• Such special places would be carried in TIGER® as KGL
5.34	**D**	Geocode by building matching	●					◔	×	×	×	×	

5.0 Provide Geographic Products and Support Services													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
5.35	**D**	Geocode by intersection matching	●					○	×	×	×	×	
5.36	**D**	Improve working relationship with the U.S. Postal Service; consider working at local Post Office level			●			○	×	×		×	
5.37	**D**	Include a “CATI Geocoding” module				●		○		×			
5.38	**D**	Include a “Match to References Sources” module				●		○		×			
5.39	**D**	Include a “Regional Office and CAUS Staff Field Geocoding” module				●		○		×			
5.40	**D**	Include an “Address Range Geocoding” module				●		◕		×			
5.41	**D**	Provide ability to automatically enter the next available map spot number to the next map spot			●			○	×	×			• Would like to be able to enter 1 and 95 and have the computer evenly space all others that are between them
5.42	**D**	Provide documentation on Decennial Master Address File extracts	●					◑	×	×	×		
5.43	**D**	Provide functionality so that respondents can spot their addresses on a map sent to them and have the information geocoded into the system	●			●		○				×	• Expect back 50 percent of maps mailed out
5.44	**D**	Provide functionality to trigger a process to geocode new addressees in the DSF				●		◕		×			

5.0 Provide Geographic Products and Support Services													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
5.45	**D**	Provide historical ZIP Code information	●					○	×	×		×	
5.46	**D**	Provide regeocoded year 2000 MAF every year to reflect updated legal boundaries			●		●	○	×	×		×	
5.47	**D**	Provide system to store spatial and geographic rules	●					○	×	×		×	• This will prevent special exceptions being written into the software
5.48	**D**	Provide the ability to create TIGER/Line® products on the fly	●					○	×				
5.49	**D**	Provide the ability to geocode substructures	●					◑	×	×			• This becomes necessary for shopping centers; individual stores in the mall change over time—need to geocode individual businesses
5.50	**D**	Provide the effective geocoding of city-style addresses	●					◑	×	×	×	×	• Noted in meeting that this cannot be done currently
5.51	**D**	Support the process of having a field geocoding clerk identify unknown areas on a map				●		◕				×	
5.52	**D**	Use commercial tools that support large volume map making	●					○	×	×		×	
5.53	**D**	Use commercial-based software to draw maps	●					◔		×	×	×	

5.0 Provide Geographic Products and Support Services													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
5.54	**O**	Build maps dynamically over the Internet	●					○		×	×		
5.55	**O**	Provide ability to temporarily erase or suppress map overlays	●				●	○	×	×			• Need to be able to get to the base map
5.56	**O**	Provide public interactive geocoding functionality on-line	●					◔	×	×	×	×	
5.57	**O**	Report activities by Type of Enumeration Area	●					◔	×	×		×	

6.0 Meet Requisite Standards													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
6.1	**E**	Follow International Standards Organization (ISO) standards	●					○		×	×	×	
6.2	**E**	Follow Open GIS Consortium (OGC) standards	●					○		×		×	
6.3	**D**	Follow Federal and state government standards	●					◔	×	×	×	×	

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.1	**E**	Allow for easier update of TIGER information by non-Census-controlled sources	●					◔	×	×	×	×	
7.2	**E**	Allow for multiple users to access the Address Reference File (ARF) information at the same time	●					○	×	×	×		
7.3	**E**	Allow the Geographic Catalogue to support current and previous geography data	●					◔	×	×		×	
7.4	**E**	Apply more project management principles	●					◔	×	×	×	×	
7.5	**E**	Attract qualified information technology (IT) professionals	●					◔				×	
7.6	**E**	Create metadata that will detail information regarding any change to the MAF/TIGER data bases	●	●	●	●	●	○	×	×	×	×	• Three interviewees stated that metadata was necessary to help with data accuracy • Concern was stated in two interviews that by opening the database to everyone, data accuracy would decrease

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.7	E	Create the ability to use a commercially available tool to query the MAF/TIGER data bases	●					○	×	×	×	×	• Interviewee suggested commercially available software such as Excel or use of an open database connectivity (ODBC)
7.8	E	Decrease difficulty in changing entities	●					○	×	×		×	• Mentioned that system is limited to VAX/Virtual Memory System (VMS) platforms and alpha machines • MAF and TIGER I/O files are difficult to extract
7.9	E	Decrease Geography Division's reliance on programmers	●					◔		×			• Programmers are overburdened
7.10	E	Decrease new software development time	●					◔	×	×		×	
7.11	E	Decrease production time of TIGER/Line® products	●	●	●			◔	×	×		×	• Create on the fly for internal use
7.12	E	Develop a system that does not require check-in/check-out partitions and partition splits	●	●				○	×	×	×	×	
7.13	E	Develop network monitoring tools	●					◑	×	×	×		
7.14	E	Develop programs for directory storage	●					◑	×	×			

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.15	**E**	Develop software to edit and maintain geographic data files	•					◑	×	×		×	• Geographic area software, geographic data log, and GUSX mentioned
7.16	**E**	Develop standard queries	•					○	×	×	×	×	
7.17	**E**	Employ GPS data to provide more accurate geocoding	•					○	×	×	×	×	
7.18	**E**	Enhance system safeguards to protect TIGER data	•					◕	×	×	×	×	
7.19	**E**	Ensure that data is easier to maintain	•					○	×	×	×	×	
7.20	**E**	Ensure that data on new system is protected	•					◕	×	×	×	×	
7.21	**E**	Ensure that system data can be deleted	•					◔	×	×		×	
7.22	**E**	Ensure that the new system is intelligently distributed and does not rely on substantial file extraction practices	•					◑	×	×		×	
7.23	**E**	Ensure that the new system provides at least comparable level of functionality	•					○	×	×			
7.24	**E**	Ensure that there is central management of all software interfacing with TIGER® and the MAF	•					◔	×	×		×	
7.25	**E**	Ensure that TIGER data is topologically accurate	•					◑	×			×	
7.26	**E**	Ensure that TIGER® is at least as accessible as it is now	•					◑	×	×			
7.27	**E**	Erase working copies of files when finished using them	•					◑	×	×		×	

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.28	**E**	Geocode 100 percent of the United States	•					◕	×	×			
7.29	**E**	Grant permission for locals to update TIGER®	•					○	×	×	×	×	
7.30	**E**	Improve communication between the technical and geographic sides of the Geography Division	•					◑		×	×	×	
7.31	**E**	Improve planning and scheduling	•					◕	×	×	×	×	
7.32	**E**	Improve software reusability	•					◔		×		×	
7.33	**E**	Improve system planning and coordination	•					○	×	×	×	×	
7.34	**E**	Improve training for new system users	•					◑	×	×	×	×	
7.35	**E**	Include in estimation module estimates for coding rates and contribution of each geocoding option				•		○		×		×	• Samples of addresses to be provided by sponsor
7.36	**E**	Include in estimation module information on completion schedule, cost, and coding				•		○		×		×	• Information to be provided by sponsor
7.37	**E**	Increase knowledge base of the Geography Division's overall flow, picture, plan	•					◑	×			×	• Noted that priorities shift and no one, unless directly involved, understands what the process is
7.38	**E**	Increase system flexibility	•					○	×	×	×	×	
7.39	**E**	Increase system portability	•					◕	×	×	×	×	

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.40	**E**	Integrate the U.S. Postal Service files completely with databases	•					◕	×	×	×	×	
7.41	**E**	Introduce standards	•					◑	×	×	×	×	
7.42	**E**	Maintain a current MAF/TIGER data base	•	•		•	•	◕	×	×	×	×	• Suggested data be less than 60 days old
7.43	**E**	Maintain historical data	•		•			○	×	×	×	×	• Allows for reconstruction of base information at any point in time
7.44	**E**	Mandate reporting	•					○	×	×			• Geography Division deliverables take precedence over Geographic reporting
7.45	**E**	Minimize system's labor intensiveness	•					◔	×	×	×	×	
7.46	**E**	Minimize/delete delays in getting promised products to the public	•					◔		×	×	×	
7.47	**E**	Provide a better way to organize and classify data	•					◔	×	×	×	×	
7.48	**E**	Provide ability for a MAF browser to bring up the TIGER data base	•					○	×	×	×		
7.49	**E**	Provide ability for small batch processing	•					◔	×	×			
7.50	**E**	Provide adequate flexibility to allow adding data fields to MAF/TIGER that does not require a restructuring	•					○	×	×			

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.51	**E**	Provide adequate time for quality assurance	●					◔	×	×		×	
7.52	**E**	Provide automatic recovery functionality that can be initiated by clearly defined conditions	●					◔	×	×		×	
7.53	**E**	Provide capability to track the map level production status				●		○		×		×	
7.54	**E**	Provide concurrent file access	●	●				○	×	×	×		
7.55	**E**	Provide data exchange functionality	●					◑	×	×	×	×	
7.56	**E**	Provide documentation on MAFUF	●					◔	×	×		×	•
7.57	**E**	Provide documentation on the Master Address File Geocoding Office Resolution	●					◔	×	×		×	
7.58	**E**	Provide fast database querying capabilities	●					◑	×	×	×		
7.59	**E**	Provide for easier system programming	●					○	×	×		×	
7.60	**E**	Provide for the historical comparison of census tracts	●					◕	×	×		×	
7.61	**E**	Provide for the integration of the MAF and TIGER data bases so that a boundary change will automatically update the corresponding geocoding	●					○	×	×	×	×	• This is the most essential requirement for accurate and correct geocoding and products (like LUCA)
7.62	**E**	Provide functionality to have the MAF queried and manipulated directly via the Internet by the Bureau and external organizations	●	●	●		●	○	×	×	×	×	
7.63	**E**	Provide functionality to reference the PLSS	●					○	×	×	×	×	

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.64	**E**	Provide interactive MAF processing	●					○	×	×	×		
7.65	**E**	Provide log file of why systems fail by documenting and analyzing previous system failures	●					◑	×	×		×	
7.66	**E**	Provide MAF extracts for the Decennial Master Address File at the county level	●					◕	×	×	×		
7.67	**E**	Provide means for continuing ongoing work while transitioning to modernized system	●					◕	×	×	×	×	
7.68	**E**	Provide method for assessing quality of MAF data	●					◔	×	×	×	×	
7.69	**E**	Provide method of distributing Local Update of Census Addresses participant addresses to appropriate geographic entity	●					●					
7.70	**E**	Provide more structured development	●					◔		×		×	
7.71	**E**	Provide proper system documentation	●		●			◔	×	×		×	
7.72	**E**	Provide same geocoding in both TIGER and MAF data bases	●					◑	×	×		×	• MAF has geocodes but not the same as TIGER®. MAF has census geocode (block) and a TIGER geocode (could be different) until the TIGER® is updated

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.73	**E**	Provide self-benchmarking capability for the MAF	●					○	×	×		×	
7.74	**E**	Provide standards of data quality to include reference materials, specs, and procedures	●		●			○	×	×	×	×	
7.75	**E**	Reduce the learning curve of database functionality and underlying data structures	●					○	×	×	×		• Use of commercially available software suggested to reduce this learning curve
7.76	**E**	Remove data partitions	●		●	●	●	○	×	×		×	
7.77	**E**	Retain qualified IT professionals	●					◔				×	• Term appointments (through September 30, 2001) hurt retention • Salary limitations hurt retention
7.78	**E**	Standardize software used in field operations	●					○		×	×	×	• There are 12 different software versions in the field
7.79	**E**	Support ad hoc MAF queries for new surveys and sample expansions				●		◕	×	×	×		

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.80	**E**	Support address breaks and address range extracts	●		●			◕	×	×			• In accordance with federal regulation (Title 13) the Bureau can divulge only address ranges, not individual addresses
7.81	**E**	Support Congressionally mandated programs	●		●			◑	×	×	×	×	• Home Mortgage Disclosure Act, Enterprise Zones, Empowerment Zones, etc.
7.82	**E**	Support density assessment maps	●					◑	×	×			
7.83	**E**	Support interfaces for the extraction from and input into MAF and TIGER data bases	●					◑	×	×	×		
7.84	**E**	Support plots for internal products such as GUSX in the geographic database	●					◕	×	×			
7.85	**E**	Support rollback by time stamp	●		●			○	×	×	×	×	
7.86	**E**	Support the 1990-derived products of the Geographic Areas Branch	●					◔	×	×	×	×	
7.87	**E**	Support the automated geocoding of Puerto Rico	●					◔	×	×	×	×	
7.88	**E**	Support the community address update system with DSMD	●					○	×	×			

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.89	**E**	Support the input of yearly American Community Survey update information	•					◕	×	×	×		
7.90	**E**	Support the storage of any addressing system	•					◔	×	×			
7.91	**E**	Support the storage of historical geographic data for historical comparison purposes	•					◑	×		×	×	
7.92	**E**	Support the twice yearly MAF extracts for the American Community Survey				•		◕	×	×			• Starting in July 2002, full country coverage is expected
7.93	**E**	Support the use of the Bureau addresses for government contact data	•					◑	×	×	×		
7.94	**E**	Support the use of spatial databases from local and tribal governments	•					◔	×	×	×	×	
7.95	**E**	Support TIGER/Line® products in the geographic database	•					◕	×	×			
7.96	**E**	Support use of the U.S. Postal Service DSF	•					◕	×	×			
7.97	**E**	Support variety of methods of input for boundary changes, to include those from TIGER/Line®, DEX files, and updates from school districts	•					◔	×	×	×	×	
7.98	**E**	Synchronize the MAF and TIGER identification numbers to link geocoding	•		•			◔	×	×		×	
7.99	**E**	Test modernized system as developers build system			•			○	×	×		×	• Q/A must be incorporated into system development

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.100	**E**	Understand MAF inputs before running processes	●					◔	×	×		×	
7.101	**E**	Update read file as soon as the live file is updated	●					◕		×		×	
7.102	**E**	Use CATI geocoding for sponsor-provided telephone numbers or addresses with a telephone number provided by a commercial source				●		◔	×	×			• This approach works faster and cheaper than a field operation • OMB approval needed
7.103	**E**	Use CAUS field staff to locate and map spot ungeocoded address				●		○				×	• Staff uses local sources • For travel efficiency, the uncoded addresses must be grouped into compact areas
7.104	**E**	Validate map spots	●					◕	×		×	×	• Use GPS/imagery
7.105	**D**	Allow for the storage of rules pertaining to geographic information for maintaining relationships	●					○		×		×	
7.106	**D**	Automate conversion to accurate coordinate system	●					◑	×	×		×	
7.107	**D**	Combine MAF and TIGER data bases into a single database	●					○	×	×		×	
7.108	**D**	Develop generic TIGER input/output (I/O) routines	●					○	×	×		×	

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement: U.S. Census Bureau Customers						21st Century MAF/TIGER Enhancements				
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.109	**D**	Develop in-house, Web-based tool to download MAF data	●					○	×	×	×	×	
7.110	**D**	Develop procedures for extracting data at the lowest geographic levels	●					◔	×	×		×	
7.111	**D**	Develop strategy for updates from local and tribal entities that usually do not update databases	●		●			◔	×	×	×	×	
7.112	**D**	Document differences between versions of TIGER/Line®	●					○	×	×		×	• Metrics may include miles of paved roads
7.113	**D**	Eliminate in-house software limitations	●					○		×			
7.114	**D**	Employ GPS data to assist with accuracy assessment	●					○	×	×	×	×	
7.115	**D**	Ensure development machines are available	●					◔		×		×	
7.116	**D**	Ensure that data Czar position is staffed by qualified, senior personnel	●		●			○	×	×		×	
7.117	**D**	Ensure the GPS and imagery are at the same level of accuracy	●					○	×	×		×	• Row house example
7.118	**D**	Geocode addresses to latitude and longitude coordinates	●					○	×				
7.119	**D**	Geocode by building matching	●					◔	×	×	×	×	
7.120	**D**	Geocode by intersection matching	●					○	×	×	×	×	
7.121	**D**	Have a data Czar with ultimate authority and responsibility for deciding what information can be entered into the database	●					○	×	×	×	×	

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.122	**D**	Improve support for different versions of software	●					○		×		×	• This has not been done previously because of space limitations
7.123	**D**	Include "data previously extracted" flag on MAF data	●					○	×	×		×	
7.124	**D**	Include business and system process reengineering	●					◔	×	×	×	×	
7.125	**D**	Include system templates on the Web	●					○	×	×	×	×	• Control system noted as one of the largest areas of bottlenecks
7.126	**D**	Include Zip Code information in the MAF	●					◑	×	×	×	×	
7.127	**D**	Integrate quality assurance into the production process	●					◕	×	×		×	
7.128	**D**	Keep copy of all data that was changed or overwritten					●	○	×	×	×	×	
7.129	**D**	Maintain a commercially available database engine	●		●			○		×			
7.130	**D**	Maintain a date of receipt for all data received into the system			●			○	×	×	×	×	
7.131	**D**	Maintain rules regarding what data system will and will not accept	●					○	×	×		×	• This would follow the model of the Post Office
7.132	**D**	Maintain up-to-date list of participant program contact information	●					◑		×	×	×	

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.133	**D**	Make entire database available to public, less Title 13 information	●					◕	×	×	×		
7.134	**D**	Make raw TIGER information accessible to the public	●					◕	×	×	×	×	
7.135	**D**	Make TIGER® more available to non-mapping users on alternate processing environments	●					○		×	×		
7.136	**D**	Produce Boundary and Annexation Survey in a wholly electronic format	●					○	×	×	×	×	
7.137	**D**	Produce generated products in a more modern fashion (i.e. nonbatch processing)		●	●			○		×		×	
7.138	**D**	Produce map of Puerto Rico in Spanish	●					◑		×			• Parts of the map are in Spanish
7.139	**D**	Produce thematic maps quickly	●					◔	×	×	×	×	
7.140	**D**	Provide ability to automatically enter the next available map spot number to the next map spot			●			○	×	×		×	• Would like to be able to enter 1 and 95 and have the computer evenly space all others that are between them
7.141	**D**	Provide additional resources to the quality assurance of Geography Division products	●	●				◔		×		×	
7.142	**D**	Provide alternate symbolization on map products	●					◕	×	×			• Currently, supporting files go out with maps

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.143	**D**	Provide better instruction on how to use GUSX	●					◑	×	×		×	
7.144	**D**	Provide better operation controls over releases of TIGER®	●					○	×	×		×	
7.145	**D**	Provide capability to log on once for access to entire system				●		○		×	×	×	
7.146	**D**	Provide data in a seamless manner so that moves between states/counties are not apparent			●			○	×	×	×	×	
7.147	**D**	Provide documentation on Decennial Master Address File extracts	●					◑	×	×	×		
7.148	**D**	Provide documentation on Local Update of Census Addresses via the intranet	●					◑	×	×	×		
7.149	**D**	Provide easy correction functionality to map and geography linkage	●					◔	×	×		×	• Want to be able to correct both simultaneously
7.150	**D**	Provide for easier training on MAF/TIGER	●					○		×		×	
7.151	**D**	Provide functionality to collect school grade ranges for every district			●			○	×		×		
7.152	**D**	Provide historical ZIP Code information	●					○	×	×		×	
7.153	**D**	Provide improved file navigation capability	●					◕		×		×	
7.154	**D**	Provide information on the entire Navajo Nation on a single disk					●	◕		×	×		
7.155	**D**	Provide MAF/TIGER data on multiple platforms to external customers	●				●	○	×	×	×		• At a minimum, must be able to access data on Mac and PC platforms

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.156	**D**	Provide regeocoded year 2000 MAF every year to reflect updated legal boundaries			●		●	○	×	×		×	
7.157	**D**	Provide system to store spatial and geographic rules	●					○	×	×		×	• This will prevent special exceptions being written into the software
7.158	**D**	Provide the ability to geocode substructures	●					◑	×	×			• This becomes necessary for shopping centers; individual stores in the mall change over time—need to geocode individual businesses
7.159	**D**	Provide the capability to produce ad hoc queries and user defined reports				●		○		×	×		
7.160	**D**	Provide the effective geocoding of city-style addresses	●					◑	×	×	×	×	• Noted in meeting that this cannot be done currently
7.161	**D**	Provide TIGER® in a format that is easily loadable to local systems	●					◑	×	×	×		
7.162	**D**	Provide undo functionality	●					○	×	×		×	
7.163	**D**	Remove dual classification on "live" and "nonlive" TIGER data			●			○	×	×			• Suggested removal of nonlive data completely

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.164	**D**	Restructure applications individually	●					○	×	×		×	
7.165	**D**	Standardize geographic level of produced products	●					◑	×	×		×	• Some are county based, others state based, etc.
7.166	**D**	Support digital file exchange	●	●				◔					
7.167	**D**	Support extracts from ZIP Code tab areas	●					◔	×	×	×		
7.168	**D**	Support plots for internal products such as ARF in the geographic database	●					○	×	×			• ARF is used for matching responses that cannot be modified using TIGER® • ARF extracts information from TIGER® because current TIGER data base is not designed to support their needs
7.169	**D**	Support the LandView product suite that derives boundaries from TIGER/Line® geographic data base			●			◕	×	×			

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
7.170	**D**	Understand information as to the dependencies across activities and dependencies between and within activities				●		○		×	×	×	• Currently, Geography Division encodes dependencies among activities into TIGER/MAF batch processes as they are uncovered • Such information, if made available to the RCCs, would allow them to understand where a partition lies in a batch process
7.171	**D**	Use commercial tools that support large volume map making	●					○	×	×		×	
7.172	**D**	Use commercial-based software to draw maps	●					◔		×	×	×	
7.173	**D**	Use E911records to improve records of housing unit counts					●	◔	×	×	×	×	
7.174	**D**	Use workflow management tools	●					◔	×	×	×	×	
7.175	**O**	Build maps dynamically over the Internet	●					○		×	×		
7.176	**O**	Correct data from previous censuses to make historic information spatially correct	●					○	×	×		×	• This references rollback requirement

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.177	**O**	Create utility program	●					○	×	×			• Consider maintaining UNIX
7.178	**O**	Develop a Geography Advisory Committee (composed of external Census support organizations) to help with the development of the modernized system					●	○	×	×	×	×	
7.179	**O**	Perform a comprehensive review of the 2000 Census to determine the accuracy of the MAF/TIGER data bases			●			○	×	×		×	
7.180	**O**	Provide ability to temporarily erase or suppress map overlays	●				●	○	×	×			• Need to be able to get to the base map
7.181	**O**	Provide MAF/TIGER information spatially to facilitate better reporting	●					○	×	×	×	×	
7.182	**O**	Provide on-line documentation on TIGER and public code information	●					○	×	×	×	×	
7.183	**O**	Provide public interactive geocoding functionality on line	●					◔	×	×	×	×	
7.184	**O**	Replace data entry screens with point and click capability				●		○		×		×	
7.185	**O**	Report activities by Type of Enumeration Area	●					◔	×	×		×	
7.186	**O**	Standardize operating procedures that create a quality check before local information is entered into the system	●	●				○	×	×	×	×	
7.187	**O**	Support the high-level mapping of Canada and Mexico	●					○	×	×	×		

7.0 Improve Accuracy and Responsiveness of Geographic Support													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
7.188	**O**	Support ZIP+4 files	•					◕	×	×			
7.189	**O**	Use an Object-Oriented programming language	•					◕	×	×			
7.190	**O**	Use commercial ZIP Code data	•					◑	×	×			
7.191	**O**	Use imagery to improve data quality	•		•			○	×	×		×	

8.0 Other Observations													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	**Level of Importance**	**Detailed Requirement (by Major Category)**	**GEO Program Managers**	**Field Operations Managers**	**Program Sponsors/ Analysts**	**Census Documents**	**External Users/ Customers**	**Base Program**	**MAF and TIGER Data**	**New Processing Environment**	**Geographic Partnerships**	**Improve Quality**	**Comment/Issues**
8.1	**E**	Decrease duplication of activity in the Geography Division	●					◕	×	×		×	
8.2	**D**	Develop custom-developed software to feed back into systems for continuous improvement	●					○	×	×		×	
8.3	**D**	Develop software to report and track Q and A findings	●					○	×	×		×	
8.4	**D**	Improve communication between branches within the Geography Division	●					◔				×	
8.5	**D**	Increase knowledge base of staff performing quality assurance	●		●			◑				×	
8.6	**D**	Produce planning tools for project management	●					○	×	×	×	×	
8.7	**D**	Provide updated project tracking tool	●					○		×		×	
8.8	**D**	Stop spending large amounts of time on customization of short-lived applications	●					○	×	×			
8.9	**D**	Stop placing programming concerns over the concerns of geographic products	●					○	×	×			
8.10	**O**	Move the responsibility of the MAF outside of the Geography Division					●	○			×		• MAF is about housing, not addresses, and needs to be worked by people who understand housing groups

8.0 Other Observations													
			Source of Requirement						21st Century MAF/TIGER Enhancements				
			U.S. Census Bureau Customers										
Requirement Number	Level of Importance	Detailed Requirement (by Major Category)	GEO Program Managers	Field Operations Managers	Program Sponsors/ Analysts	Census Documents	External Users/ Customers	Base Program	MAF and TIGER Data	New Processing Environment	Geographic Partnerships	Improve Quality	Comment/Issues
8.11	**O**	Stop writing detailed specifications for software development	●					◔		×		×	• Stated need to develop "on the fly" because of time constraints
8.12	**O**	Support international census needs with the modernized MAF/TIGER			●			○	×	×	×	×	

APPENDIX B. MOBILE COMPUTERS AND GLOBAL POSITIONING SYSTEM

Beginning with the first census in 1790, paper maps have been vital to placing the enumerator at the doorstep of each resident. The U.S. Census Bureau has been improving its processes for 200 years and has called upon technology to help make the census more efficient and economical. One of these improvements for the next census will be the use of mobile computers with electronic maps.

The challenge for the Bureau is ensuring that every person in America is counted once, and only once. For the past two decennial censuses, the census questionnaire mailings to the public have been receiving response rates of about 65 percent. An enumerator must visit the remaining 35 percent of the housing units to collect the census data. Mobile computers and the Global Positioning System (GPS) can support this door-to-door collection effort. The new direction of technology will allow enumerators to electronically validate the location of housing units that need to be visited. In addition, mobile computers could allow for collected census questionnaire information to be input directly while the enumerator is at the housing site and for the information to be transmitted directly to the Regional Office.

Mobile computers for the enumerators will contain the housing addresses of the housing units, written and interactive map directions to each housing unit, data sheets to post collected questionnaire data, and a voice and data communication device for relaying data. The device will include encrypted security measures to prevent census data from being intercepted.

This chapter provides information on the technology for mobile computers and GPS. An exhaustive list of these devices is not presented; rather, the capabilities of both industries are represented.

B.1 Mobile Computer Systems Overview

A convergence of powerful, inexpensive hardware; standardized communication protocols; and innovative software has provided the Bureau with several options for deploying geographic information system (GIS) layers in the field. Information technology (IT) industry analysts agree that it is unlikely that field GIS applications have reached more than a small percentage of their potential market. Even among organizations with a field computing strategy, most do not yet include a significant GIS component. The reason for this untapped potential is that GIS applications are complex. They are difficult to deploy in an office setting and even more so in the field. Large, complex data structures and equally complex user interfaces do not easily lend themselves to field use.

To develop effective field systems, developers are bringing together state-of-the-art components from various enabling technologies, including wireless communications networks, Internet-based data delivery protocols, and a new generation of field-ready hardware and software applications. Fundamental technical challenges must also be met with an integration of specific technologies to create a reliable and functional mobile engineering environment.

In the field, enumerators with little or no exposure to computers will use one or two highly specialized applications. The leap from paper to sophisticated new applications can present major challenges in retraining and developing new fieldwork processes. Field-ready applications must be designed to maximize ease of use and functional efficiency in the field. Applications must be task-specific, simple to operate, and built to avoid potential errors.

Hardware is just beginning to effectively address the particular needs of field workers. The price point for hardware that is suitable for the rigors of the field is also declining. Field systems must be used in remote locations miles from the nearest building, electrical outlet, or phone socket. They must be able to withstand exposure to rain, snow, mud, dirt, and extreme temperatures.

In the near term, bandwidth will continue to be the primary limiting factor. Limits of 19.2 kilobits per second (Kbps) are suitable for dispatch and simple work-order data; but larger data files, like geospatial imagery and vector-based maps, are still accessed better via CD ROM or hard drive. That situation will improve dramatically during the next 2 to 5 years. Organizations are working to deliver low-cost broadband systems, optimizing their products for small-packet data formats that can function comfortably under low bandwidth constraints.

Mobile computing requires a robust server and specific applications that integrate with key organization and operations management systems. As the backbone of a mobile information environment, the server must be scalable and powerful enough to support multiple field users. Data gathering will take place on handheld computers. The actual processing will occur at a site powerful enough to provide the processing horsepower that is lacking in the field. For several applications, the server will also need to support analysis and workflow management processes. To accommodate users, the server must be able to provide results of these processes as quickly and transparently as possible. Such mobile systems will also need to integrate with other field technologies such as the GPS.

B.1.1 Software

The software products profiled below represent only a sample of the mobile GIS mapping and asset management systems. Environmental Systems Research Institute (ESRI), PenMetrics, Condor Earth Technologies, MapInfo, and Tadpole Technology are software vendors within this market and have products that meet the enterprisewide and fully scalable requirements of the MAF/TIGER Modernization initiative.

B.1.1.1 ESRI ArcPad. ArcPad software is an easy-to-use, lightweight, low-cost solution for handheld and mobile mapping and GIS. ArcPad is designed to enhance portable touch-screen computers with intuitive mapping, GIS, and GPS functionality. ArcPad makes field data collection fast and easy, improves data accuracy, and provides immediate data availability and validation. Figure B-1 illustrates the capabilities and functionality of the ArcPad.

ArcPad leverages existing mapping and GIS software systems and databases. When a user is finished editing data in the field, changes and additions can be uploaded into the master database in the office. ArcPad supports a multilayer environment with industry-standard vector map and raster image display that includes the ability to view aerial and satellite imagery. ArcPad allows users to create custom input forms from their existing GIS database for use in the field.

Figure B-1. Capability and Functionalities of the ArcPad

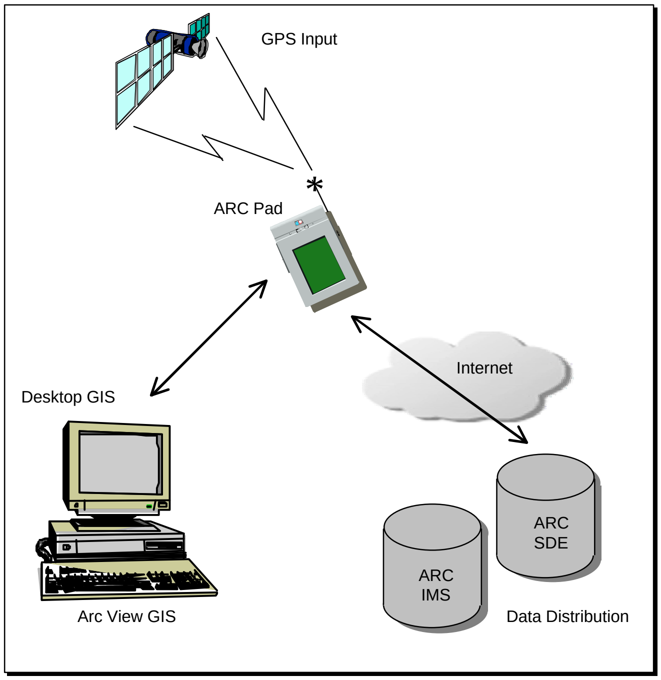


Image courtesy of ESRI, Inc. (Redlands, CA)

ArcPad can also be integrated with an optional GPS or Differential Global Positioning System (DGPS). ArcPad enhances the GPS input by not limiting position information display to numerical coordinates but displaying it with a moving crosshair on an actual map. This enables users to view their exact location in the context of a map or image.

ArcPad can use data directly from an individual's desktop or an organization's enterprise GIS system without the need to convert to unique portable formats. ArcPad uses vector data in ESRI shape-file format (as used by ArcInfo, ArcView GIS, ArcIMS, and other ESRI software programs). ArcPad directly supports the use of the following raster image formats: JPEG, MrSID (compressed images), Windows Bitmap, and CADRG raster maps.

The ESRI ArcPad software costs $500 per license and the user, at an additional cost, defines the hardware options.

B.1.1.2 PenMetric's FieldNotes for Mobile Mapping. FieldNotes for Mobile Mapping (FNMM) eliminates paper map sheets by permitting field enumerators to take digital maps into the field on pen-based or laptop computers. FNMM is a complete, portable, enterprisewide GIS software suite with powerful and easy-to-use map-viewing functionality, pop-up graphical user interface (GUI) windows, and the ability to integrate add-on modules. The GPS module allows the user to acquire and display coordinate and positional information by providing a software/hardware interface linking the GPS receiver to the mobile computer running FNMM.

Applications that do not require maps or GIS data, such as the current U.S. Census 2000 short- and long-form questionnaires, can also be integrated into FNMM. This product is a user-customizable, 32-bit application that provides the enterprisewide performance and flexibility to construct non-map-based mobile field automation solutions. The application does not require intensive programming scripts or macros to assemble custom templates for data gathering activities.

FNMM software costs approximately $1,500 to $2,000 per license. Hardware is an additional cost.

B.1.1.3 Condor Earth Technologies PenMap. PenMap is a real-time, surveying, mapping, and GIS data collection software application that operates on any personal computer (PC), laptop, or pen computer running on Windows. PenMap has been integrated into several field data collection instruments, including reflectorless laser range finders, digital cameras, environmental sensors, and GPS receivers. Information collected in the field can be viewed and edited as it is collected, thereby eliminating unnecessary and costly remobilization.

With PenMap, the user is not limited to one instrument or measurement method as is the case with other systems. User-defined data entry forms and menu buttons permit job-specific data entry. Raster and vector background images can be imported to PenMap and taken to the field.

The PenMap software costs about $1,500 per license and runs on hardware using the Windows operating system. Hardware is an addition cost.

B.1.1.4 MapInfo MapXtend. MapInfo's MapXtend is a wireless spatial technology that enables organizations to deliver their enterprise applications over the Internet to field workers. Applications built with MapXtend wireless spatial technology could allow enumerators to access live MAF/TIGER data via handheld Personal Digital Assistants (PDA) for faster and more efficient scheduling and servicing operations. MapXtend uses MapInfo's entire suite of Internet geocoding, routing, and mapping technologies to provide location-based intelligence for the enterprise field worker. MapXtend extends two-way wireless communications linking a handheld client and a server over the Internet. Figure B-2 shows the functionality of the MapXtend.

MapInfo's MapXtend costs approximately $1,500 to $2,000 a license. Hardware is an additional cost.

Figure B-2. Functionality of the MapXtend

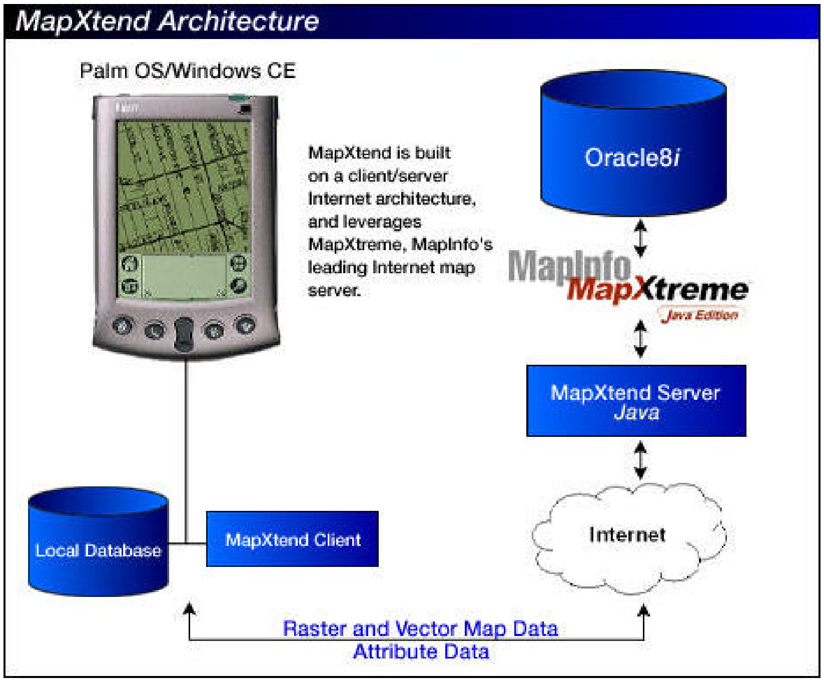


Diagram Courtesy of MapInfo Corporation (Troy, NY)

B.1.1.5 Tadpole Cartesia. Tadpole Technology has developed the Tadpole Cartesia field information system. Cartesia is a blend of Java software tools that take advantage of organization workflow to the mobile environment. The Cartesia is a field information system, combining simple-to-use, specialized software with rugged hardware that integrates critical relational database management system (RDBMS) information and detailed GIS mapping data and delivers it into the hands of field enumerators. The Cartesia spatially enabled field information system provides field teams on-line access to regional databases via digital cellular or conventional telephones from designed-for-the-field handheld computers.

The Cartesia software interfaces with organizations' existing systems using intranet, Java, Common Object Request Broker Architecture (CORBA), and RDBMS connectivity tools to assemble and feed customized, task-specific job files to field teams. Cartesia-EFS works with a GIS program to extract detailed mapping information and makes this information available where it is needed. Once assembled by Cartesia interfaces, job files are made available locally to field enumerators to view and manipulate processed information. This includes full capability to interact with maps and forms-based text. Status messages can be used to track work in progress. As each task is completed, job information can be uploaded from the pen-based computer to organization headquarters where Cartesia-EFS returns it to the appropriate enterprise system. In the case of GIS, Cartesia provides a separate quality assurance application allowing GIS controllers to audit the quality of the captured information before it is committed to the enterprise system.

The cost of the Cartesia software is approximately $1,500 to $2,000 per license. Hardware is an additional cost.

B.1.2 Hardware

The hardware products profiled below represent only a sample of the mobile GIS mapping and asset management hardware systems. Xplore Technologies, Compaq, Xybernaut, Palm Computing, and Tadpole-RDI are hardware vendors within this market and have products that have the potential to meet the requirements of the MAF/TIGER Modernization initiative.

Costs vary widely among the hardware products according to their capabilities, ruggedness, and functions. At the high end, the products are built to withstand the rigors of varied environments and are expandable to accommodate add-on functions. In addition, they may be highly specialized, suiting the needs of a specific user with little or no customization. At the low-end, the units are generally hand- or palm-held and engineered for many purposes. They are not ruggedized and may have limited capability for optional functionality.

B.1.2.1 Xplore Technologies Rugged GenSys Pen Computer. Xplore's GeneSys Systems, shown in Figure B-3, is a rugged field computer. The Ultra Hi-Brite Color display can display maps, data, and forms, whether in direct sunlight or dimly lit conditions. Users interact with the GeneSys using a passive stylus or their finger. Intel Pentium processors, Peripheral Component Interconnect (PCI) architecture, and Xplore's own "Performance Xcellerators" provide the power needed. The structure is an ergonomic, hardened magnesium/aluminum housing, which is sealed and protected from the elements. The Xplore is ruggedized and meets International Standards Organization (ISO) 9002 manufacturing program standards. Xplore communications options are flexible and simple to implement. Wireless fax/modems, spread spectrum technology, and GPS are all available via Xplore's unique XPod Expansion Modules, as shown in Figure B-4.

Figure B-3. Xplore GeneSys System

Image Courtesy of Xplore Technologies Corp (Austin, TX)

Figure B-4. GeneSys XPod Expansion Module

Image Courtesy of Xplore Technologies Corp (Austin, TX).

The XPod can be attached to the GeneSys Rugged Mobile Pen Computer to provide options for new technologies. The patent-pending XPod Expansion Module houses native communication ports.

The hardware cost per unit is approximately $ 7,000 to 9,000. Software costs are separate.

B.1.2.2 Xybernaut Mobile Assistant IV with Datria's VoCarta. Xybernaut Mobile Assistant IV is a wearable computer, and Datria's VoCarta is a speech-to-data software product that makes collecting enumerator data easier and quicker. Speech-empowered mobile computing can assist in improving field productivity with fast and accurate data collection, real-time non-stop data input, and freedom from traditional computing limitations. Xybernaut's Mobile Assistant IV allows for true PC portability. The Datria's VoCarta enhances multimedia data collection by incorporating speech with a multitude of sensors while freeing workers from pen scripting and keyboard typing.

The lightweight system, shown in Figure B-5, gives enumerators the freedom to move easily while collecting information. Whether for MAF/TIGER updating, locating housing/business units, or editing boundary delineation, the Mobile Assistant IV and VoCarta integrated solution would permit real-time inspection and quality checks with simultaneous data input and updating. Voice input and the lightweight touch screen will maintain smooth, continuous data entry while updating the database, reporting items, and providing complete and formatted reports of the work performed.

The hardware cost per unit is approximately $7,000. Software costs are separate.

Figure B-5. Xybernaut Mobile Assistant IV

Image Courtesy of Xybernaut Corporation (Fairfax, VA)

B.1.2.3 Tadpole-RDI J-Slate. The Cartesia J-Slate diskless, rugged handheld pen computer, as shown in Figure B-6, has been specifically designed for field workers. The J-Slate provides field enumerators with timely, accurate, and job-specific data and detailed maps.

Figure B-6. Tadpole Cartesia J-Slate

Image Courtesy of Wind River Systems (Alameda, CA)

Featuring a 10.4-inch daylight viewable color display and touch-screen technology, the battery-powered J-Slate computer is built around the high-performance StrongARM processor. This architecture is ideally suited to a pen computer running graphic-intensive applications. J-Slate's efficient power consumption enables the system to run for a complete workday without access to a power source.

The J-Slate has configuration options that include integrated cellular phone, radio, and GPS as well as modem and Ethernet connections. Keyboard and vehicle docking cradles are also available.

The hardware cost per unit is approximately $7,000. Software costs are separate.

B.1.2.4 Compaq Aero 2180. The Compaq Aero 2180 is a palm-held device that comes with 24 megabytes (MB) of memory for extra applications and storage space. The device comes with QMenu, QUtilities, File Explorer, and other programs. The color depth on the Aero 2180 (65,536 colors) gives it a clear, crisp screen that is ideal for maps, color photographs, and other color information. Its reflective screen makes it viewable in all lighting conditions, even direct sunlight (see Figure B-7).

The Compaq Aero 2180 operates with Microsoft Windows CE. Applications include character recognition, digital keyboard, voice recorder, e-mail in box, task pad, contacts, calendar, calculator, Active Desktop (customizable desktop), notepad, active synch (to synchronize with information on PC), mobile channels (for downloading specific Web content), and back-end functionality to connect to remote computers.

The cost per unit is approximately $450. Mobile mapping software costs are separate.

Figure B-7. Compaq Aero 2180

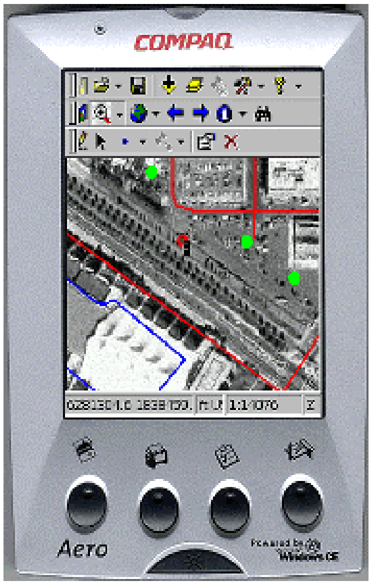


Image courtesy of ESRI, Inc (Redlands, CA)

B.1.2.5 Palm Vx. 3Com's new Palm Vx, shown in Figure B-8, comes with 8 MB of random access memory (RAM) for storage and the latest version of the Palm operating system (OS), which is designed to offer significant performance and functionality improvements over the older Palm OS architecture.

The Palm Vx has features and functionality that include four programmable application buttons, a backlit display, infrared port, and rechargeable lithium-ion batteries. The included docking cradle serves as both a battery charger and a conduit for synchronizing with a PC.

The Palm Vx's extra storage offers users the freedom to operate mobile, enterprisewide third-party applications such as MapInfo's new MapXtend field-ready GIS software, while leaving room for contact information. The Palm Vx can store approximately 10,000 addresses; 3,000 appointments; 3,000 to-do items; 3,000 memos; and 400 e-mails.

Figure B-8. Palm Vx

Image Courtesy of Palm, Inc. (Santa Clara, CA)

In addition, 3Com improved the operational aspects of the Palm Vx by integrating it with the latest update of the Palm OS, version 3.3. The major improvements in this iteration focus on faster data transfer via the docking station and extended infrared functionality.

The cost per unit is $400. Mobile mapping software costs are separate.

B.1.2.6 Palm VII. Palm VII's wireless Internet product is iMessenger, a wireless e-mail client for sending and receiving messages (see Figure B-9). Although it is probably a good device for the functions for which it was designed, it has severe limitations for the needs of MAF/TIGER Modernization. By combining it with wireless connectivity software with its low 2 MB of RAM, there is virtually no free space to run a mobile GIS data collection solution with this unit. Also, an issue exists with the Palm VII's service coverage area. While Palm.Net's wireless service blankets the most populated parts of the country, rural users are likely out of range. This lack of coverage would hamper the Census Bureau's mission of mapping and collecting attribute data in suburban and rural areas, where gaps in service areas reduce production efficiency.

The cost per unit is $500. Mobile mapping software costs are separate.

Figure B-9. Palm VII

Image Courtesy of Palm, Inc. (Santa Clara, CA)

B.2 Global Positioning System Overview

The U.S. Census Bureau has been conducting research on the capabilities of GPS during the past decade. The need for accurate housing units and street intersection locations has been the driving force for incorporating GPS data into TIGER files. The accuracy of GPS coordinates has been under review. With the potential introduction of imagery in support of TIGER alignment, comparable accuracy of GPS and imagery will be necessary. If their accuracies are not comparable, the discrepancies between TIGER data and GPS data would be obvious to all users and raise questions about accuracy. In addition to ensuring the accuracy of TIGER data, the GPS capabilities would assist enumerators in locating housing units in the field without running repeat visits. With the elimination of selected availability on May 1, 2000, the accuracy capabilities are at 3–4 meters on the average. Selected availability is intentional Department of Defense (DoD) interference to degrade the GPS signal to commercial GPS receivers, for security reasons. This section provides details of the GPS capabilities that can support the modernization of MAF/TIGER.

GPS manufacturers are interested in many consumer markets, but portable GPS packages clearly target GIS and mapping users. Most GPS mapping systems consist of handheld receivers or backpack systems that link GPS receivers to external data controllers, collectors, or PCs running data collection and real-time mapping applications. The packages more frequently use real-time differential (D)GPS because many systems now incorporate smaller, lighter, and less-expensive integrated beacon receivers and antennas. Therefore, operators can achieve greater GPS precision through real-time differential correction via free or subscription-based services.

Depending on the system, collected GPS data can be stored in internal memory or add-on memory cards. However, a variety of powerful data loggers and rugged field computers (many pen-based) are used for mapping and asset management applications. Personal assistants (like Palm Pilots) and small Microsoft Windows CE-based handheld devices also are used. Software programs for GPS mission planning, data collection, editing, post-processing, and conversion are becoming better integrated, more functional, and easier to use. In addition, user interfaces are improving, and more flexible and substantial GPS data analyses are possible, allowing GPS users to assume more control and produce reliable GPS-derived GIS and computer-aided design (CAD) products.

GPS-based tools for GIS data collection are well established and well suited to applications where precise position and detailed attribute information are required. They operate under a few constraints, because they need to be used in relatively open areas to receive the GPS signal, but offer the highest levels of positional accuracy, coupled with ease-of-use, low cost, and the ability to interface to complex attribute collection software.

The GPS/GIS field procedure is simple: while the field worker describes the attributes of a feature (using keypad, stylus, or barcode wand), the GPS receiver computes its position. This information is stored on the field computer and subsequently downloaded for export to a GIS via an industry-standard interchange format. In the case of a point feature, the GPS data may be reduced to a single position. However, in the case of polylines and polygons, each feature has a number of associated GPS positions and other attributes.

GPS technology is still relatively new and evolving rapidly. Many of the improvements in GPS/GIS systems are related more to packaging, ergonomics, and data flow than to intrinsic improvements in the GPS components themselves. When GPS-based systems were used solely for mapping and GIS data collection, data flow was one-way, from the field to the office. However, with the advent of data maintenance capabilities, data can flow in both directions.

Wide-area DGPS augmentation systems provide free DGPS signals for the Americas, Europe, and Asia. In the future, most DGPS services at the meter level will be free of charge, with only premium-accuracy services carrying a subscription charge. At the same time, DGPS receivers have become more affordable. Such receivers have been integrated into some GPS systems so that only a single antenna is required to receive GPS signals and DGPS corrections. This level of integration reduces cost and makes systems both smaller and lighter.

B.2.1 GPS Handheld and Integrated Units

The mobile GPS/GIS solutions profiled represent only a sample of the GPS/GIS integrated spatial information management systems on the market. However, each of the profiled vendors offers a wide range of GPS/GIS integrated solutions beyond what can be covered in this report.

Furthermore, Trimble, Baker GeoResearch, and Ashtech/Magellan are top-tier hardware vendors within this market and have specific products that meet the requirements of the MAF/TIGER Modernization Initiative.

B.2.1.1 Trimble GeoExplorer 3 and Pathfinder Pro XR/XRS. The GeoExplorer 3 handheld GPS mapping and GIS data collection/maintenance system is small and light, weighing only 1.4 pounds including the battery. The unit can be connected directly (no cable required) to Trimble's Beacon-on-a-Belt (BoB) differential correction receiver to use free U.S. Coast Guard corrections. The unit includes a 160-by-160 pixel graphical liquid crystal display (LCD) with two levels of backlighting. Users can upload their own vector GIS data into the unit to use as background information on a real-time display to verify and update positions and attributes. About 32,000 positions (1 MB of data) can be stored in the unit's internal memory, so users may have to perform frequent downloads of collected GPS data depending on the type of data they are collecting. However, battery life is long enough to permit a full day of GPS-based mapping. Trimble Pathfinder Office software is included, providing support for GPS mission planning, data import/export, differential correction, plotting, and data dictionary creation.

The cost of the GeoExplorer 3 is $4,500.

Trimble's Pathfinder Pro XR/XRS systems offer more advanced, but less portable, options for GIS data collection. The Pro XR is a 12-channel unit with integrated beacon receiver, and the XRS can receive beacon broadcast signals or differential corrections from wide-area satellite providers. The units can be used with the Trimble data collectors, or they can be connected to an external computer running real-time mapping software. The Aspen real-time mapping software works with all Pathfinder and 4000 series survey receivers, and it runs on notebook or pen-based computers.

The cost of the Pathfinder Pro XR/XRS is $12,000.

B.2.1.2 Baker GeoResearch GeoLink for Windows 95/98/NT and GPS Workhorse. GeoLink V6.0 (which also is referred to as PowerMap) operates with Windows 95/98/NT operating system. GeoLink incorporates a live, moving map display, including multiple vector and raster background maps, to help build map data with user-defined attributes for GIS. GeoLink for Windows 95/98/NT has been updated and expanded to use 32-bit architecture.

Collected GPS data can be saved directly into ESRI shapefile format and/or exported to several different GIS and CAD formats. GeoLink also uses open database connectivity (ODBC), so georeferenced and non-georeferenced data can easily be imported from external databases. GeoLink runs on a user-supplied computer connected via a serial connection to a GPS receiver, GeoLink and supports several different GPS receivers.

The cost of the GeoLink V6.0 is $1200. Computer hardware is not included.

GPS Workhorse is an eight-channel, single frequency (L1) real-time-capable receiver built on Motorola's Oncore receiver, which includes innovative algorithms for satellite tracking in tree canopies and urban canyons. With an added serial port, the receiver can become real-time differential capable to receive wide-area DGPS signal broadcasts.

The cost of the GPS Workhorse is $1200, excluding computer hardware.

B.2.1.3 Ashtech Inc./Magellan Corporation Reliance Field Asset Management System. Ashtech offers a line of Reliance GPS packages to support GIS/mapping and asset-management applications, including the Reliance Submeter (35-centimeter post-processed accuracy, $6,495–$8,495 or $7,495 real time). Prices differ depending on whether the system is bundled with a handheld PDA computer or a rugged pen-based computer. The Reliance RT systems also come bundled with an Ashtech BR2 beacon receiver for real-time differential corrections via U.S. Coast Guard beacons. Real-time DGPS accuracy is 45 centimeters. Non-RT Reliance systems can have real-time capability via appropriate receivers and/or antennas.

All the packages feature Reliance Processor software, which includes modules for mission planning; data collection session management; feature file creation and editing; way-point editing; map viewing of previously collected data; post-processing differential correction; reprojection; and reformatting/export to CAD, GIS, and external database formats. The Reliance PenMap system includes a compact pen computer that runs PenMap real-time mapping software. With this system, background vectors and raster maps can be viewed while GPS data is collected in real-time. The PenMap software is designed to operate like a CAD product, with capabilities for symbolizing and annotating GPS-derived features. PenMap also can integrate GPS data with data from an array of third-party instruments, including laser range finders and imaging devices.

The cost of the Reliance Submeter is between $6,500 and $7,500, depending on the need for real-time data results and computer hardware.

B.3 Laser Range Finders

Laser mapping devices measure distance by precisely measuring the flight time of several pulses of laser light to a target and back to the measuring device. Laser mapping instruments can vary in construction and functional capabilities, but they all have a laser diode, a laser rangefinder's primary sensor, which measures distance to a remote object. Most vendors sell their lower end rangefinders with the basic distance/ranging measuring capability. The higher end laser rangefinders include an inclinometer, which is used to measure the instrument vertical angle, and a compass that measures the horizontal angle of the instrument with respect to magnetic north. These two sensors are required if precise, x, y, and z coordinates are to be calculated for targets from a known ground control point or DGPS receiver integrated with the laser rangefinder.

Laser rangefinders are useful for mapping high-volume features, especially point features such as housing units that are difficult to reach. Housing units under heavy tree canopy or between tall structures can be located and mapped with laser rangefinders. The effective range in which most laser rangefinders can operate is between 5 feet and 2,000 feet. Higher end rangefinders can have a maximum ranging distance of 32,000 feet using an external reflector to enhance the signal return from a longer distance to target.

Laser rangefinders do have limitations around water and in bright conditions. In both these cases, light scattering properties of water and bright sunshine severely hamper the return signal to the rangefinders. Environmental factors such as fog, haze, and dust can also limit the effectiveness of precise measurement. In addition, flat nonreflective surfaces that absorb light particles will be harder to map than a reflective surface that reflects light particles.

The estimated cost of low-end rangefinders is $3,000 to $4,500; the high-end products range from $4,600 to $12,000.

B.4 Summary

The joint capabilities of mobile computers and GPS will enable the U.S. Census Bureau to provide small computer devices to the enumerators, allowing for a more accurate and efficient collection of housing unit locations. This capability will allow enumerators to be tasked via computer, directed to the correct housing unit, assisted with the collection of data, and to upload or download information to the server. The combined capabilities of these devices should reduce enumeration time and ensure more accurate collection of data.

Note: *The discussion of specific vendors and products in this document is for the sole purpose of offering representative examples of products available for specific technology needs. This information is not intended as a promotion of any specific vendor or product. Additionally, the vendors and products discussed herein should not be considered exclusive from other potential vendors and products in the event of future procurement requirements.*

APPENDIX C. IMAGERY ASSESSMENT

One of the major goals of the MAF/TIGER Modernization initiative is to improve the accuracy of the TIGER data base. Throughout the requirements interview process for this initiative, more than 50 organizations have stated their need for improved accuracy of TIGER data. One of the options for improved accuracy is to use aerial or satellite imagery to align existing coordinates with man-made features and existing housing units and place them in the TIGER data base. Improving the accuracy of coordinates will streamline the U.S. Census Bureau operations for gathering accurate housing unit location information in relation to the accurate base map, which will enable enumerators to relocate those structures in the field for follow-up operations. In addition, a more accurate TIGER base map will facilitate digital data exchange with Federal, tribal, state, local, and private partners.

The imagery needs for meeting Bureau requirements are based on the locations of the U.S. population. Those locations within highly condensed urban areas require imagery that will easily distinguish one housing unit from another. Those areas in rural America where housing units are hundreds of feet, if not miles, apart require imagery that distinguishes the road and the structure at a specific location. Additional considerations when choosing the imagery collection systems that best support the TIGER accuracy needs include imagery resolution, accuracy, type, and cost.

The imagery resolution directly affects the accuracy and cost. Spatial resolution refers to the size of the picture element (pixel) that makes up the scan line of the imaging system. The spatial resolution may or may not correlate to the smallest object or ground feature that can be seen on an image. The resolution of current satellite and aerial systems that would be useful to the TIGER modernization process ranges from a minimum of 8 meters in urban areas and 20 meters in rural areas. If these resolutions can be increased to 1 meter for urban areas and 5 meters for rural areas, the accuracy of the TIGER base map will be increased exponentially.

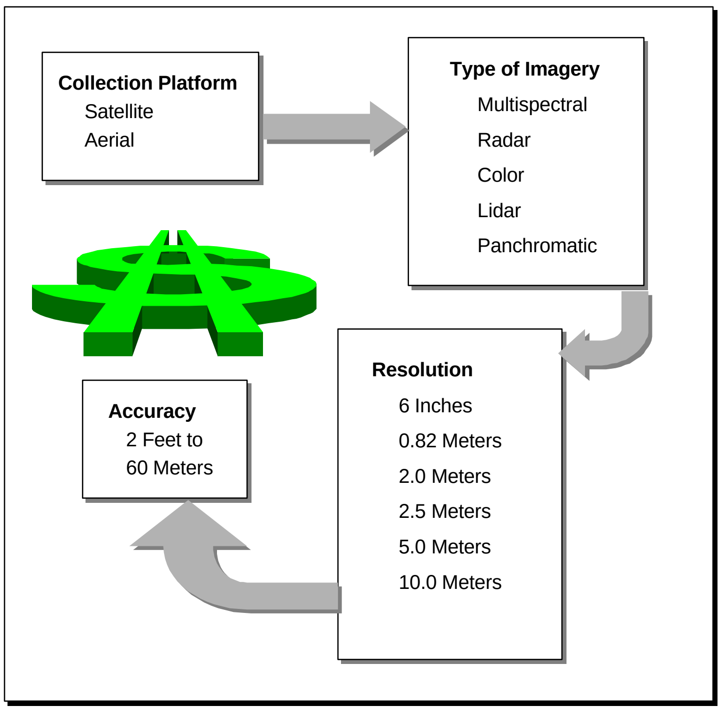
The accuracy of the data reflects a position on the ground to its actual location on the Earth. Within urban areas, this accuracy would need to be to less than 8 meters if the housing units are to be depicted on the correct side of the street with a high degree of accuracy. This 8 meters translates to about 26.2 feet from the centerline of almost all roads and streets in an urban environment. This figure is based on half the road or street width (10 feet) plus the right-of-way (10 feet) and minimal distance to the front door (4 feet). In rural areas this accuracy would not need to be as stringent; 20 meters, or 65.6 feet might be acceptable. If the image shows a road and housing unit that are potentially off by 65.6 feet and the current TIGER data is aligned to it, the image would reveal the correct position of the structure in relation to the location of the road. However, at this gross accuracy the actual structure may appear to be on the side of a cliff or in water.

The type of imagery for the MAF/TIGER Modernization is affected by alignment needs. The purpose of collecting imagery is to align TIGER data to it. The process to accomplish this is still being reviewed. Numerous commercial vendors have been contacted, and most have different approaches to solving this problem. The details of alignment or feature extraction between the imagery and the TIGER data base are discussed in Appendix E.

The imagery type needed depends on the approach to alignment/feature extraction. Some contractors claim the extraction can be done using panchromatic imagery (black and white); others claim the extraction requires color imagery; while yet others say it requires multispectral imagery. Feature extraction can be done using any of the mentioned imagery types. The success of the feature extraction depends on selecting the proper imagery type for the feature of interest. Eventually each imagery capability will have to be reviewed.

The cost of the imagery is based on choices between capabilities of each imagery function. Figure C-1 depicts the type, platform, resolution, and accuracy of imagery. The choices in each box are aligned by cost, with the first choice being the most expensive and the last choice the least expensive.

Figure C-1. Basis for Imagery Cost



The cost of imagery data is dependent on the resolution of the collection and the accuracy required of the imagery. This chapter discusses the imagery sources and attributes and their applicability to the MAF/TIGER Modernization initiative.

C.1 An Imagery Plan

The best imagery options involve collecting data of the entire United States and its territories at 1-meter resolution or covering the densely populated areas (approximately 10 percent of the United States) at 1-meter resolution and the remaining rural areas (approximately 90 percent) at 5-meter resolution. The collection of imagery is only the first step in increasing the accuracy of the TIGER data and represents only part of the cost. Imagery processing is more expensive and time consuming than imagery collection. The level of processing correlates to the accuracy needed. Precision processing, which is the most costly, provides the minimum accuracy that would be needed for aligning TIGER data to imagery. Following imagery collection and processing, TIGER data would be aligned to the imagery. The majority of this process would be automated. TIGER data that did not align satisfactorily with the imagery would be aligned manually. The best approach to manual alignment is using color imagery because of its natural colors. When all of the TIGER data is aligned correctly to the imagery, the TIGER files would be accurate to within at least 8 meters in urban areas and 20 meters in rural areas, assuming 1-meter and 5-meter imagery collections were used. The 1- and 5-meter figures are at the high end. It is estimated that these numbers could be lowered to a nationwide accuracy of 5 meters or less with new technologies in collection and particularly in processing that will be available in the very near future. All possibilities will be reviewed in this chapter.

C.2 Imagery Sources

Many sources of imagery are available to improve the accuracy of TIGER data, including—

- Satellite Imagery Providers
  - Commercial
  - Federal
- Airborne Imagery Providers
  - Commercial
  - Federal
- Potential Satellite Imagery Partners
  - Federal
  - State and Local

The commercial systems fall roughly into two market groups. The first group concentrates on competing for shares of the aerial remote sensing market that uses high-spatial resolution scanners with smaller swath widths. Several of these systems have panchromatic sensors with less than 1-meter resolution. The second group concentrates primarily on resource management, which requires much wider swaths and a greater spectral range in return for lower resolutions (10–20 meters).

C.2.1 Satellite Imagery Providers

More than 20 remote sensing satellite systems, with widely varying resolutions, are scheduled to be in operation by the year 2003. These include commercial systems that are designed for niche markets, with many emphasizing high-spatial resolutions and near-real-time receiving capabilities. Imagery data is also available from U.S. civil and military organizations. These include data from the National Oceanic and Atmospheric Administration (NOAA), the United States Geological Survey (USGS), the National Imagery and Mapping Agency (NIMA), and the National Reconnaissance Office (NRO).

C.2.1.1 Commercial Satellites. By the year 2003, as many as five U.S. satellites and two foreign satellites could have 1-meter imagery capability. Satellites with 2.5-, 5-, and 10-meter resolution capabilities will also exist. Satellite imagery providers and their specific capabilities are discussed in this section.

C.2.1.1.1 Space Imaging. Space Imaging launched its IKONOS satellite in September 1999. Space Imaging also has the U.S. marketing rights to a constellation of remote sensing satellites already on orbit. These include the Landsat 4 and 5 satellites, the IRS satellites, Canada's RADARSAT, and the European Space Agency's Radar Satellite. Space Imaging also sells sub-meter resolution aerial imagery for applications requiring a high degree of detail and accuracy. IKONOS and IRS-1C capabilities are shown in Table C-1.

Table C-1. Capabilities of Space Imaging Satellites

SATELLITE CAPABILITIES	IKONOS	IRS 1C/1D
Sensor Type	Panchromatic and multispectral imagery	Panchromatic and multispectral imagery
Resolution	.82-meter panchromatic 3.2-meter multispectral	5.8-meter panchromatic 23-meter multispectral
Coverage	Nationwide	Nationwide
Availability/Design Life	January 2000/7 years	August 1998/5 years
Accuracy—Panchromatic	Horizontal accuracy 4 meters Vertical accuracy is TBD Ground control is TBD	Horizontal accuracy 25 meters Vertical accuracy is TBD Ground control is TBD
Correction Options	Geometric and orthorectified	Geometric
Scene Size	13 km x 13 km panchromatic and MSI	70 km x 70 km panchromatic 140 km x 140 km multispectral
Image Cost	$33 per square kilometer each for panchromatic and multispectral	$2,500 per panchromatic scene $2,500 per multispectral scene
Number Scenes/Cost CONUS	57,000/$316 million	1,966/$4.9 million; multispectral imagery, 492/$1.3 million
Number Scenes/Cost Urban	5,700/$31.6 million	197/$490,000; multispectral imagery, 50/$130,000
Number Scenes/Cost Rural	51,300/$284.4 million	1,769/$4.41 million; multispectral imagery, 442/$1.17 million
Number Scenes/Cost Alaska	8,740/$48 million	302/$755,000; multispectral imagery, 76/$190,000
Company Archives Available	Since January 2000 on WWW	Since August 1995 on WWW
Restricted Use—One Buyer	Only buyer has access to data	Only buyer has access to data

C.2.1.1.2 ORBIMAGE. ORBIMAGE launched its first satellite in April 1995. This satellite, Orbview-1, provides dedicated weather-related imagery and meteorological products to the National Aeronautics and Space Administration (NASA). The company's second satellite, Orbview-2, was launched in August 1997 and provides images of land and ocean surfaces to commercial customers at a resolution of 1 kilometer. Orbview-3 and 4 are scheduled to launch in 2001. The features of Orbview 3 and 4 are listed in Table C-2.

ORBIMAGE is completing the design of the high-resolution Orbview satellites that will provide a 1-meter high resolution panchromatic (black and white) capability, a 4-meter multispectral (color and infrared) capability and, in the case of Orbview-4, an 8-meter hyperspectral capability. Orbview-3 will have a revisit time at nadir of about 23 days. By adding Orbview-4, this revisit rate will be reduced to nearly 12 days. Orbview-3 and Orbview-4 are identical except for the addition of the hyperspectral collection onboard Orbview-4.

Table C-2. Capabilities of ORBIMAGE Satellites

SATELLITE CAPABILITY	ORBVIEW-3	ORBVIEW-4
Sensor Type	Panchromatic and multispectral imagery	Panchromatic and multispectral imagery
Resolution	1-meter panchromatic 4-meter multispectral	1-meter panchromatic 4-meter multispectral
Coverage	Nationwide	Nationwide
Availability/Design Life	2001/5 years	2001/5 years
Accuracy	Horizontal accuracy 4 meters Vertical accuracy is TBD Ground control is TBD	Horizontal accuracy 4 meters Vertical accuracy is TBD Ground control is TBD
Correction Options	Geometric and orthorectified	Geometric and orthorectified
Scene Size	8 km x 8 km panchromatic and multispectral	8 km x 8 km panchromatic and multispectral
Image Cost	$33 per square kilometer each for panchromatic and multispectral imagery	$33 per square kilometer each for panchromatic and multispectral imagery
Number Scenes/Cost CONUS	150,500/$316 million	150,500/$316 million
Number Scenes/Cost Urban	15,000/$31.6 million	15,000/$31.6 million
Number Scenes/Cost Rural	149,000/$284.4 million	149,000/$284.4 million
Number Scenes/Cost Alaska	231,000/$48 million	231,000/$48 million
Company Archives Available	After launch on WWW	After launch on WWW
Restricted Use—One Buyer	Only buyer has access to data	Only buyer has access to data

C.2.1.1.3 EarthWatch Incorporated. EarthWatch's QuickBird systems will feature high resolution, precise geolocational accuracy, large area collection, and variable imaging collection times. Locational information collected by QuickBird's star trackers and an onboard Global Positioning System (GPS) will enable EarthWatch and its customers to produce high quality and accurate large-scale maps without the need for ground control points. QuickBird-1 will be the only system in the new generation of high-resolution satellites with a non-sun-synchronous, 66-degree, medium-inclination orbit. This orbit provides the benefit of viewing Earth features at various times of day under different sun angles and weather patterns. In certain geographic areas where clouds routinely gather at the same time each day, QuickBird-1's early-, mid-, and late-day imaging will afford users the opportunity to view features normally obscured by cloud cover. Table C-3 lists the QuickBird capabilities.

Table C-3. Capabilities of EarthWatch Satellites

SATELLITE CAPABILITIES	QUICKBIRD-1	QUICKBIRD-2
Sensor Type	Panchromatic and multispectral imagery	Panchromatic and multispectral imagery
Resolution	.82-meter panchromatic 3.2-meter multispectral	.82-meter panchromatic 3.2-meter multispectral
Coverage	Nationwide	Nationwide
Availability/Design Life	Summer 2000/5 years	Summer 2001/5 years

SATELLITE CAPABILITIES	QUICKBIRD-1	QUICKBIRD-2
Accuracy	Horizontal accuracy 4 meters Vertical accuracy is 4 Meters Ground control is 10 meters DEM	Horizontal accuracy 4 meters Vertical accuracy is 4 meters Ground control is 10 meters DEM
Correction Options	Geometric and orthorectified	Geometric and orthorectified
Scene Size	22 km x 22 km panchromatic and multispectral	22 km x 22 km panchromatic and multispectral
Image Cost	$25 square kilometer	$25 square kilometer
Number Scenes/Cost CONUS	19,900/$241 million	19,900/$241 million
Number Scenes/Cost Urban	2,000/$24.1 million	2,000/$24.1 million
Number Scenes/Cost Rural	17,900/$216.9 million	17,900/$216.9 million
Number Scenes/Cost Alaska	3,053/$40 million	3,053/$40 million
Company Archives Available	After launch on WWW	After launch on WWW
Restricted Use—One Buyer	Only buyer has access to data	Only buyer has access to data

C.2.1.1.4 Resource21. Resource21is a remote sensing information company that plans to provide time-critical resource management and broad-area land surveillance products to users. Although the system parameters have been optimized for global production of agriculture needs, the benefit is a system concept that can contribute greatly to the information needs of the U.S. Census Bureau. The Resource21 system can supply a tip-off capability for monitoring new construction in fast-growing areas.

The planned space collection segment will consist of three 5-band, 10-meter multispectral satellites that can cover the entire North American region twice a week with nadir coverage. Derived information can be produced and delivered electronically. The Resource21 satellite system is planned to be fully operational in 2003. The first satellite launch is expected to occur in the fall of 2002, and the final launch is to be completed by mid-2003. A fourth satellite will be built as a backup.

The conceptual Resource21 system features pixel sizes of 10 meters visible near-infrared (VNIR) and 20 meters shortwave infrared (SWIR). This requirement is driven by the need to detect, locate, and characterize vegetation and other materials and determine their conditions at densities as low as 4 percent over areas as small as 1/40th acre. For the U.S. Census Bureau, this capability would make it possible to accurately and automatically detect and locate objects with a spatial extent on the order of 2 meters by 2 meters over large regions. For example, buildings under construction that do not precisely match their surrounding environment (e.g., roads, vegetation, sand, and trees) could be located within 5 meters. With the use of advanced computing technology, it is possible to extract even more information from small patches of geospatial information derived from high-quality multispectral data. For example, the spectral presence of construction equipment, disturbed soil, road paving materials, and metals could be used to request street addresses and names from other organizations.

The impact of better revisit times, spatial performance, and radiometric performance would be a quantum leap in large-area change detection and analysis capability with important benefits. Subtle changes developing over intervals as short as 1 day and as long as several decades could be accurately characterized. Broad areas could be searched automatically in the spectral domain for events and conditions of interest. Table C-4 shows the Resource21 capabilities.

Table C-4. Capabilities of Resource21 Satellites

SATELLITE CAPABILITIES	RESOURCE21
Sensor Type	Multispectral imagery
Resolution	10-meter VNIR 20-meter SWIR
Coverage	Nationwide weekly
Availability/Design Life	2003/5 years
Accuracy	Horizontal accuracy is 5 meters Vertical accuracy is TBD Ground control is TBD
Correction Options	Geometric
Scene Size	195 km x 250 km multispectral
Image Cost	$80,000 yearly subscription $800,000 for yearly coverage
Number Scenes/Cost CONUS	200/$4,000 = $800,000
Number Scenes/Cost Urban	20/$4,000 = $80,000
Number Scenes/Cost Rural	180/$4,000 = $720,000
Number Scenes/Cost Alaska	105/$4,000 = $420,000
Company Archives Available	Information only; no imagery
Restricted Use—One Buyer	Free-to-use information

C.2.1.1.5 SPOT Image Corporation. SPOT Image Corporation of Reston, Virginia, is the U. S. subsidiary of SPOT IMAGE, S.A., of Toulouse, France. The SPOT satellite Earth Observation System was designed by the French and developed with the participation of Sweden and Belgium. The system has been operational since 1986 with the launch of SPOT 1. Three satellites are operational currently: SPOT 1, SPOT 2, and SPOT4. SPOT 5 is scheduled for launch in late 2001.

SPOT 5 will offer the main capability of interest to the U.S. Census Bureau (i.e., maintaining the scene size and increasing resolution). This new generation satellite will provide an enhanced performance using the latest technologies, including a highly innovative acquisition process called Supermode. This new process will allow SPOT 5 to acquire images using a 2.5-meter sampling grid while maintaining SPOT's 60 km by 60 km swath. This capability allows for maintaining the scene size and increasing the resolution, which could prove useful in rural areas of the United States. Further evaluation of th capability and cost of 2.5 meter SPOT imagery is needed. Table C-5 lists the SPOT capabilities.

Table C-5. Capabilities of SPOT Image Satellites

SATELLITE CAPABILITIES	SPOT 1–4	SPOT 5
Sensor Type	Panchromatic and multispectral imagery	Panchromatic and multispectral imagery
Resolution	10-meter panchromatic 20-meter multispectral	5-meter pan resampled to 2.5 5-meter panchromatic 10-meter multispectral
Coverage	Nationwide	Nationwide
Availability/Design Life	1986/10 years	Late 2001/10 years
Accuracy	Horizontal accuracy is 5 meters Vertical accuracy is 15 meters Ground control is 3 ARC Second	Horizontal accuracy is 5 meters Vertical accuracy is TBD Ground control is TBD

SATELLITE CAPABILITIES	SPOT 1–4	SPOT 5
	DEM	
Correction Options	Geometric and orthorectified	Geometric and orthorectified
Scene Size	60 km x 60 km panchromatic and multispectral	60 km x 60 km panchromatic and multispectral
Image Cost	$1 per square km panchromatic; .50 per square km multispectral	TBD
Number Scenes/Cost CONUS	Panchromatic, 2,675/$9.63 million; multispectral, 2,675/$4.82 million	TBD
Number Scenes/Cost Urban	Panchromatic, 268/$.96 million, multispectral, 268/$.48 million	TBD
Number Scenes/Cost Rural	Panchromatic, 2,407/$8.67 million; multispectral, 2,407/$4.4 million	TBD
Number Scenes/Cost Alaska	Panchromatic, 411/$1.477 million; multispectral, 411/$.75 million	TBD
Company Archives Available	Currently on WWW	2001 on WWW
Restricted Use—One Buyer	Federal use permitted	Only buyer has access to data

C.2.1.2 Future Commercial Systems. In about the 2004 time frame, the commercial satellite companies will be launching their second-generation satellites. As the resolution improves, satellite imagery will be able to distinguish more at a higher accuracy.

C.2.1.2.1 Half-Meter Imagery Resolution. Higher resolution imagery can increase resolution to about 18 inches with a potential for horizontal accuracy to 2 meters. Each of the 1-meter satellite companies has requested permission from the Department of Commerce to launch a half-meter system in 2004. The advantage of higher resolution data is that even more details would be seen. This resolution could have some uses in large metropolitan areas where more details are needed to distinguish structures. The disadvantages of higher resolution imagery include increased processing time, cost, and potentially more detail than the U.S. Census Bureau needs.

C.2.1.2.2 Quarter-Meter Imagery Resolution. One of the satellite companies has requested permission from the U.S. Government to launch a quarter-meter resolution system. A great deal of controversy has arisen over this much detail being collected and provided to the public. At this resolution, the commercial satellites are nearing the capabilities of the military satellites. A satellite with this capability may not be allowed to be built. This much detail would probably far exceed Bureau needs.

C.2.1.3 Federal (Military and Civil) Imagery Providers. The U.S. Government has launched many satellites over the years that could support MAF/TIGER modernization; however, in most cases these satellites are not available for various reasons. For example, the military satellites are focused on national defense and can be used only when all other sources have been exhausted. The NOAA/USGS/NASA Landsat 7 satellite data are available rather cheaply but have a low resolution. Information on these capabilities is provided below.

C.2.1.3.1 The Intelligence Community. The intelligence community provides access to classified remote sensing data through the USGS National Civil Applications Program (NCAP). National military satellite systems data are available to civil agencies under the following strict guidelines:

- The satellite systems are not to be used for monitoring citizens.
- The satellite systems are not to be used by law enforcement officials.
- The satellite systems are not to be used in any manner that could be perceived as intelligence gathering.
- The civil community needs are second to intelligence needs.
- Tasking must be on a not-to-interfere basis.
- Requesting classified imagery must be considered only as a last resort. Commercial collection opportunities must be exhausted prior to this type of tasking.
- Data is available only at the Secret level.

Information derived from images collected by military satellites is called an imagery-derived product (IDP). Examples include photographs, maps, line drawings, and statistical data. All U.S. Federal Government agencies may request and use classified imagery to derive information only when use of commercial sources is infeasible. IDP may be used to revise TIGER data even though the image is classified. The image itself may not be used.

C.2.1.3.2 NOAA/USGS/NASA Landsat 7. On April 15, 1999, the latest member of the Landsat family was launched into orbit. No other remote sensing system, public or private, can compete with Landsat in global change research or in civil and commercial applications. Landsat 7 imagery will be very useful in determining changes that occur during the time the MAF/TIGER modernization is taking place. For example, Landsat 7 imagery collection in 1999 and again in 2003 of the United States and its territories could be purchased at the beginning of the TIGER modernization project. From this data, major changes in construction could be located and emphasis placed on ensuring that the higher resolution imagery includes these areas. These change scenes could reduce the amount of high-resolution images needed. Table C-6 shows Landsat 7's capabilities.

Table C-6. Capabilities of Landsat 7 Satellites

Satellite Capabilities	Landsat 7
Sensor Type	Panchromatic and multispectral imagery
Resolution	15-meter panchromatic 30-meter multispectral
Coverage	Nationwide
Availability/Design Life	1999/5 years
Accuracy	Horizontal accuracy is 250 meters Vertical accuracy is N/A Ground control is N/A
Correction Options	Geometric
Scene Size	185 km x 175 km
Image Cost Per Scene	$695
Number Scenes/Cost CONUS	298/$207,000
Number Scenes/Cost Urban	30/$20,700
Number Scenes/Cost Rural	268/$186,300
Number Scenes/Cost Alaska	46/$31,970
Company Archives Available	Currently on WWW
Restricted Use—One Buyer	None

C.2.1.3.3 Other Civil Agencies. Because the USGS and the United States Department of Agriculture (USDA) have been using imagery for many years, they have acquired large amounts of imagery. The Bureau of Land Management (BLM), Federal Emergency Management Agency (FEMA), Department of Transportation (DOT), and Environmental Protection Agency (EPA) all have the need for imagery.

Despite overlapping needs for national-level data, civil agencies are unable to provide much, if any, data because of the differences in their missions and imagery needs. The USGS has areas within the United States that have accurate digital orthophoto quarter quads (DOQQ) that exceed the requirements of the Bureau. However, the USGS has data on only about two-thirds of the United States and of this two-thirds, a third is older than 4 years. USGS wants to assist the Bureau in its MAF/TIGER Modernization but will not have the resources to collect imagery required by the Bureau in a timely manner. The USDA has imagery over portions of the United States at 10- to 30-meter resolutions. It has been using Landsats 4, 5, and 7, as well as SPOT imagery, for its imagery needs. The other civil agencies have requirements for imagery in support of their tasks and are beginning to determine how to obtain it. For example, the FEMA Mitigation Directorate has identified imagery requirements with horizontal postings of 5 meters, with 2 1/2-meter horizontal accuracy at the 95 percent confidence level and a vertical accuracy of 30 centimeters at the 95 percent confidence level. Often, the needs of the civil agencies are not commensurate with each other.

C.2.2 Potential Satellite Imagery Partners

The National Digital Orthophoto Program (NDOP) has become the focal point for imagery needs within Federal and state governments. Through this organization, imagery can be purchased that supports many other organizations and often at an inexpensive rate. But once again, it must be understood that these purchases do not satisfy the needs of all users. A Bureau purchase could be orchestrated through the NDOP for maximum distribution of the data.

State governments have inquired about partnerships with the U.S. Census Bureau in imagery purchases. The major drawback to these partnerships is the time frame of the availability of financing for the modernization projects. The states are looking for partnerships soon, and the Bureau monies for imagery may not be available for several years. A concise review of the states' imagery data and the age of that data is provided in Table C-9. Very little state imagery data comes from a satellite; nearly all is from aerial imagery.

C.2.3 Airborne Imagery Providers

The number of airborne imaging systems in operation is substantially greater than the number of satellite systems. Compared to the satellite sensors, airborne systems achieve higher spatial resolution, some in submeter resolution. Airborne collection parameters are more flexible than satellite collection parameters. In addition, variable collection altitudes permit imaging under cloud cover. The airborne platforms are also responsive to changing priorities and are easily relocatable, making custom collection requirements feasible.

Although airborne collection at lower altitudes improves spatial resolution, it limits the size of the area covered. Larger areas require multiple flight lines for coverage. Mosaicking of multiple flight lines is usually required to obtain coverage of a large area of interest. Since illumination may vary considerably during the collection mission, the spectral quality of a mosaic may vary from flight line to flight line due to time intervals between flight lines. These variances can be compensated for through atmospheric corrections. A comparison between aerial and satellite imagery is provided in Table C-7.

Table C-7. Comparison of Airborne and Satellite Collection

CHARACTERISTICS	AIRBORNE	SATELLITE
Altitude	Low (up to 30,000 meters)	High (above 175 km)
Swath Width	Narrow (many strips for large area)	Wide (single strip for large area)
Resolution	Varied to meet needs	Fixed by satellite orbit
Illumination	Variable over large areas	Constant over large areas
Revisit	Frequent/flexible	Fixed by orbit
Responsiveness	Relocatable	Fixed by orbit
Flight Path	Variable—user defined	Fixed by orbit

The number of aerial companies located throughout the United States is too large to list. These numerous aerial collectors provide imagery to Value-Added Resellers (VAR) that process imagery data and add value, such as map building and geographic information system (GIS) overlaying.

Aerial imagery has advantages over satellite imagery as outlined in Table C-7. To better compare aerial imagery with satellite imagery for Census needs, Table C-8 is provided.

Table C-8. Capabilities of Aerial Imagery

CAPABILITIES	AERIAL IMAGERY
Sensor Type	Panchromatic and multispectral imagery
Resolution	6-inch panchromatic 6-inch multispectral
Coverage	Regional
Availability/Design Life	Continuous
Accuracy	Horizontal accuracy to inches Vertical accuracy to inches Ground control is DEMS, Lidar, etc.
Correction Options	Geometric and orthorectified
Scene Size	Varies for panchromatic and multispectral
Image Cost	$15 per square kilometer to $40 per square kilometer
Number Scenes/Cost CONUS	Variable—Dependent on Vendor and System
Number Scenes/Cost Urban	Variable—Dependent on Vendor and System
Number Scenes/Cost Rural	Variable—Dependent on Vendor and System
Number Scenes/Cost Alaska	Variable—Dependent on Vendor and System
Company Archives Available	Currently on WWW
Restricted Uses	Some have restrictions

C.2.3.1 MAPPS. An organization that provides both collection and value-added services is the Management Association for Private Photogrammetric Surveyors (MAPPS). The member firms are engaged in surveying, photogrammetry, satellite and airborne remote sensing, aerial photography, hydrography, aerial and satellite image processing, GPS and GIS data collection and conversion. The intention of the association is to complement the technical programs of the American Congress on Surveying and Mapping (ACSM), the American Society for Photogrammetry and Remote Sensing (ASPRS), Urban and Regional Information Systems Association (URISA), the Geospatial Information and Technology Association (GITA), or other organizations, and to be a forum and voice for the business aspects of mapping and GIS.

C.2.3.2 NDOP. The NDOP is a Federal organization that assists in obtaining imagery of users' areas of concern. The NDOP attempts to involve all levels of government in the Federal purchases of imagery. In the past, most purchases have been arranged through the USGS DOQQ program, but now more Federal and state agencies require imagery. The Federal agencies have been banking on the notion that the cost for DOQQ maintenance will be less than the current USGS fixed price for DOQQs. Declining Federal budgets might still be able to maintain a significant level of DOQQ production with lower costs. If the paradigm for orthophotos moves to high-resolution data, DOQQ maintenance costs will rise extensively.

Should the MAF/TIGER Modernization Project receive funding, and should it use imagery to realign TIGER®, this imagery could be available to state and local governments to support their imagery/GIS programs.

C.2.4 Potential Airborne Imagery Partners

State and local governments are potential airborne imagery partners.

C.2.4.1 State and Local Governments. State and local governments have begun to use imagery in conjunction with their GIS data. Most often, these governments find it difficult to afford the cost of the imagery or they do not have the expertise to support the use of imagery data. Usually, only the larger cities have become successful with integrating imagery and GIS. Examples of the needs and uses of two cities and two states are provided below.

C.2.4.1.1 Atlanta, Georgia. Atlanta believes that a continuous, on-line supply of accurate geographic information is of major importance to a modern society. The cost of creating and maintaining accurate and current digital data remains prohibitive to many organizations. Developing these databases by numerous institutions in a fragmented, wasteful, redundant, and noninteroperable manner is no longer acceptable economically and socially. This is why the Metro Atlanta Region has organized the Open Geodata Consortium. This group facilitates the regional on-line spatial data cataloging, data sharing between members, and spatial/geographic indexing for e-commerce portals by diverse industries and government agencies. This program saves millions of taxpayer dollars through integration of digital maps produced and maintained. It serves the regional community by serving an open consumer marketplace and by offering easy access to geographic information critical for government decision making, public safety, regional planning, transportation, utilities, real estate, property tax assessment, and many more basic aspects of everyday life. The imagery that supports this region is 6-inch resolution. All imagery and GIS information is available for purchase from the consortium.

C.2.4.1.2 San Diego, California. San Diego City and County jointly initiated the Regional Urban Information System project in response to the increasing complexity of delivering efficient and effective municipal services to the residents of the region. The goals of the project are to improve productivity; reduce costs; provide access to accurate, timely information for decision making; and improve service to citizens. More than 200 layers of GIS information are available to county and city agencies through a distributed network. Revenue is generated from the sale of imagery and geographic products that cover all of San Diego County.

C.2.4.1.3 New York State. The GIS Clearinghouse was established to disseminate information about New York's statewide GIS coordination program and to provide access to the New York State GIS metadata repository. The metadata repository was created to provide one central location where state agencies and local governments can list GIS data sets. The clearinghouse also provides information on links to GIS education and training and other state and Federal GIS resources that include data from USGS and the U.S. Census Bureau. Select imagery data is available in this clearinghouse.

C.2.4.1.4 Alaska. The State of Alaska was imaged once by NASA between 1978 and 1984. This coverage is at 1:60,000 scale, with color infrared, and 1:100,000-scale panchromatic imagery. Through the Alaska Geographic Data Committee, which is made up of 48 organizations (10 of these are native groups), a proposal is being developed that would provide 5-meter imagery of the entire state and 1-meter coverage of all urban areas and villages. The project will be completed over a 4-year period between 2002 and 2006. The committee has requested and received support from the U.S. Census Bureau's Geography Division for this collection. This data, if collected, could be used to realign MAF/TIGER data in Alaska.

C.2.4.1.5 State Coverage. Table C-9 lists state imagery coverage under the National Aerial Photography Program (NAPP) from the oldest to the newest collections and provides the size of the state in square miles. The type of collection (black and white or color infrared) is also provided.

Table C-9. NAPP Collections

STATE	SQUARE MILES	MOST RECENT COVERAGE	B&W VS. COLOR IR
Alaska	738,633	1978	MSI
Vermont	9,614	1992	CIR
Colorado	104,091	1993	B&W
California	158,706	1994	B&W
Florida	58,664	1994	CIR
Iowa	56,275	1994	B&W
Nebraska	77,355	1994	B&W
Nevada	110,561	1994	B&W
New York	49,108	1994	CIR
Oregon	97,073	1994	B&W
Virginia	40,767	1994	CIR
Wyoming	97,809	1994	B&W
Connecticut	5,019	1995	B&W
Massachusetts	8,284	1995	B&W
Missouri	69,697	1995	B&W
New Jersey	7,788	1995	CIR
North Dakota	70,703	1995	B&W
Oklahoma	69,956	1995	B&W
Rhode Island	1,213	1995	B&W
Kansas	82,278	1996	B&W
Maine	33,266	1996	B&W
Minnesota	84,402	1996	B&W
Mississippi	47,689	1996	B&W
Montana	147,045	1996	B&W
South Dakota	77,116	1996	B&W
Texas	266,806	1996	CIR
Washington	68,138	1996	B&W
West Virginia	24,231	1996	CIR

Table C-9. NAPP Collections (Cont'd)

STATE	SQUARE MILES	MOST RECENT COVERAGE	B&W VS. COLOR IR
Alabama	51,705	1997	B&W
Arizona	114,000	1997	B&W
Kentucky	40,409	1997	B&W
New Hampshire	9,278	1997	B&W
New Mexico	121,592	1997	CIR
Tennessee	42,144	1997	B&W
Utah	84,900	1997	B&W
Delaware	2,045	1998	B&W
Idaho	83,564	1998	B&W
Illinois	56,345	1998	B&W
Indiana	36,186	1998	B&W
Louisiana	47,752	1998	CIR
Maryland	10,461	1998	B&W
Pennsylvania	45,308	1998	B&W
Wisconsin	56,153	1998	B&W
Georgia	58,909	1999	CIR
Michigan	58,527	1999	CIR
North Carolina	52,669	1999	CIR
Ohio	41,330	1999	B&W
South Carolina	31,113	1999	CIR
Arkansas	53,156	2000	CIR
Hawaii	6,395	2000	CIR
Total	**3,766,227**		

C.2.5 Organizational Participants

Numerous organizations would directly benefit from the modernization of MAF and TIGER®. Some of these groups have been informed about this project and have provided input into the requirements listing in Appendix A. None of these groups are involved in the collection or production of imagery. They have memberships that are interested in the final outcome of the modernization. Names of organizations that have been notified of this project are provided below.

The Federal Geographic Data Committee (FGDC)
Urban and Regional Information Systems Association (URISA)
American Congress on Surveying and Mapping (ACSM)
The American Society for Photogrammetry and Remote Sensing (ASPRS)
Open GIS Consortium (OGC)
National Civil Applications Program (NCAP)
National Association of Counties (NACo)
National State Geographic Information Council (NSGIC)
The Association of American Geographers (AAG)
The Management Association for Private Photogrammetric Surveyors (MAPPS)
The Geospatial Information & Technology Association (GITA)

C.2.6 Imagery Distribution Restrictions

In Tables C-1 through C-8 the restrictions placed on the imagery is shown to be limited to the initial buyer, in most cases. Currently all of the satellite imaging companies and a large percentage of the aerial imaging companies provide their lowest costs based on the imagery being restricted to the actual buyer. If the imagery is to be distributed to additional users, the cost of the imagery goes up. In only a few cases will the imagery data become the property of the buyer. Usually it remains the property of the imaging company that collected the data and for an additional fee they distribute its use. The U.S. Census Bureau has several options to handle these restrictions.

- Negotiate the fees for greater distribution. The higher the distribution levels the higher the costs.
- Negotiate purchasing imagery from providers allowing the providers to sell the imagery for a period of time (i.e., 2 years). At the end of the time period, the imagery rights for restricted distribution would be lifted, and the imagery would be distributed freely by both the buyer and the imagery company. During the restricted time period, the buyer could use the data but not distribute it.
- Some aerial imaging companies allow the buyer to own the rights to the imagery data. When issuing requests for proposal (RFP), the Bureau should ensure that the RFP includes language supporting public domain rights for the imagery. If some companies can allow for public domain rights, others may be forced to relax their restrictions.
- Off-the-shelf imagery is cheaper and can have fewer restrictions on it. A purchase of this type of data would allow for the development of the MAF/TIGER imagery baseline. The imagery update restrictions to the baseline would require negotiation. This approach would limit the amount of imagery data under restrictions.

The imagery providers have stated that all the rights to imagery are negotiable. The restrictions on imagery can be worked out to ensure all MAF/TIGER imagery support data is available to the public.

C.3 Imagery Capabilities

Many variables should be considered when deciding which of the various types of imagery and data to use for the MAF/TIGER modernization and which vendors or sources to use as resources. For instance, there are many different types of imagery, all with advantages and disadvantages. The amount and accuracy of visible detail varies from type to type and from vendor to vendor. The currency of the imagery and the areas covered by the imagery also are factors in making decisions about sources. This section outlines the various attributes and capabilities of imagery.

C.3.1 Sensor Types

The two general classes of remote imaging sensors are active and passive. Most satellite and aerial imagery is collected from passive sensors, which include those that do not provide their own illumination but detect and record reflected sunlight or emitted light energy from features, objects, and materials. Most passive sensors are not capable of imaging through clouds or other atmospheric conditions that significantly degrade the transmission of light energy.

Active sensors emit, detect, and record electromagnetic energy reflected from the surface of a feature or object. These sensors can complement the passive sensors by providing information on terrain features that are not accessible by sensors that are dependent on the sun's energy. An example of a commonly used active sensor is radar. Many frequencies used by radar imaging systems can penetrate cloud cover, providing some measure of all-weather remote sensing. Three passive and two active sensors are described below.

- **Panchromatic Imagery—**Panchromatic imagery is collected by a digital sensor that measures energy reflectance in one wide portion of the electromagnetic spectrum or band. Panchromatic data is represented as black and white imagery. This type of collection constitutes a majority of collection for large urban areas. Panchromatic imagery can be collected at a few inches on aerial platforms and .82 meters on satellite platforms.
- **Color Imagery—**Color imagery offers three types of products—natural, false color (infrared), and multispectral. The natural color, also known as true color, is composed of three colors of the visible spectrum (blue, green, and red). When properly processed, the color rendition closely approximates the original scene as viewed by the human eye. If manual feature alignment of extraction were required, the natural color would be an aid.
- **Multispectral Imagery—**Multispectral imagery is collected by a digital sensor that measures reflectance in many bands. One set of detectors may measure reflected visible red energy, for example, while another set measures near-infrared energy. Two separate detector arrays may even measure energy in two different parts of the same wavelength. These multiple reflectance values are combined to create color images. Current multispectral remote sensing satellites measure reflectance in three to seven different bands simultaneously. Multispectral imagery can be collected at less than 1 meter using aerial platforms and at about 4 meters using satellite platforms. Multispectral imagery will be invaluable in determining changes over time for the MAF/TIGER Modernization.
- **Radar Imagery—**Radar sensors are active imaging systems, transmitting a radar signal in the microwave portion of the spectrum and measuring the strength and other characteristics of the return signal after it reflects off the Earth's surface. Radar imagery conveys feature information that differs in some ways from spatial and spectral data. Because radar is active and operates in longer wavelengths, it can acquire images through clouds, fog, haze, and darkness. This collection system is available on aerial platforms at 1-meter resolution and will be at 3-meter resolution on satellites in 2004. Its use for MAF/TIGER Modernization may be in locating roads and structures in areas with constant cloud cover, such as Hawaii.
- **Lidar Imagery—**Lidar (Light Detection and Ranging), another active imaging system, works like any ordinary radar except that it sends out narrow pulses or beams of light rather than broad radio waves. A receiver system times, counts, and processes the returning light. Lidar inertial reference systems and global positioning satellite system data can be combined to produce highly accurate airborne laser mapping. Because the Lidar is an active sensor system, it offers advantages and unique capabilities, such as the

ability to penetrate a forest canopy to map the floor. For MAF/TIGER Modernization, Lidar's greatest value could be in determining heights of ground features and buildings.

C.3.2 Resolutions

The quality of information extracted from remotely sensed images is strongly influenced by the spatial and spectral resolution of the sensor. As mentioned previously, the spatial resolution refers to the size of the smallest object or ground feature that can be distinguished in an image. Determining the size of the object that needs to be seen and then finding the imagery that has resolution sufficient to identify and locate the object has a bearing on project cost because the more detailed an image is, the more expensive it is.

The most common descriptive terms for spatial resolution are ground resolution or ground sample distance (GSD). An image with a GSD of 30 meters, for example, will not usually allow for detection of an object of less than 30 meters long. Detection depends on the data, feature, and background. Subpixel detection is common. It may be possible, however, to detect linear features such as roads, railroads, power line rights-of-way, and bridges.

C.3.2.1 Resolution of Less than 1 Meter. Imagery with a resolution of less than .82 meters is currently available only from aerial platforms and classified military satellites. The capabilities of aerial systems can provide a resolution of inches. This very high resolution coincides with a small coverage area and a very large digital file size. Imagery at submeter resolution may exceed the needs of MAF/TIGER Modernization. When contemplating this type of coverage, the cost of the coverage must be considered. The one place that submeter imagery could prove advantageous is in large metropolitan cities. More detail could provide more accurate information in these areas. An example of submeter resolution imagery is provided in Figure C-2. Note the details of the structures, automobiles, and tennis courts.

C.3.2.2 Resolution of 1 Meter. Imagery with a resolution of 1 meter is readily available from aerial and satellite providers. One-meter imagery can help identify and map features larger than 1 square meter, such as highway lanes, bridges, bus shelters, and fence lines. It can help in identifying structures and new construction and in differentiating between types of buildings and homes. This level of information would greatly assist the extraction of data in urban areas for the MAF/TIGER Modernization. Figure C-3 represents a typical 1-meter image.

C.3.2.3 Resolution of 2.5 Meters. Imagery with a resolution of 2.5 meters will be available only from the SPOT 5 satellite when launched in 2001. This resolution comes about through a process known as Supermode, where two 5-meter images are offset by half a pixel. The intermediate values are interpolated, and during the restoration step, an image is automatically generated with a sampling grid of 2.5 meters. It is not yet known if these processes will prohibitively increase the cost of the 2.5-meter data. The imagery at this resolution would be very effective in coverage for rural areas. Figure C-4 is an example of 2.5-meter resolution imagery.

Figure C-2. Submeter Resolution Imagery

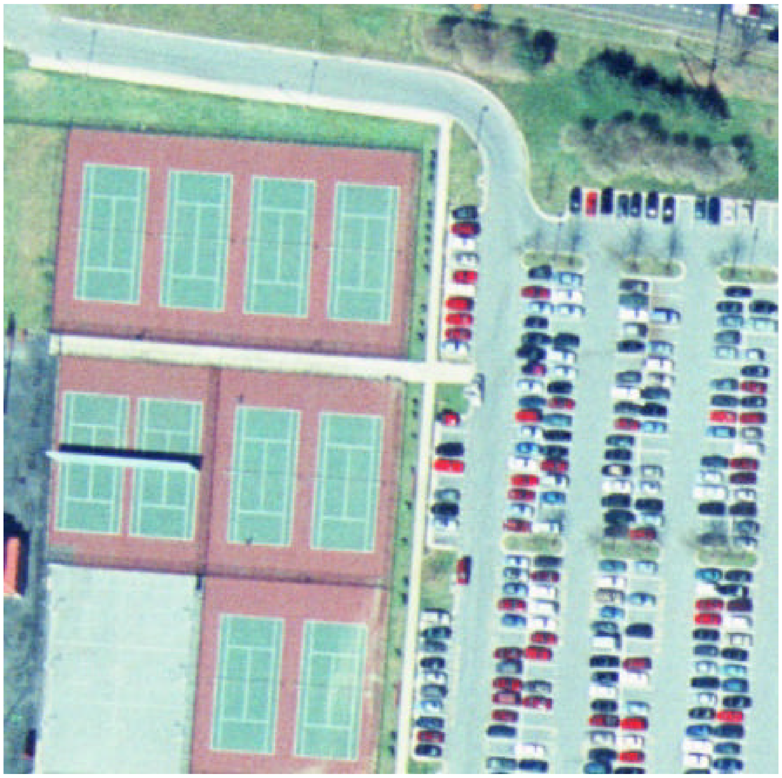


Figure C-3. 1-Meter Imagery

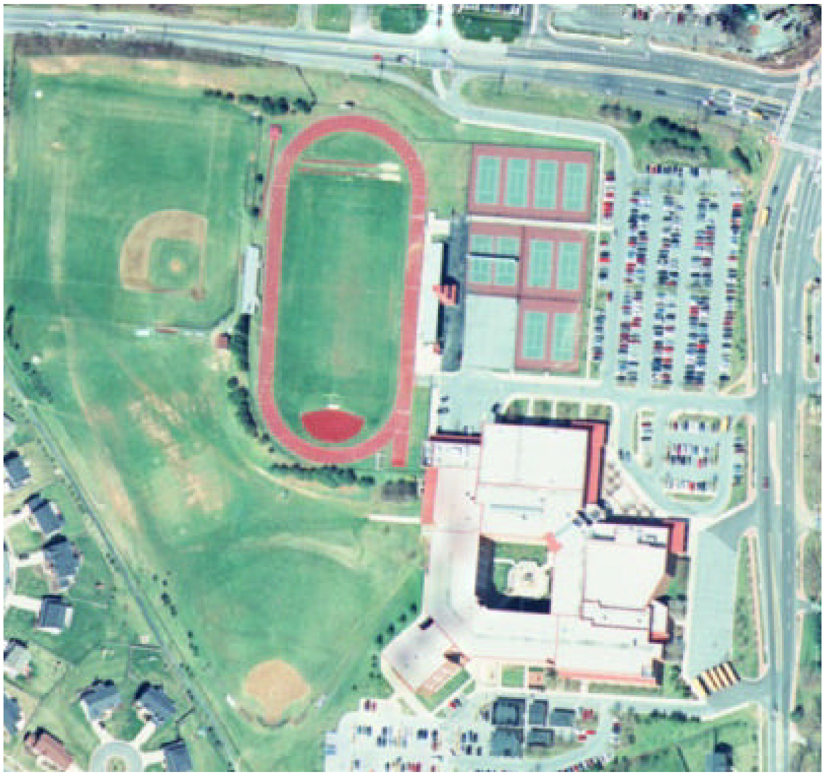


Figure C-4. 2.5-Meter Imagery

C.3.2.4 Resolution of 5 Meters. Five-meter resolution imagery has been provided by the Indian Remote Sensing (IRS) satellite for many years and will be available from the SPOT satellite next year. This resolution covers large portions of the Earth's surface at a low price, but at a less sharp resolution. This type of imagery could be efficient for coverage of rural areas where structures are not located close together. Accuracy would diminish at this level, but it would be satisfactory for structure matching and alignment with TIGER. An example of 5-meter imagery is shown in Figure C-5.

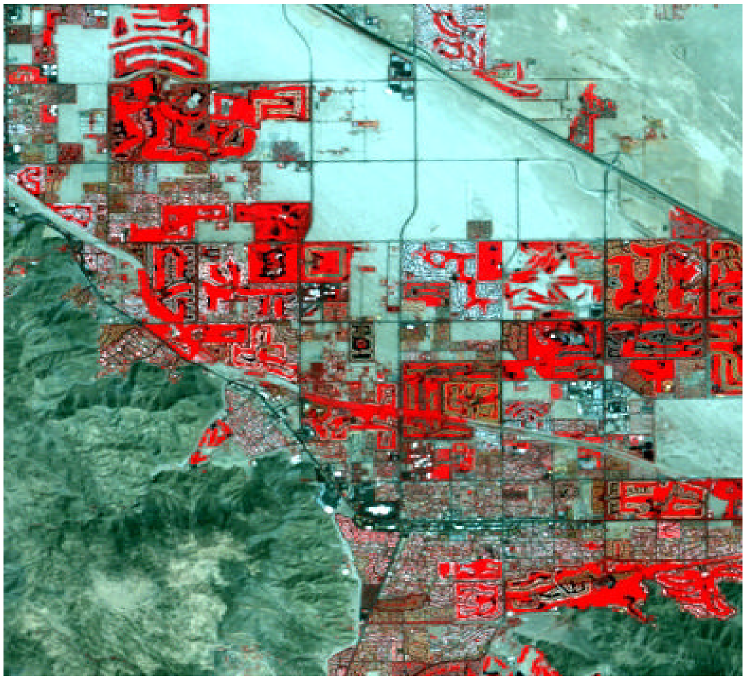
C.3.2.5 Resolution of 10 Meters. Resource21, a U.S. system that is to be launched in 2003, and Rapid Eye, a European system, also to be launched in 2003, can provide this capability. Its primary use is change detection. A 10-meter example is shown in Figure C-6.

Figure C-5. 5-Meter Imagery

Figure C-6. 10-Meter Imagery

C.3.2.6 Resolution of 30 Meters. Thirty-meter data is available only with Landsat 7. Use of this imagery could assist in targeting changes in urban sprawl between 1999 and 2003 and would focus the use of high-resolution collections in the initial stages of the modernization. The imagery would provide coarse coverage but would detect major changes in and around cities. The 30-meter resolution Landsat 7 imagery provides minimal value for MAF/TIGER Modernization beyond initial change detection. Figure C-7 is an example of 30-meter resolution imagery.

Figure C-7. 30-Meter Imagery



C.3.3 Coverage

The major reason for imagery coverage is to increase the accuracy of the TIGER data by aligning it with the imagery. Because the TIGER data includes all 50 states and the U.S. territories, imagery would be needed for all of these areas for proper alignment. The cost of collecting and processing imagery data and the alignment of this data with TIGER data must be considered when determining the coverage needed. To reduce cost, imagery could be gathered for the densely populated areas at 1 meter and the rural areas at 5 meters. The cost of collecting imagery at 5-meter resolution is considerably less than collecting imagery at 1-meter resolution.

C.3.3.1 The United States and Territories at 1-Meter Resolution. The cost of 1-meter imagery collection and processing for all 3.8 million square miles of the United States would exceed $275 million and require more than 3 years to complete. However, new processing technologies are being developed that would provide imagery data of the entire nation at 1 meter for less than $30 million and could be completed in less than 2 years. Further investigation of the technological breakthrough that would provide this capability at this low price would be required. For no additional cost, the data could also be provided to any other user.

C.3.3.2 Rural Areas at 2.5- or 5-Meter Resolution. The SPOT satellite offers 2.5- and 5-meter resolution and IRS offers 5-meter resolution imagery. The 2.5-meter SPOT data is experimental and the cost is not yet known. The IRS 1C data for the rural areas is estimated to cost about $8 million. The horizontal accuracy of the 5-meter imagery is approximately 20 meters, but if it were aligned to TIGER®, it would allow for the alignment of the structures in rural areas to be located on the correct side of the road. With the 5-meter data, the homes and barns should be distinguishable. Mobile home parks would be easily seen, but individual mobile homes in the countryside may not be as easily distinguished.

C.3.3.3 Change Detection. Ten-meter resolution imagery would be very efficient in detecting change. This resolution offers a large amount of coverage for a lower price. Change detection is an automated process that any image processing system can perform. The process requires two images of the same geographic area acquired at different times. After the images are rectified, or matched, the system compares the values of corresponding pixels in the two images and determines which values are different, indicating some change in ground features over the intervening time. Resource21 and Rapid Eye can provide this capability fully automated. Construction can be detected and requests sent to other Bureau organizations for the street name and address. To determine large-scale changes within the United States, a Resource21-type capability is the least expensive and most accurate system. Change detection with 30-meter resolution is too coarse and would miss smaller development projects. Change detection with 4-meter resolution would cost millions of dollars and would require years to complete.

C.3.3.4 Seamless Coverage. If 1-meter data were purchased and processed from one or more vendors, seamless imagery would be part of the purchase price. If the 1-meter and 5-meter plan were instituted, a seamless processing would be implemented to reduce the abrupt falloff between resolutions. One company has suggested making the United States seamless at 5 meters and then, in the areas where 1-meter data exists, boring down for more detail. If some form of seamless coverage is not implemented, difficulties with alignment of MAF/TIGER data could result in the areas where the 1- and 5-meter data meet.

C.3.4 Currency (or Age of the Imagery)

The older the imagery, the lower the price; but the older the imagery, the less accurate the changes. If the objective is to collect data as a baseline and update the baseline over time (i.e., years), older imagery is acceptable. If the objective is to immediately produce an accurate baseline as the updates to MAF/TIGER occur, older imagery is unacceptable. Areas that are expanding with construction need to be monitored closely and this expansion noted. Initially, the baseline could accept 3-year-old imagery and, allowing for a 4-year development period, would provide a baseline that was no older than 7 years when implemented.

C.3.5 Accuracy of the Imagery

Positional accuracy is defined by International Standards Organization/Technical Committee (ISO/TC) 211 as the accuracy of the position of features. The absolute positional accuracy of a feature is relative to the feature's position on the Earth's surface. Ground truth, while usually not cited for positional accuracy, can be considered to be the absolute position on the Earth's surface. Positional accuracy is often spoken of in terms of allowable deviation from ground truth. Resolution and accuracy are two characteristics of imagery that taken together greatly contribute to the positional accuracy of an image. Resolution is the size of the smallest object that can be distinguished on an image, while accuracy usually refers to the image's positional accuracy. Usually, accuracy is expressed in terms of pixels, which means that a 10-meter resolution image may have a 1-pixel accuracy, and an object in the image may be displaced by as much as 10 meters in any direction. For most images, raw data will have some degree of geometric and radiometric correction, which simply means that distortions caused by the sensor itself have been removed. The accuracy needed to make MAF/TIGER more accurate is a minimum of 8 meters. Several types of corrections are available to achieve better accuracy.

C.3.5.1 Geometrically Corrected Data. Geometrically corrected data is needed to correct variables in the image acquisition, such as airborne pitch, roll, and yaw; satellite altitude; satellite velocity; and rotation and curvature of the Earth. These variables are known as systematic because they are predictable distortions and can be corrected in software models. Other distortions, such as terrain displacement, are scene-dependent and require additional information to provide higher geometric accuracy. Geometric correction includes system- and precision-corrected data. The systematic level of correction has the least geometric accuracy. The image data is not tied to a physical ground reference point but is based on the best information as to sensor location. Systematic correction is subject to large errors as local elevation moves away from mean sea level and as local terrain relief is introduced. For precision corrections, all the systematic distortions are removed when ground control points are added to tie the image to the Earth. Precision correction is subject to large errors as local elevation variations increase.

C.3.5.2 Orthorectified Data. Orthorectified data is corrected using both systematic and precision corrections, with the addition of an elevation model to correct terrain displacement. This level of correction offers the highest level of accuracy and remains accurate regardless of elevation or terrain relief. It is also the most difficult to achieve because of the limited availability of adequate resolution elevation data. The highest orthorectified accuracy is about 4 meters with satellites and less than 4 meters with aerial imagery. More research needs to be done but it appears that orthorectified imagery will be required for the accuracy needs of MAF/TIGER.

C.3.5.3 Nadir Coverage. Nadir coverage from imagery sensors can have either a fixed or adjustable viewing geometry. If the sensor operates with a fixed or rigid viewing angle, it looks directly down (nadir) from the satellite and acquires images of only what is directly below the ground track. An adjustable sensor swivels from side to side or back and forth while acquiring images inside and outside of the ground track. The farther off nadir the image is taken, the more problems with shadows and masking that must be taken into consideration. Because the object of MAF/TIGER Modernization is to account for all structures, this masking is not acceptable. To date, an image that is off-nadir up to 12 degrees seems to be acceptable. Further review will be necessary.

C.3.5.4 Ground Control Capabilities. Ground control refers to known accurate points on the Earth. These points are expressed in horizontal (latitude and longitude) and in vertical (elevation above mean sea level) measurements. These points can come from GPS data, surveyor’s data, or Digital Elevation Model (DEM) data. The DEM data is digital information that provides a uniform matrix of terrain elevation values. It provides basic quantitative data deriving terrain elevation, slope, and surface roughness information.

C.6 Imagery Feasibility and Availability

Imagery has advantages and disadvantages. Imagery of the entire Nation is less expensive in comparison with alternative ways of aligning TIGER data, such as national GPS. Aerial imagery is readily available for most of the national coverage needs.

C.6.1 Imagery Advantages

Both aerial and satellite imagery can help with the MAF/TIGER Modernization. The aerial imagery can provide cheaper and faster high-resolution coverage over larger portions of the Nation. The satellite imagery can assist in change detection and can be used to target imagery collection needs from the change detection tip-offs. Imagery has a cost and timeliness advantage over GPS for aligning MAF/TIGER. With respect to landowners, it is also a nonintrusive collection process.

C.6.2 Imagery Limitations

Imagery requires an extensive time to collect, process, and align to MAF/TIGER. Cloud cover hampers passive sensors, haze, variations in illumination, and darkness. Large growth areas that hide structures in trees also limit imagery collection. Data collectors need to be seasonally timed to mitigate these limitations. In addition, certain sensors have better resolution and accuracy than others do. Aerial imagery has fewer restrictions dealing with clouds, fog, and haze. Airplanes can fly under the clouds, and the fog and haze in the image can be removed. Figure C-8 shows an image that was imaged from 1,700 feet in a fog bank, using a digital camera. Figure C-9 shows how effectively the processing technique removed the fog and haze.

Figure C-8. Submeter Image Collected in Fog Bank

Figure C-9. Effective Processing Technique Removed Fog and Haze

C.6.3 Imagery Archives

The total amount of imagery that is required to cover the 3.8 million square miles of the United States will require many terabytes of storage to archive. Storage options include holding and maintaining imagery data at a U.S. Census Bureau site or at a commercial company. Several companies have proposed the archiving of imagery data as part of their support to the project. A single archive company with backup locations, or archives by each imagery provider should be considered. To ensure seamless imagery, one archiving company would be a better choice. These options must be analyzed further.

C.7 Collection Management

To collect the baseline and maintain the accuracy of the data, a form of collection management should be implemented. Collection management requires a software tool that would enable users to make informed decisions in assessing the availability of existing imagery, predicting when new imagery can be collected, supporting the ordering of imagery, and performing imagery collection studies and assessments. All of this would be done via the World Wide Web. This software would be capable of accounting for a dozen satellites or aircraft and monitoring up to 3,000 scenes per day.

This tool must have user-definable options that make it easy enough for the casual user to check for an image and complex enough for the experienced user to access detailed information. This system would need the capacity to query existing imagery holdings to determine whether the user's requirements can be met by existing imagery or whether new imagery must be collected.

If a system must be tasked to meet user-specified requirements, the software needs to automatically generate access opportunities. The software would be used for collection planning analysis, available data coverage display, and satellite and aerial modeling and visualization. Area definition would be made through the use of annotations or a simple mensuration capability, such as entry of coordinates or the use of a computer mouse. The software must accurately model the orbits and sensor characteristics of multiple constellations for commercial imaging satellites and aircraft.

C.8 Summary

Imagery is available from half a dozen satellite companies and the many aerial imaging companies within the United States This imagery can be collected in many formats and processed to any level of accuracy needed. All of the information gathered on these imagery collection firms indicates that cost is the main differentiator among the firms. At the 1-meter level, both aerial and satellite imagery can provide the needed data, and these methods cost about the same for imagery collection. Satellite companies would require more than 3 years to collect all of the needed data because of the scene sizes and the cloud cover over the image sites. Aerial companies could collect it sooner because they have access to more aircraft, and they can fly under the clouds. The processing of the data is where the cost divergence arises. New technology provides the potential to reduce the time frame for processing, thereby reducing the cost.

This chapter has provided detail on known satellite companies that can provide the imagery levels required to support the U.S. Census Bureau task. The aerial collection companies are too numerous to have been outlined here. Aerial imagery is already available and can be obtained at a cheaper price than new collections. New collections can be completed in a more timely manner using aerial methods than by using satellites. The aerial collection cost would be about the same as satellite, but the processing and seamless correlation would be less. Additional research must occur to determine optimum choices.

Note: *The discussion of specific vendors and products in this document is for the sole purpose of offering representative examples of products available for specific technology needs. This information is not intended as a promotion of any specific vendor or product. Additionally, the vendors and products discussed herein should not be considered exclusive from other potential vendors and products in the event of future procurement requirements.*

APPENDIX D: CURRENT SYSTEM ENGINEERING BEST PRACTICES

For any system engineering process, the basic goals are to develop quality systems that support the mission of the organization, fulfill customer and user requirements, are responsive to technology improvements, and are cost-effective. To meet these goals, organizations must apply sound engineering principles, appropriate life-cycle development methods, and continuous improvement disciplines to the software development process. Current system engineering best practices focus not only on these traditional principles, methodologies, and disciplines but also on an emerging trend toward use of rapid systems development approaches and systems integration practices to create system environments that are adaptable to future technological developments.

This document will address best practices in traditional, efficient system engineering, along with emerging best practices that focus on faster development and on integrating new development with existing system environments.

D.1 Key Issues for the U.S. Census Bureau's MAF/TIGER Modernization Study

Best practices may vary, depending on an organization's environment and key issues. For the U.S. Census Bureau's MAF/TIGER Modernization initiative key issues are—

- **Providing user functionality—**To ensure that any system engineering effort will produce results that enhance user functionality, it is critical that the requirements for user functionality be well defined and addressed as a high priority. The best practices, methods, and products selected should support this objective.
- **Lowering software development costs—**With any new development effort, it is important to select practices that improve efficiency in development, resulting in lowered development costs.
- **Increasing development speed to meet schedule constraints—**In most cases, applying efficient system engineering practices will result in a product's being produced on time. However, because of the risks of unforeseen delays associated with software development, it may be necessary to employ schedule-oriented practices to ensure that time constraints are met.
- **Managing risks—**Risks are an ever-present factor in developing new systems to support business processes. It is important to develop a risk management plan that will identify, manage, and mitigate the risks associated with the Bureau's MAF/TIGER Modernization effort.

Integrating new development into legacy systems—The integration of new technology into an existing system environment can be the greatest challenge of a system engineering effort. Selecting practices that will coordinate the integration process is critical to the success of the Bureau's MAF/TIGER Modernization effort.

The practices, methods, and products that will best suit the U.S. Census Bureau will depend on the results of this initial assessment of key issues.

D.2 Current System Engineering Best Practices

Best practices in system engineering fall into five general areas: software engineering methodologies, software life-cycle development, continuous process improvement, risk management, and rapid application development (RAD).

D.2.1 Software Engineering Methodologies

Selection of an appropriate software engineering method is essential to the success of any software development process. The two most widely recognized engineering methods are information engineering (IE) and the object-oriented (OO) method. The method selected will affect the tools and procedures used throughout the system development life cycle.

- **Information Engineering—**Information engineering is a data-driven architectural approach to planning, analyzing, designing, and implementing software applications. Information requirements are thoroughly analyzed before the beginning of the design and build stages of application development. The IE method provides a framework to ensure that applications will work together within an enterprise, rather than building stand-alone systems to solve problems in isolation.
- **Object Oriented—**Object-oriented methods approach development from the perspective of objects rather than procedures or data. Each object contains its own set of procedures and data. OO delivers major benefits in the area of software reusability; but OO programming is difficult, and this remains an expert's technology.

D.2.2 Software Life-Cycle Development Methods

A software life-cycle development method is a model that describes and establishes order in all activities involved in the development of a software product. To ensure that all effort is leveraged to produce efficient results, it is essential to select an appropriate life-cycle development model at the beginning of any application development project. An effective life-cycle model can increase development speed, improve quality, enhance project tracking and control, minimize overhead and risk exposure, and improve client relations. In contrast, an ineffective life-cycle model can decrease productivity and increase development time and costs. To select the most effective life-cycle model, it is important to understand system requirements at the outset and to determine whether they are likely to change. It is also important to understand the system architecture and the likelihood of its changing. Other factors to consider are the model's reliability, the amount of planning and design effort required, project risk, schedule constraints, progress visibility to customers and management, and the level of sophistication that managers or developers must have to use the model successfully. Well-established life-cycle development methods are the waterfall model, the spiral model, and prototyping. The advantages and limitations of each of these methods are described below.

- **Waterfall Model—**The waterfall life-cycle model comprises an orderly sequence of steps for software development, from the initial software concept through system testing. It provides an organized, disciplined process with high design integrity. This model allows early detection of errors during the low-cost stages of the project. One drawback is that the waterfall process takes a relatively long time (needs may change, opportunity costs are high). In addition, it is difficult to get user participation and to determine user needs without a prototype. The waterfall model is also often criticized because a step must be 100 percent completed before the next step can begin. This model works best for projects with well-defined requirements, such as the release of a new version of an existing product.

- **Spiral Model—**The risk-driven spiral approach to software development identifies and studies risks in a series of learning cycles by using iterations and evolutions. The spiral model has four major components: planning, to determine objectives and scope; risk analysis, to identify and mitigate risks; engineering, for development of the product through the testing phase; and customer evaluation, to assess products of the engineering element. This approach avoids early overspecification, produces a limited-functionality system early, and focuses on major areas of uncertainty. It provides ample opportunity for developers and users to learn about the product and to reduce uncertainty and risk. The learning cycles also give management the opportunity to stop the project in a planned manner if the technology is insufficient. However, the spiral model approach requires tight discipline to avoid "coding and fixing." The spiral model is best suited for internal development within a developer's company.
- **Prototyping—**The prototyping life-cycle model allows for development of the system concept as a project progresses. The most visible aspects of the system are usually developed first and demonstrated to the customer, with feedback from the customer incorporated as development of the prototype continues. The primary disadvantage of this model is the unknown number of iterations before customer acceptance of the product, which makes it impossible to establish a firm schedule at the project's outset. This model is especially useful in projects with rapidly changing requirements or in cases in which developers are not confident about the system architecture.

D.2.3 Continuous Process Improvement Program

For more than a decade, credible software process improvement programs have been based on the Software Engineering Institute (SEI) Capability Maturity Model (CMM), a framework that describes the key elements of an effective software development process. CMM lays out an evolutionary improvement path that software organizations can follow to maximize their capabilities. It incorporates five levels of maturity, as shown in Table D-1, and defines practices for planning, engineering, and managing software development and maintenance. When implemented, these practices can improve the ability of an organization to meet cost, schedule, functionality, and product quality goals for software development projects.

Table D-1. Capability Maturity Model

LEVEL	CHARACTERISTICS	KEY PROCESS AREAS
5. Optimizing	Continuous process capability improvement	• Process change management Technology change management • Defect prevention
4. Managed	Quantitative measurement of process and qualitative management of product	• Software quality management • Quantitative process management

LEVEL	CHARACTERISTICS	KEY PROCESS AREAS
3. Defined	Software processes defined and institutionalized	• Peer review • Intergroup coordination • Software product engineering • Integrated software management • Training program • Organization process definition • Organization process focus
2. Repeatable	Management controls in place; stable planning and product baselines; still dependent on individuals for new products	• Software configuration management • Software quality assurance • Software subcontract management • Software project tracking and oversight • Software project planning • Requirements management
1. Initial	The software process is characterized as ad hoc or even chaotic. Success is very dependent on individual effort	

D.2.4 Risk Management

Implementing an active risk management strategy is a critical best practice for all system development projects. Management of risks related to scheduling is particularly important when a software product must be delivered within a tight time frame. Scheduling constraints may also increase the risks associated with product quality and cost. For instance, as a product advances through development, there may be a risk that product quality will be compromised to meet a deadline. One way of mitigating this risk is to involve the end users and customers throughout the development cycle. A comprehensive risk management strategy will involve identifying and prioritizing risks, implementing risk mitigation strategies, and monitoring schedule, product quality, and cost risks throughout the project.

D.2.5 Rapid Application Development

RAD is based on the concept that software can be produced faster and with higher quality through implementation of development strategies that focus on increasing efficiency and speed. Typically, this approach employs techniques to closely link the system developers with system users and customers during the development cycle. Increasingly, organizations are using the rapid development approach to produce high-quality software.

Although RAD comprises many of the same elements as traditional application development, some practices, such as joint application development (JAD) sessions and prototyping, have become more commonly associated with a rapid development approach. The core components of the RAD approach are—

- Early, thorough definition of project objectives
- Facilitated JAD sessions involving system users and customers, to identify system requirements
- Prototyping and early, reiterative user testing of designs

- Reuse of software components
- A rigidly paced schedule that defers design improvements to the next product version
- Full-time team assignments and full-time user participant(s).

Using a RAD approach for the U.S. Census Bureau's MAF/TIGER Modernization effort will improve the speed with which applications are produced and will provide multiuser capabilities, improve data processing and data exchange capabilities, and increase software reusability and quality assurance. Application of fundamental software development practices and use of a suitable system development tool support achievement of these goals.

D.2.5.1 Joint Application Development

JAD is a requirements-defining methodology in which systems requirements are established through facilitated sessions attended by customers, business owners, executives, and developers. This method is considered a best practice for use in the RAD approach because it gathers requirements faster and more efficiently than do traditional methods. In addition, because this method involves customers and business users in the planning and design phases, requirements are less likely to change later in the development process, minimizing risks to costs and schedule. The success of this method depends on effective leadership of the sessions, commitment of appropriate participants, and successful consensus-building within the group.

D.2.5.2 Reuse of Software Components

A common practice in RAD is to plan for the reuse of frequently used software components to facilitate faster assembly of new programs. Software reuse has been known to produce dramatic time savings and is a growing practice in application development. Establishing a long-term strategy for the selection and preparation of appropriate components increases the potential for success. A reuse strategy should be coordinated with other software-development fundamentals.

D.2.5.3 Schedule-Oriented Practices

Applying practices aimed specifically at achieving schedule goals is a core component of rapid application development. Schedule-oriented practices include those that improve development speed, mitigate schedule risk, and emphasize the visibility of progress. The types of schedule-oriented practices chosen will depend on the project's specific development needs. If the goal is to develop the application faster, the focus should be on practices that increase speed. If the goal is to improve the customer's perception of development speed, the focus should be on practices that increase the visibility of progress. All schedule-oriented practices are best applied after the initial establishment of fundamental development and system engineering practices.

D.2.5.4 Rapid Application Development Tools

The most common type of tool for supporting rapid development of software applications is the computer-aided software engineering (CASE) tool. CASE involves the use of a computer-assisted method to organize and control the development of software and is especially useful for large, complex projects involving several software components and users. The use of CASE tools in software development can be particularly helpful for improving design quality in a database-intensive environment. CASE tools allow designers and programmers to move easily from analysis, to design, to implementation by automating the creation of diagrams defined in software methodologies. A CASE tool commonly supports structured methods and object-oriented methods by providing a graphical user interface (GUI), notational capabilities, error and consistency checking, code generation, and support for the Internet environment. Frequently used CASE tools are Microsoft's Visual Studio, Sybase's PowerDesigner, and Oracle Designer.

Fourth-generation languages (4GL) can also be used to support high productivity for rapid application development. These languages support the development of throwaway user-interface prototypes, a common practice in RAD. The most widely used 4GLs for rapid development are Microsoft's Visual Basic, Sybase's PowerBuilder, and Borland's Delphi 5.

D.2.5.5 Representative Products to Support Rapid Application Development

Representative products that support high productivity for RAD are listed below. This product list provides examples of market share tools used for RAD; however, this list is not comprehensive, and these products should not be construed as the only products from which the Geography Division might draw in the event of a procurement.

- **Microsoft Visual Studio—**The Visual Studio development suite supports developer productivity and design with integrated features across all the popular programming languages. Visual Studio includes a set of development tools for building reusable applications by assembling components in Visual Basic, Visual C++, and Java. For large systems and distributed applications, Visual Studio offers enterprise database development and design tools, team development support, development life cycle, and development and test versions.
- **Oracle Designer—**Oracle Designer provides an integrated solution for application development that uses a RAD user interface. It provides support for all phases of the software development life cycle, from business modeling to deployment. Oracle Designer has a repository-driven approach that allows components to be used for rapid development of scalable, cross-platform, distributed applications.
- **Borland Delphi 5—**Delphi 5 is a 4GL that provides high-productivity development of Web and database applications in a Windows environment. It combines a visual development environment with compiler and database technology to provide an object-oriented, client/server development environment. It manages the entire development cycle with more than 150 reusable components, the Advanced Project Manager, and a 32-bit compiler to support faster delivery of applications.
- **Sybase PowerBuilder Enterprise 7.0—**PowerBuilder Enterprise 7.0 offers a family of rapid application and 4GL tools for client/server and Internet application development. For distributed application development, it provides integration support for execution in Sybase Enterprise Application Server and support for building component object modeling components for deployment to Microsoft Transaction Server.
- **Forte Software's Forte Application—**The Forte application environment is an integrated set of object-oriented tools for building, deploying, and managing high-end

client/server applications. Forte provides a complete set of object-oriented tools, including a GUI designer, an object-oriented 4GL, a comprehensive set of class libraries, and a repository to support team development. Forte automates the process of deploying an application in a specific client/server environment with the use of GUIs, operating systems, relational database management systems (RDBMS), and networks. In addition to generating applications that are ready for target deployment environments, Forte includes capabilities for distributed execution and a set of application management tools.

D.3 Use of Commercially Available Software

An alternative to custom development of applications is the use of commercially available software products. This software has the advantage of being immediately available to provide users with at least some valuable capabilities. The disadvantage of commercially available software is that it usually cannot satisfy every customer need. This problem is sometimes solved by the release of new versions that offer more functionality. Before selecting a commercial software product, an organization must examine how well each package will satisfy user requirements and to evaluate such issues as compatibility with current system architecture, costs, and training needs. In addition, organizations should note that commercial software is not recommended for applications in which significant changes are expected. If requirements are not clearly defined or are not expected to remain static, commercial software may not be a good alternative to RAD.

D.4 Integrating New Development with Legacy Systems

Over the past few years, organizations have attempted to increase productivity, improve efficiency, and reduce expenditures by implementing business-critical software applications. In many cases, these systems were built to serve one purpose and one set of users. Typically, these systems were custom built using the technology of the day, much of which was proprietary and may have used nonstandard data storage technology. This stove-piped approach to system development has resulted in uncoordinated systems architectures in which each system and its interfaces must be maintained manually, requiring excessive resources. This lack of coordination also precludes effective system modernization to address legacy systems migration.

D.5 Enterprise Application Integration (EAI)

EAI is an emerging approach to linking applications and information for unrestricted data sharing within a distributed computing environment. EAI provides the tools necessary to completely integrate an organization’s business processes, data, applications, and platforms by use of legacy, middleware, and commercially available solution sets. The benefits of using EAI are—

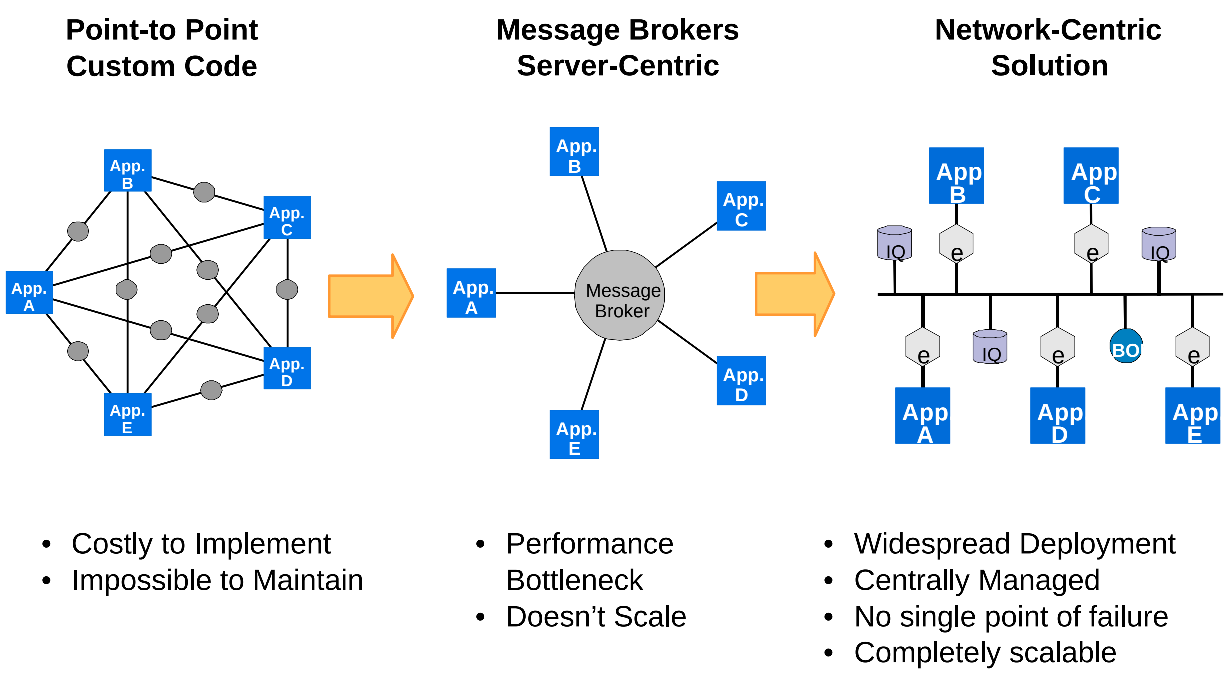
- **Flexibility for changing standards**—EAI provides the ability to seamlessly change interfaces when requirements of data trading partners require modification, without affecting existing interfaces and operations.
- **Flexibility for implementation—**EAI can be implemented methodically in a system-by-system process without the “big BANG!” implementation of the past.

- **Cost-benefit**—EAI reduces fixed costs and variability for legacy application maintenance work because of reduced network traffic on lower bandwidth.
- **Consistency of information**—EAI gives internal and external users a consistent view of an organization’s information.
- **Facilitates modernization efforts**—EAI greatly facilitates the modernization of an organization’s systems.

D.5.1 Middleware

In distributed computing, middleware is usually a means of connecting clients to servers without having to negotiate many operating systems, networks, and resource server layers. Middleware hides the complexities of the underlying operating system and network to facilitate the integration of various enterprise systems. Although there are a variety of middleware products offering various solutions, middleware generally falls into three categories: point-to-point custom code, message brokers server-centric various solutions, and network-centric solutions. Figure D.1 depicts the architectures for the various solutions and lists some key aspects of each.

Figure D.1. Architecture for Various Solutions



- **Point-to-Point**—Custom Code examples of point-to-point solutions are remote procedure calls (RPC) and message-oriented middleware (MOM). Both technologies provide standard application program interfaces (API) across hardware and operating system platforms and networks.
  - RPCs are a synchronous solution that halts an application’s processing while invoking a remote function procedure on the server. RPCs are typically useful only for high-speed networks.

  - MOM is asynchronous meaning that an application's processing is not suspended when the application involves the middleware API. The MOM solution is most useful for transaction-oriented applications that need to cross many platforms. The advantage of MOM is that it has the flexibility of a message-based system that can be adapted to most applications.
  - The disadvantage of point-to-point solutions like MOM and RPCs is that they integrate information only between single applications and therefore have limited utility for multipoints integration solutions. Point-to-point solutions are also costly to implement and difficult to maintain.

- **Message Brokers Server-Centric Solutions—**Message brokers provide a central integration point to facilitate the sharing of information between one or more integration points (such as network, middleware, applications, systems). Message brokers integrate multiple business applications, including new, old, legacy, centralized, or distributed applications, using a hub and spoke type of architecture. Although the message broker's primary role is to provide a central point of integration, it can also enhance business functions and translate data into usable forms for each application. Message brokers have the advantage of being able to integrate several systems, but are still considered an immature technology in which bottlenecks are a frequent problem and lack of scalability is an issue.
- **Network-Centric Solution—**Network-centric technology provides integration for sharing data in a distributed environment. Commonly known as distributed object technology, network-centric solutions make reusable program components available to other computers in a distributed network. Two models for distributed object technology are the Distributed Component Object Model (DCOM) and the Common Object Request Broker Architecture (CORBA). Both of these models allow sharing of data and application logic and provide a central clearinghouse for enterprise information.
  - DCOM is Microsoft's approach to the network-centric solution. It provides program interfaces in which client program objects can request services from server program objects on other computers in a network. DCOM can work on a network within an enterprise or on other networks outside the public Internet.
  - CORBA is a set of standards for creating, distributing, and managing distributed program objects over a network. In this model, an interface broker, known as the object request broker, allows programs at different locations and developed by different vendors to communicate in a network environment. CORBA is considered the standard architecture for distributed object technology. Currently, CORBA and Microsoft are developing a gateway approach to allow component object modeling and CORBA client objects to communicate with each other.
  - Network-centric solutions have the advantage of allowing sharing of data over multiple locations while providing centralized management to monitor transactions processed between multiple system platforms.

D.6 Summary

There are many challenges in selecting best practices for current system engineering efforts. First and foremost, best practices must be selected to meet an organization's unique needs.

Establishing the key issues for a system engineering effort will drive the selection and application of appropriate best practices. Second, identifying and applying efficient and effective engineering principles will significantly influence the success of all phases of system engineering, from requirements analysis through development and testing. In addition, applying new best practice trends and emerging technologies in system engineering will facilitate the integration of new technology into existing system environments. Finally, identifying and managing the risks involved in a system engineering effort are critical to ensuring that all the applied best practices will produce acceptable results.

Note: *The discussion of specific vendors and products in this document is for the sole purpose of offering representative examples of products available for specific technology needs. This information is not intended as a promotion of any specific vendor or product. Additionally, the vendors and products discussed herein should not be considered exclusive from other potential vendors and products in the event of future procurement requirements.*

APPENDIX E. COMMERCIAL GEOSPATIAL SOFTWARE

To perform its mission in the past two decades, the U.S. Census Bureau's Geography Division has written custom geospatial software to create and maintain its geographic database, produce mapping products, control and track production progress, extract data for various reports, and geocode housing units. No commercial software equivalents existed at the time the custom software was written. The Geography Division was breaking new ground in building TIGER® and its processing components. However, in the intervening time, and with much credit due to the Geography Division, a geospatial software industry sprang up to meet the needs of a fast-growing market for geographic data. Today, commercial geospatial software exists that could meet many of the functions needed by the Geography Division to complete its requirements.

To continue meeting its responsibilities, the Geography Division has maintained a significant investment in custom-written software. The costs of prolonging the life of these applications is becoming increasingly expensive. Today, computer hardware and storage devices are exponentially less costly. Commercial software with much of the necessary functionality can be purchased for less than the cost of writing and enhancing original code, and the maintenance and testing of the software is borne by the vendor.

There are many other good reasons to migrate to commercially available geospatial software. In addition to replacing current functionality performed by the Geography Division, there are additional functions, not currently used by the Geography Division, that could be used to their advantage. Most geospatial software vendors today have developed software that can be used over the Internet, which is very much in the plans of the division. Imagery and image processing software exists to provide accuracy never before envisioned as practical for TIGER®. Several imagery software vendors also have feature extraction applications that can semiautomatically recognize geographic features by processing the image file. A final, and important, reason to modernize is that a large proportion of state, local, and tribal governments and the general public use these readily available geospatial tools. Data exchange in both directions will be greatly improved.

Despite the wide variety of geospatial tools available today, it is unlikely that any one package or any suite of packages will fulfill all the geoprocessing needs of the Bureau. Whatever tools are selected, the tools will need to have easy-to-use application programming interfaces to facilitate custom coding, be interoperable with other tools, and work on a variety of operating systems.

E.1 Commercially Available GIS Software

Since the inception of readily available geospatial data, the Geographic Information System (GIS) industry has grown dramatically. Many vendors now provide GIS software and services. Of them, Environmental Systems Research Institute (ESRI) and Intergraph hold the largest share of the market. However, there are many smaller companies that provide high-quality and niche applications. This section compares the products of several representative geospatial processing companies. A summary of the companies' capabilities is contained in the summary section of this chapter.

E.1.1 Autodesk

Autodesk has a long, strong history as a computer-aided design (CAD) company. In recent years, however, it also has developed a suite of products for GIS and CAD/GIS data integration. Autodesk World Version 2.5 for Windows NT/95/98 is an open GIS/desktop mapping product. It is open not only in its ability to be customized and to interact with Microsoft Office products—by means of object linking and embedding (OLE) automation, application program interface (API), and integrated Visual Basic for Applications—but also in terms of data access. Because Autodesk is designed to work seamlessly with all AutoCAD data formats, common vector GIS formats from the ESRI, MapInfo, Intergraph, and others can be integrated without conversion. Other formats can be incorporated as a Dynamic Link Library (DLL) by third party software. Autodesk World uses Microsoft's Jet Engine from Access as its internal database and connects to external databases using Open Database Connectivity (ODBC), Data Access Objects, or other drivers. Version 2.01 includes an integrated Earth Resource Mapping ER Mapper image engine for improved image access and manipulation.

In addition to AutoCAD World, the AutoCAD Map 2000 product is designed to simplify the process of creating and maintaining GIS data and maps and merging data with existing CAD data sources throughout an enterprise.

Another application, Autodesk MapGuide, is a bundle of three packages designed for rolling out GIS data over the Internet. The first package is the free MapGuide Viewer (for NT/95/98, Macintosh, and Sun UNIX platforms), which is a Java plug-in or ActiveX control for viewing and interacting with spatial data using a Web browser. The second package, MapGuide Author (NT/95/98), is a program for creating and controlling the appearance and access restrictions for Web applications. The third package, MapGuide Server, an NT-based Web server product, brokers requests and delivers maps and data to users working with MapGuide Viewer.

Autodesk's recent acquisition of Vision Solutions should speed the development of scalable\enterprise solutions for large communications, utilities, and government organizations, specifically by integrating Autodesk's products with Vision's Oracle-based server technologies.

The approximate cost per license is $2,600 to $7,800, depending on configuration.

E.1.2 Bentley Systems

MicroStation GeoGraphics is a geoengineering package specifically engineered to facilitate integration with Bentley System's MicroStation 95 CAD software. GeoGraphics is offered as one of six configurations of MicroStation/J, Bentley System's new enterprise engineering modeling solution. The "J" stands for Java, which is used as the native development environment to create engineering or GIS applets termed "mapplets" using MicroStation/J's new Java superset called Java Macro Development Language (JMDL). MicroStation/J remains fully compatible with earlier MicroStation editions but adds the capability of writing Java applications that can run on any computer anywhere, regardless of operating system, within a company's intranet.

The approximate cost per license is $3,500 to $18,800, depending on configuration.

E.1.3 Environmental Systems Research Institute (ESRI)

For more than 20 years, ESRI has provided a wide range of GIS products for the geospatial business. Its widely used ArcView GIS product is designed to offer simple user interface and customization tools, Avenue and Dialog Designer. Although originally envisioned as a “mini GIS package,” ArcView has evolved to have a full set of functions that most GIS users need. There are many additional components that can add more functionality to ArcView. ESRI’s fully functional GIS package, Arc/Info, is in its eighth release. One of the most important changes in version 8 is the addition of a new, object-oriented data model with continued support for the earlier versions.

Arc/Info’s GeoData Objects are stored in Microsoft Access format, unless the installation has implemented a Spatial Data Engine (ArcSDE) extension, which allows users to store objects in their choice of supported relational database management systems (RDBMS) and provides better management of long database transactions. Although Arc/Info will continue to run on several operating systems, ESRI has optimized Arc/Info Version 8.0 to run on Windows NT. Windows NT users can access three easy-to-use toolbar-based applications built on Microsoft’s Component Object Model (COM) technology.

In addition, ESRI’s ArcMap allows users to edit, display, query, and analyze map data and design map layouts. ESRI’s ArcMap is significantly easier to use than the previous ArcEdit or ArcPlot tools. ArcCatalog provides for easy location, browsing, and management of spatial data. ArcToolBox facilitates processes such as automated and interactive data conversion, transformation, overlay analyses, and buffering.

The approximate cost per license is $3,500 to $20,000, depending on configuration.

E.1.4 Intergraph

Another company that has been around since the beginning of the GIS industry is Intergraph, headquartered in Huntsville, Alabama. In 1992, Intergraph moved its application development out of the proprietary UNIX environment and embraced Windows NT. Now the company is highly competitive with the release of its GeoMedia line of GIS products. These GeoMedia products integrate geospatial data from Intergraph, other GIS vendors, and from within Oracle’s Spatial Cartridge, providing a needed level of interoperability.

Intergraph’s Graphical Data Objects (GDO) technology was modeled after Microsoft’s Data Access Object and Remote Data Object methods of data access. GDO essentially is an OLE extension that provides data servers to access external data formats without conversion. There are no proprietary languages or data formats in these products. GeoMedia products also are completely customizable; menu items can be manipulated, custom commands defined, and user applications developed. In all, GeoMedia provides at least 17 ActiveX controls, 9 data servers, and more than 90 programmable automation objects.

Intergraph leverages the open architecture of GeoMedia by extending the Intergraph suite of geospatial software into NT-based products for the Internet. GeoMedia’s WebMap appeared about the same time as the desktop products. Most recently, GeoMedia Web Enterprise was offered as an open Internet server solution for spatial and network analysis. Intergraph also has a full range of image and image processing software.

The approximate cost per license is $3,500 to $12,500, depending on configuration.

E.1.5 Laser-Scan

Laser-Scan, a British Company, develops UNIX and Windows NT-based GIS products for such geospatial markets as defense, national mapping and charting agencies, and businesses. Laser-Scan offers a wide range of applications, all of which were developed using the object-oriented Gothic Applications Development Environment (ADE), now in version 3.2. Gothic stores its intelligence, self-validation methods, display characteristics, real-world behaviors, and rules in the data, not within the application. This database schema results in thin, easily customizable application layers that work on fat data layers.

Integrator, a toolkit layer that surrounds Gothic, is available in Windows and Java versions. Integrator Windows Edition allows the power of Gothic to be integrated with other desktop applications and allows users to build their own task-specific applications. Integrator Java Edition uses Java/Common Object Request Broker Architecture (CORBA) technologies to enable data access and analysis over the Internet or intranet.

Laser-Scan's database schema is especially optimized for Web applications to perform complex analysis and processing entirely on the (fat) server machine, leaving (thin) clients free to handle small Java applets.

The approximate cost per license is $3,000 to $15,500, depending on configuration.

E.1.6 MapInfo

MapInfo, headquartered in Troy, New York, also provides several geospatial software tools. Professional 5.5 is used by companies in almost every GIS industry and in public-sector organizations from local to national levels as a desktop mapping/GIS product. The product includes a built-in geocoder and arrives bundled with more than 400 megabytes (MB) of data.

The company places a strong emphasis on business geographics. MapInfo's newest release provides enhanced business graphics and charting capabilities and seamless integration of Oracle 8i and IBM DB2, in addition to Informix and Sybase connectivity. The company also offers ProPress, an optional raster-image processor that greatly speeds the process of producing high-quality hard-copy maps of any size.

A free viewer program, ProViewer, also is offered to facilitate sharing of maps and data throughout an organization. MapInfo's MapX package provides for customization. MapX is an ActiveX control that allows developers to embed mapping functionality into new or existing applications using Visual Basic, Visual C++, or other fourth-generation development languages. Also available is an enterprise version that integrates MapX with SpatialWare, MapInfo's server technology for managing spatial data within a database like Oracle or Informix. MapInfo also offers the MapXtreme mapping server to facilitate Web-based mapping. MapXtreme is available as a 100 percent Java server or as an ActiveX-based server for Windows NT. Both versions support full access to RDBMS servers.

The approximate cost per license is $3,500 to $18,800, depending on configuration.

E.1.7 Oracle Spatial

Oracle Spatial provides an integrated set of functions and procedures that supports the storage, access, and analysis of spatial data in a relational database. Oracle Spatial, combined with Oracle8i, enables spatial data to be stored, accessed, managed, and manipulated in the same manner as structured data. With Oracle8i, spatial data can be stored in relational tables or as objects—abstract data types (ADT). This new object data type is supported directly by the core Oracle8i database engine and provides the capability to store all locational (geographically referenced) information within a database server without having to resort to custom-built external indexes and functions to improve performance. In addition, users of spatial data have access to standard Oracle8i features, plus enhanced features such as increased database size limits and improved backup and recovery operations.

Oracle Spatial supports the management of spatial and attribute data in one physical database, thereby controlling processing overhead and reducing the complexity of coordinating and synchronizing disparate sets of data. Users define and manipulate spatial data through structured query language (SQL) and gain access to standard Oracle features such as an n-tier architecture, object capabilities, data management utilities, and Java stored procedures. In addition, a geocoding framework supports address matching, storage, and retrieval of geocoded spatial point data, as well as within-distance query capability, from within Oracle8i databases.

Oracle Spatial (Version 8.0 and 8i) has passed Open GIS conformance testing for “Simple Features for SQL.” Oracle is an active member of the Open GIS Consortium (OGC) and is also working with SQL3 standards to define SQL extensions for handling spatial data. Oracle is an active member of the International Organization for Standardization Technical Committee, Geographic Information and Geomatics (ISO/TC 211), and the American National Standards Institute (ANSI) Technical Committee X3L1 GIS standards groups. Oracle’s participation in these initiatives increases the likelihood that its product lines will adhere to the evolving standards and support interoperability with other commercially available GIS products.

E.1.8 SmallWorld

SmallWorld provides a scalable, object-oriented GIS development environment called SmallWorld GIS. SmallWorld is most commonly used in electrical and gas utilities, telecommunications, and public-sector service providers (water distribution and local governments) markets.

Although SmallWorld GIS has its own database, the company has invested effort in optimizing communications among heterogeneous databases, including Oracle, and across computer platforms regardless of communications protocol. SmallWorld GIS is especially good in applications that require long-transaction handling.

SmallWorld's architecture is designed to take advantage of client/server-based computing. It does this by smart caching in the SmallWorld client. For wide area networks, such as the U.S. Census Bureau network, each site can use a local cache to ensure that data is transferred from the master database only once, regardless of the number of users who need it. Alternatively, if communications links are unreliable, a replica of the master database can be used and synchronized periodically. The result is that a large number of users can be supported without difficulty on a single local area network.

The core SmallWorld technology can also be used to build application servers that support thin clients, such as field units and Web browsers. This capability extends the system's inherent scalability to provide a complete solution for deploying applications across the enterprise.

SmallWorld directly uses many common GIS formats and is integrated with Safe Software's Feature Manipulation Engine products for converting data. SmallWorld is built on an object-oriented development environment called SmallWorld Magik, which can be used alone or with Visual Basic, Visual C++, PowerBuilder, Java, and other object-oriented software to customize and enhance the product.

SmallWorld has released its new Spatial Intelligence product, designed to compete in the desktop mapping/GIS market. This new product has a Windows-type user-friendly interface but maintains consistency with SmallWorld's other products.

SmallWorld also offers a Web-mapping product called SmallWorld Web. SmallWorld Web makes it possible for customers to access SmallWorld data from Web browsers and eliminates the need for a memory-intensive spatial data application on every desktop.

The approximate cost per license is $4,000 to $20,000, depending on configuration.

E.2 Feature Alignment and Extraction Software

If imagery is used to increase the accuracy of TIGER®, specialized geospatial software will be required to perform feature extraction and alignment. The difficulties of extracting features from imagery are many and vary with quality and resolution of the imagery. However, several companies offer feature extraction and alignment capabilities. All of the alignment tools are experimental with high potential. The U.S. military has been trying, with limited success, to automatically align features to imagery for more than 20 years and has spent millions of dollars. The military requires much closer alignment of imagery and GIS data than the Geography Division requires.

No fully automated feature alignment tools are commercially available to align TIGER data to imagery. Those listed below have demonstrated the capability to provide some automation of this process. This list is not all-inclusive but represents those found during the past 3 months to have potential in this area. Further research should continue to determine the best capable feature alignment software tool.

E.2.1 Innovative Solutions Group

Innovative Solutions Group has developed semiautomated feature extraction software that includes applications for bodies of water, medium to large buildings and housing units, and roads. The software tool, FeatureFinder, is an integrated tool that allows the user to extract selected topographic and manmade features from orthorectified digital imagery to support automated map generation or revision by feature layer.

Automated extraction of water-related features from orthorectified digital imagery is the most advanced application developed to date by Innovative Solutions Group. The water-related applications include capabilities for automated extraction of shoreline features of rivers and coastlines, as well as the extraction of features of lakes, ponds, and tidal wetlands. The module for extracting features of buildings (small, medium, and large) is currently being updated. The module for extracting features of roads, railroads, and other manmade linear features is being developed and will be available in late 2000.

FeatureFinder performs an assessment of the loaded images and estimates the effectiveness of the extraction. This tool saves time for the expert imagery analysts (IA) and geospatial imagery (GI) analysts and supports less skilled IA and GI personnel in these tasks. FeatureFinder is a stand-alone product that works in conjunction with an installed GIS. FeatureFinder has an operator interface that allows the IA or GI to select the features to extract and generate a vector file that can be displayed and compared to existing data for map revisions, or used as a primary data source for generating new maps in a GIS. The GIS receives the extracted feature layer file from FeatureFinder and realigns previous vector data on the image to allow editing of the vector files.

E.2.2 ERDAS

ERDAS has developed a combination of tools in addition to its IMAGINE Professional suite, including Expert Classifier and Subpixel Classifier, to create a fully customizable and scalable feature realignment tool that could be used for the MAF/TIGER Modernization Initiative.

IMAGINE Professional Versions 8.3/8.4 contain rule-based sophisticated tools for remote sensing and complex analyses and are built around the IMAGINE Expert Classifier, which is integral to ERDAS IMAGINE 8.4, enabling users to graphically build expert or knowledge-based systems for image classification, post-classification refinement, and advanced spatial modeling. Expert systems take advantage of high-resolution imagery by incorporating rules for evaluating spatial relationships along with contextual information (proximity to other features, altitude, aspect, shape, etc.), doing much of the analysis for the user.

IMAGINE Subpixel Classifier is a supervised classifier designed to detect and report whole and subpixel occurrences of a specific Material of Interest (MOI) in multispectral imagery. IMAGINE Subpixel Classifier derives a signature for a specific material, groups all pixels in an image as either containing the material or not, and classifies all pixels that contain the material into classes based on how much of the material they contain. This component is essential in distinguishing roadways from sidewalks, rooftops of housing units, and other areas where feature realignment is required. It is a fully integrated, add-on application module to ERDAS IMAGINE.

E.2.3 Pixxures/Litigation Systems, Inc.

Using customized commercially available tools and a blend of proprietary software, Pixxures/Litigation Systems is developing a graph-based feature realignment system. This technique offers a method for realigning vector data by extracting mathematical graphs as "match points" that can be found in both raster and vector data. This technique is possible because of a patented technique for converting any mathematical graph into a unique numeric code. This code and attendant topological and coordinate information can be loaded into a database and used to match similar graphs taken from other raster or vector images.

E.2.4 Other Research and Development Efforts

Several companies and organizations, including Space Imaging, ENVI/Research Systems, Visual Intelligence Systems, and Rochester Institute of Technology, are working on semiautomated feature extraction/realignment tools. Material from these companies and others will be provided later this year or as the individual organizations develop a solution that could potentially be integrated into the MAF/TIGER Modernization Initiative.

E.3 Feature Alignment Process

The process of aligning TIGER data and imagery will be handled by feature alignment software that will be purchased from a commercial vendor. A specific approach will be determined by the contractor providing the software. In a potential process, vector data graphs will be captured in a straightforward manner by extracting intersecting vectors at their point of intersection, such as the crossing of two roads. Raster graphs will be created and extracted through a twofold process. First, items of interest will be automatically highlighted in the raster image using proven techniques such as pixel classification or edge detection. Second, the highlighted features will be converted into mathematical graphs. Each graph will be distilled into a unique numeric code that reflects the graph structure in terms of edges and vertices.

Once the graphs are identified in the raster data, they will be compiled in a database that contains a mathematical code for the graph and geometric coordinate information. A similar database will be built for the graphs extracted from the vector data. The graphs from the raster and vector data will be compared and matched. Graphs will be matched by graph numeric code and relative position. When a sufficient number of graphs have been matched, the images will be registered. After registration has been achieved, the detailed vector information will be compared to the detailed extracted features from the raster data.

The product of this comparison is the determination of which vectors can be corroborated by data from the raster image. In those cases where a match is made, the existing alignment will be confirmed or an automatic realignment performed. If a mismatch occurs or no corresponding data is found, visual inspection will be used to evaluate the condition and determine the nature of realignment required. The value of the technique is its ability to reduce significantly the amount of human labor involved in vector realignment.

E.4 Summary

Commercial geospatial software will be instrumental in the MAF/TIGER Modernization program development. The GIS software and feature alignment software will provide options for the increased capabilities of mapping, geocoding, maintenance, and TIGER data alignment. The majority of the required GIS software and feature alignment software will come from commercially available sources.

Table 5-1 presents a summary of the functions available in the software products discussed in this chapter. Research of the capabilities of these GIS software products will provide the best match to the Geography Division.

Table 5-1. Commercially Available GIS Software Matrix

VENDOR/SOFTWARE	AUTOMATED MAPPING	AUTOMATED GEOCODING	DATA EXTRACTION	DATABASE MAINTENANCE
AutoDesk World V. 2.5	X		X	X
Bentley System's MicroStation/J	X		X	X
ESRI ARC/Info V 8.0	X	X	X	X
ESRI ArcView V 3.2		X	X	
ESRI ArcSDE			X	X
Intergraph GeoMedia	X	X	X	X
Laser-Scan GIS	X		X	X
MapInfo ProV 5.5	X	X	X	X
Oracle Spatial	X		X	X
SmallWorld GIS V 3.0	X	X	X	X

Note: *The discussion of specific vendors and products in this document is for the sole purpose of offering representative examples of products available for specific technology needs. This information is not intended as a promotion of any specific vendor or product. Additionally, the vendors and products discussed herein should not be considered exclusive from other potential vendors and products in the event of future procurement requirements.*

APPENDIX F. COMMERCIAL DATABASE SOFTWARE

Among the many goals of the MAF/TIGER Enhancements initiative is updating the technology to support and enhance data processing and storage. A robust database management system will be crucial for storing and updating address list and geographic information.

F.1 Relevance to MAF/TIGER Modernization

The modernization of the MAF/TIGER system will have to fulfill key requirements to handle the volumes of information that are collected from the census and other sources. Because data collectors are scattered all over the country, distribution, portability, and mobility are crucial to the functionality of the system. Multiple users must have access to the system simultaneously, and updates must occur uniformly throughout the system. Because of the varied technical backgrounds of the users, ease of use is critical to ensuring the accuracy of data entry and extraction. Additionally, the database system must integrate with GIS and other geospatial processing software.

F.2 Key Issues

There are four issues regarding commercial database implementation: geospatial data management, database distribution, database optimization, and database backup strategy.

F.2.1 Geospatial Data Management

Of particular importance to the MAF/TIGER modernization effort is the ability of a commercial database package to manage geospatial data. Traditionally, relational database management systems (RDBMS) are not capable of storing and retrieving spatial data in any acceptable manner. Geospatial data usually need to be stored in a scheme where objects that are close to one another on the ground are close to one another in the database. Additionally, queries on geospatial data frequently take the form of finding objects that are within a given circular or rectangular area, finding objects that lie within a specified distance of another object, determining if a point falls inside a given polygon, etc. The relational database model does not lend itself well to solving such problems.

Most large-scale geospatial data installations have taken the approach of assigning identifiers to spatial objects, storing their attributes in a RDBMS, and custom developing a storage and retrieval system to manage the spatial objects. In recent years, RDBMS vendors have recognized the geospatial (and temporal) limitations of their products and have begun to produce extensions to address geospatial processing needs.

The current trend in the geospatial data industry appears to be for RDBMS vendors to provide just enough functionality to store and retrieve spatial objects in an efficient manner, and for GIS vendors to provide functional geographic processing modules. Both IBM and Oracle have released extensions to their database packages to manage geospatial data, and the major GIS vendors such as Environmental Systems Research Institute (ESRI), MapInfo, Intergraph are developing "middleware" products that connect to the spatial extensions of the database packages. The Open GIS Consortium (OGC), a consortium of database and GIS vendors, is developing industry standards for this approach.

F.2.2 Database Distribution

Fueled by the growth of the Internet, the popularity of distributed computing has affected even the most traditional mainframe based applications; databases are no exception. The concept of a distributed database has become a reality, because of its increased reliability and feasible cost of implementation.

Database distribution, or distributed database is defined as a single database, artificially partitioned into subsets of data, and physically placed on separate machines to achieve performance and/or scalability needs. Despite partitioning, the data continues to be viewed as a single-image, a single database schema.

Although distributed databases offer improved performance and scalability, the issue of data consistency is a significant concern. Because the database is distributed, different users can access it without interfering with one another. However, the database management system (DBMS) must periodically synchronize the scattered databases to make sure that they all have consistent data. Data synchronization is the process whereby a database defined by a single data model, which is distributed in physical implementation, is unified and brought into a synchronized state. The database may be synchronized for only a short period; however, the success of the application relies on a process of reconciliation at certain key intervals, therefore ensuring data consistency.

Some considerations for synchronization include—

- Transfer of net changes only
- Field-level synchronization
- Mappers to support field-level synchronization in a heterogeneous environment
- On-line vs. store-and-forward synchronization
- Scheduled and on-demand synchronization
- Managerial tools to establish and change users' data profiles.

Another form of distributed databases involves data replication. Replication is a copy management technology, not a process. Using database logs, a snapshot of data is taken at the point in time and copied to the target system(s). There is no comparison or reconciliation of changes.

F.2.3 Database Optimization

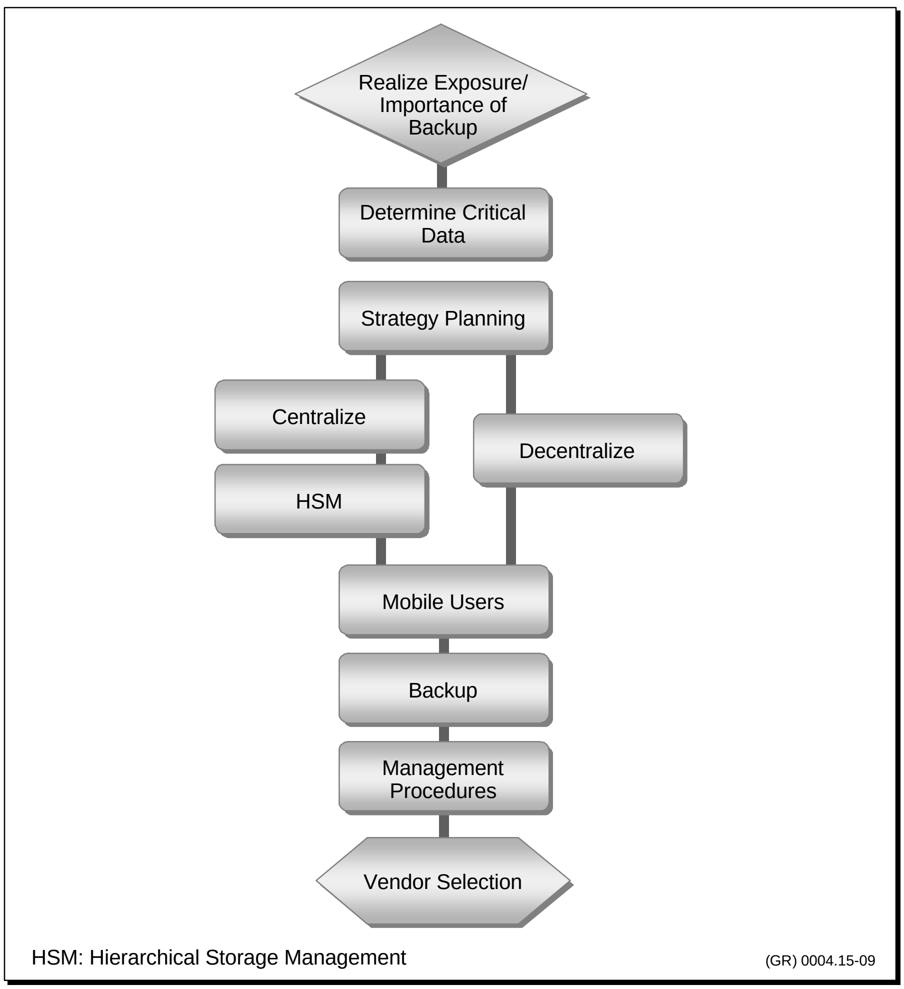
Systems need to be technically tuned to run databases efficiently. When a database application performs any function, it pulls various parts of tables, queries, or files directly into the system's memory. As more resources are loaded into memory, the speed of the database decreases exponentially. Generally, increased memory provides a greater performance boost than increased processor power, and upgrades should be considered in that order.

Proper user training can also improve database performance. When users are able to articulate the type of reports they want to generate, databases can be properly indexed to increase performance. However, without proper training, users are often forced to compile their own indexing schemes without knowledge of correct procedures, leading to a decrease in database performance.

F.2.4 Database Backup Strategy

As with any mission-critical application, data backup and recovery is an essential risk mitigation step. To determine the level of coverage required from a backup strategy, two factors must be considered: cost of the backup strategy and the value of the data. The cost of implementing a sound backup strategy with broad coverage must be weighed against the degree of risk and the value of the data. The total cost of the backup strategy includes more than just the cost of hardware and software; the administrative costs of designing the strategy and ensuring compatibility should also be considered.

Figure F-1. Backup Strategy Process Flow



Source: Gartner Group

Figure F-1 illustrates the backup strategy process flow. The first step is to determine what data is critical and where it resides. After critical data is pinpointed, a determination must be made whether to back up that data centrally or to adopt a decentralized strategy. Although decentralized backup can be more expensive than centralized backup, it is necessary in some situations.

F.2.4.1 Data Classification. An essential part of the backup planning strategy is data classification. Often, data cannot simply be assessed by monetary value alone. Instead of using a monetary value to estimate the data value, enterprises should measure the data on business impact. If the data is lost or corrupted, how will it affect business operations? The classification scheme should identify multiple levels of data criticality. Each level should have guidelines, procedures, and technology investments to the level justified by the risk of loss, the impact of the loss, and the cost of recovery.

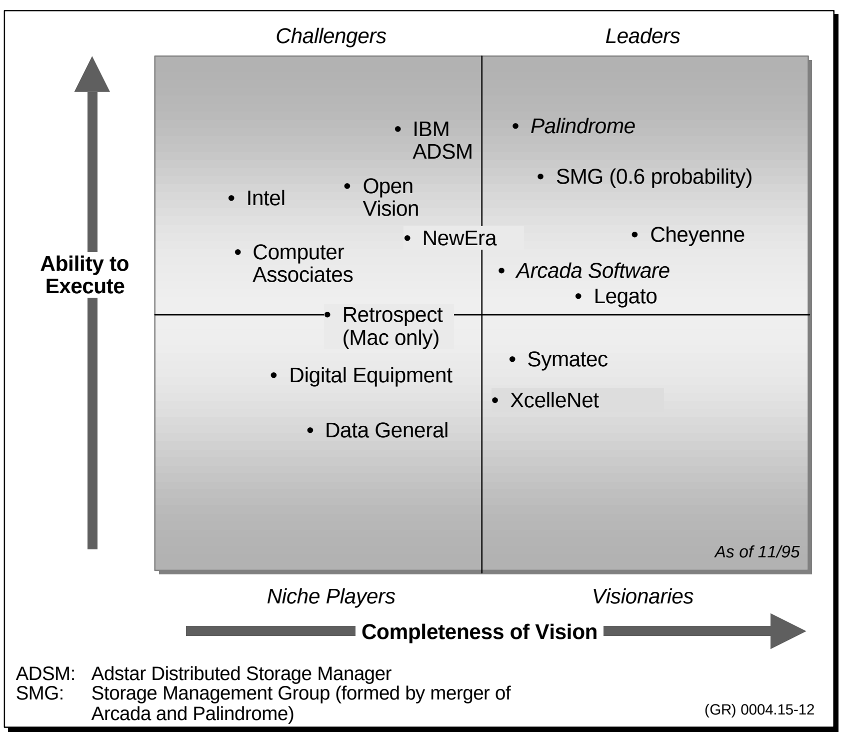
Data classification can be done in two ways:

- Through a rigorous audit program, where data is physically cataloged and backup is administered automatically

On an ad hoc basis, where the parameters and procedures for classifying data are published, and the owners of the data are solicited to execute appropriate data protection and backup programs.

Figure F-2 illustrates a matrix of major storage backup providers.

Figure F-2. Storage Backup Providers



Source: Gartner Group

F.3 Database Implementation Best Practices

The following database best practices are recommended and used by top vendors and service providers, including Microsoft.

- Enterprises should map out their current data topology along with their 5-year projected topology. The topology should highlight the key corporate data and their locations, define performance requirements for each DBMS, and document the enterprise's overall data landscape.
- The database architectural design should be finalized before products are selected to ensure that the solution is not constrained by product limitations.
- Database administrators must plan capacity management and have a strategy for accepting information that goes beyond the planned-for capacity. Vendors recommend reclaiming unused disk space by defragmenting.
- Data replication should be kept at a minimum. The more instances of data there are, the greater the cost of ownership. Additionally, the probability of data inconsistency is increased significantly with replication.
- Backup hardware is available that can optimize restoration performance and backup procedures. Database administrators are advised to back up every night or at the close of business, monitor backups, and test and practice backup and restoration practices regularly.
- The database administrators are advised to track the database engine and pay attention to error messages.
- Information systems (IS) organizations should aim to simplify their DBMS environment by reducing the number of products they support. Most enterprises simply must support multiple databases (both legacy and new), so it is more realistic to define a short list of supported DBMSs than to support only a single DBMS solution.
- All data access requirements should be identified to ensure the flow of data updates can be established.

F.4 Database Tools

This chapter reviews the latest versions of database tools currently used by large-scale enterprise organizations. Included is a brief description of the intended design and function of each tool followed by the advantages and disadvantages to using each tool.

F.4.1 Development and Implementation Tools

F.4.1.1 IBM VisualAge. IBM VisualAge is a family of development toolkits designed to aid application developers in migrating from traditional coding practices to visual development. Designed with easy, ongoing integration with other IBM software products, VisualAge offers support and scalability for almost any development environment. VisualAge products support traditional and object-oriented development and promote emerging Web standards, such as the eXtensible Markup Language (XML) and Java. Furthermore, toolkits within VisualAge can convert developed applications to almost any platform, including Windows NT, AIX, AS/400, and S/390 environments.

Because VisualAge relies heavily on Java to support multiple platforms, it is subject to the same drawbacks as Java. The most noted Java performance issues are reported on both the client and server side, but the specific issues surrounding the performance problems are dependent on the Java implementation itself. While IBM is working with Sun Microsystems to continuously improve Java performance on their specific platforms, potential customers should factor the performance issue to their equation in their selection process. Another drawback is the heavy hardware resources VisualAge products need to perform at acceptable speeds.

F.4.1.2 Microsoft Visual Studio. Microsoft's Visual Studio 6.0 is designed for creating enterprise-level Web applications. Available in two versions (Professional and Enterprise), Visual Studio incorporates Microsoft's Distributed interNet Applications Architecture (DNA), for creating Web-based and, in many cases, cross-platform applications using the Component Object Model (COM)/Distributed Component Object Model (DCOM). According to Gartner Group's DataPro, MS Visual Studio is probably the most complete package available, with excellent connectivity to all sorts of databases and legacy applications, built-in transaction processing and monitoring, and desirable features like remote debugging for Web sites and third-party database code. A major drawback of Visual Studio is that it only supports Microsoft's component and operating system standards, which limits it to only Windows NT, and certain IBM mainframe platforms. Furthermore, Visual Studio does not support Common Object Request Broker Architecture (CORBA) objects, which will detract a core group of developers from using this product.

F.4.1.3 Oracle Developer. Oracle Developer 6.0 is a design and construction tool that supports rapid application development (RAD) of client/server database applications. It incorporates a set of builders (Form Builder, Reports Builder, etc.) for creating forms, reports, charts, queries, database objects, and procedures, as well as a project builder for team development. In addition, Developer's wizards automate the creation of an application through a point-and-click interface, enabling users to build and extend applications with reusable components stored in an Object Library. A major strength of Oracle Developer is its support for multiple development approaches, including advanced business process modeling, business process reengineering (BPR), information engineering (IE), and RAD. Developer also supports many databases on different platforms, including Microsoft SQL Server, Sybase, Informix, and of course, Oracle 8. Oracle Developer has a major drawback. It lacks integrated enterprise system management tools such as configuration management and workflow automation, although those can be obtained through third-party vendors.

F.4.2 Commercial Database Engines

F.4.2.1 Sybase Adaptive Server Enterprise. The latest version of Sybase Adaptive Server Enterprise (ASE) is designed to support the requirements for applications of Web integration with the addition of a Java Virtual Machine (JVM) and expanded support for XML. Java has replaced SQL as Sybase's common language processor for database environments. ASE stores and runs Java objects in a relational database, which is run by an internal Java database connectivity (JDBC) driver running as part of the database server that enables JavaBeans to access relational data directly.

Java offers the advantage of applications running faster, more securely, and more reliably, which makes Java an appropriate application for large-scale on-line transaction processing (OLTP), electronic commerce, and data-warehousing. Java can be used to add user-defined functions, stored procedures, types, and behaviors that are accessible to any SQL client and others. Also, Java is more portable than SQL. It can be ported across all tiers of a multitier application for both clients and servers. The development of new client applications is much easier with Java.

Because improved querying is a priority of MAF/TIGER Modernization, it should be noted that Java maximizes the productivity in OLTP environments by providing optimized thread management to allow for prioritized queries. However, Java needs to be combined with a data description language, whereas SQL is both a data description language and a data retrieval and manipulation mechanism. Sybase ASE is tightly integrated with XML, which enables the database to exhibit features, such as backup and recovery, persistent storage, and integration into the enterprise.

F.4.2.2 Oracle8i. Oracle8i is Oracle's latest version specifically designed for application deployment on the Internet. Like Sybase, a JVM has been added as an alternative to pipeline structured query language (PL/SQL), which enables JVM to access multiple data types over the Web.

Oracle8i provides two options for managing complex data: interMedia and iFS. The interMedia option supports audio, video, and image data for Web applications and authoring tools. The iFS option provides a single data store for indexing, managing, and querying data of virtually any type (relational, nonrelational, or both).

Oracle8i provides features for improving database availability such as index rebuilds and automated standby data support for disaster situations. Oracle offers the same benefits of Java (as stated above for Sybase) including portability, optimized thread management, and the capability of data-warehousing. Oracle8i places a lightweight directory address protocol (LDAP) compatible directory on the relational database itself, which offers a higher degree of scalability and availability that is not possible with traditional directories (can store up to 500 million entries and support 10,000 concurrent users). Users must be connected to the network, however, to run these applications, whereas Java solves this by employing a single-base code for both on-line and off-line operations.

The advantages of using Oracle8i include enhanced query performance through hash partitioning and composite partitioning. The high-availability features that allow for a number of options for disaster recovery are beneficial for data integrity. Oracle8i offers scalability, a top priority for Internet users, allowing tens of thousands of concurrent users to connect over multiple network protocols. Finally, Oracle is recognized as a market leader in database technology.

F.4.2.3 SQL Server 7.0 (Microsoft Corporation). Microsoft SQL Server offers database design and development architecture suitable for mobile applications, which is a critical component of MAF/TIGER Modernization and may improve data accuracy.

The advantages to using Microsoft's SQL Server 7.0 consist of the following: SQL Server retains the ease-of-use elements to which Windows users are accustomed. Database developers may use SQL's Visual Data Modeling tools, which provide a graphical interface, or the well-known Access interface, because Access programs are compatible. Like MS Windows, SQL Server is easy to administer and is tightly integrated with NT system administration facilities.

SQL Server offers Dynamic Database Sizing that alters the database size automatically, eliminating the need to enlarge or create a new device where the databases reside. SQL Server has a distributed heterogeneous query that can be used to access data from multiple sources at one time. SQL Server packages all the data into a single result set and returns the answer to the client personal computer (PC). For improved mobile usage, a distributed merge replication enables a remote system to periodically connect to a central server and synchronize table changes.

Finally, users have indicated that SQL Server is less expensive to own than other DBMSs. In terms of manageability and volume capability, SQL Server has achieved the record for supporting the highest number (21,000) of concurrent users with a specified response time via the PeopleSoft Human Resources Management System (HRMS) benchmark.

One of the reported limitations of SQL Server 7.0 is limited support to Windows platforms, which may limit its scalability in certain environments. Another limitation is that SQL Server has fewer vertical market applications (a few hundred) as compared to competitors Oracle and DB2 that have tens of thousands of independent software vendors (ISV) and business partners. Users have also reported that automatic shutdown occurs when the database detects corruption or a hardware problem that cannot be repaired during runtime, affecting the reliability and availability of the system.

Additional limitations are as follows: not all SQL Server features are available on all platforms. Upgrades may require significant disk storage investment, requiring a user to run two different sets of database management tools until the conversion is complete. SQL Server requires Internet Explorer 4.0 to operate. SQL Server lacks integrated hybrid on-line analytical processing (OLAP) backup utilities.

The SQL Server and Microsoft Office 2000 combination can bring the basic query, reporting, and OLAP technology to all of the company's users, including remote users as a low-cost, entry-level data-mart platform. SQL Server's ability to successfully scale to the enterprise level is questionable because it is dependent on the scalability of Windows 2000. SQL Server is recommended for small- to medium-size businesses and for global enterprises that need to extend decision support over the Web.

F.4.2.4 IBM DB2 Universal Database. IBM has brought the capabilities of a mainframe database to minicomputer platforms in the form of Release 5.0 of DB2 Universal Database (DB2

UDB). Version 6.1 enables Internet and mobile computer users to take advantage of DB2 capabilities. One of the improvements to the UDB server is XML and Java Stored Procedure Builder support, which is a graphical tool for rapid development of server-side logic in Java. A strength of the DB2 is the database administrator tool because of its browser-like interface that guides database administrators step-by-step through performance configuration, tool setup, and database monitoring and management.

DB2 has received praise from developers for its scalability, programmability, and multimedia handling. One of DB2's offerings includes the ability to perform parallel database searches, dramatically reducing lapsed time for query processing. The time reduction is the result of the system's capability of running a single query on multiple nodes of a parallel processing system. The system administrator defines a set of rules that the query governor follows allowing for customizability. DB2 UDB Enterprise Extended Edition (EEE) can support inter- and intra-query parallelism in which individual queries are processed in parallel across multiple processors resulting in significantly reduced response times.

Database administrators have claimed that DB2 is easy to administer with the help of graphical tools and SmartGuides for installation and maintenance. Version 6.1 also allows database administrators to manage large groups of satellites (laptops) with similar data configuration requirements. One configuration change to the group's model satellite can be propagated to all satellites in that group.

Additional strengths of DB2 are that database administrators can access object linking and embedding (OLE) DB data sources including Jet, Access, SQL Server, and Index Server from Microsoft, Intersolv Connect OLE DB for Notes, and ISG Navigator. The optimizer of DB2 UDB EEE is responsible for generating the parallel query execution techniques for all SQL statements. It determines how all data is accessed and updated. The optimizer ensures that even poorly written SQL performs optimally.

Some limitations of the system are that it lacks fundamental object-oriented characteristics (inheritance, polymorphism and CORBA support). DB2 requires a professional database administrator to maintain local DB2 servers. In comparison with competitors, DB2 uses a different version of JVM for different operating systems unlike Oracle that uses its own JVM tailored to Oracle database requirements.

F.4.2.5 NCR Teradata RDBMS (NCR Corporation). NCR Teradata is a parallel RDBMS optimized for decision-support applications. It is designed to support scalable growth and extremely large databases at a manageable cost.

Teradata supports parallelism by recognizing that certain steps are not dependent on each other and can be executed in parallel. Teradata also supports pipelining that allows a new task to start before previous tasks are completed. Other offerings of Teradata include an extensive administrative tool suite that is user friendly; provides automatic, transparent recovery from hardware failure; and automatically rebalances the system workload. Databases can be backed up directly on the Teradata platform or on a mainframe.

F.4.2.6 Informix Dynamic Server (Informix Software, Inc.). The Informix Dynamic Server is designed for high-volume applications, where linear scalability is essential. Informix Dynamic Server provides one solution that by-passes operating system limitations to take advantage of multiprocessor architectures to meet the scalability requirement.

Informix's strengths include dynamic shared memory that is pooled among a group of virtual processors, allowing Informix to be configured to automatically add more memory to its shared memory to process client requests more efficiently. Shared memory reduces fragmentation, avoiding degradation of memory use over time. Shared memory segments are automatically added as needed but can also be added by the administrator while the database is running.

Asynchronous input/output (I/O) is used to speed up I/O processing, which is usually the slowest part of database processing. This eliminates waiting for an I/O to complete before beginning work on another service request. Informix Dynamic Server is also capable of reading, joining, and updating tables on different computer systems during a single transaction in which data is shared between multiple databases. This supports consistency throughout databases when certain data is modified.

Informix Dynamic Server also offers star schema. Star schema support provides a query-centric view of the database by allowing users to formulate digital satellite system (DSS) queries and perform multidimensional analysis. Advanced indexing for decision support offers several options to improve data retrieval for a wide array of queries. Such indexes minimize storage space while offering significant gains in data access abilities.

Optimized load performance has been improved to provide rapid data load performance, which is a factor that is critical to a data warehouse where data loads and unloads occur regularly.

Informix also has its limitations. Development of multimedia applications requires low-level programming to application program interfaces (API) to direct data between the database and the application. Economically and technically, Informix holds a relatively small market share compared to its competitors, which results in a lower number of skilled information technology workers to support the company's products.

F.5 Summary

Comparative testing of database products using MAF/TIGER data is recommended to determine the most efficient software.

Note: *The discussion of specific vendors and products in this document is for the sole purpose of offering representative examples of products available for specific technology needs. This information is not intended as a promotion of any specific vendor or product. Additionally, the vendors and products discussed herein should not be considered exclusive from other potential vendors and products in the event of future procurement requirements.*

APPENDIX G. INTERNET-ENABLED APPLICATIONS

Implementing Internet connectivity allows computer applications to be widely distributed, and accessed anywhere, by anyone with access to the Internet. In the U.S. Census Bureau's modernization effort, providing Internet access to the MAF/TIGER data bases will modernize the collection methodology and increase efficiency of updating data in the MAF/TIGER data bases. Internet applications also enhance data integrity by reducing unnecessary data replication and eliminating stale data. Key considerations and best practices for developers implementing Internet applications are provided in this document.

G.1 Relevance to MAF/TIGER Modernization

Internet applications can aid the U.S. Census Bureau in achieving the following goals of the MAF/TIGER Modernization.

- **Expand Level of Participation in Partnership Programs—**Internet-enabled applications will afford greater accessibility to MAF/TIGER partners by providing a user-friendly, standardized interface through which the partners can submit their data.
- **Modernize Data Collection and Processing Methodology—**Use of remote access and live updates will enable field agents to update MAF/TIGER in real time, reducing human error and improving the accuracy of MAF/TIGER data.
- **Institute Systematic Evaluation Process—**Enabling MAF/TIGER access via the Internet will allow field agents to compare field data against the data within the databases. If a discrepancy exists, the agent can resolve the incongruity instantly.

G.2 Key Issues

There are three primary issues in enabling applications for Internet access: security, capacity and size planning, and remote access.

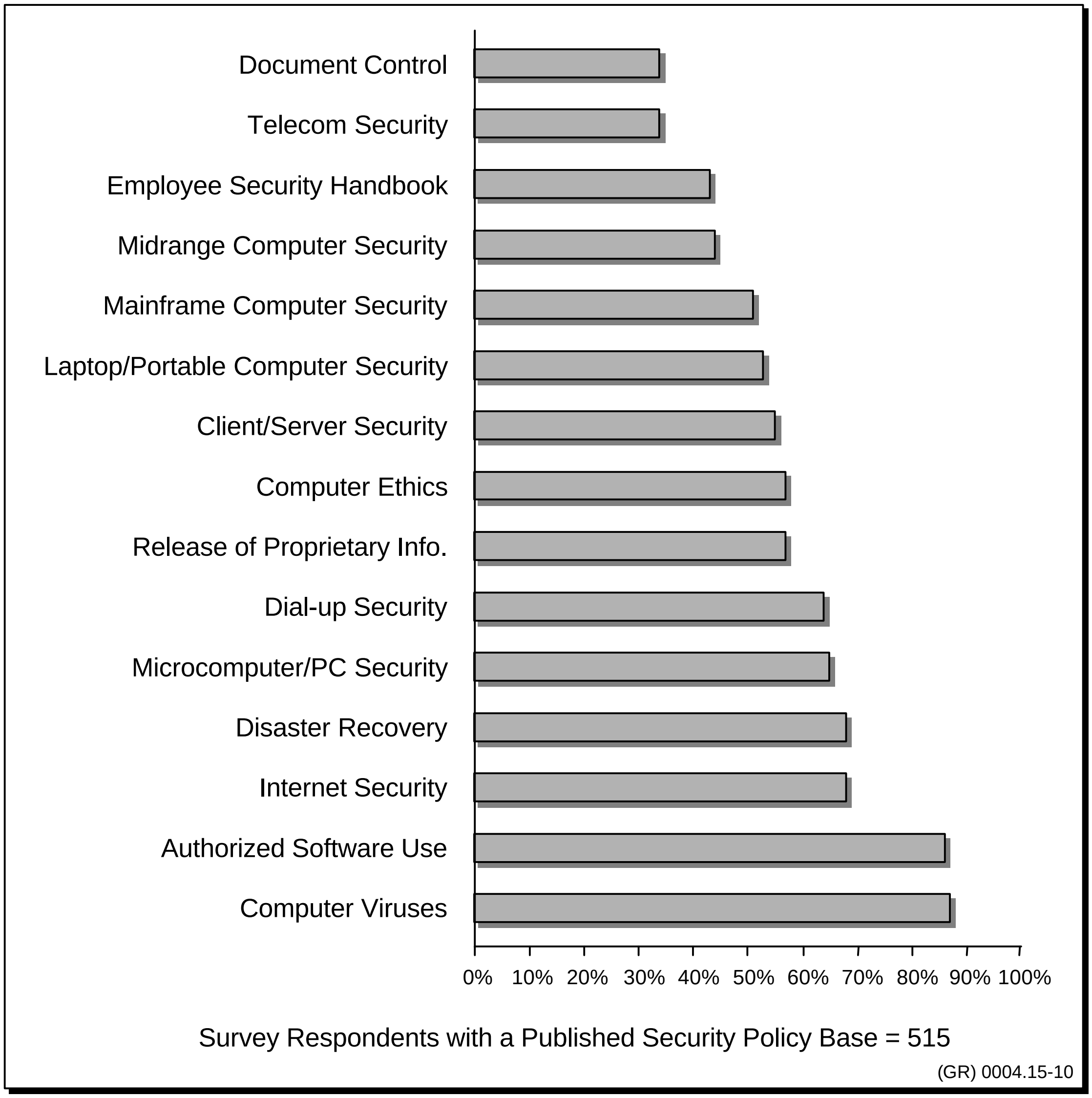
G.2.1 Security

Security has always been a key issue in the information technology (IT) community, but until recently it was often only an afterthought in the systems planning and design process, not an integral part of the systems planning and design phase. As a result, security was implemented ad hoc, without a formal plan, often after systems were in place. This approach left systems without a solid security foundation and left them vulnerable to attacks.

The recent attacks on popular Web sites have changed the way the IT community views security. Instead of designing and implementing security to blanket a system, security planning is now integrated into the design process itself.

The security plan defines the environment needing protection, plus the level of coverage desired, and identifies the initiatives or solutions needed to accomplish the desired results. A recent survey of 515 organizations indicated the range of topics that are covered by such security plans, as shown in Figure G-1.

Figure G-1. Topics Covered by Security Policies

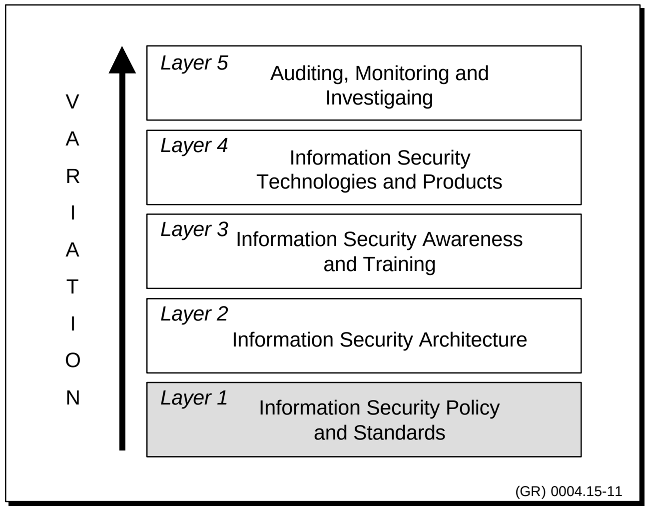


Source: Gartner Group Datapro information security survey, 1998

G.2.1.1 Security Best Practices. No single piece of hardware or software can provide complete security for a system. Instead, IT security is a blend of applied policies, best practices, and technology coexisting to ensure a secure environment. Although no system is completely secure, the following best practices for Web site security can greatly reduce security risk.

- **Create a Security Policy Framework.** The security policy framework shown in Figure G-2 consists of five levels of security practices and incorporates the policy statement, roles and responsibilities, and risk management and analysis. Without a security framework, tools and technologies will have little long-term effect on risk management.

Figure G-2. Five Layers of a Security Policy Framework



Source: GartnerGroup

- **Provide Security Awareness Training for Webmasters**. The Webmaster, who typically has front-line access to the firewall, proxy server, and Web usage and tracking tools, is in an ideal position to identify potential vulnerabilities and breaches in the security perimeter. Therefore, it is imperative that Webmasters receive adequate security training.
- **Designate a Web Security Administrator.** The Web security administrator should have responsibility for examining scripts, directory trees, and Java executables. Misconfigured scripts and improper directory permission settings are common points of vulnerability and targets for hacker security breaches.
- **Maintain Administrative and Password Control Procedures.** This practice includes creating secure password policies, maintaining active and accurate log files, and disabling any unused features.
- **Centralize Web Servers.** Server administration, authentication, and access control lists can best be secured from a central location.
- **Offload Critical Security Services.** Most enterprises will benefit from offloading transaction-intensive security services, such as authentication and certification, from their core Web sites. Offloading these services increases the speed with which sites can process Web requests and increases the security of the site's perimeter.

G.2.1.2 Security Vendors and Products. The following are examples of security-related products and solutions.

G.2.1.2.1 Cisco PIX Firewall. Cisco introduced the Private Internet Exchange (PIX) Firewall series in 1994. The PIX Firewall is a high-performance firewall that uses stateful packet filtering. The PIX Firewall is essentially a firewall appliance—it has its own integrated hardware and software solution (Intel hardware and proprietary operating system [OS]). The PIX Firewall is not UNIX or NT based but is based on a secure, real-time embedded system, known as the adaptive security algorithm. One strength of the PIX Firewall is certification by the International Computer Security Association (ICSA). In addition, in March 1999, the PIX Firewall received certification for compliance with the Common Criteria for Protection Profile for Firewalls for low-risk environments from the National Security Agency. A drawback of the current PIX Firewall Manager (PFM) is that it is awkward to use for setting up all but the simplest security access policies, which are based on services, hosts, and networks. Another drawback is that, with PFM, reordering rules in the security policy is not a simple procedure—the original list of rules has to be removed, and then a new rule list must be inserted.

G.2.1.2.2 Network Associates Gauntlet Firewall. Gauntlet's new firewall technology, Adaptive Proxy, is based on adaptive proxies that use a combination of packet filtering and secure proxy technology. New features for Gauntlet include McAfee virus content scanning functionality and enhanced performance. Gauntlet integrates virus scanning at the proxy level and includes options for enabling e-mail, Internet, and File Transfer Protocol (FTP) virus scanning. Strengths of the Gauntlet Firewall include support for multiple encryption methods, MD5 and SHA-1 authentication, and automatic and dynamic key exchange between virtual private network (VPN) hosts. In addition, Gauntlet has an extensive range of proxies, plus plug-and-play proxy support, which enables new proxies to be added and updated as they become available. Limitations include the nonintuitive design of the management interface. Reporting and alerting are adequate, but output is basically log entries, requiring use of a third-party product to sort or filter the information. Gauntlet is available either as a stand-alone product or as a component of Net Tools Secure.

G.2.2 Capacity and Size Planning

On October 1999, the Strategis Group reported that the on-line population of the United States had hit 101 million, which represented a 20 percent increase over 6 months. With such rapid growth in the Internet population, a main concern for most systems planners has become capacity and size planning. Without proper capacity and size planning, most Web sites will experience significant decrease in performance and even downtime during surges in user volume.

The usual way to discuss load on a Web site is in terms of "hits," also known as Hypertext Transfer Protocol (HTTP) operations. A hit is a single request from the Web client to the Web server. A given Web page can receive multiple hits in a single user visit, one for each element on the page. Daily peaks in hit volume average three to five times greater than the average daily load. Many U.S. sites experience their greatest loads from 2 p.m. to 5 p.m. EST. Preparing a site for a volume five times its average daily load will enable it to handle most transient spikes but is not enough for unexpected surges. Proper load balancing of servers and adequate bandwidth will enable a Web site to function normally even during unexpectedly heavy network traffic.

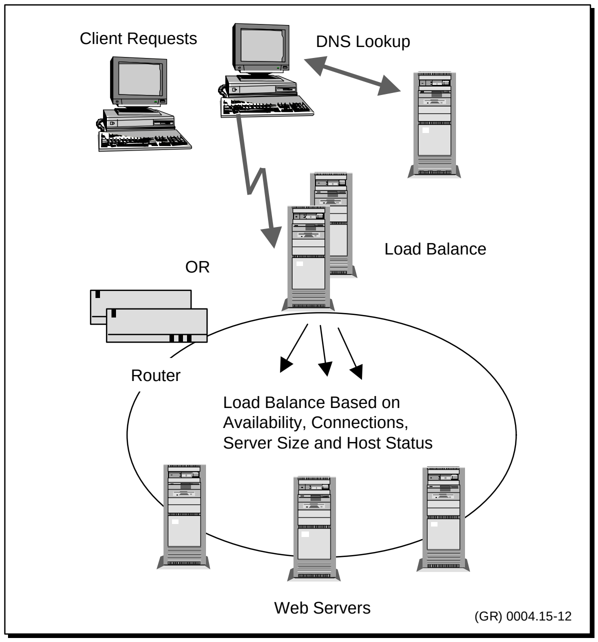
The most basic approach to achieving higher levels of Web-based application availability involves redundancy. Means of providing redundancy include multiple points of presence, redundant servers (and content), and use of an uninterruptible power supply for power outage protection. The typical Web application architecture involves multiple front-end Web servers taking client requests and serving them directly.

G.2.2.1 Solutions for Server Load Balancing. Three solutions for server load balancing are round-robin domain naming system (DNS), intelligent load-balancing technologies, and mirror sites. Each has costs, advantages, and drawbacks, which are described below.

G.2.2.1.1 Round-Robin DNS. This solution assigns multiple Internet Protocol (IP) addresses to a single domain name. When a client makes a uniform resource locator (URL) request, the DNS response shuffles the list of addresses returned to the client, thus redirecting the connection to a succession of available machines. Round-robin DNS can be implemented at nearly zero cost; however, it does have drawbacks. A major drawback is that round-robin DNS does not consider differences in performance among machines on the list and does not know the current status of any of the machines (any one of which might be inoperable at any given point.) Caching of IP addresses on the client side also can defeat the round-robin cycling by forcing the end user to list the same server repeatedly.

G.2.2.1.2 Intelligent Load-Balancing Technologies. Intelligent load-balancing technologies are network-based (router or bridge) or server/software-based solutions that redirect client connection requests from the load balancer to a specific server (see depiction in Figure G-3). The load balancer is configured with policy information about the Web servers to which it sends requests, such as maximum number of connections and the relative server size. By evaluating policies and current connections, this solution routes traffic to less utilized servers. Some load balancers support multiple geographic locations and can redirect requests according to the client's location, based on the topology of the Internet. The main drawback of intelligent load-balancing technologies is cost of implementation.

Figure G-3. Intelligent Load Balancing Representation



Source: GartnerGroup

G.2.2.1.3 Mirror Sites. The mirror site solution is based on establishment of multiple servers with redundant content. Mirror sites leave load balancing to the individual end user by requiring the user to select from a list of available sites. Although inconvenient for daily use, mirror sites can form part of an overall contingency plan to deal with peak loads on the system. Mirror sites can be implemented at a minimal cost.

G.2.2.2 Capacity and Site Planning Best Practices. The following site planning best practices should be followed:

- Site managers should configure their sites to handle at least twice the largest expected peak use, to cope with traffic growth and unexpected surges in demand.
- Site managers should upgrade to the most recent release of their HTTP server software and make sure operating systems are configured according to relatively well-known tuning parameters.
- The server should be zero hops, or at most one hop, away from a national Internet backbone and should have multiple redundant connections using different backbone providers.
- IT architects should implement a flexible and robust site architecture. A good architecture partitions the functionality of the site into distinct boxes that can be monitored and managed separately, and more important, scaled separately. This architecture allows for triage in the event of unexpected surges in demand. In a crisis, site administrators can rebalance the architecture by temporarily removing servers for less important functions and redeploying these boxes for more critical functions.
- Web designers should minimize unneeded design features. The home page should minimize the number of Java applets, ActiveX components, and multimedia files it uses.

- Web designers should redesign Web pages. Experienced Web designers have a repertoire of techniques for streamlining Web page design while maintaining a high level of visual impact. Redesign of a dated Web site not only will freshen the look of the site, but also can improve the perceived speed of the site, often by a factor of two to five.

G.2.3 Remote Access

A direct method of Internet-enabling applications (not involving the World Wide Web) is to provide remote access to the application server itself. With this method, instead of accessing a Web page, the U.S. Census Bureau employee (or partner) would remotely, but directly, access the MAF/TIGER application server by *remote control access, remote node access*, or a combination of the two.

G.2.3.1 Remote Control Access. Remote control access involves dialing into a modem-equipped or local area network (LAN) attached server to access its files or applications. This method requires special remote control software running on two computers, the remote client and the host server. As long as the software is installed on both machines, the remote computer assumes the capabilities of the host. Everything on the host's screen is mirrored on the remote client.

Popular remote control access packages include Citrix Metaframe, Microsoft Windows NT Terminal Server Edition (TSE), and Wyse Winterm.

G.2.3.1.1 Citrix Metaframe. Citrix Systems, Inc., is considered a pioneer in remote control access software. Basic functionality of Metaframe includes the ability to display 32-bit Windows applications on both Windows and non-Windows clients. A primary advantage of Metaframe is that it has proven scalability. The software's proprietary Independent Computer Architecture (ICA) protocol has been licensed to many clients, including chief competitors Microsoft and Wyse Technologies. Additional features, including load balancing and support for Sun Solaris, further influence most enterprises to deploy Citrix Metaframe. A drawback of Citrix Metaframe is its high licensing cost.

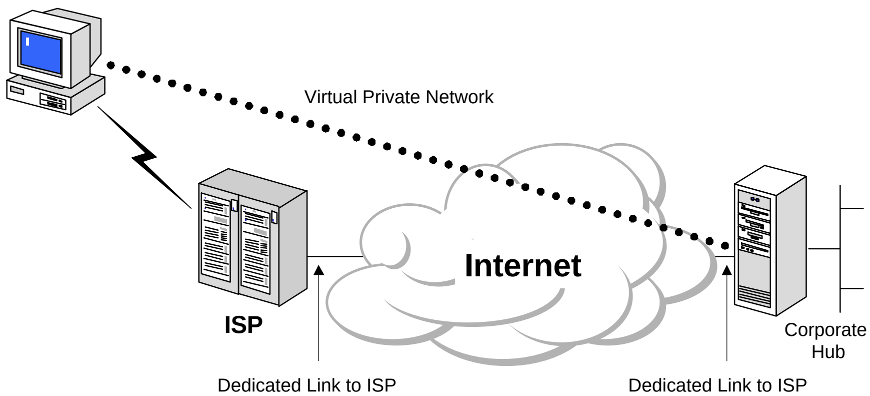
G.2.3.1.2 Microsoft Windows NT Terminal Server Edition. Microsoft's Terminal Server Edition (TSE) is a time-sharing environment for displaying Windows applications (with the Windows 9x graphical user interface [GUI] appearance). The server executes all user interface logic and applications. The only data transferred between the client and the server are the keystrokes and mouse movements needed to display information on the screen.

G.2.3.1.3 Wyse Technologies Winterm. Wyse is the volume leader in Windows terminals, with almost 40 percent of the market. Unlike its competitors, Winterm is both a hardware solution and a software solution. In addition to offering similar functionality to Citrix Metaframe and Microsoft Windows NT TSE, Wyse Winterm offers emulation for UNIX mainframes and compatibility with its proprietary operating system, the Winterm OS. A further advantage is that Winterm costs are lower than those of both of its chief competitors.

G.2.3.2 Remote Node Access. With the remote node approach, the computer that the user dials into attaches to the LAN as a remote node. The remote node architecture is very similar to the client/server architecture. The remote system acts as a client, and a communications server provides access to the desired network resources. The remote computer does not control another computer as it does in the remote control method; instead, it runs regular applications as if it were directly attached to the LAN. One common method of remote node access is the use of VPNs.

G.2.3.2.1 Virtual Private Networks. The rapid growth of remote access is driving many enterprises to build VPNs. A VPN connects the components of one network over another network as shown in Figure G-4. From the user's perspective, the VPN is a point-to-point connection between the user's computer and a corporate server. The nature of the intermediate Internetwork is irrelevant to the user because it appears that the data is being sent over a dedicated private link. VPN can be used to provide remote access to corporate resources using the public Internet while maintaining privacy of information.

Figure G-4. Using VPN to Connect a Remote Client to a Private LAN



Source: Microsoft VPN Whitepaper

Table G-1 provides a list of some VPN vendors and their available products.

Table G-1. VPN Vendors and Products

VENDOR	PRODUCT
3Com	PathBuilder
Axent Technologies	PowerVPN
Cisco Systems	Enterprise VPN
InfoExpress	VTCP/Secure
Intel Network Systems (formerly Shiva)	VPN Express
Lucent Technologies	VPN Gateway
Nortel Networks	Contivity Extranet Switch
WatchGuard Technologies	Remote User VPN

Source: GartnerGroup

G.2.3.3 Remote Access Best Practices. The following are remote access best practices:

- **End-User Training**—Remote users must be given more extensive technical training than their in-office counterparts because of the increased difficulty of obtaining help in remote locations.

- **Help Desk Availability**—Remote workers often do not have fixed schedules. If these individuals cannot get technical support when they need it, the enterprise has lost the value of having these personnel work outside the office.
- **Problem Ownership—**Responsibility for problem resolution should be placed with the support staff, not with the end user. Regardless of who is ultimately responsible for fixing the end user's problem, handling and routing of the problem must be the responsibility of the help desk.
- **Standard Configuration of Software and Hardware—**Remote users should use standard equipment and software. Standard configurations limit the number of incompatibilities and ease the troubleshooting efforts of the help desk.
- **Frequent File Backups—**Mobile and portable devices are inherently more susceptible to loss, theft, or damage. Therefore, a more robust (automated) backup solution is essential for remote and mobile users. Also, a traveling user's laptop (or other device) should never contain irreplaceable data.
- **Spare Laptops and Other Mobile Devices—**A pool of spare laptops and other mobile devices must be available for express mail to any remote or mobile employee who faces serious problems that require direct attention. Under no circumstances should the user be required to mail the device to the help desk and wait for repairs, which would cause at least 2 days of worker downtime.

APPENDIX H
ACRONYMS AND ABBREVIATIONS

4GL	Fourth-Generation Language
AAG	Association of American Geographers
ACF	Address Control File
ACS	American Community Survey
ACSM	American Congress on Surveying and Mapping
ADE	Applications Development Environment
AFF	American FactFinder
ALRB	Address List Review Branch
API	Application Program Interface
ARF	Address Reference File
ASA	Adaptive Security Algorithm
ASE	Adaptive Server Enterprise
ASPRS	American Society for Photogrammetry and Remote Sensing
BAS	Boundary and Annexation Survey
BIDC	Business and Industry Data Center Program
BLM	Bureau of Land Management
BOB	Beacon-on-a-Belt
BSA	Basic Street Address
CAD	Computer-Aided Design
CASE	Computed-Aided Software Engineering
CATI	Computer-Assisted Telephone Interviewing
CAUS	Community Address Update System
CIC	Census Information Center
CMM	Capability Maturity Model
COB	Cartographic Operations Branch
COM	Component Object Model
CORBA	Common Object Request Broker Architecture
CNStat	Committee on National Statistics
CSB	Computer Support Branch
DADS	Data Access and Dissemination System
DEM	Digital Elevation Model
DEX	Digital Exchange
DGPS	Differential Global Positioning System
DLL	Dynamic Link Library
DMAF	Decennial Master Address File
DMD	Decennial Management Division
DNS	Domain Naming System
DOQ	Digital Orthrophoto Quadrangle
DOQQ	Digital Orthophoto Quarter Quadrangle

DOT	Department of Transportation
DSCMO	Decennial Systems and Contracts Management Office
DSF	Delivery Sequence File
DSMD	Demographic Statistical Methods Division
DSSD	Decennial Statistical Studies Division
EEE	Enterprise Extended Edition
E.O.	Executive Order
EPA	Environmental Protection Agency
EPB	Economic Programs Branch
ESRI	Environmental Systems Research Institute
FEMA	Federal Emergency Management Agency
FGDC	Federal Geographic Data Committee
FIPS	Federal Information Processing Standards
FLD	Field Division
FNMM	Field Notes for Mobil Mapping
FTP	File Transfer Protocol
GAB	Geographic Areas Branch
GASB	Geographic Areas Systems Branch
GB	Gigabyte
GBF	Geographic Base File
GBF/DIME	Geographic Base File/Dual Independent Map Encoding
GDO	Graphical Data Objects
GEOCAT	Geographic Catalogue
GI	Geospatial Imagery
GIS	Geographic Information System
GITA	Geospatial Information & Technology Association
GPBS	Geographic Planning and Budget Staff
GPP	Geographic Program Participants
GPQA	Geographic Products Quality Assurance Team
GPS	Global Positioning System
GRaSS	Geospatial Research and Standards Staff
GRF	Geographic Reference Files
GSD	Ground Sample Distance
GSS	Geographic Support System
GTT	Geographic Technologies Team
GUI	Graphical User Interface
GUS	Geographic Update System
GUSB	Geographic Update Systems Branch
GUSX	Geographic Update System for X-windows
HHES	Housing and Household Economic Statistics Division
HRMS	Human Resources Management System
HTTP	Hypertext Transfer Protocol

I/O	Input/Output
IA	Imagery Analyst
ICMA	International City/County Management Association
ICSA	International Computer Security Corporation
IDP	Imagery Derived Product
IE	Information Engineering
IP	Internet Protocol
IRS	Indian Remote Sensing
ISO	International Standards Organization
ISO/TC	ISO Technical Committee
ISP	Internet Service Provider
ISV	Independent Software Vendor
IT	Information Technology
JAD	Joint Application Development
JDBC	Java Database Connectivity
JMDL	Java Macro Development Language
JVM	Java Virtual Machine
KGL	Key Geographic Locations
Km	Kilometer
LAN	Local Area Network
LCD	Liquid Crystal Display
LDAP	Lightweight Directory Access Protocol
Lidar	Light Detection and Ranging
LUCA	Local Update of Census Addresses
MAF	Master Address File
MAFGOR	MAF Geocoding Office Resolution
MAFOB	MAF Operations Branch
MAFUF	MAF Update File
MAPPS	Management Association for Private Photogrammetric Surveyors
MB	Megabyte
MOI	Material of Interest
MPT	MAF Products Team
MSB	Matching Systems Branch
MSC	Mapping Sciences Committee
MST	Mapping Services Team
NACo	National Association of Counties
NAPP	National Aerial Photography Program
NAS	National Academy of Sciences
NASA	National Aeronautics and Space Administration
NATaT	National Association of Towns and Townships

NCAI	National Congress of American Indians
NCAP	National Civil Applications Program
NCES	National Center for Education Statistics
NCSL	National Conference of State Legislatures
NDOP	National Digital Orthophoto Program
NGA	National Governors Association
NIMA	National Imagery and Mapping Agency
NLC	National League of Cities
NOAA	National Oceanic and Atmospheric Administration
NPC	National Processing Center (Jeffersonville, IN)
NRC	National Research Council
NRO	National Reconnaissance Office
NSDI	National Spatial Data Infrastructure
NSGIC	National States Geographic Information Council
ODBC	Open Database Connectivity
OGC	Open GIS Consortium
OLAP	On-Line Analytical Processing
OLE	Object Linking and Embedding
OLTP	On-Line Transaction Processing
OMB	Office of Budget and Management
OO	Object-Oriented
OS	Operating System
PCI	Peripheral Component Interconnect
PCS	Production Control System
PDA	Personal Digital Assistant
PIX	Private Internet Exchange
PL/SQL	Pipeline Structured Query Language
PLSS	Public Land Survey System
POB	Production Operations Branch
POP	Population Division
PRED	Planning, Research and Evaluation Division
PRS	Problem Referral System
PSS	Products and Services Staff
RAD	Rapid Application Development
RAM	Random Access Memory
RCC	Regional Census Center
RDBMS	Relational Database Management System
RFP	Request for Proposal
RO	Regional Office
SDC	State Data Center
SDE	Spatial Data Engine
SEI	Software Engineering Institute

SQL	Structured Query Language
SWIR	Short Wave Infrared
TB	Terabyte
TBD	To Be Determined
TEA	Type of Enumeration Area
TIGER	Topologically Integrated Geographic Encoding and Referencing
TIP	TIGER Improvement Program
TMB	TIGER Mapping Branch
TMU	Targeted Map Update
TOB	TIGER Operations Branch
TSB	TIGER Systems Branch
UDB	Universal Database
UML	Uniform Modeling Language
UOB	Update Operations Branch
URISA	Urban and Regional Information Systems Association
URL	Uniform Resource Location
USDA	United States Department of Agriculture
USGS	United States Geological Survey
USPS	United States Postal Service
VAR	Value-Added Resellers
VMS	Virtual Memory System
VNIR	Visible Near Infrared
VPN	Virtual Private Network
WWW	World Wide Web
XML	extensible Markup Language