An Integrated Multi-users Application Using a Centralized SDE Database

Yingjie Wang and Jay E Spurlin

Abstract

The Geographic Areas Analysis and Delineation System (GAADS) is being developed to facilitate the analysis and creation of collection geographic entities for the 2010 Census. The system is designed to support multiple concurrent users to remotely conduct reading and writing operations on a centralized database built using ArcSDE and Oracle. GAADS allows the users to import data, delineate entities, visualize and analyze entity properties, perform various quality assurance checks and mark error exceptions. GAADS has been used for the delineations of Assignment Area (AA), Crew Leader District (CLD), Field Operation Supervisor District (FOSD), and Local Census Office (LCO)/Early Local Census Office entities for the 2008 Census Dress Rehearsal and 2010 Census.

The tools and commands built in GAADS allow users to execute consistent database transactions and maintain data integrity and accuracy. The system design and implementation considerations also include the challenges of performance enhancement and data security management. This presentation will give you a brief description of the design and implementation of GAADS, as well as the lessons learned in the production using the application.

1. Introduction

The delineation processes for 2010 Census generally include batch delineations and interactive delineations. The Geography Division of the Bureau of the Census is responsible for creating the delineation software that will be used by the Field Division (FLD) and the Geography Division (GEO) for preparing for both the 2008 Census Dress Rehearsal and the 2010 Decennial Census.

The GAADS application contains the tools and commands that are designed to aid users during the interactive delineation processes. The interactive delineation processes include the delineation of geographic entities and attributes for different field operations during the Census. Some of these operations require batch delineation processes, and some of them require interactive delineation processes only. In many cases, the entities created from the automated batch processes are not optimal for field operations. Therefore, the tools and commands built in GAADS can be used to evaluate and adjust the entities in order to best fit the FLD's operational needs and take advantage of local knowledge in the Regional Offices.

The following sections describe the system design, implementation method, architecture, and the underlying database of GAADS.

2. Terms of Reference

The relational description of production workspace, user workspace, project, delineation layers, delineation objects, and the application platform are shown in Figure 1.

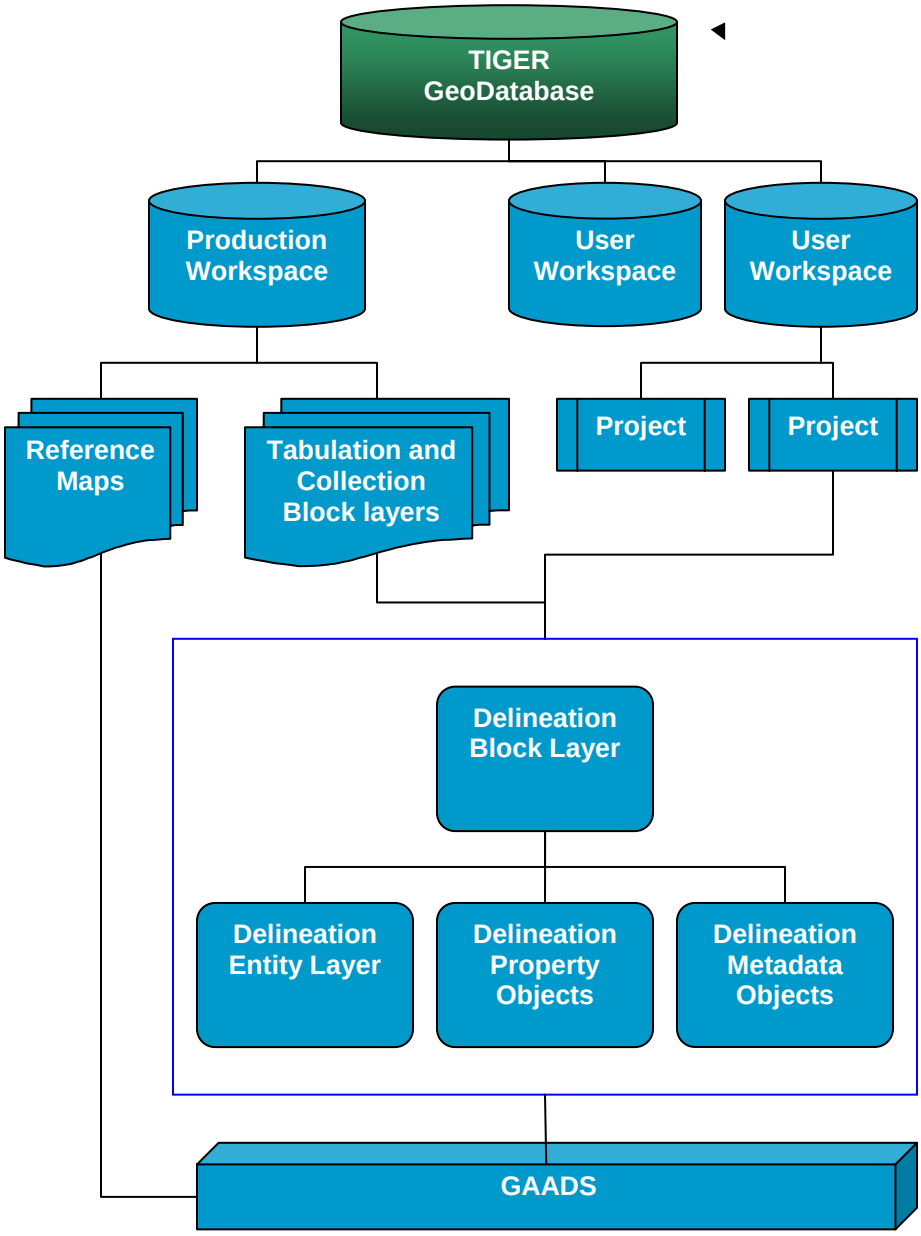

Figure 1. The concepts of production workspace, user workspace, project, delineation layers, and delineation objects.

2.1 User Workspace and Project

A user workspace, which is an ArcSDE dataset, is the physical separation of delineation data layers from other reference and production data layers. The delineation data layers are stored in the production database but under a different dataset that is defined as the user workspace. A project in a user workspace contains the layers that are created by a user for a delineation task. For example, for the Address Canvassing Assignment Area delineation (AC AA), a project will contain an extracted Block layer and the delineation entity boundary layer. The delineation entity boundary layer will contain the delineation boundaries of the AC AA entities.

2.2 Delineation Layers

The structure of the delineation layers is automatically constructed during the project creation process. A project contains the Block layer, delineation entity layer, delineation entity property table, and delineation metadata objects. The Block layer is built from the production Block layer with some added fields. Initially, the delineation entity layer contains no data but only the feature class structure that is needed for the delineation process. Different delineations have different structures of delineation layers, including the Block layers and the delineation entity layers. For example, LCO delineation layers are different from AA delineation layers. The delineation layers are populated in different phases of a delineation process.

2.3 Reference Maps

The required reference layers, referred to as reference maps, are created and stored under a production database schema that can be used for read-only access by the delineation users. Examples of reference layers are the Military Installation and American Indian Reservation layers. The purpose of loading reference maps is to provide a reference to the users for delineating geographic areas in the delineation layers. Users are able to load any reference maps from the database into ArcMap at any time during a delineation process.

3. System Architecture

3.1 System Overview

There are dependencies among the delineation processes and between the delineation processes and databases. Most of the delineations require the most current MAF/TIGER Database (MTDB) data, updated housing units and Address Characteristic Type (ACT) codes. The TIGER geodatabase that is built from a recent snapshot of the MTDB is the working database for all the interactive delineations. The AA delineation for most operations depends on the completion

of the LCO delineation, the CLD delineation depends on the completion of the AA delineation, and the FOSD delineation depends on the CLD delineation. Figure 2 shows the dependencies of most delineation processes.

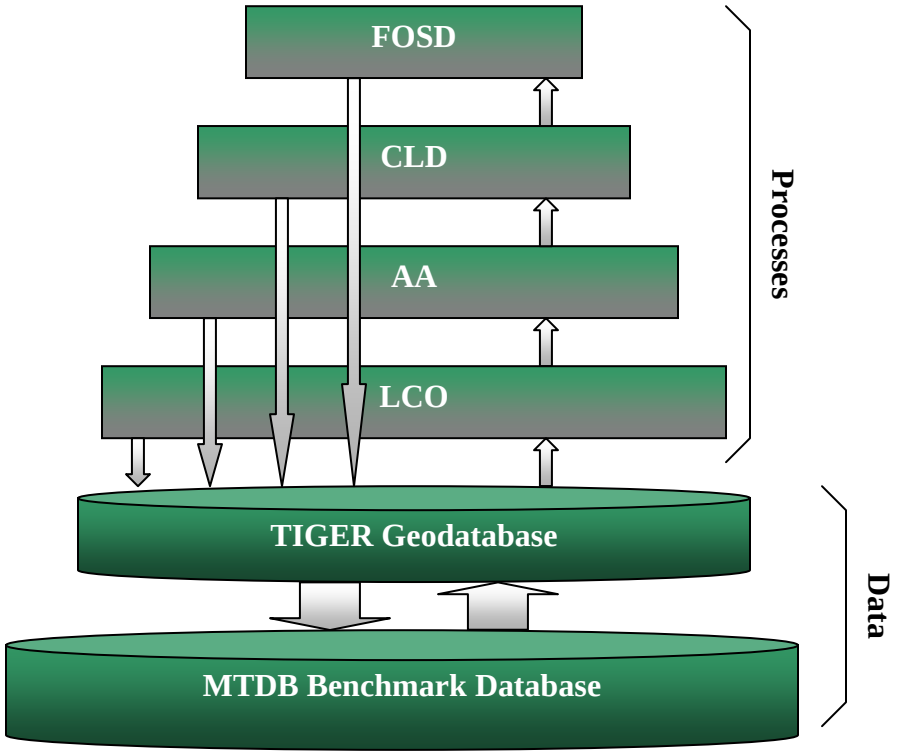

Figure 2. The dependency of delineation processes.

Figure 3 shows the dependencies of the Address Canvassing delineation processes.

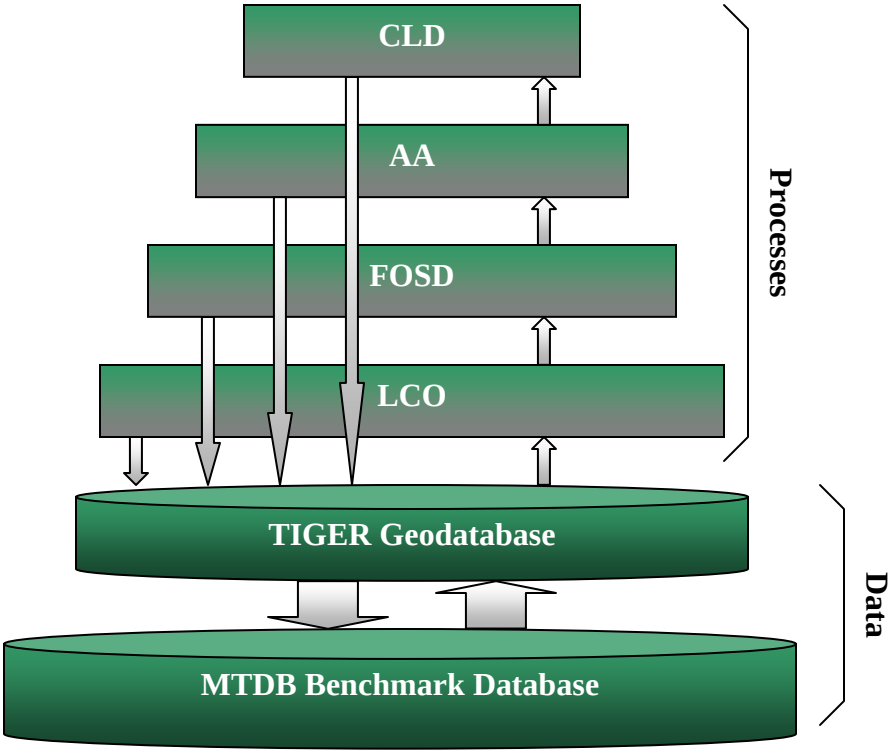

Figure 3. The dependency of Address Canvassing delineation processes.

The GAADS application is designed to have the capability and flexibility of conducting the interactive delineation processes for different Census operations in one integrated system. The goal of the system is to provide users with a flexible, scalable, friendly, efficient, and easy-to-use application. The system architectural approach includes the considerations of database integrity and performance, which are effectively handled through the data management strategic design.

The tools and commands built into the system allow users to perform standard database transactions and automatically maintain data integrity. Using the system, only valid data can be written into the database. When the database is taken from one state to another, the consistency with the database rules is automatically maintained. The design considerations also include the concepts that explain the essential components of the system and how they work together. The system supports non-versioned editing and multiple users through the relational database system supported by ArcSDE and Oracle Spatial where the data layers are built, registered, and stored.

3.2 Architectural Strategies

3.2.1 Processes and Data Flow

To initiate a delineation task, a user identifies or selects a subset of Blocks based on state code, county code, and/or LCO or ELCO codes. The selected Blocks will be loaded from the production Block layer into the user's delineation Block layer in the user workspace. The necessary Blocks can be imported at once or in multiple sessions.

The concepts of user workspace and project are used to manage the delineation layers and activities. A delineation process can only be conducted on the defined delineation layers within a user workspace. The delineation work, or user's edits in the user's workspace, can be validated by various predefined validation modules. The errors found through the validations can be either fixed or marked as exceptions using an error exception marking tool. After properly handling the errors, users can submit changes to the associated production data layer in the database. Figure 4 shows the architecture, processes, and data flow of GAADS.

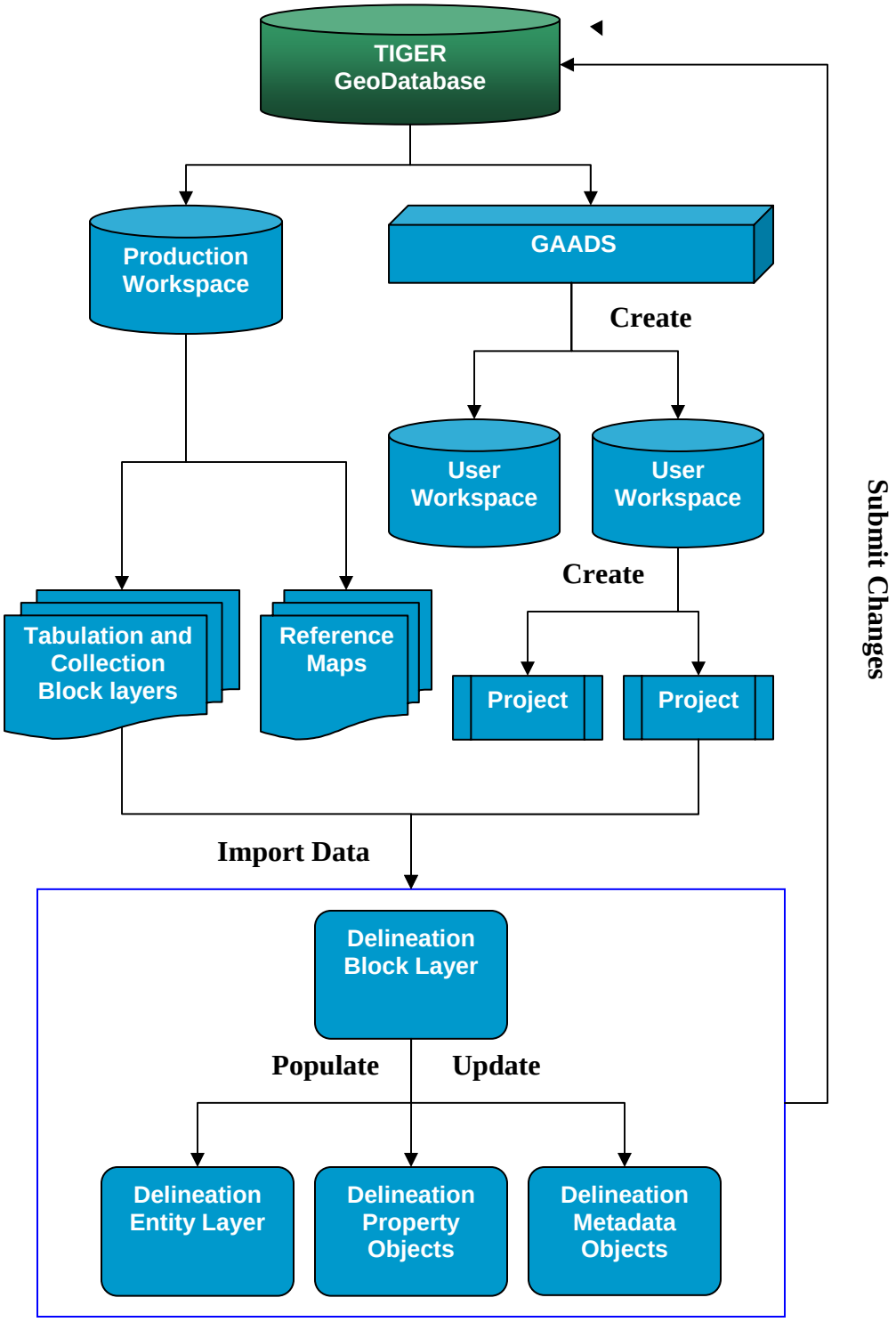

Figure 4. The architecture, processes, and data flow of GAADS.

3.2.2 The Design Considerations

1) **Maintain database integrity, consistency, and accuracy**
   The delineations are conducted on the delineation layers in a user workspace, which is separate from the production layers. The production database is updated through a programmed process that is fully proven. This methodology ensures the data always complies with the database rules.

2) **Improve performance**
   Users only work on the subset of a production database layer that is necessary for a specific delineation task of an operation. The production layer can be either a Census Tabulation Block layer or Collection Block layer, which depends on the timing of Census operation. Since SQL procedures can not be applied to ArcSDE versioned data, the ArcSDE versioning capabilities must be disabled and removed from GAADS to improve data processing performance for operations such as updating columns, inserting records, deleting records, data statistics, and complex computations. GAADS has been redesigned and reengineered to embed SQL procedures and execute editing against non-versioned database layers, in order to perform standard database transactions.

3) **Conduct QC in user's workspace**
   Before submitting changes, a user can validate the delineation layers in a user's workspace based on project specifications.

4) **Support multi-users**
   The application is designed to support multiple users. Multiple users can concurrently work in different user workspaces on the same database.

5) **Support normal and supervisory users**
   Normal users can only view another user's project. Supervisory users can view and edit another user's project, which includes updating the Block layer, updating the delineation property table, populating the delineation entity layer, validating delineation work, and submitting changes. The operations such as creating a workspace, creating a project, deleting a workspace, deleting a project, and importing blocks, can only be performed by the data owner.

6) **Show delineation property data**
   Delineation property information such as housing unit count, entity area, and block count can be updated and displayed during and after a delineation process.

4. Data Structure

Production data layers and reference maps such as tabulation block layer, collection block layer, military installation layer, and American Indian Reservation

layer are extracted from the MAF/TIGER Database. These extracted data layers are loaded and built in Oracle Spatial, registered with ArcSDE for Oracle, and made available to authorized users. The delineation Block layer, entity layer, property objects, and metadata objects are user-created data layers or objects that are stored in a user's project under a user's workspace. As mentioned before, a delineation Block layer is a subset of the corresponding production Block layer. In other words, a delineation Block layer contains only a small part of the production layer, which has been selected by the user. Delineation layers are editable, and they are the source data for submitting changes to the production block layer. A delineation entity layer contains the boundaries of all the delineated entities. A delineation property object or table (_PT, _PC, or _PF) contains the summarized information of all the delineated entities, such as housing unit count and entity area. The metadata objects include the table (_XT) that contains the delineation type and the layer extent of the project and the table (_CV) for LCO delineation that stores LCO numbers and names.
Reference maps can only be used for reference purposes; they can not be edited.

5. Implementation

The GAADS application is developed based on the architectural design described in the previous sections. The interface, commands, and tools of GAADS are implemented using *ESRI ArcObjects (ArcGIS 9.x)*. The GAADS interface is incorporated into ArcMap and consists of toolbars, buttons, and dialog windows. The tools and commands can be used to conduct the predefined tasks of interactive delineations. The system component objects are built independently even though they share public libraries and custom functions. This approach makes the application flexible and scalable. Different delineations or different operations can be easily integrated into GAADS. The application is designed to support multiple concurrent users at different levels of privileges through ArcSDE schemas built on top of an Oracle database.

6. Conclusion

GAADS was initially designed and implemented on ArcGIS 9.1 using an ArcSDE Versioning approach. The Regional Office staff of the Bureau of Census utilized GAADS for the delineations of the 2008 Dress Rehearsal Address Canvassing Universe (TEA [Type of Enumeration Area] Phase 1), Production Rate Area and FOSD, and AA/CLD entities. The customers were not satisfied with the performance of the early releases of GAADS. A lot of effort had been made to improve the performance, but significant improvements could not be achieved on most of the modules. After ArcGIS 9.2 was released, the option to edit non-versioned data became available. By default, ArcMap edit sessions are set to perform versioned edits, which means that the user can only edit data that has been registered as versioned. To perform non-versioned edits on non-versioned data, versioned editing must be turned off in ArcMap. This new capability of ArcGIS 9.2 provided the opportunity to significantly improve the performance of

GAADS. A great deal of work has been done to change the GAADS design and implementation from the versioned editing to non-versioned editing. Most of the GAADS modules have been rewritten with embedded SQL procedures. After reengineering the software, the performance has truly improved. Additionally, the GAADS application and the ArcGIS software suite are installed on a Citrix server instead of the user's workstation. The Citrix server and the database servers are based at the headquarters of the Bureau of Census. The remote users at the Regional Offices use the Citrix server to conduct their delineation tasks. This hardware architecture reduces the cost of data transmission on the network. During the development and reengineering of the application, the algorithms used in constructing the tools and commands have been optimized.

Using the application will significantly enhance delineation efficiency and maintain database accuracy and integrity. To a certain degree, users can share their projects, either ongoing or completed, for the delineation tasks on which they are currently working. Normal users can view the projects of other users and supervisory users can view and edit the projects of other users. Therefore potential duplicate efforts are eliminated.

Acknowledgements

Special thanks to the Bureau of the Census for contractual support. Development support was also provided by Sabre Systems, Inc.

About the Authors

Yingjie Wang is a contractor of the Bureau of Census. Dr. Wang's expertise includes Software Engineering, GIS, Algorithms and Applications Development, and Spatial Data Modeling. His major focus is in the areas of implementation of advanced GIS technology in federal and local government. He is a certified GIS professional by the Urban and Regional Information Systems Association.

Corresponding Address:

2200 Columbia Pike, #502,
Arlington, VA 22204
Phone: 301-763-8976 (office)
Email: yingjie.wang@census.gov
ywang@razavi.com

Jay E Spurlin leads the Application Software Team in the Spatial Data Software Branch, Geography Division, U. S. Census Bureau. He has been with the Census Bureau for over 15 years. His expertise includes Database Design, Software Engineering, GIS, and project management. His major focus is the creation of applications that support the Bureau's goals by implementing the

loading, maintenance, and extraction of spatial data in the MAF/TIGER Database.

Phone: 301-763-8972 (office)
Email: jay.e.spurlin@census.gov