US CENSUS BUREAU MAF/TIGER PRODUCT DATABASE: IMPLEMENTATION IN THE REDESIGNED MAF/TIGER SYSTEM ENVIRONMENT*

A. A. Danso

Geography Division, U.S. Census Bureau, 4700 Silver Hill Rd, Washington DC 20233 – ama.a.danso@census.gov

KEY WORDS: MAF/TIGER System, Topology, Spatial Data Storage, Geospatial Product Creation, Geodatabase, Address Data Products

ABSTRACT:

The accuracy, quality, and on-time delivery of geographic, spatial, cartographic, and address data products are integral to accomplishing the mission of the U.S. Census Bureau. These products support current surveys, decennial censuses, estimates programs, partnership programs, and the public need for geospatial data. For the 2010 Census field operations, the US Census Bureau's Geography Division (GEO) delivered over 17 million unique map sheets, 10 Geographic Reference Files (GRFs) containing 6.7 million blocks, and eight full MAF Extracts of over 144 million addresses. These products are generated from the Master Address File/Topologically Integrated Geographic Encoding and Referencing (MAF/TIGER) Product Database (PDB) that is created from the current MAF/TIGER transaction database (MTDB) through a Benchmarking process. The MTDB was developed as part of the MAF/TIGER Redesign project, and was a major step forward for GEO's management of spatial and address data. The system utilized Commercial Off-The-Shelf (COTS) software including Oracle Spatial and Topology Data Model. The design included persistent topology and a seamless national database that integrated the spatial and address data. It facilitated real time updates and maintenance of business rules. However, to effectively support continuous flow-based on-time delivery of data products, a multi-level geospatial database design approach was adopted. The design resulted in the implementation of a complimentary database to the current MTDB known as the MAF/TIGER Product Database (PDB). The PDB utilizes a denormalized database design approach where tables from the current MTDB are combined, derived fields are stored, and there is no explicit storage of topology. The PDB also provides a data structure that supports geospatial COTS application tools including ESRI's Geodatabase for spatial data visualization, analysis, and research.

1. INTRODUCTION

The U.S. Census Bureau Geography Division (GEO) redesigned MAF/TIGER system environment utilizes a multi-level geospatial database design model. The multi-level model effectively supports real time spatial and non spatial data updates, maintenance of business rules, and flow-based processing. The first part of the multi-level design model is the implementation of a current transaction-based MAF/TIGER database (MTDB). The second level involves a benchmarking process to produce a Benchmark Database. The third level is the implementation of a MAF/TIGER Product Database (PDB) that serves as the source for all product creation.

1.1 The Topologically Integrated Geographic Encoding and Referencing System (TIGER)

The Topologically Integrated Geographic Encoding and Referencing (TIGER) system is a street center-line "digital map" (geographic database) of the entire United States, Puerto Rico, and the associated Island Areas. The data content of TIGER includes the spatial representation, names, and other attributes for a variety of features including roads, hydrography, railroads, legal, and statistical geographic areas, address ranges, ZIP codes, and structure locations. The TIGER database provides a centralized repository for all the geographic information needed to support census and survey data collection, data tabulation, data dissemination, geocoding services, geographic and statistical analysis, and map production.

1.2 Master Address File (MAF)

The Master Address File contains an accurate and up to date inventory of all known living quarters in the United States and Puerto Rico. The content of the MAF includes permanent MAF ID, mailing address, location address, census block code, structure coordinate, source, and history data. The MAF provides support for field data collection efforts, address canvassing operations, and questionnaire deliveries to each residence. It also provides the basis for data tabulation.

The subsequent sections of this paper will describe the component systems making up the multi-level geospatial

* This paper reports the results of work undertaken by U.S. Census Bureau staff. This paper is released to inform interested parties of research and to encourage discussion. Any views expressed are those of the author and not necessarily those of the U.S. Census Bureau.

database design model of the redesigned MAF/TIGER system i.e. the current MTDB, the Benchmark, and the Product Database. Also included will be discussion of design considerations, product creation, processing challenges, and how the Census Bureau has been successful in applying best practices and process improvement methods to support customer needs.

2. MAF/TIGER TRANSACTION DATABASE (MTDB)

The first part of the MAF/TIGER system multi-level geospatial database model is the current MTDB. The redesign of MAF/TIGER resulted in an integrated system where both the spatial and non spatial address data elements are combined into a single repository. The MTDB provides a seamless national integrated spatial and address database utilizing a commercial object-relational database management system. Oracle object-relational database management system is used to store all data elements. Spatial data management including spatial operations and indexing utilizes Oracle Spatial. Oracle Spatial Topology Data Model stores and manages geographic features and the topological data structures on which they are built.

The current MTDB logical model employs database normalization along with storage of persistent topology to effectively support update operations, data integrity, and maintenance of business rules. The design decisions supporting a normalized database model for spatial data storage and persistent topology for the current transaction MTDB (Galdi, 2005) facilitate:

- Rapid spatial data retrieval
- Spatial data editing and clean up
- Improved data consistency
- Management of shared geometry (Hoel, 2003)
- Definition and enforcement of data integrity rules (Hoel, 2003)

In addition to the above benefits, the storage of persistent topology in the current MTDB provides:

- Efficient storage of spatial data
- Reduction or elimination of redundancy of spatial data
- Easier implementation of certain spatial business rules in the database
- Improved management of hierarchical geographic relationships (Ramage and Woodsford, 2002)
- A more straightforward way to address coordinate precision and tolerance issues that can lead to gaps or slivers (Egenhofer et. al., 1989)

Although the design decision to go with a normalized database design with persistent topology for the current transaction MAF/TIGER database effectively supports maintenance of business rules and performance with real time updates interactively and in batch mode, it may not be necessarily optimal for mapping applications (Trainor, 2003), data extraction, and creation of other geographic products.

2.1 The Need for a Benchmark

The US Census Bureau business process is cyclical in nature. This requires periodic updates to the current MTDB. The geographic boundaries and spatial alignments for legal entities such as state, county, minor civil division, and incorporated places occur through the annual Boundary and Annexation Survey (BAS). The administrative area updates occur through the Redistricting Data Program and the School District Review Program. The MAF data is updated bi-annually using the US Postal Service Delivery Sequence File (DSF). For the decennial census, updates to both the spatial and address data occur through numerous operations including Local Update of Census Address (LUCA), Address Canvassing, and various updates obtained during enumeration and verification activities. Census Bureau-defined statistical geographic areas such as census tract and census block are updated prior to each census through the Bureau's Participant Statistical Areas Program (PSAP) and Tribal Statistical Areas Program (TSAP).

The updates to the current MTDB occur on a flow basis to allow on-time product creation and delivery to meet data needs. To that end, products are generated from a snapshot of the data content at a point in time, typically twice a year. The result of taking a snapshot of the current MTDB data content is the creation of the MTDB Benchmark. Furthermore, the transaction database supports 24 X 7 batch update processing and 12 X 6 interactive updates, meaning it must be tuned to support these processes. Many indexes that might be useful for product creation are not allowed because they slow down transaction processing. A benchmark allows running of a series of edits to make sure the benchmark is valid and internally consistent. In effect, the benchmark allows the Bureau to have a stable, edited database to effectively support product creation. In the benchmark, any columns not maintained on a real-time basis in current MTDB can be calculated. Also, indexes can be created without fear of disrupting the update process.

2.2 The Need for a Product Database

Having a benchmark of the current MTDB built on the same schema structure with persistent storage of topology in a separate database instance provides a stable and flexible environment for product creation. However, these measures alone are not enough to provide the efficiency and throughput needed to support timely creation of the vast array of address, spatial, geographic, and cartographic products required for the decennial census and other Census Bureau ongoing programs. Therefore, a Product Database is built from the Benchmark. A PDB is needed to efficiently support product creation because:

- Product creation from current MTDB is problematic due to the large number of database joins required related to the normalized structure.
- The transaction database avoids storing redundant data, which is efficient for updating, but leads to duplicative effort during the creation of a massive array of products. The same data must be calculated over and over if the transaction database is the source of product creation.

- There is difficulty creating products using the transaction database that could be constantly changing. There would not be one "data as of" date for a series of products.

Most products are created wholly or partially from the PDB. Due to the cyclical nature of the update process and to effectively support the continuous flow-based processing to meet product delivery requirements, a Benchmarking Process and Product Database design became essential in the redesigned MAF/TIGER system environment.

3. BENCHMARKING PROCESS

The benchmarking process results in the creation of the MTDB Benchmark and subsequently the Product Database. It effectively constitutes the second and third parts of the multi-level geospatial database model. The process involves the transfer of data from the current MTDB on one server to a database on another server applying a series of edits to ensure data integrity. The benchmarking process encompasses the update cycles of the MTDB, pre-benchmarking, and benchmark creation steps leading to the Product Database creation. The benchmarking process that forms the basis of the multi-level MAF/TIGER system geospatial database model is depicted in Figure 1.

On a flow basis, an individual county is extracted out of the current national MTDB, transferred to a Product Database server, merged into a target benchmark schema and then loaded into the Product Database schema to support product creation. For a successful workflow management, GEO has determined county/equivalent area as the base unit of work for the benchmarking operation because the county unit is well-defined and quite stable. Stable means that the universe of units and their spatial extent does not change often. The base unit of work is the base unit of control for the benchmarking process.

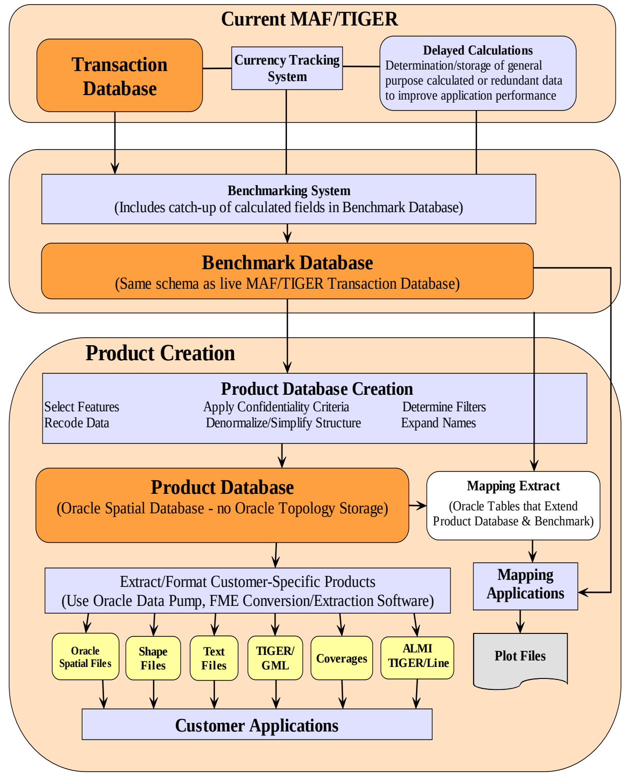

Figure 1. MAF/TIGER System Multi-level Geospatial Database Design Model

3.1 MTDB Updates

Each benchmarking cycle begins with a number of MTDB updates to the address, geographic, and spatial data resulting from field operations and other sources. During the decade, the Census Bureau conducts a number of surveys such as the American Community Survey, Boundary and Annexation Survey, and current surveys that require the Geography Division to maintain the most up to date spatial, geographic, and address data to support the Bureau's mission. The MAF is also updated twice a year using the US Postal Service DSF as a source. The MTDB data updates occur on a flow basis by county. There are huge amounts of data to process. For example, during the 2010 Census Address Canvassing operation, over 150 million addresses were updated in MTDB over a period of 2 months. The geographic update and verification programs, as well as updates of street feature, coordinate, and address information must all complete for the Bureau to meet the collection, tabulation, and product dissemination needs. For the spatial updates supporting the 2010 Census data collection and address canvassing operations, over 730,920 Assignment Areas, 5668 Crew Leader Districts, and 748 Field Operations Supervisor Districts were delineated in the MTDB.

3.2 Pre-Benchmarking

The pre-benchmarking step includes a series of processes to ensure integrity of the source data prior to the extract process. The release criteria for the readiness of a subject county to start the benchmarking process are established. The criteria include adjacency checks to ensure integrity of the entire benchmark. To get the boundaries of the counties to spatially align without gaps in the benchmark, the subject county is spatially locked once all the MTDB updates have occurred for the county and its adjacent counties. The subject county proceeds through a series of data edits, including legal values checks, business rule checks, geometry, and topology validation prior to being extracted and merged into the benchmark schema. The process utilizes the concept of delayed calculations where computations of certain attributes are deferred to support performance during real time update operations in the current MTDB.

3.3 Concept of Delayed Calculations

Delayed calculation is a concept used in the MAF/TIGER system environment to support performance during transaction processing in the current MTDB. The MTDB includes attributes such as area measurement, internal points, boundary edges, and geometric representation that are independent of (and redundant with) the underlying topology. These attributes are derived from other columns. To calculate their values on the fly during update operations would cause enormous slowdown during processing. Therefore, the calculations of these attributes are delayed and computed during the benchmarking process. Employing the concept of delayed calculations alleviates performance hindrances during transaction processing to the current MTDB. Computing this data prior to the extract of the county from the current MTDB during the benchmarking process ensures the most up to date values are derived and copied to the benchmark. These calculated values facilitate product creation, but are not required for maintenance activities in the current MTDB.

3.4 Benchmark Database

The Benchmark Database resides on separate hardware and database servers from the current MTDB. This configuration allows the Census Bureau to realize some of the benefits associated with distributed processing such as improved performance and throughput. The distributed processing environment also facilitates the ongoing transaction processing in the current MTDB while product generation can continue. The Benchmark is a replica of the current MTDB on a separate server. It provides static reference copies of the MAF/TIGER data. Each static copy is made to meet the data requirements of the products it serves. Thus, the MAF/TIGER Benchmark is the supplier of logically consistent and complete MAF and TIGER data. The MAF/TIGER Benchmark serves the MAF/TIGER product applications that support Census Bureau programs. These programs include:

- Decennial Census
- Economic Census
- American Community Survey
- Current Surveys
- Boundary and Annexation Survey
- Special Census Program
- Population Estimates and Projections
- Administrative Records Program
- Public TIGER/Line Shape Files

3.5 Benchmark Creation

The benchmark creation step includes the data extract, data pump, merge, data quality checks, and data integrity checks. The target benchmark schema can be sub-national as well as national. The sub-national benchmarks along with their corresponding Product Database schemas are typically state-based. The intent for the sub-national schemas is to facilitate flow processing. The national schema allows for seamlessness to create products for the entities that cross sub-national partitions, such as American Indian Areas or metropolitan and micropolitan statistical areas. The edits performed on the benchmark include legal values checks, universe checks, and validation of the geometry and topology. Some delayed calculations are also performed in the benchmark to ensure that the values accurately reflect the spatial extent of the features loaded in the benchmark. The Product Database load begins only after successful completion of the quality checks on the benchmark for the subject county.

4. PRODUCT DATABASE IMPLEMENTATION

The creation of the Product Database represents the last part of the multi-level model. The design approach adopted for the PDB includes database denormalization, use of Oracle Spatial geometry storage exclusive of the Oracle Topology Model, storage of product-specific calculations, and considerations for industry geospatial COTS tools. The goal for the PDB implementation is to create an infrastructure to easily manipulate data into more customer-oriented form with minimal processing and optimal performance. The PDB can be considered in some respects as a data warehouse because of the large volume of data to maintain and the need for data mining tools to manipulate the data. For the 2010 Census data collection products, the size of the Product Database and Benchmark created from the current MTDB updates which include the address canvassing operation, boundary, feature, and geographic area updates was 3.5 terabytes.

The PDB is loaded with spatial, geographic, and address data. Quality assurance is integrated into the PDB load to ensure that integrity of the data content is maintained. Some of the quality checks include validation of the geometry column, area measurements, required feature relationships, and coverage checks to assure there are no gaps or overlaps in those features.

A mapping extract database is also created to provide support for mapping operations. This cartographic database extends the benchmark and PDB by adding tables specific to mapping application needs.

4.1 Use of Denormalization

The current MTDB logical design uses normalization to efficiently support the update operations and maintenance of business rules for data integrity. However, to attain better throughput and performance, denormalization is used in the design and implementation of the PDB. For major applications and large data warehousing type databases, it may be necessary to deviate from normalization where needed to support optimal mining of large volumes of data (Koch and Looney, 1997). Thus the PDB uses denormalization to simplify the current MTDB database structure and allows combination of data attributes to be stored directly in the database. Examples of denormalized data stored in the PDB include:

- Combination of the feature name table with elemental feature table to allow for feature names to appear on the spatial features tables.
- Combination of the MAFUNIT, ADDRESS, FEATNAME and STRUCT tables to allow for output of a consolidated address record that includes a single preferred mailing address, a single preferred location address, and a preferred structure coordinate.
- Combination of the FACE (area) and related tables from MTDB into a single PDFACE table
- Combination of rail, road, and hydrography edge attributes with the EDGE table to form the PDEDGE table.

Employing the concept of denormalization in the design of the PDB has greatly eased the manipulation of data into more customer friendly form.

4.2 Spatial Data Storage

Unlike the current MTDB, the design of the PDB uses Oracle Spatial without the Oracle Topology Data Model. The PDB uses Oracle Spatial and the sdo_geometry data type. The PDB relies primarily on the Oracle Spatial data storage. The reasons for this approach include overhead involved in managing and copying data in the Topology Data Model. In the current MTDB, the topological structure, including the nodes, edges, areas and their relationships is over 4 billion records. Updates to any of the underlying topological tables (node, edge, area) requires complex processing which can be extremely time consuming.

Most of the advantages of using Oracle Topology Data Model in the current MTDB relate to data updates and maintenance, which are relevant in MTDB but not in the PDB. In the PDB, there are no ongoing updates once the database is fully loaded through the benchmarking process. In the current MTDB, Oracle Topology Data Model stores the coordinates only for the topological primitives, i.e., nodes, edges, and areas. Features are then defined by linking them to one or more primitives. In the PDB, Oracle Spatial uses the special Oracle defined data type, 'sdo_geometry' which stores all the coordinates in order for any spatial feature, whether it is a point, line or area. The sdo_geometry attribute is calculated as part of the delayed calculation process and it is stored in the PDB to readily support data retrieval and spatial data analysis.

4.3 Storage of Derived Data

The PDB was designed to store derived data attributes in the database. The derived data pertains to attributes that are calculated for selected features based on certain filters provided by customers. In some cases, data elements are re-coded, confidentiality criteria are applied, or feature names are extended in the PDB. Some examples of derived data include decennial address filters, American Community Survey (ACS) address filters, MAF/TIGER Feature Classification Code (MTFCC), expansion of feature name types and directions, calculation of housing unit counts, and designation of primary versus alternate name. Storing product-specific calculations facilitates large volumes of data retrieval and product creation and avoids massive recalculations by application software. The values are calculated once, stored and made available to all product creation software.

4.4 Geospatial Data Visualization COTS Tools Consideratons

Another aspect of the PDB design is the flexibility to allow geospatial COTS tools to interface with the database structure. In particular, the use of Oracle Spatial exclusive of the Oracle Topology Model allows the PDB to be registered as an ESRI Geodatabase. The registration of the PDB as an ESRI Geodatabase makes data available through ESRI tools which are widely used in the GIS industry to support spatial data visualization and analysis. To support the Geodatabase registration, here are some of the design considerations made:

- Separation of the MTDB Elemental Feature tables into three tables respectively in the PDB: point, areal and linear features.
- Conversion of some point geometry fields from SDO_GEOMETRY to text (latitude/longitude) in order to avoid multiple SDO_GEOMETRY columns in a single table.
- Removal of some user-defined data types such as collections.
- Specification of lengths for NUMBER fields.
- Addition of a GEOID column to the geographic area tables.

The capability to register the PDB as a geodatabase has served the Census Bureau's needs well. For delineation of collection geography to support the 2010 Census, an ESRI ArcGIS-based application, the Geographic Areas Analysis and Delineation System (GAADS) was developed using the MAF/TIGER PDB registered as an ESRI geodatabase as the data source. GAADS is a geospatial database application that provides the capabilities for analyzing and creating Census collection geographic entities. The GAADS has been used by the Field Division (FLD) and GEO for various operations to delineate the following collection geography entities:

- Local Census Office (LCO)/Early Local Census Office (ELCO)

- Type of Enumeration Area (TEA)
- Assignment Area (AA)
- Crew Leader District (CLD)
- Field Operations Supervisor District (FOSD)

GAADS was developed in house using the PDB generated from the Initial Universe Control and Management (UC&M) benchmark. The production and reference data layers are extracted from the benchmark of the MTDB. The extracted data layers are loaded into the PDB, an Oracle Spatial database, registered with ArcSDE as a geodatabase and made available to authorized application users (Wang et. al., 2009).

5. PRODUCT CREATION – GEOGRAPHIC, SPATIAL, CARTOGRAPHIC, AND ADDRESS DATA

The quality, on-time creation and delivery of geographic, spatial, cartographic, and address data products constitute the ultimate end results of the redesigned MAF/TIGER system multi-level geospatial database model. The current MTDB, benchmark and PDB along with the benchmarking process provide the infrastructure and architectural support for product creation. The products created include extracts for:

- Spatial Data – ESRI Shapefiles, Oracle dump files
- Geographic Data - Geographic Reference files, SAS datasets, Text Files
- Address Data - MAF Extracts
- Cartographic and Map Products – PDF files

The extracts and the formats are created to meet specific customer product needs. Use of Oracle Data Pump, Safe Software's FME tool, and other conversion, extraction or transformation software are used to mine the data to meet product requirements. The types of file extract products include Oracle dump files, ESRI Shapefiles, SAS Datasets, Text Files, TIGER/GML, ESRI Coverages, and Map files.

For the 2010 Census field operations, GEO created and delivered over 17 million map sheets, 10 Geographic Reference Files (GRFs) containing 6.7 million collection blocks, and 8 full MAF Extracts of over 144 million addresses to support the operation. These products all utilized the MAF/TIGER PDB. The delivery mechanism utilized for the products was dependant on the customer processing environment. For decennial processing support systems, Oracle Advance Queues package is used to facilitate product deliveries. To support the public geospatial data needs, the TIGER/Line shapefiles are made available online for interested parties to acquire. In recent years, the Census Bureau's Partnership Programs have also extended their data exchange programs with the local, state, and federal partners by making data products accessible through a secured internet site.

6. PROCESSING CHALLENGES

Implementing the multi-level geospatial database model does provide enormous benefits and improved efficiency in meeting the Census Bureau's mission and data product needs. However, such a model also does have its own challenges. The major challenge is the size of the databases and the volume of processing needed to complete in order to meet product delivery deadlines. The need for optimal performance with the large volume of data to process is a constant challenge. Highly tuned database queries and monitoring are critical. Optimal configuration of the servers, network connectivity, and additional hardware are required at times to support the processing environments.

In addition, operating in the Oracle Spatial and Topology Data Model environment has proven benefits but challenges as well. The Census Bureau for many years has operated using its own proprietary TIGER system and custom code. In that respect, there are bound to be functionalities the Census Bureau has become accustomed to but are not readily available in a COTS product. One challenge the Bureau has is that in the proprietary TIGER database, coordinates were stored as integers with 6 implied decimals. However, with Oracle Topology Data Model, coordinates are stored as real numbers (Oracle number data type) and treated in the Topology Manager JAVA Application Programming Interface (API) with double precision. Occasionally, this additional precision can create issues with data integrity and add to the complexity of snapping and comparison algorithms. To circumvent the situation, the Census Bureau developed a custom API to truncate coordinates to 6 decimal places before writing to the database.

To further complicate the issue, Topology Data Model stores the end points of an Edge twice - once on either end of the geometry for the edge, and also on the node records corresponding to the ends of the edge. Although the coordinates for the nodes are guaranteed under the Topology Model to be identical to the edge endpoints, they may appear to be slightly different depending on how many decimal places are considered and on the environment in which the coordinates are examined.

Another challenge involves registration of the PDB as an ESRI Geodatabase, which adds a mandatory column to the PDB tables that serves as a record identifier for the ArcSDE Engine. Making the column mandatory poses a slight inconvenience in the event that a reload of parts of the PDB is needed if the Geodatabase has already been registered. In such circumstances, the Geodatabase will have to be unregistered first before the reload. The Geodatabase registration process also imposes some overhead to the overall performance of the PDB due to the increasing number of objects stored in the Oracle data dictionary.

7. KEYS TO SUCCESS

A number of best practices and process improvements in GEO have led to the successful implementation and maintenance of the MAF/TIGER system and product creation. Extensive automated edits are built into the benchmarking process. Edits are performed during updates, pre-benchmarking, and benchmarking steps. Well-defined quality checks of the PDB load and product creation processes exist. Subject matter experts within GEO perform user acceptance testing for sample products before full production commences. An established change control and configuration management process is in place to assure that changes are well documented, reviewed, analysed, approved, and implemented with well known impacts and constraints. A dedicated test suite environment is utilized to ensure that all new patches to fix any spatial and topological issues identified and accepted by Oracle Corporation are first executed with a set of pre-defined test cases.

In addition, a well-defined workflow process and delivery mechanism with automated submission allows processing to continue 24 hours/7 days per week. The process accounts for the workflow management and deals with issues pertaining to adjacencies which are very critical to ensure consistency and data integrity of all the systems. To support the continuous process improvement efforts, each new benchmarking cycle feeds off the lessons learned from the previous cycle. A repository of issues is maintained during the benchmarking cycle to track issues and resolutions. The data forms the basis for the lessons learned sessions so that the subsequent cycles will avoid similar issues.

8. CONCLUSION

The multi-level geospatial database design model adopted in the redesigned MAF/TIGER system environment has served the Census Bureau and the public's needs for address, geographic, and spatial data products quite well despite the challenges. The flow-based processing required for supporting the cyclical nature of the Bureau's update process in order to meet product deadlines necessitated the implementation of a system to allow processing to occur in a distributed environment: a real-time transaction-based environment and a static environment for product generation. The creation of the benchmarking process and the PDB has allowed the Census Bureau to establish an infrastructure for a mission critical national geospatial data source that is open, based on industry standards, and allows home grown applications to be replaced by Commercial Off-The-Shelf software (COTS).

The PDB provides flexibility. The data structure allows for ease of integration with varied spatial and address data visualization and analysis tools. It opens the door for Web-based technology development and tools to manipulate and access the data using Rapid Application Development and Agile software development methodologies. The PDB can support emerging technologies such as Enterprise GIS, Cloud Computing, Open GIS Consortium Web Feature Service (WFS), and Web Map Service (WMS). Innovative solutions in the area of GIS can be developed in the future to mine the PDB to readily serve TIGER data over the World Wide Web, all due to the intelligent design employed in the implementation of the benchmarking process and the creation of the PDB in the redesigned MAF/TIGER system environment.

REFERENCES

Egenhofer, Max J., Andrew U. Frank, and Jeffrey P. Jackson, 1989, A Topological Data Model for Spatial Databases. Design and Implementation of Large Spatial, Lecture Notes in Computer Science, pp. 409, 271-286.

Galdi, David E., 2005, Spatial Data Storage and Topology in the Redesigned MAF/TIGER System, U.S. Census Bureau, Washington, DC. http://www.census.gov/geo/mtep_obj2/topo_and_data_stor.html (accessed 16 Nov. 2009).

Hoel, Erik, Sudhakar Menon, and Scott Morehouse, 2003, Building Robust Topologies. In Advances in Spatial and Temporal Databases Proceedings of the 8th International Symposium on Spatial and Temporal Databases. SSTD 2003. Santorini Island, Greece, Springer-Verlag Lecture Notes in Computer Science, pp. 2750.

Kainz, Wolfgang, 2004, Geographic Information Science, Version 2.0, http://www.geografie.webzdarma.cz/GIS-skriptum.pdf (accessed 1 April, 2010).

Koch George, Loney Kevin, 1997. *Oracle Press Oracle8: The Complete Reference*. CA, USA pp. 30, 873

Oracle Corporation, 2004, Oracle® Spatial Option and Oracle® Locator Datasheet - Location Features in Oracle Database 10g, http://www.oracle.com/technology/products/spatial/htdocs/data_sheet_9i/10g_spatial_locator_ds.html (accessed 1 April. 2010).

Ramage, Steven and Peter Woodsford, 2002. The Benefits of Topology in the Database, http://spatialnews.geocomm.com/features/laserscan2/ (accessed 1 April. 2010).

Trainor, Timothy, 2003, U.S. Census Bureau Geographic Support: A Response to Changing Technology and Improved Data. Cartography and Geographic Information Science, 30(2), pp. 217-223.

Wang, Y., Anacker, S. and B. Kammela. 2009. The Geographic Area Analysis and Delineation System for 2010 Decennial Census. U.S. Census Bureau, Washington, DC, USA.