Implementing Automation to Transition from Librarian to a Versioned Geodatabase

Author(s)
Jonathan Pickus, Jason Schwartz, Keath Long

Abstract
The Las Vegas Valley Water District AM/FM/GIS Division had utilized a Librarian based GIS database to manage its infrastructure data. Over the years, multiple processes were implemented to enhance as-builting and GIS database maintenance activities necessary to keep pace with the Las Vegas Valley's skyrocketing growth rate. Successfully transitioning to the Spatial Database Engine (SDE) data model meant developing new automation methods that complemented the Division's successful production methods while leveraging SDE's strengths. This included the ability to support both traditional "asbuilting" using CAD and GIS data management using advanced SDE functionality such as versioning, conflict resolution and network validation. Key automation developed included workflow applications supporting SDE data extraction and an automated project-based data loader tool that supports updates to a fully versioned Geodatabase in a multi-user environment.

Introduction

The transition from the ESRI Librarian GIS data management system and associated "coverage" data model to a Spatial Database Engine (SDE) geodatabase can be challenging due to differences in the underlying data model and user application interface. The Las Vegas Valley Water District (i.e. District) AM/FM/GIS Division (i.e. Division) undertook this implementation to gain benefits in GIS database quality assurance and control, network modeling, seamless data structure and enhanced environment for user access and editing.

The AM/FM/GIS Division manages an as-built GIS database that describes the configuration and characteristics of the District's water transmission and distribution network. This system exceeds 3,500 miles of pipeline, 80,000 valves, 9,500 fittings, 26,000 fire hydrants along with reservoirs, pumps stations, wells and other system appurtenances.

The District is the largest water purveyor in southern Nevada, serving over 290,000 accounts, one million people and a service area in excess of 500 square miles. The customer and infrastructure growth rate has averaged 6% growth over the last 12 years and is continuing with over 6,500 new accounts being added to the system monthly. This dynamic and continuing growth has required the AM/FM/GIS Division to institute numerous automation processes to ensure that the database reflects as-built conditions and to provide that information to District staff in a timely fashion.

The use of automated processes and production techniques weighed heavily on how best to transition from Librarian to an SDE Geodatabase. The implementation objectives included the need to maintain production throughout the transition and incorporate systems that are highly productive and efficient. The Division had benefited from using both GIS and CAD software for maintaining GIS content and this strategy became a central component in the deployment of SDE. This paper discusses the approach and technology employed to successfully implement a fully versioned and transactional SDE Geodatabase in a heterogeneous and open environment.

Legacy System

The Division managed its geographic data using a typical Librarian database structure that tiled geographic layers based on square mile sections. Several applications were deployed that automated tile management, map production and GIS data editing. Feature attributes were maintained within a set of Oracle tables based on a common primary key. Layers included pipelines, valves, hydrants, fittings, meters, cathodic stations, wells, reservoirs, pump stations and annotations.

Extensions were made to this system to support both GIS and CAD systems (Figure 1). As-builting tasks were performed by a team of technicians using CAD software, while map production tasks, analysis and spatial data management were addressed using GIS software. Systems were deployed that supported GIS data editing within the CAD environment, including modification, insertion of spatial features back into Librarian and attributes into respective Oracle tables. Managing these updates from the CAD environment is a key difference from most other Librarian implementations and significantly impacted the methodology used in the transition to the SDE geodatabase.

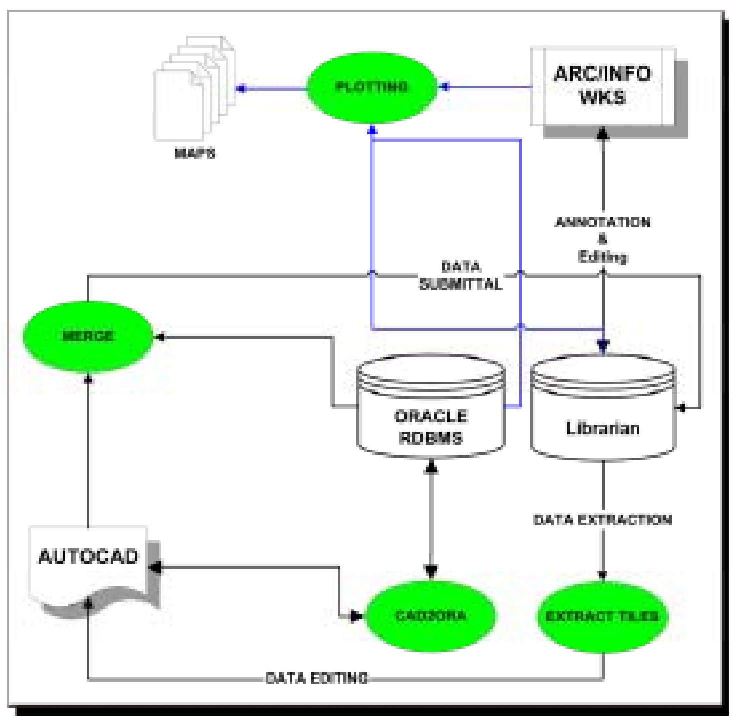

Figure 1 - AM/FM/GIS used Librarian to manage its GIS data. Extensions to this system were developed to support as-builting by CAD technicians.

Implementation Methodology & Approach

The Division has benefited from a high level of automation and efficiency gained from integrating GIS and CAD technology and staff to manage the central GIS database. The advantages of the SDE Geodatabase were also recognized, especially in the areas of network modeling, validation and support for long-transactions. Four "implementation" objectives were identified that supported the benefits of an "open" environment integrated within the SDE geodatabase system model. These were:

- Goal 1: Minimize loss of production throughout the implementation and transition to the SDE geodatabase.
- Goal 2: Maintain a heterogeneous environment that continues to allow both GIS & CAD teams to maintain the GIS database.
- Goal 3: Incorporate automation to maintain operational efficiency.
- Goal 4: Develop a system that supports a fully versioned transitional networked geodatabase.

A technical strategy was then developed, first by identifying technical processes that needed to change versus those that could remain intact throughout the transition. It was decided that the implementation would concentrate on the SDE database design, data flow between the GIS and CAD technical teams, versioning and network management. Map production would be left as is for the time being and the use of ArcMap for editing would be limited until after the final implementation. The map production systems, based on Arcplot and workstation Arc/Info, were highly automated and could be readily adapted to access SDE. The decision to continue to use Arcplot for plotting also postponed the coverage annotation conversion problem. The Division actively manages over 800,000 annotations and initial tests indicated performance problems with this amount of data being loaded and managed in the District's SDE database.

Conversion Milestones

The conversion schedule was divided into six milestones (Table 1 & Figure 2). The first steps implemented the SDE database and maintained and built systems that supported both Librarian and SDE in parallel. This involved building applications that automated data loading, versioning and reconciliation into the SDE database. The second milestone instituted methods that continued to support as-builting and GIS content management by both GIS and CAD teams. Converting the map production systems to access SDE and produce the District's standard map products was the objective of the third milestone. The final conversion steps were validating the SDE attribute and geometry data quality and switching the primary data entry point from Oracle corporate tables to SDE business tables.

Table 1 *Implementation Milestones*
1. Design SDE database
2. Implement SDE as a parallel systems (Librarian & SDE)
3. Convert editing environment to support GIS & CAD Teams
4. Convert plotting environment to access SDE
5. Validate SDE database against Librarian
6. Switch to SDE business tables as primary attribute data source

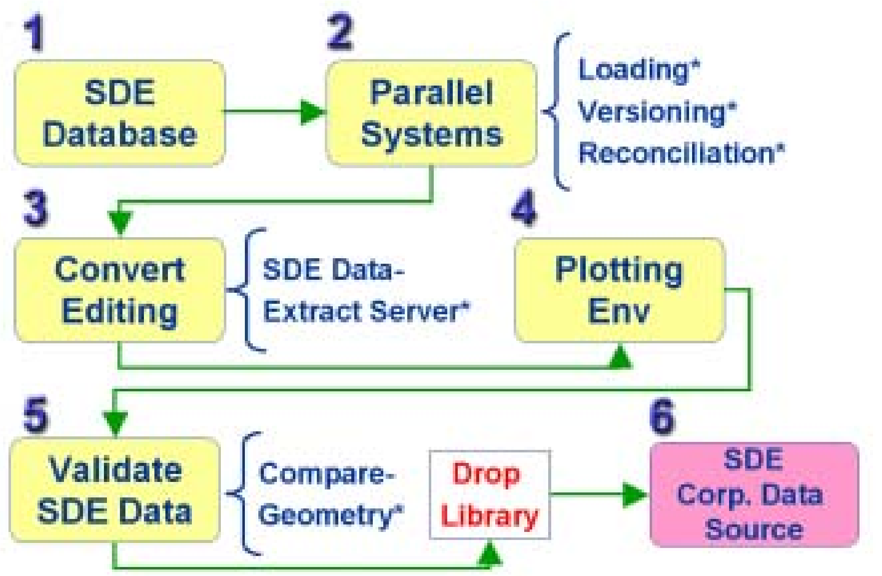

Figure 2 – The technical conversion followed these six phases.

Setting up Parallel Systems

Concurrent SDE and Librarian systems were developed and maintained respectively to test and evaluate the conversion methodology, validate data integrity and associated applications and minimize impacts to production systems. The steps involved in setting up the SDE system included designing the database and developing applications that automated data loading, versioning and reconciliation.

The geodatabase design team modified the ESRI supplied water utilities data model to meet District business requirements. Staff used Visio CASE tools to generate the model in Oracle and provided the automation needed to reiterate this process as the model was tested against functional and performance requirements. These initial tests looked at drawing and editing performance, data loading, domains, relationship classes and the network.

Automated Data Loading

Since a majority of the workload flowing into the Division is on a per-project basis, the decision was made to manage data on that same basis. A single project involves the water distribution system for a subdivision, building or a major construction contract. The Division handles between 50 and 100 projects and work orders per month. Each of these projects needs to be "as-built" and loaded into the geodatabase in a timely fashion. It became clear that manual "loading" of data into SDE on a repeated basis would negatively impact production and would be prone to human error especially when considering controlling versions and transaction management. A key application was developed to completely automate data loading. "ProjectLoader" was designed to automatically load data into the proper feature classes, perform version management and perform a host of quality assurance and control tasks (Figure 3). The system also boasts a modular design that supports newly added feature classes. This is due to its tight integration with Arc Catalog and other configuration specifications stored in a RDBMS.

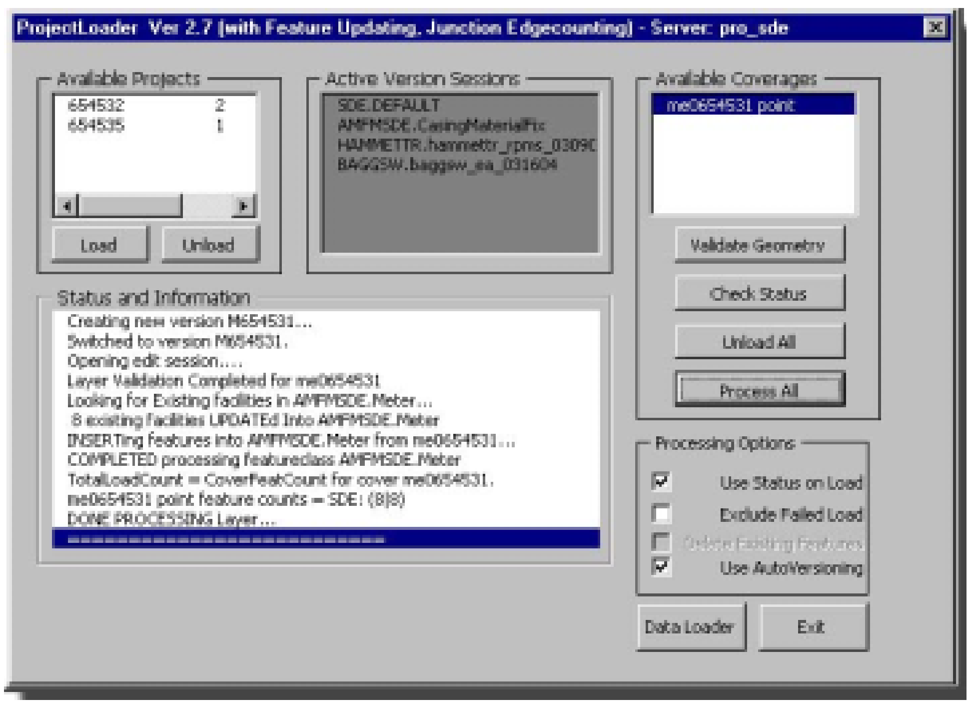

Figure 3 - ProjectLoader automates the loading of data into the SDE geodatabase. It supports versioning, performs various quality assurance and control steps and supports network connectivity of modified features.

ProjectLoader performs the following functions:

- Identifies and loads Arc/Info coverage "project" data into ArcMap.
- Automatically assigns and tracks versions.

- Identifies which data is to be loaded into which feature class.
- Performs attribute data validation including checks for valid domains and subclasses.
- Adds new features and replaces the geometry of modified features to preserve network connectivity, conflict resolution and relationship classes.
- Generates 4-way junctions needed to identify pipeline intersections on maps.
- Checks for load success.
- Validates geometry by comparing source data against loaded features.
- Logs all activity in a database.

Feature Geometry Replacement

Feature replacement is done when an SDE feature has been exported and modified by an “external” team and then sent back through ProjectLoader. This capability is key to allowing other teams to continue to support the as-builting of GIS content. It also reduced the number of work processes that needed to be re-engineered and instead allowed the Division to concentrate on the data flow component between the GIS and CAD teams.

Feature replacement solved a number of issues with loading “external” data into a geodatabase. This included network connectivity, maintaining relationships classes and supporting conflict detection.

Network connectivity is maintained for modified features when its altered geometry is replaced. This is done automatically through the implementation of the “iNetworkFeature” and “iRow” interfaces in ArcObjects.

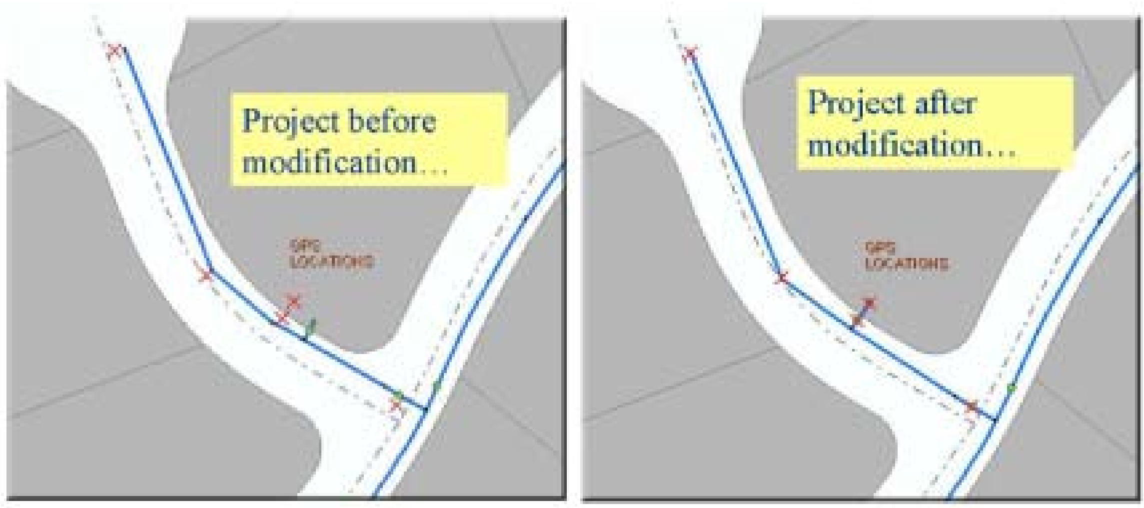

Figure 4 - Feature geometry replacement functions within ProjectLoader allow features to be modified outside of ArcGIS while still maintaining network connectivity.

Auto Reconcile and Posting

The number of "projects" and versions processed by the Division on a regular basis necessitated automation of version reconciliation and posting to the geodatabase. "AutoReconcile" is an application based on ESRI "BatchReconcile" sample application. "AutoReconcile" was customized from "BatchReconcile" to meet specific District requirements (Figure 5). This included automated database connections, working with the Division's particular version naming conventions and performing additional quality assurance checks on source data. AutoReconcile also helps maintain accurate tracking of which features have been exported from the SDE geodatabase for as-builting by external teams.

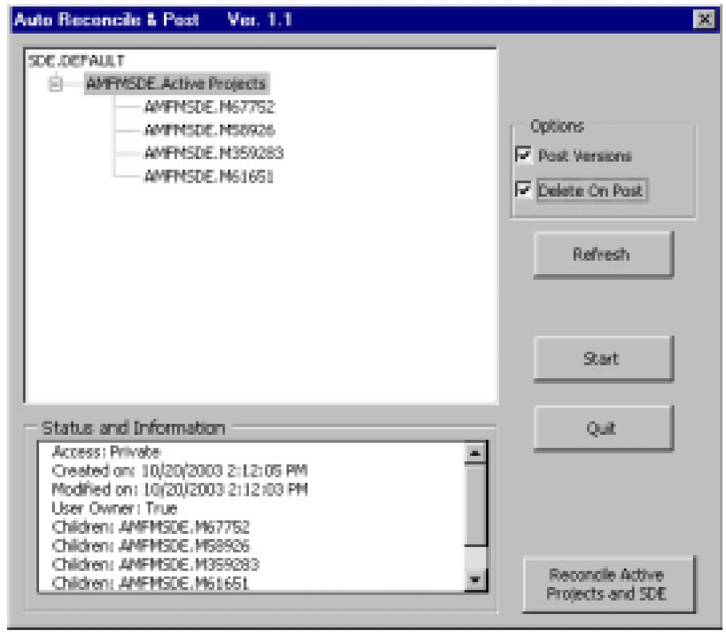

Figure 5 - Auto Reconcile & Post automates version reconciliation and posting to the SDE geodatabase.

Open "Editing" Environment

The Division has benefited from having both GIS and CAD teams perform GIS data conversion tasks. This conversion effort engineered an environment where "asbuilting" tasks can be performed by technicians using native CAD software in a similar fashion as before.

The "old" method involved a set of processes that extracted a tile from Librarian, converted the GIS features to CAD blocks and entities and tracked the feature's primary keys in a database. GIS and CAD software were made "aware" of these features, essentially "locking" these features from further manipulation until they were reinserted back into the Library.

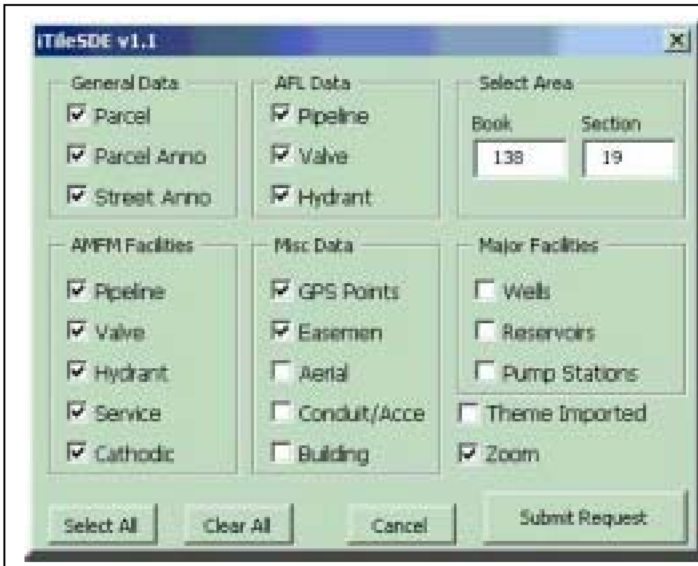

Figure 6 - iTileSDE brokers requests from CAD technicians to deliver SDE data as formatted dwg layers.

This set of processes was invoked client-side from the CAD technician, which delivered data as a “dwg” file directly into their software environment. Applications were built within the CAD software that allowed the technician to update attributes directly into the Oracle table that housed the GIS feature’s attributes.

The reengineering of this process meant developing a similar system that worked with the SDE geodatabase. The resulting application “iTileSDE” was designed as a “service” that brokered requests from CAD technicians to export feature classes from SDE into CAD (Figure 6). The service implements workstation Arc/Info “LayerExport” and renders the output into CAD blocks and entities organized into standard layering. “LayerExport” proved much more efficient that implementing the “iExportOperator” interface within ArcObjects.

Feature Locking and Versioning

Feature locking is a process where SDE features, that have been exported for editing, are tracked by the District’s GIS editing work processes. This ensures that staff does not perform simultaneous edits and that work is not lost due to SDE version conflicts.

SDE’s standard functionality supports backend conflict resolution through the use of versions. Conflicts occur when edits are made to the same features within different versions. Comparisons are made to the “A” (adds) and “D” (delete) tables for both geometry and feature attributes that have undergone edit changes. SDE conflict detection automatically identifies and highlights the differences between versions during reconciliation (Figure 7). This process

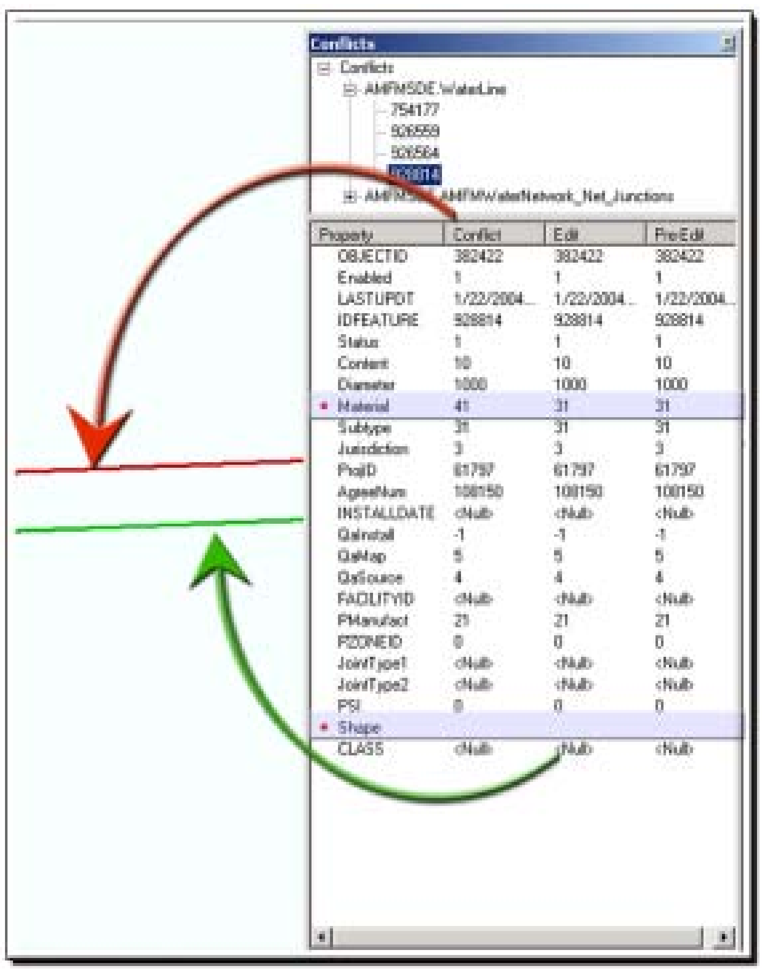

Figure 7 - SDE's conflict resolution identifies differences between features. Users must select the edit state to keep.

requires a user to choose which edit state to keep, and which to discard.

Versioning provides back-end conflict resolution. Its benefits are that it is integrated within SDE. However, it does require manual assessment and work can be lost when one or more feature edits are discarded due to a version conflict.

Feature locking provides front-end resolution. It prevents loss of work by avoiding conflicts before they occur. The other benefit is that it does not require any manual oversight. However, the system can delay work by forcing staff to wait for edits to be completed (i.e. a feature they wish to edit has been locked for another editing session). It is also a custom environment that adds more complexity to the GIS and CAD editing environment. Feature locking continues to be supported in the Division's SDE editing environment to reduce conflicts before they can occur and to streamline the GIS data loading process.

Plotting Environment

The Division produced over 11,500 standard maps in 2003 supporting the planning, approval and maintenance of the District's water distribution system. The plotting system is based on a set of Arc Macro Language (Aml) routines running ArcPlot. Maplex is also used to generate annotations and maps. The system proved highly efficient and productive.

ArcGIS introduced a very different plotting model and interface which raised a significant conversion challenges. There are also some compatibility issues to contend with, mainly in the differences between coverage annotation and geodatabase annotation. For example, ArcPlot cannot access geodatabase annotation for plotting, and ArcGIS does not support editing of any coverage features.

The Division had to decide whether to redesign and deploy an entirely new plotting system while simultaneously re-engineering and implementing the SDE database conversion and workflow processes. Several tests were conducted to better understand the task at hand including performance and compatibility, which resulted in the following conclusions:

1. The ArcMap VBA interface provides a powerful tool to automate standard plotting and produce highly professional map output. The Division was able to produce several plotting tools in a relatively short time frame that demonstrated the flexibility of the ArcMap interface.
2. Performance suffered when creating plots in ArcMap that relied upon joins and relates to external tables. This poised a significant obstacle as most of the Division's standard maps relied upon corporate engineering and business database systems to theme features and control map output.
3. The Division actively manages over 800,000 annotations with various attributes. A conversion program called "Anno2Geo" was written to convert these annotations to geodatabase format while preserving key attributes. However,

drawing performance suffered for some annotation layers, concluding that more analysis was needed to understand SDE tuning characteristics for geodatabase annotation.

4. Maplex was not yet fully compatible with SDE.
5. It would be much easier to convert the plotting environment once the SDE environment was in production.
6. ArcPlot could easily be converted to read SDE layers without significant modifications to the plotting methodology (Figure 8).

It was decided, based on these conclusions, to delay the plotting conversion and continue to use the Division's Arcplot-based plotting systems. This simplified the overall transition to SDE and significantly reduced impacts to production.

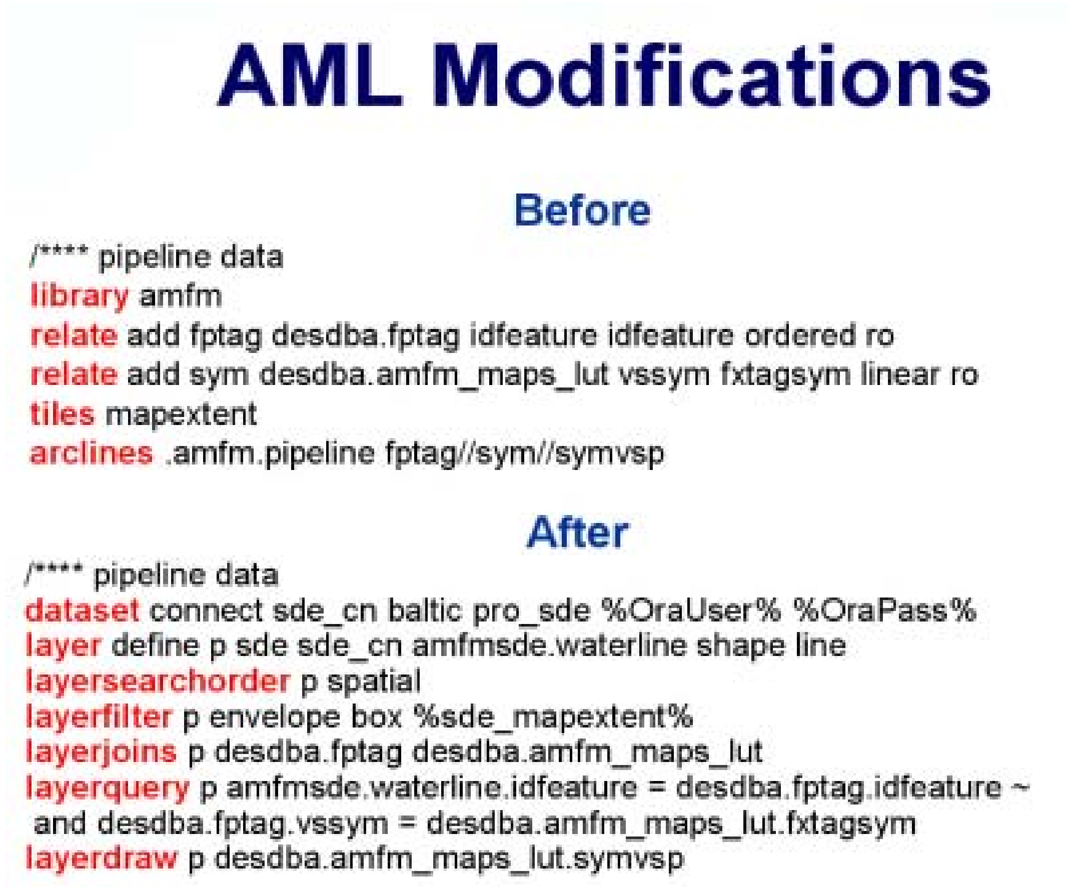

Figure 8 - The Division's AML/ArcPlot based map production system was converted to access SDE by modifying the underlying AML code. Delaying the re-engineering of the plotting system simplified the implementation to SDE and maintained map-production throughout the transition.

Data Validation

The validation of SDE Data was performed 90 days after the implementation of the SDE-based workflow systems, illustrated on page 4, (i.e. phase 2 and 3 of the technical conversion). Entire layer comparisons where made against the Librarian database, which was being maintained using the Division's existing coverage-based production systems. Data assessment was also performed against specific projects to further validate the automated SDE loader applications (i.e. ProjectLoader). Having systems operate in parallel (i.e. Librarian and SDE) provided a true benchmark for validating SDE and the various applications used to manage the SDE data.

Validating Feature Class Attributes

A method was required to compare attributes between the new SDE feature classes and the original Librarian layers with Oracle attributes. Structure query language and "multiversion" views where used to compare and identify differences using foreign keys to establish joins between tables. "Multiversion" views, created using ESRI procedures, are views over the SDE tables for a feature class that resolves rows added and deleted for a version. Several attribute differences and missing features where identified using this approach. A root-cause analysis revealed both process flow and application logic inconsistencies were responsible for the differences found. As a result, the data were corrected and the processes improved.

Validating Feature Geometry

Comparing feature geometry between SDE and Librarian layers was used to validate the initial SDE data load and maintenance procedures and ensure that the SDE data truly reflected the spatial information managed in Librarian. An application named "SDEChkProj" was written to perform geometry comparisons. Geometry comparisons are done against points (i.e. valves, hydrants, fittings, etc) and lines (i.e. pipelines). Geometry comparisons were done using topological operators and spatial filters to identify Librarian and SDE features within a prescribed tolerance (Figures 9 and10). The application was run against all features in SDE and Librarian and has been incorporated into "ProjectLoader" to ensure accuracy of all SDE data loads.

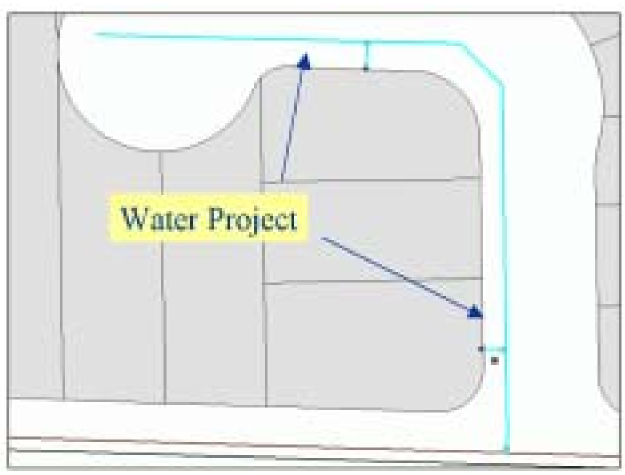

Figure 9 - SDEChkProj was used to compare geometry loaded into SDE against the source data to validate processes. Figure 10 demonstrate the approach.

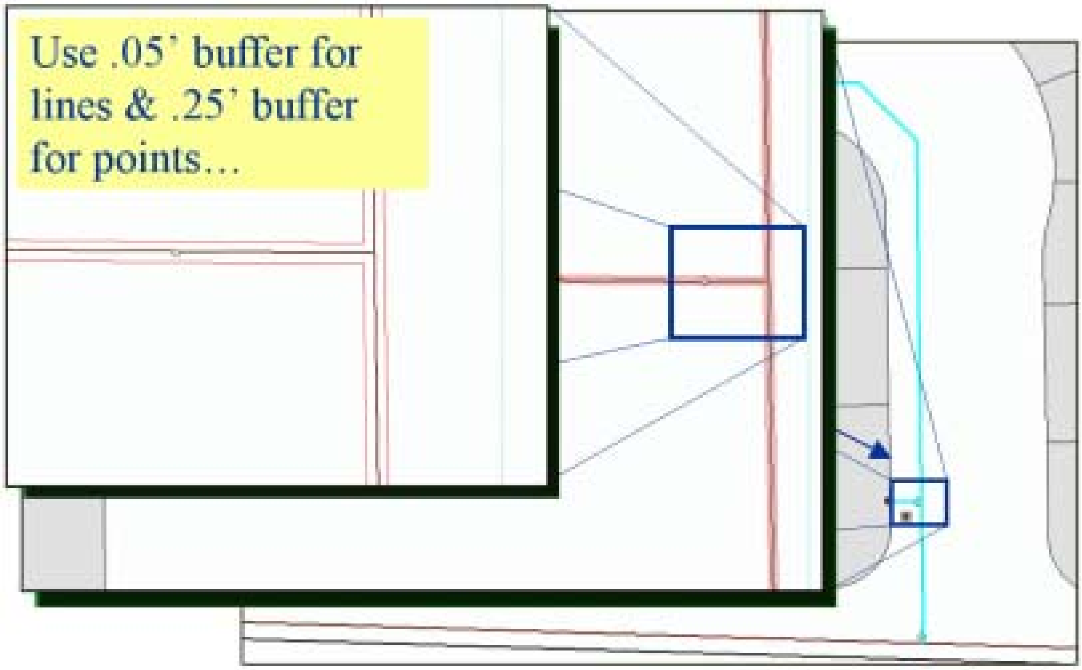

Figure 10 - A topological buffer was used to select and compare SDE and coverage features with a prescribed set of tolerances. Features must be completely within the "buffer" to pass the test.

Dropping the Library

The Library was dropped from production once each feature passed the data validation steps and the "SDEChkProj" application was incorporated into the "ProjectLoader" application. At this juncture, the Division had implemented a fully functional SDE database and had built and tested:

- Automated data loading systems supporting versioning, long transactions and reconciliation.
- An open data-editing environment allowing GIS and CAD teams to maintain the GIS database using native software.
- A modified plotting environment that accessed SDE feature layers for map production.
- Additional quality assurance and control procedures to ensure the accuracy and completeness of SDE data.

The next step was to re-engineer the GIS data central access point for corporate users.

Corporate Data Source

Before the conversion began, all GIS attribute data was stored in Oracle tables. Database primary keys were used to join spatial data, stored in Librarian, with the attributes stored in Oracle tables. Various applications were developed that directly accessed these Oracle tables for reporting and data maintenance tasks. The objective of this phase was to "switch" the target for accessing GIS attributes from native Oracle to the SDE feature class business tables.

Creation of SDE Business Tables

Additional SDE business tables were created to hold GIS attribute modifications created by the CAD team. These "staging" tables were used to temporarily maintain feature's attributes before being loaded into SDE by "ProjectLoader." The staging table record are deleted once the feature was loaded into SDE.

Corporate Business Tables

The various Oracle tables that used to hold attributes for Librarian layers are now kept synchronized with the respective SDE feature class tables. Table privileges make certain that these tables are only updated through a SDE synchronization processes. This supports compatibility with other applications that access these tables.

Where We Are Now

The Division has a fully maintained versioned SDE database that supports both CAD and GIS data management. Fully automated data loading, version management and many QA/QC functions have been deployed that have maintained the efficiency and effectiveness of Division staff and spatial data flow processes. Figure 11 illustrates the new production environment.

The implementation approach taken had significantly minimized impacts to production throughout the transition. Maintaining parallel GIS systems provided data validation and a framework for testing and deployment of applications and processes in a controlled environment that supported the step-wise re-engineering of specific GIS processes.

Next Steps

Additional steps are planned now that the SDE production environment is stable and in production. The next development phase will add validation and topology rules to refine and improve the network. This will involve customizing the ArcMap editing environment for network management and upgrading "ProjectLoader" to perform and support validation. Finally, the Division will continue to assess the benefits of production plotting in ArcMap and the performance costs incurred as a result of converting coverage annotation to geodatabase annotation.

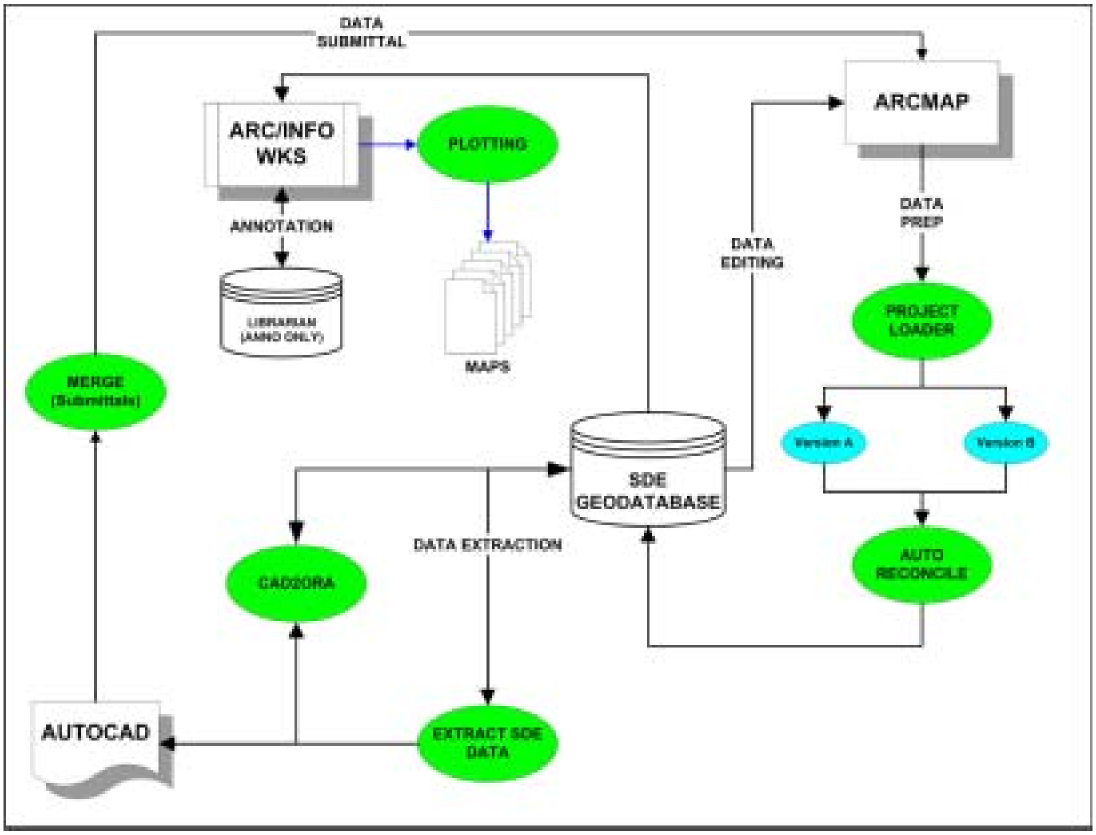

Figure 11 - The resulting SDE implementation supports both GIS and CAD clients. Key applications include CAD integration, automated project loading, versioning and reconciliation.

AUTHOR INFORMATION

Jonathan Pickus
Manager, AM/FM/GIS Division
Las Vegas Valley Water District
Las Vegas, NV. 89153
Phone: 702-259-8224
Email: jon.pickus@lvvwd.com

Jason Schwartz
Senior GIS Analyst, AM/FM/GIS Division
Las Vegas Valley Water District
Las Vegas, NV. 89153
Phone: 702-258-3826
Email: jason.schwartz@lvvwd.com

Keath Long
Senior GIS Analyst, AM/FM/GIS Division
Las Vegas Valley Water District
Las Vegas, NV. 89153
Phone: 702-822-8467
Email: keath.long@lvvwd.com