## SCS Case Study

# Improving Data Quality and Query Performance with a Data Warehouse

### *The Client*

Trean Corporation (pronounced "Tree-on") is an independently owned entity providing insurance management services, insurance and reinsurance consulting and reinsurance placement



services. Trean and its subsidiaries manage 4 insurance carriers and 1 captive insurance entity. One of the insurance carriers under management is an affiliate of Trean with an A.M. Best's "A-" positive rating and provides fronting services and risk sharing with insurers, captives and risk retention groups.

Website: http://www.treancorp.com/

### *The Challenge*

Trean Corporation has been growing quickly over the past several years, as has the amount of data it maintains. Their primary reporting databases were designed and implemented several years ago, when the company had a considerably smaller and less complex data footprint. There was little data cleansing while the data was being loaded from XML files into the database. As more reports have been requested over time, more logic was required to return the data being requested. This consumed more resources on the server, slowing everything down. Some reports were taking too long to run, testing the limits of users' patience and causing application timeouts. This necessitated the implementation of data caching monthly, adding a layer of complexity to their solution. Due to this caching, monthly processes became difficult to track across loads for different customers and made report creation overly difficult.

As an insurance management company, Trean Corporation receives XML files from numerous sources, built from different policy and claims management systems and different development teams. These differences create small (or not so small) issues in the XML that are being saved to the database. As reports are being written and tested, those issues are coded into the logic for each problem area the report touches. Not all of the issues are known or properly caught in the logic, leading to incorrect results being returned to the user, which caused skepticism about the correctness of the data.

This situation is not unique to Trean Corporation. Superior Consulting Services has encountered challenges like these in businesses of all sizes. Staff at Trean recognized their current solution was not going to be adequate for the long run. Consequently, they reached out to Superior Consulting Services to implement a solution to solve three main problems.

- Improve data quality
- Simplify report queries
- Decrease report execution times

To get a "game plan" developed, the first steps were to interview Trean's IT Data Team to better understand their current processes and databases, as well as what they would like to see for an overall solution. Prior to SCS working with this team, they worked with stakeholders throughout the company to define a list of measures and attributes they would like to see in reports. With the list of fields, SCS consultants also reviewed the 10 most frequently used reports, gathering additional fields stakeholders had not included in their list. The fields were mapped to the physical database tables and fields, to find the true source of the data. Doing so provided the scope for the project, with a solid idea of which tables needed to be exported from their current reporting database into a new operational data store (ODS), before being loaded into the data warehouse.

## *The Solution*

Based on the requirements, the project was split into two phases, one for their workers' compensation business and the other phase for the remaining business. From a data perspective, the majority of their data is for workers' compensation policies and claims, and it is stored and calculated differently than the other lines of business. Two SSIS solutions were created to load the ODS, one loading data from their existing system, which has a lot of data validations already in place. The other SSIS solution loads data from XML files into the ODS, including some data cleansing and then saving the data for historical purposes.

The next step was to create a data mart for workers' compensation data. For about half of the premiums and claims, dollar amounts associated with them are sent in the XML as an accumulated amount, generally for the life of the policy or claim. The remaining data sent from other systems is for just the last completed month. As the data is loaded into the data warehouse, the accumulated amounts are re-calculated, removing previous months, so it is just the incurred amounts for the current month. By transforming the data once, as it's going into the data warehouse, all of the incurred amounts are consistent, simplifying the report queries and future analytical processes.

The second phase of the project was to implement the same solution for the remaining lines of business, with data going into a separate ODS and data mart. The approach allows the focus of the ETL development to be on part of their business, and not have to account for every case in one spot. While it takes longer to implement, the overall ETL logic is simplified, as the developers do not need to account for conflicting rules between the two groups of business lines. While there are some differences in the design of the data marts, they are similar, making it reasonable to find the same information in both marts.

The final piece of the puzzle is the Enterprise Data Warehouse (EDW), combining the data from the two data marts. Once populated, the EDW gives users a complete view of the policies Trean Corporation manages. For example, users are able to see if a particular type of policy is exceeding a loss ratio threshold or not, the what percentage of their overall portfolio is defined by that type of policy, and much more.

## *The Result*

SQL Server Reporting Services reports that were re-written to use the data marts or the enterprise data warehouse generally run in seconds, rather than minutes.  Users have seen a 70 – 90% decrease in report execution times with this solution, compared to their previous solution.  The decrease in execution time is directly related to the decrease in query complexity, consolidating databases, and normalizing data.  Instead of needing to hit 10 tables or views and then transform the data into the form users are requesting, those transformations are now in the ETL process, allowing the developers to write queries against few tables and in a more straightforward manner.  Additionally, with the improved performance of their reports, Trean no longer needs to use data caching.

## *Contact*

**Superior Consulting Services, LLC**

350 West Burnsville Parkway, Suite 550

Burnsville, MN 55337

952-890-0606

www.teamscs.com