White Paper

One-Click Automation Framework
For Data Warehouse Testing

Amit Kurchania
Ganesh Parthasarathy
Ravish Hande

Tavant Technologies
PEOPLE. PASSION. EXCELLENCE.
Abstract

Testing is an integral part of the design lifecycle of any software system. While there are a number of software testing tools in the market, none of them address all aspects of data warehouse (DW) testing.

This paper presents some of the challenges in data warehouse testing and focuses on Tavant’s One-Click Data Warehouse Test Automation Framework, which addresses these challenges. The paper focuses on the logical components of the Test Automation framework and highlights its key features. It concludes with the business benefits realized by customers in using this framework.

Why Data Warehouse Testing?

A data warehouse provides an integrated view of an organization’s data and delivers actionable, timely and reliable business information for decision making. Poor quality data in the data warehouse can lead to wrong decisions affecting the organization’s performance. Hence, Data Warehouse Testing assumes a significant role in the long term dependability in decisions.

Challenges in DW Testing

> **Huge volumes of data**: While transactional systems involve testing individual records, data warehouse testing involves huge volumes of data from heterogeneous sources, which significantly impact performance and productivity.

> **High cost of data quality**: A data warehouse is primarily used for taking tactical and strategic business decisions. Every defect that slips into production transforms into high costs for the organization when you are looking at huge volumes that can amplify a small error.

> **Large scope of testing**: A transactional system may have a limited number of valid combinations of use cases. Since a data warehouse is aimed at supporting any number of views of data, the number of possible combinations of use cases is virtually unlimited.

> **Substantial business knowledge**: Transactional system testing focuses on program code, while data warehouse testing is directed at data. A good understanding of the data is required to ensure figures provided to the business user’s queries are correct.

What to Test?

> **Source data**: Source data testing includes verification of source files for data availability, data duplication and format validation.

> **Warehouse data**: Warehouse data testing includes verification of data in the warehouse for data completeness, data correctness, business logic transformation, primary key, foreign key, surrogate key references.
One-Click Automation Framework for Data Warehouse Testing

> **User interface front-end:** Front-end testing comprises activities such as user security, data security, performance, functionality and report disbursement.

**Why a DW Test Automation Framework?**

Current software testing tools in the market do not cover all aspects of data warehouse testing. The huge volumes of data loaded into a data warehouse make exhaustive manual comparison of data impractical. These factors motivated the Business Intelligence team at Tavant Technologies to embark on the initiative to build a DW Test Automation Framework.

**One-Click DW Test Automation Framework**

Tavant’s One-Click DW Test Automation Framework is an easy-to-use and comprehensive tool designed to test high-volume data warehouse applications.

**Overview**

One-Click DW Test Automation Framework integrates seamlessly into the Extract, Transform and Load (ETL) process and ensures all test cases are called in from the ETL process. It maintains a log of all test runs and a configurable workflow process to ensure that all stages of ETL are tested and verified before subsequent data loads are triggered.
One-Click Automation Framework for Data Warehouse Testing

It comprises a dashboard to view results of different test runs to perform advanced analysis such as trend analysis, statistical analysis and flow analysis. The report framework ensures all front end user interface components are verified by querying the data from the database and matching it with reports and dashboards.

To ensure data quality, the framework includes:

> Statistical Analysis like 3-Sigma Analysis and Moving Average analysis
> Trend Analysis for data over time
> Flow Analysis to analyze whether the threshold for data loss is acceptable
> Count Analysis to ensure data completeness and data accuracy

Features

> **Automation**: Provides the ability to automate and schedule test runs. Once configured, the processes can be triggered with a single click.
> **Early detection**: Ensures early detection of data quality defects and improves overall data quality.
> **Graphical view**: A graphical view of the test case runs with trend analysis for easy interpretation.
> **Traceability**: Stores data about all test runs for easy traceability.
> **Reusability**: Reusable in different implementations with minimal code changes.

Business Benefits

> **Lower total cost of ownership**: Accelerated testing process results in shorter testing cycles and hence reduces the overall effort for test execution. The framework can also be configured.
> **Improved data quality**: Automation of test processes ensures reduced human errors and improved data quality.
> **High customer satisfaction**: Improved data quality results in high customer satisfaction.

Tavant successfully implemented the framework for a leading Digital Video Recorder manufacturer and reduced the overall testing effort by 80 percent as compared to manual testing.
About Tavant Technologies

Tavant Technologies is a specialized IT solutions & services provider that leverages its expertise to provide impactful results to its customers. We have leveraged our unrivaled capabilities and domain insights to create game changing results for leading businesses across chosen industry micro-verticals. We are known for our long-lasting customer relationships, engineering excellence and passionate employees. Founded in 2000, we are headquartered in Santa Clara, California and service customers across North America, Europe, and Asia-Pacific.

About the Author

Amit Kurchania works as a Senior Lead - Quality Assurance at Tavant Technologies. He has over 7 years of experience in the area of software testing. He specializes in test automation of data warehouses using varied technologies.

Ganesh Parthasarathy is a Senior Architect at Tavant Technologies. With over 9 years of experience in business intelligence domain, his expertise spans business intelligence consulting, data modeling and system design. Ganesh has worked with customers across the globe.

Ravish Hande works as a Manager - Quality Assurance at Tavant Technologies. He has led many Quality related initiatives at Tavant, focusing on data warehouse and business intelligence.