

**MIGRATION GUIDE**

# Automated migration from SQL Server to Yellowbrick, the data warehouse for distributed clouds

Using Systech DBShift™ to migrate SQL Server based data & analytics environments to Yellowbrick

## Introduction

Many organizations rely on Microsoft SQL Server as their data and analytics platform. However, SQL Server was designed for transactional workloads and lacks the massively parallel processing (MPP) architecture required for doing interactive, ad hoc analytics at scale. In turn, organizations are looking to modernize their data warehouse and accelerate the migration to a platform that can support their analytics initiatives.

Yellowbrick Data and Systech have partnered to deliver a modernization and migration solution that leverages automation to help organizations move from SQL Server to Yellowbrick Data Warehouse. This guide provides an overview of migration activities and migration considerations.

## Modernization drivers

Legacy modernization means updating all or some of the data management and analytics stack (i.e., data warehouses, data platforms and applications) to better support business goals and processes. Business leaders drive innovation at their companies, but they need new technologies to enable and support this innovation in a real-world environment. They require fast applications, systems that support connectivity, and platforms that bring everything together. Most older IT stacks fail to meet those modern needs.

### Key platform modernization drivers

---

**Business Fit**



Inability to support the platform needs of modern analytic workloads in an efficient manner

**Business Value**



Lack of business value, quality of support, information delivered from the systems, etc.

**Agility**



Inability to scale at the pace of ever-changing business needs leading to unacceptable risk & cost

**Cost**

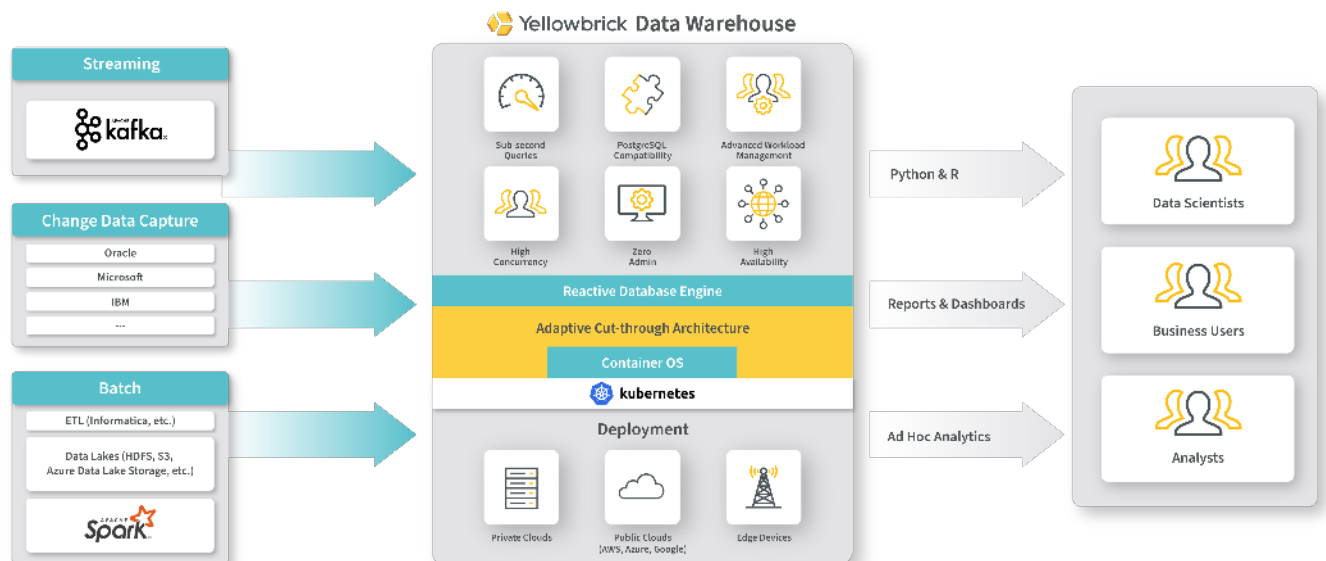


TCO of operating, maintaining, and changing the platform is high

## Modernizing your data warehouse with Yellowbrick

Yellowbrick Data Warehouse is an advanced, massively parallel (MPP), SQL database designed for the most demanding batch, real-time, ad hoc, and mixed workloads. It can run complex queries at up to petabyte scale across numerous nodes, with guaranteed sub-second response times. Yellowbrick was conceived with the goal of optimizing price/performance. It's not uncommon for customers to see their workloads run tens or hundreds of times faster at a fraction of the cost compared to cloud-only or legacy data warehouses.

Yellowbrick continuously implements new hardware (e.g., NVMe and flash memory) and software (most recently Kubernetes) protocols in an adaptive “cut-through” architecture that ensures best performance in every environment. Yellowbrick combines these advances with smart thinking about storage formats and indexing, and add on top a modern, standards-based database interface that's familiar to users (PostgreSQL) for ecosystem compatibility. The result is a modern, quickly provisioned, and easy-to-use data warehouse that knocks the socks off rivals in price/performance economics and can be deployed across distributed clouds (private data centers, public clouds, and edge networks) – with all instances, databases, and users managed through a simple, unified control plane (Yellowbrick Manager).



## Replace SQL Server with Yellowbrick for 100X performance at a fraction of the cost

Replacing SQL Server databases with Yellowbrick Data Warehouse brings the power of interactive, ad hoc analytics to thousands of users, simplifies operations, and reduces costs.

- **Analyze data 100X faster:** Designed for highly optimized MPP analytics, Yellowbrick enables real-time, ad-hoc queries for thousands of users, far beyond what SQL Server can achieve.
- **Consolidate disparate databases:** Yellowbrick Data Warehouse can scale from single TBs to multiple PBs, offering a way to combine multiple SQL Server databases into a single, easy-to manage instance. Or, easily replicate data from SQL Server to Yellowbrick for Change Data Capture scenarios.
- **Simplify management:** With Yellowbrick, unlike SQL Server, there's no need for manual time-consuming tasks like query tuning or building indexes. PostgreSQL compatibility ensures access to common skill sets.
- **Count on rock-solid reliability:** Yellowbrick's Advanced Workload Management offers granular control of workloads to prioritize concurrent queries across massive amounts of data, bringing added reliability and performance.
- **Grow along with the business:** Yellowbrick instances can be expanded easily without downtime, eliminating any impact on operations.
- **Fit into the Windows world:** Yellowbrick integrates with Microsoft Active Directory, and for cloud deployments, functions as a Private Link Service in Azure. Plus, PowerBI and a full range of other enterprise ecosystem tools are supported.

### Customer Spotlight

## BMW Group Financial Services

BMW Group Financial Services is one of the leading financial services providers in the automotive sector, serving customers worldwide.

### Challenge

The customer's SQL Server-based analytics and reporting solution was at the edge of performance and scalability limits for use cases like funding analysis, payments aggregation, and calculating portfolio internal rate of return.

### Results with Yellowbrick

BMW Group chose Yellowbrick to replace SQL Server as its analytics and reporting platform.



### Test results including:

- 152X faster queries on average
- 70% reduction in required data storage space
- Successful integrations with Informatica, Tableau, and SAS

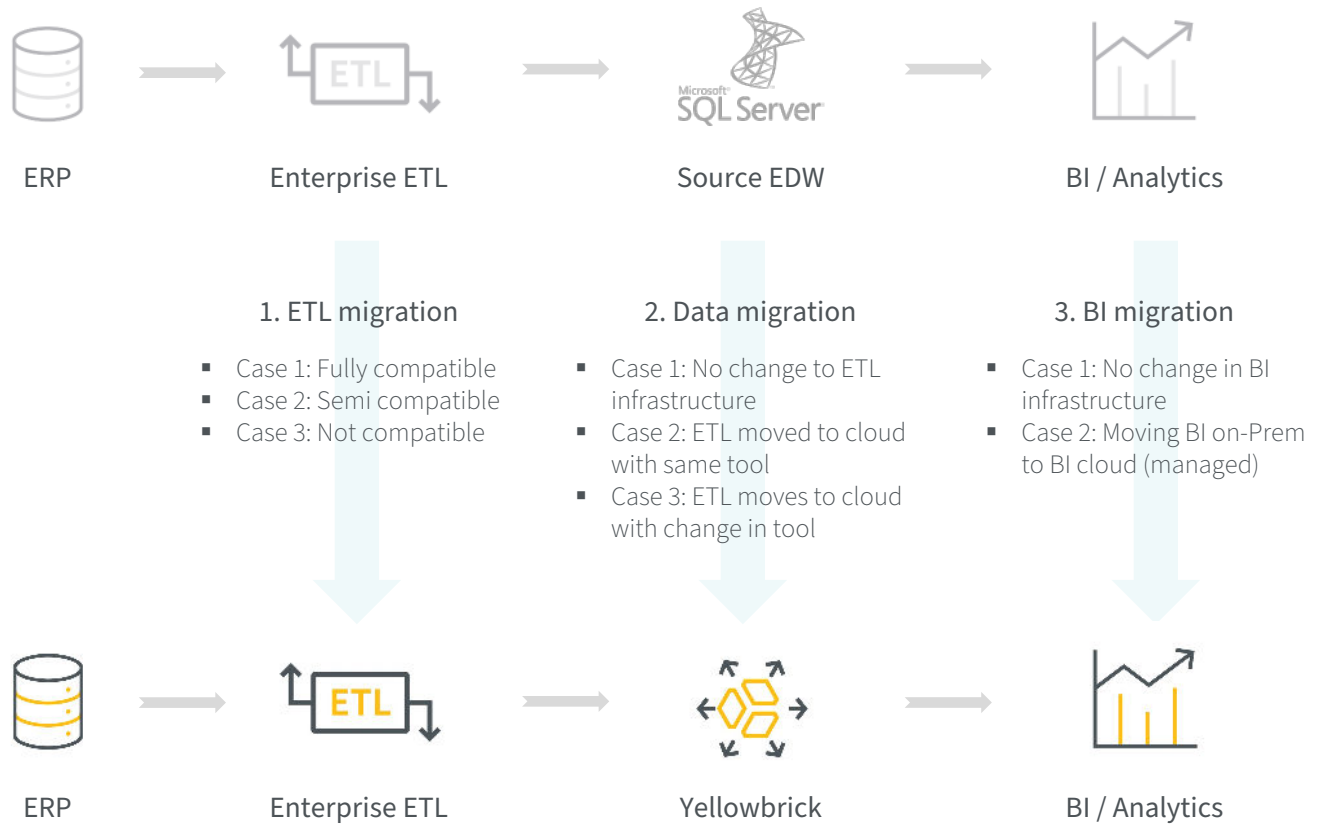
**“We’re confident that Yellowbrick’s ability to quickly analyze large amounts of data and offer new insights will help us deliver on our goals.”**

Ian Smith  
CEO, BMW Group Financial Services North America

## A holistic modernization and migration approach

Undoubtedly, modernizing the data management and analytics environment provides huge business value. Next, it seems pertinent to highlight the effort and tasks involved in modernizing and migration of the environment. The diagram below outlines a typical analytical ecosystem.

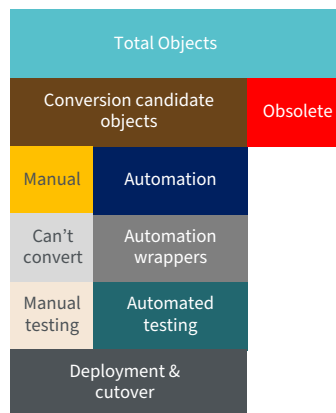
The data storage layer in the middle is where the data warehouse and data lake(s) are stored. To the left of the data storage layer there are data pipelines that bring in data from various sources internal and external to the organization, transform and load it into the data warehouse and data lakes. To the right side of the data storage layer there are dashboarding, visualization, and reporting applications that read data from the data warehouse and data lakes. There could also be other downstream applications that consume data from the data storage layer. Modernization involves the migration of these 3 layers into Yellowbrick.



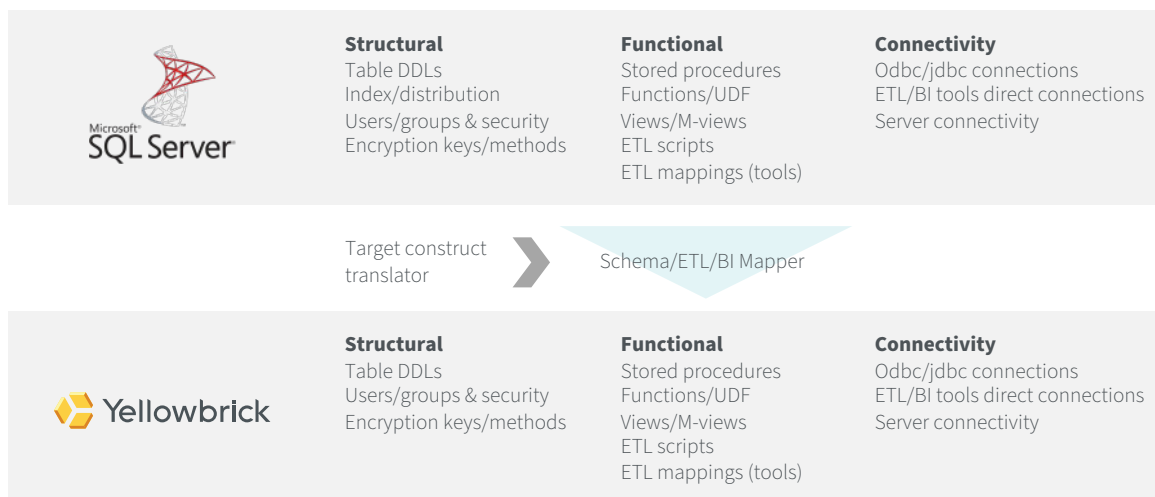
## Migration considerations

The migration of these 3 layers requires systematic planning and execution. For each of these 3 layers, inventory analysis should first be performed to identify objects that need to be migrated. Migration should be taken as an opportunity to clean up obsolete, not required, and unused objects. The next step once the objects have been identified is to map each of the source data warehouse (SQL Server) components, both functional and structural to target data warehouse (Yellowbrick) environment. Migration should then be carried out based on these mappings followed by detailed validation, UAT, and sign-off.

**Inventory analysis**  
Done for database objects,  
ETL objects, and BI reports



The DB Objects migration involves structural migration from SQL Server to Yellowbrick. SQL Server specific SQL functions and statement syntaxes need to be identified and converted to ANSI SQL or Yellowbrick specific functions. Each SQL Server database object including but not limited to databases, schemas, tables, columns (data types), indexes, triggers, and stored procedures need to be mapped and converted to Yellowbrick objects. Next comes the migration of connections used in data pipeline platforms and dashboard and visualization tools from SQL Server to Yellowbrick.



## Migration methodology

The phases of migration outlined in the previous section can further be detailed into a migration roadmap and approach document. Based on the maturity of the data management and analytics environment, the complexity of migration varies. More mature environments employ all possible processes to extract insights from data and these need to be migrated to Yellowbrick. Migration becomes more involved both based on the number of objects to be migrated as well as the complexity of these objects.



### Inventory Analysis

- ✓ Detailed objects inventory
  - ✓ SQL Server databases, schemas and objects
  - ✓ ETL ecosystem
  - ✓ Reporting / dashboard ecosystem
  - ✓ Downstream systems
- ✓ Determine automation level (70%+)
- ✓ Migration roadmap & detailed plan
- ✓ Define impact criteria
- ✓ Identify in-flight projects and define data & code synchronization mechanism
- ✓ Testing & in-depth validation strategy



### Setup Yellowbrick On-Premise / Edge / Cloud

#### Configure Network & Security

- ✓ Configure user access
- ✓ Configure firewall rules & network controls
- ✓ Configure an inbound traffic protocol(HTTPS,RDP)
- ✓ Configure a local persistent storage and a database repository services

#### Configure Container & Abstraction Services

- ✓ Configure the necessary container & abstraction services

#### Create Static IP for VPC

- ✓ Allocate & associate a new private IP address to the instance

#### Validation

- ✓ Perform validation of all the configured services



### Impact Analysis

- ✓ Identify objects that needs to be migrated based on the impact criteria
- ✓ Identify the data / code base which needs to be re-written/refactored

#### Data Migration

- ✓ Modify / re-create database objects in the Yellowbrick environment
- ✓ Re-architect / migrate data management
- ✓ Jobs(SQL, NoSQL, Hadoop, Data Pipes etc)
- ✓ Modify connection parameters
- ✓ Apply changes to the new environment

#### Application Migration

- ✓ Re-architect / migrate application repository i.e. reports, portals, dashboards, microservices etc.
- ✓ Apply changes to the new environment

#### In-Flight Projects Management

- ✓ Perform code sync at multiple check points & dual coding wherever needed
- ✓ Perform final data sync
- ✓ Production freeze



### Validation

- ✓ Perform In-depth data validation & functionality validation for identified reports
- ✓ Perform multiple parallel runs of data pipelines for testing data accuracy
- ✓ Perform in-depth testing of migrated code based for all objects affected by impact criteria

#### Monitoring & Resource Utilization

- ✓ Set up all the log capture mechanism
- ✓ Setup monitoring mechanisms for Resource utilization and billing alert

#### Deploy in Production

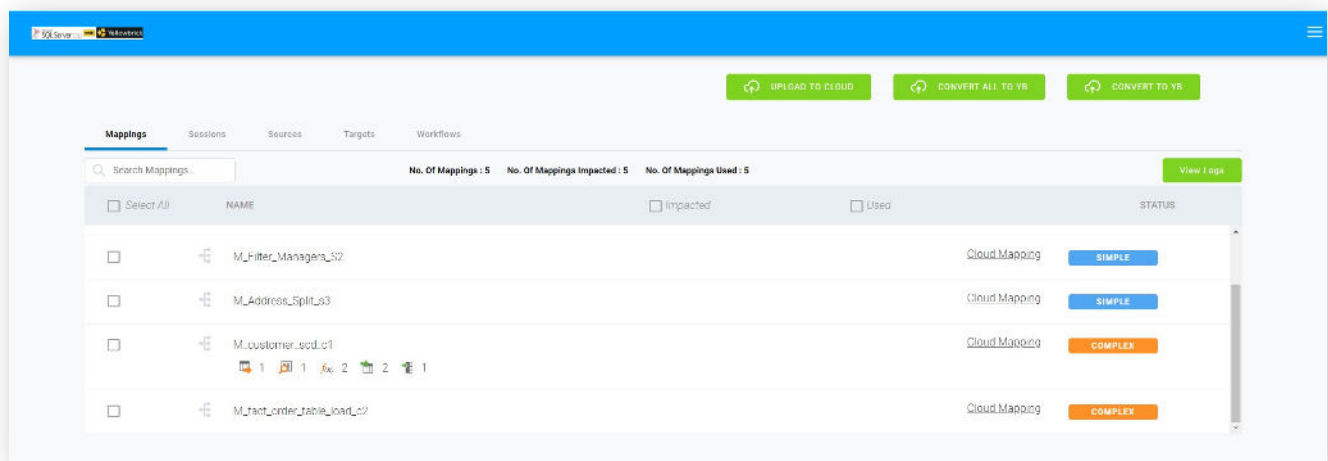
System's migration methodology is well proven and accelerates the migration from SQL Server to Yellowbrick. Each of these phases will have well-defined inputs (artifacts, process outputs) and produce outputs that will be fed to the subsequent phase.

## Automated migration using Systech DBShift™

Systech’s DBShift™ utility accelerates data warehouse modernization by automating the discovery, migration and validation phases of the migration. Automation scripts will be deployed to identify objects that need to be migrated. These SQL Server objects are automatically mapped to Yellowbrick objects. In addition, DBShift™ can perform automation of migration of the data pipelines built on platforms like Informatica from SQL Server to Yellowbrick. DBShift™ can achieve at least 70% automation, reducing migration efforts and cost considerably.

Key features of DBShift™:

- ✓ Guided strategy & assessment to evaluate SQL Server environment and select the right migration approach to meet business needs (e.g., Re-Host / Re-Architect / Re-Factor).
- ✓ Automated database, data pipeline, and application migration utilities to ensure a swift, secure, and smooth transition to cloud.
- ✓ Containerization & orchestration of workloads for a higher level of baked-in scalability, monitoring and fault tolerance.
- ✓ DevOps methodology for continuous integration & delivery.
- ✓ Monitoring & logging for security and performance (APM).





## Summary

This document explores the various benefits of migrating your data and analytics environment from Microsoft SQL Server to Yellowbrick Data Warehouse. Specifically, this migration guide highlights the modernization drivers, migration approach, migration considerations, migration methodology, key features of DBShift™, and the benefits that Yellowbrick offers.

To learn more about Yellowbrick Data, visit [yellowbrick.com](https://yellowbrick.com) to book a demo or sign up for a free test drive.

## About Systech

Systech Solutions delivers measurable value, fast. They accomplish data and analytics initiatives 2x faster for 2x lower cost than most providers. With an unsurpassed technical breadth and depth along with 25+ years of industry experience, Systech is a leader in their field, having executed thousands of data strategy and management projects for businesses across verticals with a 100% success rate. To learn more about Systech, please visit: [www.systechusa.com](https://www.systechusa.com).