D.E. Shaw Group Replacing SAP IQ – Why Yellowbrick?

Yellowbrick: the Next Best Step for SAP IQ End of Life

Hello,

There's only one K8s based Distributed Data Platform that affords the world's highest performing investment and technology development firms flexibility in deployment, simplicity in management, and unrivaled performance at scale.

In order to power business-critical mixed workloads, Yellowbrick has rebuilt the data warehouse from the ground up. Our customers seamlessly migrate production workloads in-place, consolidating siloed data and utilizing Yellowbrick as their conduit to the cloud, or in some cases back to an on-prem configuration.

We look forward to briefing your team on our technologies and sharing why many call Yellowbrick a competitive advantage. Deconcentrate your deployment risk while meeting strict SLAs and do more with your data than ever imagined.

Mike Hinke

Mike Hinke
Director of Sales
mike.hinke@yellowbrick.com

William Huttner
Sales Development Representative
william.huttner@yellowbrick.com

Contents

The End of Life for SAP IQ

For decades, SAP IQ has enjoyed popularity with many organizations, especially Financial Services Industry players (FSIs). However, it is clearly approaching end-of-life. SAP is now only promising updates for bug fixes and security modifications. Many SAP IQ customers are seeking the simplest, least expensive, and most scalable analytics database platform that can be quickly deployed and substituted for SAP IQ (cloud or on-premises). SAP IQ is not viewed as an industry-leading, best-of-breed modern analytics database for either on-premises or cloud usage.

Why Not SAP HANA as a SAP IQ Replacement?

An "SAP IQ business continuity strategy" conversation with SAP almost always yields the recommendation to consider SAP HANA. However, HANA is not an optimal choice for a general-purpose analytics database and/or data warehousing platform (at least where SAP's ERP is not involved).

SAP HANA is a high-performance, clustered database that provides sub-second response to analytic queries against multimodel data, on-premises and in the cloud. HANA also handles traditional OLTP transaction processing. However, as a replacement for SAP IQ, we urge organizations to exercise caution. There are several reasons a modern analytical and data warehousing architecture, such as Yellowbrick can be advantageous:

1. HANA is an in-memory database, meaning that across all queries being executed at any point-in-time, all datasets being analyzed (along with all transient, internal data structures created to perform the analysis) must fit in DRAM memory at all times—or queries will abort. As datasets and workloads grow, it requires costly, memory-rich servers, and a growing number of them in the HANA cluster (each requiring separate licenses or subscriptions).

    It is preferable for real-time query response to occur without queries ever failing and to be achieved by the analytical database platform automatically, making the best use of the hardware's available "memory hierarchy" (DRAM, SSD, rotating media, and Cloud Object Store). This is the case with some modern analytical database choices, including Yellowbrick.

2. Like SAP IQ, HANA requires that DBAs declare indexes that match user queries to achieve promised performance. This introduces a significant database maintenance burden and an undesirable dynamic between data analysts and DBAs that impedes productivity. This is all too familiar for traditional RDBMSs (and SAP IQ), namely. Experienced DBAs must understand what query patterns are in use, declare the indexes correctly, and define new indexes quickly as queries evolve.

    Additionally, when data loading occurs, declared indexes must be populated by HANA at data load-time, increasing its workload and slowing data load speeds considerably. In analytical environments, the speed with which freshly loaded data becomes visible to queries is very important. These challenges are mitigated by sidestepping the need for any indexes, a characteristic of Yellowbrick.

3. With HANA, one must explicitly declare the choice of whether to store data, at load-time, in HANA's rowstore vs. its columnstore. Two notes for all analytical databases:

(a) sourcing a query's needed columns from referenced tables is much more efficient from column-based storage than from a rowstore; but (b) when loading data as singleton rows or in small batches, the fastest possible data load speed is achieved into a rowstore, not a columnstore.

These factors present a dilemma in analytical database architecture, for which HANA's solution of each table having to be declared row- or column-oriented (not both) is highly problematic.

Yellowbrick navigates this problem successfully with an important innovation. Large batches of bulk-loaded data go directly into the columnstore at massive speed, but loading of individual rows and small batches (which is the typically-sized data loaded, for example, through a JDBC interface) are automatically routed to a rowstore (from which the data is immediately and transparently visible to queries, along with data in the columnstore). Then, Yellowbrick automatically and transparently migrates all recently loaded data from the rowstore for each table into its corresponding columnstore—thereby giving the best of both worlds in terms of performance with no effort on the application developer's part.

As a consequence of these challenges, by comparison to Yellowbrick, HANA's Total Cost of Ownership (TCO) and time-to-deployment can be quite high. To create a different incentive for SAP IQ users to follow SAP's preferred migration path to HANA, SAP is reportedly considering a feature that enables IQ users to avoid the need to migrate their data into HANA in a "big bang." HANA's memory-based query and processing engine would provide backward-compatible access to existing IQ data files, by adding IQ components into the HANA stack that manages a complex "level of indirection" to

IQ's disk and index management layers. Apart from the hardware cost and run-time overhead of this approach to every query, this baroque "merger" between the products would add to the three challenges with HANA's manageability and TCO listed above.

From a historical perspective, experience suggests that there comes a time when it is appropriate to migrate from a legacy database platform, such as SAP IQ, to a modern, durable database architecture that is economical, fit-for-purpose, and will last for decades. Yellowbrick makes that architectural transition easy; see the Customer Success Story below.

Yellowbrick vs. SAP IQ

For the reasons discussed, SAP HANA does not always offer the best migration path for SAP IQ users. Instead, many FSIs have migrated to the Yellowbrick modern cloud data warehouse. This has reduced query time to seconds, accelerated time to value and insight, and reduced costs through a predictable subscription model. The next section presents some findings from a FSI organization.

This section compares the key capabilities of Yellowbrick and SAP IQ: why it is the obvious next step for SAP IQ users and how Yellowbrick extends the attributes of SAP IQ.

Time to Value

An area of real improvement with Yellowbrick is that organizations are able to speed query and trading reports performance. Data loads into Yellowbrick within minutes versus hours in SAP IQ. Queries run in seconds, not minutes.

Database Structure

An IQ server, either single node or cluster (multiplex), only supports a single database instance. This implies the entire structure must be handled as a single entity and to support multiple applications that may have no common structures. Yellowbrick overcomes this

limitation. When new nodes are added to a cluster, data redistribution across the new nodes happen automatically in the background with little impact to query processing; hence, capacity expansion is viewed as an online activity.

Database Replication

SAP IQ uses Replication Server for database and object replication. This product is almost without competition within the OLTP database market, yet it has never fully been supported by IQ. Inside IQ, the development, maintenance, and control of a replicated structure is complex and both time and space constrained. Replication is an inherent component within Yellowbrick, which offers improvements to SAP IQ users, including full support for unidirectional, asynchronous replication. Also, replication can be between on-premises and cloud instances, as desired, and is transactionally consistent with data written to the target in one transaction to guarantee consistency for all users.

Data Loading

SAP IQ was developed for bulk loading in the large volume table via the SAP Sybase IQ LOAD TABLE SQL command. There is a mechanism for individual row loading, but it is complex and limits performance, resulting in query performance degradation.

Yellowbrick's modern architecture spans all the structures from device drivers up to the higher functions of the database, which is much faster at passing data into the database and providing SQL support for row-level inserts and bulk loads. With this modern architecture, SAP IQ customers easily load data in parallel from multiple files, further reducing the complexity for IQ customers who have grown accustomed to the benefits of bulk loading.

Operational Maintenance

SAP IQ is known as a database system that requires very little operational maintenance. Yet, as the system has become more complex and size and requirements have increased, a growing amount of

operational control and maintenance is required. SAP IQ's DBAs are also retiring so resources, on the aging technology, are becoming scarce. On the other hand, engineers and staff who are familiar with PostgreSQL are plentiful. Yellowbrick's open standards using PostgreSQL allows for an easy migration that offers an integrated operational and maintenance package that streamlines installation, upgrade, and control.

ETL Processes

Yellowbrick is compatible with traditional on-premises ETL tools, like Informatica PowerCenter and IBM DataStage, even in the cloud. Purpose-built to avoid vendor lock-in, Yellowbrick uses open standards to ensure portability of data and workloads. SAP IQ requires changes to the ETL and load process to the database and the older ETL processes would likely need to be replaced in a product migration. It is expected that these "replacements" would not need modification, simply implementation.

Cost

Customers quickly find Yellowbrick's subscription pricing model is more cost-effective than the leading competitors, including SAP IQ. With other leading data warehouse vendors, the more you use, the more you spend. With Yellowbrick, as the number of queries run increase, the cost per query is reduced significantly––the more you use, the more you save.

The following table presents a summary of features that make Yellowbrick a solid replacement for existing SAP IQ customers:

Features	Yellowbrick	SAP IQ
Public Cloud, Private Cloud, Hybrid Cloud, or On-Premises Deployment	✓	
Best-of-Breed Data Load Performance	✓	
Best-of-Breed Query Performance	✓	
Best-of-Breed System Scalability (e.g., as data volumes and # of users grow)	✓	
Best-of-Breed Query Concurrency (i.e., Workload) Management	✓	
No Need to Declare Indexes	✓	
Multi-Database Support	✓	
Data Encryption at Rest with Zero Performance Impact	✓	
Fully Transactional Behavior for Multiple Writers and Readers-vs.-Writers	✓	
Performance Scalability	✓	
Dynamically (Transparently) Redistributes Data Across Additional Nodes When Capacity is Added	✓	
Architecture Emphasizes Minimal Human Maintenance	✓	
Superior Integration with ETL and BI Tools (By Virtue of SQL Compatibility with Postgres)	✓	
Simple, Cost-Effective Subscription Pricing	✓	
High-Availability with Built-In Replication Services	✓ (Built-In)	**Partially Supported (Add-On Component)**
Ready Availability of DBA's	✓ (Postgres)	**Retiring**
Proven Columnstore Relational Database	✓	✓
Compliant with SQL Standards	✓	✓
Designed for Analytics Workload on Large Data Sets	✓	✓
Deploy On-Premises and On Infrastructure as a Service	✓	✓
Data Compression	✓	✓
Custom Optimizer	✓	✓

Financial Services Industry Customer Success Story

One of the leading market makers in the financial services industry that executes billions of daily trades for third parties migrated from SAP IQ to Yellowbrick. In this example, Yellowbrick builds on the capabilities of SAP IQ allowing this market maker to gain faster and richer insights in minutes, while also saving millions of dollars.

SAP IQ was used for historical trade analysis; over time, this data increased to billions of daily trades and, as a result of developer investment, IQ was unable to handle this growth. Historical analysis was limited to a quarter period, making quarter on quarter or annual analysis impossible. Performance was slow with trading reports taking an entire day to be generated adding regulatory risk.

After evaluating a number of leading providers, Yellowbrick was chosen as the next modern analytics platform, replacing SAP IQ for historical trade analysis .

Testing was completed in only three weeks and the results included:

- Up to 728x faster trading reports than with SAP IQ
  – With SAP IQ: 19 hours
  – With Yellowbrick: 94 seconds
- Ad hoc queries for regulatory requests run in seconds instead of hours, allowing for on demand request of CAT compliance reports
  – With SAP IQ: One query analyzing 440 million rows executed in hours
  – With Yellowbrick: The same query executed in 4 seconds
- Business users can rapidly access trillions of rows for historical trade analysis
  – With SAP IQ: Data loading in four hours, queries running in minutes
  – With Yellowbrick: Data loading in four minutes (~15 TB/hr) and queries running in seconds
- Transforming the data allowed for querying of more up-to-date data, faster
  – With SAP IQ: Transforming the data (ELT) takes 5+ hours
  – With Yellowbrick: ELT takes four minutes

Summary

Yellowbrick is designed to deliver high performance to the mixed workloads that modern enterprises require. It can support thousands of users running mixtures of real-time transactional atomic inserts; bulk loads of data; ELT workloads; interactive queries; ad hoc queries; and long-running, batch-oriented reports. While today's mixed workloads require a high-performance solution and as the number and diversity of workloads increase, it is critical for FSIs to specify, control, and automate the work that matters most.

As a result of the lack of investment in SAP IQ from the developer, many FSIs find that their legacy analytic database platform used for analyzing and reporting trades, investments, risk, and compliance is no longer a fit for their growing needs. Additionally, the ongoing shift towards cloud-hosting and cloud-native solutions are forcing a huge transition. At the same time, companies must contend with higher data volumes and velocities and meet the challenges of heightening security and regulatory requirements.

To achieve success, companies need all tools and technologies to work in concert. This is Yellowbrick's strength—designed from the ground-up to work together with maximum performance and scalability and ease-of-integration with minimal effort or ongoing maintenance. Yellowbrick's deployment and migration journey to the cloud, on-premises, or a hybrid of the two, is fast, efficient, and simple.

Yellowbrick is a modern Data Warehouse that delivers high performance on massive volumes of live data with the fastest and most efficient execution necessary for making critical, better informed business decisions with predictable and controlled costs.

Intrigued? Find your biggest, most complex data set and most important business challenge and invite us in for a Yellowbrick test drive. We'll show you how high-speed data analytics can redefine your business opportunities.

Try our free 7-day test drive:

yellowbrick.com/test-drive