

Automating Data Pipelines

Qlik's Data Integration Platform for Federal Agencies



INTRODUCTION

How Federal Agencies leverage data for decision making is changing. Historical transaction reporting and static batch data analyses are giving way to predictive data science, streaming analytics, machine learning (ML), and artificial intelligence (AI). Agencies are embracing these new paradigms for more informed, real-time decisions to impact business, operations, and agency performance.

Federal data platforms are reaching such levels of data volume and complexity that traditional integration tools no longer suffice. The scale, speed, and low latency necessary to satisfy Agency requirements requires a more innovative integration solution.



That solution is Qlik Data Integration. The foundation for modern, efficient data pipelines, Qlik Data Integration provides automated, real-time, and universal data integration across all major data lakes, streaming systems, databases, data warehouses, and mainframe systems and helps automate the creation of new data warehouses and lakes both on-premises and in the cloud. Leveraging Qlik Data Integration, Agencies can deliver more data, in real time, and in analytic ready form.

Qlik Data Integration enables agile analytics all while reducing time, effort, and cost. One example of this is with a large Federal ERP customer who shortened their real-time analytics delivery schedule from 6 months to 1 week due to Qlik's no-code, drag and drop UI. That same customer also saw a 10x decrease in time for processing financial transactions compared to the ERP vendor's native toolset.

Qlik Data Integration is different. It streams data at high speed and in real time from source to target, with each data pipeline configured and monitored in an intuitive, no-code, drag-and-drop web interface. This completely automates and accelerates end-to-end replication. With Qlik Data Integration's streamlined

and agentless configuration, administrators and data architects can quickly set up, control, and monitor bulk loads and ongoing real-time streams across the entire data platform environment.

Qlik's Data Integration Platform Capabilities

Qlik's approach for delivering world-class analytics and integration capabilities can be summarized by the following guiding principles:



SaaS Principle – The market is clearly shifting to the cloud and the consumption of SaaS services. We recognize that each Agency's journey to the cloud and the pace at which they move to the cloud is specific to their mission requirements. We are building solutions that enable that journey to the cloud.

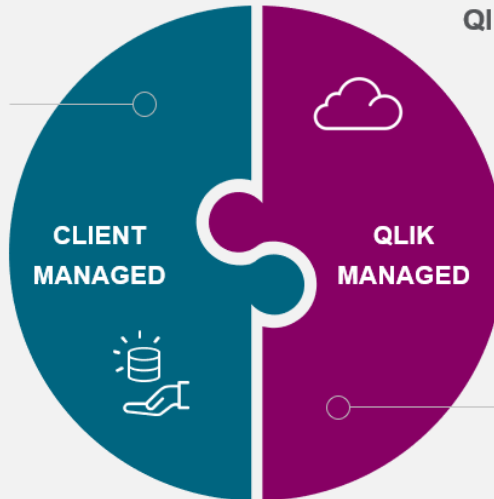
Data Principle – We are building solutions that provide the flexibility for Agencies to house their data where they want and enable them to get business value from that data regardless of location.

Customer Principle – Customers are at the center of all we do.

With these guiding principles in mind, Qlik's Data Integration Platform capabilities align with these and offer Federal agencies the flexibility they need to modernize their data strategies to leverage new and emerging capabilities.

Qlik Data Integration (QDI)

- Best of breed real-time data movement and automation
- Product suite (Qlik Replicate, Qlik Compose, Qlik Enterprise Manager)
- Targets IT personas (data architects, data engineering, IT ops, data stewards/governance)



Qlik Cloud Data Integration (QCDI)

- iPaaS - Integration Platform as a Service
- Offering is a collection of one or more data services (data movement/storage, transformation, data mart)
- Personas include both IT and LOB (citizen integrator, power users, data scientists,)
- New innovations are Cloud first

Qlik Data Integration (QDI) helps to automate real-time data pipelines for analytics. At the core of this capability is a market-leading set of change data capture (CDC) and data automation capabilities that can be deployed in any on premises or cloud data center, giving Agencies the flexibility and choice that align with our core values above. The following details these core capabilities of QDI:

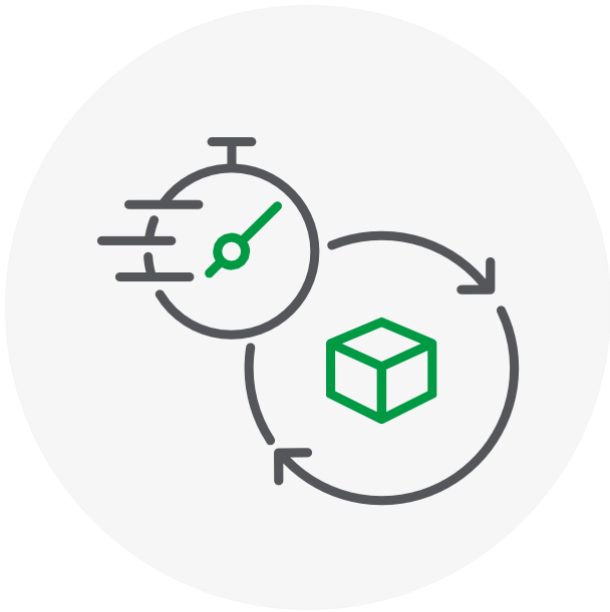
Qlik Replicate

Qlik Replicate continuously automates CDC from multiple data sources to the cloud target. Our solution helps you avoid the heavy lifting associated with manually extracting data, transferring it via API/script, then chopping, staging, and importing it.

Qlik Replicate supports all major source and target data systems for data replication. This includes data systems used by many of the major commercial enterprises and Federal customers. Years of innovation, optimization, and our strategic partnerships ensure our solution integrates data across all these platforms seamlessly.



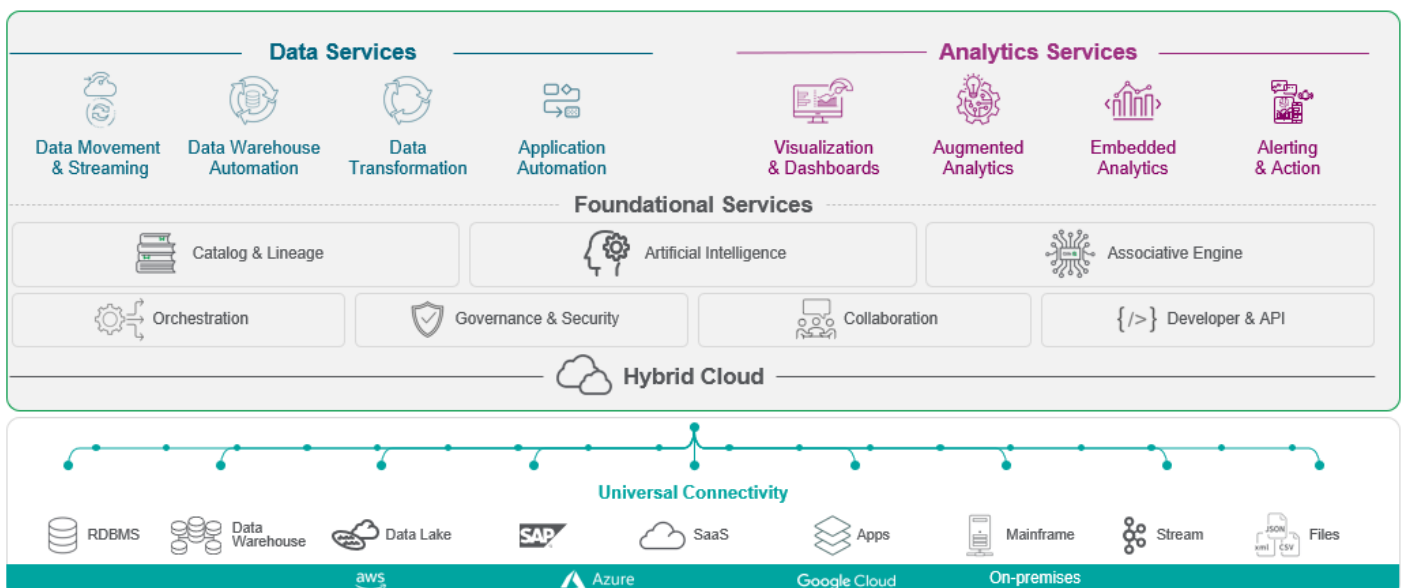
Qlik Compose



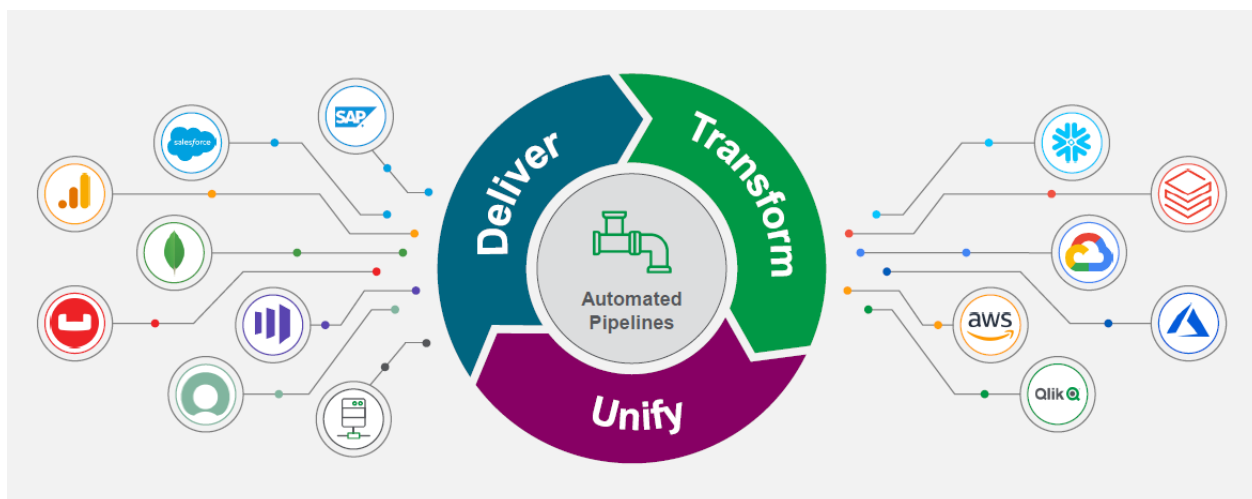
Qlik Compose automates the design, implementation, and updates of the data that Qlik Replicate streams into Data Lakes and Warehouses. It minimizes the existing manual, error-prone design processes of data modeling, ETL coding, and scripting. Deploying Qlik Compose speeds your analytics projects, helps you achieve greater agility, reduces your risk, and allows Agencies to realize the potential of their cloud investment more fully.

Qlik Cloud Data Integration (QCDI)

Qlik Cloud provides all the **Data Integration** and **Analytics** needed to transform raw data into informed action, supported by a rich set of foundational services. It's the only cloud platform designed to help organizations shift from a passive set of tools to an active system that delivers information in real-time and compels action. With this open SaaS platform, Agencies also benefit from cloud-agnostic and **hybrid** deployment options that offer maximum choice and flexibility in how and where you store and analyze data, across one or multiple clouds.

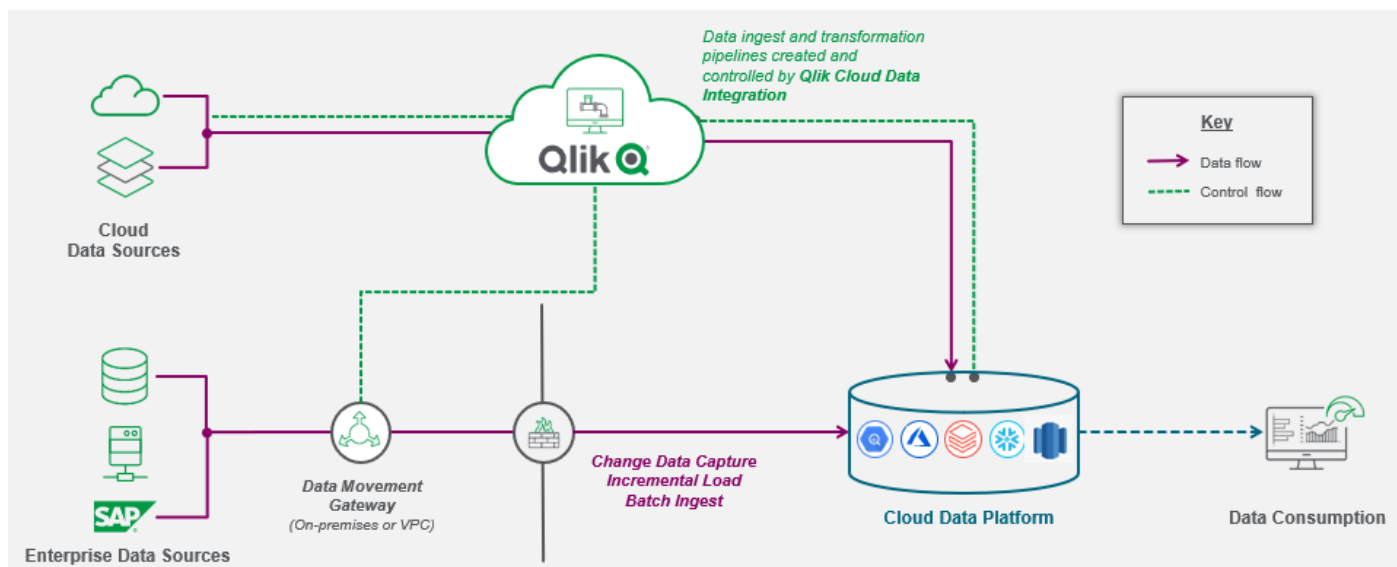


As with QDI, Qlik Cloud Data Integration (QCDI) provides a powerful data integration fabric to help you deliver, transform, and unify enterprise data in real-time via automated, governed, and reusable data pipelines.



QCDI's Data Movement Architecture

QCDI's data movement architecture and components are depicted and described below:



- **Qlik Cloud** – this is where we define, manage, and monitor all our data pipelines, projects, tasks, and connections – this is our central command and control.
- **Enterprise data sources** – Sources that reside behind firewall or inside a virtual private cloud (VPC).

- To access enterprise data sources, we offer a Data Movement Gateway, which is an agent used established secure reverse tunnel communications between the data source, Qlik Cloud, and Cloud targets.
- When streaming data from those sources via the Data Movement Gateway, data flows directly into the cloud target without doing a detour via Qlik Cloud.
- **Cloud data sources** – cloud-based applications like Salesforce and ServiceNow. Those sources are usually hosted in a public cloud and do not require the Data Movement Gateway for access.
 - When connecting to a cloud source, data does flow through Qlik Cloud.
 - The supported data delivery method today is Full Load only (CDC is on the roadmap).
- **Cloud data platforms** – Qlik’s emerging Cloud Data Platform target support currently includes the following cloud-native data platforms: Snowflake, Azure Synapse, Google Big Query, Databricks, and Redshift, as well as QVD generation for Qlik Cloud Analytics capabilities.

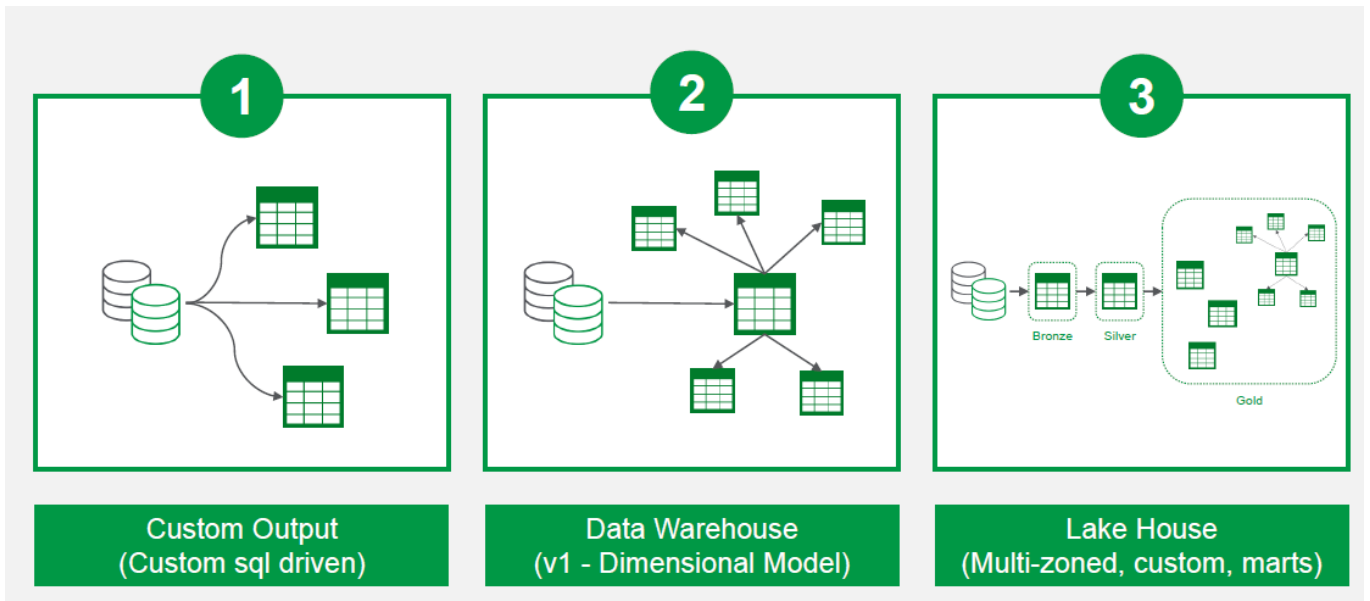
QCDI’s Transformation Capabilities

Once data is captured and stored into cloud data platform targets, QCDI can provide transformation capabilities to prepare raw data into analytics-ready datasets.

The following transformative features are offered within QCDI:

- **Onboard data** from a wide range of sources to your cloud data store.
- **Transform data** with rules and custom SQL-based transformations.
- **Automate data mart** creation, with other design patterns to follow.
- **Re-use** transformation outputs as inputs to other patterns.
- **Automate**, orchestrate, and monitor end-to-end transformation pipelines.

QCDI’s Data Transformation capabilities are depicted and described below:



- **Custom output** – Ability to define any target structure within a data pipeline.
- **Data Warehouse** – In the initial version of this capability, raw data can be transformed into dimensional star-schema models for analytic consumption.
- **Lake House** – Ability to offer the best of both worlds: Structured data marts for reliable analytic results along with enriching data lakes with filtered (silver) and analytic (gold) models.

Qlik Cloud for Government (QCG)

Qlik Cloud Government (QCG) provides the U.S Public Sector with a modern data integration and analytics platform built for speed, security, and scale. As government agencies modernize their data and analytics platforms and adopt new emerging technologies, often cloud-first strategies can be delayed due to the lack of FedRAMP authorized offerings. Qlik has addressed this issue by providing the capabilities of Qlik Cloud within a FedRAMP authorized government infrastructure. QCG offers both Data Integration and Analytics services and is currently designated as “Authorized” for both FedRAMP “Moderate” Impact Level (IL) as well as the Department of Defense (DOD) IL 2. Qlik continues to invest heavily in supporting our U.S. Public Sector with plans on expanding our certifications for QCG.

For more information about QCG, please refer to <https://www.qlik.com/us/trust>



About Qlik

Qlik's vision is a data-literate world, one where everyone can use data to improve decision-making and solve their most challenging problems. Only Qlik offers end-to-end, real-time data integration and analytics solutions that help organizations access and transform all their data into value. Qlik helps companies lead with data to see more deeply into customer behavior, reinvent business processes, discover new revenue streams, and balance risk and reward. Qlik does business in more than 100 countries and serves over 50,000 customers around the world.

[qlik.com](https://www.qlik.com)

© 2020 QlikTech International AB. All rights reserved. All company and/or product names may be trade names, trademarks and/or registered trademarks of the respective owners with which they are associated.