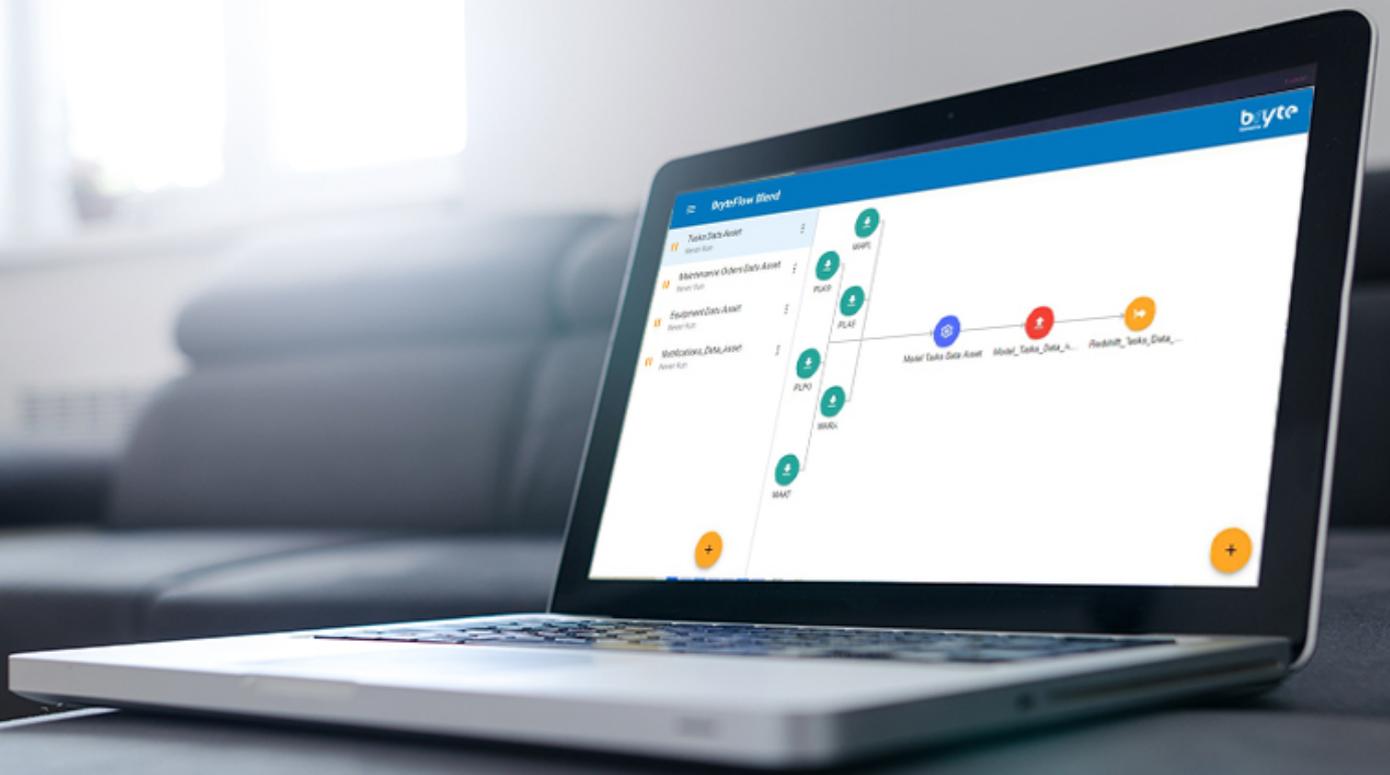


bryteflow[®]

BLEND



**SQL BASED DATA MANAGEMENT FOR AMAZON S3
WITH BRYTEFLOW BLEND**

KEY FEATURES

- Built-in SQL Workbench
- SQL code is converted to powerful Hadoop/Spark
- Complements BryteFlow Ingest
- Organize in multiple S3 folders
- Multiple levels of data maturity
- Versioning
- Monitoring and alerting
- Integration with various AWS services
- Point-and-click interface
- Auto scaling ELT
- No more bottleneck queries

BENEFITS

- Architecture enables Concurrency, Elasticity and Separation of Storage and Compute
- Spend less time monitoring ETL machines
- Cost-effective
- Use the power of Hadoop and Spark with zero coding
- Easy to use SQL interface

bryteflow[®]

BLEND

As more organisations embrace the cost-effective and scalable AWS S3 object storage as the innovation proving ground for analytics, they are looking to put this limitless power in the hands of analysts. But how will analysts transform data on a powerful Hadoop platform when they are so used to SQL and databases? How do they escape from the tyranny of data preparation and accelerate time-to-analytics?

BryteFlow addresses these challenges by providing a point-and-click SQL editor to harness the power of AWS S3. In comparison to other storage solutions, AWS S3 cents-per GB economics is too hard to ignore in an era of endless growth of data. S3 has become the de facto object storage standard for storing all spectrum of data, from raw to normalised, structured and unstructured. Customer's data from websites, call centre, CRM and ERP, each system with its own data model can now be stored in S3 and analysed in a unified manner.

Using ELT instead of ETL

With today's infrastructure technologies using the cloud, systems can now support large storage and scalable compute. Therefore, a large expanding data pool and fast processing is virtually endless for maintaining all the extracted raw data. BryteFlow Blend uses the new ELT approach which loads the data into an S3 bucket before transforming it using the power of the data lake instead of extra ELT machines that need to be managed.

This gives a few advantages like the ability to retain raw data next to the transformed data, no more bottleneck queries and auto scaling machines. Plus the data assets persist in Amazon S3 but can also be optionally exported to Amazon Redshift or Amazon Aurora.

Data Lineage

Bryteflow provides a rich graphical map that outlines all the data flows and dependancies in your transformation process. This provides users with complete clarity to understand what the data means, its source, transformation journey and final endpoint.

Built-in SQL Workbench for amazon S3

BryteFlow Blend offers a full SQL workbench to build data models without the need to move data into a relational database or data warehouse. The workbench also has the ability to schedule jobs and manage complex dependencies.

SQL code is converted to powerful and scalable Hadoop/Spark

BryteFlow Blend allows data preparation on the raw data with a built in SQL editor that converts SQL to scalable Hadoop/Spark code. This means analysts don't need to learn a new complex coding language.

Organize in multiple S3 folders

With BryteFlow Blend every table in the data sources equals to one folder in the S3 bucket. And the metadata file allows you to easily find your way in the S3 bucket.

Multiple levels of data maturity

BryteFlow Blend is an optimized DaaS environment that is designed to retain multiple levels of data maturity ranging from raw data through highly processed data marts or outputs. This enables a fast time to insight by giving users access to the type of data they need.

Seamless integration with GIT for Versioning

BryteFlow Blend offers out-of-the-box options to retain a full record history using a standard type 2 approach. This method tracks historical data and replaces old records with newer versions without deleting the previous one.

Monitoring and alerting

BryteFlow Blend integrates seamlessly with other AWS services to monitor the status of all the workflows and tasks. It can also proactively alert by configuring automatic event notification messages.

Point-and-click interface

With an easy to use point-and-click interface clients can select which tables they want to transform and export to another AWS service like Redshift or Aurora.

Auto scaling ELT

Managing an ETL/ELT tool can be very time consuming and expensive. BryteFlow eliminates the need for the analyst to anticipate and request the infrastructure team to provision the necessary compute power. The software works with auto scaling EMR clusters based on the data transformation workloads. Since all raw data is available with BryteFlow Ingest, other queries can easily continue running in the same environment to aid discovery or operational initiatives and identify the best possible data transformations that match the business requirements.

No more bottleneck queries

The ELT approach enables clients to perform large jobs faster with more machines at the same cost of doing it at a slower rate with only a few machines. This eliminates bottleneck queries and offers a zero workload conflict.

Cost is the same for 10 machines running for 2 hours as 2 machines running for 10 hours.



ABOUT BRYTE

Bryte provides an out-of-the-box software driven approach to help enterprises build high performance cloud Data Lakes and Analytics Environments. Our company mission is to make Data Access, Innovation and Analytics more pervasive, cost effective and easier than ever before by blending together the latest in innovative cloud, business intelligence and data liberation technologies. We are recognized by our clients for world class expertise with Amazon Web Services, Real Time Data Lakes and Software Driven Automation.