Qlik

COMPARISON GUIDE

Top Cloud Data Warehouses for the Enterprise

How do the big four stack up?

amazon REDSHIFT

Azure Synapse Analytics

Google BigQuery

snowflake®



For modern analytics and AI needs, look to the cloud.

For decades, data warehouses have reliably stored data for enterprise analytics and reporting. But they weren't designed to handle today's explosive data growth – or the ever-changing needs of generative AI and data consumers.

Enter the cloud data warehouse. Without the constraints of physical data centers, you can quickly grow or shrink your warehouses to adjust to dynamic requirements. And with modern cloud architectures, you can combine three essential capabilities at a fraction of the cost of traditional warehousing:

1. The power of data warehousing
2. The flexibility of big data platforms
3. The elasticity of the cloud

Which cloud data warehouse is right for you?

amazon REDSHIFT

Google BigQuery

Azure Synapse Analytics

snowflake®

In this eBook, we compare the four leading cloud data warehouses and describe our proven approach to making any (or all) of them accessible, effective, and efficient for all your data users.

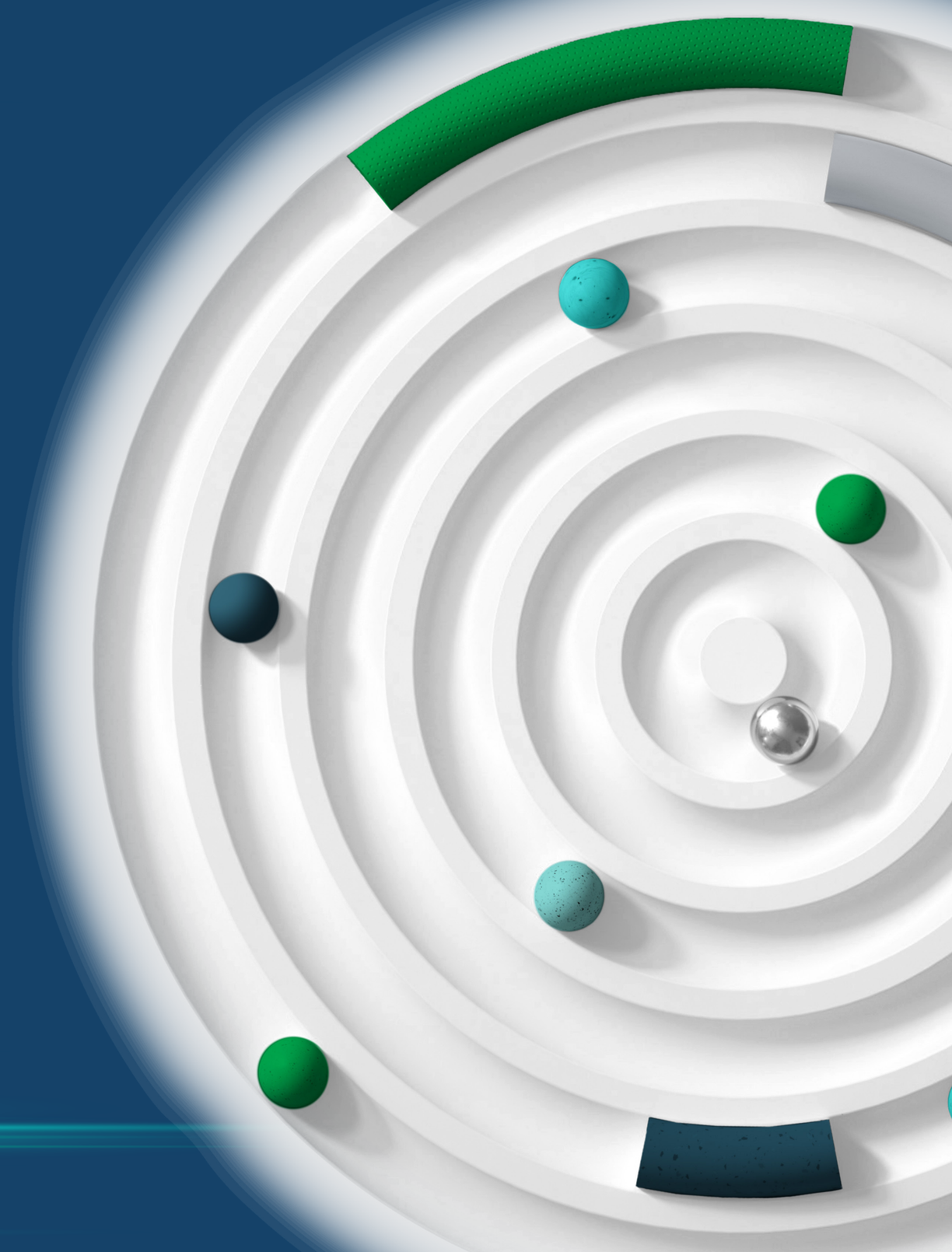
Two terms to know:

Massively Parallel Processing (MPP)

Data warehouses that support big data projects use massively parallel processing architectures to provide high-performance queries on large data volumes. MPP architectures consist of many servers running in parallel to distribute processing and input/output loads.

Columnar Data Stores

MPP data warehouses are typically columnar stores, which are the most flexible and economical for analytics. Columnar databases store and process data by columns instead of rows, which can accommodate more data in a smaller amount of memory. This arrangement makes aggregate queries – the type often used for reporting – run significantly faster.



Amazon Redshift

The first cloud data warehouse to gain widespread market share.

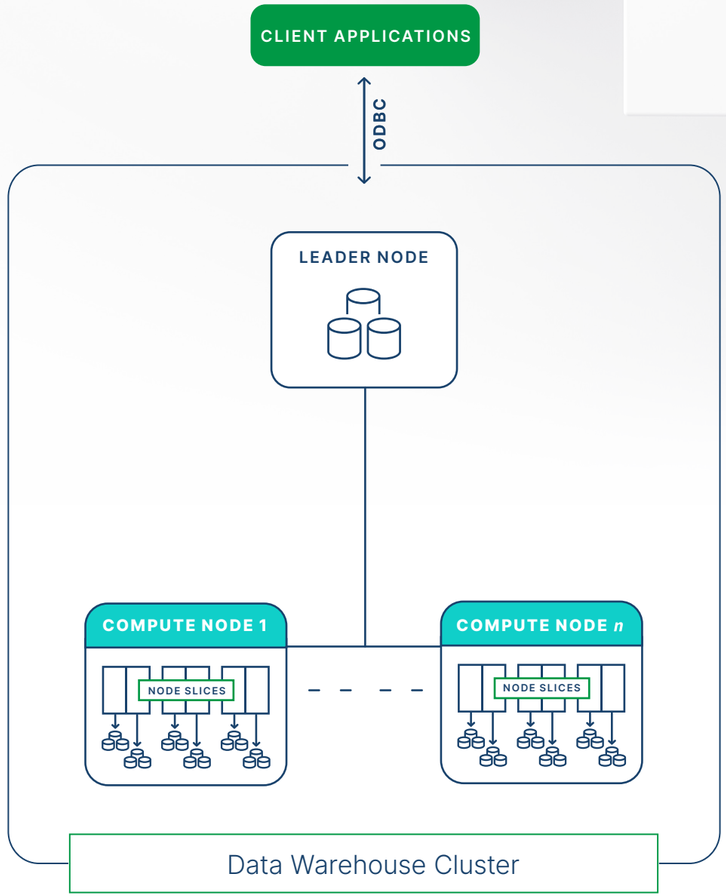
Amazon Redshift offers a powerful, scalable, and cost-effective data warehouse solution for running complex queries on large datasets, with integrations into the broader AWS ecosystem and support for a wide variety of analytics use cases.

Redshift's SQL dialect is based on PostgreSQL, which is well understood by analysts worldwide and uses an architecture familiar to many on-prem data warehouse users. You can start with just a few gigabytes of data and scale to petabytes. Regardless of the size of your data set,

Redshift delivers fast query performance using familiar SQL-based tools and business intelligence (BI) applications.

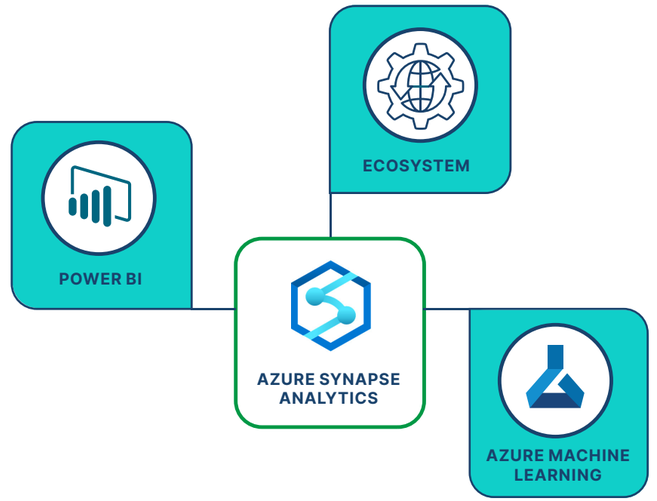
Basic Architecture

The first step to creating a Redshift data warehouse is to launch a set of nodes, called an Amazon Redshift cluster. After you provision your cluster, you upload your data set and then perform queries.



Azure Synapse Analytics

A unified experience for all your BI and ML needs.



Azure Synapse Analytics offers a single, centralized experience to ingest, prepare, manage, and serve data for your business intelligence and machine learning (ML) needs. It brings together enterprise data warehousing, AI, and big data analytics, and it gives you the freedom to query data using either serverless/on-demand or provisioned resources.

At the heart of Azure Synapse Analytics is a cloud-native, distributed SQL processing engine. It's built on the foundation of SQL Server to drive demanding enterprise data warehousing workloads. Similar to other cloud MPP solutions, Azure SQL Data Warehouse separates storage and compute, billing each separately.

Azure Synapse saves relational tables data with columnar storage and abstracts physical machines by representing compute power in the form of data warehouse units. This allows your users to seamlessly scale compute resources at will.

Taking SQL Beyond Data Warehousing

Azure Synapse aims to unify a range of analytics workloads – including data warehouses, data lakes, and ML – in a single UI. The combination of a SQL engine, Apache Spark with Azure Data Lake Storage, and Azure Data Factory gives users the power to control data warehouses/lakes and data preparation for ML tasks.

Azure Synapse allows for both vertical and horizontal scaling of the data warehouse – vertically by changing the service tier or placing the database in an elastic pool and horizontally by adding more data warehouse units.

Google BigQuery

Analyze petabytes of data at incredible speeds.

BigQuery is a fully managed, serverless data warehouse — built for Google Cloud — that automatically scales to match storage and computing power needs. Its columnar ANSI SQL database can analyze terabytes to petabytes of data at incredible speeds.

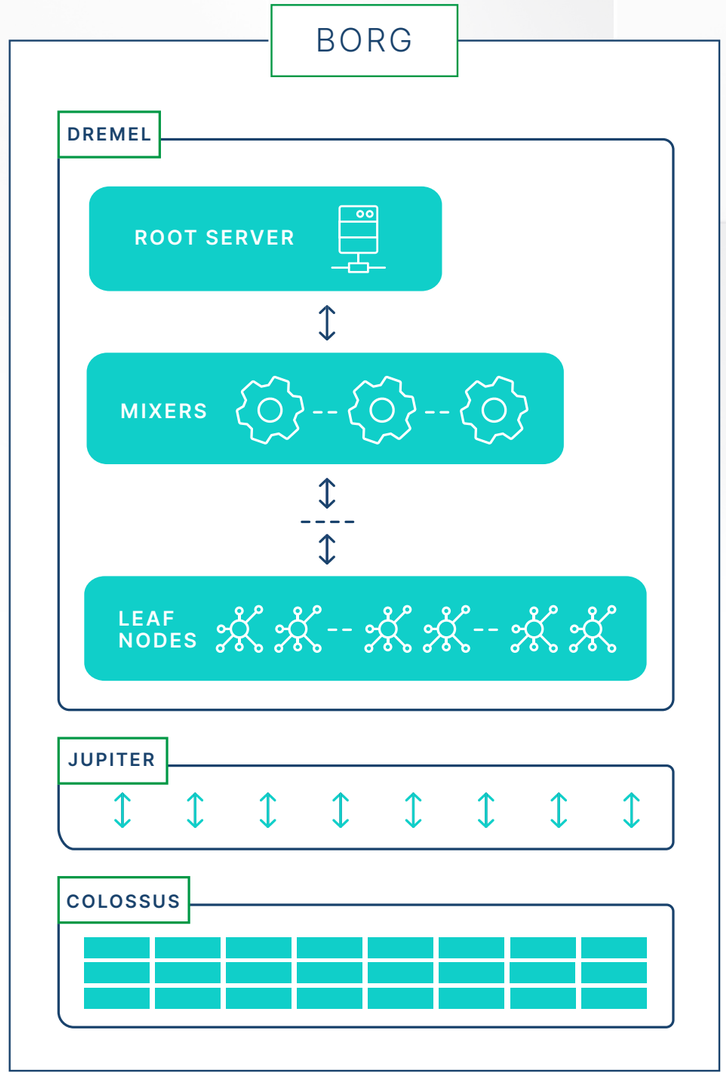
With Google BigQuery, you can:

1. Support real-time interactive dashboarding with the BigQuery BI Engine
2. Quickly build and operationalize ML models on large-scale structured or semi-structured data with simple SQL
3. Conduct geospatial data analysis using familiar SQL

Google doesn't expect you to manage your data warehouse infrastructure, which is why BigQuery hides many of the underlying hardware, database, nodes, and configuration details. And the elasticity works out of the box. To get started, you create an account with Google Cloud Platform (GCP), load a table, and run a query.

Basic Architecture

BigQuery architecture has several components. Borg is the compute layer, Colossus is distributed storage, Jupiter is the network, and Dremel is the execution engine.



Snowflake Cloud Data Platform

The first multi-cloud data warehouse.

Snowflake is a fully managed MPP cloud data warehouse that runs on AWS, Google Cloud Platform, and Microsoft Azure. When you're a Snowflake user, you can spin up as many virtual warehouses as you need to parallelize and isolate the performance of individual queries.

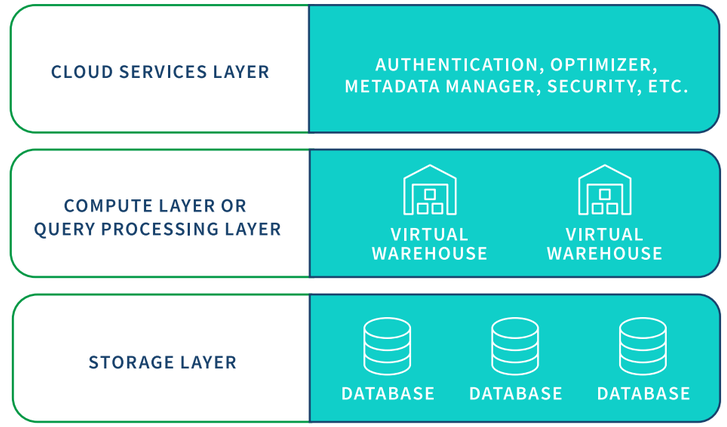
Like Google BigQuery and Microsoft Azure Synapse Analytics, Snowflake enables very high concurrency by separating storage and compute, making it possible for many warehouses to simultaneously access the same data source.

Snowflake, unlike the other data warehouses covered in this eBook, doesn't run on its own cloud. It's the first multi-cloud data warehouse available globally on AWS, GCP, and Azure. With a common and interchangeable code base, Snowflake features global data replication, which means you can move your data to any cloud, in any region – without having to recode your applications or learn new skills.

Basic Architecture

You interact with Snowflake's data warehouse through a web browser, the command line, an analytics platform, or via Snowflake's ODBC, JDBC, or other supported drivers. The platform supports ACID-compliant relational processing and has native support for document store formats such as JSON, Avro, Optimized Row Columnar (ORC), Parquet, and XML.

Snowflake's hybrid architecture is separated into three distinct layers:



Top Cloud Data Warehouses at a Glance

Features	amazon REDSHIFT	Azure Synapse Analytics	Google BigQuery	snowflake®
Concurrency	✓	✓	✓	✓
Durability	✓	✓	✓	✓
MPP	✓	✓	✓	✓
Columnar	✓	✓	✓	✓
Separates Storage and Compute	✓	✓	✓	✓
Foreign Keys	✓	✓	✗	✓
Automation	✗	✗	✗	✗
Multi-Cloud	✗	✗	✗	✓
Transaction	ACID	ACID	ACID	ACID
Elasticity	Automatic	Manual and Automatic	Automatic	Automatic
Query Language	Amazon Redshift SQL	TSQL	Standard SQL 2011 & BigQuery SQL	Snowflake SQL
Pricing	Pay-as-you-go based on nodes	Pay-per-query or provisioned capacity	Pay-per-query (data scanned)	Pay-per-usage (compute and storage separately)
Free Trial	✓	✓	✓	✓

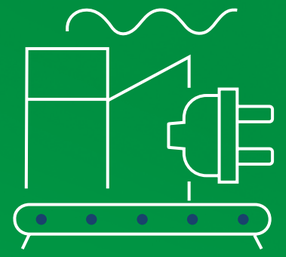
Deliver real-time data and automate transformations on any cloud.

Get set up for success with end-to-end data integration.

The cloud is now the de facto deployment model for modern analytics and AI — but enabling this doesn't just happen on its own. You need approaches and technologies that set you up to iterate quickly and deliver reliable, AI- and analytics-ready data to users fast.

That's where we come in. Qlik Talend® automates the entire data warehouse lifecycle, from ingestion through delivery. Our model-driven approach helps your data engineers design, deploy, manage, model, and catalog purpose-built cloud data warehouses in record time.

Add Qlik Talend to any cloud data warehouse you choose, and you'll be able to automate your entire data pipeline.



Need APIs? Meet the Qlik Connector Factory.

With an in-house R&D team dedicated to developing standard APIs, we're continually expanding access to and delivery of data from hundreds of SaaS applications and data sources. Our customers already benefit from over 500 existing connectors, and we're continuing to add more each year.

Learn More

A simpler, faster data warehouse lifecycle.

Qlik Talend automates the entire data warehouse lifecycle to accelerate the availability of your AI- and analytics-ready data.

The many benefits include:

1. **Real-time data ingestion and updates**
   A simple and universal solution for continually ingesting your enterprise data into popular cloud data warehouses in real time — and use it all to support AI and machine learning.
2. **Advanced transformations**
   Transform your data with powerful automation to conform data to a variety of AI or analytics use cases or architectures.
3. **Automated workflow**
   A model-driven approach for continually refining your data warehouse operations
4. **Trusted, enterprise-ready data**
   Easily discover, remediate, and share trusted data with powerful data quality and governance features — including profiling and lineage tracking.

Ready to move to agile data warehousing? We're ready to help.

About Qlik

Qlik transforms complex data landscapes into actionable insights, driving strategic business outcomes. Serving over 40,000 global customers, our portfolio leverages advanced, enterprise-grade AI/ML and pervasive data quality. We excel in data integration and governance, offering comprehensive solutions that work with diverse data sources. Intuitive and real-time analytics from Qlik uncover hidden patterns, empowering teams to address complex challenges and seize new opportunities. Our AI/ML tools, both practical and scalable, lead to better decisions, faster. As strategic partners, our platform-agnostic technology and expertise make our customers more competitive.

qlik.com

© 2024 QlikTech International AB. All company signs, names, logos, product names, and/or trade names referenced herein, whether or not appearing with the symbols ® or ™, are trademarks of QlikTech Inc or its affiliates. All other products, services, and company names mentioned herein may be trademarks of their respective owners and are acknowledged as such. For a list of Qlik trademarks please visit: https://www.qlik.com/us/legal/trademarks

1-WEB-08