SNOWFLAKE VS REDSHIFT: WHAT'S THE DIFFERENCE & HOW TO CHOOSE?



Snowflake and AWS Redshift are two popular cloud-based data warehousing platforms that offer outstanding performance, scale, and business intelligence capabilities.

In this article, we'll help you decide whether AWS Redshift or Snowflake is right for you. We'll compare these two solutions based on their similarities, differences, and use cases. We also highlight how each platform addresses common challenges faced by businesses looking to implement a data warehouse.

Choose the right data warehouse

AWS Redshift and Snowflake are two popular data warehouses. How do you know which one's right for you? First, let's briefly look at data warehouses in general.

Data warehouses (DWH) are large repositories of data, collected from different data sources, which organizations typically use for analytical insights and business intelligence. An efficient DWH relies on an architecture that offers consistency by collecting data from different operational databases, then applying a uniform format for easier analysis and quicker insights.

One of the fundamental purposes of DWH is to enable quick access to historical data and context, thereby helping decision makers to optimize strategies and improve bottom lines.

Implementing the right data warehousing solution is key to gaining a competitive advantage in

today's data-centric business world. When leveraging an efficiently provisioned business intelligence framework, a data warehouse supports business outcomes such as:

- Increased bottom line
- Efficient decision making
- Enhanced customer service
- Improved analytics

The most important characteristics of any efficiently designed data warehouse is that it:

- Has consistent schemas across different tables that return expected results against a query.
- Supports multi-table querying out of the box, which allows users to generate ad hoc reports without writing custom code or creating custom table views.

Some key factors to consider when selecting a warehousing platform include:

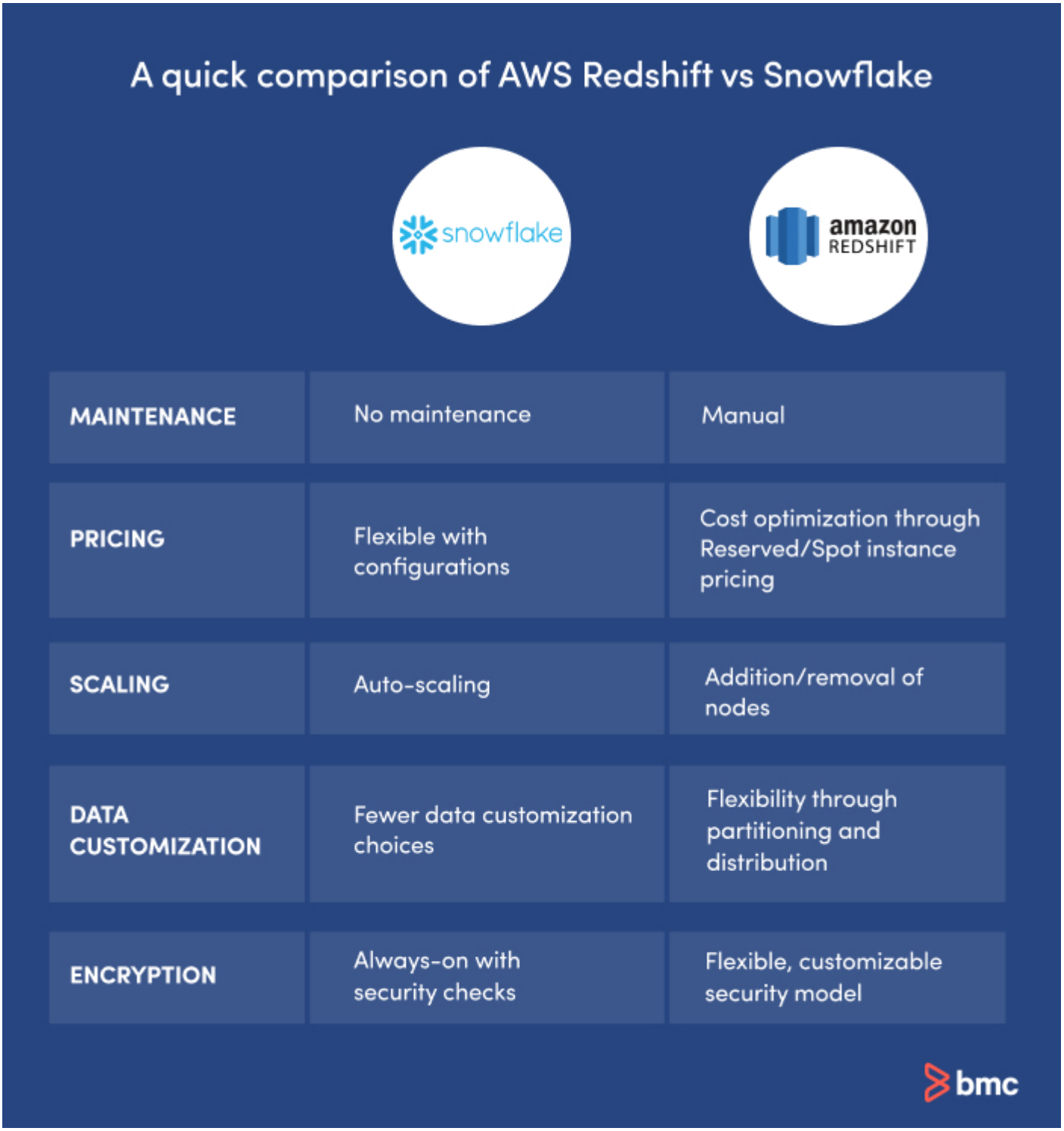
- Business goals
- Cost models
- Simplicity of integration
- Cloud readiness
- Adherence to security and compliance standards

Snowflake vs AWS Redshift

Both platforms offer similar core functionalities, such as:

- Relational management
- Security
- Scalability
- Cost efficiency

The key differences, however, are their pricing models, deployment options, and user experience.



What is Snowflake?

Snowflake is a cloud-based, software as a service (SaaS) data platform that allows:

- Secure data sharing.
- Unlimited scaling.
- A seamless multi-cloud experience.

The platform relies on a virtual warehouse framework that leverages third-party cloud compute resources such as AWS, Azure, or GCP. The option to choose high-performance cloud platforms allows real-time auto scaling to organizations who are looking to run faster workloads and process large query volumes on the elastic cloud.

As compared to legacy DWH solutions, Snowflake offers a non-traditional approach to data warehousing by abstracting compute from storage. That means data can reside in a central repository, while compute instances are sized, scaled, and managed independently.

Snowflake manages all aspects of data administration for a simpler, more flexible warehousing solution that provides various capabilities of enterprise offerings.

The Snowflake analytics platform leverages a custom SQL query engine and three-layer architecture to support real-time analytics of streaming big data. Its flexible architecture allows users to build their own analytical applications without having to learn new programming languages.

(Check out our Snowflake Guide.)

Benefits of Snowflake



When to use Snowflake

Snowflake is considered the perfect data warehouse solution for situations when:

- The query load is expected to be lighter.
- The workload requires frequent scaling.
- Your organization requires an automated, managed solution with zero operational overhead to manage the underlying platform.

Now, let's turn to Redshift.

What is AWS Redshift?

AWS Redshift is a DWH platform that uses cloud-based compute nodes to enable large-scale data analysis and storage. The platform employs column-oriented databases to connect business intelligence solutions with SQL-based query engines. By leveraging PostgreSQL and Massive Parallel Processing (MPP) on dense storage nodes, the platform delivers quick query outputs on large data sets.

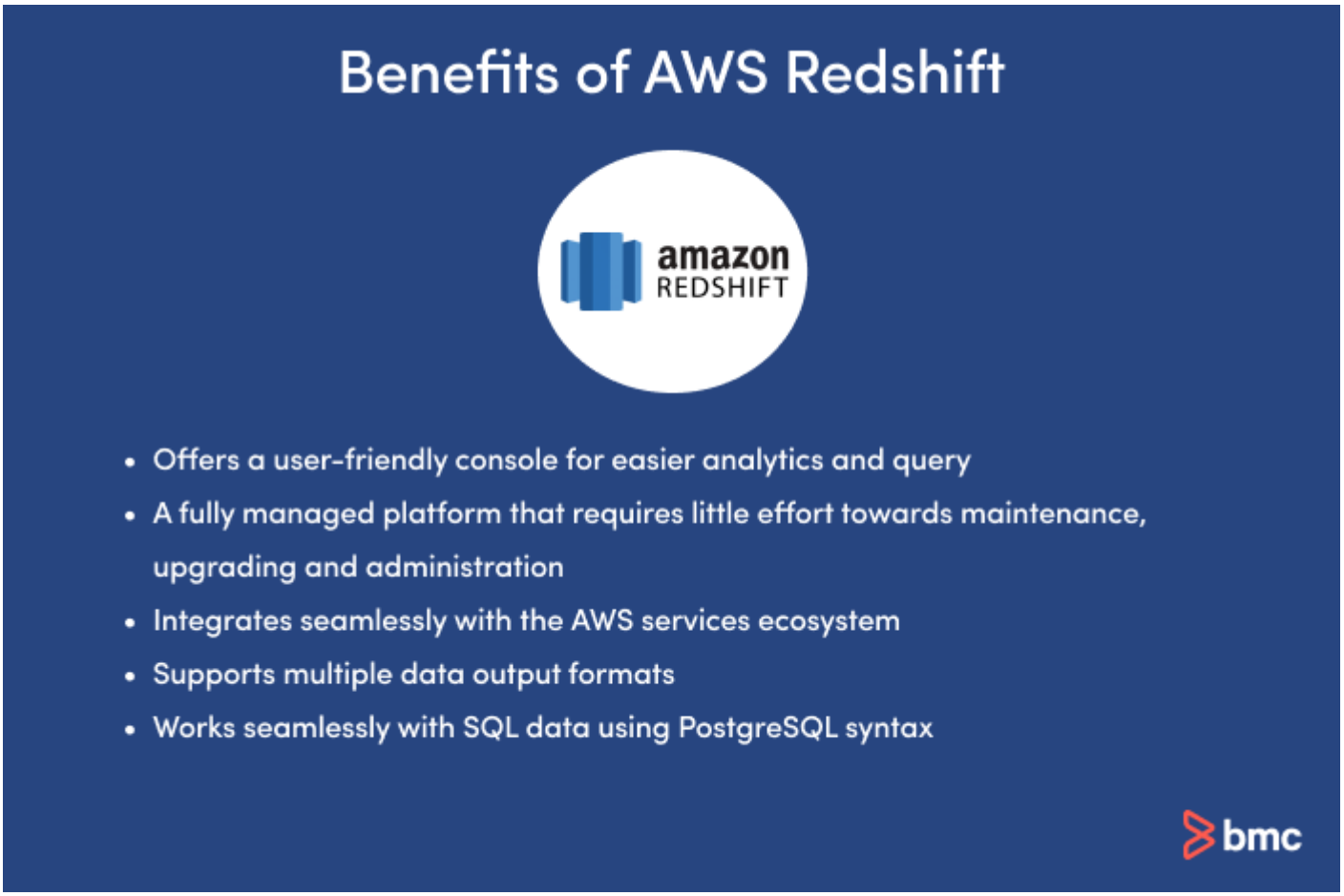
Redshift offers faster query processing and multiple options for efficient management of its clusters. These include:

- Interactively using the AWS CLI or AWS Redshift Console
- Amazon Redshift Query API
- AWS Software Development Kit

AWS Redshift is a fully managed warehousing platform that allows organizations to query and combine petabytes of data with optimized price performance. The Advanced Query Accelerator (AQUA) offers a cache that boosts query operation performance by up to 10 times, allowing businesses to gain new insights from every data point in the application/system.

(Explore our hands-on AWS Redshift Guide.)

Benefits of AWS Redshift



When to use Redshift

AWS Redshift is considered the perfect data warehouse solution for situations when:

- Your organization is already using AWS services.
- Workloads run structured data.
- The application has a high query load.

A quick comparison of AWS Redshift vs. Snowflake

Let's look at the clear differences between the two options.

Feature	Snowflake	AWS Redshift
Maintenance	Complete SaaS offering with no maintenance required	Requires some manual maintenance
Compute and Storage	Separates compute from storage, allowing flexible pricing and configuration	Cost optimization through Reserved/Spot instance pricing
Scaling	Instantaneous auto-scaling	Requires addition/removal of nodes for scaling
Data Customization	Supports fewer customization options	Supports more flexibility through features like partitioning and distribution
Security	Always-on encryption with strict security enforcement	Offers a flexible and customizable security model

Similarities between Snowflake and Redshift

- Both support MPP for faster performance.
- Both the platforms connect BI solutions to databases using column-oriented databases.
- Data in both warehouses is accessed using SQL-based query engines.
- Both Snowflake and Redshift are designed to abstract data management tasks, so users can easily gain insights and improve system performance using data-driven decisions.

Choosing Snowflake or Redshift

In the modern data-driven world, DWH solutions allow organizations to store large sets of operational data and make holistic analytical decisions to improve system performance.

DWH are designed to store vast amounts of structured or semi-structured data, to provide fast retrieval times and easy analytics.

Whichever platform you choose, enterprise orchestration for Snowflake with Control-M ensures your warehouse loading, transformation, and reporting jobs are scheduled, monitored, and governed within your broader data pipeline — regardless of cloud provider.

While both the platforms are highly popular, the choice between the two platforms depends on business demands, resources, bundled services, and specific use cases.