DATA INTEGRATION & FEDERATION

Analytics in 15

Build Apache Spark Apps and Integrate Them with Amazon Redshift

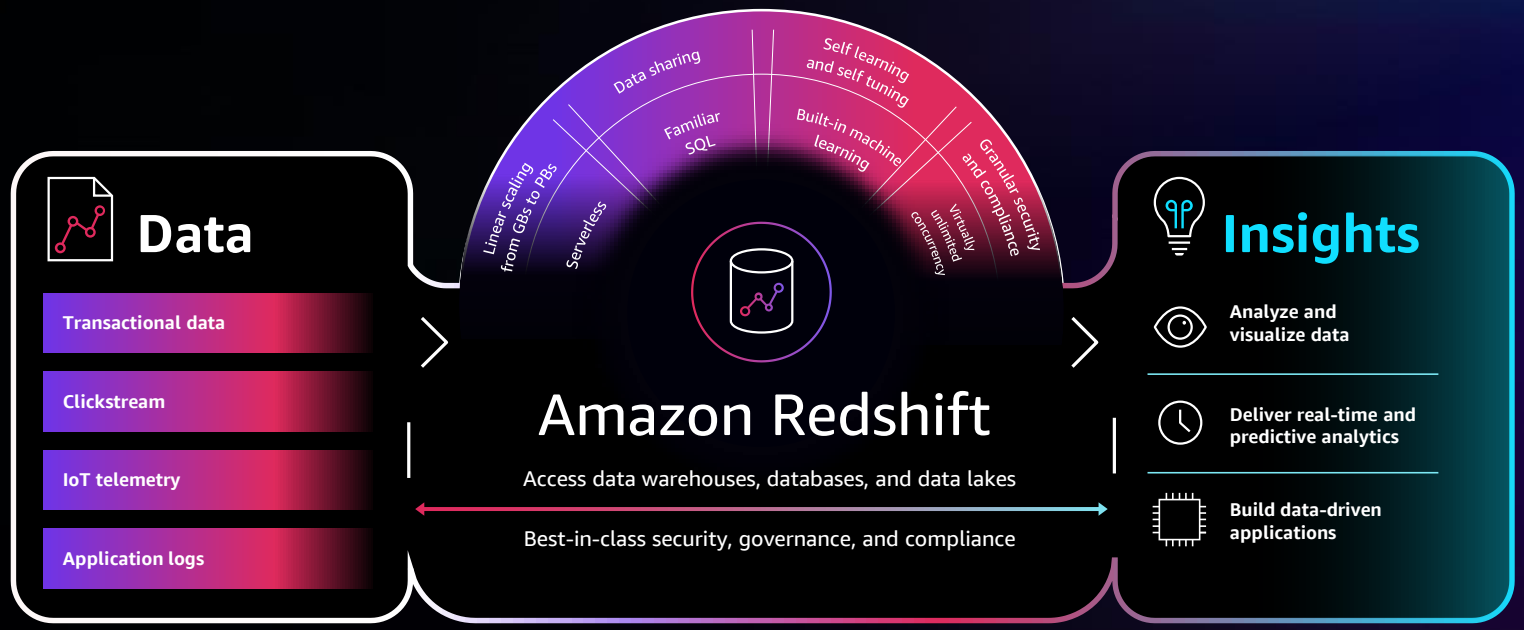
Gagan Brahmi

Sr. Analytics Solutions Architect
Amazon Web Services

© 2023, Amazon Web Services, Inc. or its affiliates.

FULLY MANAGED, PETABYTE SCALE CLOUD DATA WAREHOUSING

Amazon Redshift



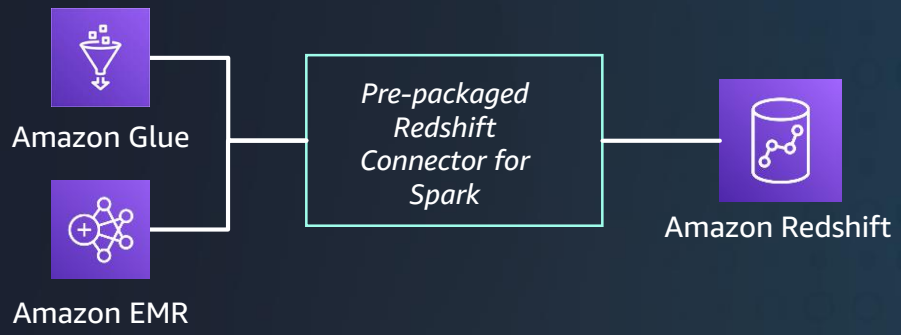
- **Price-performance at any scale**
- **Easy analytics and machine learning for everyone**
- **Analyze all your data**
- **Secure and reliable**

© 2023, Amazon Web Services, Inc. or its affiliates. All rights reserved.

[GENERAL AVAILABILITY]

Amazon Redshift integration for Apache Spark

Simplify and speed up Apache Spark applications accessing Amazon Redshift data from AWS analytics services



Author Apache Spark applications using Java, Python, Scala, with access to rich, curated data in your data warehouse

No manual setup and maintenance of uncertified versions of Spark-Redshift open-source connectors

Improved performance with only relevant data moved from Redshift to consuming applications

Improved security with IAM-based credentials

aws

© 2023, Amazon Web Services, Inc. or its affiliates.

Demo

© 2023, Amazon Web Services, Inc. or its affiliates.

Example Corp.

Example Corp. is a fictitious company, which provides ticket sales and distribution for live events across the country. An important aspect of the business is to provide tickets to end-customers as competitive prices. They plan to modernize the architecture to provide forecast for future sales which will also help determine ticket sale prices.

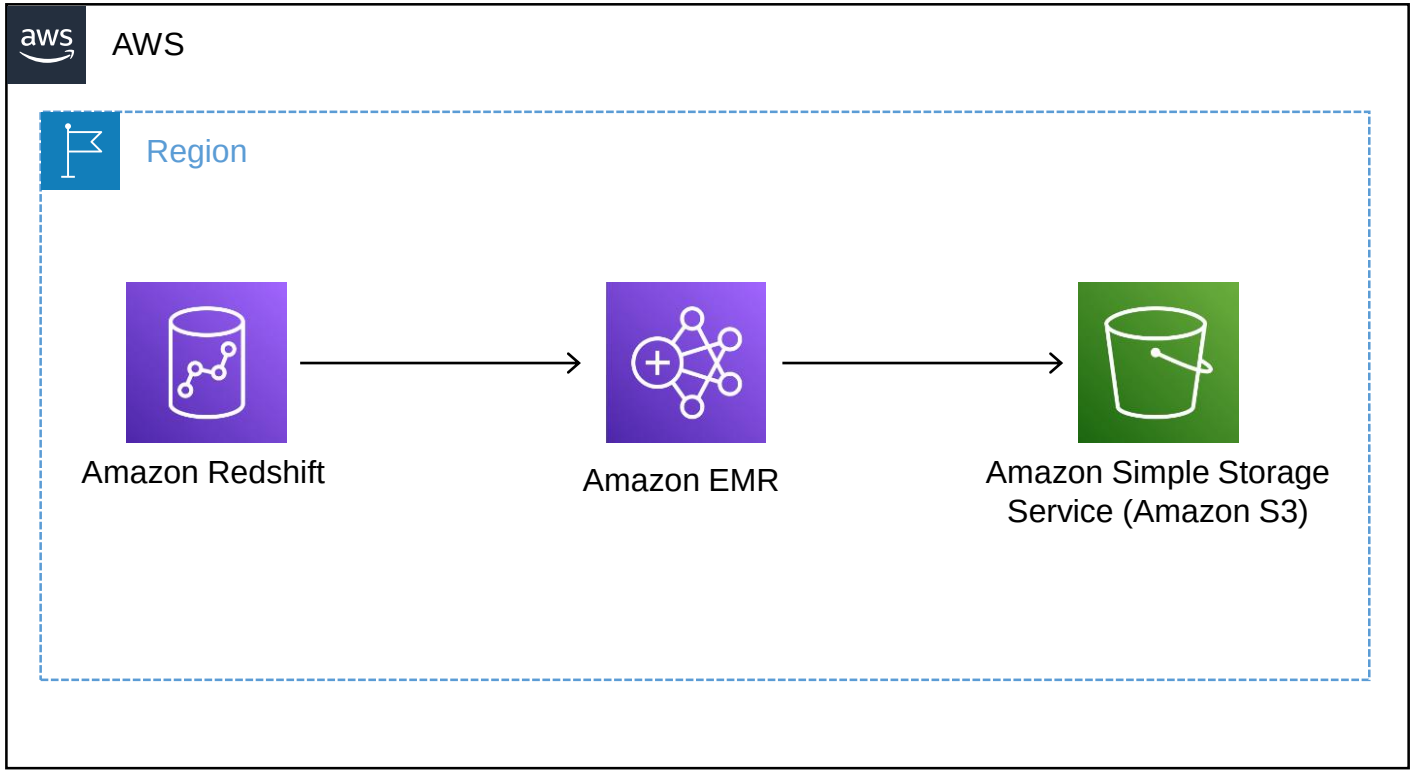
© 2023, Amazon Web Services, Inc. or its affiliates.

Consume data using Amazon EMR

Consume data from Amazon Redshift using Apache Spark with Amazon EMR

Persona: Data Engineer

Consume Amazon Redshift data using Apache Spark with Amazon EMR



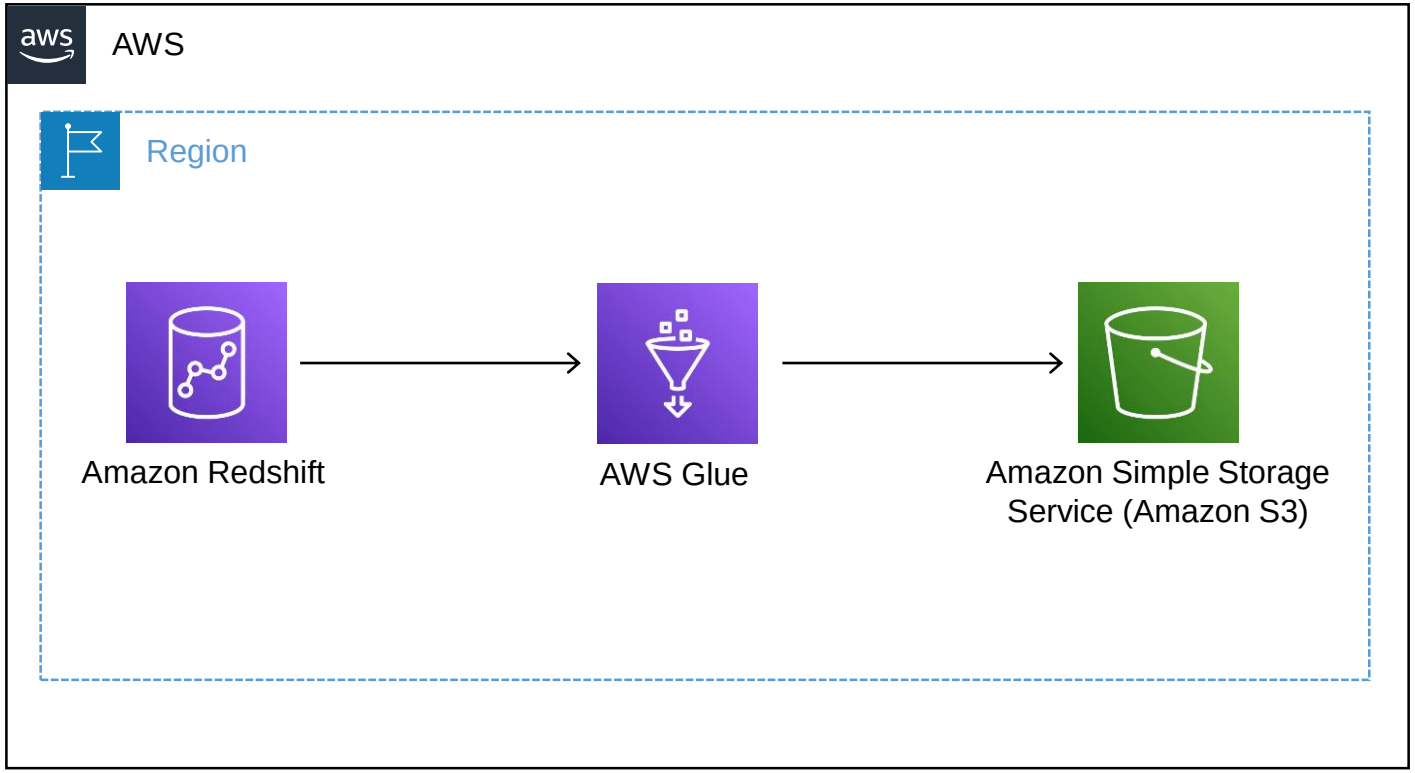
© 2023, Amazon Web Services, Inc. or its affiliates.

Consume data using AWS Glue

Consume data from Amazon Redshift using Apache Spark with AWS Glue

Persona: Data Engineer

Consume Amazon Redshift data using Apache Spark with AWS Glue



© 2023, Amazon Web Services, Inc. or its affiliates.

Demo Recap

- Get started in minutes
- No maintenance or operational overhead
- Secure access through AWS IAM-based authentication
- Faster performance with predicate pushdown capabilities

© 2023, Amazon Web Services, Inc. or its affiliates.

Amazon Redshift Streaming Ingestion

Ingest streaming data into your data warehouse for real-time analytics

Adobe LiveMe

© 2023, Amazon Web Services, Inc. or its affiliates.

PREVIEW

Amazon Aurora Zero-ETL to Amazon Redshift

Boost analytical query performance

No ETL pipelines

Real-time analytics and machine learning on transactional data

Consolidated insights from multiple Aurora databases

© 2023, Amazon Web Services, Inc. or its affiliates.

PREVIEW

Auto-copy from Amazon S3

Simplified & automated file ingestion from Amazon S3 into Amazon Redshift

Simple, low-code data ingestion

Avoid re-ingestion and manual tracking of loaded files

Easily convert your existing COPY statements into automatic ingestion jobs

Automatic ingestion of new data from Amazon S3 based on user defined configurations

© 2023, Amazon Web Services, Inc. or its affiliates.

AWS Data Exchange is fully integrated with Amazon Redshift data sharing

Goldman Sachs

moderna

Find and subscribe to **third-party data**

No data copies or movement

Automated **licensing** and contracting

Secure and **governed collaboration** with your partners

aws

© 2023, Amazon Web Services, Inc. or its affiliates.

Thank you!

To learn more, please visit:

- ✓ Amazon Redshift webpage: aws.amazon.com/redshift
- ✓ Amazon Redshift new feature page: aws.amazon.com/redshift/whats-new/
- ✓ Analytics webpage: aws.amazon.com/big-data/datalakes-and-analytics/
- ✓ Amazon Redshift Integration for Apache Spark: aws.amazon.com/redshift/features/integration-for-apache-spark/
- ✓ Amazon Redshift Integration for Apache Spark Blog: aws.amazon.com/blogs/big-data/simplify-and-speed-up-apache-spark-applications-on-amazon-redshift-data-with-amazon-redshift-integration-for-apache-spark/

© 2023, Amazon Web Services, Inc. or its affiliates.