AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

ANT308

Building analytics at scale with Amazon Athena

Saurabh Bhutyani
Sr. Analytics Specialist
AWS

Sreekanth Munigati
Sr. Analytics Specialist
AWS

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

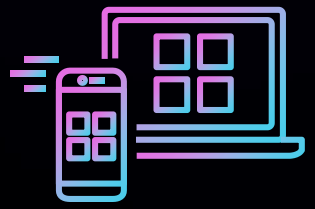
- Amazon Athena
- Common use cases
- Data integration
- Machine learning
- Mutable data lakes
- Data governance
- Q&A

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

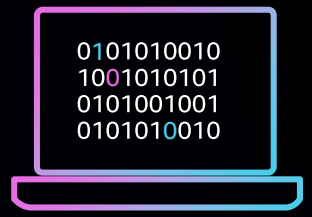
Leading companies expect more value from their data



Growing exponentially



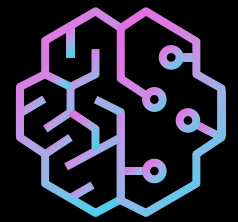
From new sources



Increasingly diverse formats



Used by many internal roles

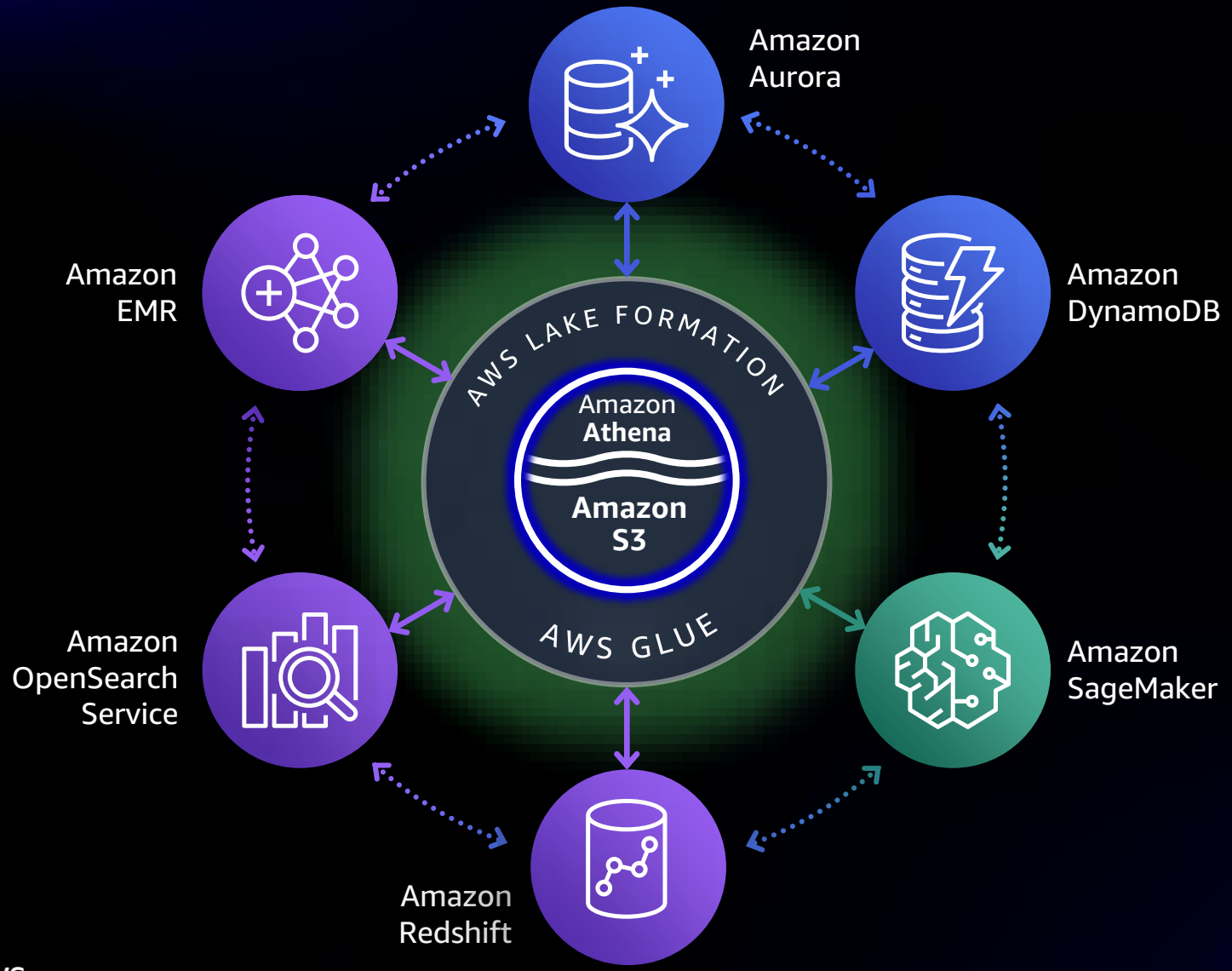


Analytics and machine learning

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data architecture on AWS



SCALABLE DATA LAKES

PURPOSE-BUILT
DATA SERVICES

SEAMLESS
DATA MOVEMENT

UNIFIED GOVERNANCE

PERFORMANT AND
COST-EFFECTIVE

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon Athena

Query data in S3 using SQL



Serverless

Quickly query S3 data without managing infrastructure, and pay only for the queries you run

Open and standard

Use ANSI SQL for querying with support for Parquet, CSV, JSON, Avro and other standard data formats

Fast interactive performance

Parallel execution to deliver most results within seconds, with no cluster management required

Cost effective

Pay only for queries run and save 30–90% by compressing, partitioning, and converting your data into columnar formats

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Simple pricing

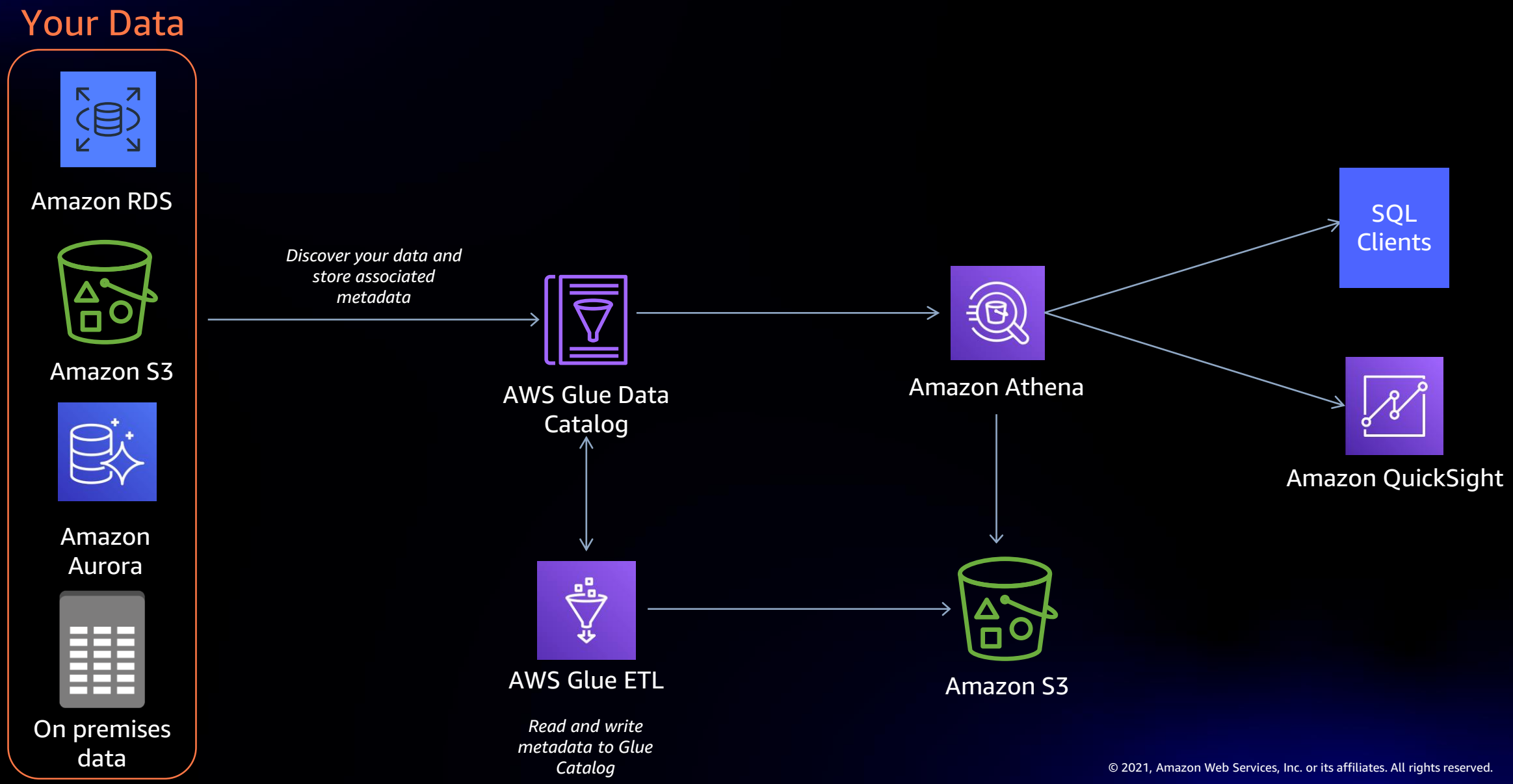
- DDL operations – No cost to customer
- SQL operations – No cost to customer
- Query concurrency – No cost to customer
- Data scanned - $5 / TB
- Standard S3 rates for storage, requests, and data transfers apply

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Athena use cases

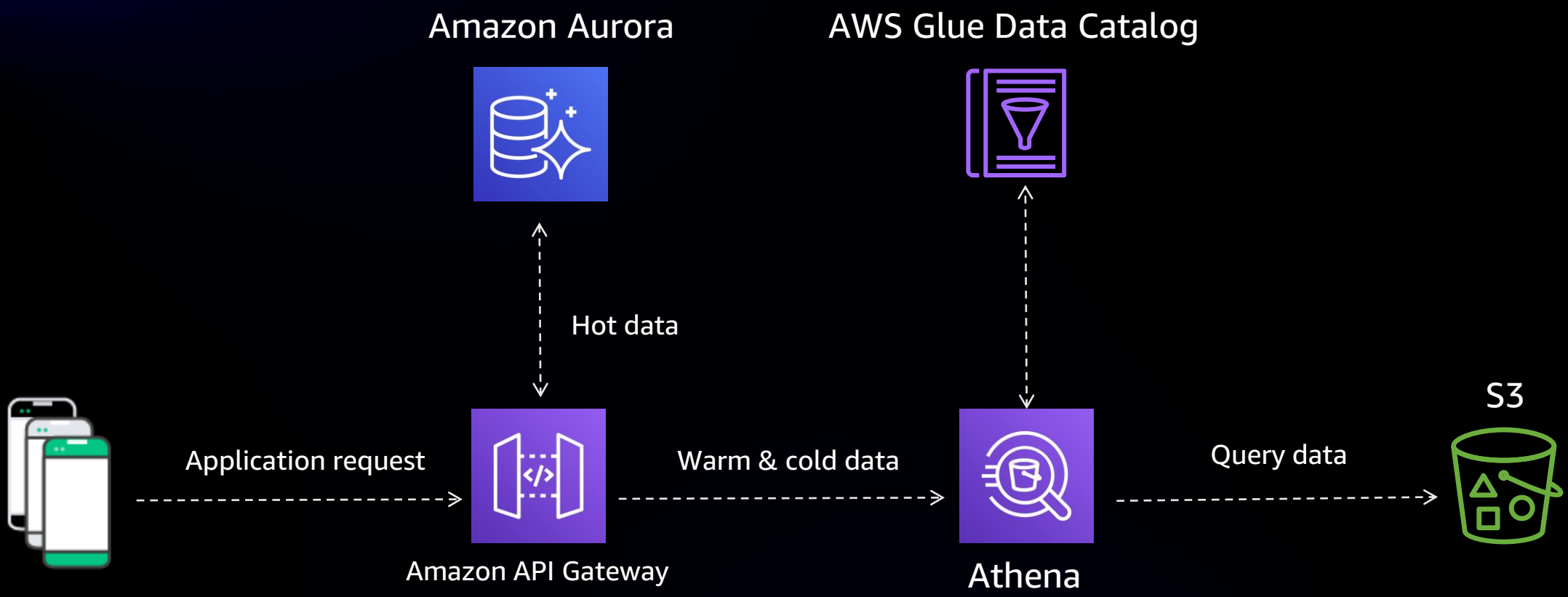
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data exploration



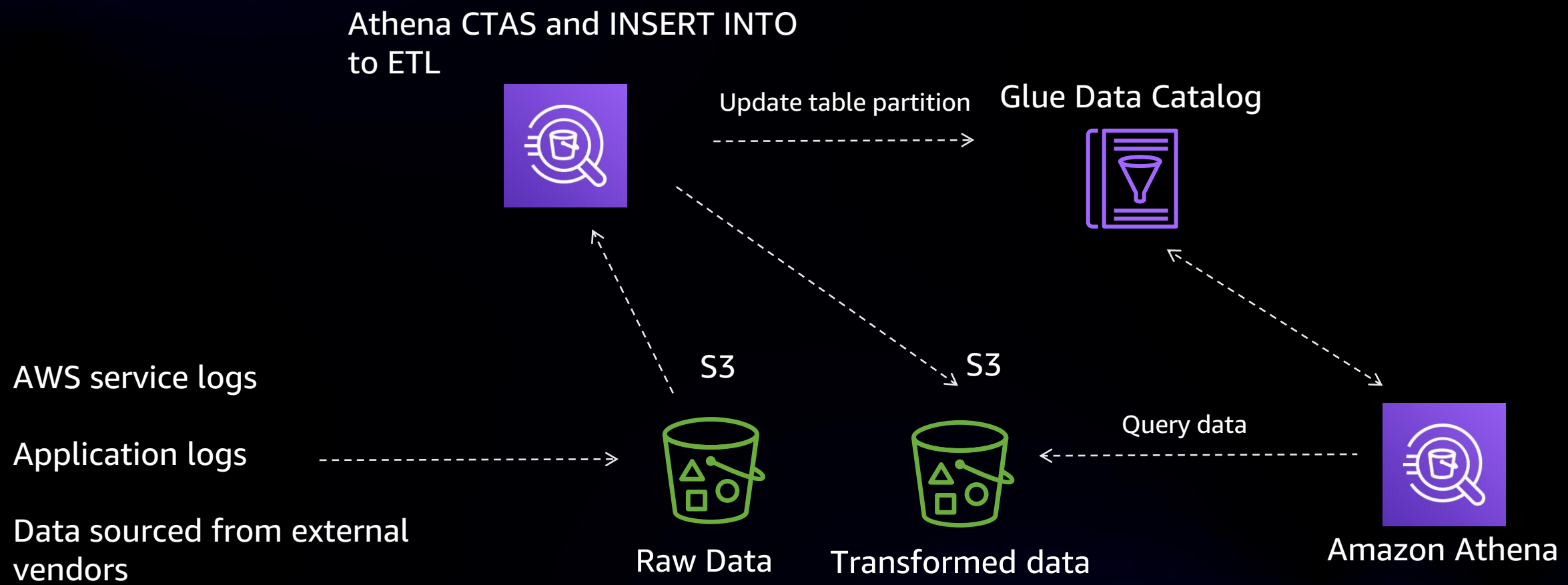
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

SaaS



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ETL and query

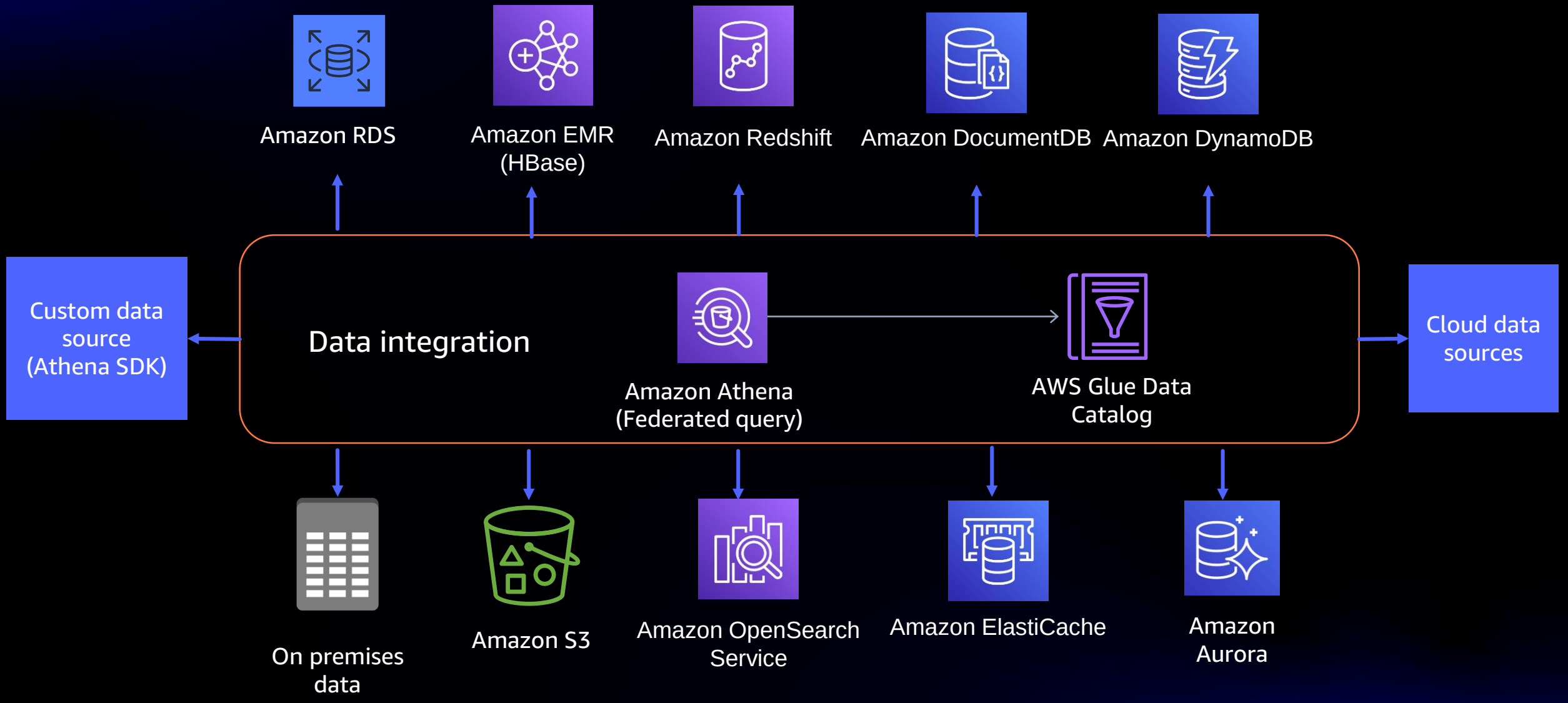


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data integration

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data integration

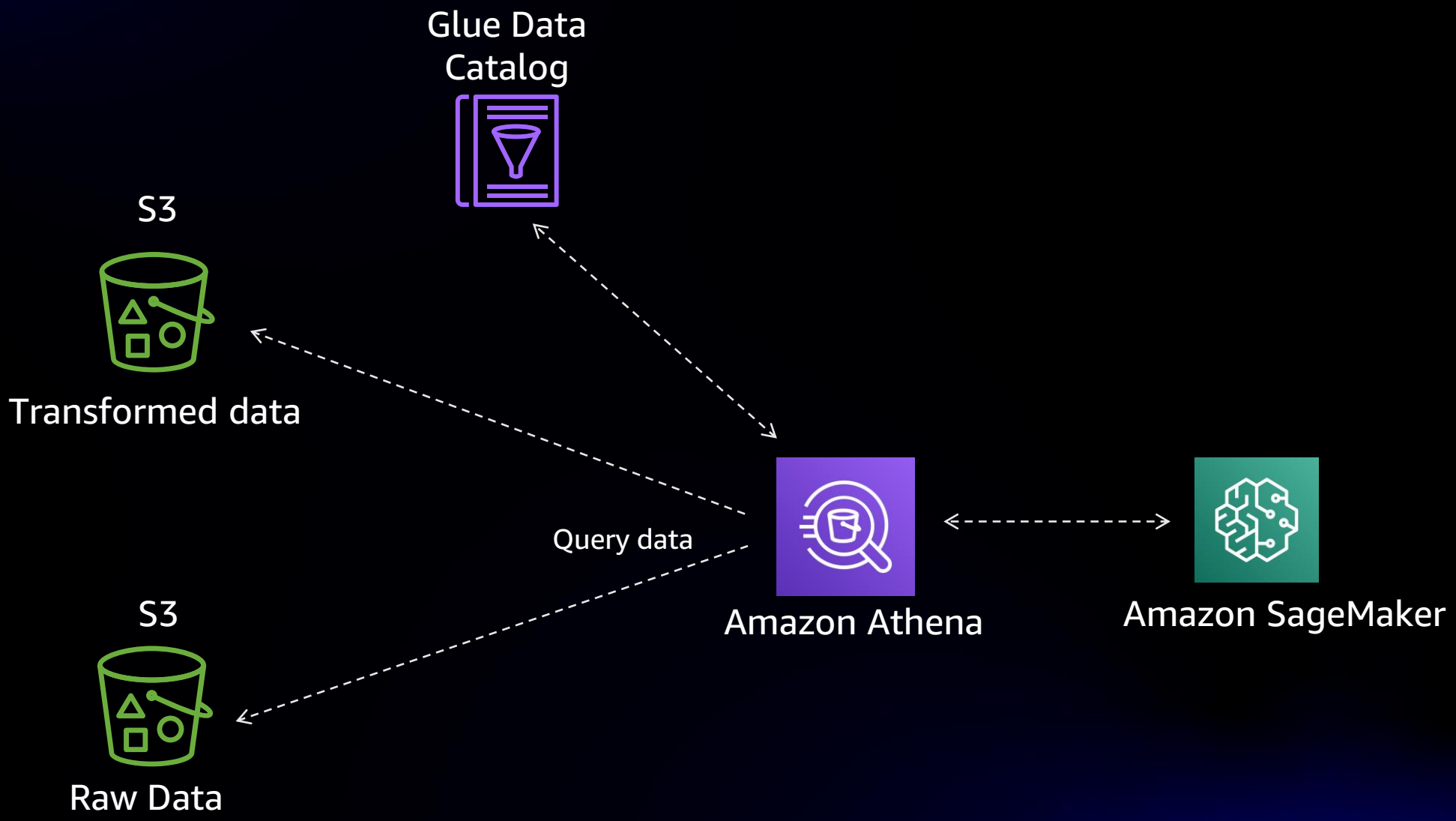


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine learning with Amazon Athena

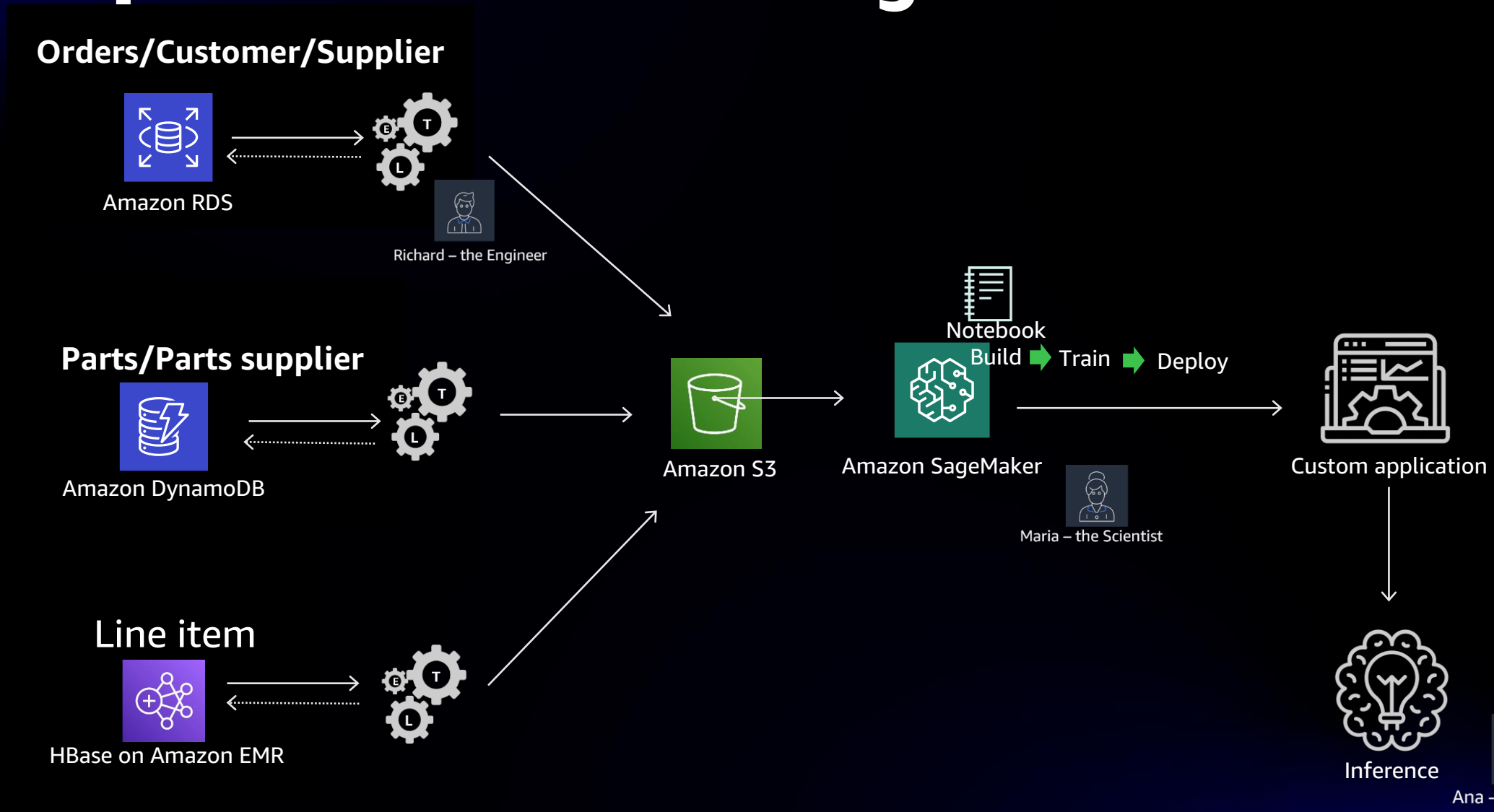
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data science exploration and feature engineering



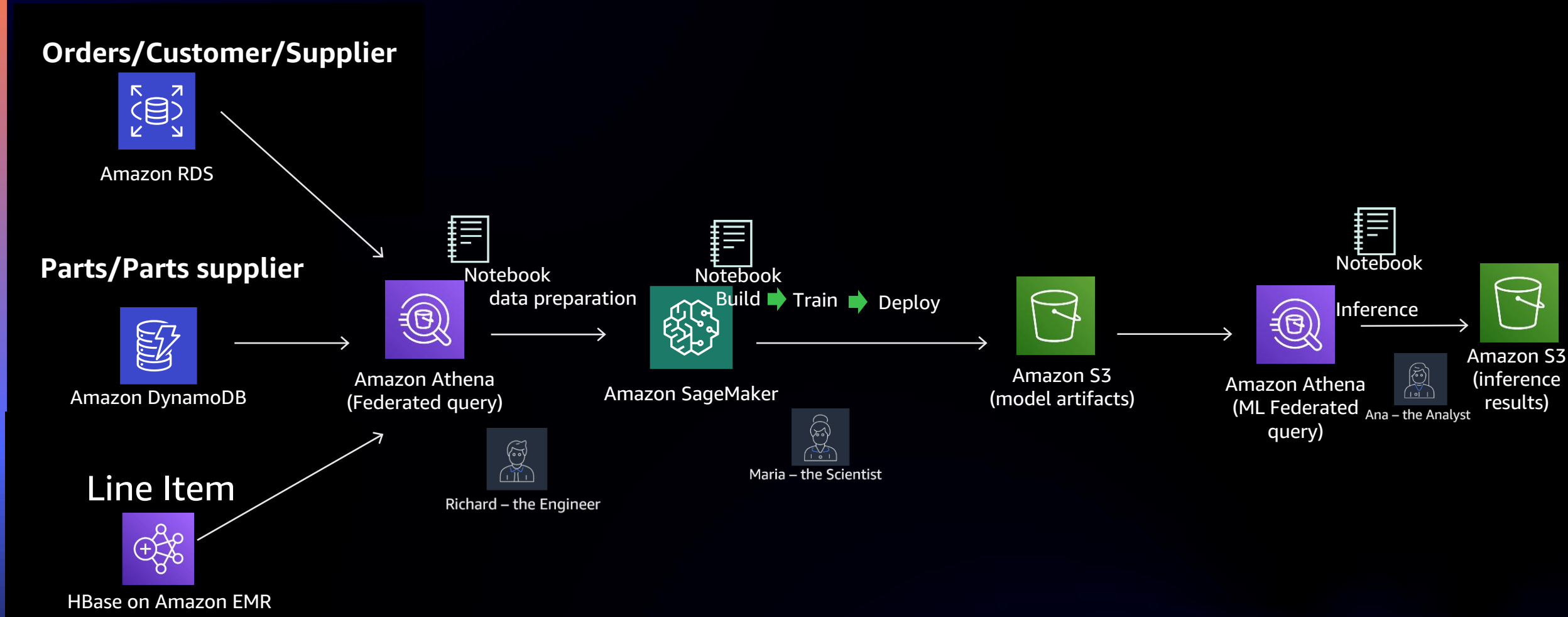
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Example: Machine learning architecture



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Operationalize ML models with Athena



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Mutable data lakes with Amazon Athena

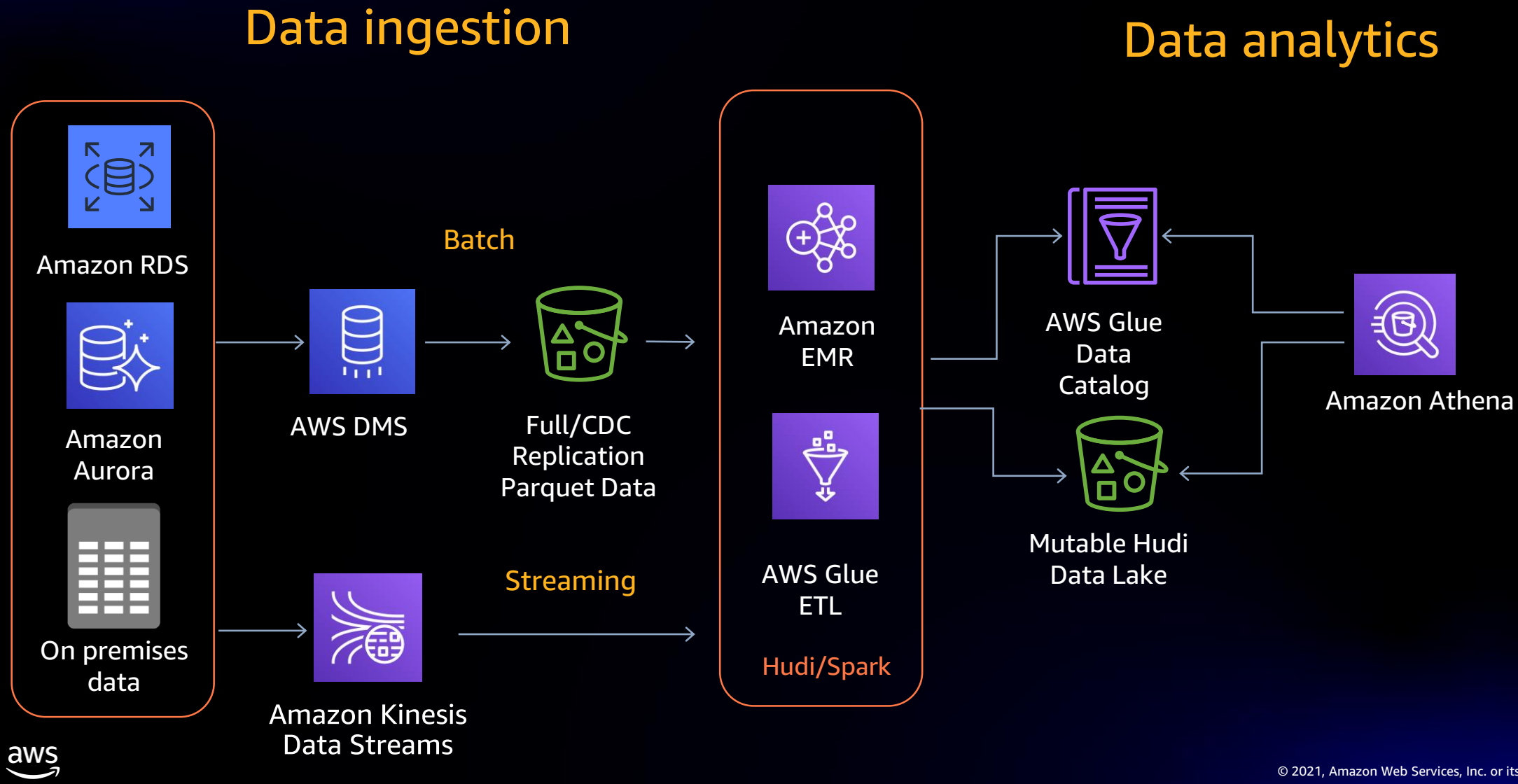
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Mutable data lakes using…

- Apache Hudi
- AWS Lake Formation governed tables (Preview)
- Apache Iceberg (Preview)

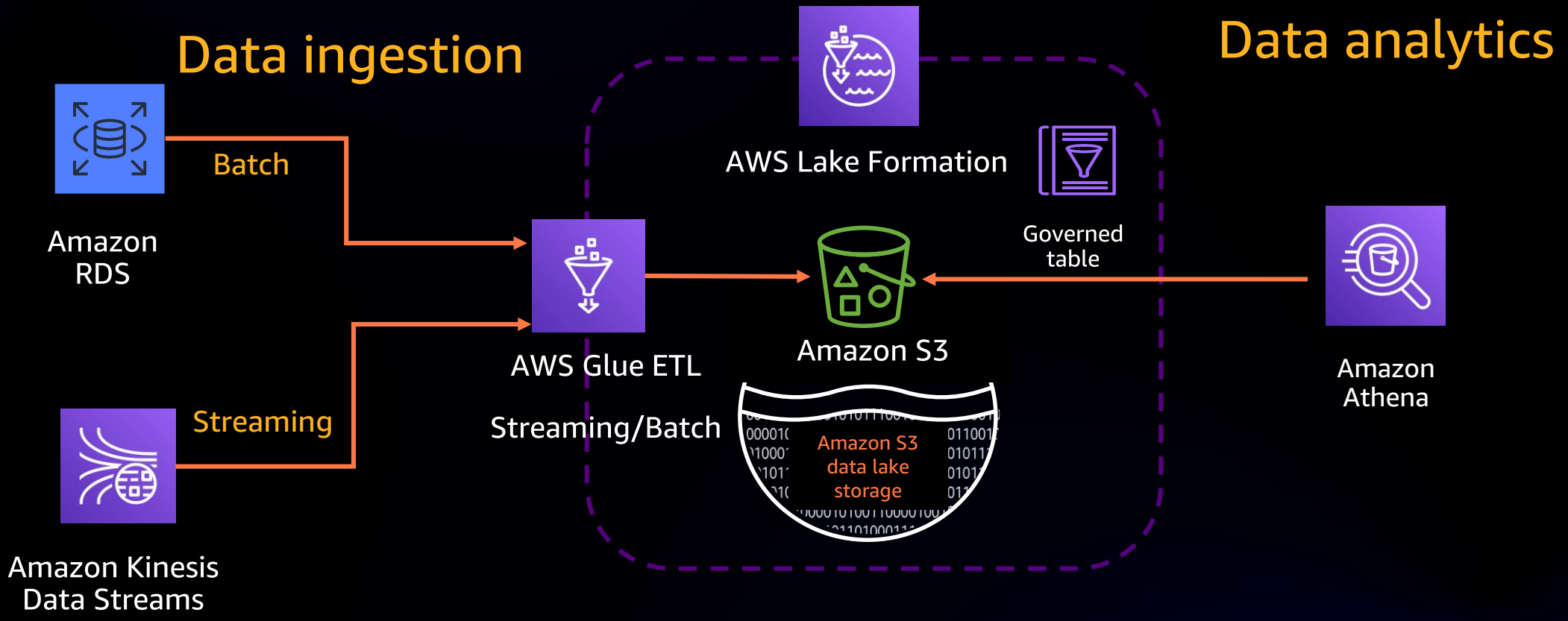
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Apache Hudi with Athena



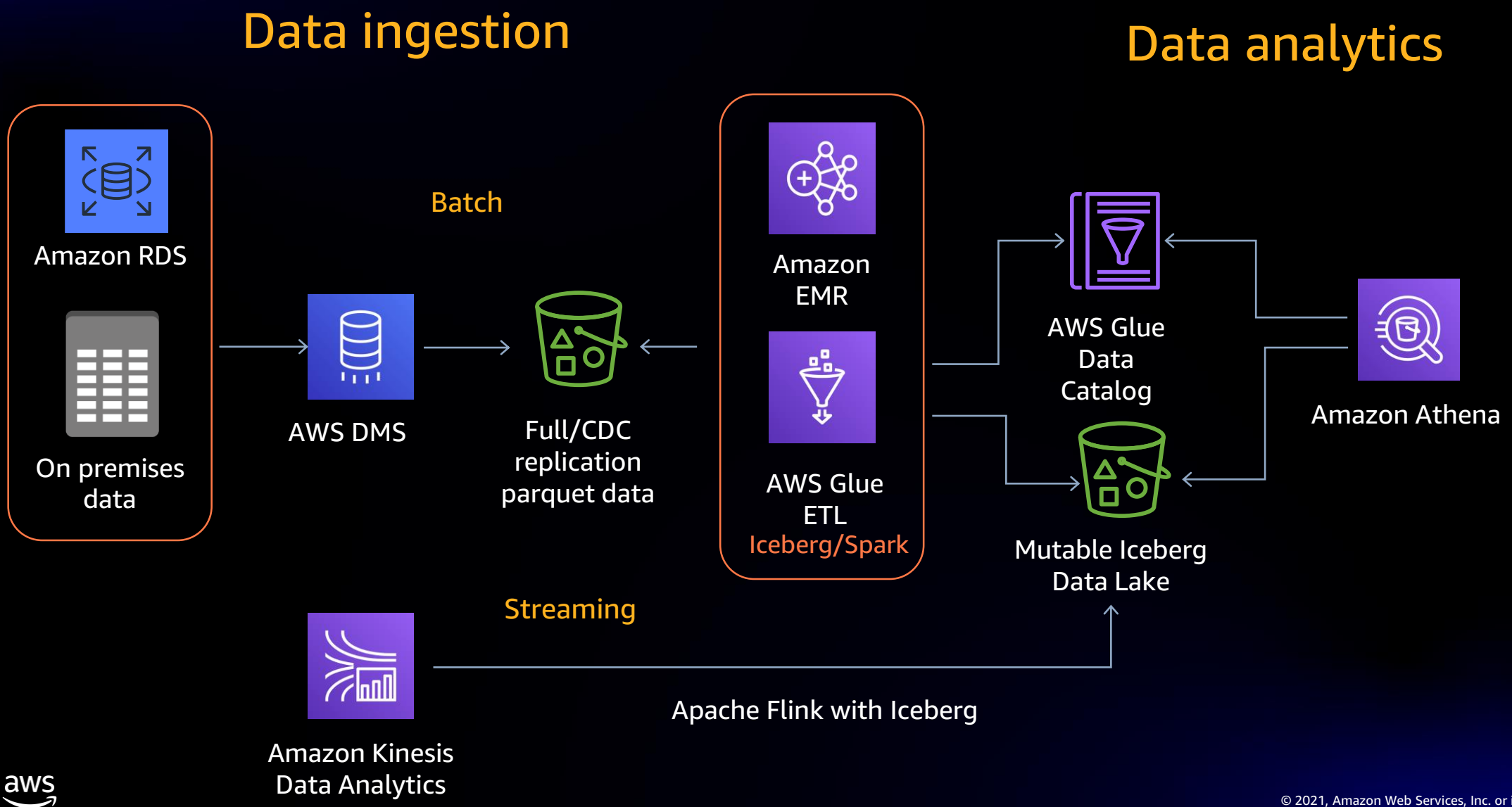
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Lake Formation governed tables



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Apache Iceberg with Athena

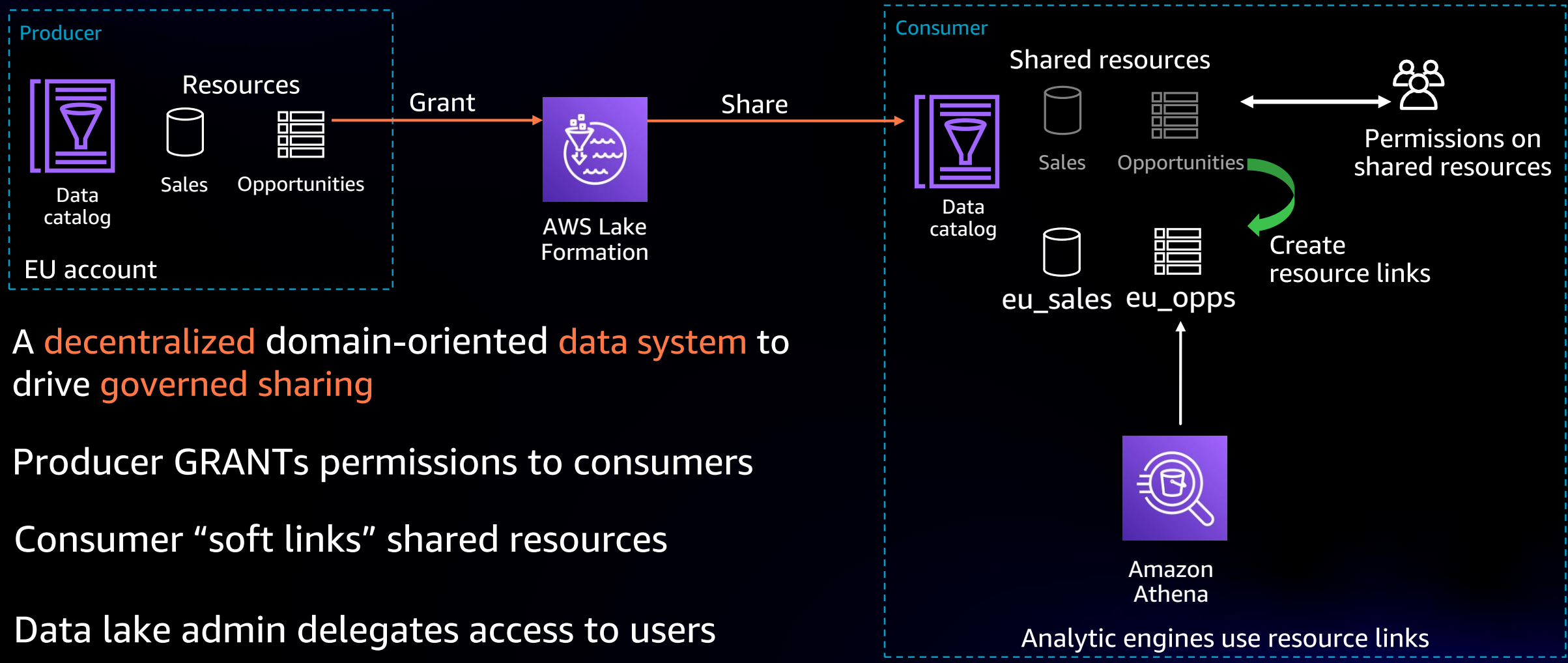


© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh with Amazon Athena

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh with Athena & AWS Lake Formation



A decentralized domain-oriented data system to drive governed sharing

Producer GRANTs permissions to consumers

Consumer “soft links” shared resources

Data lake admin delegates access to users

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.