



SUMMIT
ONLINE

Deep dive into AWS Lake Formation

Syed Jaffry

Solutions Architect
Amazon Web Services

Agenda

Why AWS Lake Formation?

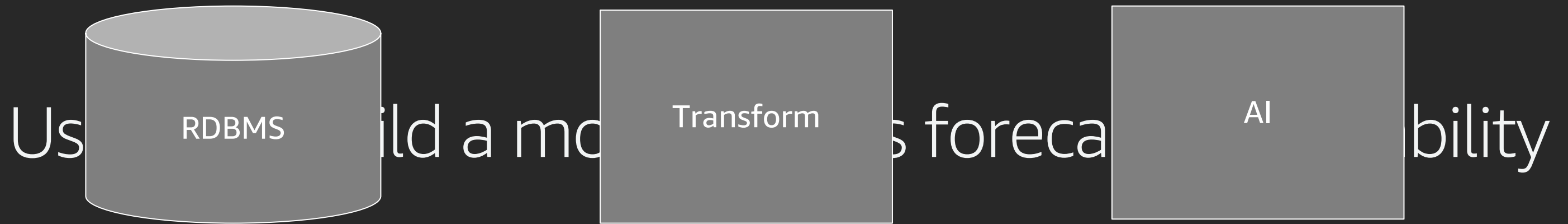
Ingest and transform

Security and access control

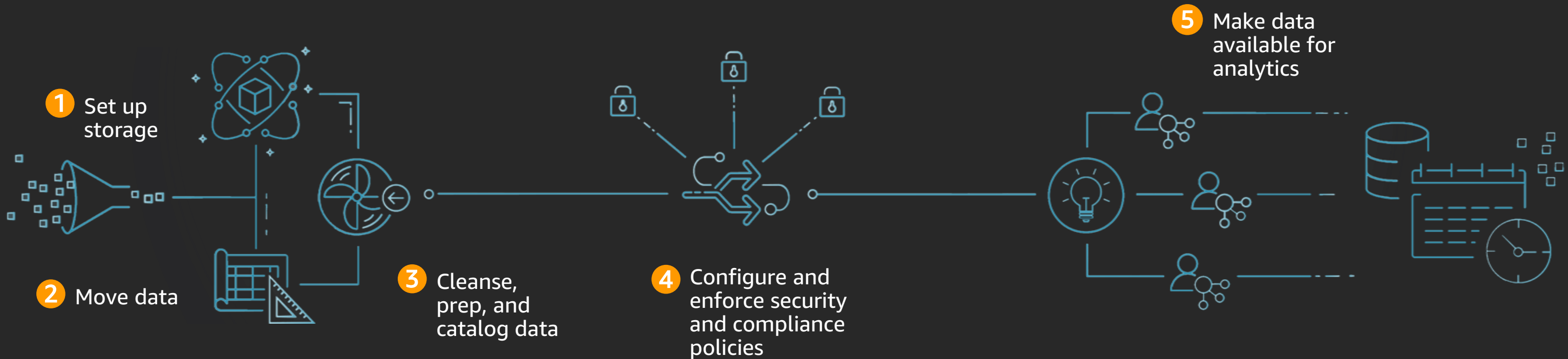
Data discovery and collaboration

Auditing and monitoring

Why AWS Lake Formation?



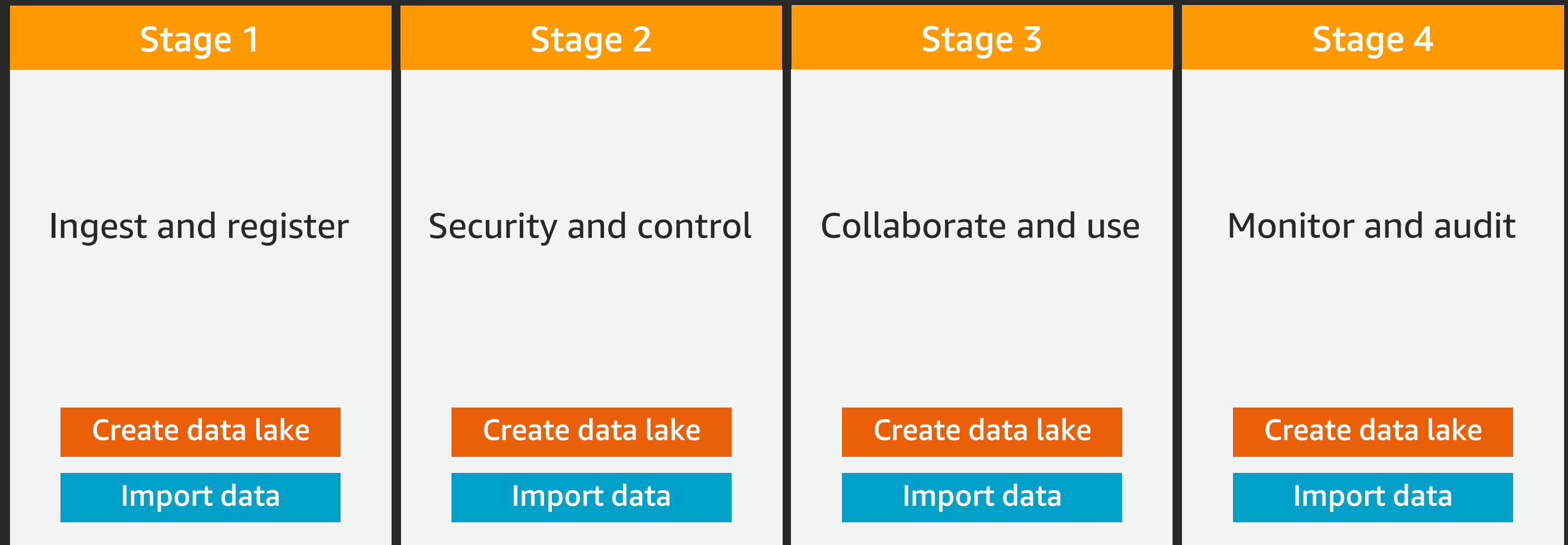
Typical steps of building a data lake



AWS Lake Formation

Overview

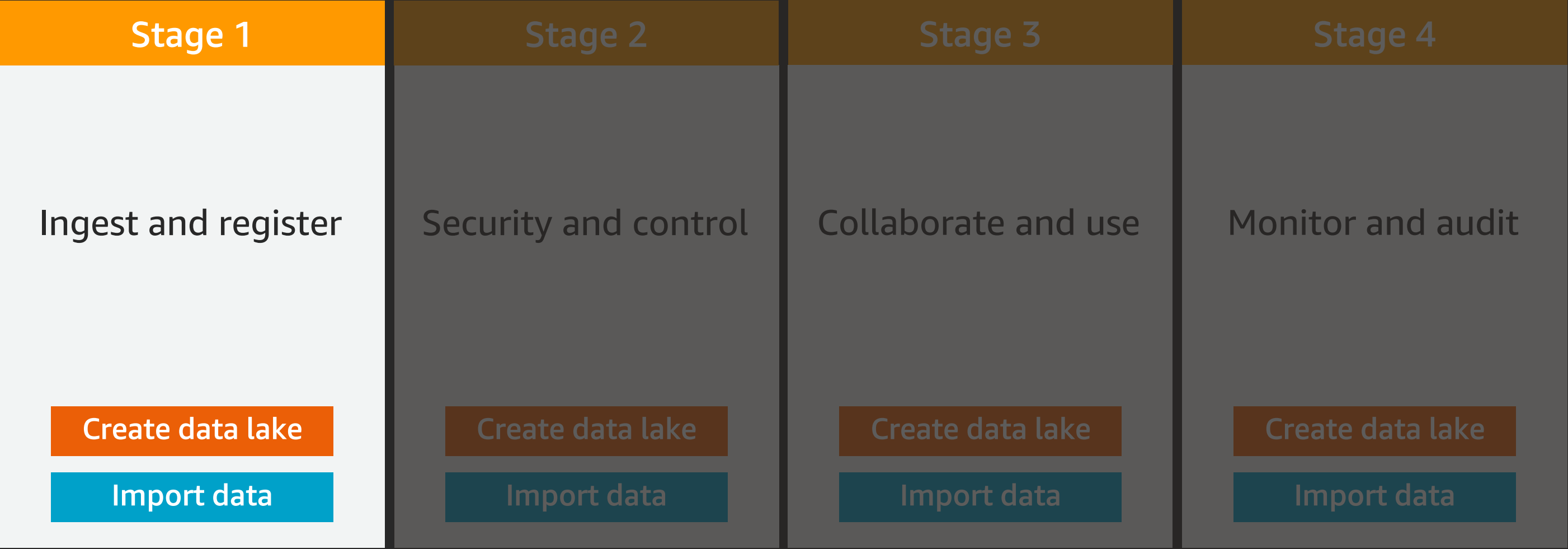
Data lake life cycle stages and activities



AWS Lake Formation

Overview

Data lake life cycle stages and activities



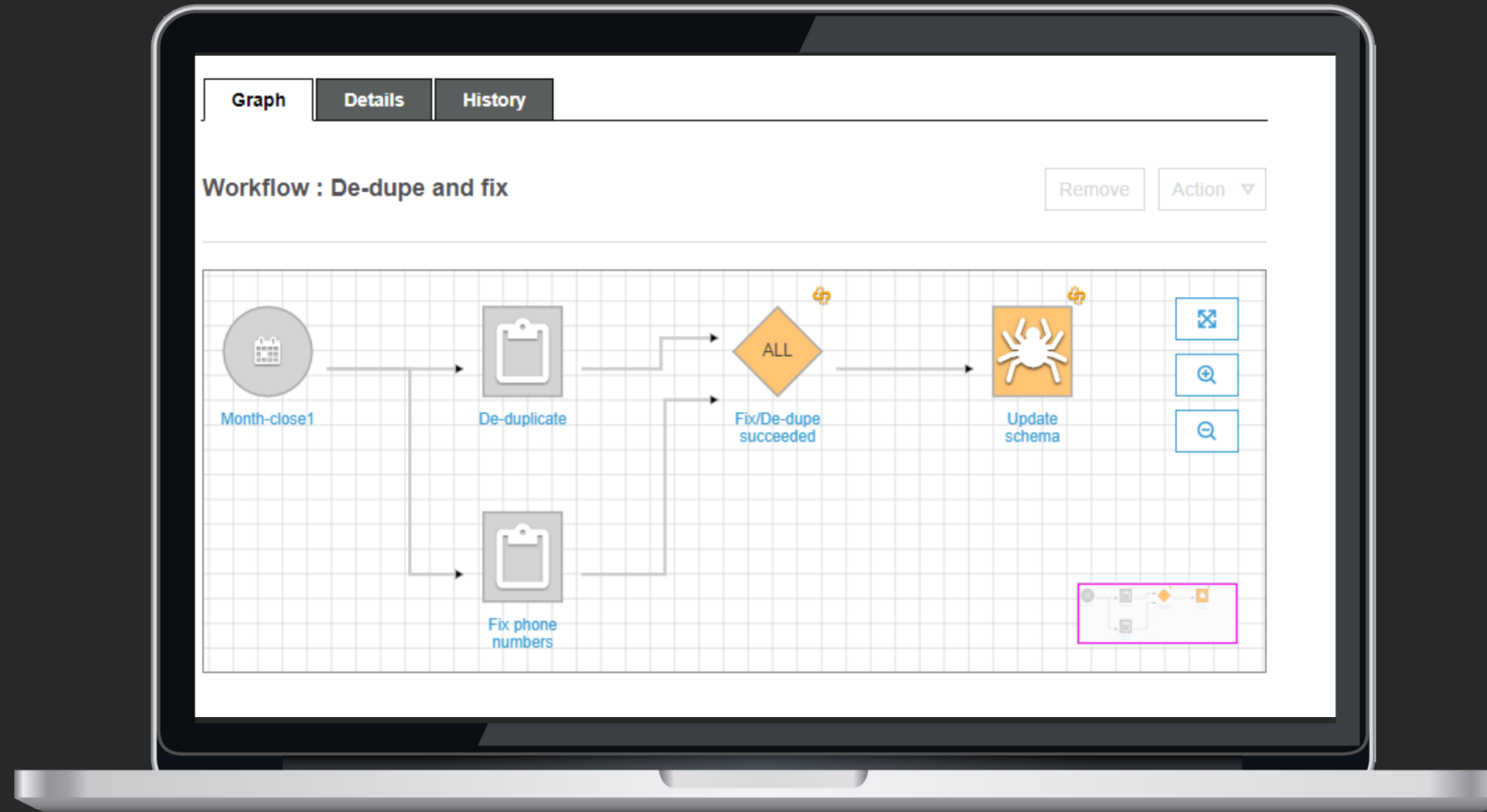
Demo

AWS Glue workflow: Orchestrate repeatable data pipelines

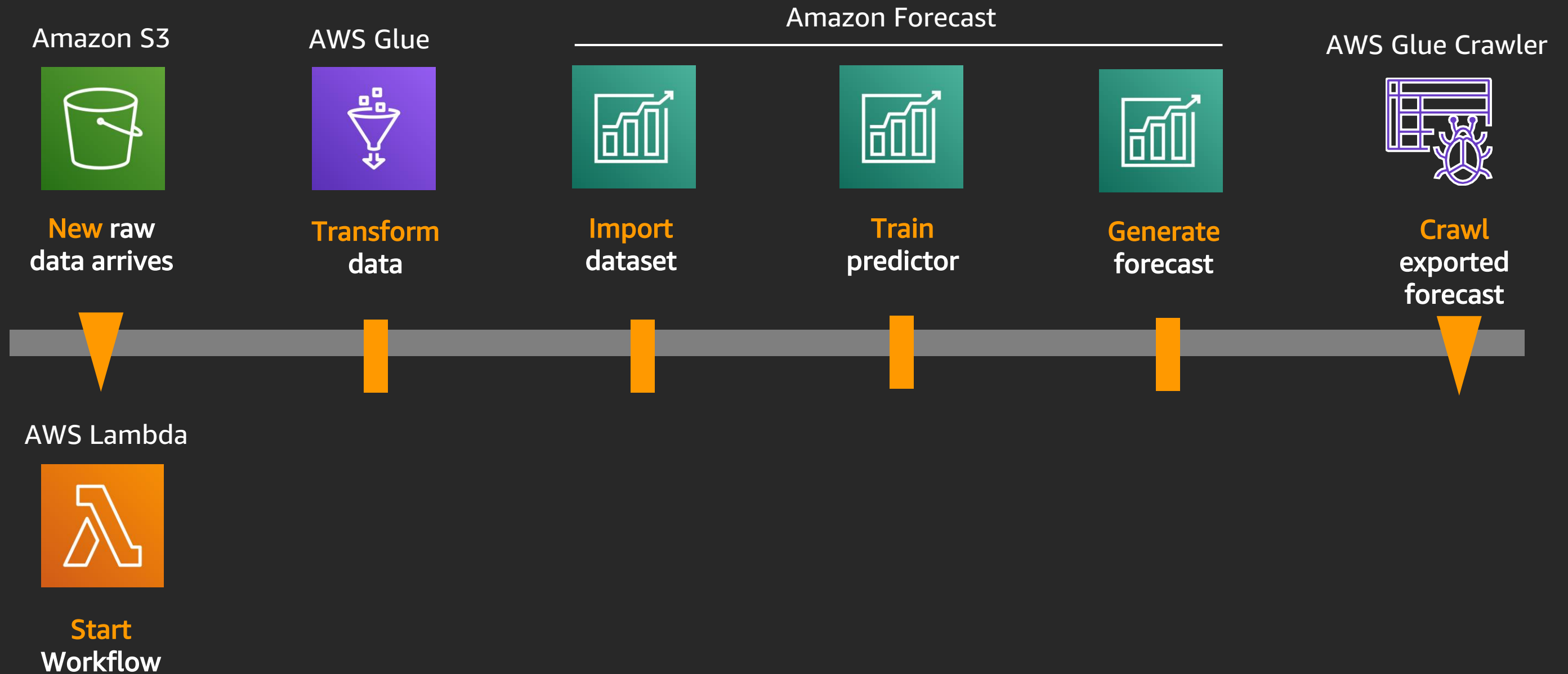
Easy way to create and visualize your business transformation rules

Allows for parameters and pipeline state to be shared across stages

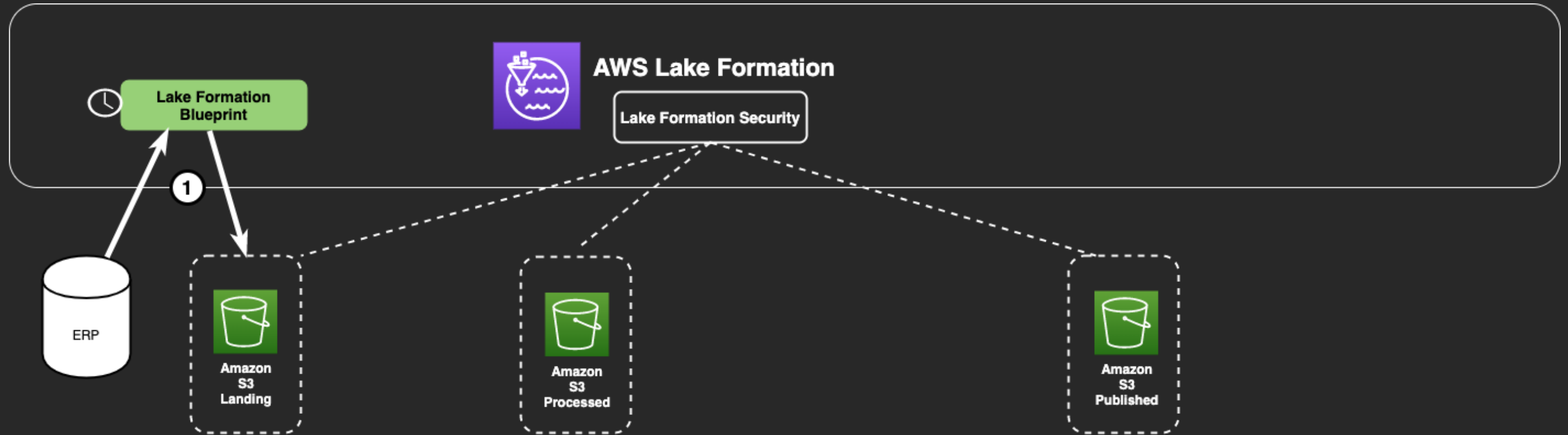
Dynamic views allow inspection of current running workflows for diagnostic and current state information



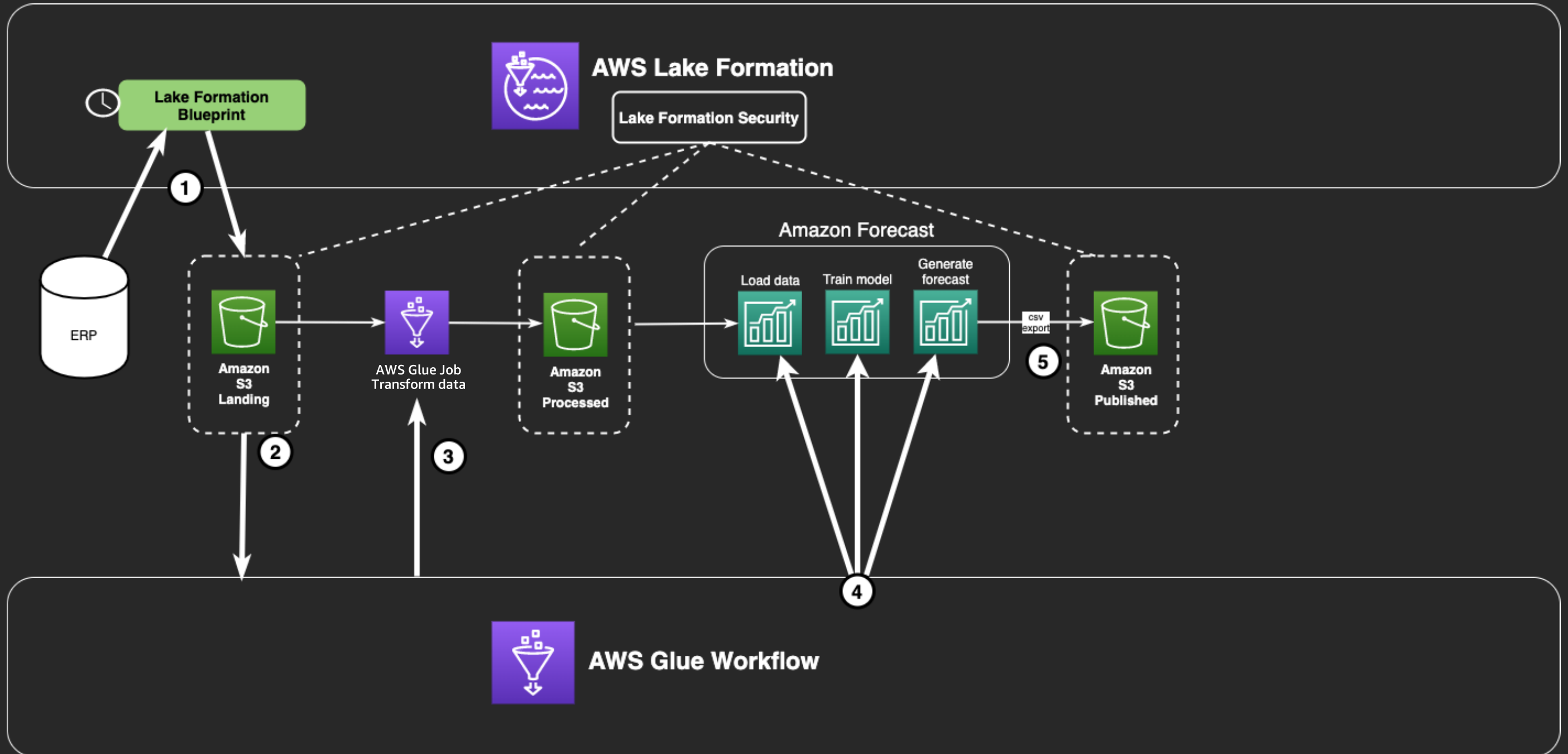
AWS Glue workflow for inventory forecast



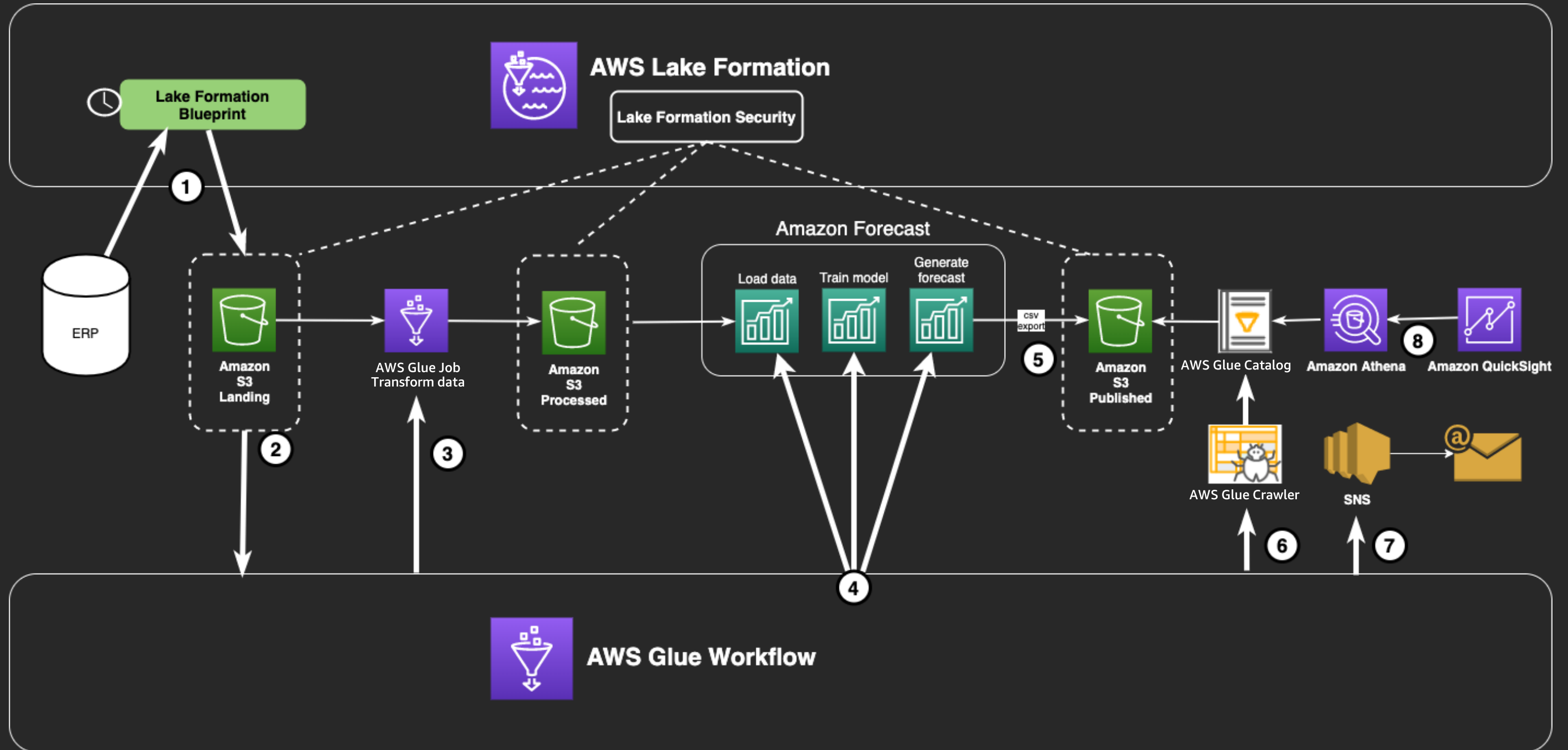
Inventory forecast end-to-end solution



Inventory forecast end-to-end solution

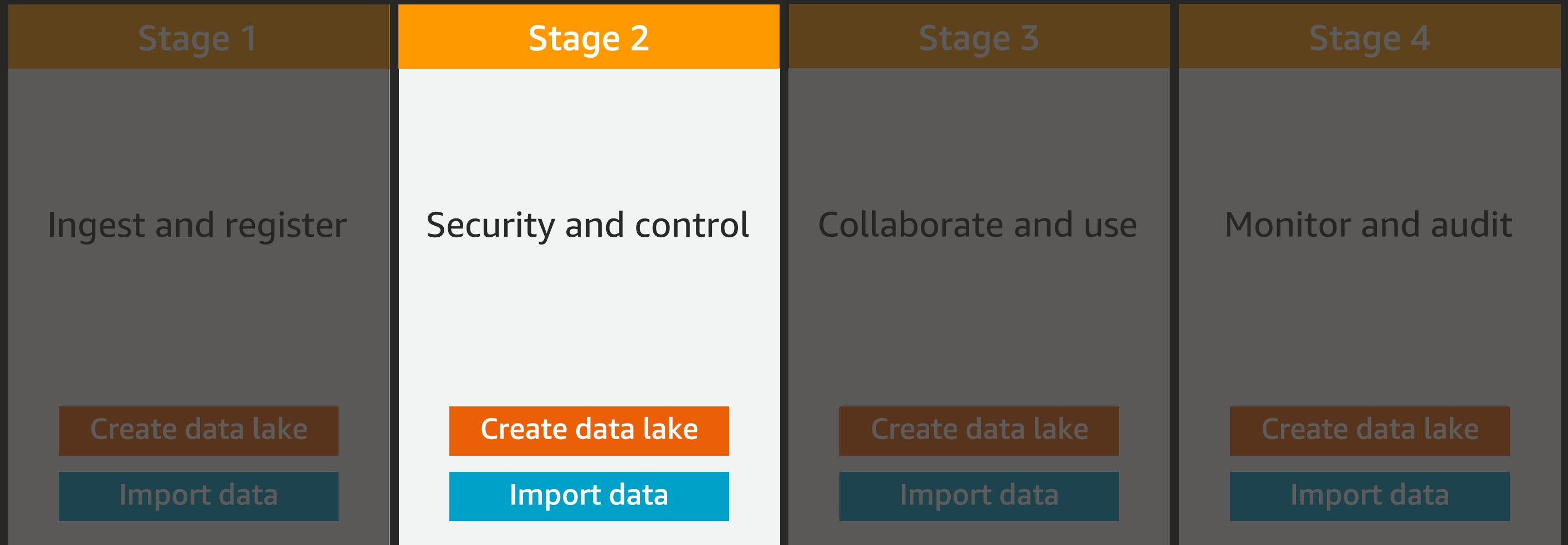


Inventory forecast end-to-end solution



Demo continued

AWS Lake Formation



Security personas in AWS Lake Formation



Data lake admins

- Run and operate the data lake
- Define secure storage boundaries
- Manage users
- Audit/optimize data lake



Data lake users

- Create, consume, and curate data sets
- Configure and manage access controls across data assets

Your project personas



Data lake admin



Retail sales manager

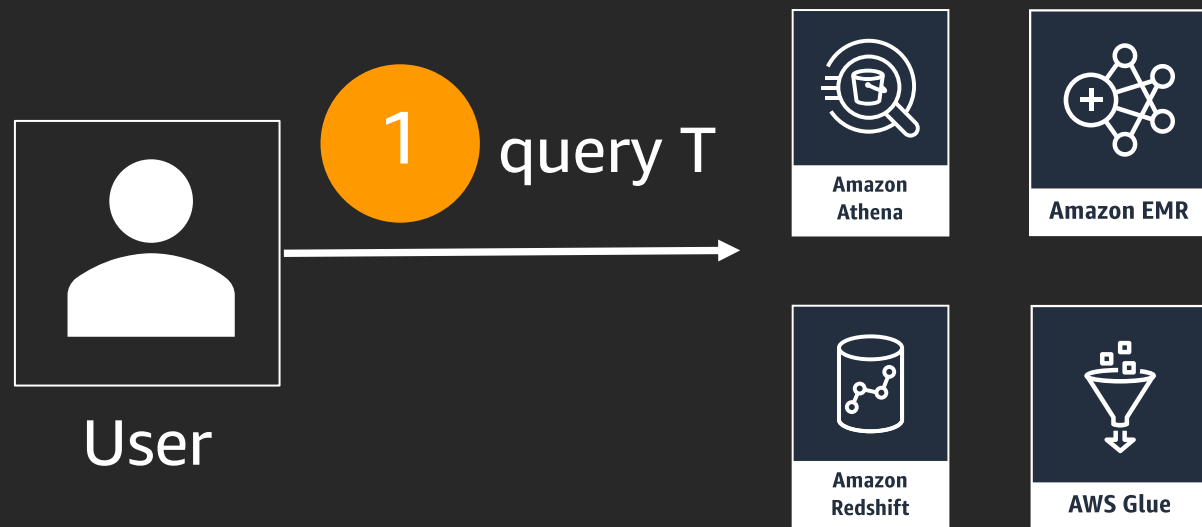


Analyst

Demo continued

Security – deep dive

End-services retrieve
underlying data
directly from S3



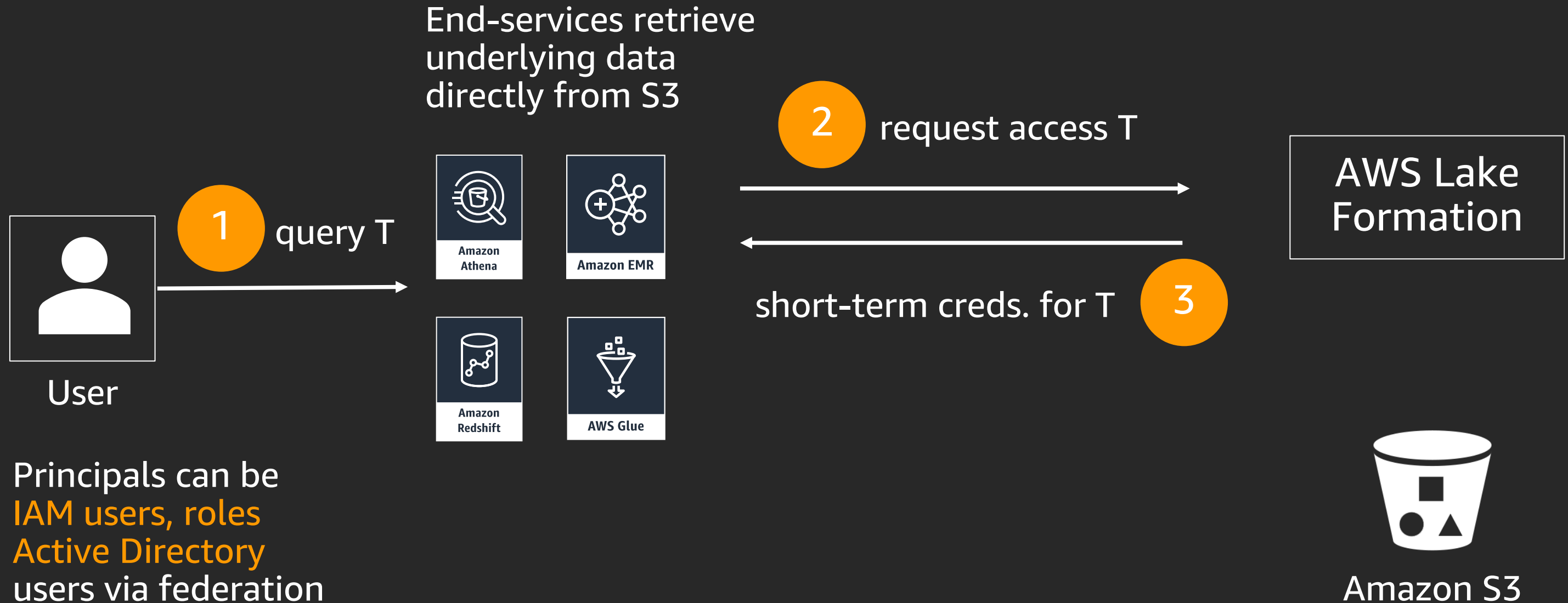
Principals can be
IAM users, roles
Active Directory
users via federation

AWS Lake
Formation

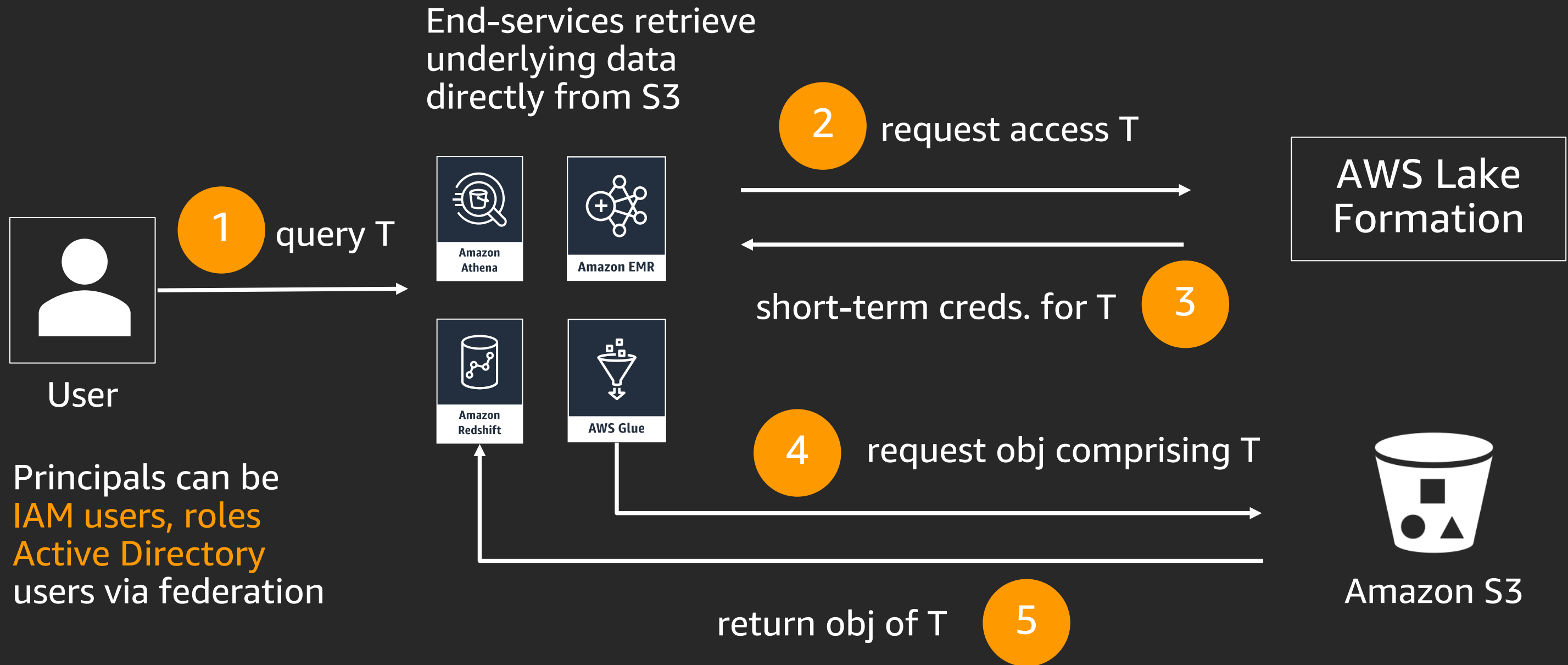


Amazon S3

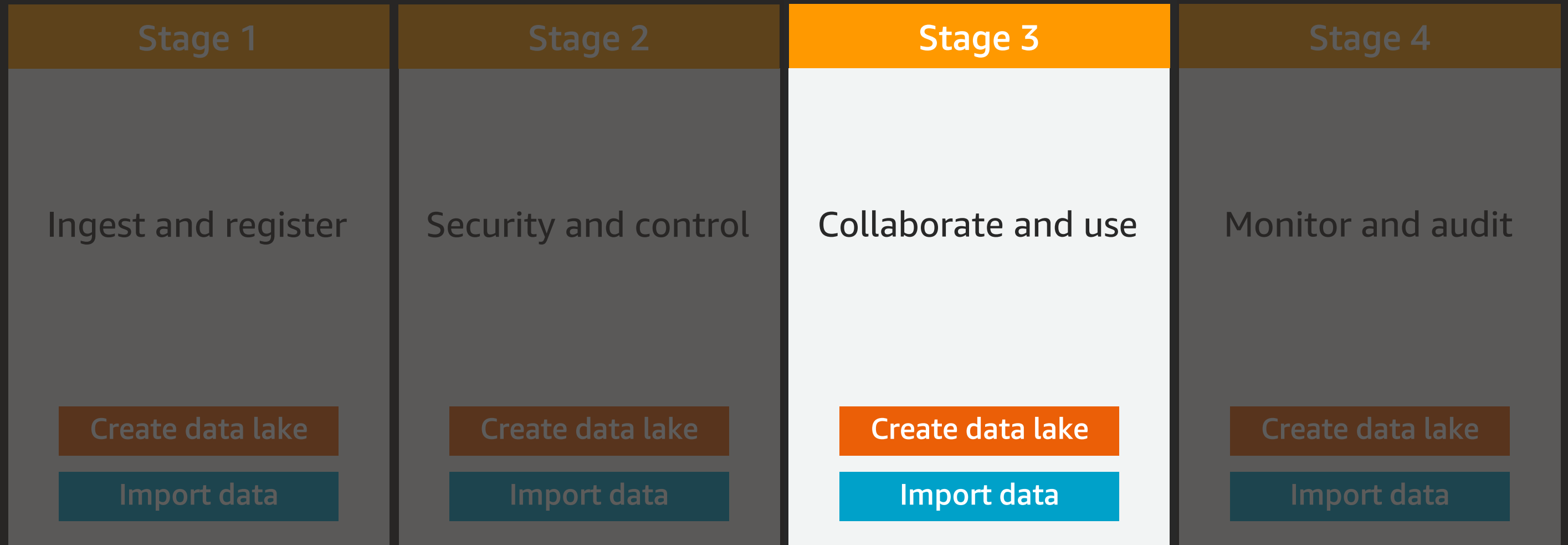
Security – deep dive



Security – deep dive



AWS Lake Formation



Demo continued

Visualizing forecasts

Amazon Athena, serverless interactive query



Used for DDL functionality

Complex data types

Multitude of formats

Supports data partitioning

(e.g., CREATE TABLE, ALTER TABLE, MSCK REPAIR)



Used for SQL Queries

In-memory distributed query engine

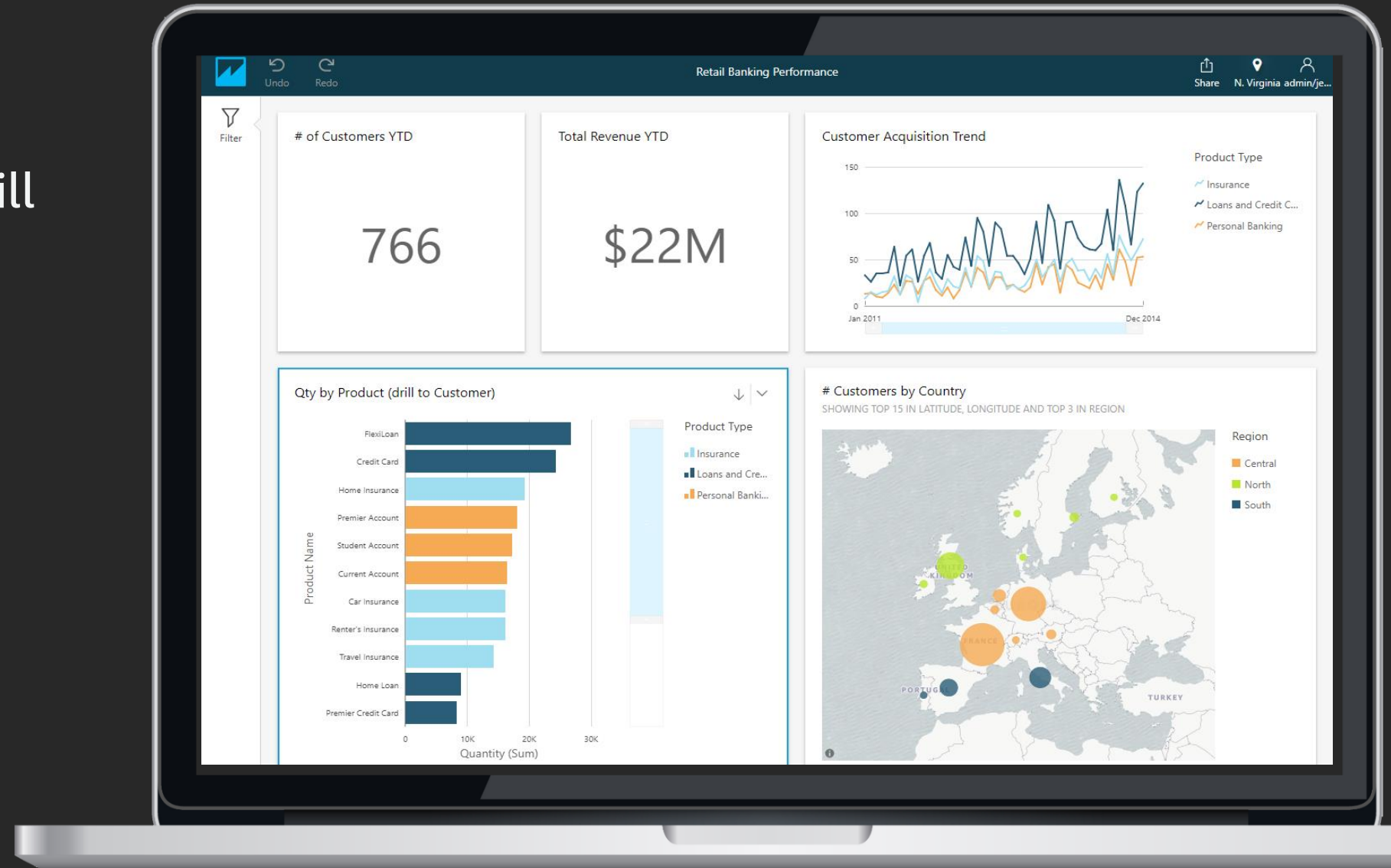
ANSI-SQL compatible with extensions

(e.g., SELECT * FROM tableName)

Amazon QuickSight

Beautiful, interactive dashboards

- Add rich interactivity like filters, drill downs, zooming, and more
- Blazing fast navigation
- Accessible on any device
- Data refresh
- Publish to everyone with a click



Data visualization with Amazon Athena

Athena ODBC/JDBC connector

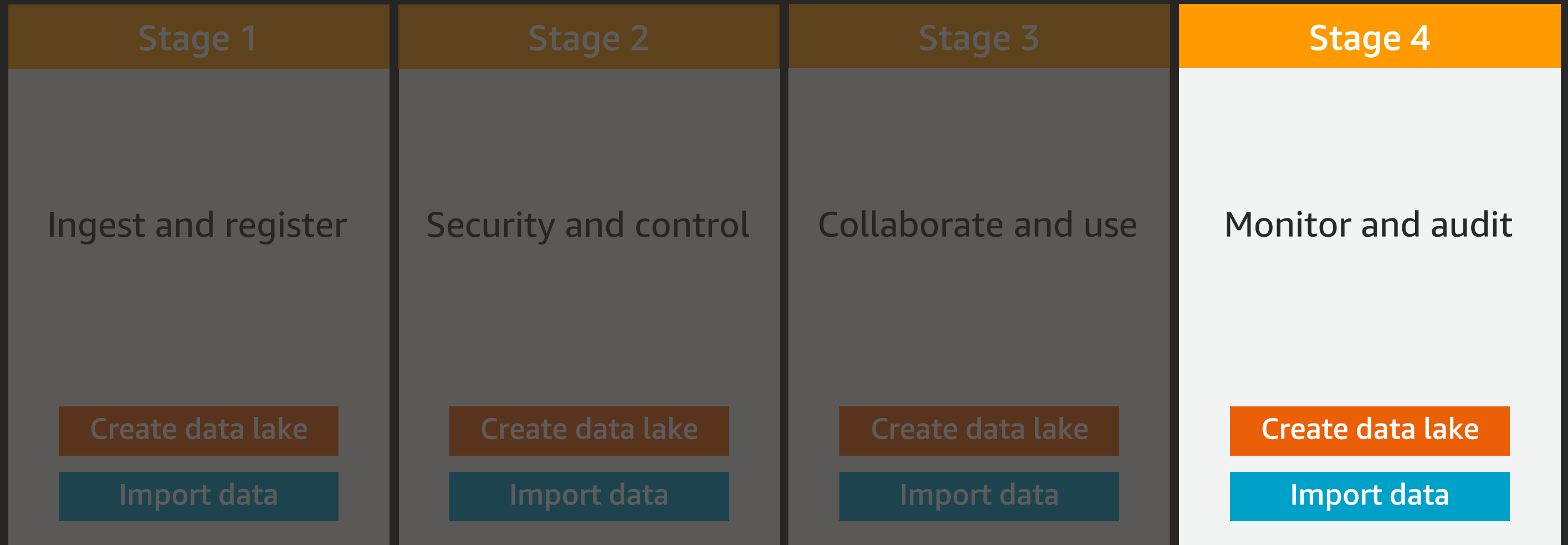


Amazon QuickSight



Power BI

AWS Lake Formation



Demo continued

Thank you!

Syed Jaffry

syejaffr@amazon.com