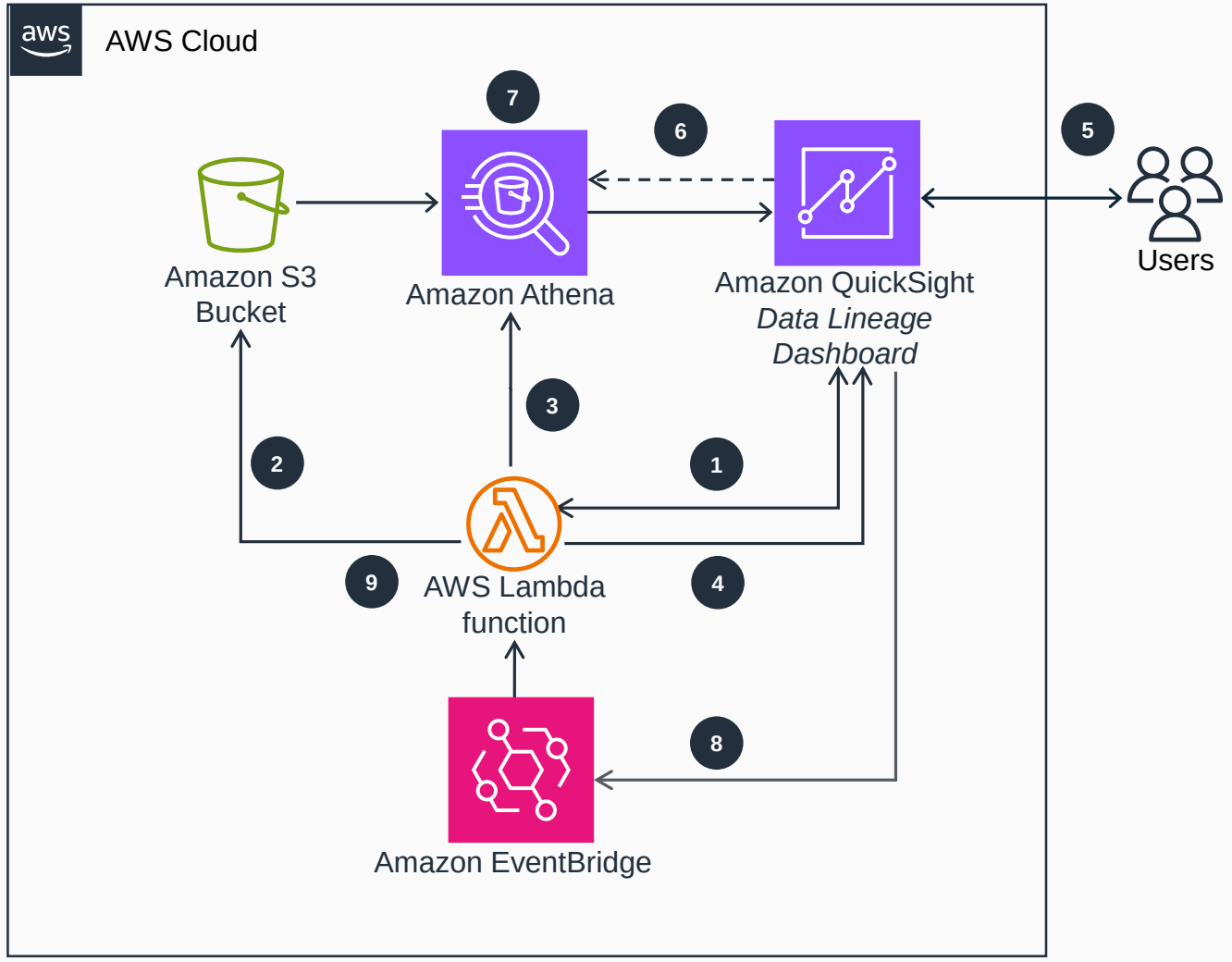
Guidance for Understanding Your Data Lineage on Amazon QuickSight

This architecture diagram demonstrates how to improve data lineage analysis in Amazon QuickSight. It accelerates dashboard optimization by automating manual data analysis. The AWS CloudFormation template also substitutes complex scripting with a simplified 15 minute setup, decreasing manual effort from weeks to minutes.



1. An **AWS Lambda** function reads the metadata for all **Amazon QuickSight** data sources, datasets, analyses, and dashboards in the AWS account.
2. The **Lambda** function stores the metadata as flat files in an **Amazon Simple Storage (Amazon S3)** bucket.
3. The **Lambda** function generates individual **Amazon Athena** tables for the metadata of data sources, datasets, analyses and dashboards.
4. The **Lambda** function constructs new datasets and a **QuickSight** dashboard called the Data Lineage Dashboard. This visualizes data lineage, objects, and resource information in **QuickSight** using the **Athena** tables.
5. Users access **QuickSight** to view the Data Lineage Dashboard and derive insights.
6. When users view the Data Lineage Dashboard, **QuickSight** issues queries to the **Athena** tables.
7. **Athena** runs the queries, accesses the data in the **Amazon S3** bucket, and returns the results to **QuickSight** to render the Data Lineage Dashboard.
8. An **Amazon EventBridge** rule invokes the **Lambda** function when users create or update a data source, dataset, analysis, or dashboard in QuickSight.
9. The **Lambda** function reads the new or revised resource metadata from **QuickSight** and adds to the **Amazon S3** bucket. The Data Lineage Dashboard can then access this added metadata through **QuickSight**.

Reviewed for technical accuracy February 16, 2024

AWS Reference Architecture