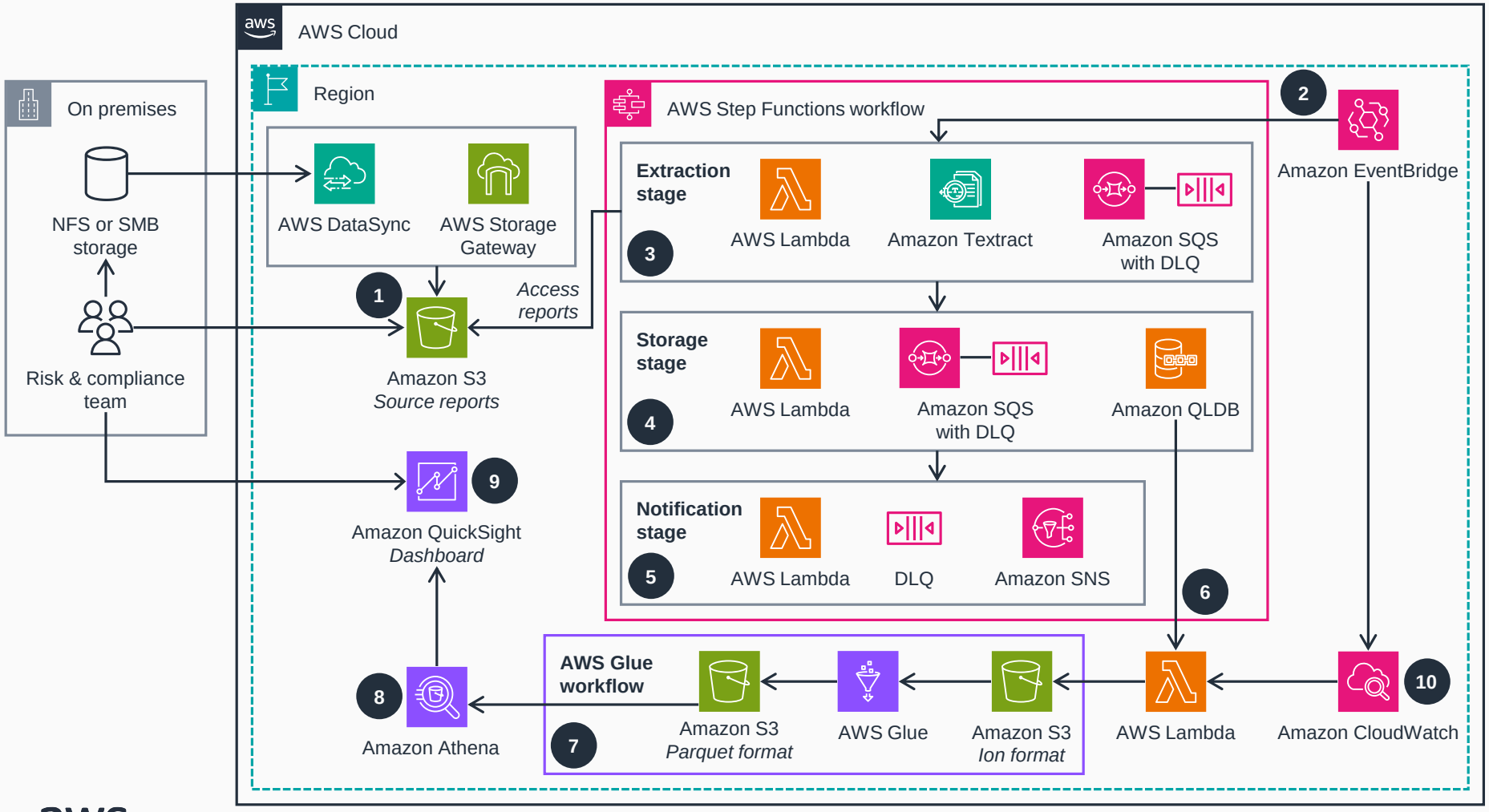
Guidance for Automating Background Checks for Reporting & Audits on AWS

This architecture diagram shows how organizations can improve background reporting using AWS serverless technology and Amazon Quantum Ledger Database (Amazon QLDB). This slide details steps 1–6 of the architecture diagram. For details on steps 7–10, go to the next slide.



1. Risk and compliance teams receive background check reports from direct sources or third-parties, which are then uploaded into **Amazon Simple Storage Service (Amazon S3)**. These reports are stored in **Amazon S3** using frontend applications or are automatically moved from the Network File System (NFS) or Server Message Block (SMB) storage using **AWS DataSync** or **AWS Storage Gateway**.

2. The reports are processed in batches using a preconfigured **Amazon EventBridge** scheduler, which initiates a three-stage **AWS Step Functions** workflow based on a defined schedule in an asynchronous fashion.

3. The report extraction stage uses an **AWS Lambda** function as an invocation type to read files from **Amazon S3**. **Amazon Textract** extracts data from the report files and stores it on **Amazon Simple Queue Service (Amazon SQS)**. **Amazon SQS** supports dead-letter queues (DLQs), which other queues can target for messages that are not processed successfully.

4. In the storage stage, a **Lambda** function (as an invocation type) reads the messages from **Amazon SQS** and stores the message payload on **Amazon Quantum Ledger Database (Amazon QLDB)**. **Amazon QLDB** maintains the entire history of data changes on individual data records in an immutable fashion for full traceability. Any messages that are not processed successfully are moved to the DLQ.

5. In the notification stage, a **Lambda** function validates messages on the DLQ and sends email notifications using **Amazon Simple Notification Service (Amazon SNS)**.

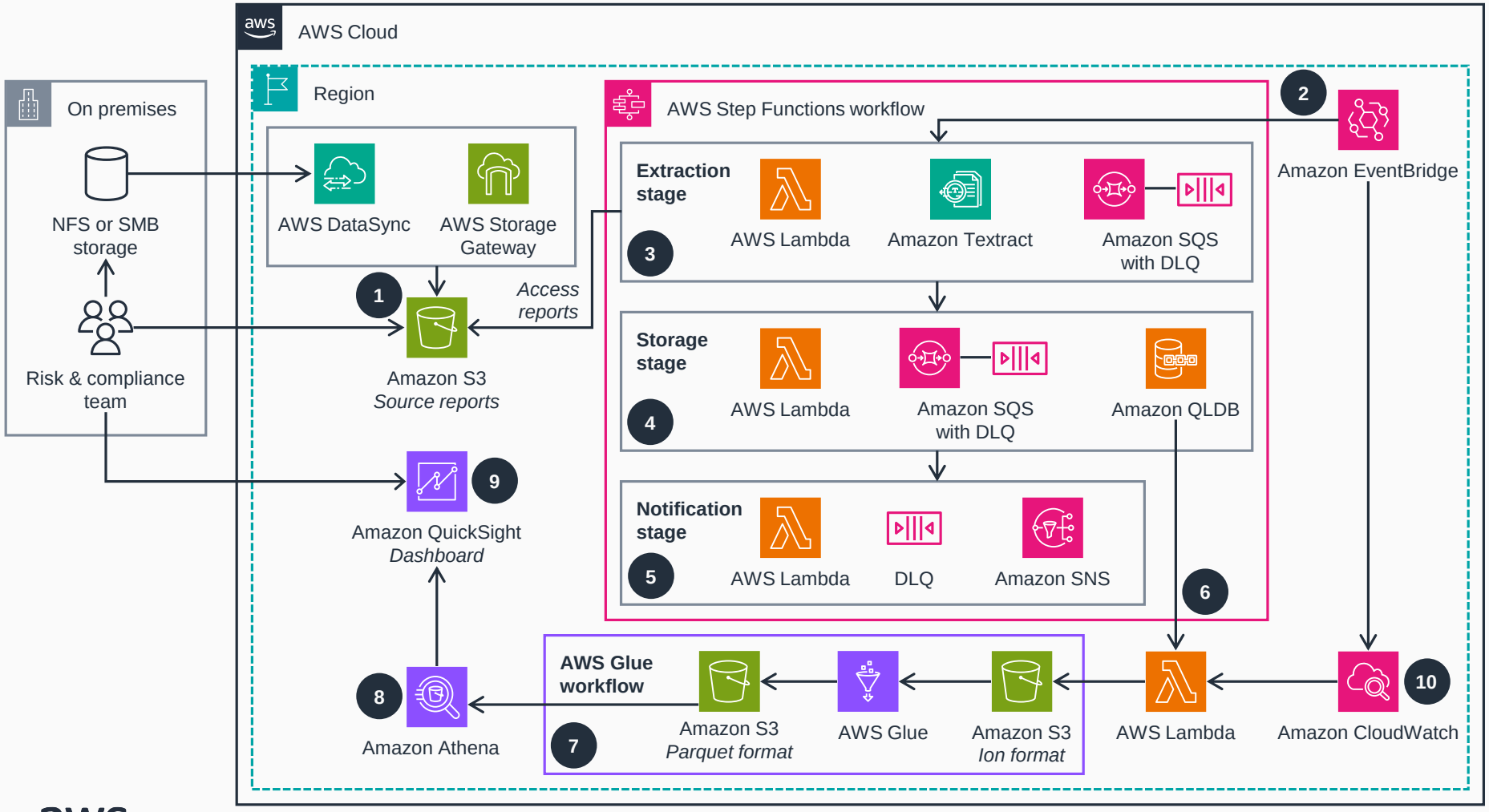
6. Depending on the required frequency for reporting, the **EventBridge** scheduler invokes the daily, weekly, or monthly data extraction from **Amazon QLDB** using a **Lambda** function in an Amazon Ion format. It then stores the data on **Amazon S3** as a raw export and invokes an **AWS Glue** workflow.

Reviewed for technical accuracy January 5, 2024

AWS Reference Architecture

Guidance for Automating Background Checks for Reporting & Audits on AWS

This slide details steps 7–10 of the architecture diagram.



7 An **AWS Glue** workflow transforms the Amazon Ion extract from **Amazon QLDB** into partitioned Apache Parquet files. An **AWS Glue** crawler reads and catalogs the Parquet-formatted version, making it available to **Amazon Athena**.

8 **Athena** views that are created on partitioned Parquet data provide data enrichment for the **Amazon QuickSight** dashboard.

9 The **QuickSight** dashboard uses **Athena** views to create business intelligence dashboards on top of background check reports.

10 **Amazon CloudWatch** is used for workflow monitoring, logging, and event tracking.

Reviewed for technical accuracy January 5, 2024

AWS Reference Architecture