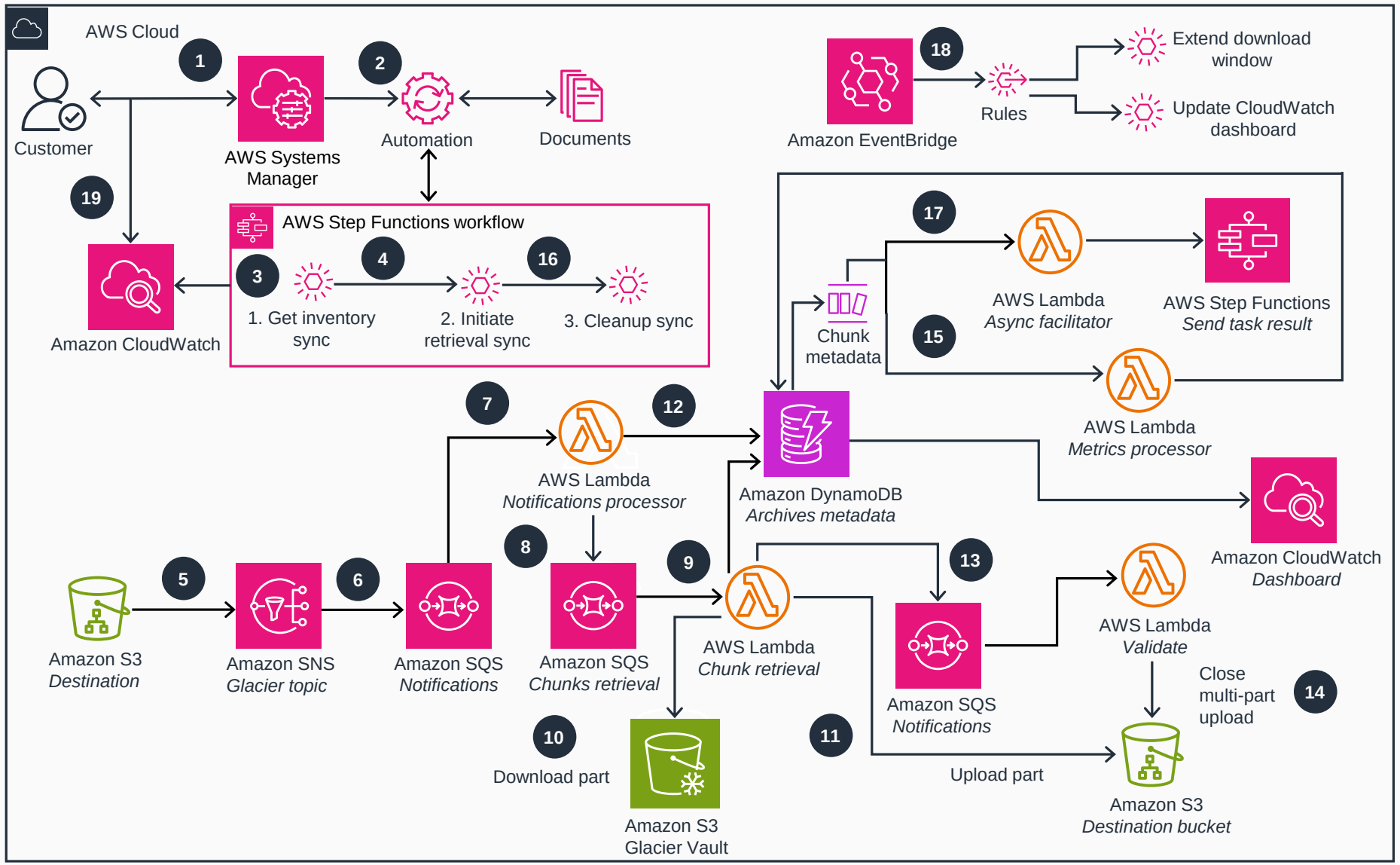
Guidance for Data Transfer from Amazon S3 Glacier Vaults to Amazon S3

This architecture diagram demonstrates how to automate and optimize the restore, copy, and transfer process of Amazon S3 Glacier vault archives by copying the vault archives to a defined S3 bucket destination and storage class. (Steps 1-9)



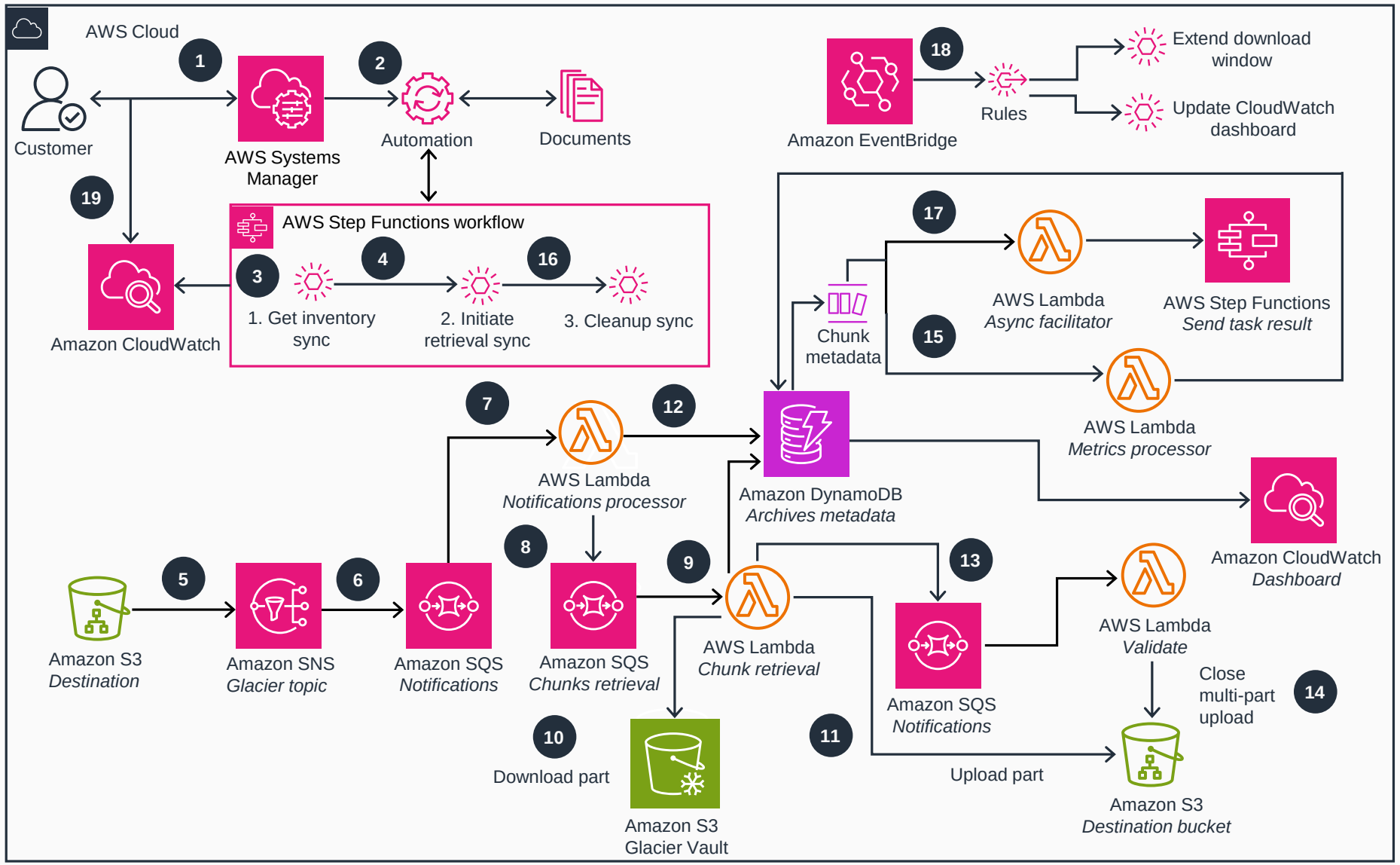
1. Invoke a transfer workflow using an **AWS Systems Manager** document.
2. The **Systems Manager** document starts an **AWS Step Functions** Orchestrator execution.
3. The **Step Functions** Orchestrator execution initiates a nested **Step Functions** Get Inventory workflow to retrieve the inventory file.
4. Upon completion of the inventory retrieval, the Guidance invokes the Initiate Retrieval nested **Step Functions** workflow.
5. When a job is ready, **Amazon Simple Storage Service (Amazon S3) Glacier** sends a notification to an **Amazon Simple Notification Service (Amazon SNS)** topic, indicating job completion.
6. The Guidance stores all job completion notifications in the **Amazon Simple Queue Service (Amazon SQS)** Notifications queue.
7. When an archive job is ready, the **Amazon SQS** Notifications queue invokes the **AWS Lambda** Notifications Processor function. This **Lambda** function prepares the initial steps for archive retrieval.
8. The **Lambda** Notifications Processor function places chunks retrieval messages in **Amazon SQS** Chunks Retrieval queue for chunk processing.
9. The **Amazon SQS** Chunks Retrieval queue invokes the **Lambda** Chunk Retrieval function to process each chunk.

Reviewed for technical accuracy April 10, 2025

AWS Reference Architecture

Guidance for Data Transfer from Amazon S3 Glacier Vaults to Amazon S3

This architecture diagram demonstrates how to automate and optimize the restore, copy, and transfer process of Amazon S3 Glacier vault archives by copying the vault archives to a defined S3 bucket destination and storage class. (Steps 10-19)



10 The **Lambda** Chunk Retrieval function downloads the chunk from **Amazon S3 Glacier**.

11 The **Lambda** Chunk Retrieval function uploads a multipart upload part to **Amazon Simple Storage Service (Amazon S3)**.

12 After a new chunk is downloaded, the Guidance stores chunk metadata in **Amazon DynamoDB** (for example, etag, checksum_sha_256, tree_checksum).

13 The **Lambda** Chunk Retrieval function verifies whether all chunks for that archive have been processed. If so, it inserts an event into the **Amazon SQS** Validation queue to invoke the **Lambda** Validate function.

14 The **Lambda** Validate function performs an integrity check against the tree hash in the inventory, calculates a checksum, and passes it to the into the close multipart upload call. If that hash is wrong, **Amazon S3** rejects the request.

15 **DynamoDB Streams** invokes the **Lambda** Metrics Processor function to update the transfer process metrics in **DynamoDB**.

16 The **Step Functions** Orchestrator execution enters an async wait, pausing until the archive retrieval workflow concludes before initiating the **Step Functions** Cleanup workflow.

17 The **DynamoDB** stream invokes the **Lambda** Async Facilitator function, which unlocks asynchronous waits in **Step Functions**.

18 **Amazon EventBridge** rules periodically initiate **Step Functions** Extend Download Window and Update **Amazon CloudWatch** Dashboard workflows.

19 Monitor the transfer progress using a **CloudWatch** dashboard.

Reviewed for technical accuracy April 10, 2025

AWS Reference Architecture