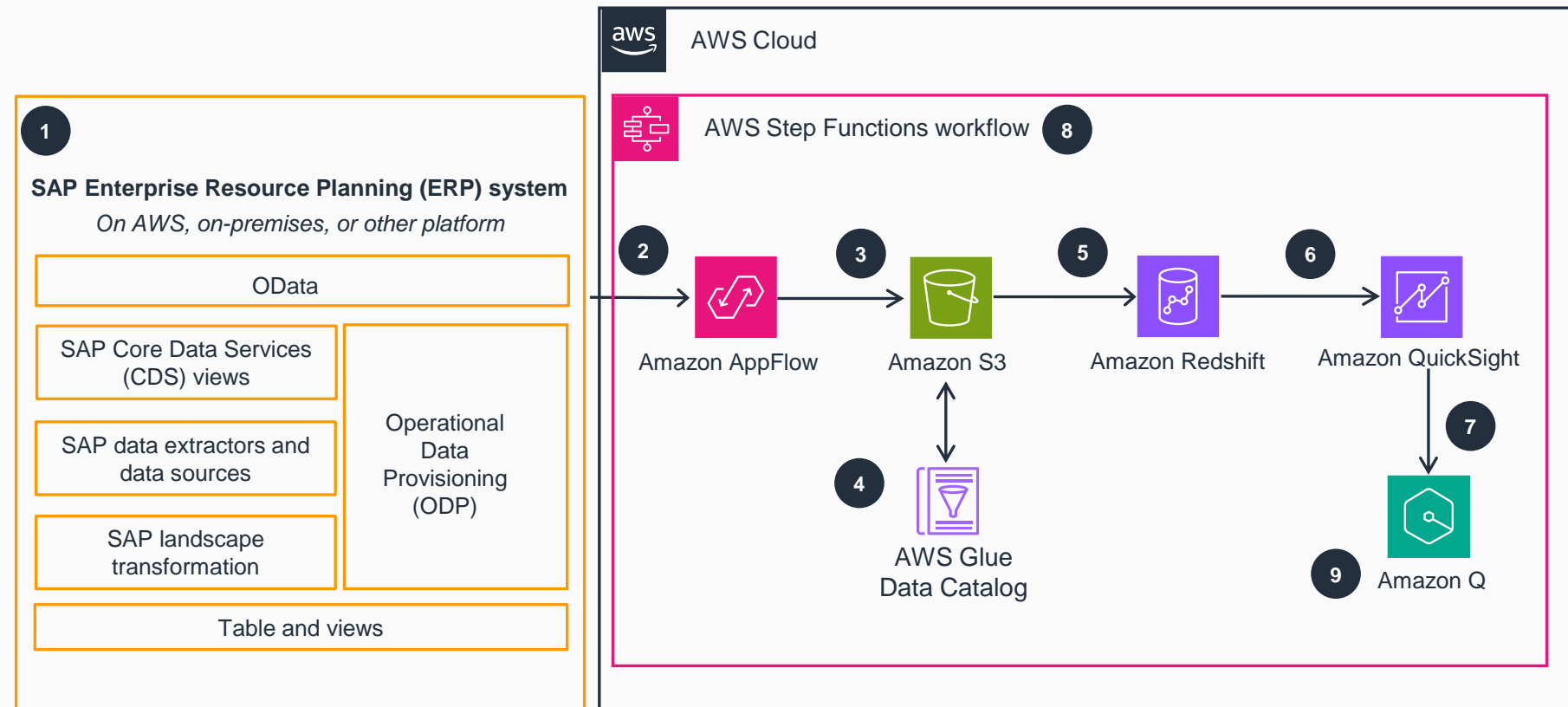


# Guidance for Building a SAP Cloud Data Warehouse on AWS

## Overview

This architecture diagram shows how to build a cloud data warehouse on AWS by extracting data from SAP using the OData protocol. You can use the data warehouse to model and combine SAP data with that of other sources loaded into the data warehouse. The next two slides show metadata replication and data marts, respectively.



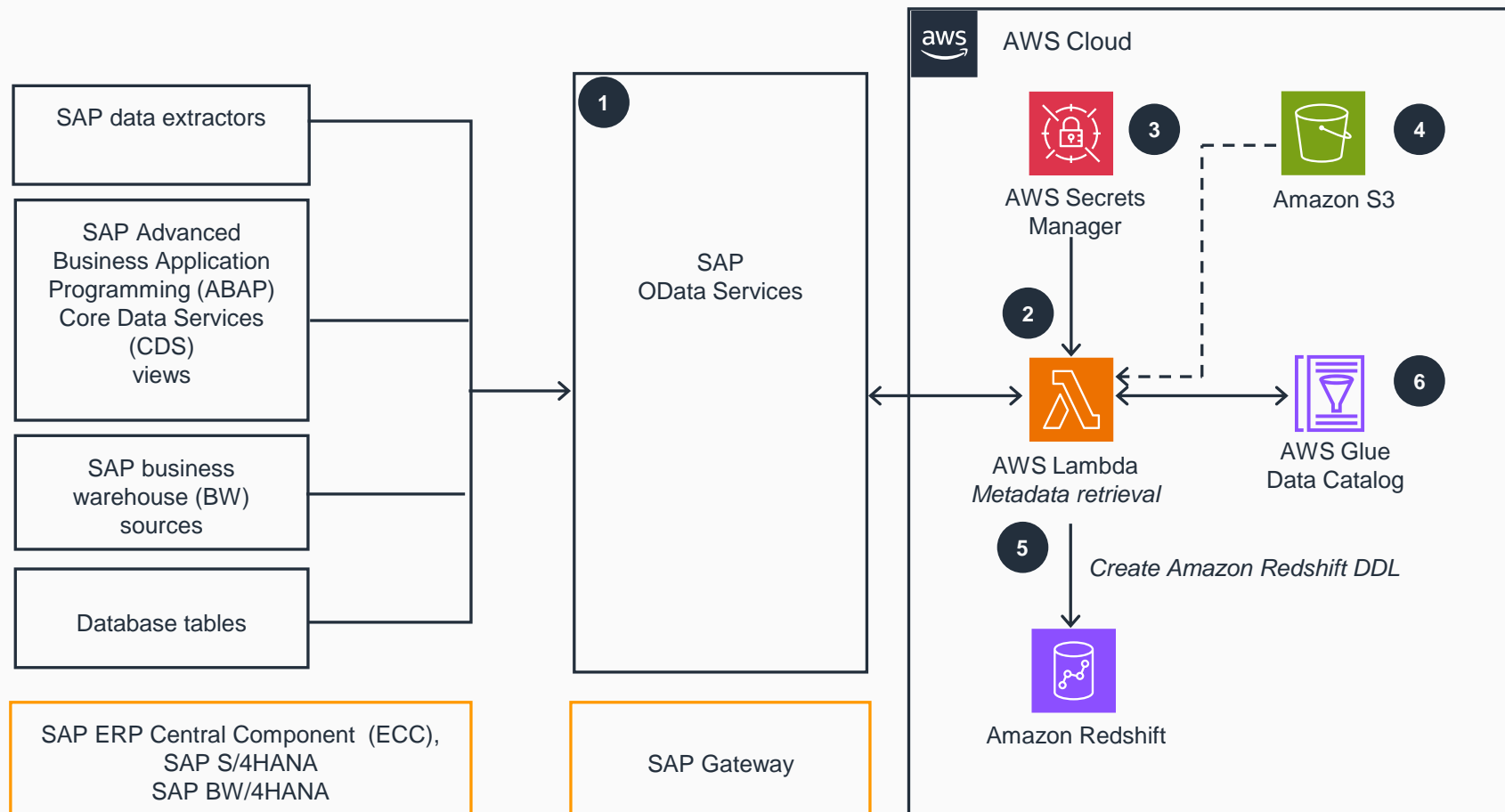
- 1 Configure operational data provisioning (ODP) for extraction in the SAP Gateway of your SAP system.
- 2 Create the OData system connection from **Amazon AppFlow** to your SAP source system. This is through **AWS PrivateLink** for SAP on AWS. You can connect with AWS through a virtual private network (VPN), **AWS Direct Connect**, or over the internet.
- 3 In **Amazon AppFlow**, create the flow using the SAP source created in step 2. Run the flow to extract data from SAP and save it to an **Amazon Simple Storage Service** (Amazon S3) bucket.
- 4 Use an **AWS Glue** crawler to create a data catalog entry with metadata for the extracted SAP data in an **Amazon S3** bucket.
- 5 Load data into **Amazon Redshift** through simple 'COPY' commands. Model the data with other non-SAP sources in your data warehouse.
- 6 Create the dataset in **Amazon QuickSight** with **Amazon Redshift** as the data source.
- 7 Create a dashboard to visualize the business data according to user requirements. Use inbuilt machine learning (ML) and insight features to help enable speed to insight.
- 8 Deploy **AWS Step Functions** for overall orchestration and alerting of your data pipelines and processes.
- 9 The generative BI capabilities of **Amazon Q** within **Amazon QuickSight** offer insights from the data.



# Guidance for Building a SAP Cloud Data Warehouse on AWS

## SAP metadata replication

This architecture uses Amazon Lambda and SAP OData to replicate metadata and create Amazon Redshift Data Definition Language (DDL) tables. A Python script with PyOdata queries SAP OData sources and generates DDL to create tables in Amazon Redshift and the AWS Glue Data Catalog.



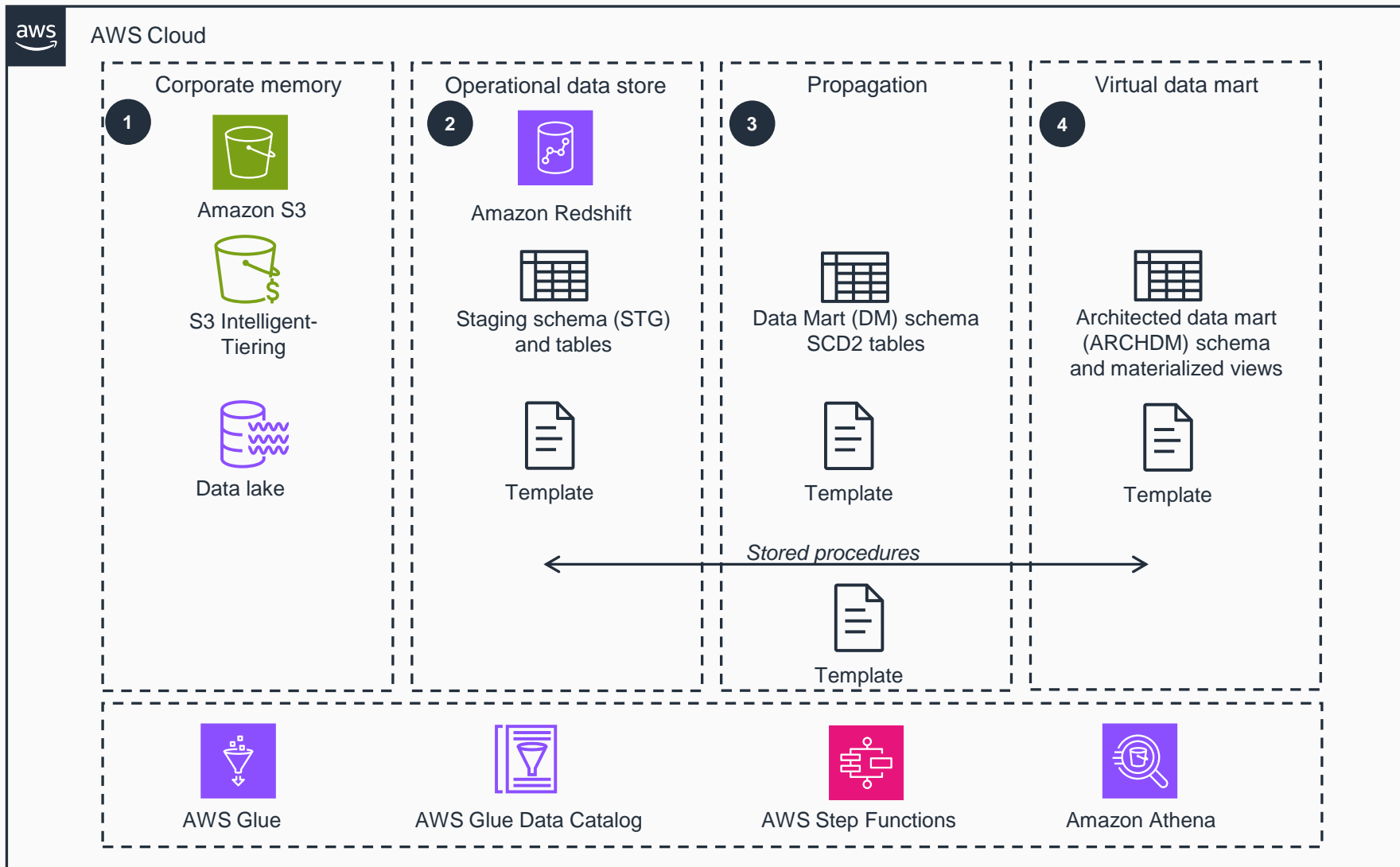
- 1 Configure ODP for extraction in the SAP Gateway of your SAP system.
- 2 Configure and run the SAP Metadata Replication **AWS Lambda** function.
- 3 **Lambda** retrieves the SAP credentials from **AWS Secrets Manager**.
- 4 **Lambda** reads the config file from **Amazon S3**.
- 5 **Lambda** creates Data Definition Language (DDL) statements and runs them in **Amazon Redshift**.
- 6 **Lambda** creates **AWS Glue** tables in the **AWS Glue Data Catalog**.



# Guidance for Building a SAP Cloud Data Warehouse on AWS

## Amazon Redshift data marts

This architecture diagram illustrates how the Redshift data mart layers are used. With the Slowly Changing Dimension Type 2 (SCD2) data modeling technique, you have the full history of your data movements that you can query.



- 1 The corporate memory layer contains all SAP extracted data in **Amazon S3**.
- 2 The operational data store layer is used to temporarily hold the data until the data is loaded into the Data Mart (DM) and the architected data mart (ARCHDM) layers.
- 3 The propagation layer includes the Slowly Changing Dimension Type 2 (SCD2) tables or tables that will contain the entire history of changes.
- 4 The virtual data mart layer provides the presentation layer through the use of materialized views in **Amazon Redshift**.

