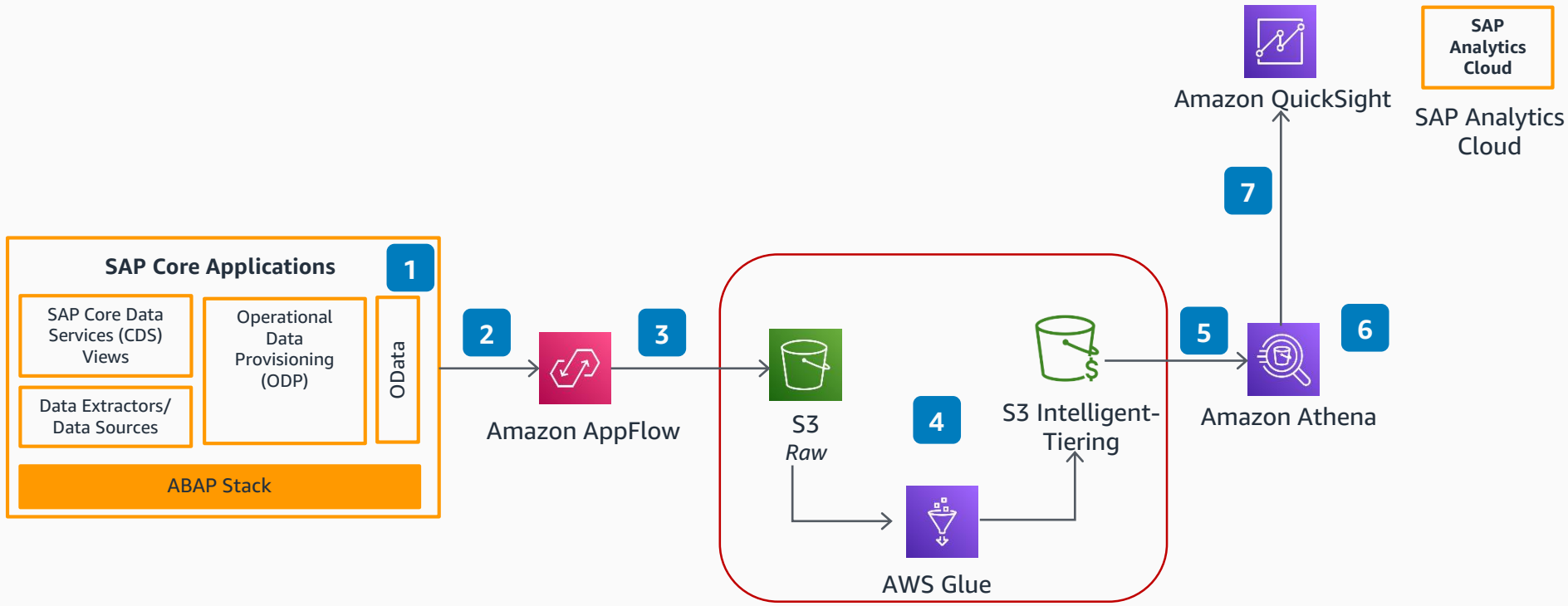
Guidance for Data Compliance and Retention for SAP on AWS

This guidance shows how to build a Reference SAP system data on AWS by extracting data from SAP before it is retired and decommissioned. The data required for reference and compliance purpose is extracted, transformed, and loaded to low-cost storage using Amazon Simple Storage Service (Amazon S3). The data in Amazon S3 is analyzed using AWS Analytics services to gain insights on the data for reference and compliance purposes.



1. Identify and configure (or activate) the data sources. Transfer from SAP application to AWS services. Then configure the Operating Data Provisioning (ODP) for extraction in the SAP Gateway of your SAP system.
2. Create the OData system connection from **Amazon AppFlow** to SAP source system. This can be via **AWS PrivateLink** for SAP on AWS/connect with AWS via 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 to an **Amazon Simple Storage Service** (Amazon S3) bucket.
4. Use **AWS Glue DataBrew** to cleanse and optimize the format of your SAP data. Save the transformed data in another **Amazon S3** bucket. With **AWS Glue** crawler, create a data catalog entry with metadata for the extracted SAP data in another **Amazon S3** bucket. Use the **Amazon S3 Intelligent-Tiering** storage class to automatically optimize storage costs.
5. Use **Amazon Athena**, an interactive query service, to analyze the data in **Amazon S3** using standard SQL.
6. Visualize the dataset in **Amazon QuickSight** with **Amazon Athena** as the data source.
7. Create dashboards to visualize the audit and compliance data requirements. For joint visualization, extend the archive data using SAP Analytics Cloud (SAC).

Reviewed for technical accuracy October 26, 2022

AWS Reference Architecture