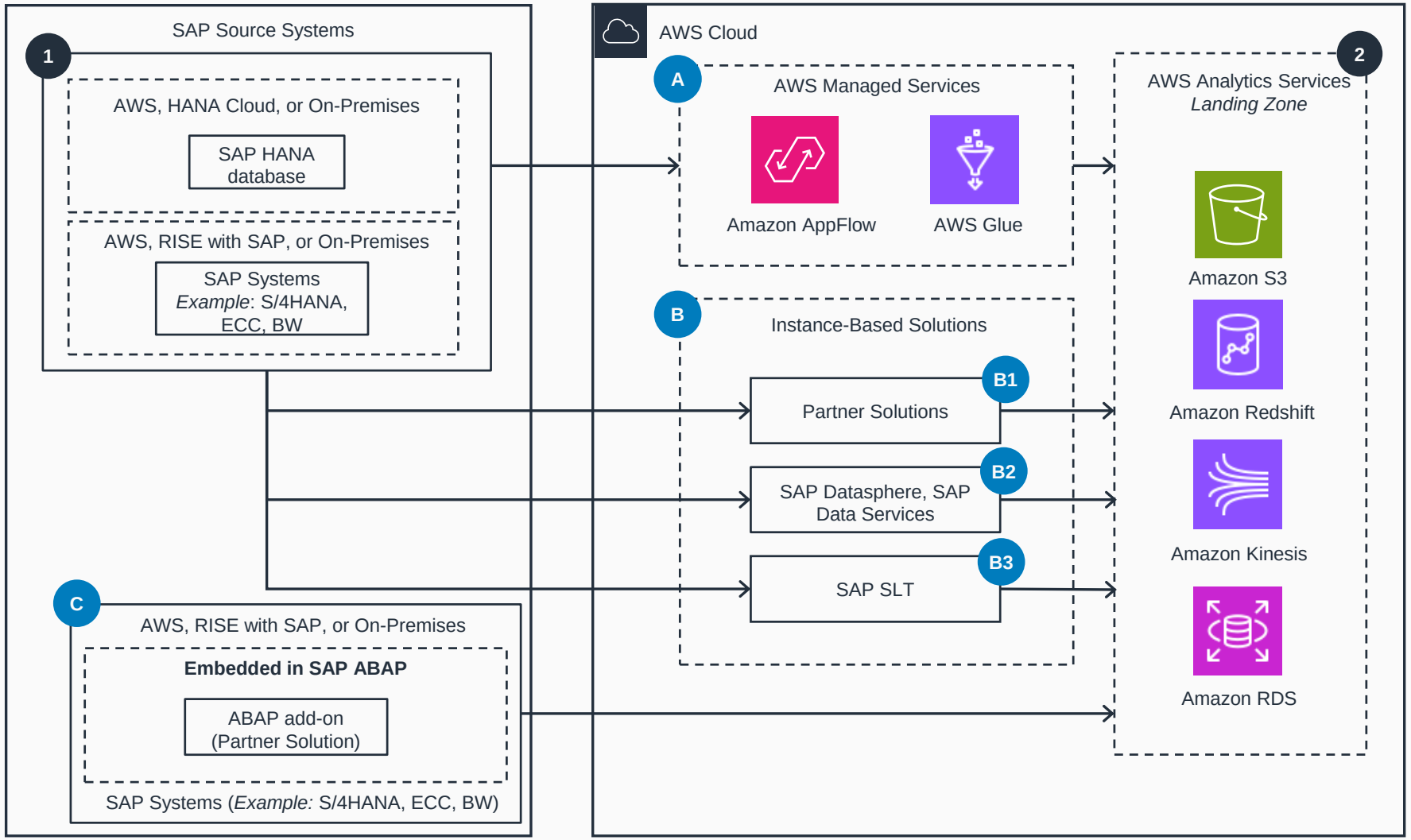
Guidance for SAP Data Integration and Management on AWS

Overview of Architecture Patterns

This architecture diagram shows the pattern options for ingesting SAP systems to AWS. These patterns complement SAP supported mechanisms (such as OData, ODP, SLT, BTP) using AWS services, SAP products, and AWS Partner Solutions.



1. SAP data hosted on SAP RISE, SAP HANA Cloud, AWS, or on-premises systems can be extracted in real-time or batch and full or incremental mode from SAP NetWeaver systems, such as SAP ERP Central ECC, SAP S/4HANA, or SAP BW. It can also be extracted with SAP HANA Database using options such as:
   - A. AWS Managed Services
   - B. SAP and other AWS Partner Solutions with dedicated instances
   - C. AWS Partner Solutions embedded in SAP NetWeaver

Data Integration Options

A — **AWS Glue**, a serverless data integration service, and **Amazon AppFlow** SAP OData connector offer application-level data extraction.

B1 — AWS Partner Solutions such as BryteFlow SAP Data Lake Builder and Qlik Replicate offer instance-based solutions for comprehensive data ingestion scenarios.

B2 — Using SAP native integration, SAP Datasphere or SAP Data Services sends data to **Amazon Simple Storage Service (Amazon S3)** or **Amazon Redshift**.

B3 — SAP SLT replication engine supports replicating data to **Amazon Relational Database Service (Amazon RDS)** using a database connection. AWS Partner Solutions such as Syntax CxLink support streaming data to **Amazon S3** and **Amazon Kinesis** using ABAP add-on for SAP SLT.

C — AWS Partner Solutions embedded in SAP NetWeaver, such as SNP Glue, offer point-to-point data replication from SAP NetWeaver-based source system to the AWS Cloud.

---

2. Data extracted from SAP can be landed in AWS services, such as **Amazon S3**, **Amazon Redshift**, **Amazon Kinesis**, or **Amazon RDS**, combined with non-SAP data, further processed, and analyzed using AWS analytics and generative AI services.

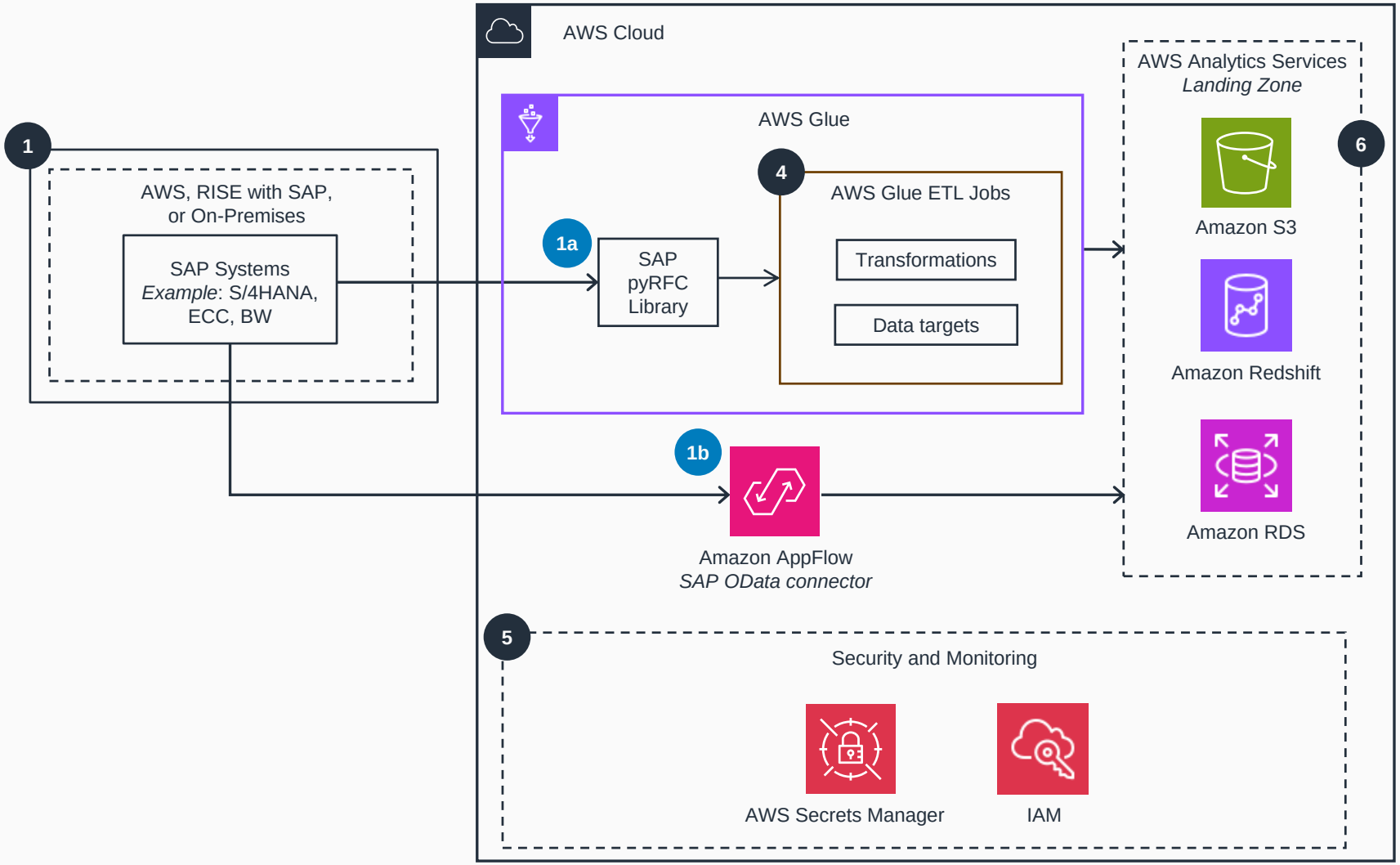
Reviewed for technical accuracy July 9, 2024

AWS Reference Architecture

Guidance for SAP Data Integration and Management on AWS

A. AWS Managed Services

This architecture diagram shows how to ingest SAP data to AWS using AWS Glue.



1. Use the following **AWS Managed Services** options to extract data from SAP:
   - A. **AWS Glue** SAP PyRFC Library for application-level extraction (requires custom design for change data capture)
   - B. **Amazon AppFlow** SAP OData connector (in-built SAP ODP change data capture)

   Design change data capture for **AWS Glue**-based methods using appropriate fields that indicate changed records, such as data change timestamp.

1a. Invoke SAP RFC to extract SAP data using SAP PyRFC library and **AWS Glue** Python modules.

1b. Configure the OData service in the source SAP system, configure the SAP connection using SAP OData connector for **Amazon AppFlow**, create a flow, and schedule the flow or run it on-demand to extract SAP data.

4. **AWS Glue** performs data transformation such as join, union, aggregate, filter, renaming field, dropping fields, adding timestamps, or custom transform.

5. **AWS Secrets Manager** stores credentials. **AWS Identity and Access Management (IAM)** is used for access management and role configurations.

6. Choose destination AWS services, such as **Amazon S3**, **Amazon Redshift**, or **Amazon RDS** as the data target. Data extracted from SAP can be combined with non-SAP data, further processed, and analyzed using AWS analytics and generative AI services.

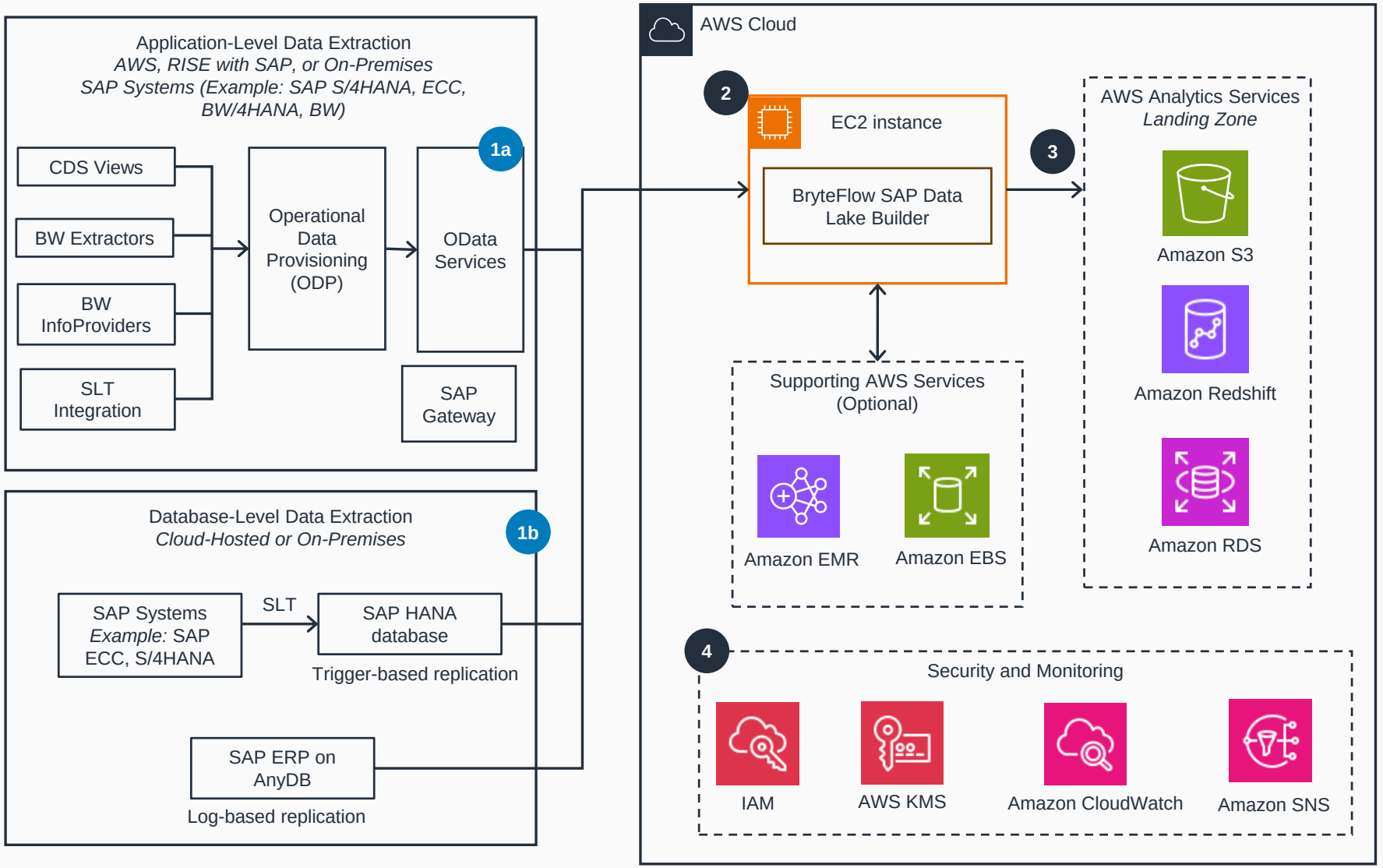
Reviewed for technical accuracy July 9, 2024

AWS Reference Architecture

Guidance for SAP Data Integration and Management on AWS

B1. AWS Partner Solution by BryteFlow

This architecture diagram shows how to ingest SAP data to AWS using the Partner Solution: BryteFlow SAP Data Lake Builder.



1a For application-level data extraction, configure SAP OData services based on CDS views, BW extractors, BW InforProviders, or HANA information views.

1b Database-level data extraction (which requires SAP license that allows database access) uses a trigger-based (SAP HANA database) or log-based mechanism (Oracle, SQL, DB2) to replicate data.

2 The AWS Partner Solution, BryteFlow SAP Data Lake Builder, provides application- and database-level SAP data extraction with change data capture to the AWS Cloud.

BryteFlow SAP Data Lake Builder is available as a pre-configured Amazon Machine Image (AMI) on **AWS Marketplace**. Follow instructions to configure AMI on an **Amazon Elastic Compute Cloud (Amazon EC2)** instance.

3 The captured initial and changed data is ingested by BryteFlow SAP Data Lake Builder software running on an **EC2** instance to AWS analytics services.

Append and update/insert ("upsert") to **Amazon S3**, **Amazon Redshift**, and **Amazon RDS** are supported in this Guidance. **Amazon S3** upsert operations require additional services, such as Amazon EMR and Amazon Elastic Block Service (Amazon EBS). Data catalog and portioning of the schema is configured.

4 BryteFlow SAP Data Lake Builder uses **IAM**, **AWS Key Management Service (AWS KMS)**, **Amazon CloudWatch**, and **Amazon Simple Notification Service (Amazon SNS)** for security, monitoring and alerts.

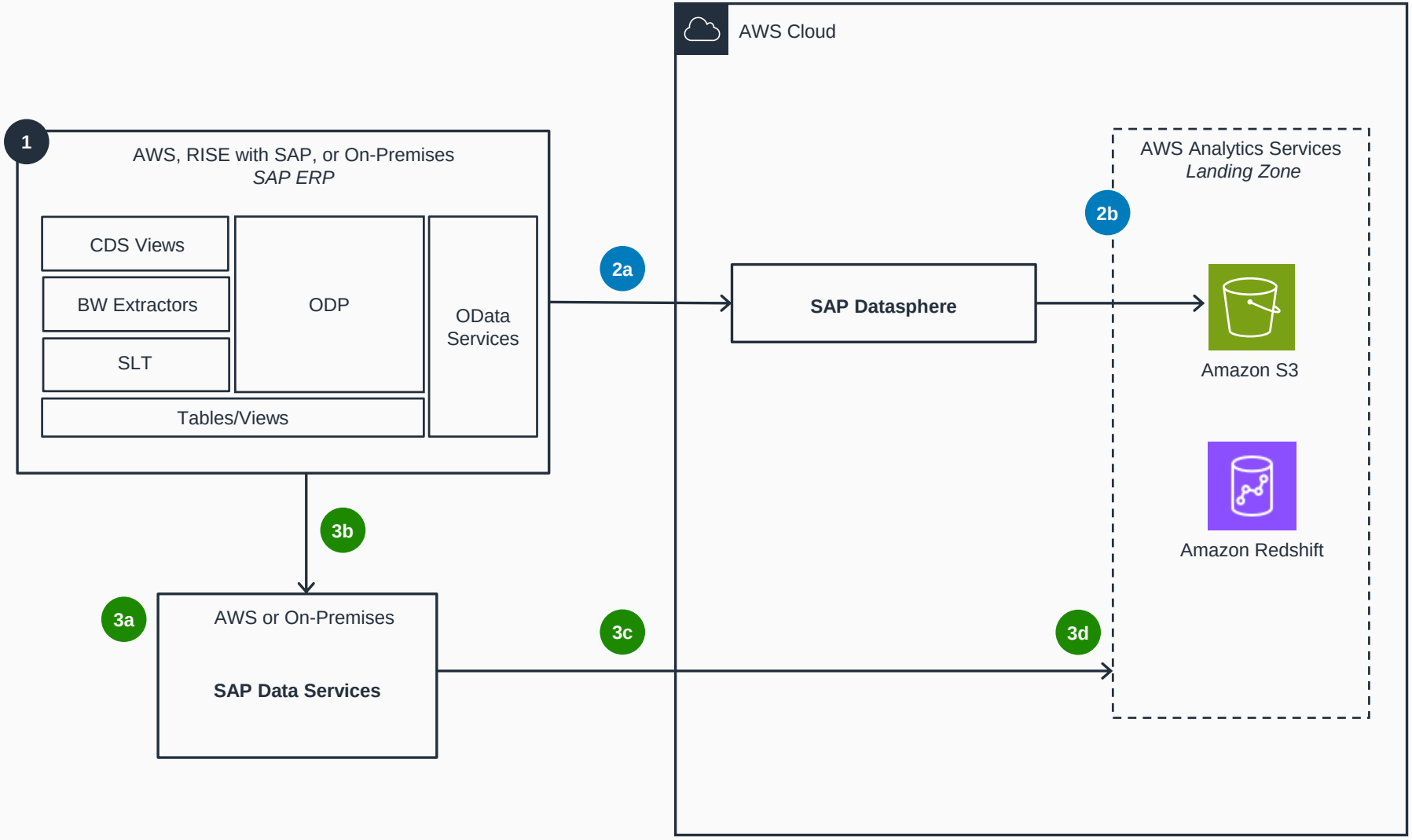
Reviewed for technical accuracy July 9, 2024

AWS Reference Architecture

Guidance for SAP Data Integration and Management on AWS

B2. SAP Datasphere and Data Services

This architecture diagram shows how to ingest SAP data to AWS using SAP Datasphere or SAP Data Services.



1 Data from SAP ERP hosted on RISE, AWS, or on-premises can be extracted using:
A. SAP Datasphere
B. SAP Data Services

SAP Datasphere

2a SAP Datasphere offers various connection types, such as SAP ABAP Connections, SAP ECC Connections, and SAP S/4HANA Cloud Connections, which support RFC and ODP protocols. Refer to SAP Datasphere documentation to choose most appropriate connectivity to extract SAP data.

2b Using premium outbound integration for Amazon Simple Storage Connections, configure the SAP Datasphere replication flow to ingest data to **Amazon S3**.

SAP Data Services

3a Install SAP Data Services on an **Amazon EC2** instance or on-premises.

3b SAP Data Services offers various connections to extract data from SAP ECC data. Refer to SAP Data Services documentation to choose most appropriate connectivity.

3c SAP Data Services offers Amazon Redshift Datastore and Amazon S3 datastore to ingest data to AWS.

3d SAP Data Services offers options for Amazon S3 file location protocol, such as encryption type, compression type, batch size, number of threads, **Amazon S3** storage class, and more.

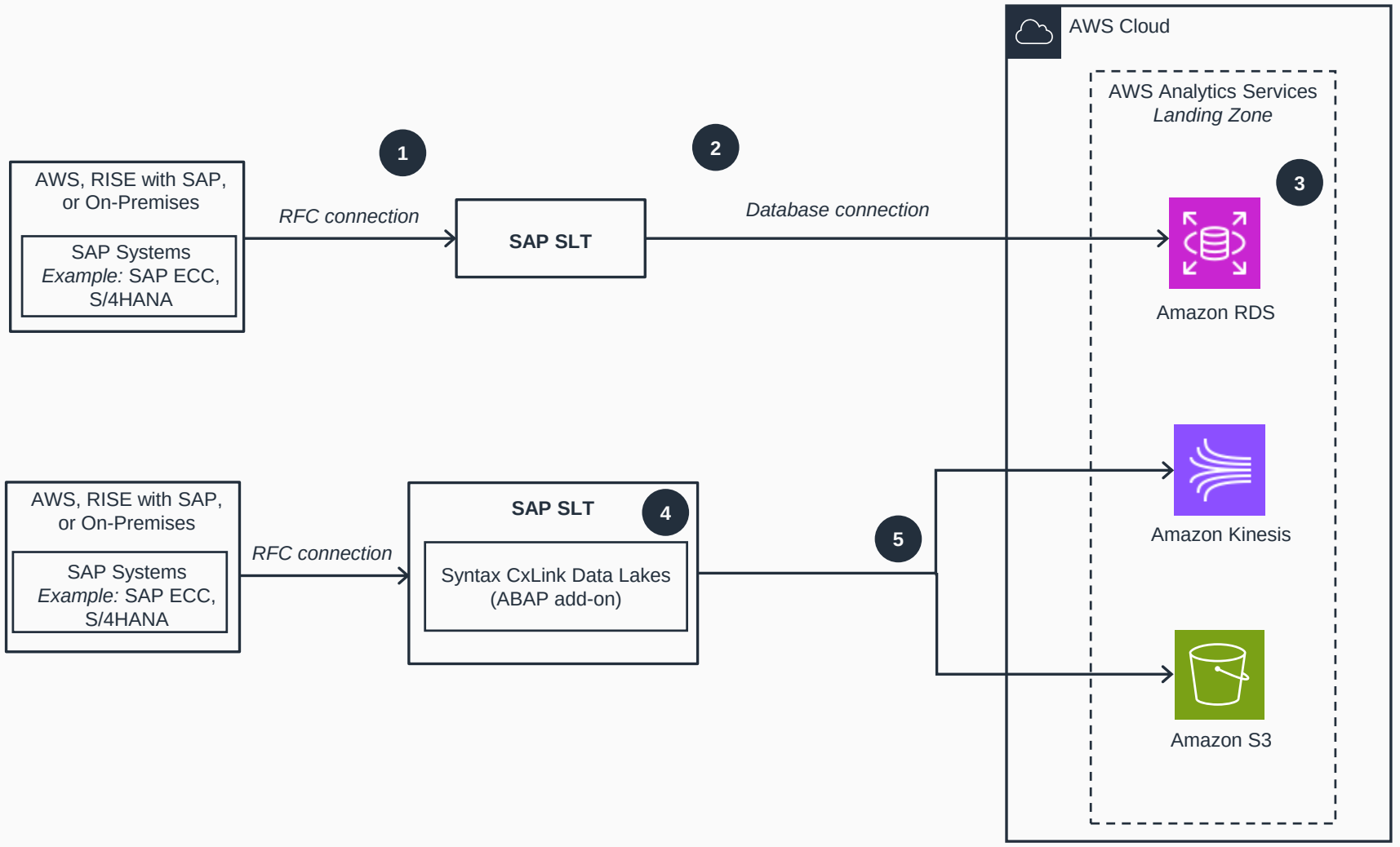
Reviewed for technical accuracy July 9, 2024

AWS Reference Architecture

Guidance for SAP Data Integration and Management on AWS

B3. SAP SLT

This architecture diagram shows how to ingest SAP data to AWS using SAP SLT.



1. Configure RFC destination in SAP SLT to the source SAP ERP system.
2. Configure the SAP SLT database connection to the destination **Amazon RDS** server using a host name, username, and password. Configure SAP SLT mass transfer ID to replicate tables (initial and incremental data) in real-time or at a scheduled frequency to **Amazon RDS**.
3. Insert, update, and delete operations are supported to **Amazon RDS**, which can be used as a landing zone for subsequent data loads to **Amazon S3** or **Amazon Redshift**.
4. For data replication to **Amazon S3** or **Amazon Kinesis**, install an AWS Partner Solution ABAP add-on, such as Syntax CxLink Data Lakes on SAP SLT Server.
5. Syntax CxLink Data Lakes replicates data in real-time or at scheduled frequencies to **Amazon S3** or **Amazon Kinesis**. Incremental data is appended to existing data.

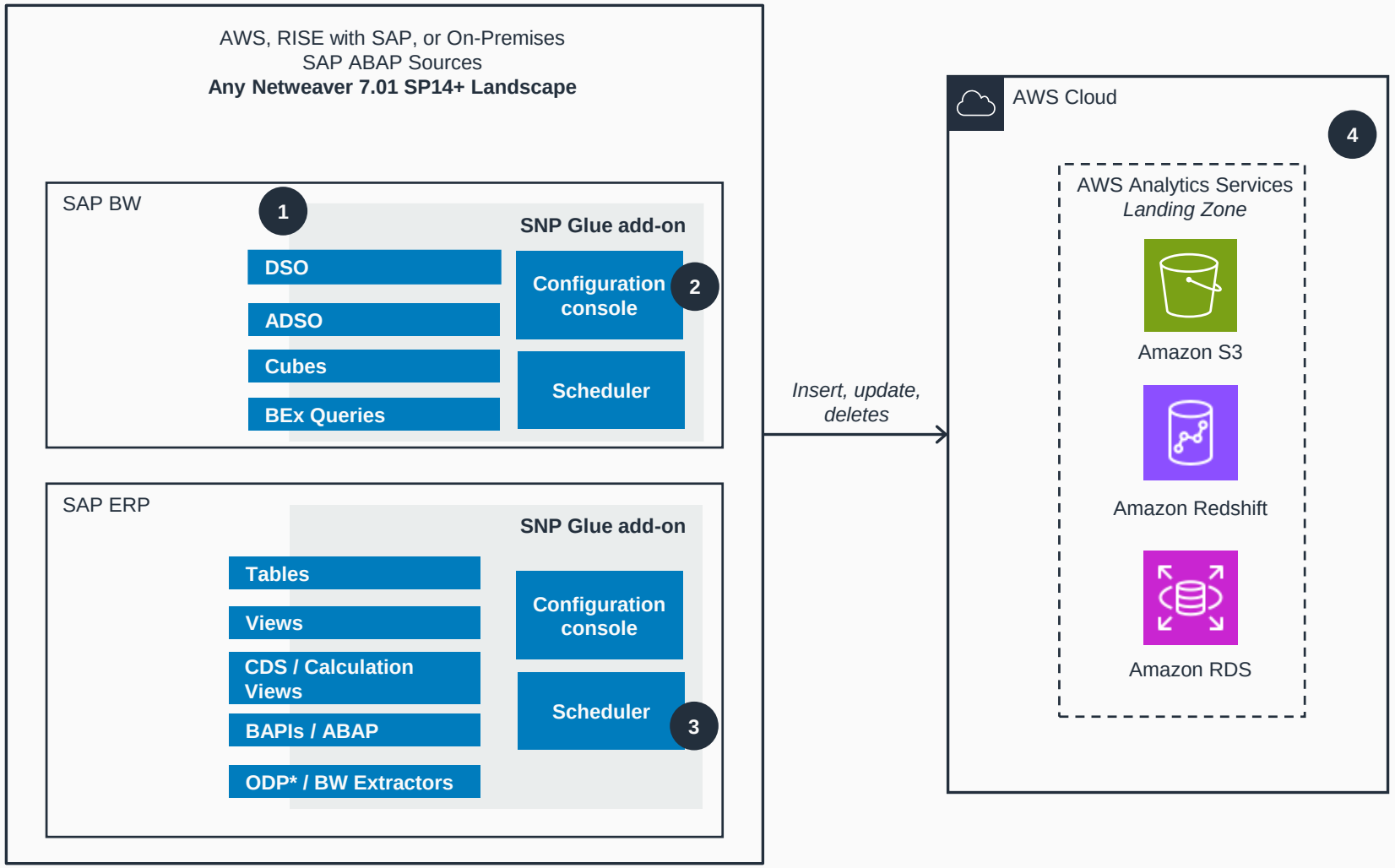
Reviewed for technical accuracy July 9, 2024

AWS Reference Architecture

Guidance for SAP Data Integration and Management on AWS

C. SAP NetWeaver Add-On Solution by SNP

This architecture diagram shows how to use SAP NetWeaver add-on solution SNP Glue to extract data from SAP to AWS.



1. Install and configure SNP Glue ABAP add-on on the SAP ABAP-based source system (such as S/4HANA, ECC, CRM, or BW) Netweaver 7.1 SP14 or higher.
2. SNP Glue configuration workbench allows selection of tables, modification of source and destination structures, data filtering, and addition of transformation rules.
3. SNP Glue scheduler allows creating flexible schedule and throttling SAP resources by limiting the maximum number of background work processes.
4. Initial and incremental data in addition to deletions are captured by SNP Glue and replicated to AWS services such as **Amazon S3** and **Amazon Redshift**.

Reviewed for technical accuracy July 9, 2024

AWS Reference Architecture