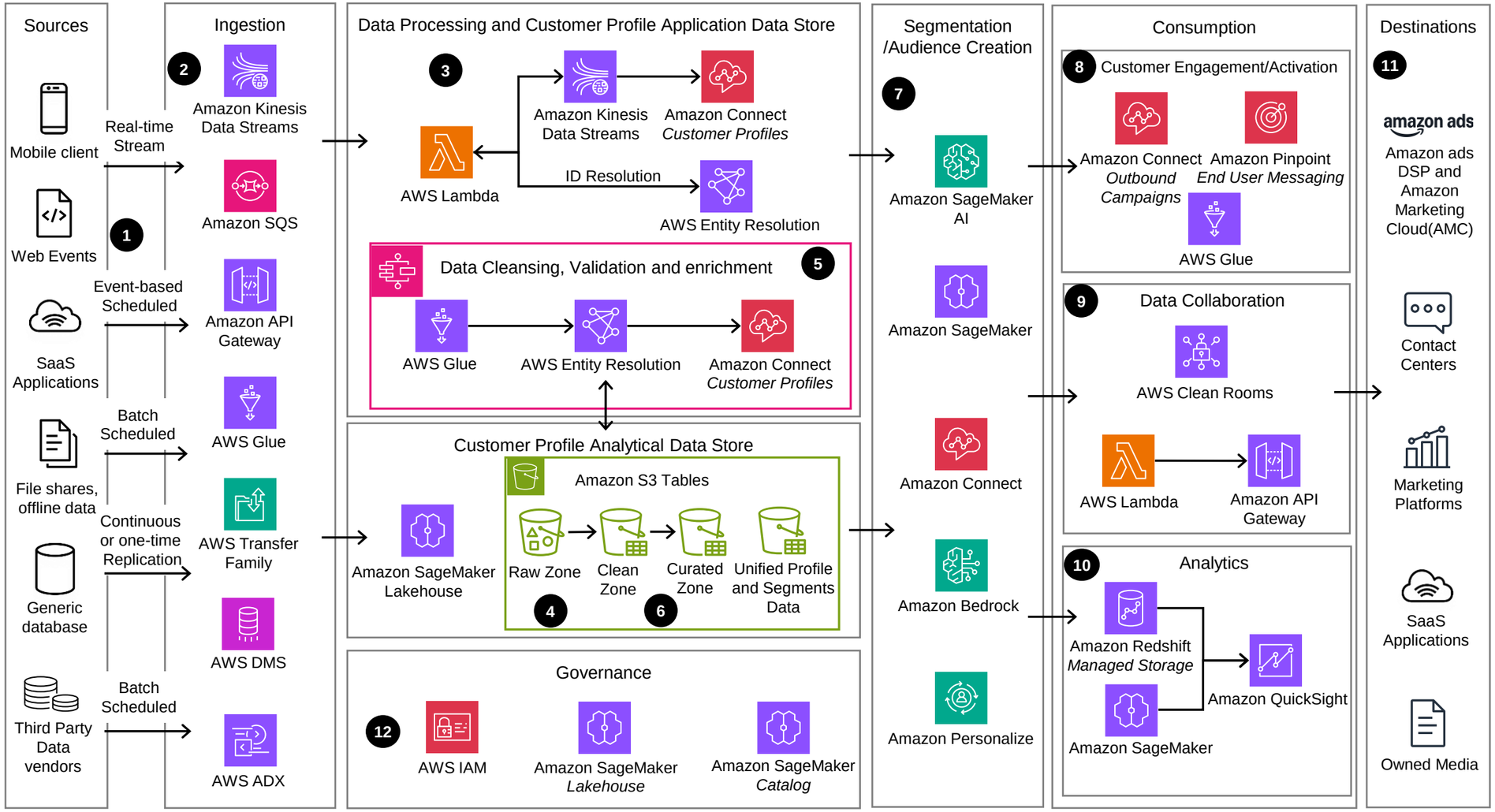
Guidance for Customer Data Platform on AWS

This architecture diagram shows best practices for building a Customer Data Platform (CDP) on AWS from a broad range of sources. It explores all the CDP capabilities - data ingestion, identity resolution, segmentation, analysis and activation and the recommended AWS managed services. This slide shows steps 1-4.



1. Build a Customer 360 profile from data sources including website, mobile application, advertising, and social media events, as well as transactional data from multiple systems of records and external data sets. This data is available for consumption in multiple formats and protocols: SaaS platform API for pull, real-time event push, batch files, Cloud Data Shares, Databases and Data Market Places.

2. Achieve near real-time data ingestion through **Amazon Kinesis Data Streams**, **Amazon Simple Queue Service (Amazon SQS)** and **API Gateway.** Batch data ingestion uses **AWS Transfer Family, AWS Database Migration Service (AWS DMS) and Amazon AppFlow.** Use the **Amazon AppFlow** Connector SDK to build custom connectors to pull data from System of Record APIs. **AWS Data Exchange (AWS ADX)** subscriptions provide access to external data in multiple modes. Use Zero ETL capability where available to move data from AWS transactional datastores.

3. In near real-time data collection, ingestion services collect data, apply real-time data transformations, resolution and store the data in **Amazon Connect** Customer Profiles. Use Generative AI-based data mapping to identify where to store the incoming data.
Use a combination of Customer Profile attributes (Master Data) and Profile objects (Transactional Data) to resolve and store required information. Use calculated attributes feature to generate actionable KPIs that support use cases like customer lifetime spend and other customer insights.
For high volume data like clickstream, use Customer Profile as a pass-through and store the most recent information essential for activating real time personalization use cases like current items in cart, last viewed item, content, etc. Use the near real-time data export feature to send data to Analytical storage.

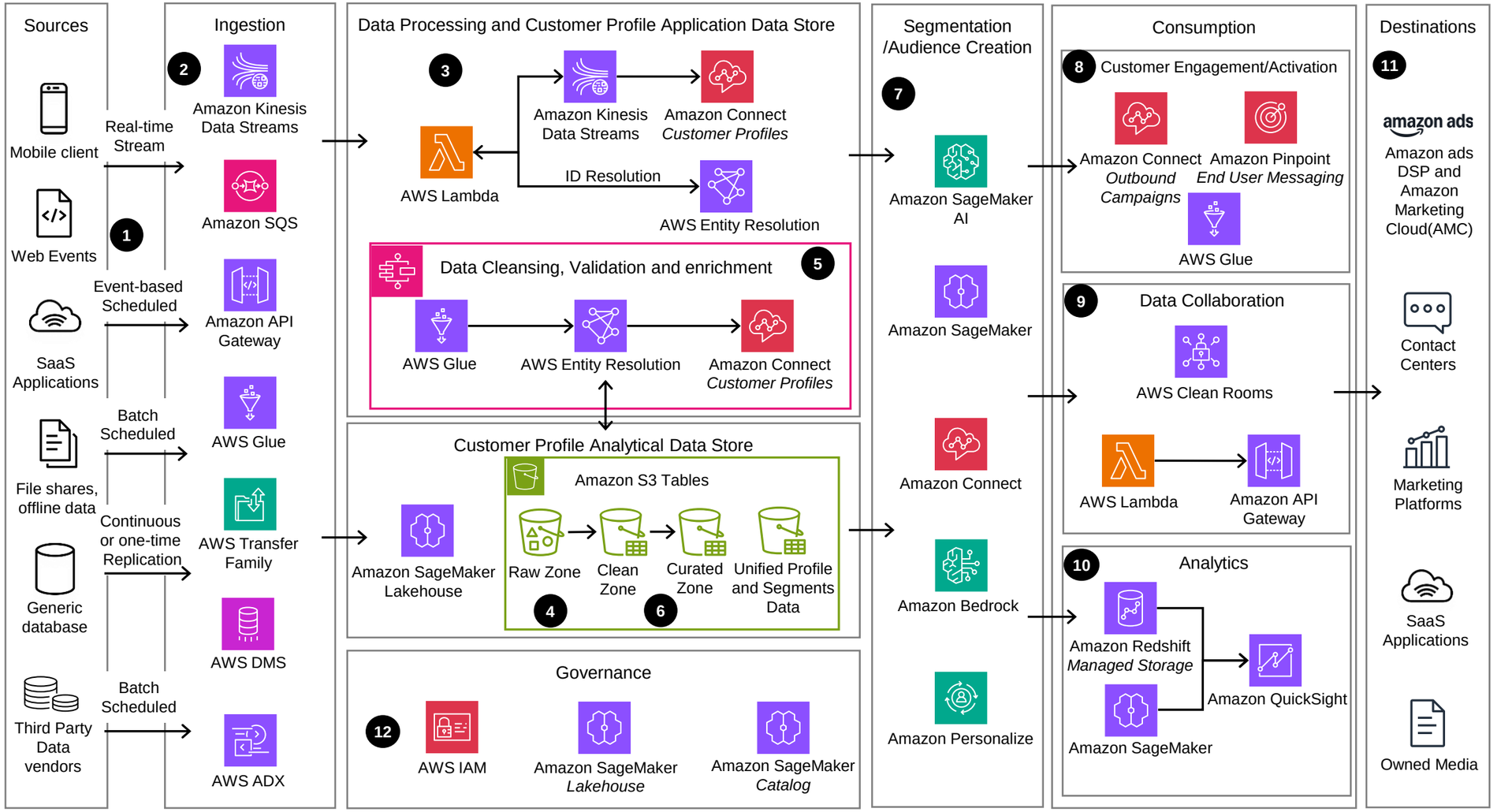
4. In Batch Data Processing, the ingestion services collect and store raw data in **Amazon Simple Storage Service (Amazon S3)**.

Reviewed for technical accuracy October 6, 2025

AWS Reference Architecture

Guidance for Customer Data Platform on AWS

This architecture diagram shows best practices for building a Customer Data Platform (CDP) on AWS from a broad range of sources. It explores all the CDP capabilities - data ingestion, identity resolution, segmentation, analysis and activation and the recommended AWS managed services. This slide shows steps 5-8.



5 For batch sources, use **AWS Step Functions** to orchestrate **AWS Glue** ETL jobs to clean and prepare the data. This data passes to **AWS Entity Resolution** service to match and link records. The resolved records persist in **Amazon S3** and loaded into **Amazon Connect** Customer Profiles. Use this common pattern for initial bulk load and/or batch sources.

6 Data processing and identity resolution workflow transient data storage occur in the logical Clean Zone created using **Amazon S3 Tables**. The logical Curated Zone in **Amazon S3** Tables store the final output of data processing for consumption.

7 Use the unified customer profile stored in **Amazon S3 Tables** for segmentation. Use **Amazon SageMaker** Unified Studio to provision the ideal analytical compute features to data engineers, analysts and scientists. Use **Amazon SageMaker AI** to create Propensity, Churn, and other probabilistic segmentation attributes. Use **Amazon Personalize** recipes to generate product recommendations for upsell and cross-sell, next best action, or offer recommendations to improve other business KPIs.
Use the **Amazon S3** import feature to send relevant segmentation attributes to **Amazon Connect** Profiles and use them in real-time/event-driven customer engagements. Use the **Amazon Bedrock** Prompt Management feature to store Generative AI prompts that allow the creation of hyper-personalized email, SMS, and push notification content. Use **Amazon Bedrock** Guardrails to store organizational policies and validate the generated contents before activating them**. Use Amazon Bedrock** Knowledge Bases to store and generate contextually relevant, personalized content.

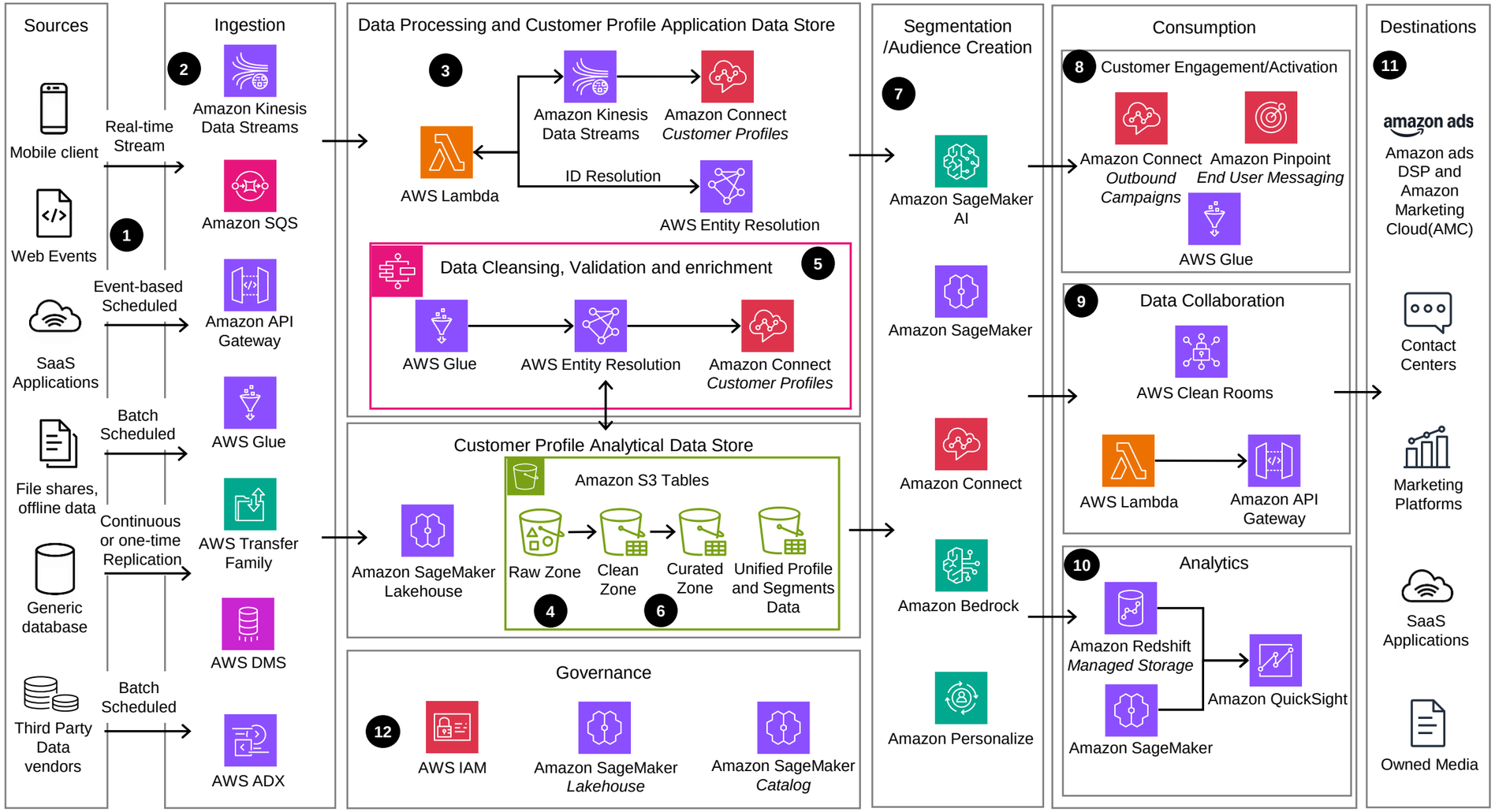
8 Use **Amazon Connect** Outbound Campaigns that access data in **Amazon Connect** Customer Profiles and create proactive multi-channel customer engagements. **Amazon Connect** uses the unified customer profile to enhance the customer experience in call centers. **AWS Glue** software-as-a-service (SaaS) applications integrations allow uploading data to advertising platforms.

Reviewed for technical accuracy October 6, 2025

AWS Reference Architecture

Guidance for Customer Data Platform on AWS

This architecture diagram shows best practices for building a Customer Data Platform (CDP) on AWS from a broad range of sources. It explores all the CDP capabilities - data ingestion, identity resolution, segmentation, analysis and activation and the recommended AWS managed services. This slide shows steps 9-12.



9 Use **AWS Clean Rooms** for privacy-enhanced data collaborations with partners to support media planning, audience activation and measurement use cases. **AWS Lambda** and **Amazon API Gateway** enable API access to Customer 360 profiles.

10 Store clean, modeled data in **Amazon Redshift** for fast and repeated queries. Make all other data in the analytical platform available for business intelligence use cases through the **Amazon SageMaker** Unified Studio SQL analytics feature. **Amazon QuickSight** offers large-scale data analysis and visualization.

11 Upload Customer 360 profile data to paid media Ad Platforms such as **Amazon Marketing Cloud (AMC)** and **Amazon Ads Demand-Side Platform (Amazon Ads DSP)** for online media targeting. Marketing Platforms and other SaaS Applications use the Customer 360 profile data for marketing and data monetization use cases. Owned Media platforms use Customer 360 profile for Website and Mobile App personalization.

12 Use **Amazon SageMaker** Catalog to create a technical and business data catalog for data discovery and data product sharing**.** Use **Amazon SageMaker** LakeHouse for fine-grained access controls on all cataloged data. Use an Apache Iceberg-compatible REST catalog to allow easy access of customer profile data to third-party tools. **AWS Identity and Access Management (AWS IAM)** securely manages identities and access to AWS services and resources.

Reviewed for technical accuracy October 6, 2025

AWS Reference Architecture