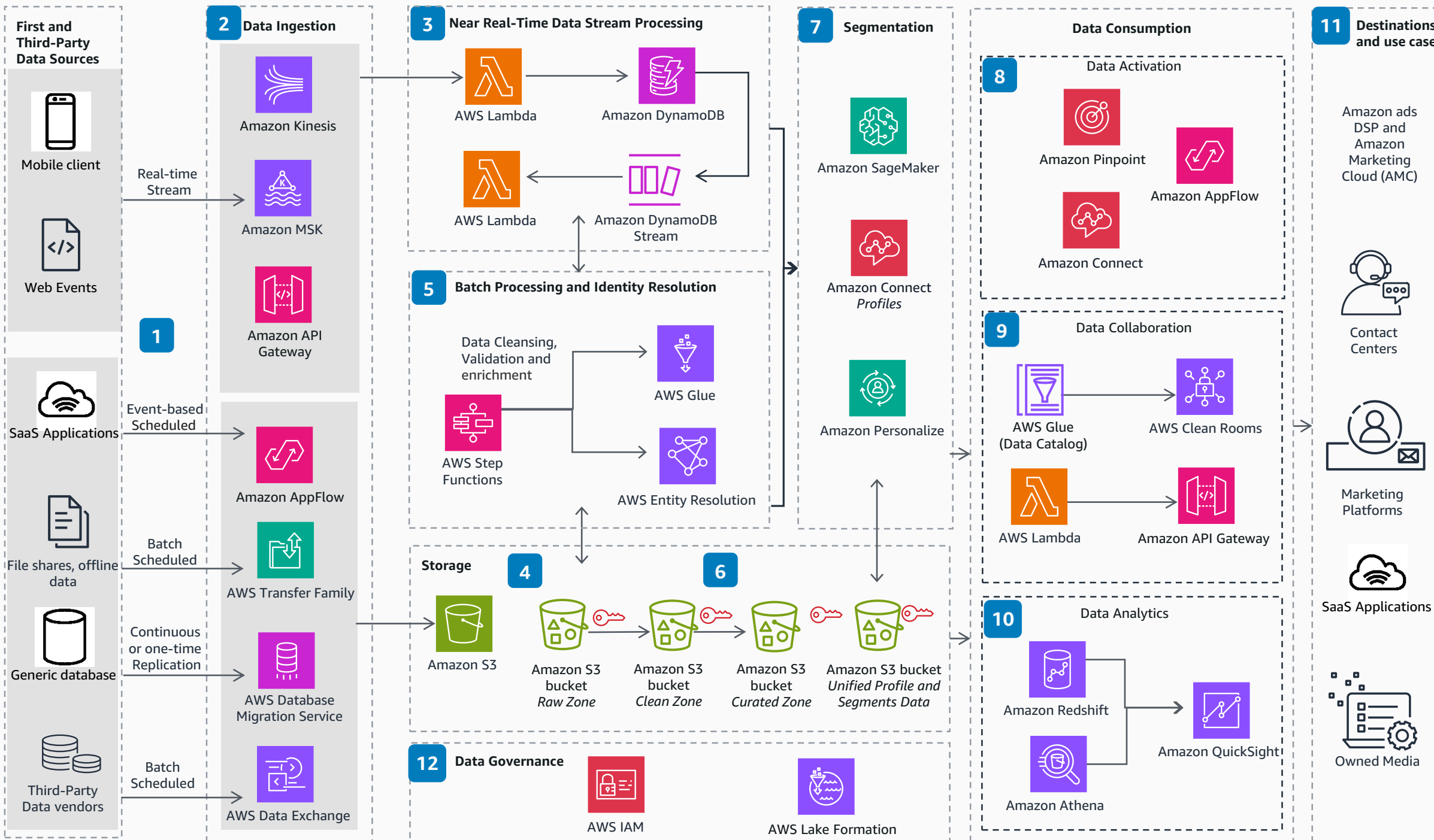


# Guidance for Customer Data Platform on AWS

This Guidance shows how to build a customer data platform with a full, 360 degree profile view of customer data. It explores each stage of building the platform, including data ingestion, identity resolution, segmentation, analysis, and activation.

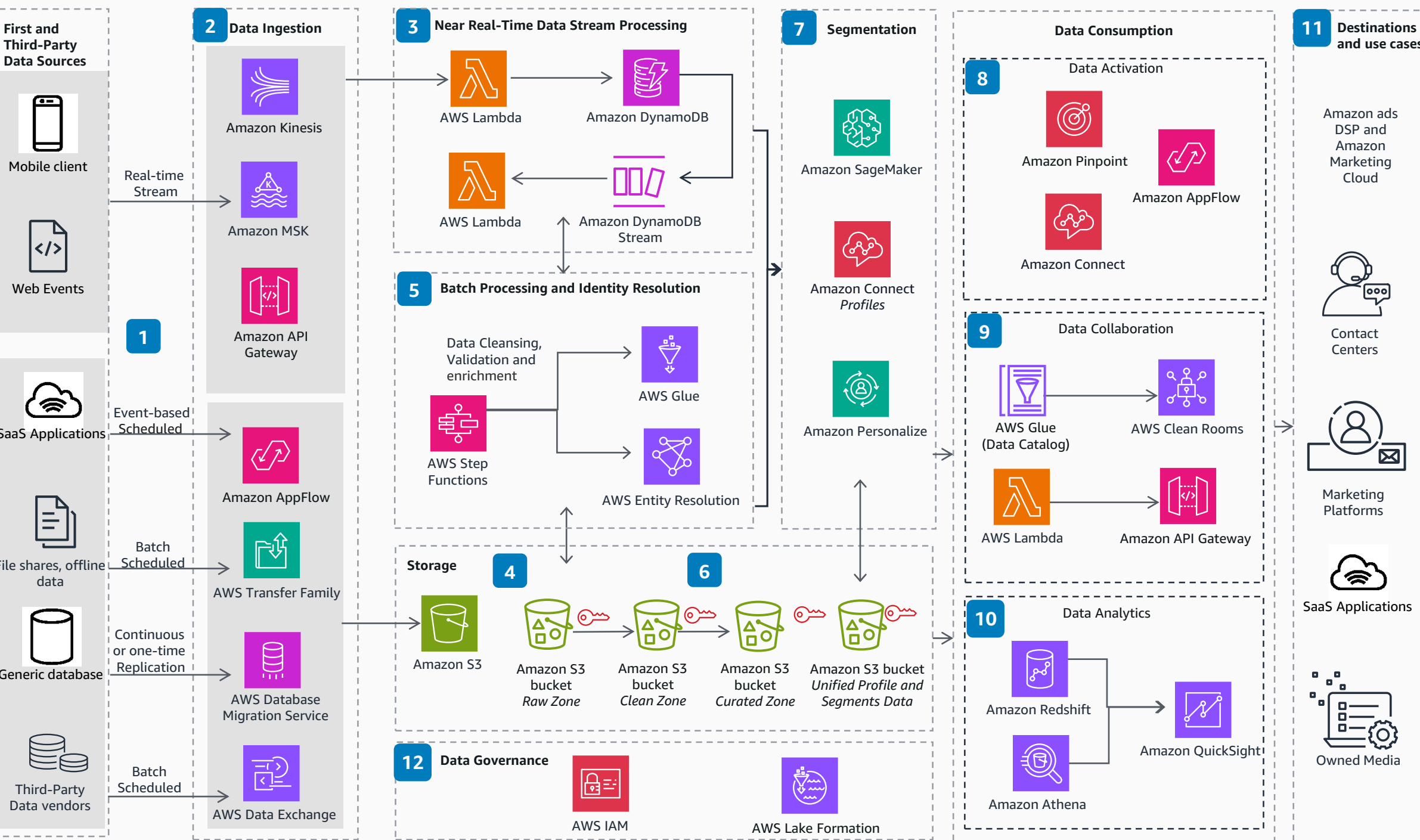


- 1 Data sources for building a customer 360 profile include website and mobile application events, advertising events, social media events, and transactional data from multiple system of records and third-party data sets. This data is available for consumption in multiple formats and protocols. For example, software as service (SaaS) applications, batch files, cloud data shares, databases, and data market places.
- 2 Near real-time data ingestion is achieved through **Amazon Kinesis**, **Amazon Managed Streaming for Apache Kafka (Amazon MSK)** and **Amazon API Gateway**. Batch data ingestion uses **AWS Transfer Family**, **AWS Database Migration Service (AWS DMS)**, and **Amazon AppFlow**. **Amazon AppFlow** Custom Connector Software Development Kit (SDK) is used to build custom connectors to pull data from system of record API's. **AWS Data Exchange** subscriptions provide access to third-party data in multiple modes.
- 3 In near real-time data stream processing, the ingestion services collect data, applies near real-time data transformations using **AWS Lambda**, and stores the data in **Amazon DynamoDB**. A **DynamoDB** stream is used to propagate data downstream in near real-time using **Lambda**.
- 4 In batch data processing, the ingestion services collect and store raw data in **Amazon Simple Storage Service (Amazon S3)**.
- 5 **AWS Step Functions** orchestrates **AWS Glue** data pipeline jobs to clean and validate data. The cleansed data is passed to an Identity Resolution workflow. This workflow is built using **AWS Entity Resolution**.
- 6 Data processing and the transient data storage for the Identity Resolution workflow uses clean zone **Amazon S3** bucket. The **Amazon S3** curated zone bucket stores the final output of data processing for consumption.
- 7 The unified customer profile is stored in **Amazon S3** and used for segmentation. Artificial intelligence and machine learning (AI/ML) models for segmentation are developed and deployed using **Amazon SageMaker**. The unified view of customer profiles for contact center applications is stored in **Amazon Connect Customer Profiles**. Next Best Item recommendations for cross sell or up sell are created from the unified customer view using **Amazon Personalize**.



# Guidance for Customer Data Platform on AWS

This Guidance shows how to build a customer data platform with a full, 360 degree profile view of customer data. It explores each stage of building the platform, including data ingestion, identity resolution, segmentation, analysis, and activation.



- 8** **Amazon Pinpoint** utilizes the unified customer profile to conduct multi-channel outbound marketing. **Amazon Connect** uses the unified customer profile to enhance the customer's experience in call centers. Audience upload to advertising platforms is done using **Amazon AppFlow** integrations.
- 9** **AWS Clean Rooms** is used for privacy enhanced data collaborations to support media planning, audience activation, and measurement use cases. The customer 360 profile is made available for API-based consumption using **DynamoDB**, **Lambda**, and **API Gateway**.
- 10** **Amazon Redshift** stores clean, modeled data for fast and repeated queries. **Amazon QuickSight** provides large-scale data analysis and visualization. **Amazon Athena** enables data exploration and querying.
- 11** Customer 360 profile data is uploaded to paid media ad platforms such as **Amazon Marketing Cloud** and **Amazon DSP** for online media targeting. Marketing platforms and other SaaS solutions use the customer 360 profile data for marketing and data monetization use cases. Media platforms use customer 360 profiles for website and mobile app personalization.
- 12** **AWS Lake Formation** defines access controls on **AWS Glue** catalog tables, columns, and rows in the data lake. **AWS Identity and Access Management (IAM)** securely manages identities and access to AWS services and resources.

