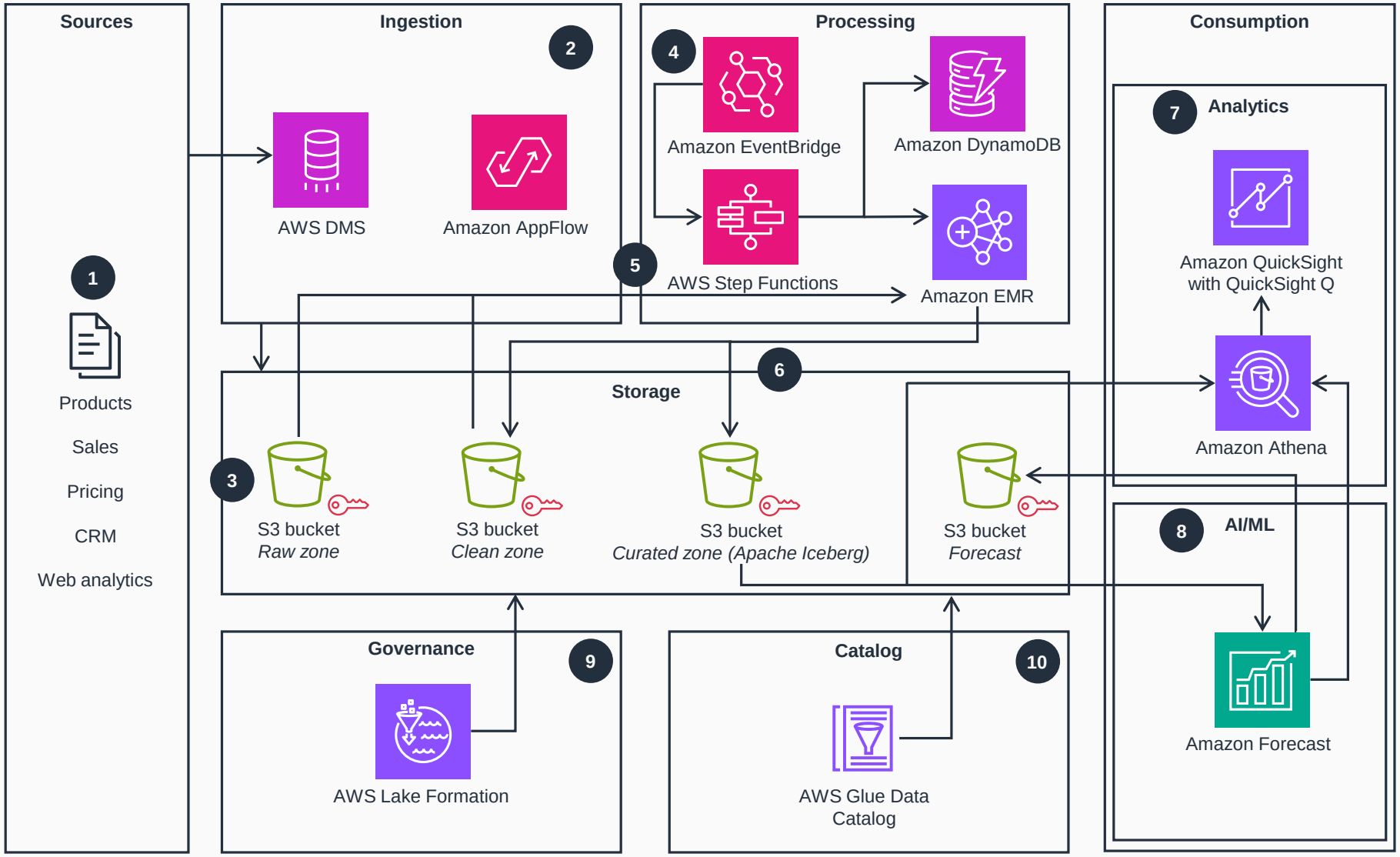
Guidance for Product 360 on AWS

This architecture diagram shows how to use AWS services for data ingestion and transformation in addition to AI-based forecasting and analytics.



1. Sources for Product 360 include product, customer relationship management (CRM), sales transaction, pricing, customer interaction, and clickstream data.
2. **AWS Database Migration Service (AWS DMS)** ingests data from database and analytical sources, and **Amazon AppFlow** ingests data from software-as-a-service (SaaS) services.
3. Ingested data is sent in its original, immutable format to an **Amazon Simple Storage Service (Amazon S3) r**aw zone bucket.
4. Data processing and pipeline orchestration is conducted using purpose-built data processing components and transformation libraries through **AWS Step Functions** and **Amazon EMR**. **Amazon DynamoDB** stores pipeline configuration and schema information about data sources.
5. **Amazon EventBridge** initiates **Step Functions**, which invokes **Amazon EMR** to transform raw data into an efficient data format (Parquet). This data is moved to a clean zone **S3** bucket.
6. **EventBridge** initiates **Step Functions** again, which invokes **Amazon EMR** to use Apache Spark-based batch and streaming pipelines to create Apache Iceberg-based data in the curated zone **S3** bucket from the clean zone **S3** bucket.
7. The analytics layer uses **Amazon QuickSight** and **Amazon Athena** to natively integrate with the curated zone for analytics, dashboards, ad-hoc reporting, and ML.
8. **Amazon Forecast** forecasts product demand, and an **S3** bucket stores forecast output. **Athena** and **QuickSight** query and visualize the forecast output.
9. Using **AWS Lake Formation**, Product 360 users will get fine-grained access to data assets in the curated zone and forecast **S3** buckets for consumption.
10. **AWS Glue Data Catalog** stores technical metadata for all data assets in **S3** buckets, which is used for querying the data assets from **Athena** and **QuickSight**.

Reviewed for technical accuracy September 25, 2023

AWS Reference Architecture