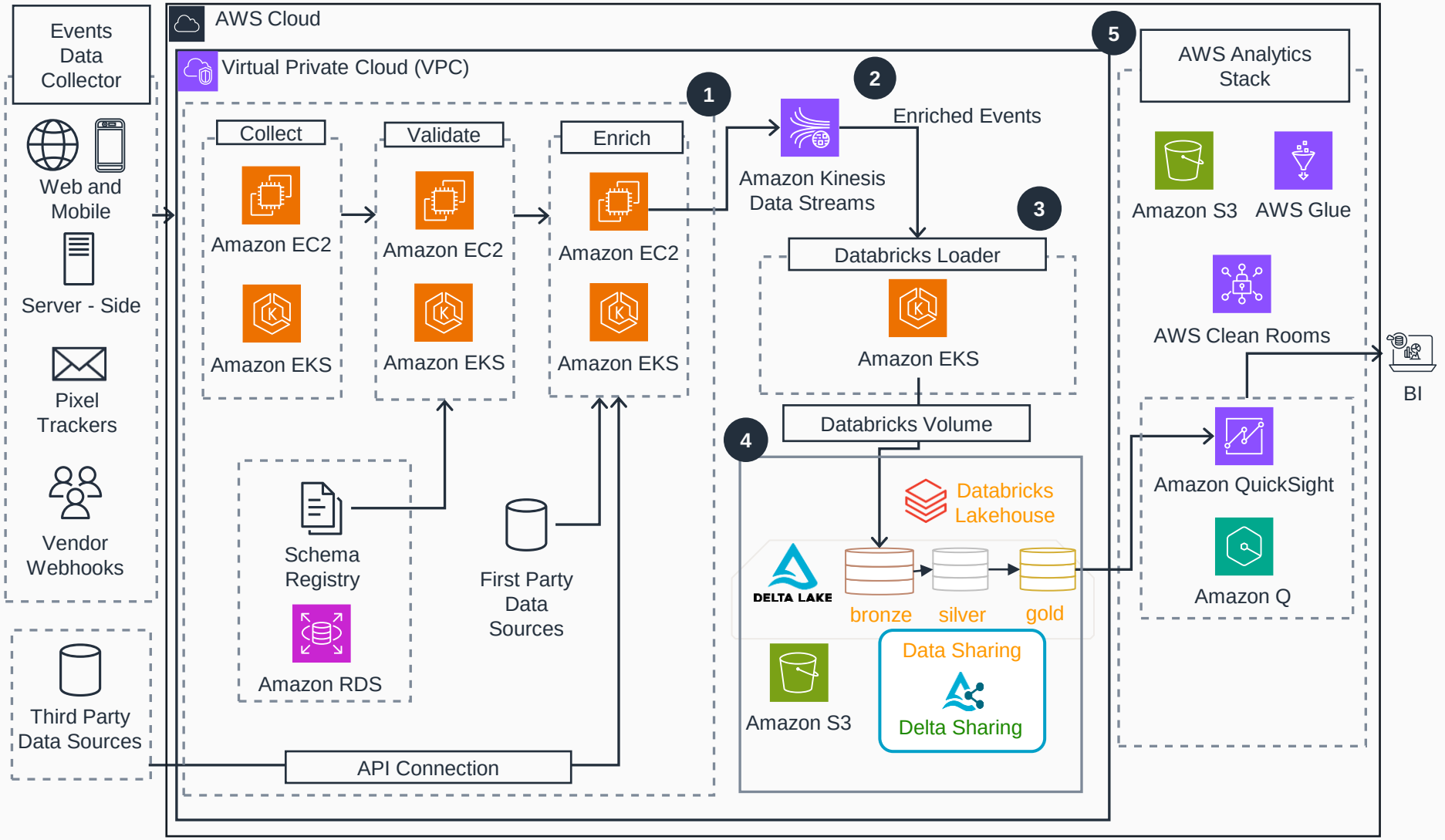
Guidance for Composable Web Analytics on AWS

This architecture diagram demonstrates how to leverage collected events using Snowplow Pipeline, along with Databricks Lakehouse and AWS Services, to deliver AI powered Business Intelligence. This slide shows steps 1-4.



1 Snowplow's Collect application automatically captures user interactions from websites and mobile apps through embedded code snippets. The application runs on Amazon Elastic Compute Cloud **(Amazon EC2)** via Amazon Elastic Kubernetes Service **(Amazon EKS)**.

The Validate application validates this event data against the predefined schema stored in the schema registry. The Information is stored in Amazon Relational Database Service **(Amazon RDS PostgreSQL)**.

The optional Enrich application uses integrations built with first-party and third-party data sources through APIs to enrich the events for better insights.

2 Post enrichment, the event data is sent to Amazon Kinesis Data Streams in near real-time.

Amazon Kinesis Data Streams delivers enriched events to a custom Databricks loader application in **Amazon EKS**.

3 The Databricks loader stores raw data to the existing Databricks volume (bronze layer) which is backed by Amazon Simple Storage Service **(Amazon S3)**, ensuring privacy and ownership.

4 Databricks Lakehouse processes the raw data using Snowplow's Data models (dbt models) on Databricks compute to create structured, analysis-ready datasets (sliver/gold layers) for better insights.

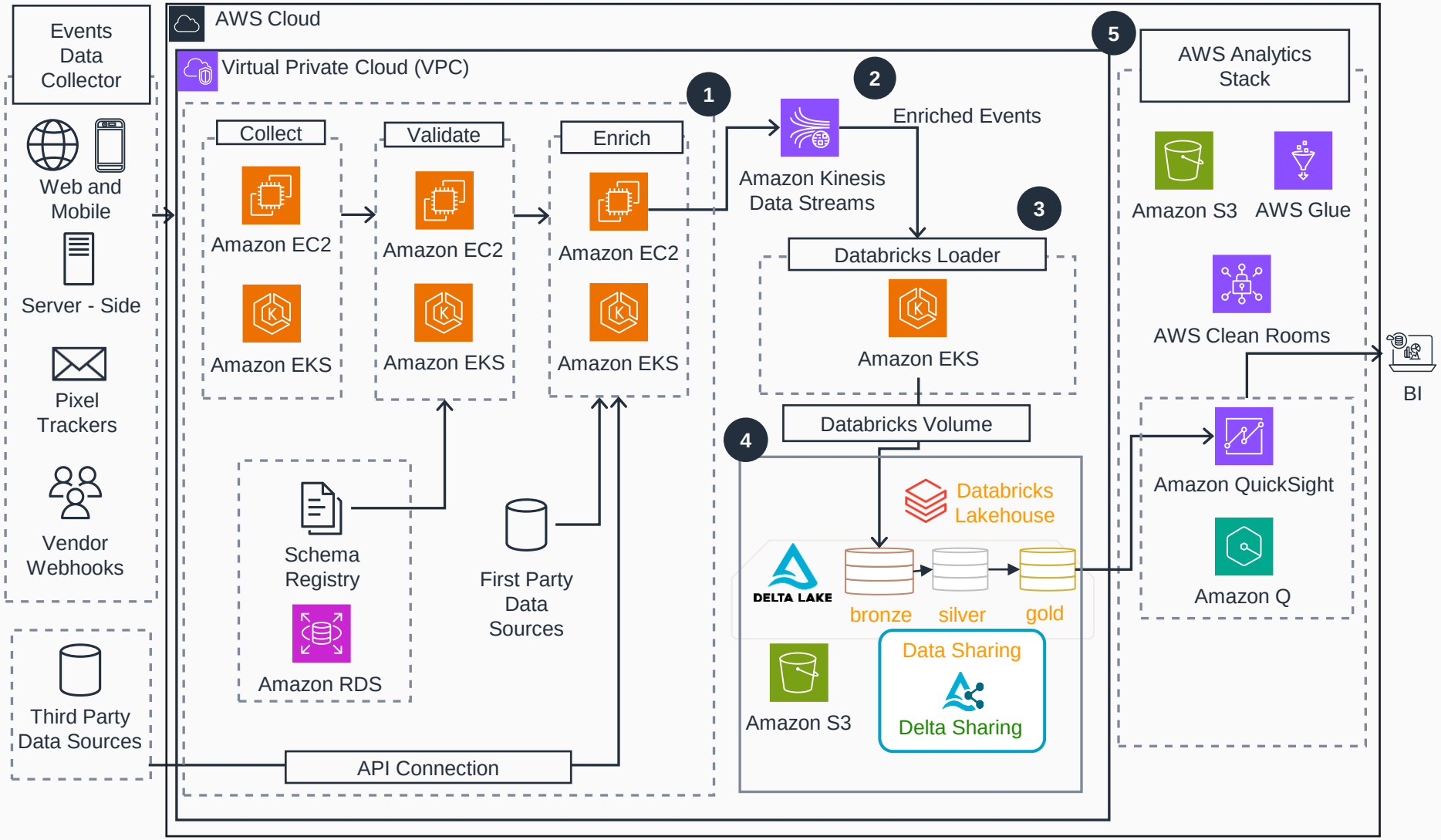
Reviewed for technical accuracy September 8, 2025

© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Reference Architecture

Guidance for Composable Web Analytics on AWS

This architecture diagram demonstrates how to leverage collected events using Snowplow Pipeline, along with Databricks Lakehouse and AWS Services, to deliver AI powered Business Intelligence. This slide shows step 5.



5 The Analytics Stack supports Generative BI capabilities for customers to unlock additional use cases using AWS Services.

- Use Amazon QuickSight for User Journey visualization. **QuickSight** provides unified Business Intelligence (BI) capabilities for Analytics and dashboarding to support various personas within the organization and externally.
- Use Amazon Q in QuickSight to create reports and dashboards, and to interact with web analytics data using natural language to analyze user journeys and identify customer segments.
- Use web interaction data in AWS Clean Rooms and collaborate with publisher partners to perform marketing conversion analysis and measurements in a privacy enhanced manner. This reduces the need for third-party trackers on the web and mobile apps required for marketing measurements.

Reviewed for technical accuracy September 8, 2025

© 2024, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Reference Architecture