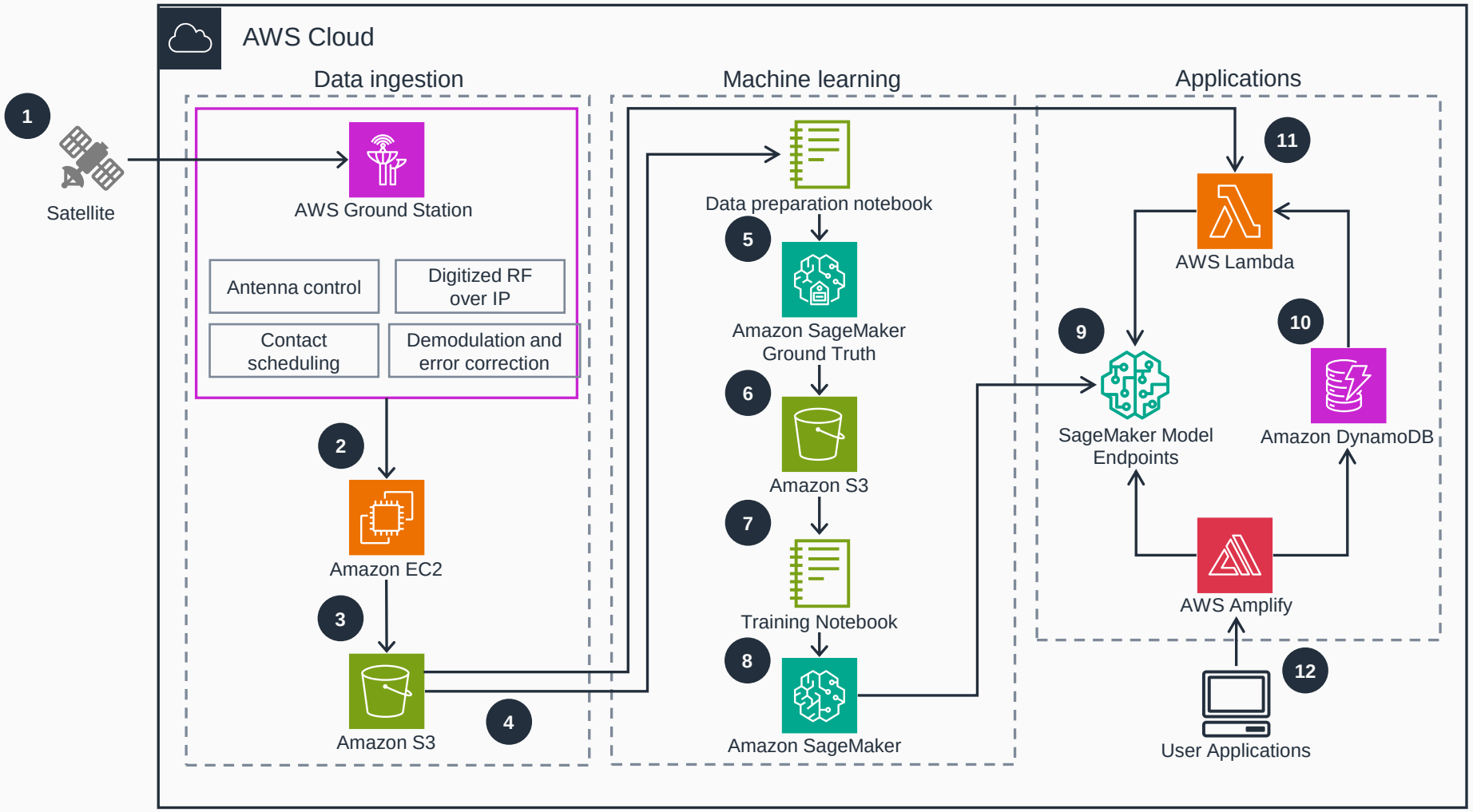
Guidance for Running Machine Learning Algorithms with Satellite Data on AWS

This architecture diagram shows how to ingest satellite imagery through AWS Ground Station using Amazon SageMaker to label data, train ML models, and report inferences to your applications.



1. Satellite sends data and imagery to the **AWS Ground Station** antenna.
2. **AWS Ground Station** delivers baseband or digitized radio frequency (RF)-over-IP data to an **Amazon Elastic Cloud Compute (Amazon EC2)** instance.
3. The **Amazon EC2** instance receives and processes the data and then stores the data in an **Amazon Simple Storage Service (Amazon S3)** bucket.
4. A data preparation notebook ingests data from the **Amazon S3** bucket to prepare the data for model training.
5. **Amazon SageMaker Ground Truth** labels the images.
6. The labeled images are stored in the **Amazon S3** bucket.
7. The notebook hosts the training algorithm and code.
8. **Amazon SageMaker** runs the training algorithm on the data and trains the ML model.
9. **SageMaker** deploys the ML models to an endpoint.
10. The **SageMaker** ML model processes the image data and stores the generated inferences and metadata in the **Amazon DynamoDB** database.
11. Image data received into **Amazon S3** automatically triggers an **AWS Lambda** function to run ML services on the image data.
12. Applications interact with **AWS Amplify** to access the ML algorithm and database.

Reviewed for technical accuracy October 29, 2024

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

AWS Reference Architecture