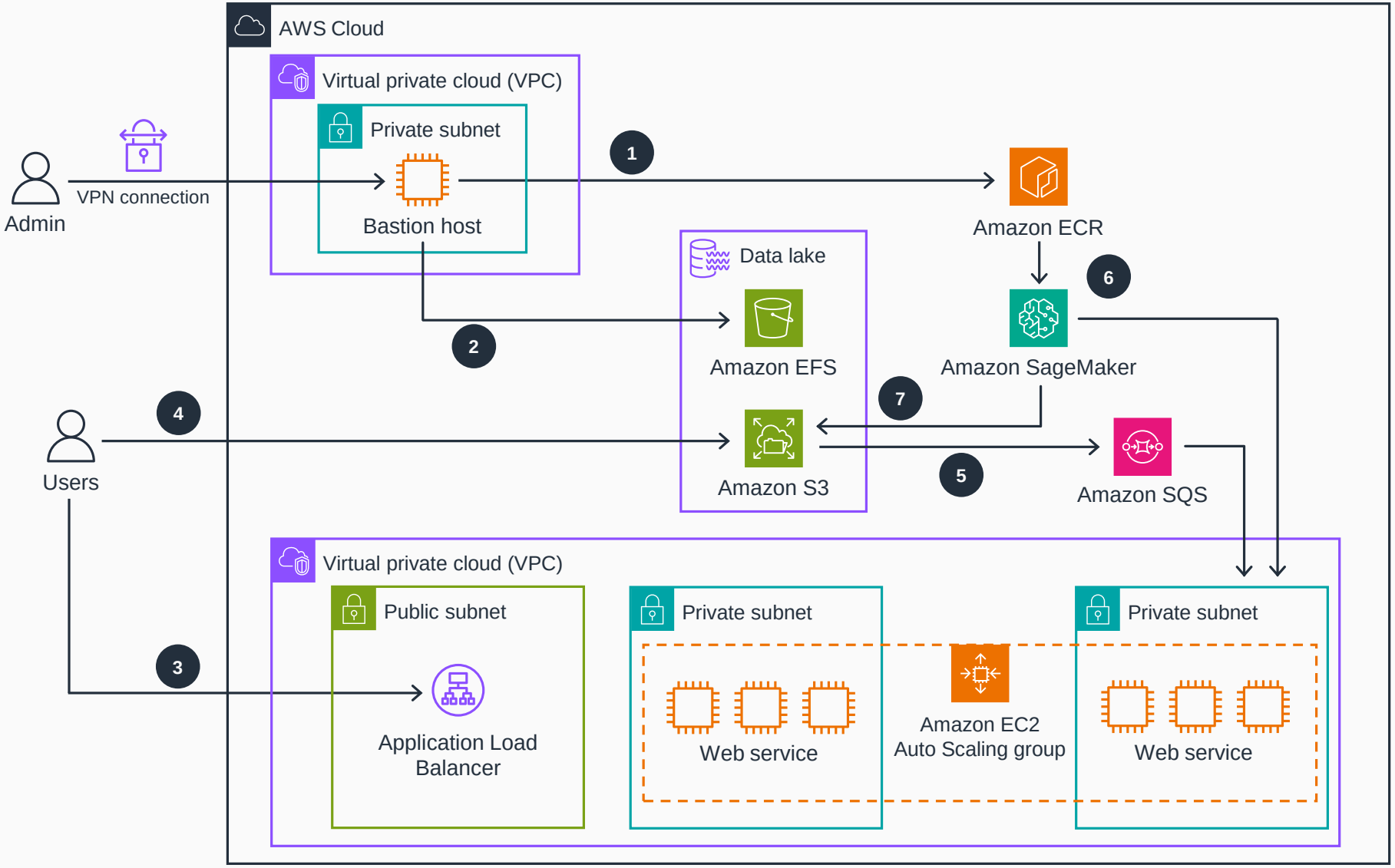
Guidance for AI Training for Smart Products on AWS

This architecture diagram shows how to build a secure and flexible multi-tenant artificial intelligence (AI) training environment for smart products.



1. The platform owner builds a custom model with a “Bring-Your-Own-Model” and registers the model in an **Amazon Elastic Container Repository (Amazon ECR)**.
2. The platform owner uploads data as training source data via the **Amazon Elastic File System (Amazon EFS)**.
3. Users log in to the web portal via an **Application Load Balancer**. Users can prepare training scripts, submit training jobs, and check training status.
4. Users upload their own datasets to **Amazon Simple Storage Service (Amazon S3)** as another source for training data.
5. Once the data is uploaded to **Amazon S3**, an event is triggered via **Amazon Simple Queue Service (Amazon SQS)** and users are notified to submit training jobs from the portal.
6. **Amazon SageMaker** pulls the image from **Amazon ECR**, and the training job is launched on the web service.
7. When the training job is done, **SageMaker** exports the output model to **Amazon S3** so users can download the model and deploy it on smart devices.

Reviewed for technical accuracy February 9, 2024

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

AWS Reference Architecture