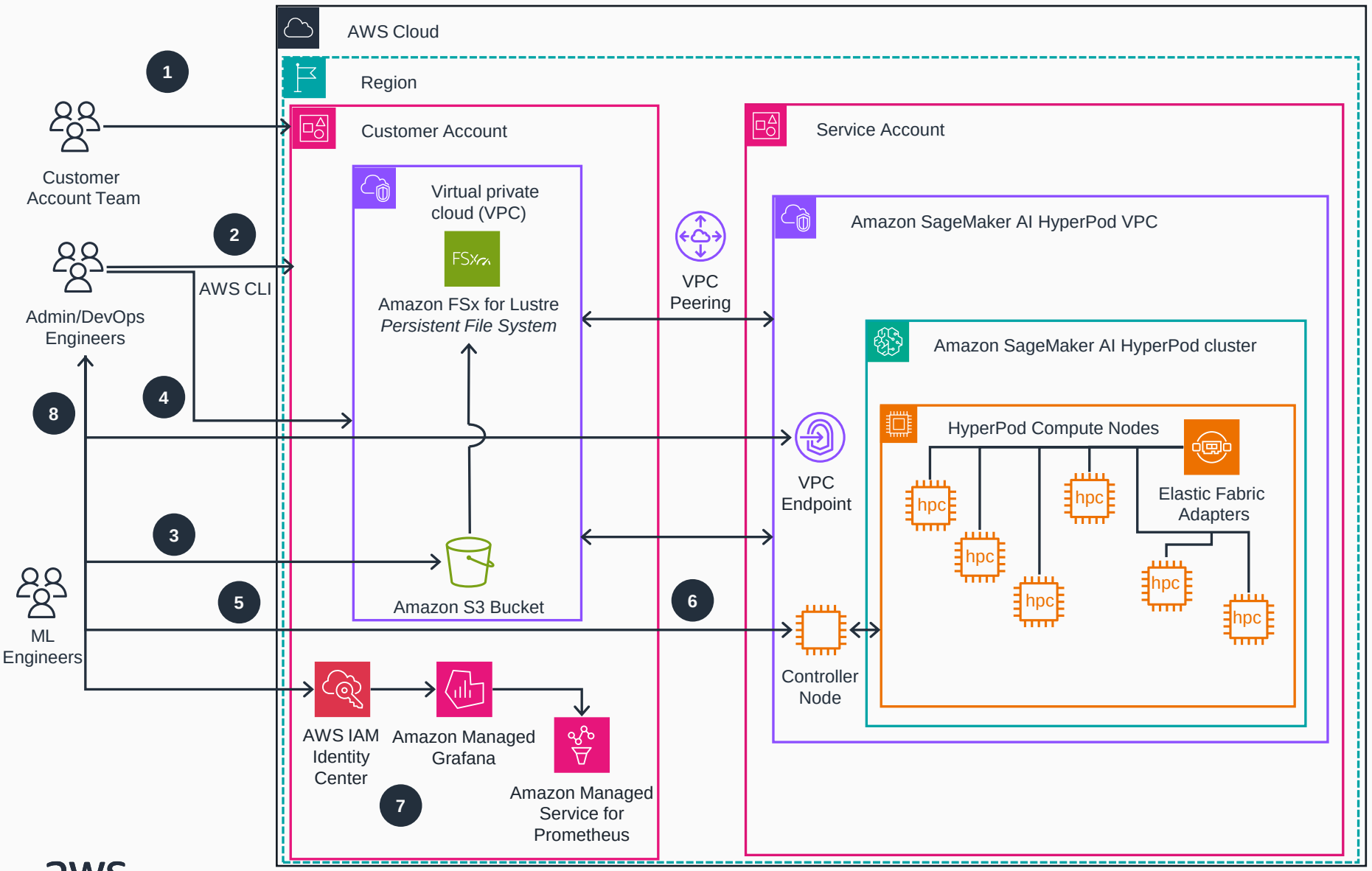
Guidance for Training Protein Language Models (ESM-2) with Amazon SageMaker AI HyperPod

This reference architecture demonstrates how to deploy Amazon SageMaker AI HyperPod clusters based on HPC (SLURM) orchestrator.



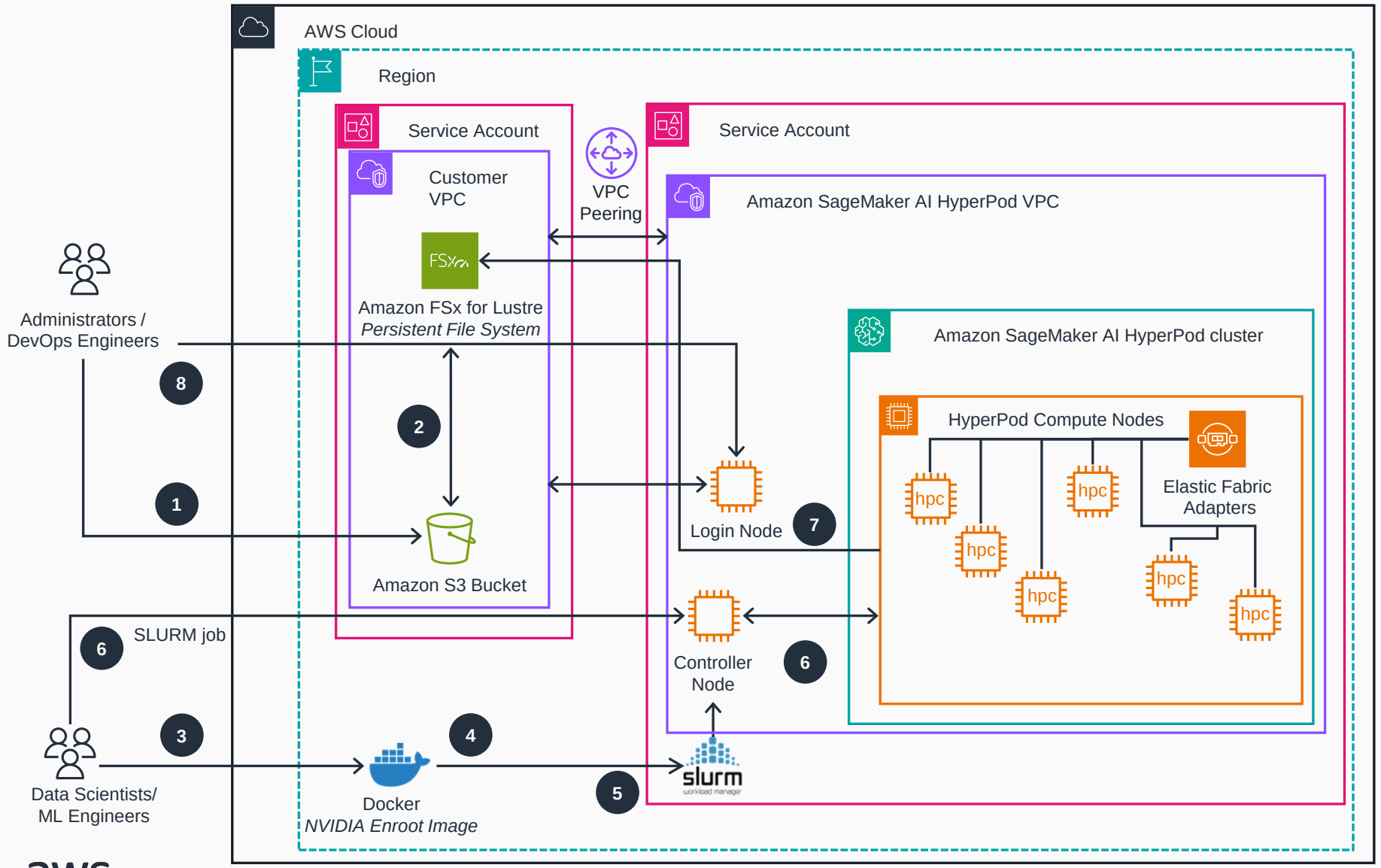
1. The account team reserves compute capacity with On-Demand Capacity Reservations (ODCR) or Amazon SageMaker HyperPod Flexible Training Plans for projected jobs.
2. Admin/DevOps Engineers use the Amazon Sagemaker HyperPod Virtual Private Cloud VPC **Amazon CloudFormation** stack to deploy networking, storage and Identity and Access Management IAM resources into the Customer Account.
3. Admin/DevOps Engineers push Lifecycle scripts to the designated **Amazon Simple Storage Service (Amazon S3)** bucket created in the previous step.
4. Admin/DevOps Engineers use the Amazon SageMaker AI CLI to provision the SageMaker AI HyperPod cluster.
5. Admin/DevOps Engineers generate key-pairs to establish SSH access to the Controller Node of the HyperPod cluster.
6. Once the HyperPod cluster is created, Admin/DevOps or ML Engineers can test SSH access to the Controller and Compute Nodes and examine the cluster.
7. Admin/DevOps Engineers configure IAM Identity Center to use Amazon Managed Service for Prometheus for collection of cluster metrics and Amazon Managed Grafana to set up the observability stack.
8. Admin/DevOps Engineers can make further changes to the cluster using the HyperPod CLI.

Reviewed for technical accuracy September 2, 2025

AWS Reference Architecture

Guidance for Training Protein Language Models (ESM-2) with Amazon SageMaker AI HyperPod

This reference architecture demonstrates how to run distributed ESM-2 model training jobs on a SLURM based HyperPod cluster.



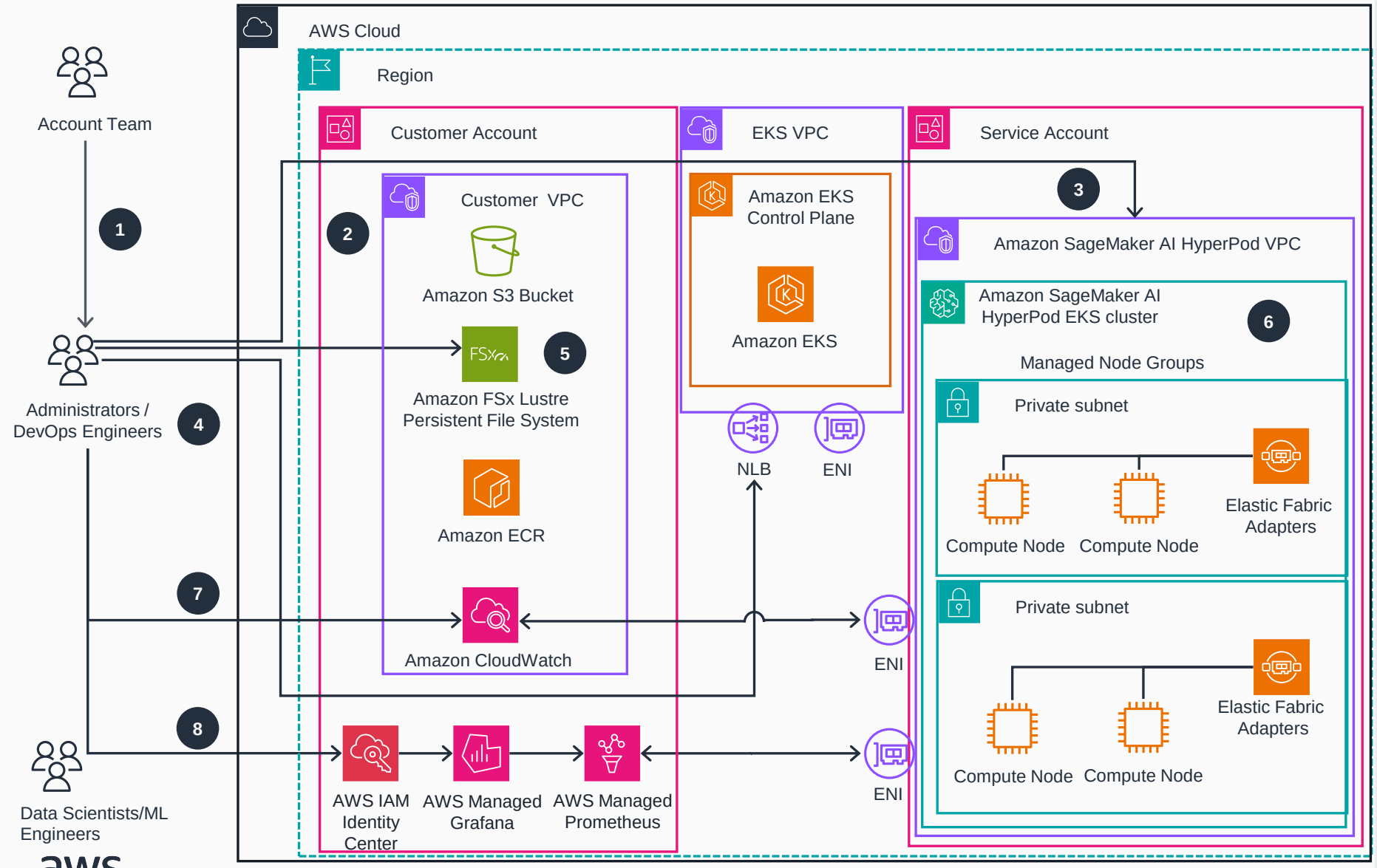
1. Admin/DevOps Engineers move their training data from on-premise to an Amazon Simple Storage Service (Amazon S3) bucket.
2. Admin/DevOps Engineers create Data Repository Associations between an Amazon S3 bucket and Amazon FSx for Lustre.
3. Data Scientists/ML Engineers build AWS optimized Docker container images with a set base image.
4. Data Scientists/ML Engineers create an NVIDIA Enroot image based on the Docker image.
5. Data Scientists/ML Engineers create a SLURM training jobs submission script using the NVIDIA Enroot image.
6. Data Scientists/ML Engineers submit model training SLURM jobs that reference the ESM dataset and use container images built in steps 3 and 4 to run on the **Amazon SageMaker AI** HyperPod compute nodes via Controller Node.
7. **Amazon SageMaker AI** HyperPod SLURM cluster compute nodes run training job tasks and write checkpoints to the shared **FSx for Lustre** file system. Data Scientists can monitor the training process via logs to determine when the training job is completed.
8. (Optional) Admin/DevOps Engineers can create Login Nodes for Data Scientists/ML Engineers to only submit training jobs but without access to making cluster configuration changes.

Reviewed for technical accuracy September 2, 2025

AWS Reference Architecture

Guidance for Training Protein Language Models (ESM-2) with Amazon SageMaker AI HyperPod

This reference architecture demonstrates how to deploy SageMaker HyperPod clusters based on Amazon EKS orchestrator.



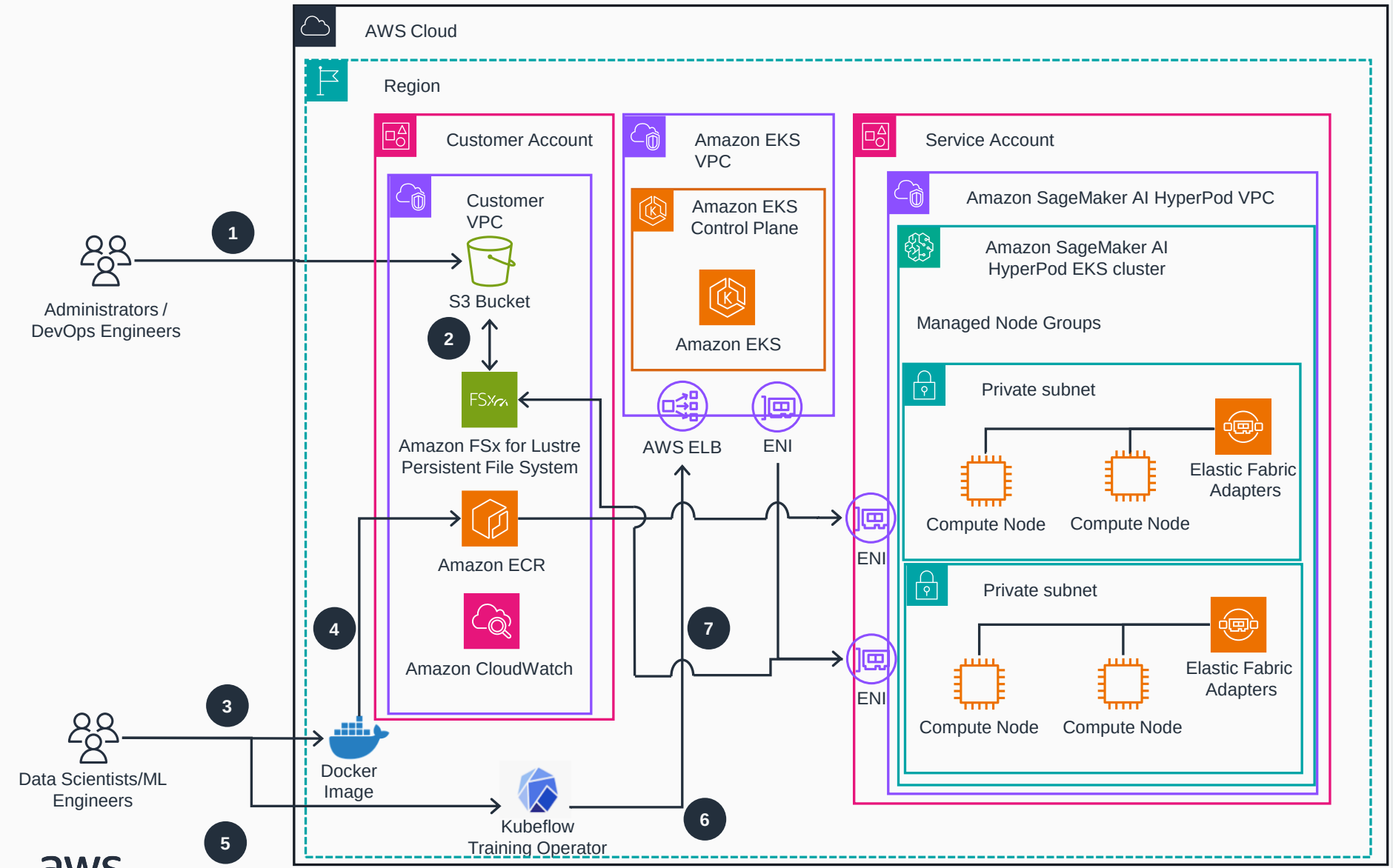
1. Account team reserves capacity with ODCRs or **Amazon SageMaker AI** HyperPod Flexible Training Plans.
2. Admin/DevOps Engineers can use eksctl CLI to provision an Amazon Elastic Kubernetes Service (Amazon EKS) cluster
3. Admin/DevOps Engineers use the Amazon SageMaker AI HyperPod VPC stack to deploy HyperPod managed node group on the **Amazon EKS** cluster.
4. Admin/DevOps Engineers verify access to **Amazon EKS** cluster API via Network Load Balancer (NLB).
5. Admin/DevOps Engineers can install FSx for Lustre CSI driver and mount that file system on the **Amazon SageMaker AI** HyperPod **Amazon EKS** cluster.
6. Admin/DevOps Engineers install Amazon Elastic Fabric Adaptor (EFA) Kubernetes device plugins connected to compute nodes for Elastic Network Interface (ENI) based connectivity cross account.
7. Admin/DevOps Engineers can configure AWS Container Insights to push Amazon SageMaker HyperPod cluster workloads metrics into Amazon Cloudwatch.
8. Admin/DevOps Engineers configure IAM to use Amazon Managed Prometheus to collect metrics and Amazon Managed Grafana to set up the observability stack.

Reviewed for technical accuracy September 2, 2025

AWS Reference Architecture

Guidance for Training Protein Language Models (ESM-2) with Amazon SageMaker AI HyperPod

This reference architecture demonstrates how to run distributed ESM-2 training jobs on an Amazon EKS based HyperPod cluster.



1. Admin/DevOps Engineers move their training data from on-prem to an Amazon Simple Storage Service (Amazon S3) bucket.
2. Admin/DevOps Engineers can create Data Repository Associations between Amazon S3 and Amazon FSx for Lustre.
3. Data Scientists/ML Engineers build AWS optimized Docker container images with a designated base image.
4. Data Scientists/ML Engineers push Docker images to Amazon Elastic Container Registry (Amazon ECR).
5. Administrators/DevOps Engineers deploy Kubeflow Training Operators to **Amazon SageMaker AI** HyperPod **Amazon Elastic Kubernetes Service (Amazon EKS)** cluster to orchestrate PyTorch based distributed training jobs.
6. Data Scientists/ML Engineers deploy Kubernetes model training manifests that reference the ESM-2 dataset and use container images built in Step 3 to kickstart model training jobs on the compute nodes.
7. **Amazon SageMaker AI** HyperPod cluster Compute Nodes write model training job checkpoints to the shared FSx for Lustre file system. Data Scientists can monitor the training process via logs to determine when training job is completed.

Reviewed for technical accuracy September 2, 2025

AWS Reference Architecture