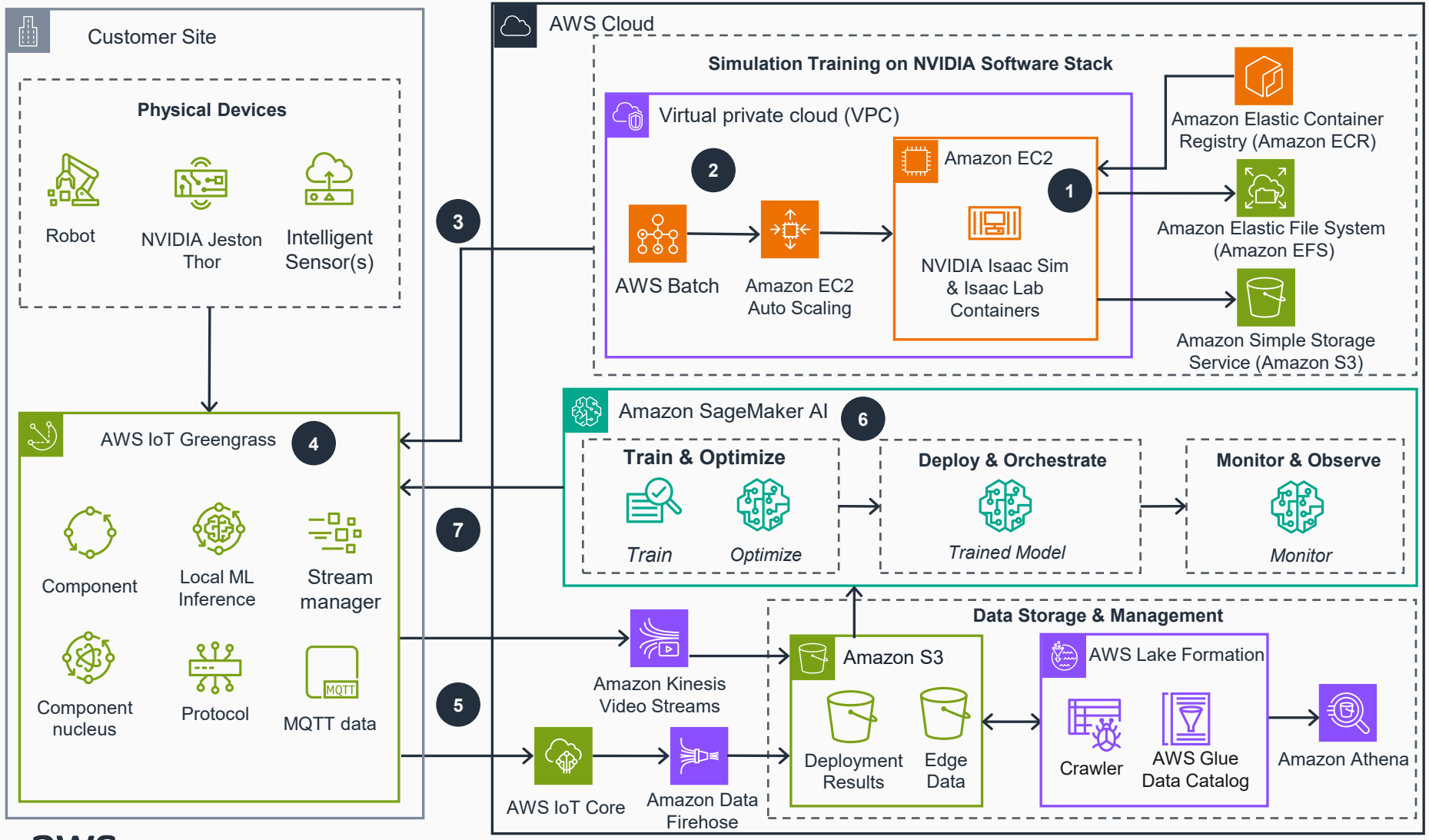
Guidance for Physical AI for Robotic Applications on AWS

This architecture enables autonomous robotics application development through reinforcement learning from physics-based closed-loop feedback systems. Models trained in simulation are deployed to robot controllers for real-world testing. Edge devices collect multi-sensor data that flows to the AWS cloud for continuous model refinement and redeployment. Customers can choose to start with either approach and iterate continuously. This slide shows steps 1-6.



1. Begin Simulation training with robotics reinforcement learning using NVIDIA Isaac Sim containers deployed on GPU-powered **Amazon Elastic Compute Cloud (Amazon EC2)** instances for modeling and NVIDIA Isaac Lab to scale training scenarios. Test physics constraints and scenarios through multiple iterations within this simulation loop.

2. **AWS Batch** orchestrates simulation workloads across GPU-powered **Amazon EC2** Auto Scaling groups to dynamically scale compute resources based on demand.

3. One-way Deployment:
The trained ML model with robot policies are deployed one-way to **AWS IoT Greengrass** running on physical controllers that interface with robots at the edge.

4. **AWS IoT Greengrass** components process real-time physics and environmental feedback data from sensors including cameras, audio, gyroscopes, force, accelerometers, contact sensors, joint encoders, position, and pressure sensors.

5. **AWS IoT Greengrass** sends MQTT sensor data through **AWS IoT Core** and **Amazon Data Firehose** to **Amazon Simple Storage Service (Amazon S3)** data lakes, while video streams flow via **Amazon Kinesis Video Streams** to **Amazon S3** for storage and management

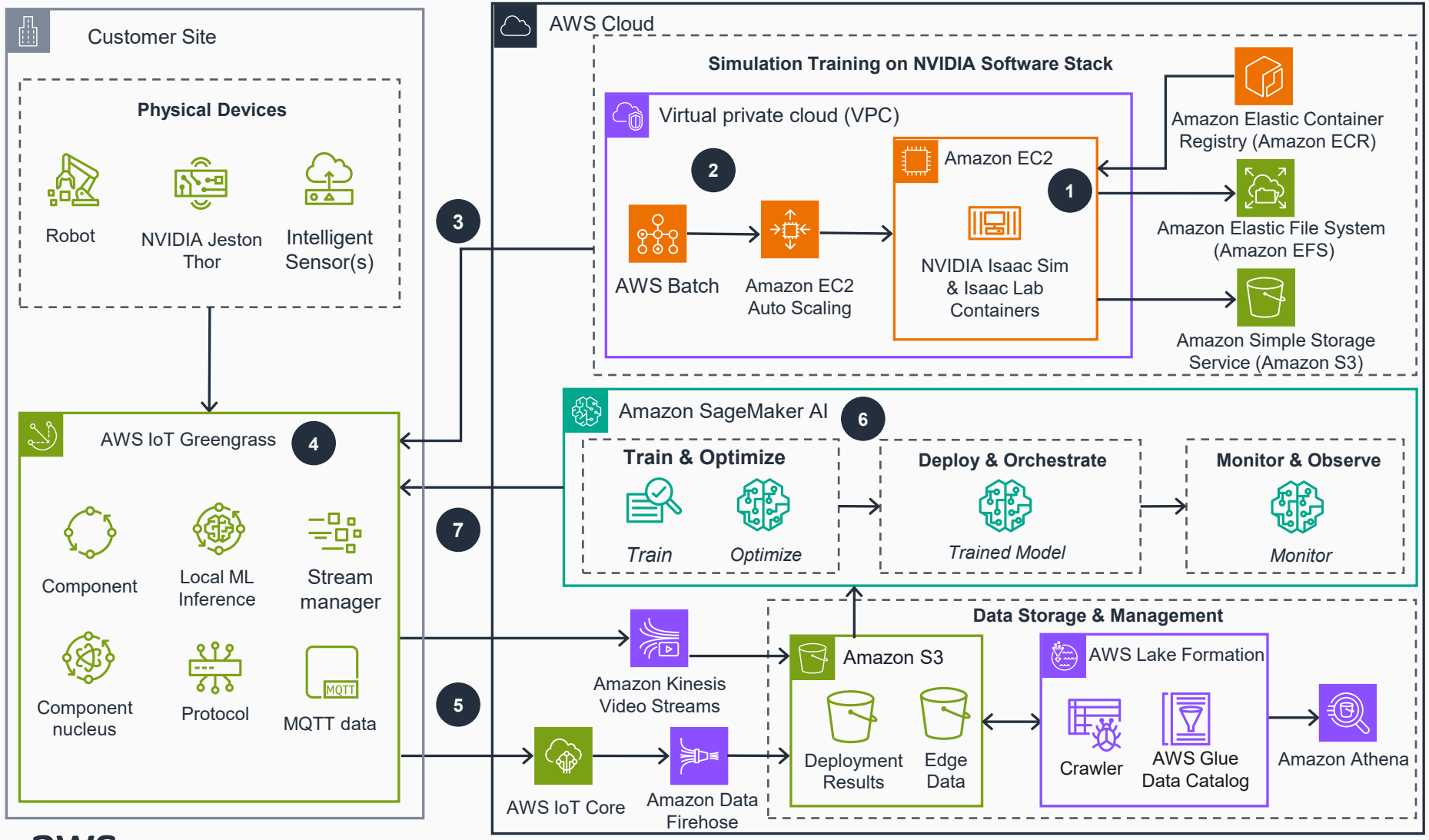
6. **Amazon SageMaker AI** processes batches of real-world data to train and/or retrain and optimize models, bridging sim-to-real gaps between NVIDIA Isaac Sim virtual simulation and actual robot operations.

Reviewed for technical accuracy Oct 17, 2025

AWS Reference Architecture

Guidance for Physical AI for Robotic Applications on AWS

This architecture enables autonomous robotics application development through reinforcement learning from physics-based closed-loop feedback systems. Models trained in simulation are deployed to robot controllers for real-world testing. Edge devices collect multi-sensor data that flows to the AWS cloud for continuous model refinement and redeployment. Customers can choose to start with either approach and iterate continuously. This slide shows step 7.



7 Continuous Deploy and Monitor: Refined ML models trained in **Amazon SageMaker AI** are deployed to **AWS IoT Greengrass** on the robot edge. Inference is performed using these models to optimize robot behavior and meet performance goals. A monitoring layer tracks metrics, detects drift, and triggers retraining. ML model iteration continues through this cycle: robots generate operational data, models are refined based on real-world performance, and improved models are redeployed.

Reviewed for technical accuracy Oct 17, 2025

AWS Reference Architecture