

I O T 3 1 7 - R

Learn the essentials of ML inference with AWS IoT Greengrass

Richard Kang

Senior WW IoT Specialist SA
Amazon Web Services

Inference @ Edge with AWS IoT Greengrass

Ways to deploy ML model to the AWS IoT Greengrass core

- AWS IoT Greengrass ML connectors
- AWS NEO-DLR Runtime
- Deploy as AWS IoT Greengrass ML resources

AWS IoT Greengrass provides runtimes and precompiled framework libraries to the hardware accelerators

AWS IoT Greengrass Machine Learning Runtimes and Precompiled Libraries

Machine learning runtimes and libraries are required for your ML models to perform inference on the edge.

Download the model type for your platform.

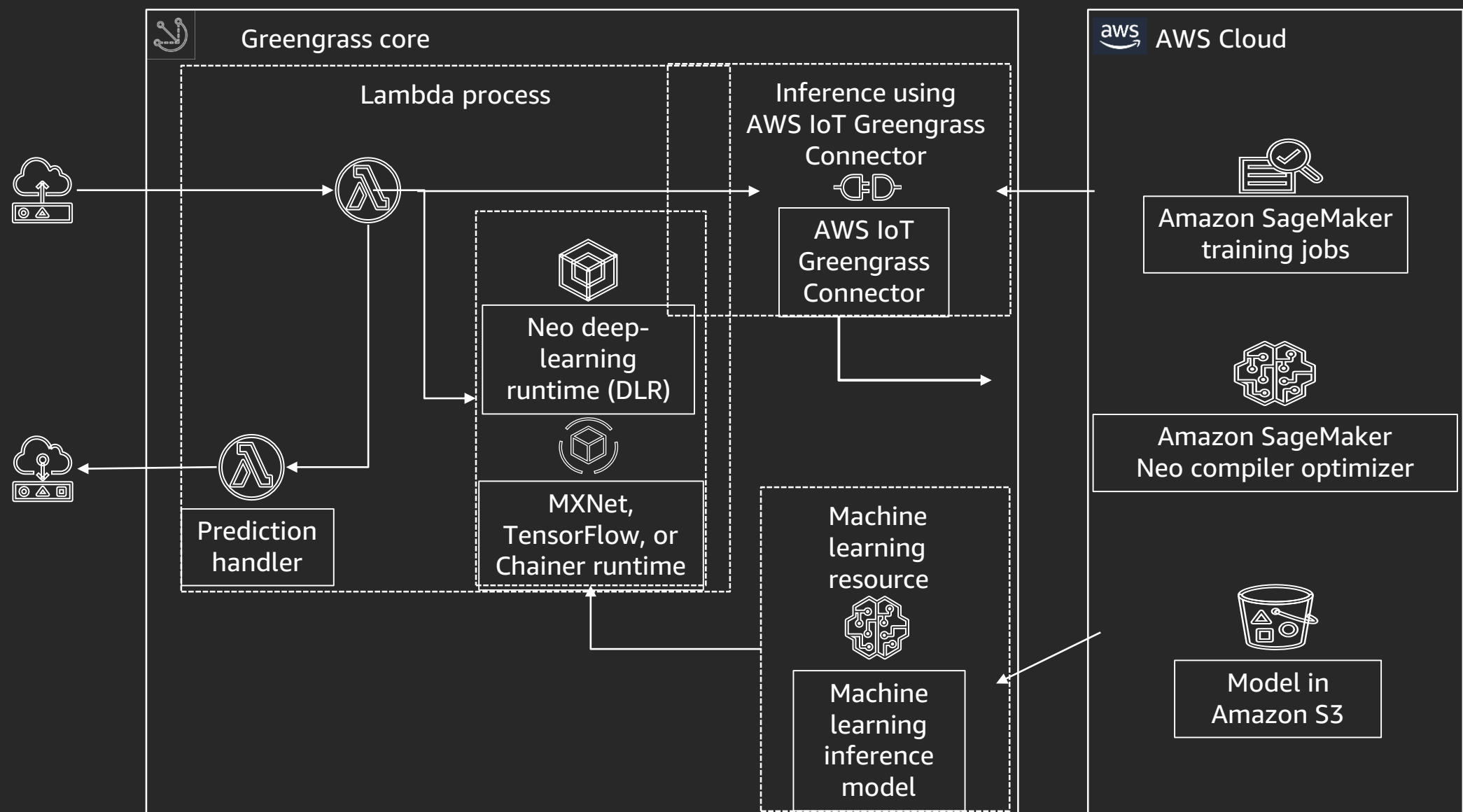
Raspberry Pi Nvidia Jetson TX2 Intel Atom

Choose the download link for your model type.

By downloading this software you agree to the associated license.

Model type	Version	License	Link
MXNet	1.2.1	Apache License 2.0	Download
TensorFlow	1.4.0	Apache License 2.0	Download
Deep Learning Runtime	1.0.0	Greengrass License	Download
Chainer	4.0.0	MIT License	Download

Overview architecture



Let's build the solution

AWS IoT Greengrass Machine Learning Inference (MLI)

We will be building the pre-trained model solution in the AWS IoT Greengrass ML Inference Solution Accelerator

Link: <https://amzn.to/33ti76C>



AWS IoT Greengrass ML Inference Solution Accelerator

Background

The AWS IoT Greengrass ML Inference solution accelerator demonstrates how to deploy and run machine learning models trained in [Amazon SageMaker](#) on a local edge device running [AWS IoT Greengrass](#), with the ability to perform local inferencing on locally-generated data while publishing the inference results to connected AWS IoT Greengrass-aware devices or AWS.

[Get solution accelerator »](#)

AWS IoT Greengrass ML Inference

Version: 1.0

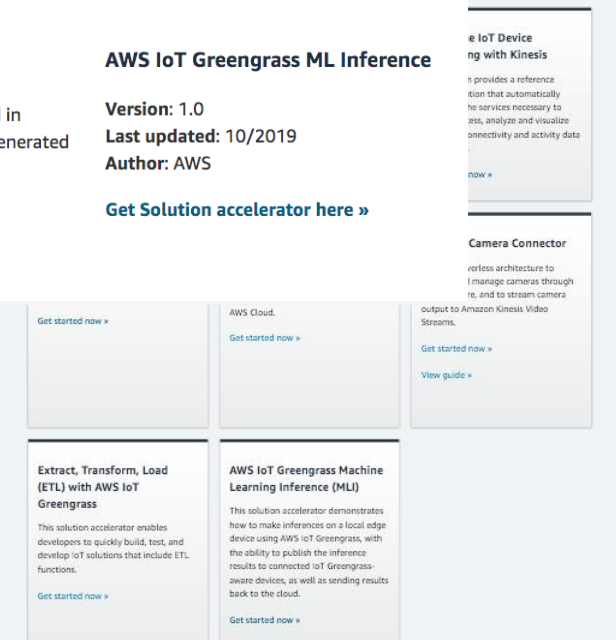
Last updated: 10/2019

Author: AWS

[Get Solution accelerator here »](#)

Greengrass Machine Learning Inference (MLI) with pre-trained model uploaded to S3

This document describes the steps in setting up Greengrass Machine Learning Inference, using a pre-trained model uploaded to S3.



Description of steps

Step 1: Upload a pre-trained ML model to S3 - <https://bit.ly/2plLs4a>

Step 2: Create the following resources using the AWS CloudFormation - <http://bit.ly/2P0JLnJ>

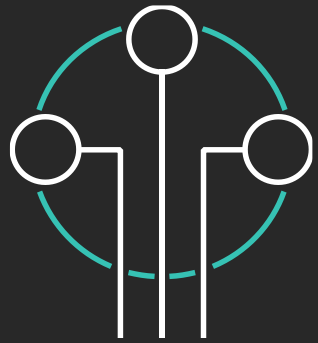
1. AWS IoT Greengrass core
2. Lambda functions for the image acquisition and prediction handling
3. AWS IoT Greengrass ML resources
4. AWS IoT Greengrass subscription to send the prediction results back to the cloud

Step 3: Creates the AWS IoT Greengrass core on an Amazon Elastic Compute Cloud (Amazon EC2) and starts the AWS IoT Greengrass core - <http://bit.ly/33rqcc2>

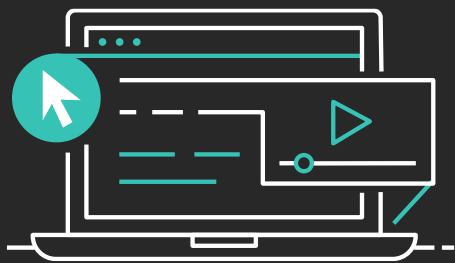
Step 4: Monitor the output from AWS IoT Core test client - <http://bit.ly/2MN87hF>

Learn IoT with AWS Training and Certification

Resources created by the experts at AWS to help you build IoT skills



Take the free digital curriculum, Internet of Things (IoT) Foundation Series, to build IoT skills and work through common scenarios



25+ additional free digital courses cover topics related to IoT, including:

- AWS IoT Core
- AWS IoT Greengrass
- AWS IoT Analytics
- AWS IoT Device Management
- AWS IoT Events

Visit the Learning Library at <https://aws.training>

Thank you!



Please complete the session
survey in the mobile app.