Implementing Deep Learning Models Using PyTorch

Janani Ravi
Co-founder, Loonycorn

www.loonycorn.com

Overview

An overview of PyTorch

Implement simple classification models using PyTorch on Databricks

Implement image classification models using PyTorch on Databricks

An overview of Horovod for distributed training on a cluster

Use Horovod to implement distributed processing for image classification

An Overview of PyTorch

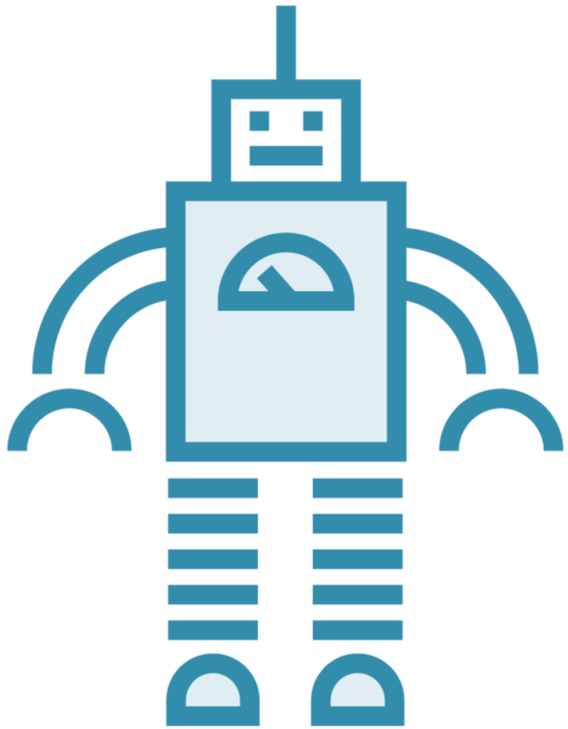
PyTorch

Open source machine learning framework that accelerates the path from research prototyping to production deployment.

PyTorch

Optimized tensor library for deep learning using GPUs and CPUs.

PyTorch



Production ready

Distributed training

Robust ecosystem

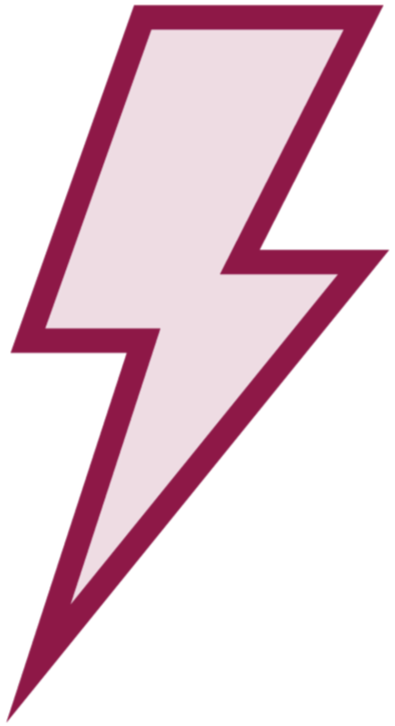
Cloud support

Autologging in MLflow is compatible with certain versions of PyTorch Lightning NOT with native PyTorch

PyTorch Lightning

PyTorch research framework which allows you to scale your models without the boilerplate code that models usually need.

PyTorch Lightning



Define models using LightningModule

Define training, test, validation datasets

Use Trainers to train modules - abstract away complexity needed for scale

Perform advanced training using Trainer arguments

Demo

Build and train a PyTorch classification model

Deploy and serve the model using classic MLflow serving

Demo

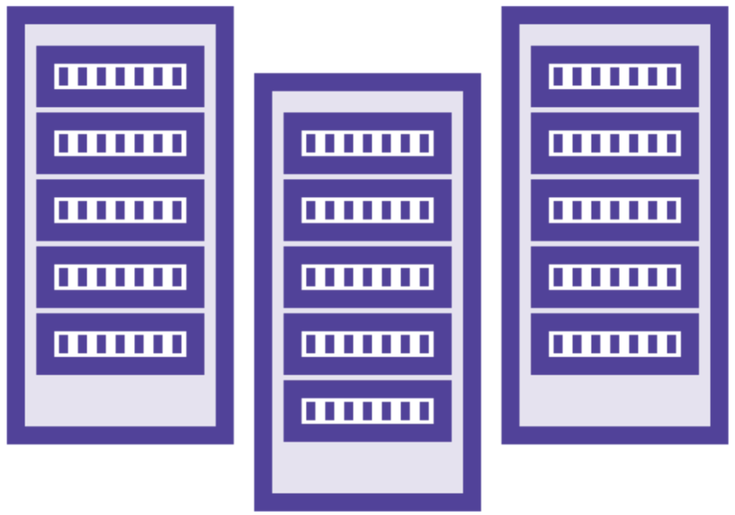
Build and train a PyTorch image classification model using MLflow on Databricks

Horovod for Distributed Processing

Horovod

Distributed deep learning training framework for TensorFlow, Keras, PyTorch, and Apache MLNext.

Horovod for Distributed Training



Originally developed by Uber

Makes distributed deep learning fast and easy to use

Used with existing training scripts to scale them up to run on 100s of GPUs

Can run on on-premises clusters or on cloud platforms

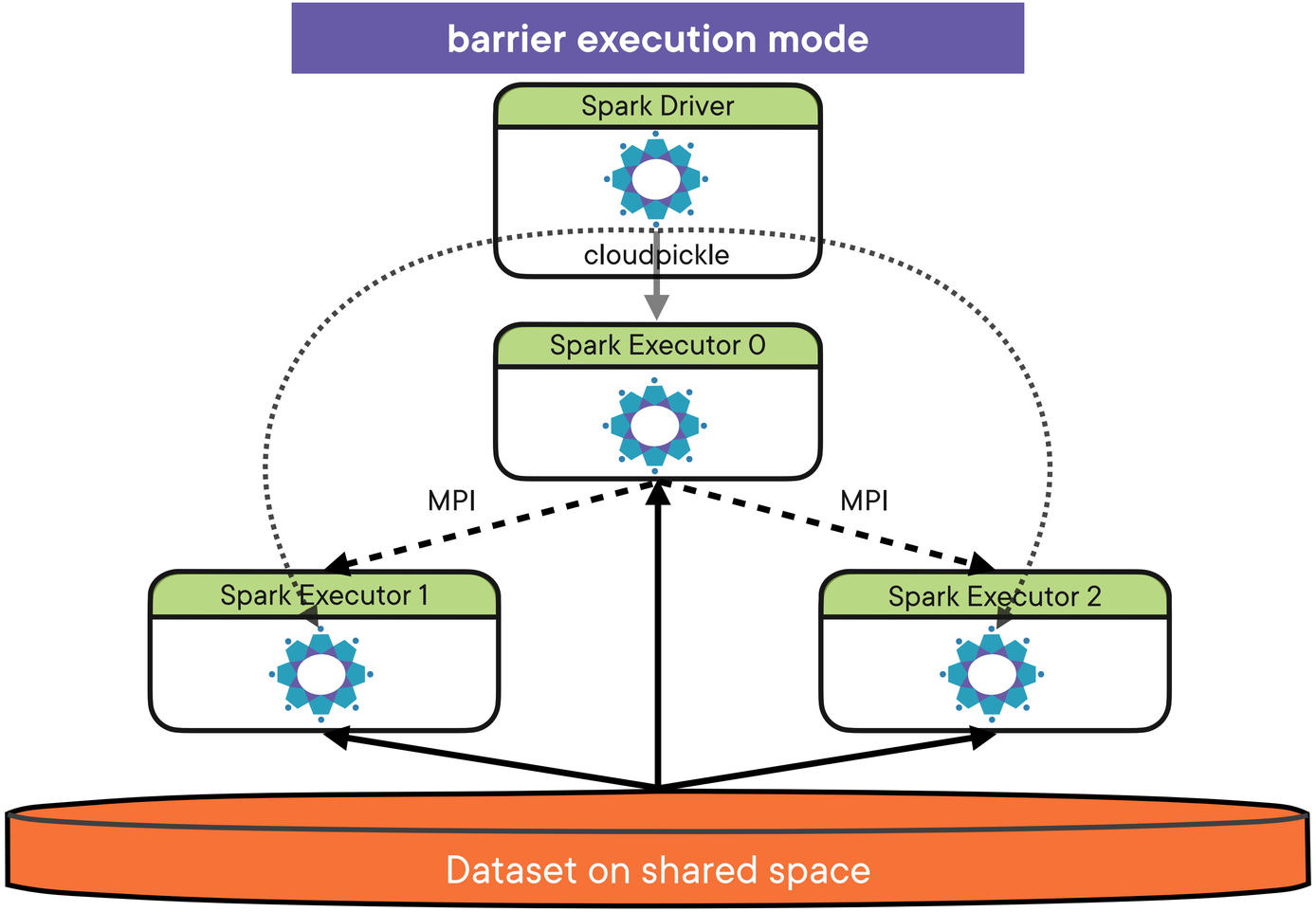
Horovod on Databricks

Horovod is integrated with Apache Spark's *barrier* mode to provide higher stability for long-running deep learning training jobs on Spark.

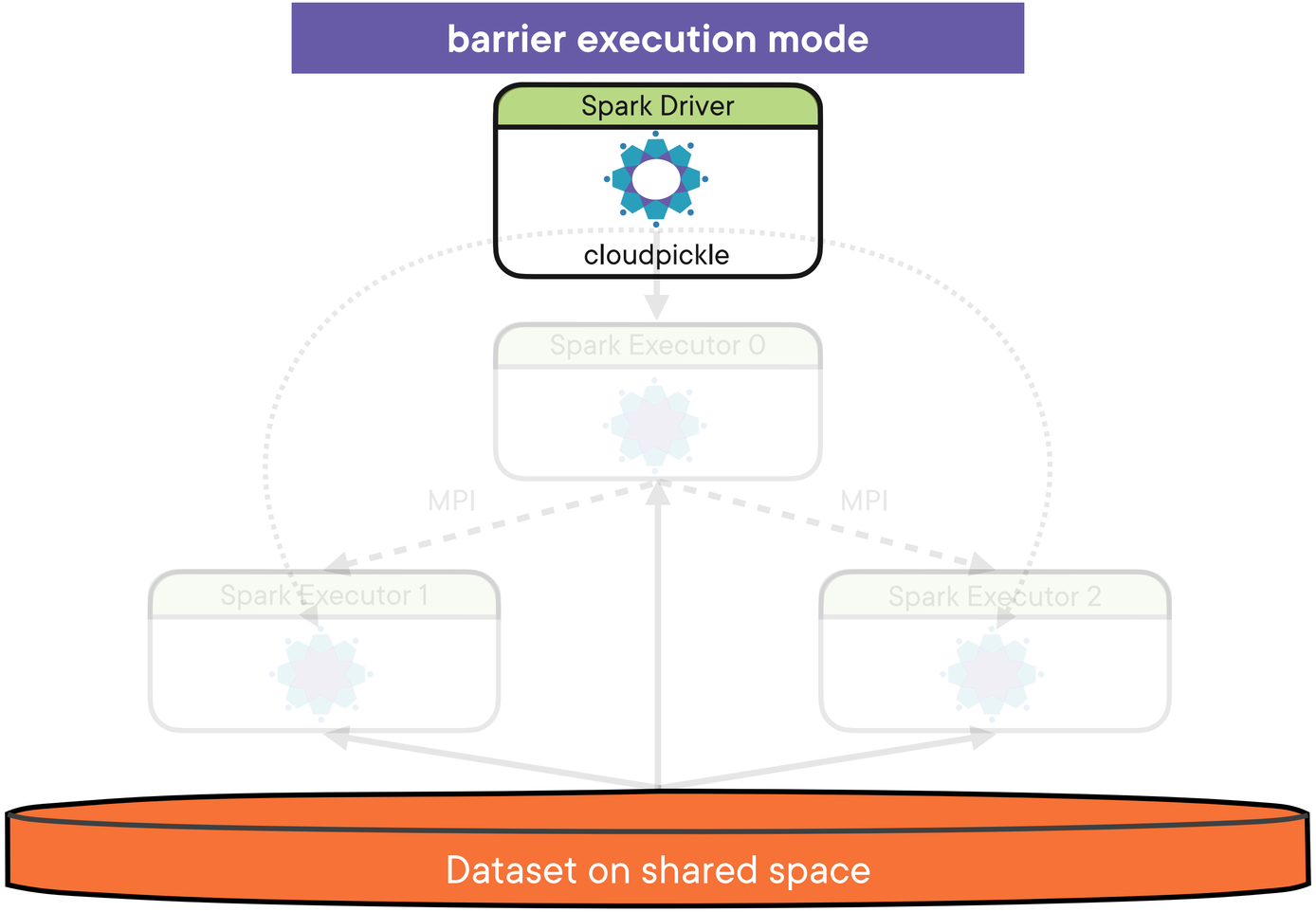
Barrier Execution Mode in Apache Spark

A scheduling model which allows users to embed distributed deep learning training as a Spark stage to simplify the distributed training workflow.

Horovod Distributed Training on Apache Spark

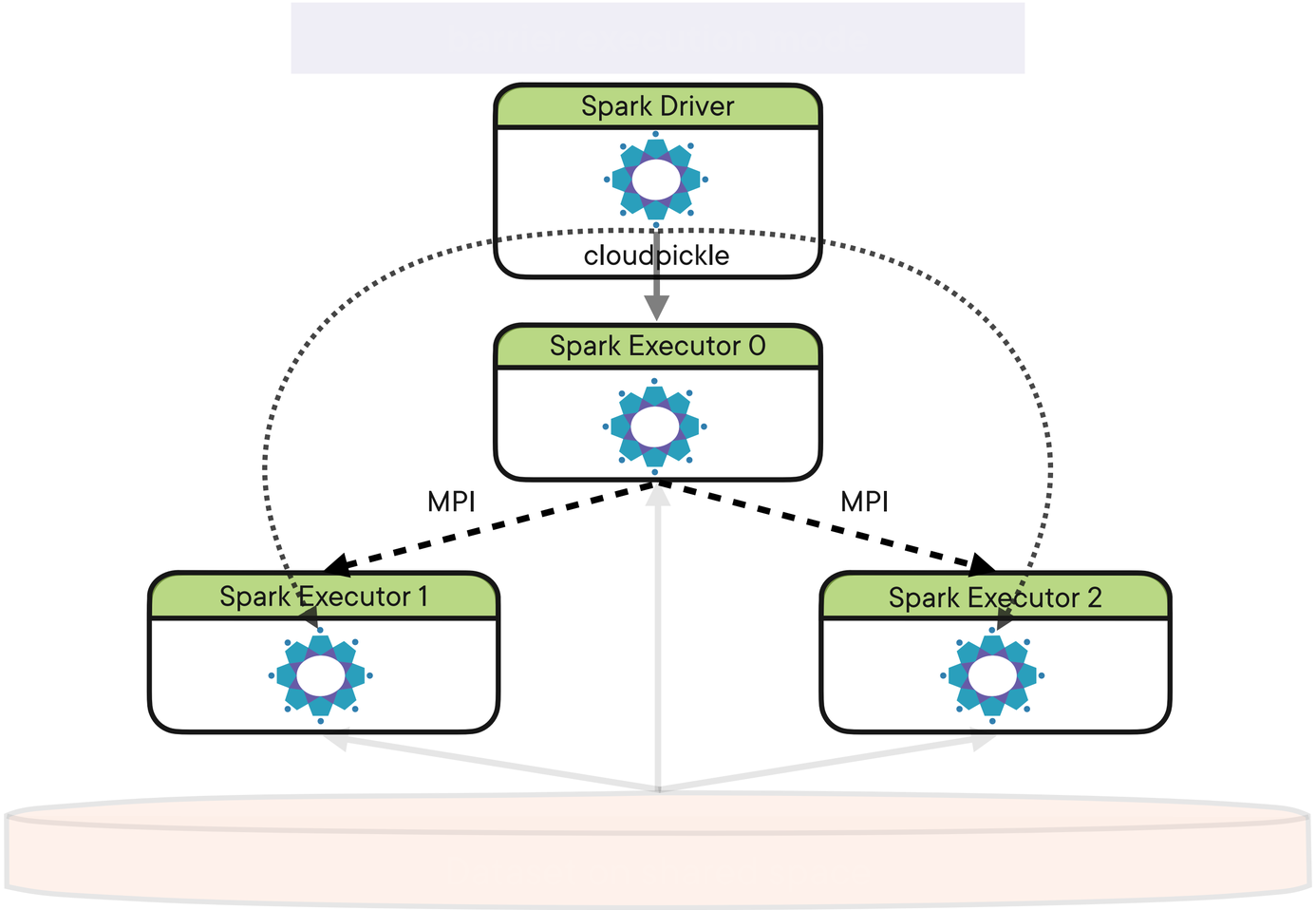


Horovod Distributed Training on Apache Spark



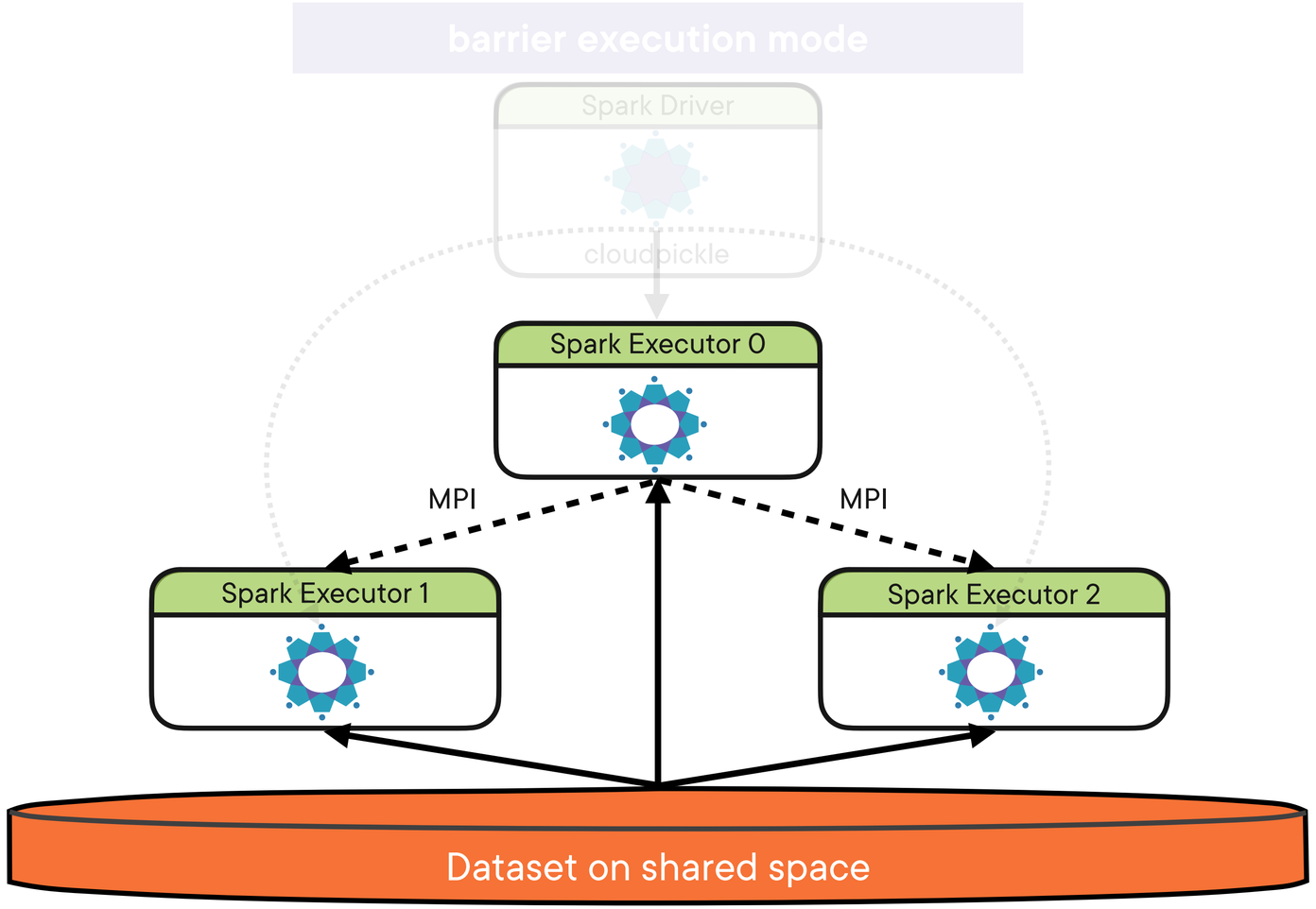
Horovod takes a Python method that contains deep learning code with Horovod hooks

Horovod Distributed Training on Apache Spark



Horovod pickles the method on the driver and distributes it to Spark workers

Horovod Distributed Training on Apache Spark



Each Python process loads the pickled user program, deserializes it, and runs it

Demo

Build and train a PyTorch image classification model using distributed training with the HorovodRunner

Summary

An overview of PyTorch

Implement simple classification models using PyTorch on Databricks

Implement image classification models using PyTorch on Databricks

An overview of Horovod for distributed training on a cluster

Use Horovod to implement distributed processing for image classification

Up Next:
Hyperparameter Tuning Using Hyperopt