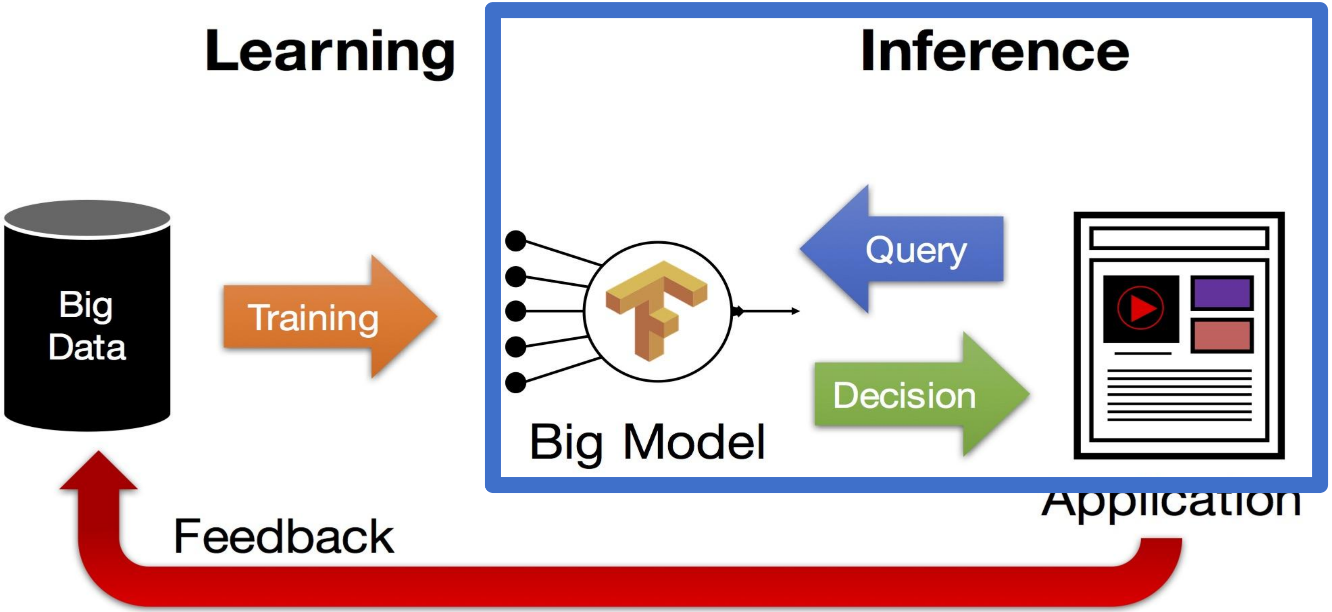
Benchmarking Distributed Training on GPUs

Exploring Performance Metrics and Inferring PyTorch DataParallel module optimizations

Team members: Tanushree Banerjee, Vedant Shah

COS 316 Final Project Video, Fall 2023

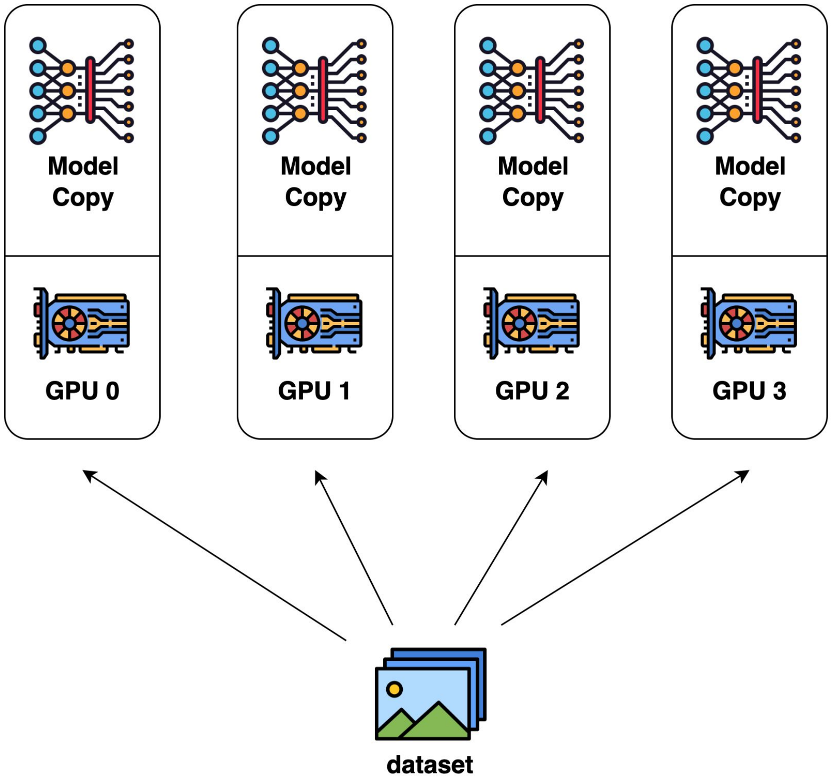
Machine Learning Life Cycle: An Overview



GPT3 175B
34 days
1024 expensive GPUs

<u>Timescale</u>: ~10 milliseconds
<u>Systems</u>: online and latency-optimized
<u>Less studied</u> …

PyTorch DataParallel



Project Goals

Benchmark distributed training on GPUs (NVIDIA A6000s)

Vary batch size, feature extractor model, number of trainable parameters, number of GPUs used

Measure computation, communication times, loss and data transfer size.

Infer optimizations made by PyTorch DataParallel module

Approach

Models: ResNet18, VGG16, SqueezeNet10, DenseNet121, AlexNet

Dataset and task: MNIST Classification

Vanilla DataParallel module with PyTorch primitives implemented explicitly:

1. Replicate
2. Scatter
3. Gather
4. Parallel apply

Metrics

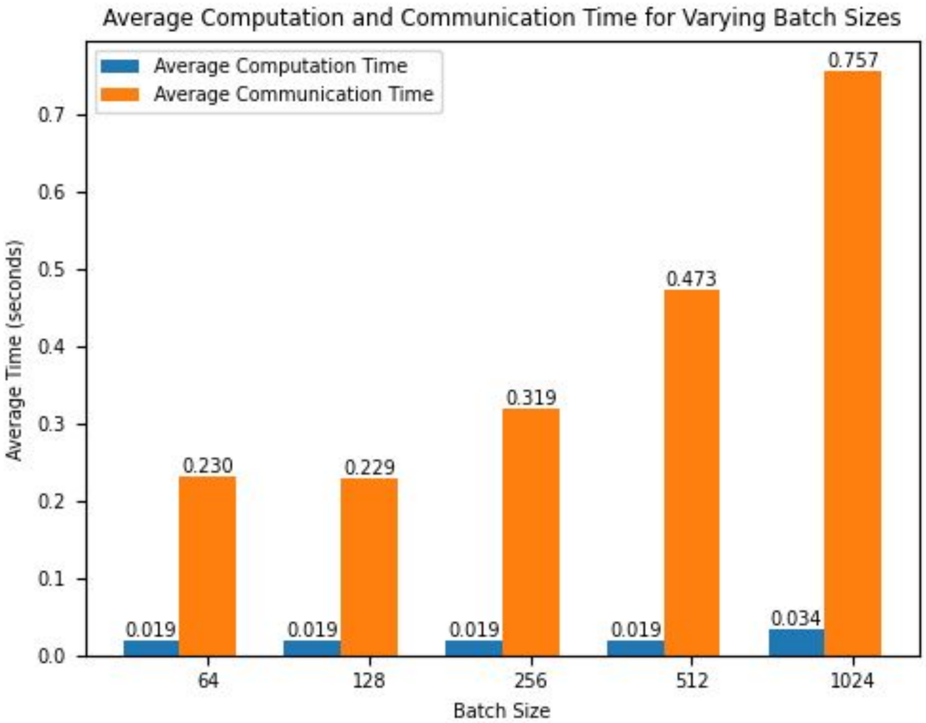
Performance evaluation metrics: computation time, communication time, data transfer size, number of epochs after which loss drops below threshold (0.04)

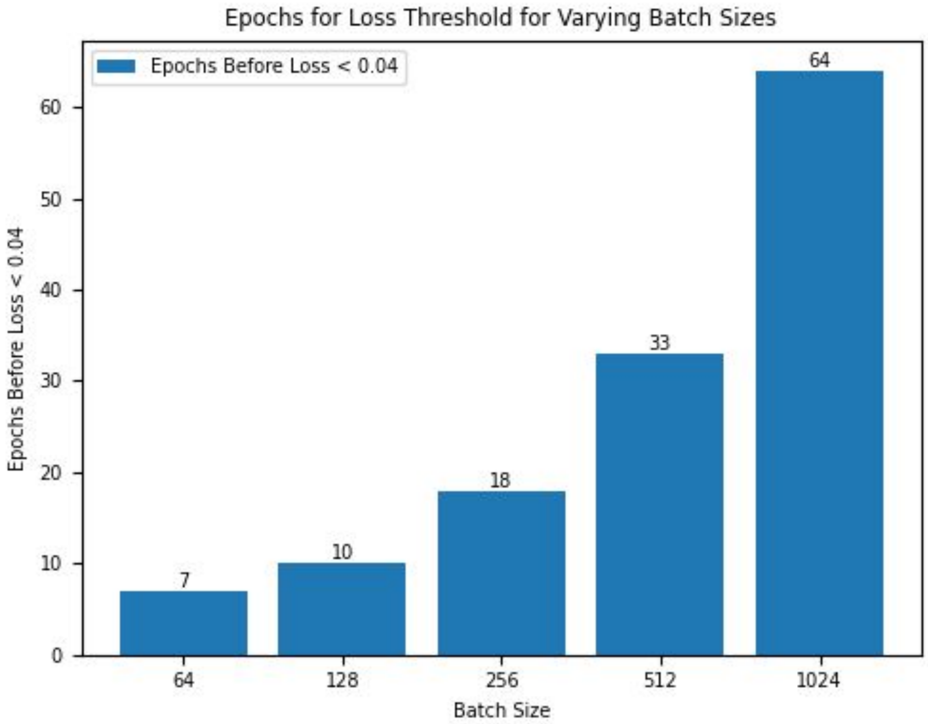
Test various configurations: feature extractor architecture, batch sizes, parallelism degrees (# GPUs), number of trainable parameters

Evaluation of VanillaDataParallel module against PyTorch's DataParallel - benchmark performance for different batch sizes

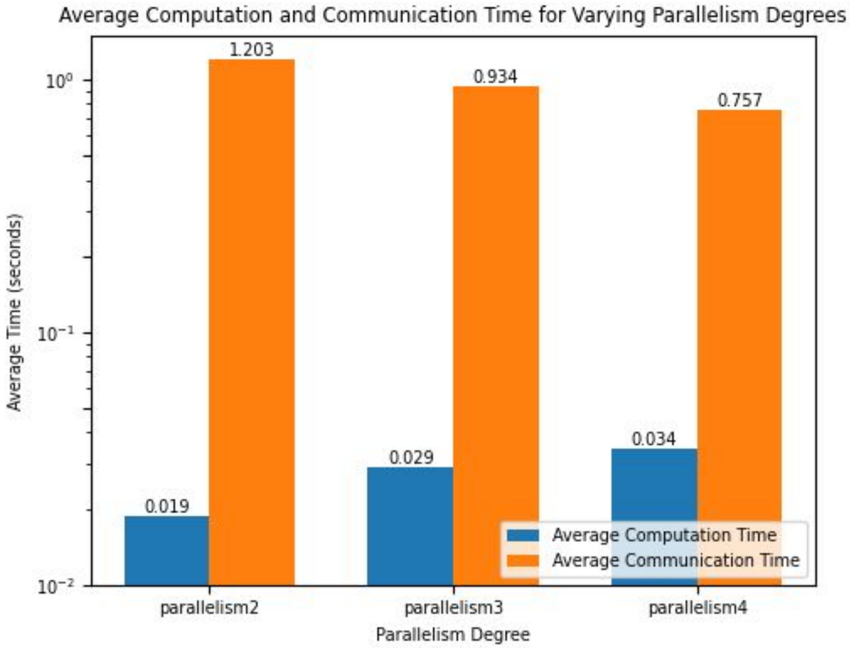
Results

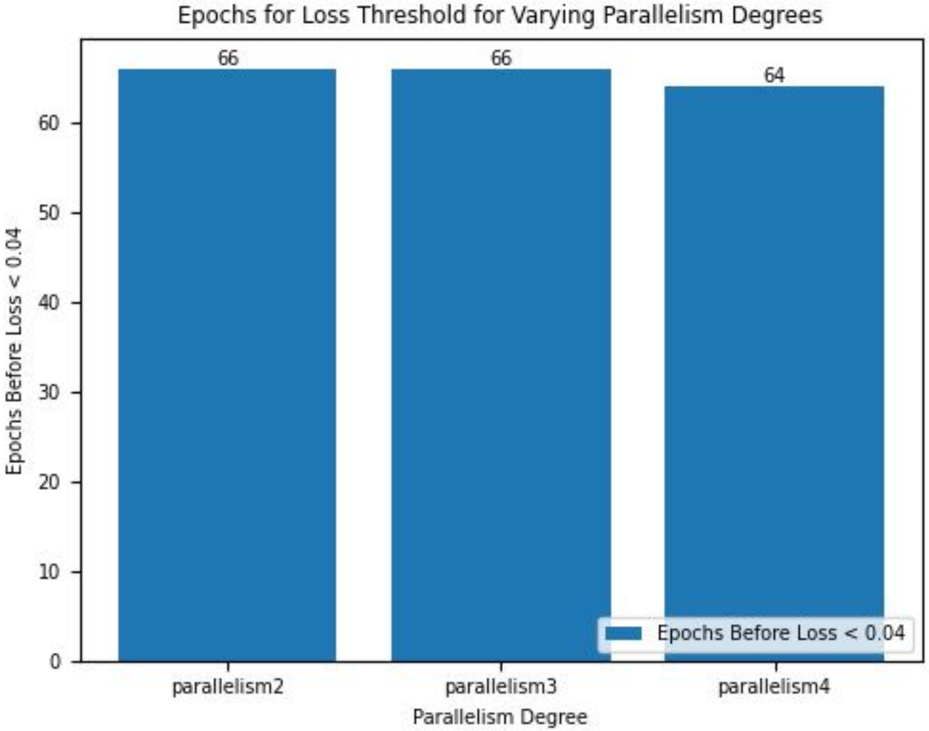
Batch_Size


Average Computation and Communication Time for Varying Batch Sizes
Average Computation Time
Average Communication Time
Average Time (seconds)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.019
0.230
0.019
0.229
0.019
0.319
0.019
0.473
0.034
0.757
64
128
256
512
1024
Batch Size

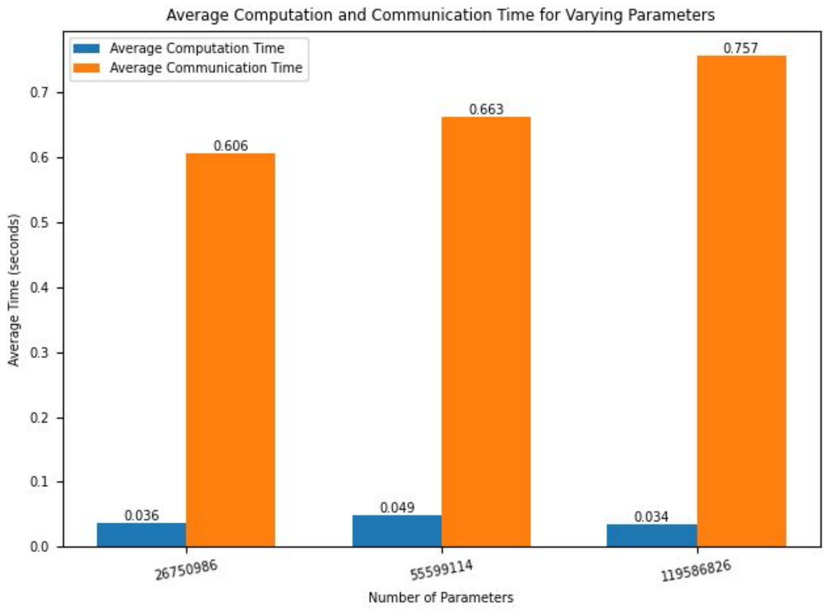

Epochs for Loss Threshold for Varying Batch Sizes
Epochs Before Loss < 0.04
Epochs Before Loss < 0.04
0
10
20
30
40
50
60
7
10
18
33
64
64
128
256
512
1024
Batch Size

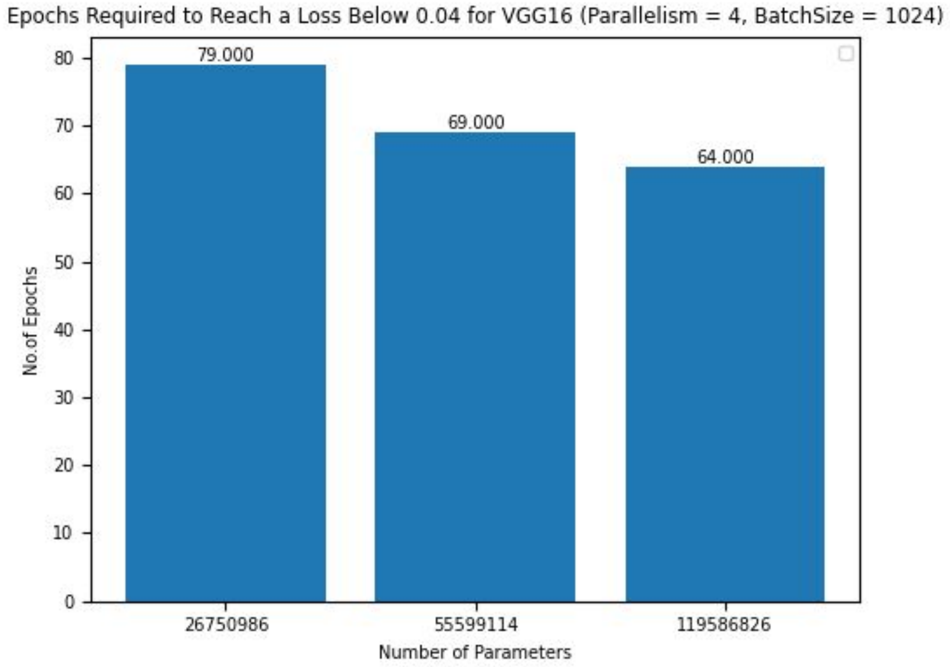
Parallelism Degree (# GPUs)



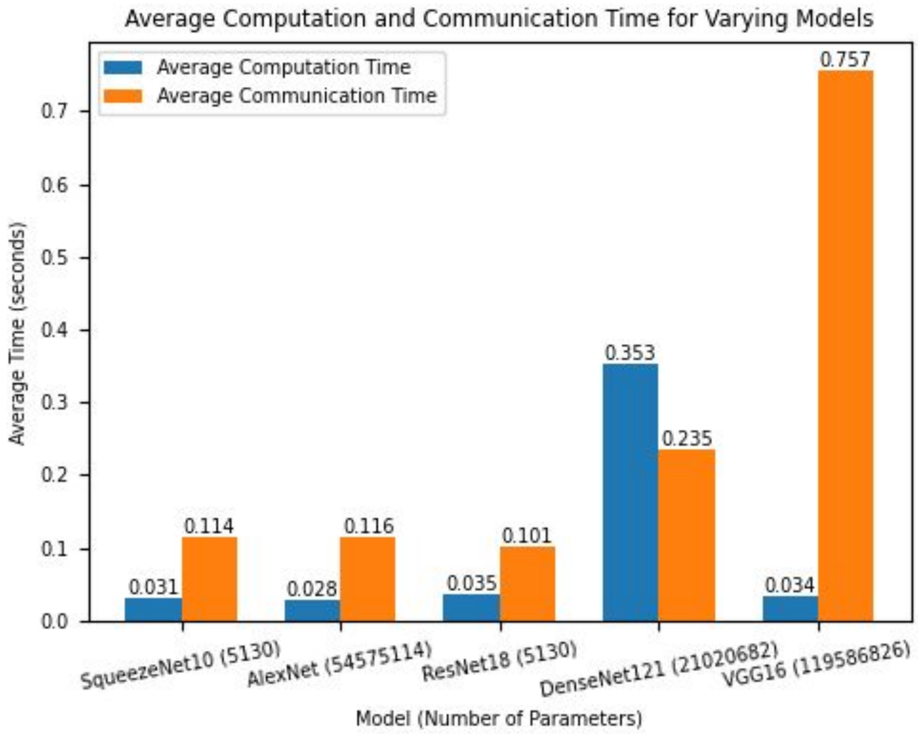


Number of trainable params

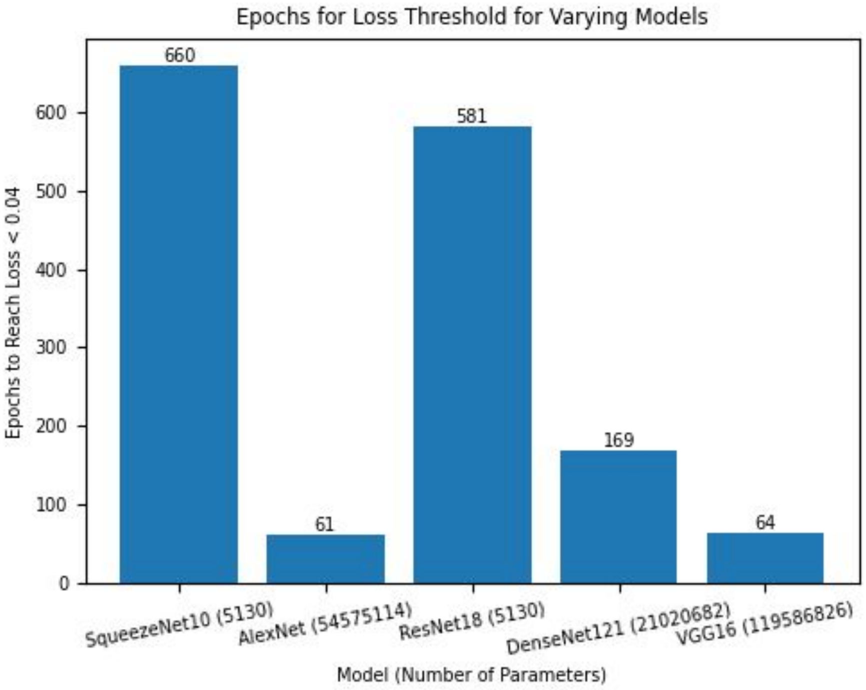




Models

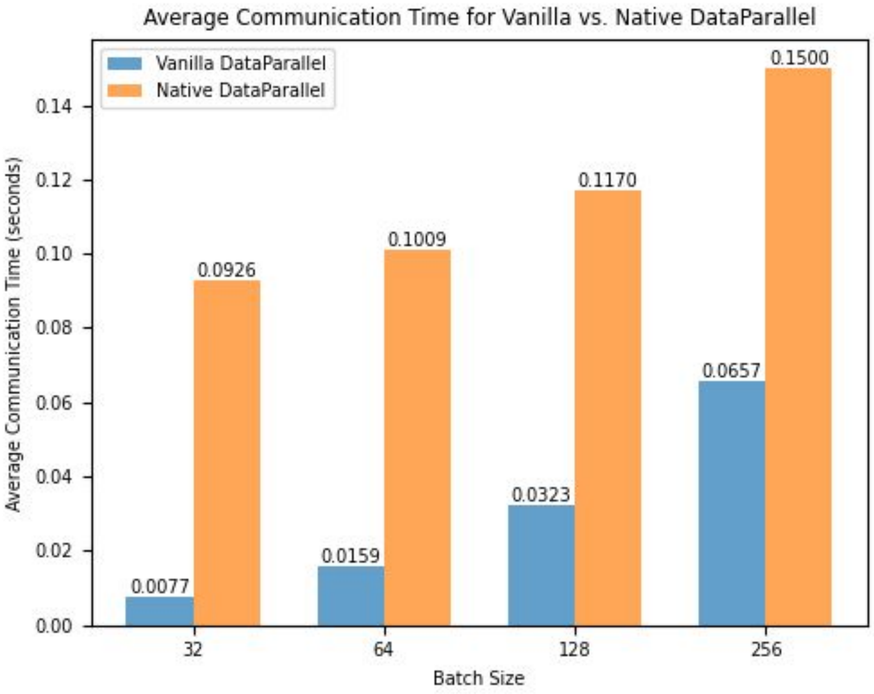


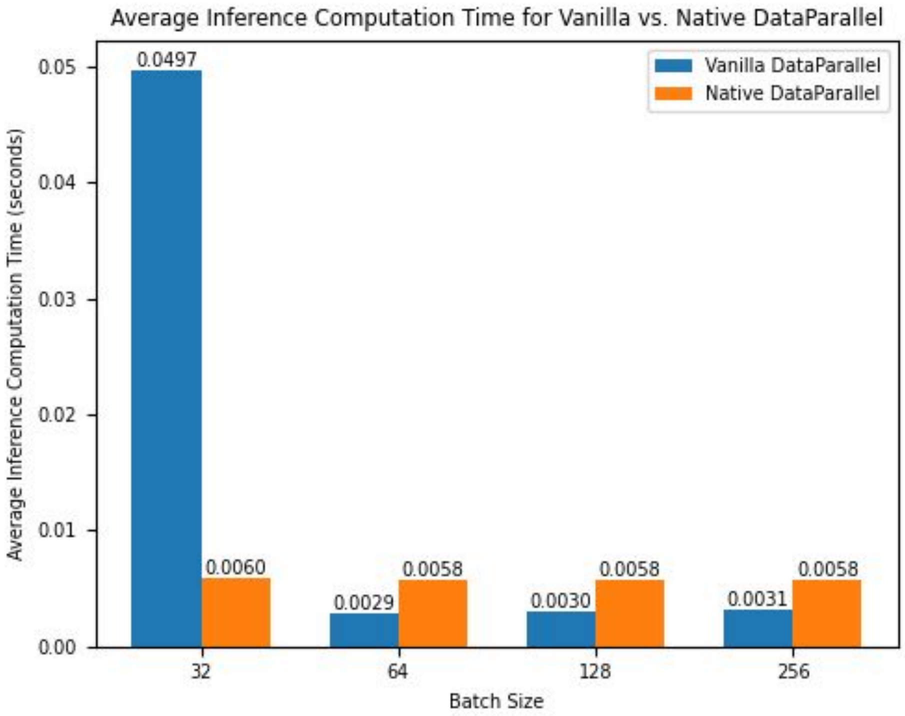
Average Computation and Communication Time for Varying Models
Average Computation Time
Average Communication Time
Average Time (seconds)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.031
0.114
0.028
0.116
0.035
0.101
0.353
0.235
0.034
0.757
SqueezeNet10 (5130)
AlexNet (54575114)
ResNet18 (5130)
DenseNet121 (21020682)
VGG16 (119586826)
Model (Number of Parameters)



Epochs for Loss Threshold for Varying Models
Epochs to Reach Loss < 0.04
0
100
200
300
400
500
600
660
61
581
169
64
SqueezeNet10 (5130)
AlexNet (54575114)
ResNet18 (5130)
DenseNet121 (21020682)
VGG16 (119586826)
Model (Number of Parameters)

Vanilla vs PyTorch Data Parallel





Thank you!