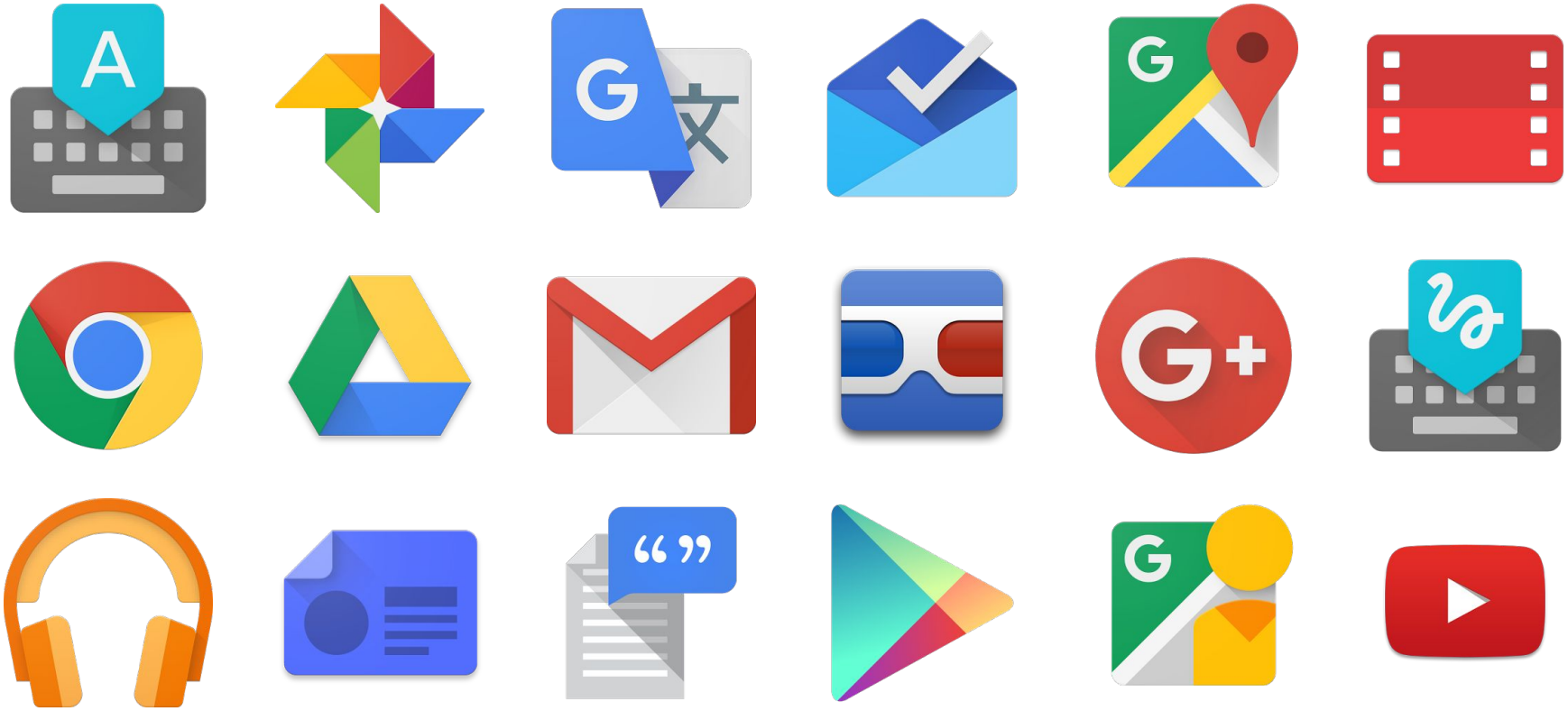
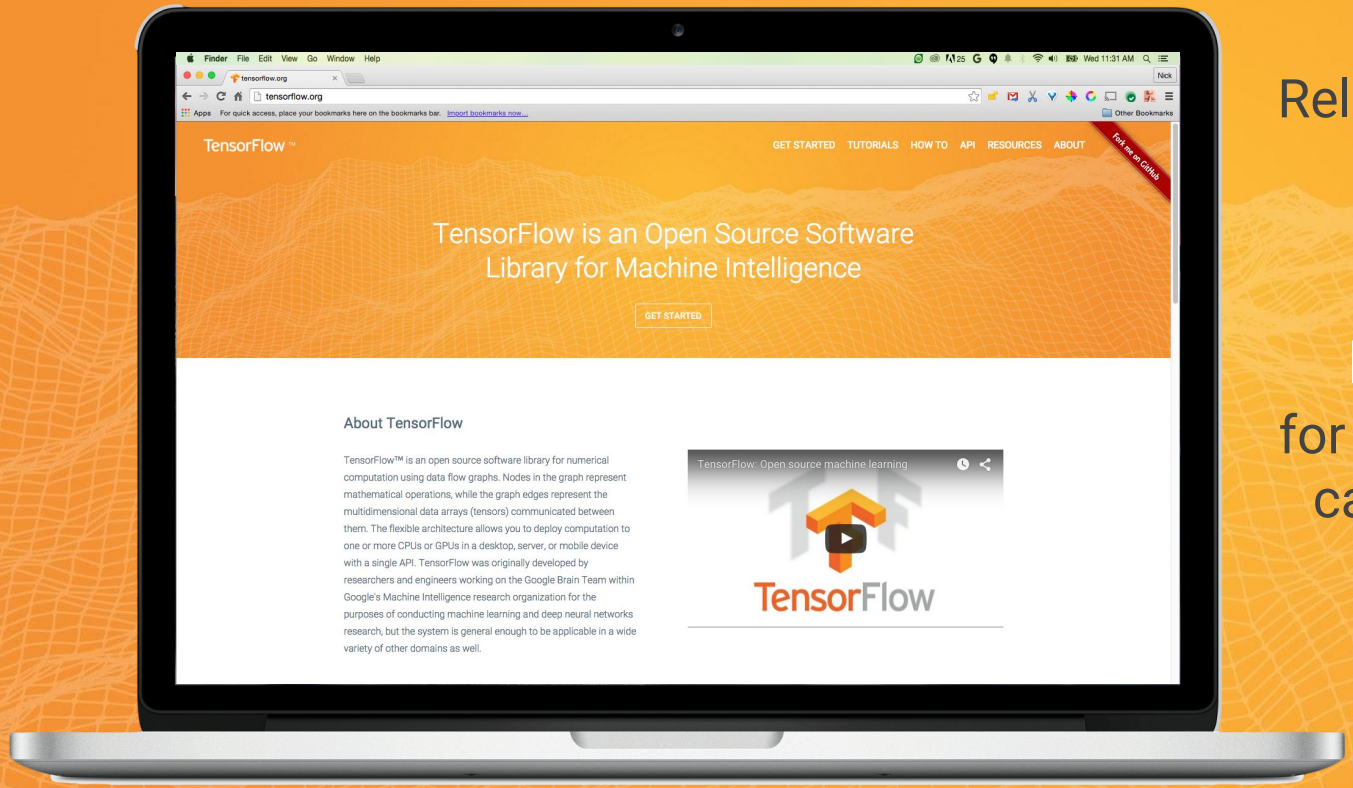
TensorFlow
Flexible, Scalable, Portable

Rajat Monga
Engineering Director, TensorFlow





Released in Nov. 2015

#1

repository

for “machine learning” category on GitHub

Some Stats

10,000+ commits since Nov, 2015

450+ contributors

1M+ binary downloads

#20 most popular repository on GitHub by stars

Used in ML classes at many universities:
Toronto, Berkeley, Stanford, ...

ARM

quantiphi

AIRBUS
DEFENCE & SPACE

CIST

CEVA

Google

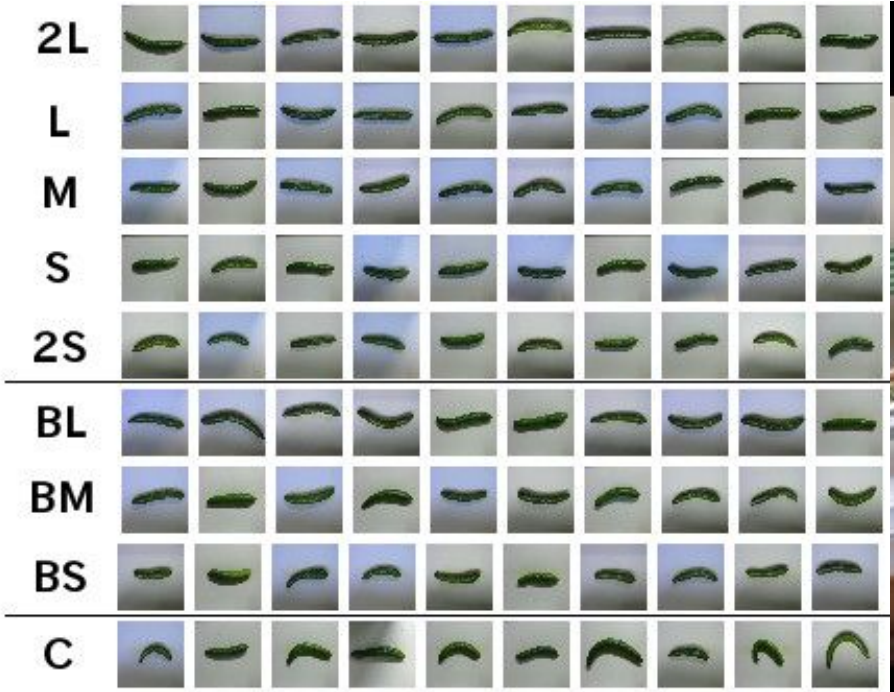
Movidius

UBER

JD.COM 京东

DeepMind

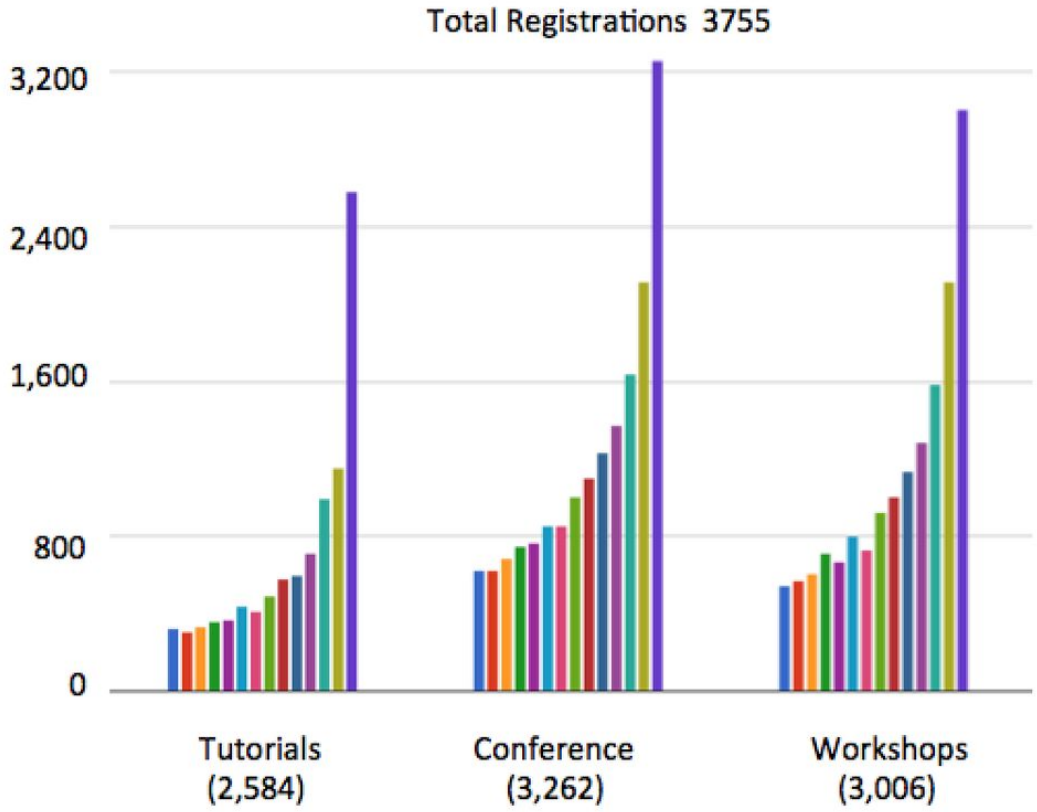
TensorFlow powered Cucumber Sorter



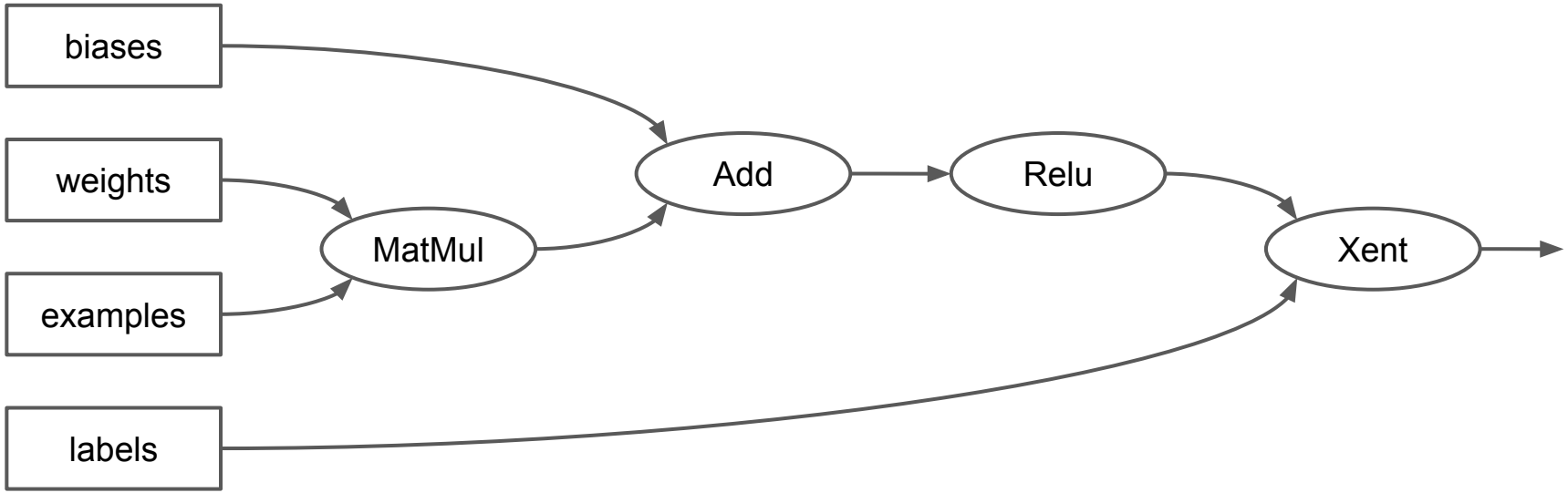
From: http://workpiles.com/2016/02/tensorflow-cnn-cucumber/

Flexible

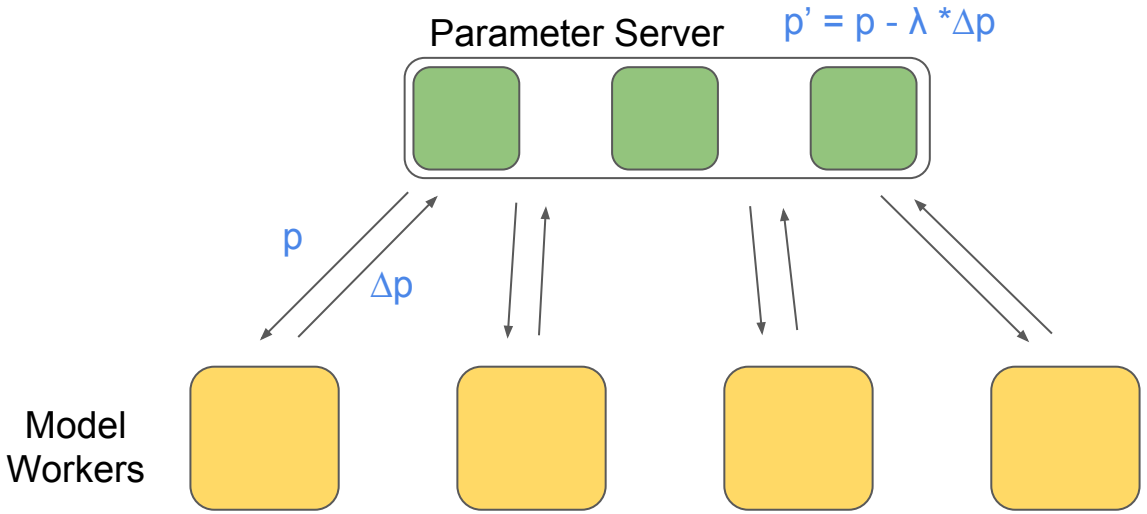
NIPS Growth


Total Registrations 3755
3,200
2,400
1,600
800
0
Tutorials
(2,584)
Conference
(3,262)
Workshops
(3,006)

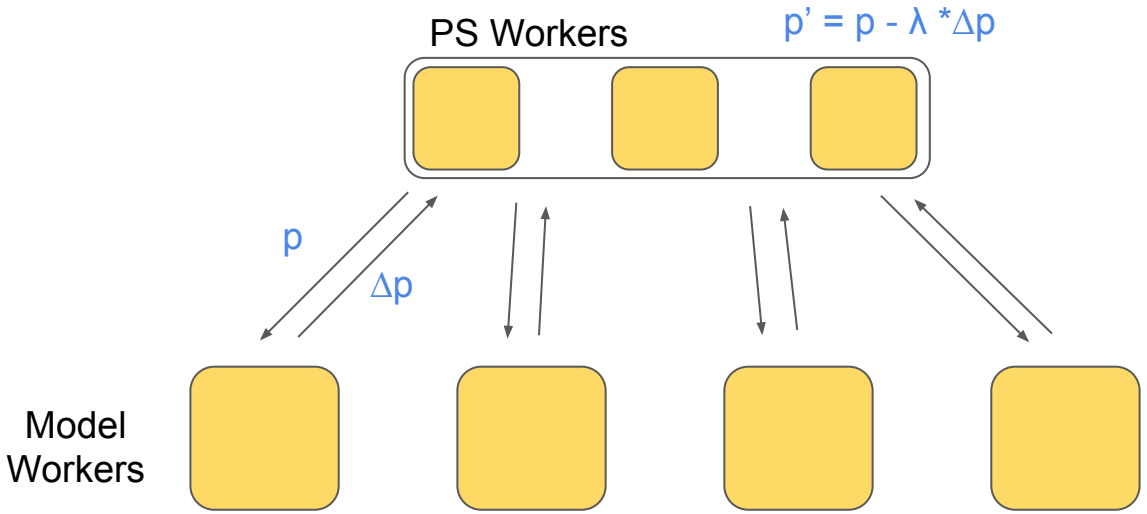
Directed graph



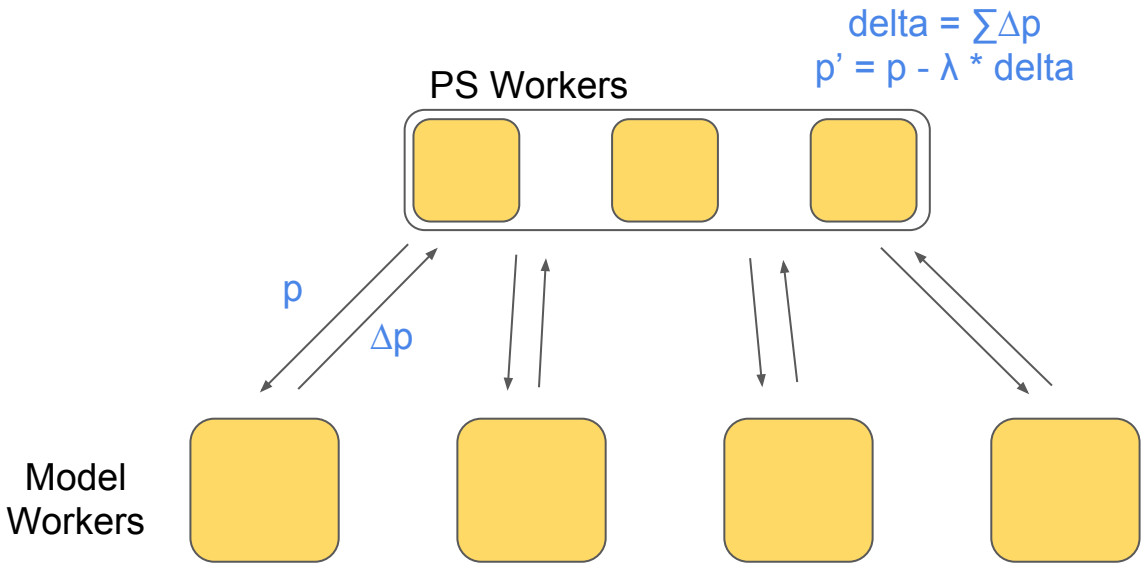
Async SGD with Parameter Server



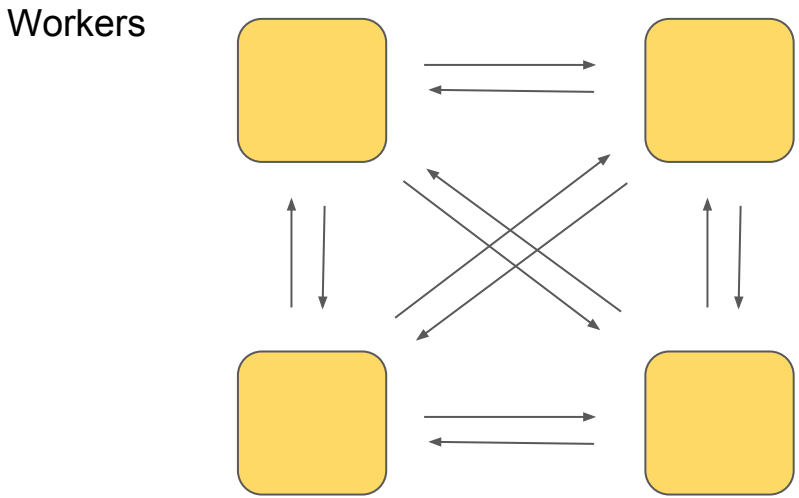
Async SGD with TensorFlow



Sync SGD with TensorFlow



Alternate version of Parameter sharing

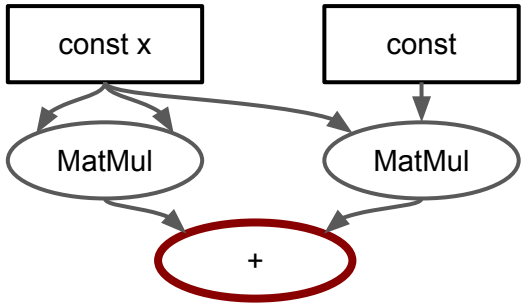


More ML Algorithms

- Clustering e.g. k-means
- Random Forests
- Logistic Regression
- Bayesian methods

Scalable

Deferred Execution

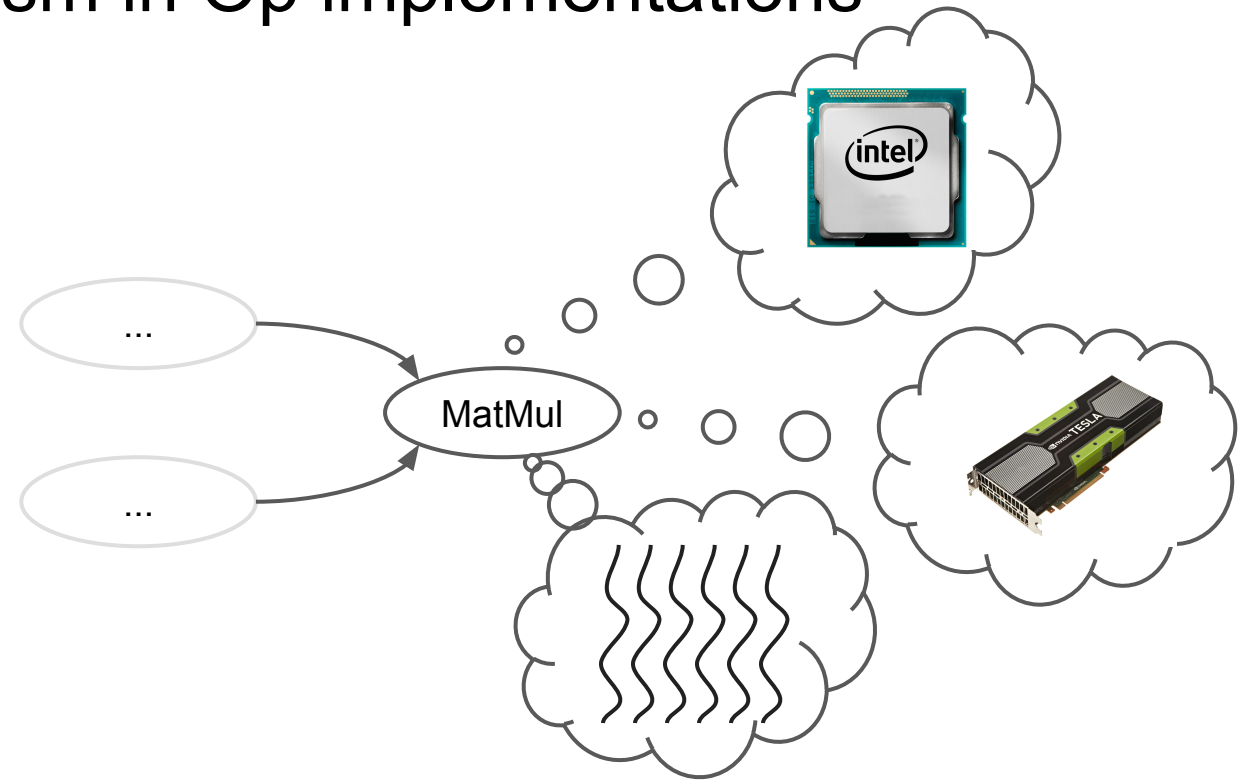


Operations build the dataflow graph; `eval()` fetches the result.

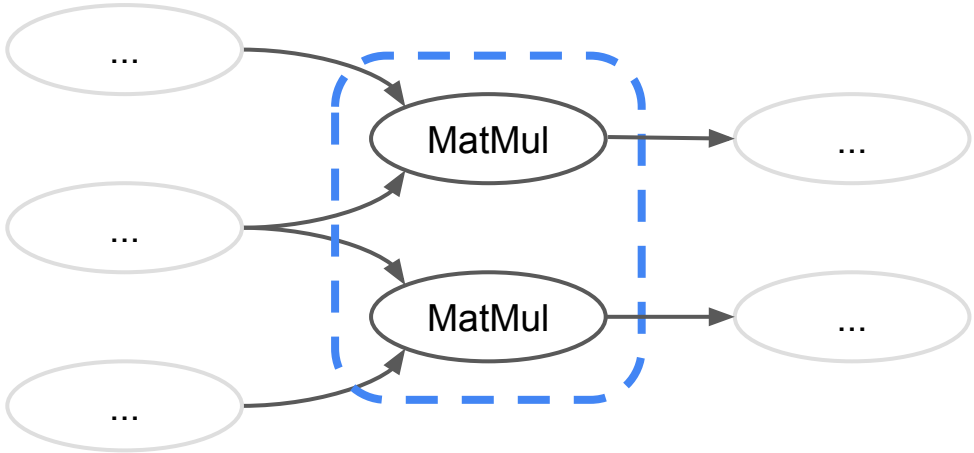
```
import tensorflow as tf

with tf.Session():
  x = tf.constant([[5, 6], [7, 8]])
  z = tf.matmul(x, x) + tf.matmul(x, [[1, 0], [0, 1]])
  # Run graph to fetch z.
  result = z.eval()
```

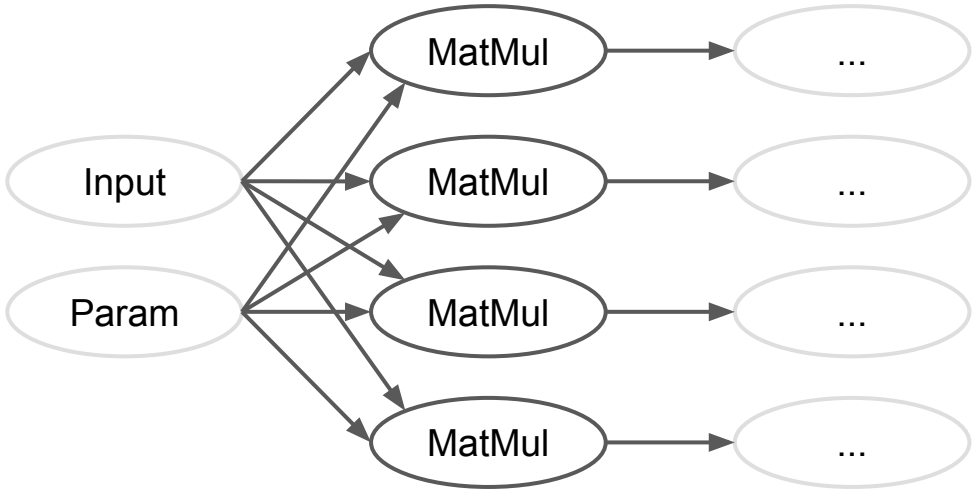
Parallelism in Op implementations



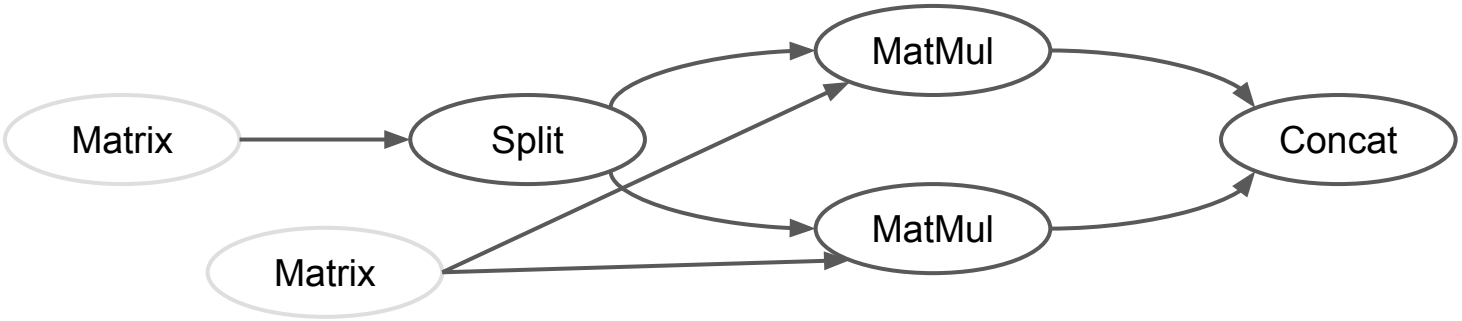
Task Parallelism in DataFlow graph



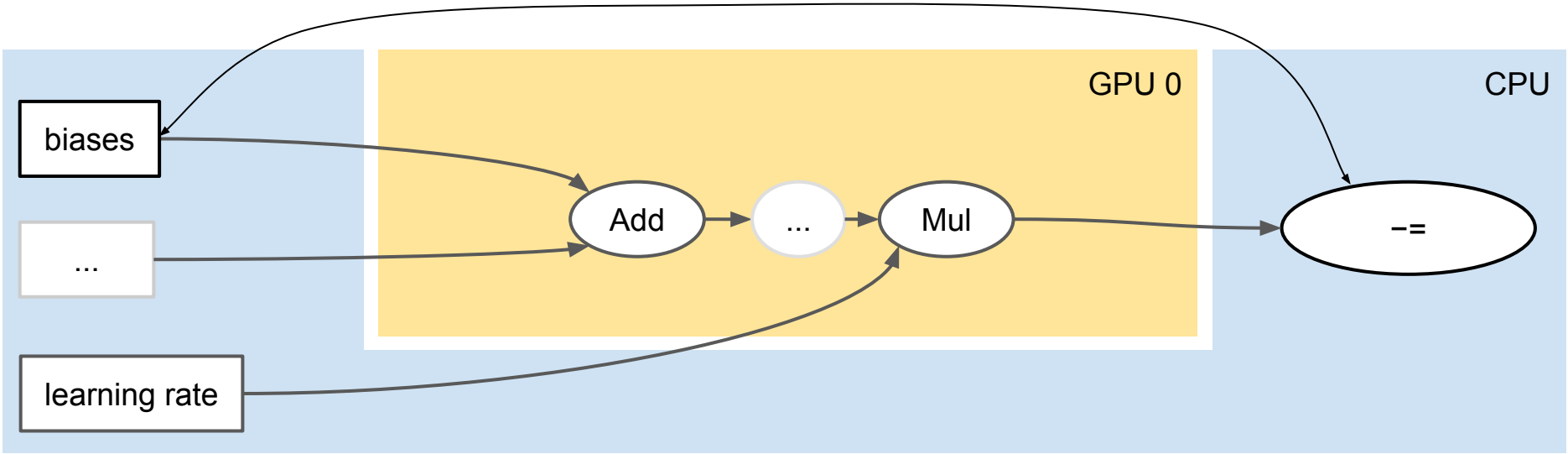
Data Parallelism



Model Parallelism

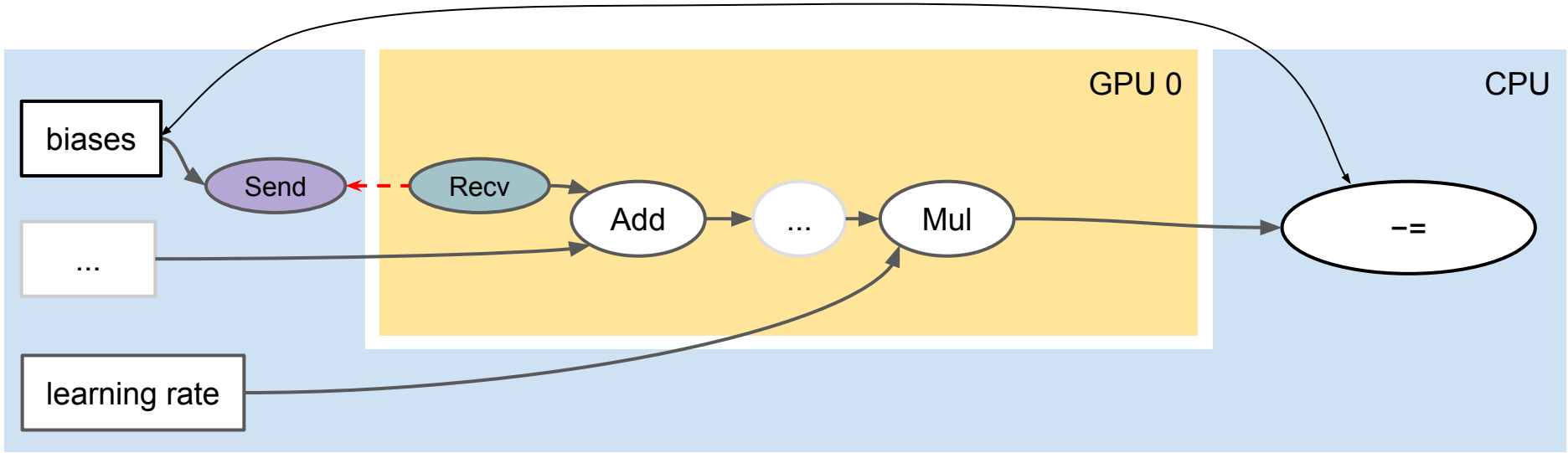


Distribution across Devices



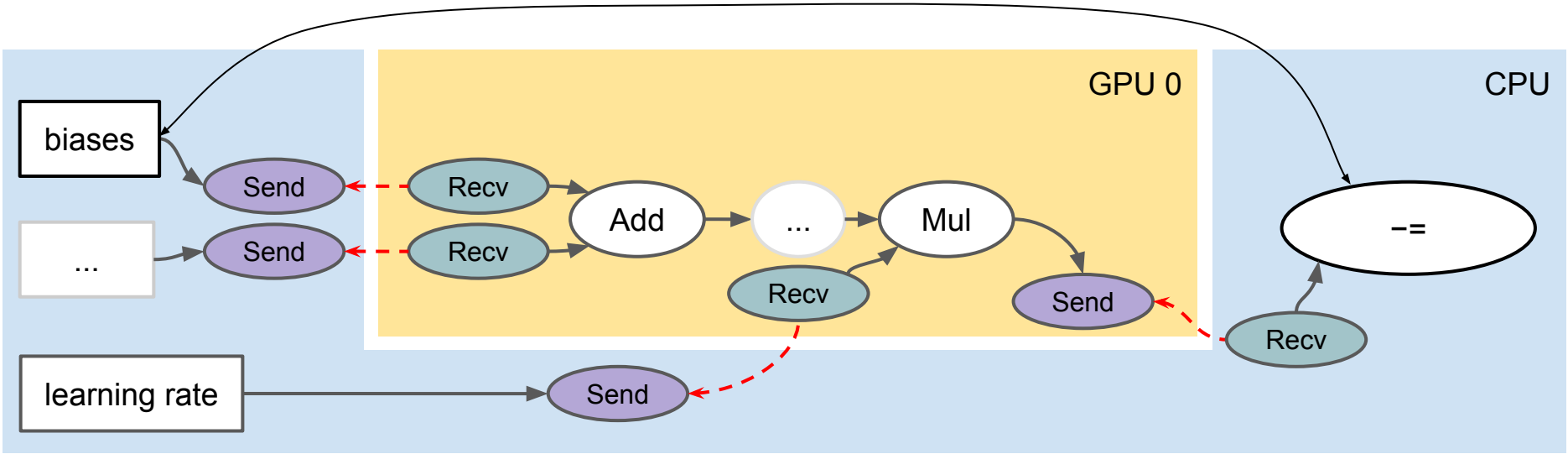
Distribution

- TensorFlow inserts *Send/Recv* Ops to transport tensors across devices
- *Recv* ops pull data from *Send* ops

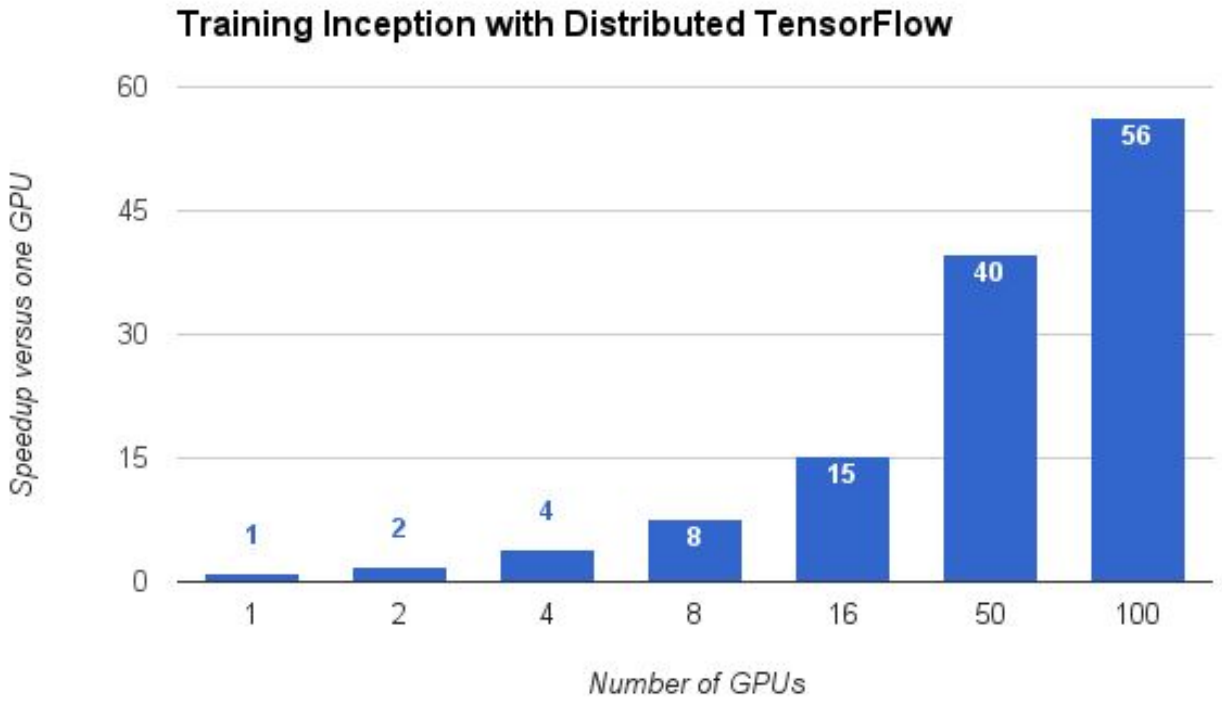


Distribution

- TensorFlow inserts *Send/Recv* Ops to transport tensors across devices
- *Recv* ops pull data from *Send* ops

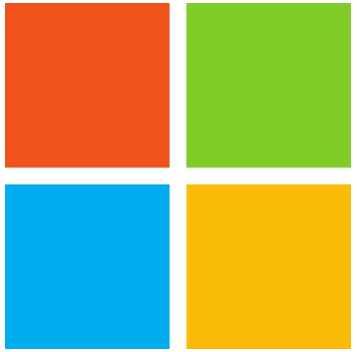


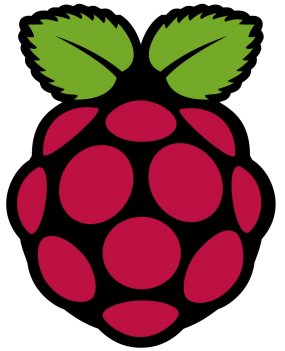
Scale across machines



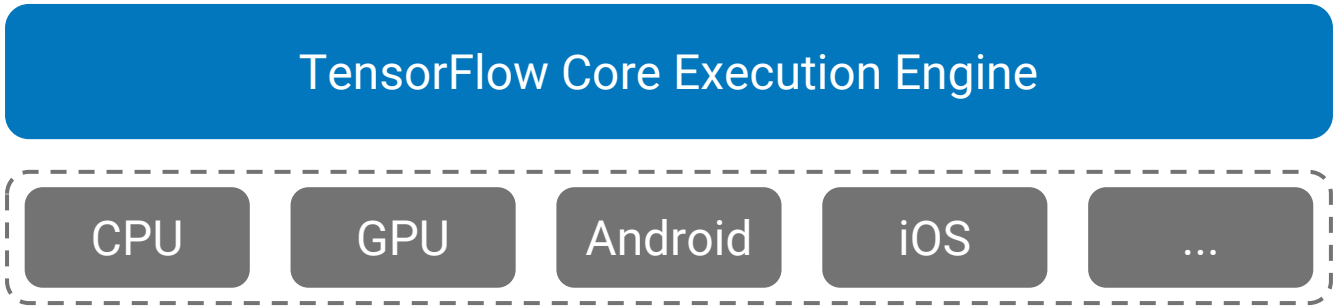
Portable

Platforms





Device Abstraction

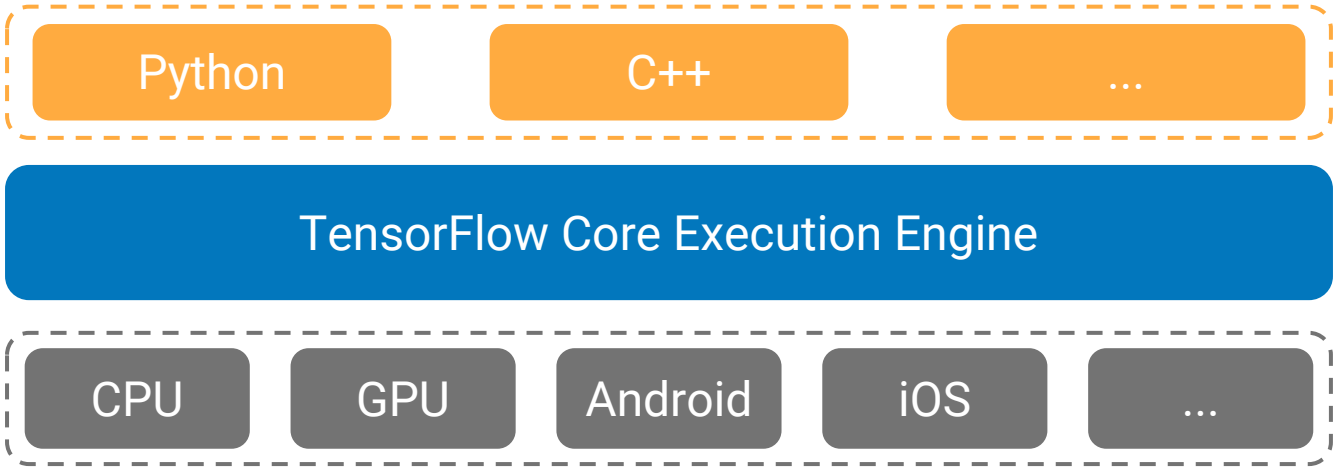


Languages

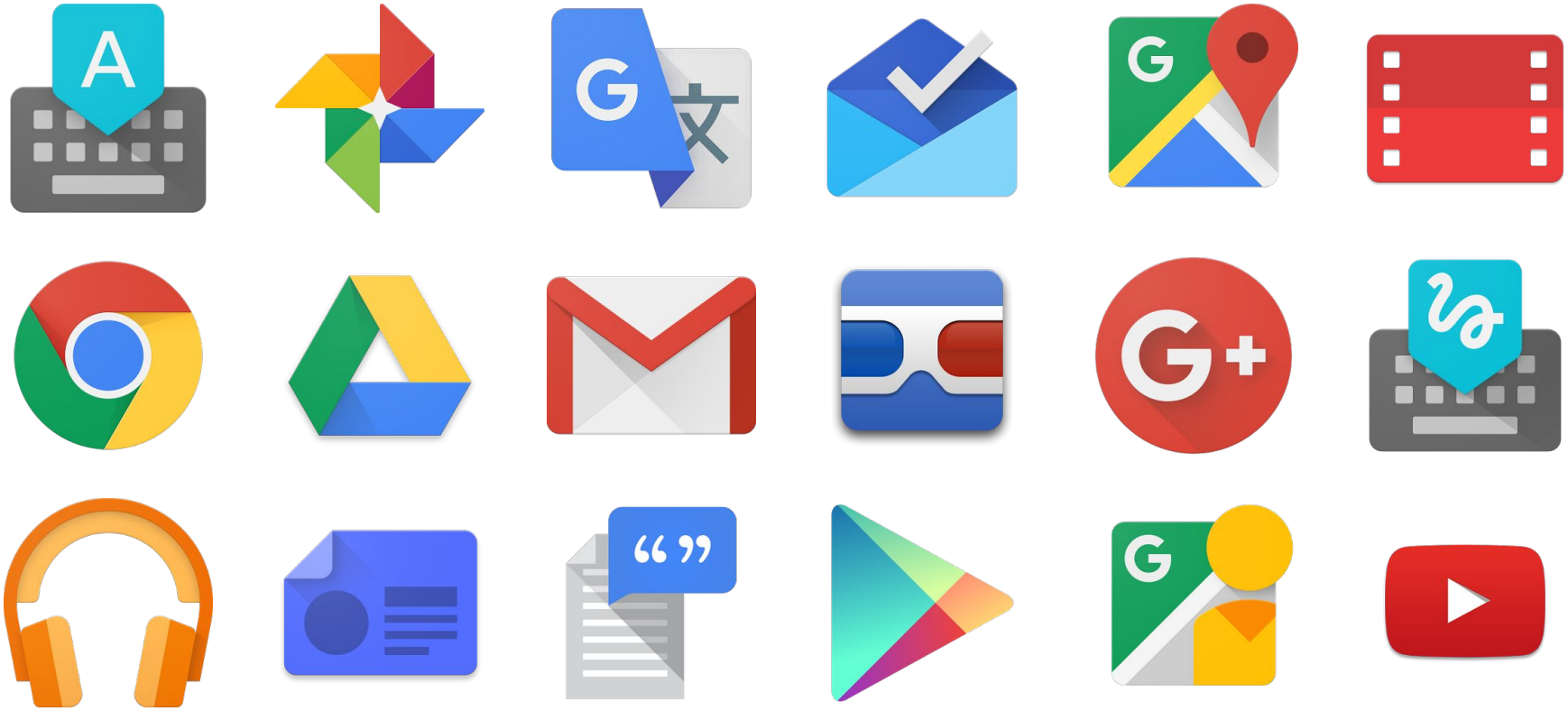
C++

Go

Language Abstraction



Now what?



2016-06-08 06:4 :45

2016-06-08 06:4 53

Thank You!

Rajat Monga
@rajatmonga