

Summary



Unified Platform

Spark SQL

Spark
Streaming

PySpark
SparkR

MLlib
spark.ml

GraphX

Spark Core

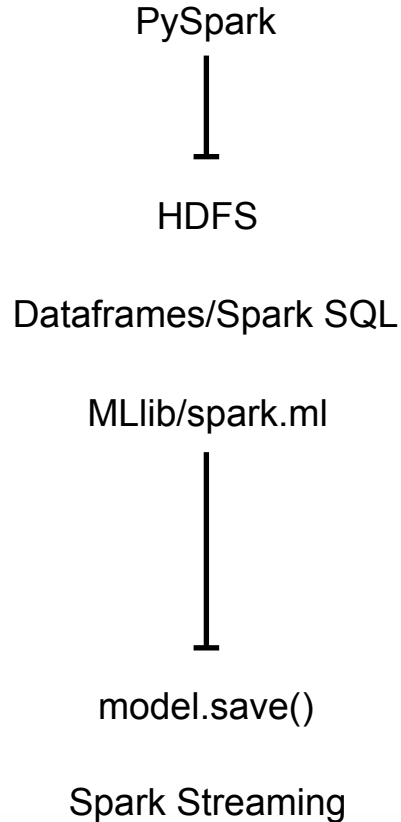
Standalone Scheduler

YARN

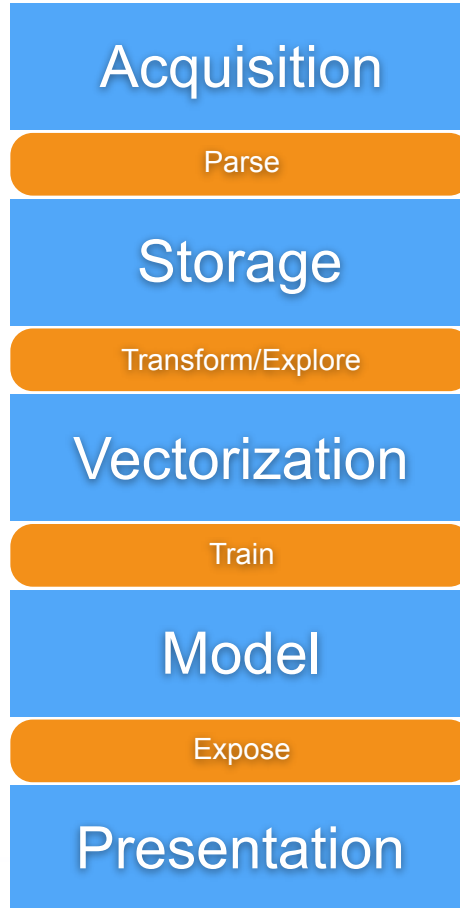
Mesos



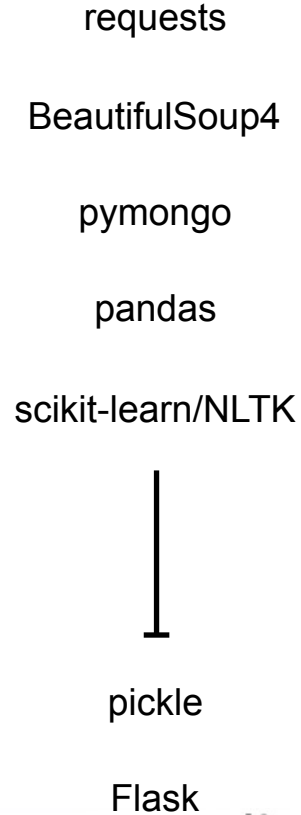
At Scale



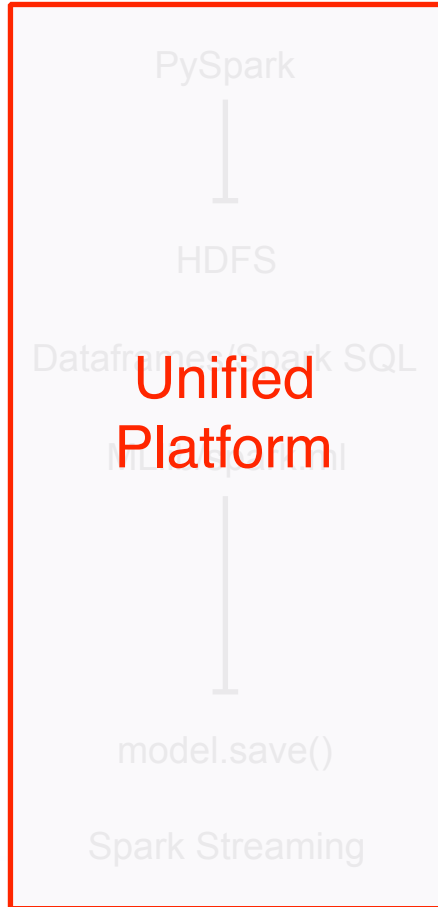
Data Pipeline



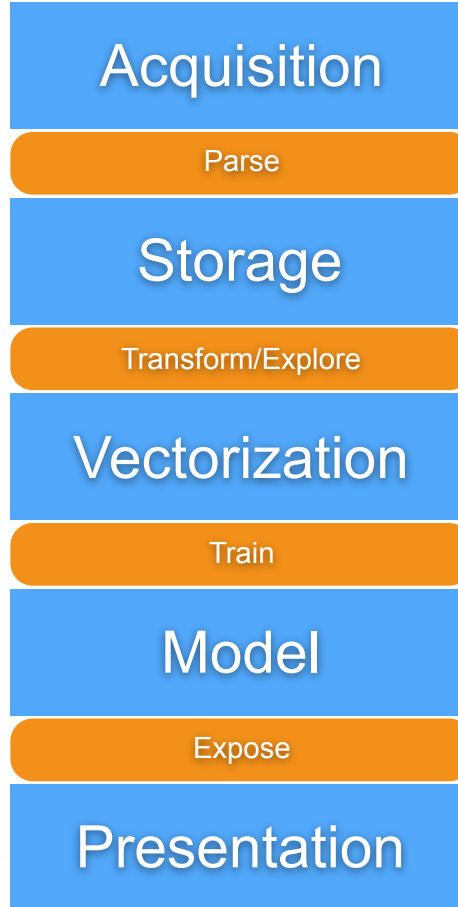
Locally



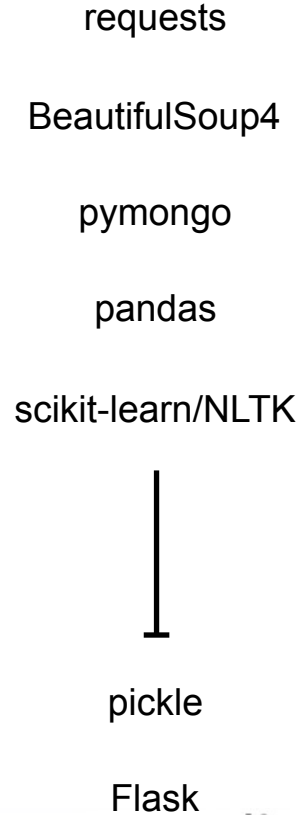
At Scale



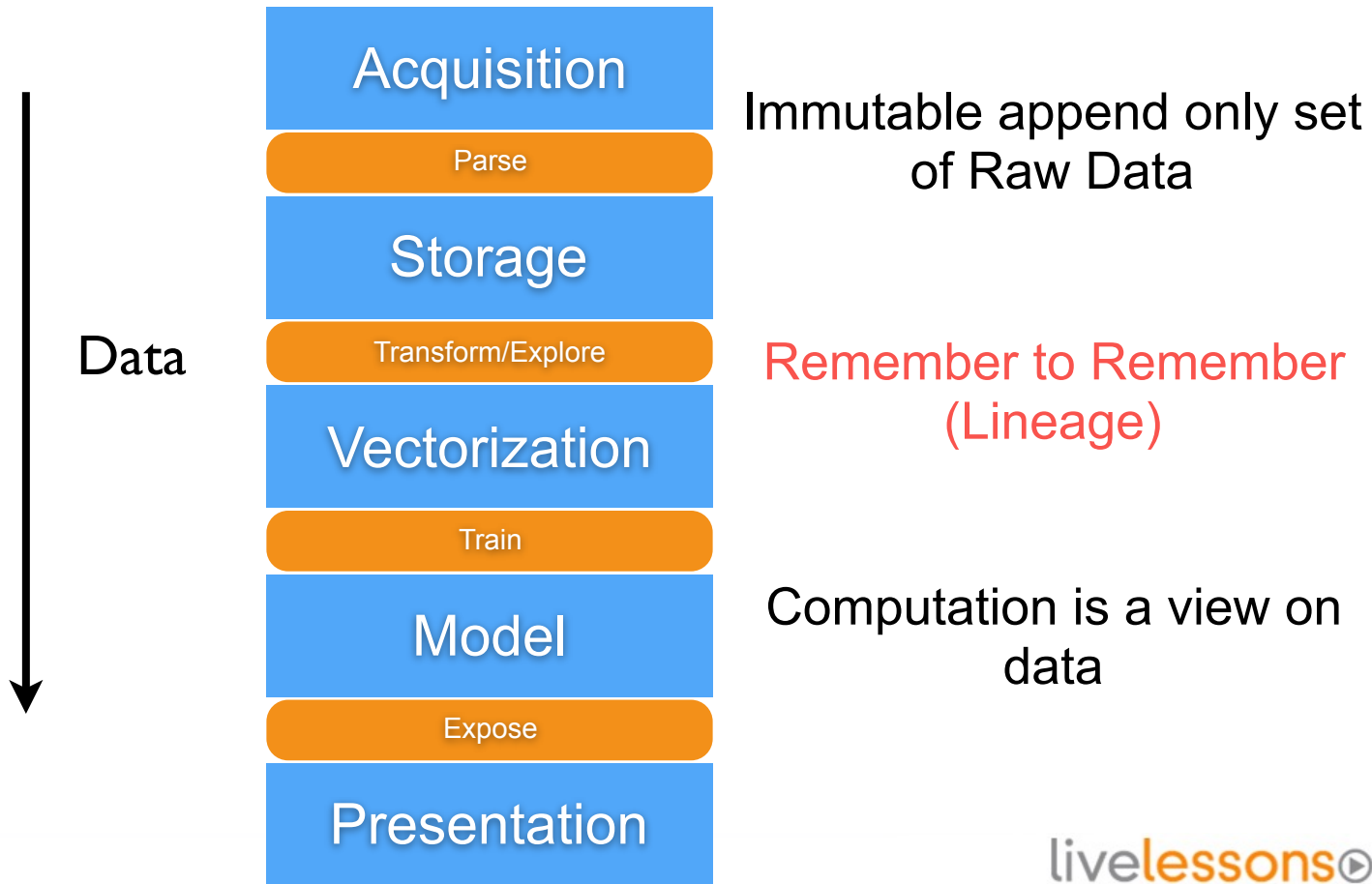
Data Pipeline



Locally



Data Pipeline



Remember: How to Scale



Next Up: Keep Learning!

