



Hadoop Ecosystem

In 20 Minutes!

Kai Voigt, Cloudera Inc.
Berlin Buzzwords, June 4th, 2012

cloudera



A Hadoop Cluster

cloudera

HDFS

MapReduce



HDFS

Java



MapReduce



Java



HDFS

SQL

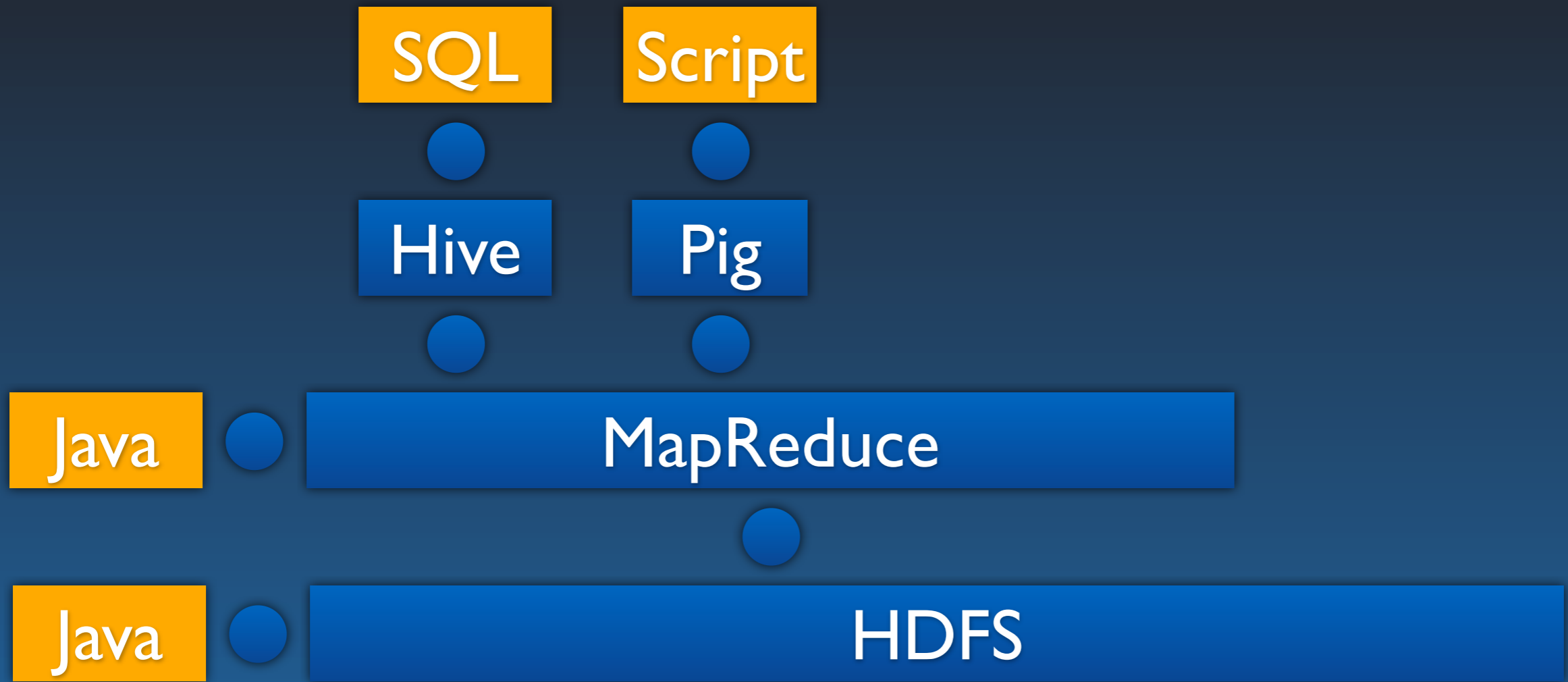
Hive

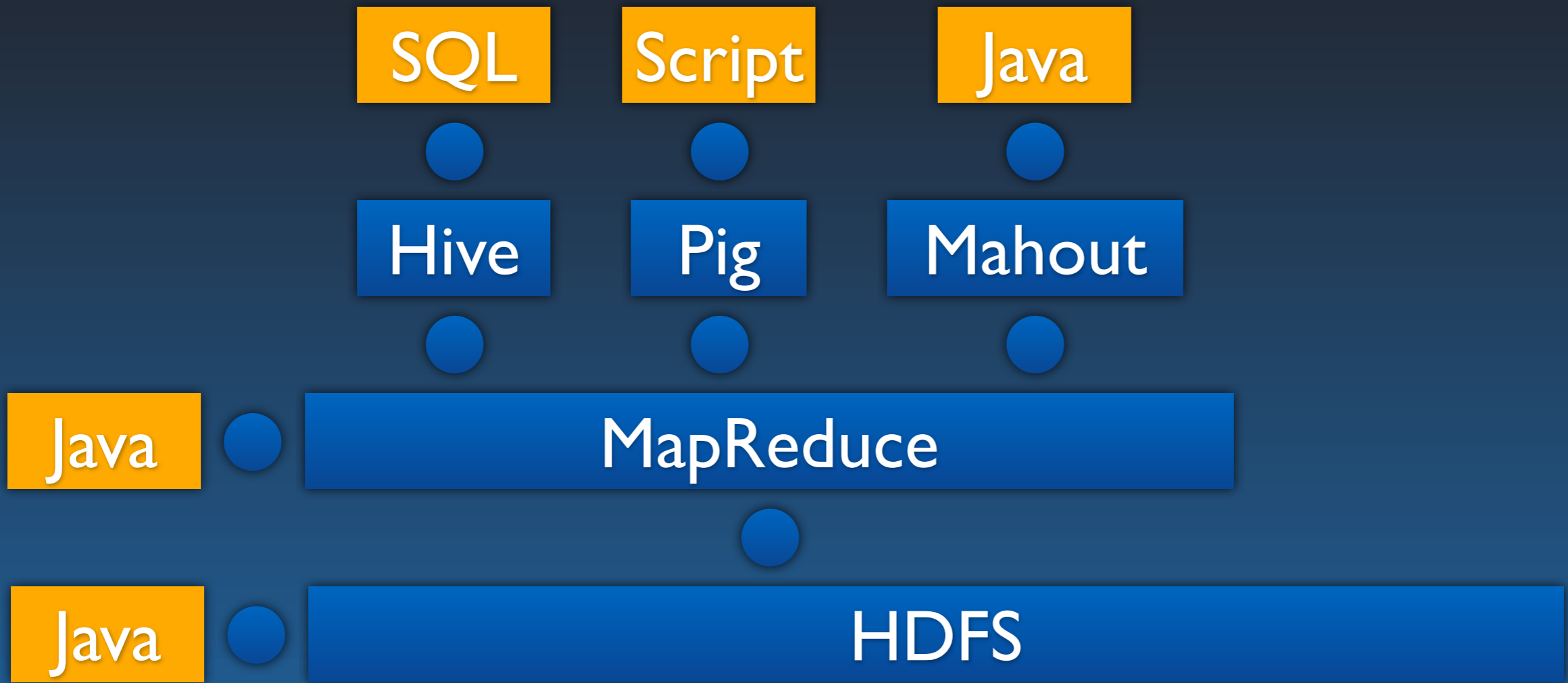
Java

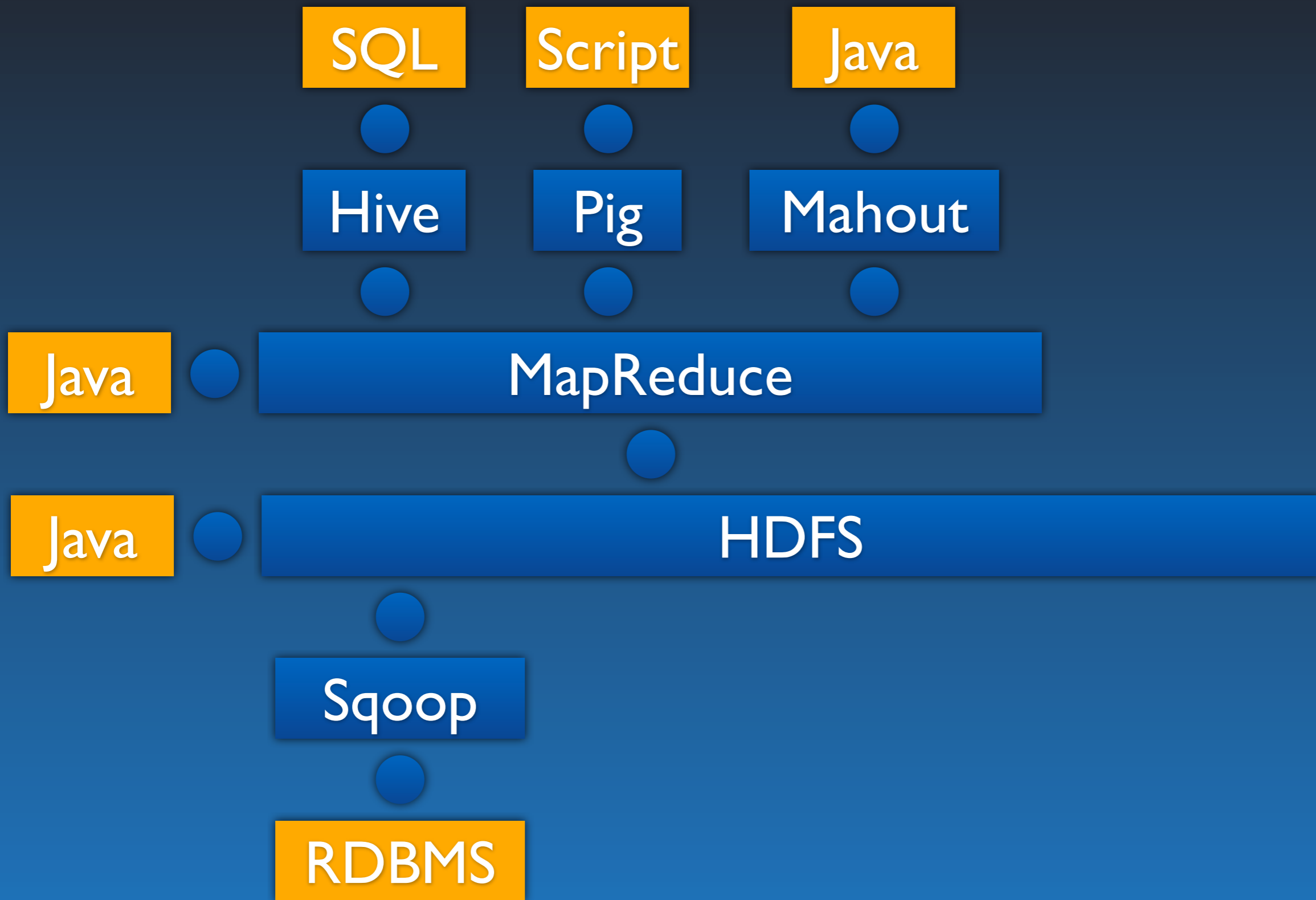
MapReduce

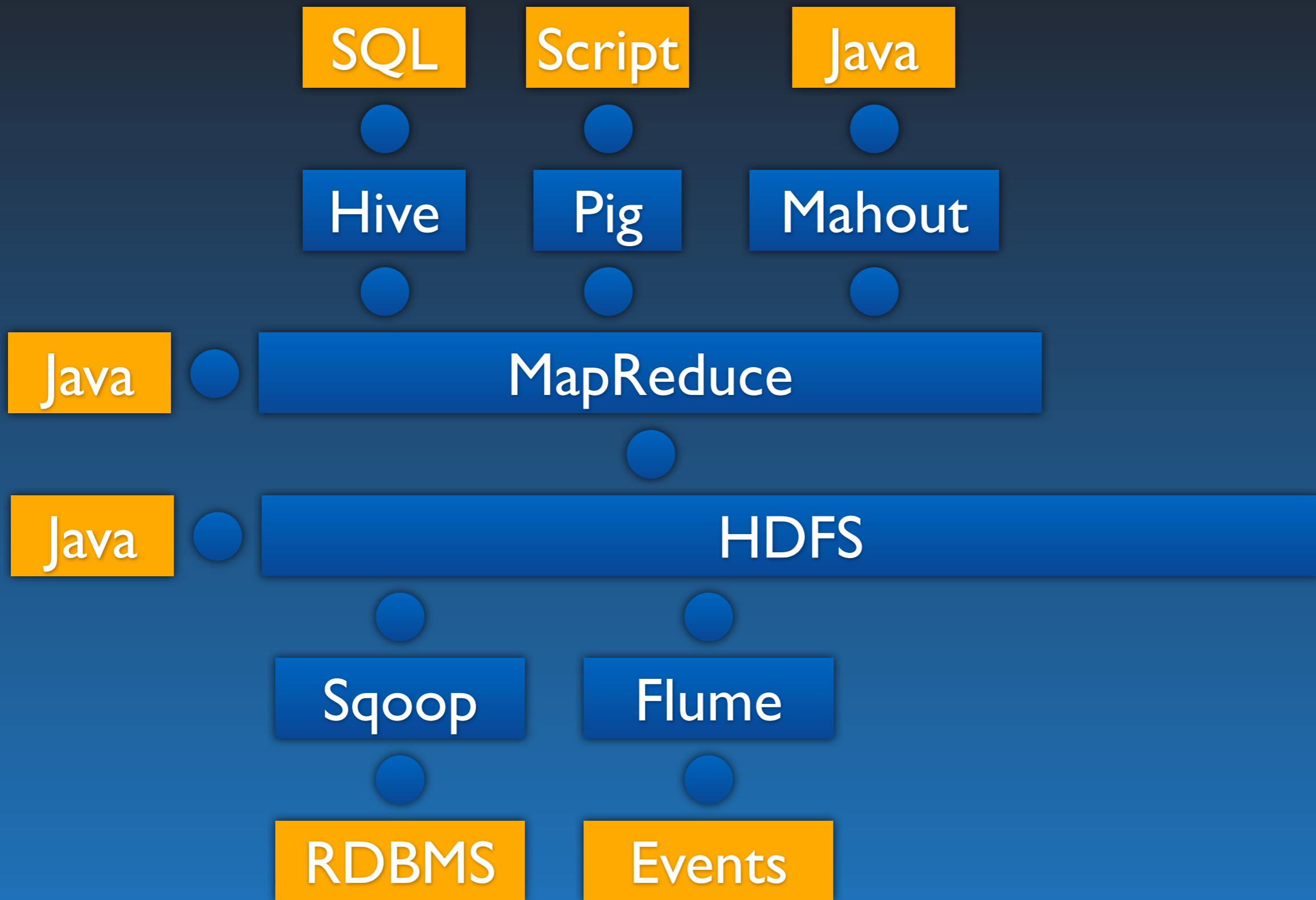
Java

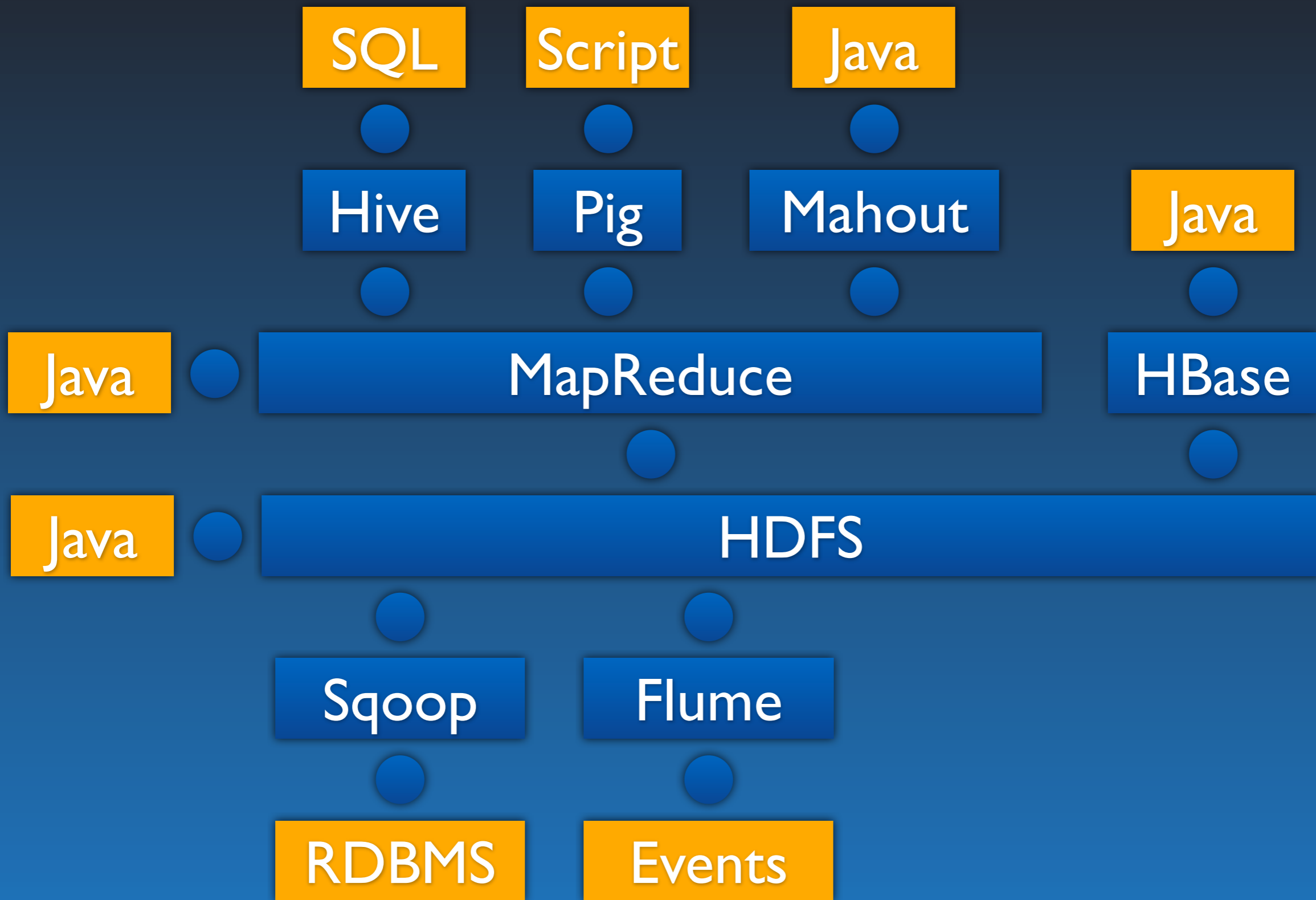
HDFS

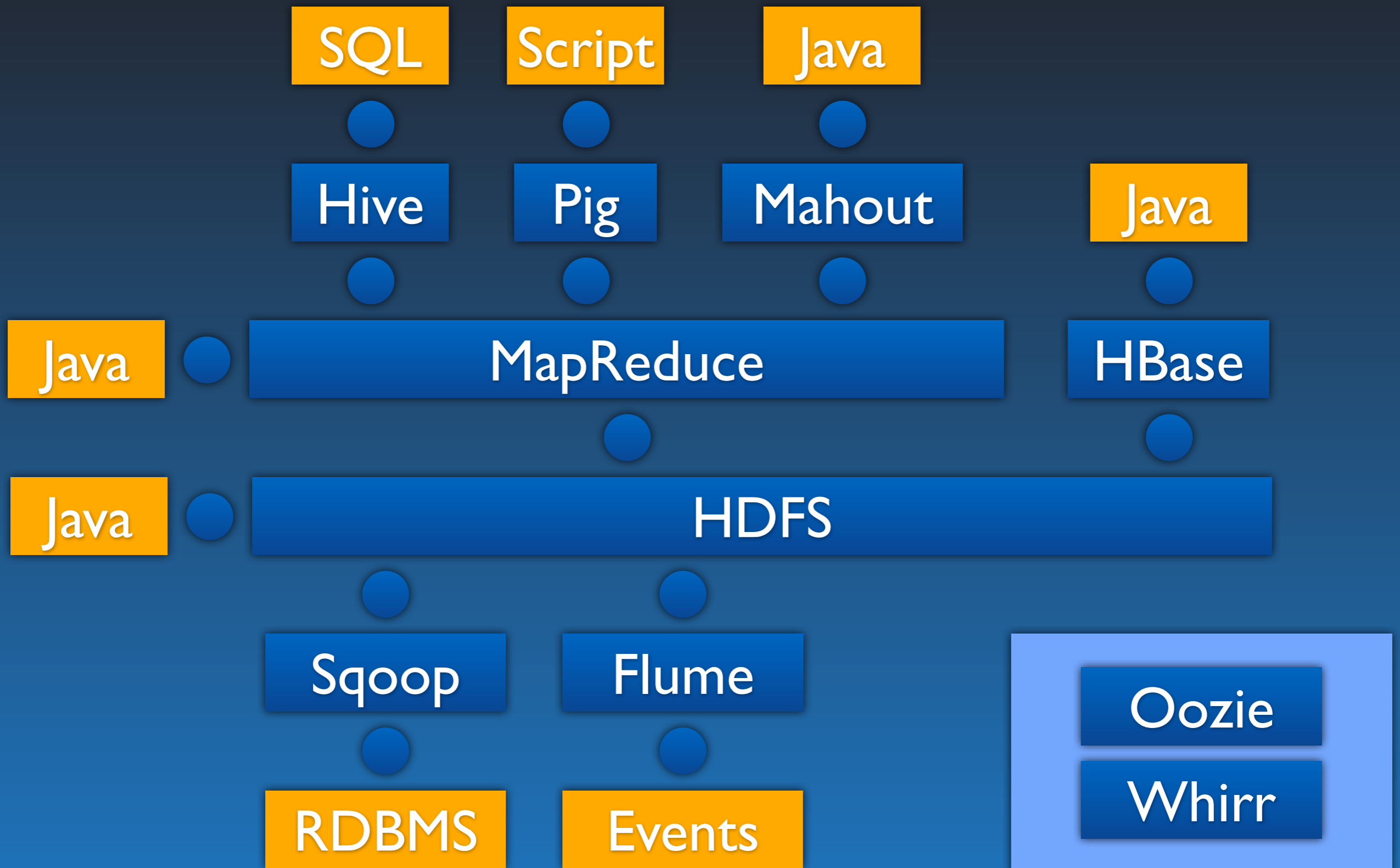












CDH 4.0



- Cloudera's Distribution Including Hadoop
- <http://www.cloudera.com/>
- Packages and Virtual Machines
- True Apache