



HDP Developer: Apache Spark

Hortonworks. We do Hadoop

A Hortonworks University Training Course

Course Outline

Day 1

- Hadoop Essentials Review
- Introduction to Spark
- Spark Hello World
- Intro to RDDs
- Intro to HDFS

Day 2

- Parallel Programming with Spark
- Spark Application Components
- Caching and Persisting Data
- Creating a Spark Application
- Spark Advanced Features and Improving Performance

Day 3

- Spark SQL
- Spark Streaming
- Spark MLlib

© Hortonworks Inc. 2011 – 2014. All Rights Reserved



Logistics

- Daily schedule and breaks
 - 9 a.m. – 5 p.m.
 - 15 min breaks every hour or so
 - 1 hour for lunch
- Restrooms
- Lunch
- Computers and the VM environment

© Hortonworks Inc. 2011 – 2014. All Rights Reserved



Introductions

Please share:

- Your name
- Job responsibilities
- Previous Hadoop experience (if any)
- Your expectation for this class

© Hortonworks Inc. 2011 – 2014. All Rights Reserved

