TRAINING AGENDA
HORTONWORKS DATA PLATFORM (HDP®) REAL-TIME DEVELOPMENT

2 DAYS
INTERMEDIATE   **DEVELOPER**

This 2 day training course is designed for developers who need to create real-time applications to ingest and process streaming data sources using Hortonworks Data Platform (HDP). Specific technologies covered includes: Apache Hadoop, Apache Kafka, Apache Storm & Trident, Apache Spark and Apache HBase. The highlight of the course is the custom workshop-styled labs that will allow participants to build complete streaming applications with Storm and Spark Streaming.

PREREQUISITES
Students should be familiar with programming principles and have experience in software development. **Java programming experience is required.** SQL and light scripting knowledge is also helpful. No prior Hadoop knowledge is required.

TARGET AUDIENCE
Developers and data engineers who need to understand and develop real-time and streaming applications on HDP.

AGENDA SUMMARY

Day 1: HDP Real-Time Architecture and Components
Day 2: Real-Time Processing with Storm & Spark Streaming

Day 1: HDP Real-Time Architecture and Components
- Real-time architecture & overview of the class
- Identify the relevant HDP components
    - Kafka, Storm, Spark Streaming
- Kafka architecture
- HBase architecture
- Storm architecture & building Storm topologies

Hands-on Labs

- Creating Kafka topics from CLI and publishing & consuming messages from Java
- Intro to Storm
- Optional: Creating & accessing HBase tables from HBase shell and from Java

Day 2: Real-Time Processing with Storm
- Integrating Storm with Kafka
- Spark ecosystem overview
- RDD Programming
- Spark Streaming
- Integrating Spark with Kafka

Hands-On Labs

- Integrating Storm with Kafka
- Intro to Spark RDD
- Intro to Spark Streaming
- Integrating Spark with Kafka

About Hortonworks

Hortonworks is a leading innovator at creating, distributing and supporting enterprise-ready open data platforms. Our mission is to manage the world's data. We have a single-minded focus on driving innovation in open source communities such as Apache Hadoop, NiFi, and Spark. Our open Connected Data Platforms power Modern Data Applications that deliver actionable intelligence from all data: data in-motion and data-at-rest. Along with our 1600+ partners, we provide the expertise, training and services that allows our customers to unlock the transformational value of data across any line of business. We are Powering the Future of Data™.

Contact

For further information visit   +1 408 675-0983
www.hortonworks.com   +1 855 8-HORTON
                                     INTL: +44 (0) 20 3826 1405

- Optional: Consuming a Kafka topic with a Storm topology and publishing results to HBase