



Hadoop for System Administrators H6C60S (GL660)

HPE course number	H6C60S
Course length	3 days
Delivery mode	ILT
View schedule, local pricing, and register	View now
View related courses	View now

This course covers the essentials of deploying and managing an Apache Hadoop cluster. The course is lab intensive with each participant creating their own Hadoop cluster using either the CDH (Cloudera's Distribution, including Apache Hadoop) or Hortonworks Data Platform stacks. Core Hadoop services are explored in depth with emphasis on troubleshooting and recovering from common cluster failures. The fundamentals of related services such as Ambari, Zookeeper, Pig, Hive, HBase, Sqoop, Flume, and Oozie are also covered. The course is approximately 60 percent lecture and 40 percent labs.

Why HPE Education Services?

- IDC MarketScape leader 4 years running for IT education and training*
- Recognized by IDC for leading with global coverage, unmatched technical expertise, and targeted education consulting services*
- Key partnerships with industry leaders OpenStack®, VMware®, Linux®, Microsoft®, ITIL, PMI, CSA, and (ISC)²
- Complete continuum of training delivery options—self-paced eLearning, custom education consulting, traditional classroom, video on-demand instruction, live virtual instructor-led with hands-on lab, dedicated onsite training
- Simplified purchase option with HPE Training Credits

Audience

Systems Administrators who will be responsible for managing and administering Hadoop clusters.

Supported distributions

Red Hat® Enterprise Linux 6

Prerequisites

Qualified participants should be comfortable with the Linux commands and have some systems administration experience, but do not need previous Hadoop experience.

Benefits to you

Hands-on coverage of Hadoop gives systems administrators the skills they need to properly deploy, manage, and maintain Hadoop clusters.

*Realize Technology Value with Training, IDC Infographic 2037, Sponsored by HPE, January 2016

Detailed course outline

Module 1: Hadoop: The Big Picture

- Data Analysis
 - Big Data
 - Hadoop Core Architecture
 - Hadoop Ecosystem
 - Hadoop Ecosystem continued
 - Running Commands on Multiple Systems
-

Module 2: HDFS

- Design Goals
 - Design
 - Blocks
 - Block Replication
 - Namenode Daemon
 - Secondary Namenode Daemon
 - Datanode Daemon
 - Accessing HDFS
 - Permissions and Users
 - Adding and Removing Datanodes
 - Balancing
-

Module 3: MapReduce

- Terminology and Data Flow
 - MapReduce Daemons
 - MapReduce Essential Configuration
 - Failure and Recovery
 - YARN
-

Module 4: MapReduce Schedulers

- Working with Jobs
 - Scheduling Concepts
 - FIFO Scheduler
 - Fair Scheduler
 - Fair Scheduler: Configuration
-

Appendix A: Installing Hadoop with Ambari

Detailed lab guide

Lab 1:	<ul style="list-style-type: none">• Running Commands on Multiple Hosts (R6)• Preparing to Install Hadoop (CDH5 R6)
Lab 2:	<ul style="list-style-type: none">• Single Node HDFS (CDH5 R6)• Multi-node HDFS (CDH5 R6)• Files and HDFS (CDH5 R6)• Managing and Maintaining HDFS (CDH5 R6)
Lab 3:	<ul style="list-style-type: none">• MapReduce (CDH5 R6)
Lab 4:	<ul style="list-style-type: none">• MapReduce Schedulers (R6)
Lab 5:	<ul style="list-style-type: none">• Installing Hadoop with Ambari (R6)

Next steps

- View our Big Data curriculum path
- Hewlett Packard Enterprise also recommends the HPE Roadmap Services for Hadoop from our Consulting team

Learn more at
hpe.com/ww/learnlinux

Follow us:



© Copyright 2015–2016 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

Microsoft is either a registered trademark or trademark of Microsoft Corporation in the United States and/or other countries. Red Hat is a registered trademark of Red Hat, Inc. in the United States and other countries. The OpenStack Word Mark is either a registered trademark/service mark or trademark/service mark of the OpenStack Foundation, in the United States and other countries and is used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation or the OpenStack community. Pivotal and Cloud Foundry are trademarks and/or registered trademarks of Pivotal Software, Inc. in the United States and/or other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries. VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. All other third-party trademark(s) is/are property of their respective owner(s).

c04577721, December 2016, Rev. 3