



HP Education Services Course Overview

Hadoop for System Administrators (GL660) (H6C60S)

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% lecture and 40% labs.

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

Why education services from HP?

- Named a 2014 IDC MarketScape leader* in IT education in Cloud, Security, Big Data, Storage, ITIL/ITSM, Project Management training, after being recognized in 2013 for being an IDC MarketScape leader in IT education overall
- Global reach through 90 training centers in 45+ languages, with access to over 800+ experienced instructors

Course title:	Hadoop for System Administrators (GL660)
HP product number:	H6C60S
Category/Subcategory:	Big Data
Course length:	3 days
Level:	Intermediate
Delivery language:	English
To order:	You can register your interest for this course online at http://www.hp.com.au/education . At the site, select the course under Big Data portfolio and you will see dates for the course. Register your interest for the date of your choice.

- Job-focused courses on HP technologies leading to HP ExpertOne certifications
- Wide range of education consulting services tailored to your specific needs to prepare you for IT transformation projects

- Flexibility to learn through a wide variety of delivery modalities: traditional ILT (Instructor-led), VILT (Virtual instructor-led), WBT (Web based training), games and simulations

* IDC MarketScape: US Cloud , Security, Big Data, Storage, ITIL/ITSM, Project Management Training Ecosystem, August 2014 Vendor Assessments, IDC #250180, IDC #250186 , IDC #250181, IDC #250182, IDC #250185, IDC #250178, #239139.

Next steps

- View our Big Data curriculum path
- HP also recommends the HP Roadmap Services for Hadoop from our HP Consulting team

Detailed course outline

Chapter 1 - Hadoop: The Big Picture

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

Lab Tasks

- Running Commands on Multiple Hosts [R6]
- Preparing to Install Hadoop [CDH5 R6]

Chapter 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

Lab Tasks

- Single Node HDFS [CDH5 R6]
- Multi-node HDFS [CDH5 R6]
- Files and HDFS [CDH5 R6]
- Managing and Maintaining HDFS [CDH5 R6]

Chapter 3 - MapReduce

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

Lab Tasks

- MapReduce [CDH5 R6]

Chapter 4 - MapReduce Schedulers

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

Lab Tasks

- MapReduce Schedulers [R6]

Appendix A - Installing Hadoop with Ambari

Lab Tasks

- Installing Hadoop with Ambari [R6]

For more information

To locate contact information and to learn more about education services, please visit our web site at <http://www.hp.com.au/education> .

