

HDP Developer: Apache Pig and Hive

AWS Guacamole Setup Guide

Rev 6





Copyright © 2012 - 2016 Hortonworks, Inc. All rights reserved.

The contents of this course and all its lessons and related materials, including handouts to audience members, are Copyright © 2012 - 2015 Hortonworks, Inc.

No part of this publication may be stored in a retrieval system, transmitted or reproduced in any way, including, but not limited to, photocopy, photograph, magnetic, electronic or other record, without the prior written permission of Hortonworks, Inc.

This instructional program, including all material provided herein, is supplied without any guarantees from Hortonworks, Inc. Hortonworks, Inc. assumes no liability for damages or legal action arising from the use or misuse of contents or details contained herein.

Linux® is the registered trademark of Linus Torvalds in the United States and other countries.

Java® is a registered trademark of Oracle and/or its affiliates.

All other trademarks are the property of their respective owners.



Become a **Hortonworks Certified Professional** and establish your credentials:

- HDP Certified Developer: for Hadoop developers using frameworks like Pig, Hive, Sqoop and Flume.
- HDP Certified Administrator: for Hadoop administrators who deploy and manage Hadoop clusters.
- HDP Certified Developer: Java: for Hadoop developers who design, develop and architect Hadoop-based solutions written in the Java programming language.
- HDP Certified Developer: Spark: for Hadoop developers who write and deploy applications for the Spark framework.
- HDF Certified Professional: for DataFlow Operators responsible for building and deploying HDF workflows.

How to Register: Visit www.examslocal.com and search for “Hortonworks” to register for an exam. The cost of each exam is \$250 USD, and you can take the exam anytime, anywhere using your own computer. For more details, including a list of exam objectives and instructions on how to attempt our practice exams, visit <http://hortonworks.com/training/certification/>

Earn Digital Badges: Hortonworks Certified Professionals receive a digital badge for each certification earned. Display your badges proudly on your résumé, LinkedIn profile, email signature, etc.





Self Paced Learning Library

On Demand Learning

Hortonworks University Self-Paced Learning Library is an on-demand dynamic repository of content that is accessed using a Hortonworks University account. Learners can view lessons anywhere, at any time, and complete lessons at their own pace. Lessons can be stopped and started, as needed, and completion is tracked via the Hortonworks University Learning Management System.

Hortonworks University courses are designed and developed by Hadoop experts and provide an immersive and valuable real world experience. In our scenario-based training courses, we offer unmatched depth and expertise. We prepare you to be an expert with highly valued, practical skills and prepare you to successfully complete Hortonworks Technical Certifications.

Target Audience: Hortonworks University Self-Paced Learning Library is designed for those new to Hadoop, as well as architects, developers, analysts, data scientists, and IT decision makers. It is essentially for anyone who desires to learn more about Apache Hadoop and the Hortonworks Data Platform.

Duration: Access to the Hortonworks University Self-Paced Learning Library is provided for a 12-month period per individual named user. The subscription includes access to over 400 hours of learning lessons.

The online library accelerates time to Hadoop competency. In addition, the content is constantly being expanded with new material, on an ongoing basis.

Visit: <http://hortonworks.com/training/class/hortonworks-university-self-paced-learning-library/>

System Requirements

The following is the recommended minimal system requirements:

- Reasonably powerful **x86/amd64 hardware**. Intel Core2 Duo or AMD Athlon Dual-Core or equivalent or above.
- RAM: At least **2GB**
- Free disk space: **2GB**
- A good consistent Internet connection: broadband, cable, DSL, etc. (wired connection preferred)
- Client machine should be connected to Internet directly without sending traffic through an HTTP proxy. Internet connection through an HTTP proxy is not supported.
- **OPTIONAL** - A good way to test the quality of Internet is using this utility:
<http://www.pingtest.net>

If it rates the internet-quality as “B” or above then it means that the Internet-connection at the premises is suitable for the training.

- The following ports need to be open for OUTBOUND access from the participant's machines to Amazon Web Services:
 - 80
 - 22 (optional)
- **Required Browser:** IE is not supported currently. Firefox, Safari, Chrome will work.

IMPORTANT: If you have any issues or questions, please send an email to training-support@hortonworks.com.

Connect to AWS VM From a Browser

Objective:	To connect AWS lab VM from browser.
Before You Begin:	<ul style="list-style-type: none">• You need to execute these steps during the class.• Make sure your hardware meets the requirements listed above.• You have port 80 open on your firewall.• The instructor has already provided an IP address for the AWS VM.

Step 1: Connect to AWS VM from browser

1.1. Get an IP address from your instructor

1.2. Open your favorite browser (IE is not supported) and type following URL:

```
http://<IP ADDRESS>/guacamole
```

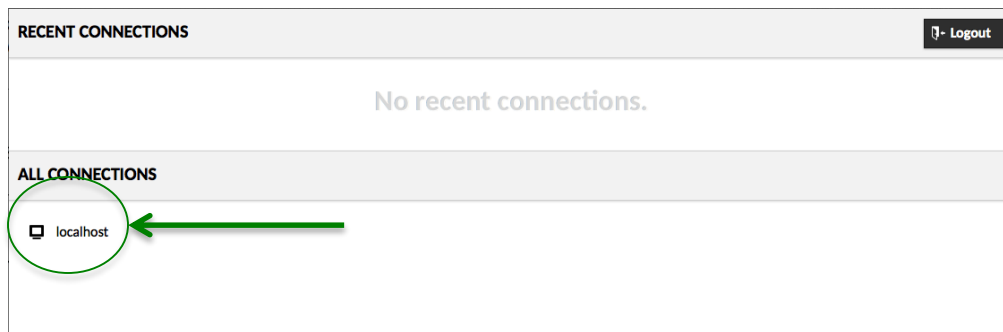
1.3. Enter User ID and Password:

User ID: remote

Password: D3skt0p

The image shows the Guacamole 0.9.5 login interface. At the top is the Guacamole logo, which consists of a stylized green and black circular icon. Below the logo, the text "GUACAMOLE 0.9.5" is displayed. Underneath, there are two input fields: the first contains the text "remote" and the second contains a series of dots representing a password. At the bottom of the form is a dark grey button with the word "Login" in white text.

1.4. Click on '**localhost**' and now you are connected to the lab VM:



Now, you will be connected to AWS VM and you should see Hortonworks Lab environment Dekstop (**one of these**):



RESULT: You should have tested the connection to the classroom VM running on AWS.

Classes Available Worldwide Through Our Partners



Study Options Worldwide

In combination with our partner providers, classes are often available in numerous locations across the world.



Private On-site Training

Hortonworks training in-house covers all of our basic coursework, and provides a more intimate setting for 6 or more students.

[Contact us for more details](#)



Learn from the company focused solely on Hadoop.



What Makes Us Different?

1. Our courses are designed by the **leaders and committers** of Hadoop
2. We provide an **immersive** experience in **real-world** scenarios
3. We prepare you to **be an expert** with highly valued, **fresh skills**
4. Our courses are available **near you**, or accessible **online**

Hortonworks University courses are designed by the leaders and committers of Apache Hadoop. We provide immersive, real-world experience in scenario-based training. Courses offer unmatched depth and expertise available in both the classroom or online from anywhere in the world. We prepare you to be an expert with highly valued skills and for Certification.