

Implementing Real-Time Analysis with Hadoop in Azure HDInsight

Lab Setup Guide

Overview

This course includes optional labs in which you can try out the techniques demonstrated in the course for yourself.

What You'll Need

To complete the labs, you will need the following:

- A web browser
- A Microsoft account
- A Microsoft Azure subscription
- A Microsoft Windows computer on which Microsoft Visual Studio with the Azure SDK has been installed
- The lab files for this course

Creating a Free Trial Azure Subscription

If you already have a Microsoft Azure subscription, you can skip this section. Otherwise, follow these steps to create a free trial subscription. You will need to provide a valid credit card number for verification, but you will not be charged for Azure services – for more information, see <http://azure.microsoft.com/pricing/free-trial-faq/>.

Note: Microsoft Azure free trial subscriptions are not available in all regions, and are restricted to a single 30-day trial per Microsoft account.

1. If you already have a Microsoft account that has not already been used to sign up for a free Azure trial subscription, you're ready to get started. If not, don't worry, just create a new Microsoft account at <https://signup.live.com>.
2. After you've created a Microsoft account, browse to <http://azure.microsoft.com> and click the **Free Trial** link. Then follow the instructions to sign up for a free trial subscription to Microsoft Azure. You'll need to sign-in with your Microsoft account if you're not already signed in. Then you'll need to:
 - a. Enter your cellphone number and have Microsoft send you a text message to verify your identity.
 - b. Enter the code you have been sent to verify it.

- c. Provide valid payment details. This is required for verification purposes only – your credit card won't be charged for any services you use during the trial period, and the account is automatically deactivated at the end of the trial period unless you explicitly decide to keep it active.

Setting Up a Client Computer

In addition to an Azure subscription, in which your HDInsight cluster will be hosted, you'll need a client computer from which you can manage the cluster and the big data processing jobs you will run on it. You can use any Internet-connected computer running Windows. The lab procedures assume you are using Windows 10, but they can be adapted for any of the following operating systems:

- Windows 7
- Windows 8
- Windows 8.1
- Windows 10
- Windows Server 2012
- Windows Server 2012 R2

You can download an evaluation edition of Windows 10 from <https://www.microsoft.com/en-us/evalcenter/evaluate-windows-10-enterprise>.

Note: If you do not have a Windows PC you can use for these labs, consider installing Windows in a virtual machine (VM), either using virtualization software on your own computer or in a virtual machine in Microsoft Azure.

Installing Microsoft Visual Studio and the Azure SDK

The Microsoft Azure PowerShell module includes cmdlets for managing Microsoft Azure services, including Microsoft Azure HDInsight. Visual Studio, when used with the Microsoft Azure SDK provides a development environment for building Azure solutions; including HDInsight Big Data processing solutions.

Install Visual Studio and the Azure SDK from the Web Platform Installer

1. On the computer where you intend to perform the labs, start a web browser and navigate to <http://www.microsoft.com/web/downloads/platform.aspx>.
2. Download the Microsoft Web Platform Installer and run it.
3. In the Web Platform Installer, on the **Products** tab, next to **Visual Studio Community 2015 with Microsoft Azure SDK 2.7.x**, click **Add**. Then click **Install** and accept the license agreement to install Visual Studio with the Azure SDK. (This will take a considerable while depending on your Internet connection, so now would be a good time to resume viewing the course content until installation is complete!)

Note: If you already have Visual Studio 2013 or 2015 installed on your computer, just install the **Microsoft Azure SDK Tools** for the version of Visual Studio you have installed.

4. When the installation is complete, restart the computer and log back in.

Note: If the installation is interrupted or fails for any reason, restart your computer and re-launch the Web Platform Installer at <http://www.microsoft.com/web/downloads/platform.aspx>.

5. On the Start menu, view all apps. Then right-click **Visual Studio 2015** (or whatever version you had previously installed) and click **Pin to Taskbar**:

Configure Visual Studio

1. On the taskbar, click the Visual Studio icon. When prompted, do not sign in.
2. Accept the default layout options, and start Visual Studio.
3. Wait for Visual Studio to start, and then close it.

Installing the Lab Files

The course materials for this course include files that are required to complete the labs.

Download the Lab Files

1. Download the lab files from <https://github.com/MicrosoftLearning/Implementing-Realtime-Analysis-with-Hadoop-in-HDInsight/raw/master/Labs/HDRTLabs.zip>.

Note: All of the handouts for this course are published in a GitHub repository at <http://microsoftlearning.github.io/Implementing-Realtime-Analysis-with-Hadoop-in-HDInsight/>. Links to individual files in this repository are provided throughout the course, or you can choose to download the entire repository as a ZIP or TAR archive from the repository project page.

2. Extract the **HDRTLabs.zip** archive you downloaded to **C:** to create the following folder structure:
3. Right-click the **C:\HDRTLabs** folder and click **Properties**. Then clear the **Read-only** option for the folder and click **OK**. When prompted, apply the changes to the folder and its subfolders and files.