

# Hadoop Fundamentals

*Lab Setup*





---

# Contents

<b>LAB 1</b>	<b>SETUP .....</b>	<b>4</b>
	1.1 SETTING UP AN ACCOUNT .....	4
	1.2 AMBARI CONSOLE.....	6
	1.3 SSH AND PUTTY .....	7
	1.4 FTP AND WINSCP FILE TRANSFER.....	8
	1.5 SUMMARY .....	11

## Lab 1 Setup

Before you can start the labs required for this course you will need to setup your workstation. This first lab will enable you to do that and ensure you have everything ready for the remaining exercises.

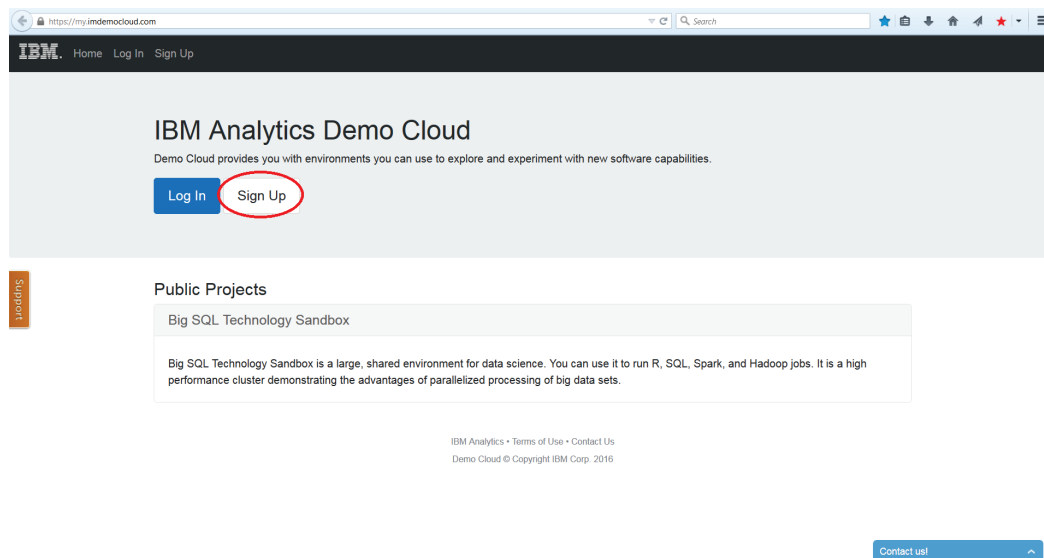
After completing this hands-on lab, you'll have:

- Have an IBM Analytics Demo Cloud Account created
- Install and be able to use the SSH and SFTP programs to connect to the cloud.
- Be able to start up Ambari, which will be used in the following labs.

Allow 30 minutes to complete this lab.

### 1.1 Setting up an account

1. Start by navigating to the IBM Analytics Demo Cloud website: <https://my.imdemocloud.com/> and clicking the Sign Up button. And fill out the required information.



2. An example of Sign Up and Log In.

The image shows two side-by-side forms: "Sign Up" and "Log In". The "Sign Up" form has fields for name (Daniel Tran), email (trand@ca.ibm.com), and two password fields. A blue "Sign up" button is at the bottom. Below the button is a link: "Already applied but did not receive a confirmation?". The "Log In" form has fields for email (trand@ca.ibm.com) and a password field with masked dots. A blue "Log in" button is below. Below the button is a "Remember me" checkbox and two links: "Forgot your password?" and "Didn't receive unlock instructions?". An error overlay is positioned between the two forms, displaying the following text: "High complexity password required", "Passwords are painful but necessary. Longer passwords are better.", "Bad password: P@ssword1!", and "If you're having trouble, try repeating your password a few times."

- After logging in, you will need to click the “Big SQL Technology Sandbox” link.

## Your Projects

**Big SQL Technology Sandbox**

Big SQL Technology Sandbox is a large, shared environment for data science. You can use it to run R, SQL, Spark, and Hadoop jobs. It is a high performance cluster demonstrating the advantages of parallelized processing of big data sets.

- Click the “Systems” tab, which will show 3 sections. We are interested in “Ambari Console” and “Connection Details”. Pressing the Launch button will open the Ambari Console. In Connection Details, the username is what will be used to log into Ambari. The server host name will be important in later steps (as well as the sample SSH command).

**Systems** Resources Feedback

BigSheets, Text Analytics, and Big SQL Console

Work with your data using spreadsheets and SQL

**Launch**

Ambari Console

Read-only view of different services and metrics on the cluster

**Launch**

Connection Details

Your **username is trand**. All services use your Demo Cloud password. Change the password in [your profile](#)

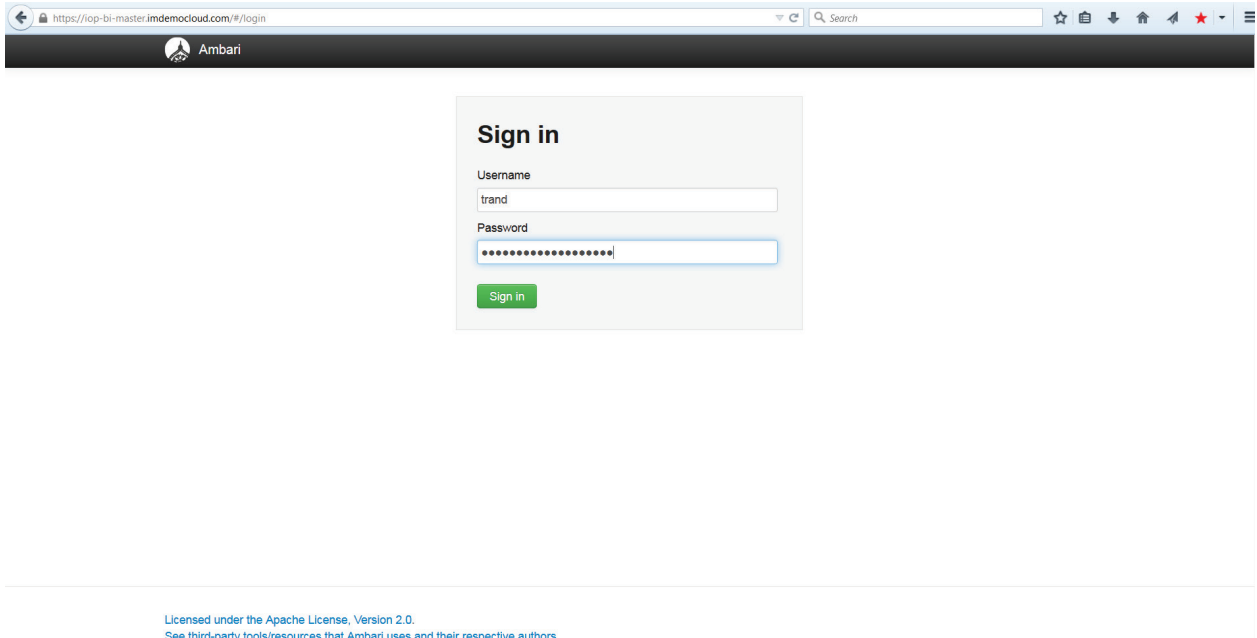
The server hostname is **iop-bi-master.imdemocloud.com**

You can use your Demo Cloud username and password to connect to the SSH command line on the cluster. Sample command:

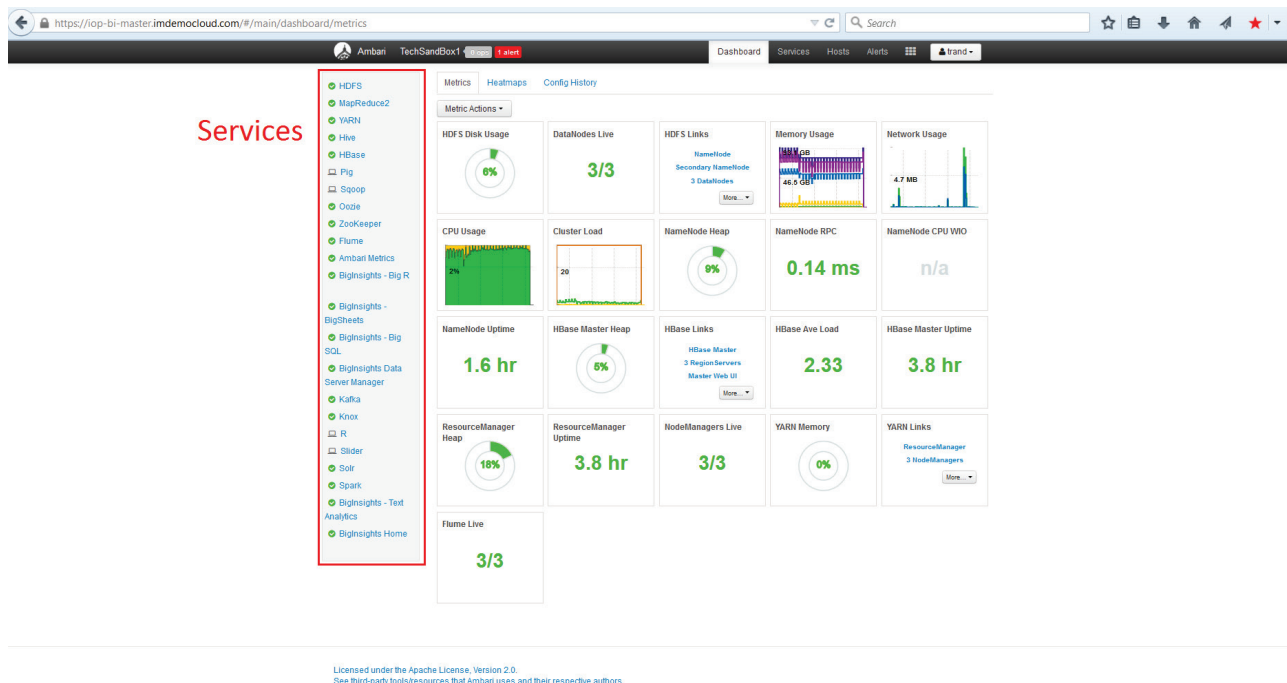
```
ssh trand@iop-bi-master.imdemocloud.com
```

## 1.2 Ambari Console

- After pressing the launch button under Ambari Console (seen in 1.1 Step 4), this tab will appear in your browser. Input your username and password (username seen in 1.1 Step 4, password is the same as the one used to sign up for the IBM Analytics Demo Cloud account).



- After logging in, the screen should look like this. Leave the Ambari Console on throughout the following labs. All of the services should be started.



## 1.3 SSH and PuTTY

**NOTE: A SSH client (e.g. PuTTY, etc.) is necessary to connect to the cloud environment. The suggested client that will be used in the following labs is PuTTY.**

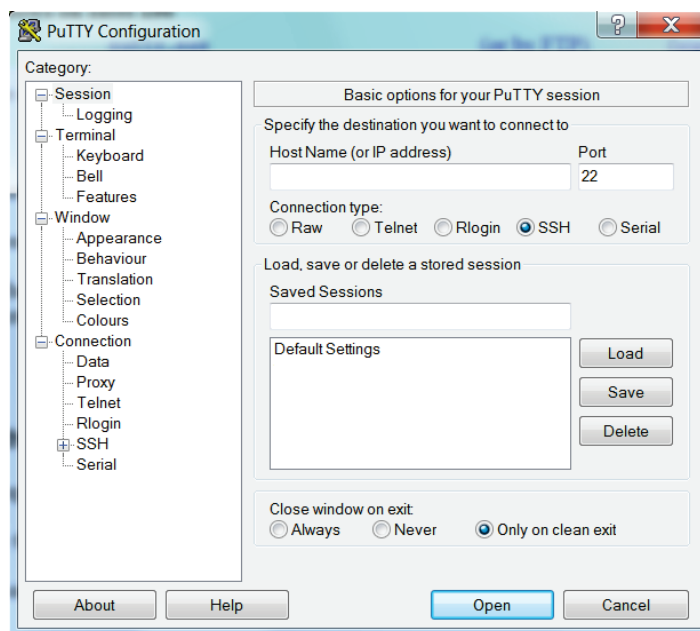
If you already have a SSH program or PuTTY installed, skip this section.

1. Navigate to [www.chiark.greenend.org.uk/~sgtatham/putty/download.html](http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html) and download PuTTY.

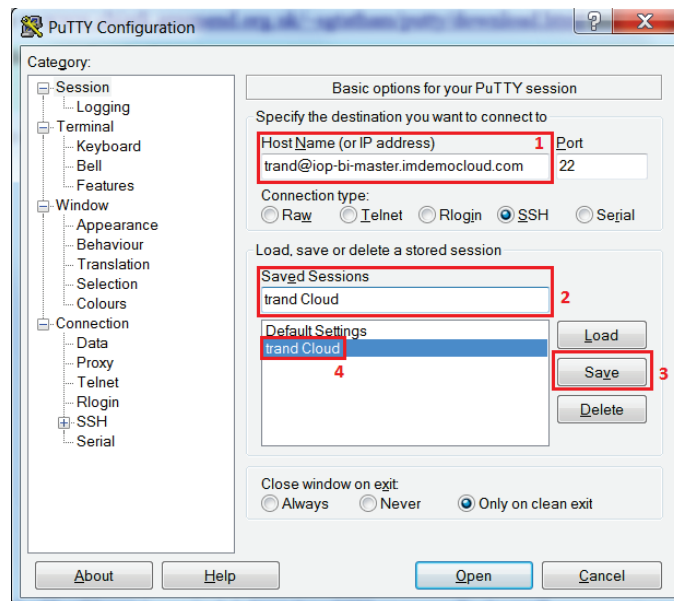
### For Windows on Intel x86

PuTTY: [putty.exe](#) [\(or by FTP\)](#) [\(signature\)](#)

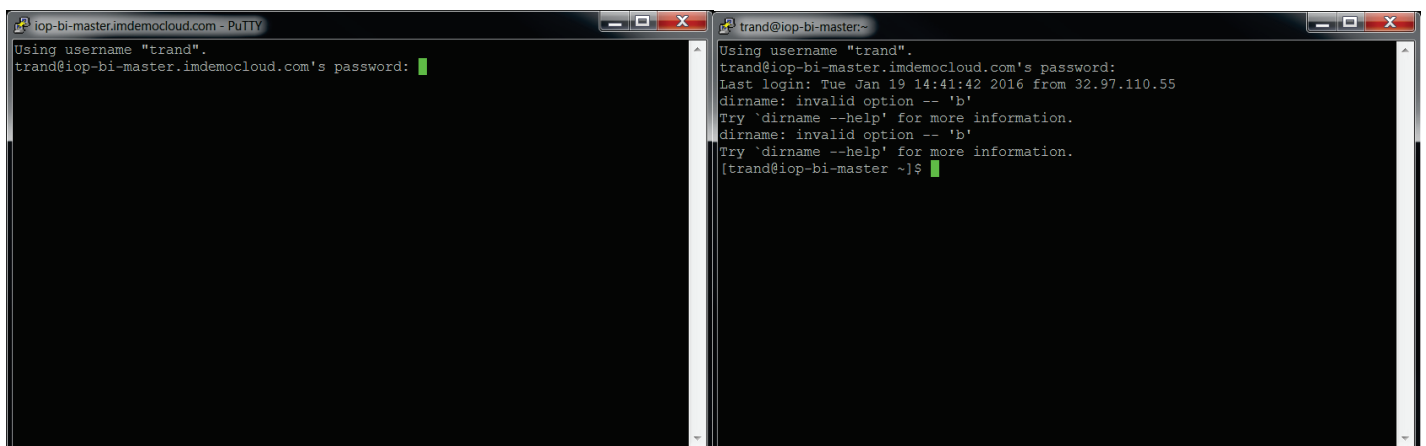
2. After the download, open PuTTY. It should look like this:



3. Type in the HostName (seen in 1.1 Step 4 as SSH command). As well, you can type in a name for the session and press the save button to store the session. Saving the session allows you to click the session next time you open PuTTY and load it without needing to retype the Host Name. Ensure the Connection type is SSH.



4. Click the session and press open. On the new screen, there is a prompt for your password. The password is the same as the one used to sign up/login to the IBM Analytics Demo Cloud website. Input the password and you should be able to run commands on the cloud!



## 1.4 FTP and WinSCP File Transfer

**NOTE:** A FTP tool (e.g. WinSCP, FileZilla, etc.) is necessary to transfer files. The suggested client that will be used in the following labs is WinSCP.

If you already have an FTP file transfer program or WinSCP, you can skip this step.

1. Navigate to <https://winscp.net/eng/download.php> and click on the Installation Package link to start downloading the program. Install and set up WinSCP.



## Download WinSCP

### WinSCP 5.7.6

**Installation package** (5.6 MB; 1,204,476 downloads to date)

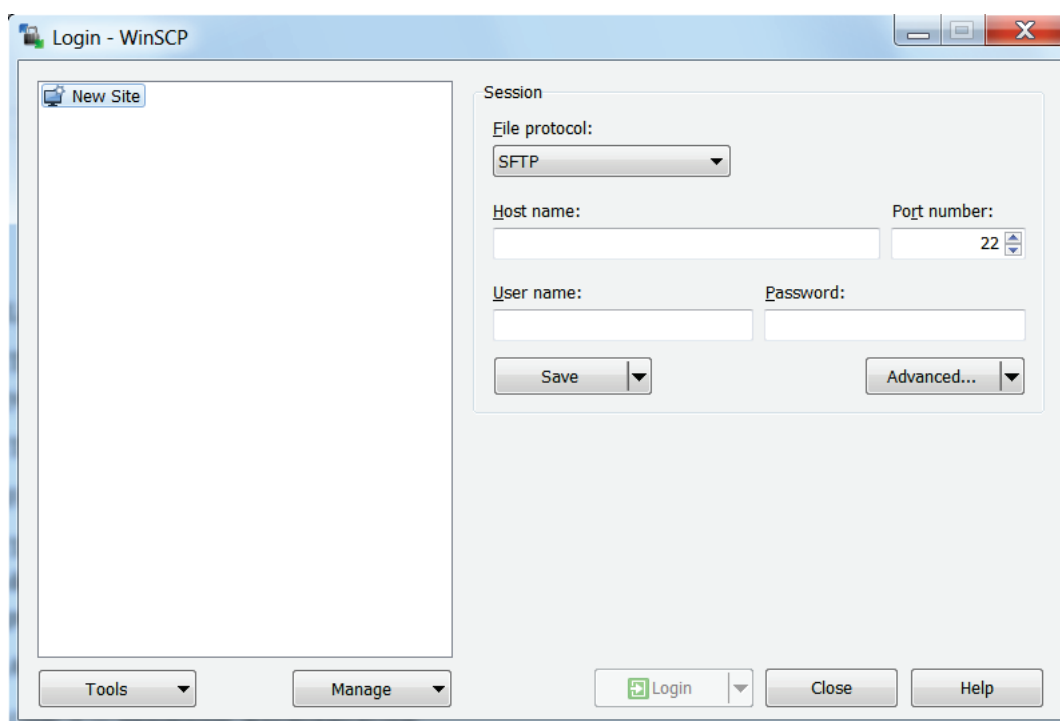
**Portable executables** (4.7 MB; 228,129 downloads to date)

**.NET assembly / COM library** (4.6 MB; 10,038 downloads to date)

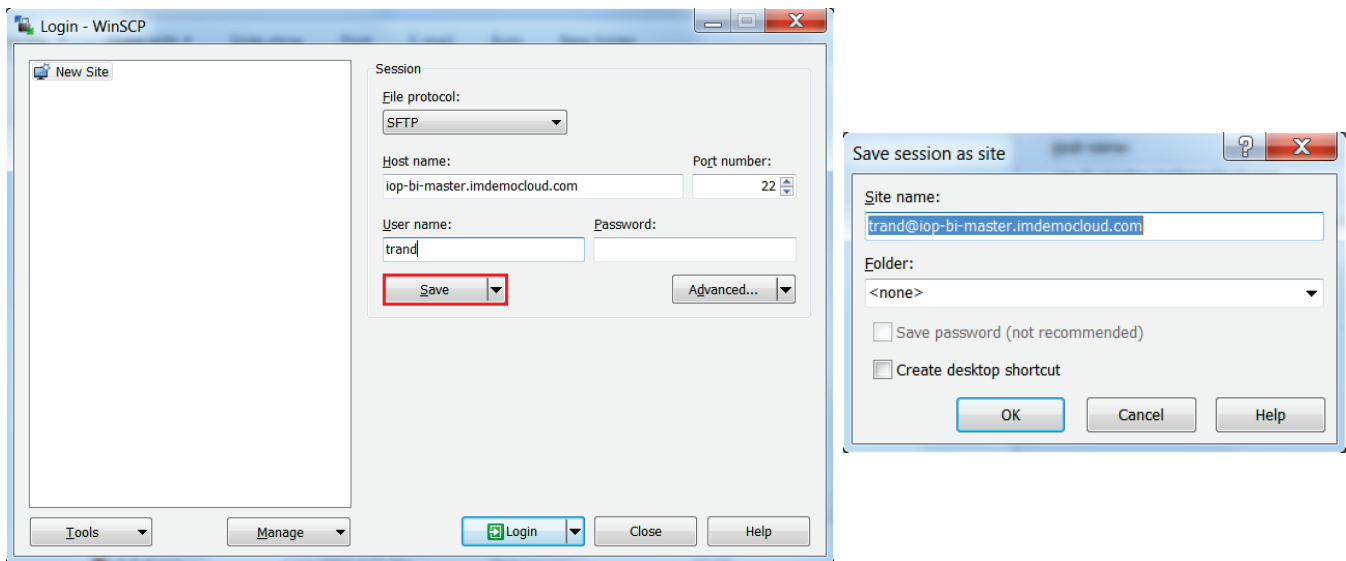
**Source code** (8.8 MB; 5,245 downloads to date)

[\[Release Notes, Checksums\]](#) [\[What's New\]](#) [\[Release Notifications\]](#)

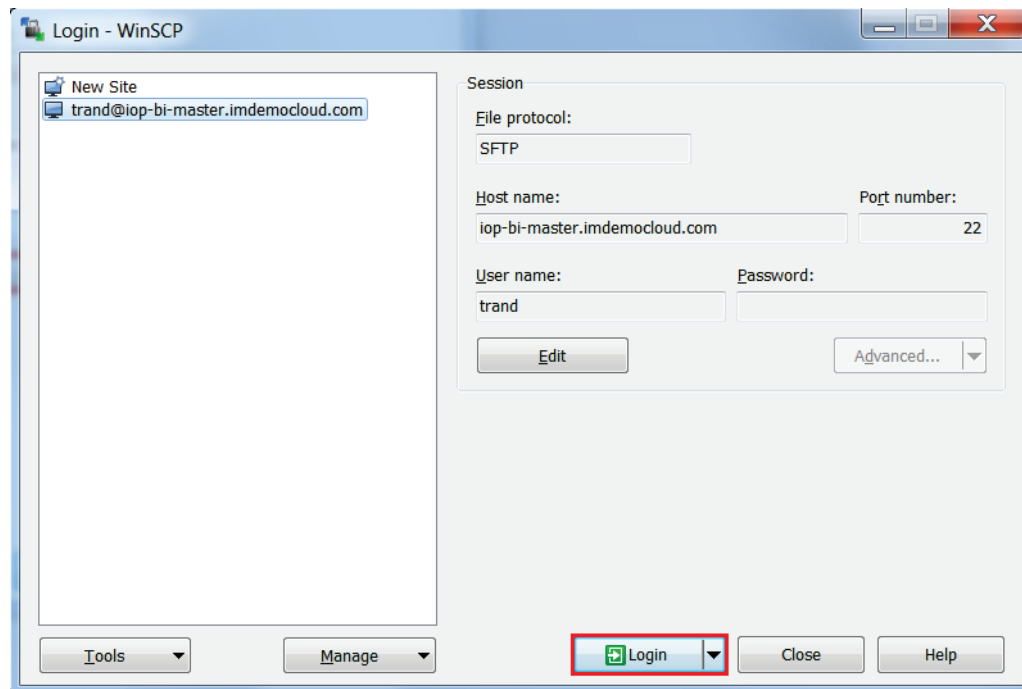
2. After finishing installation, we need to set up WinSCP for file transfer to the cloud. Open WinSCP.



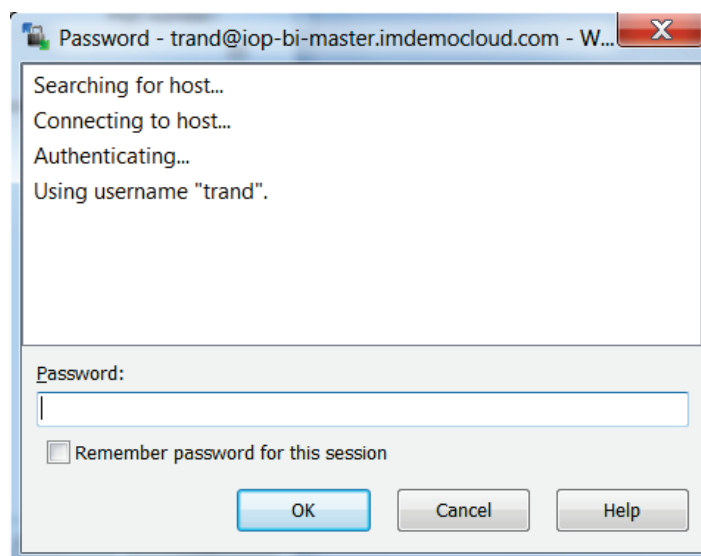
3. The Host Name is iop-bi-master.imdemocloud.com (seen in 1.1 Step 4 as Server Host Name). Set the User Name as seen in 1.1 Step 4. The Password section can be left blank, as a prompt for the password will always appear when trying to login. Press save, and it will prompt for a name for the session.



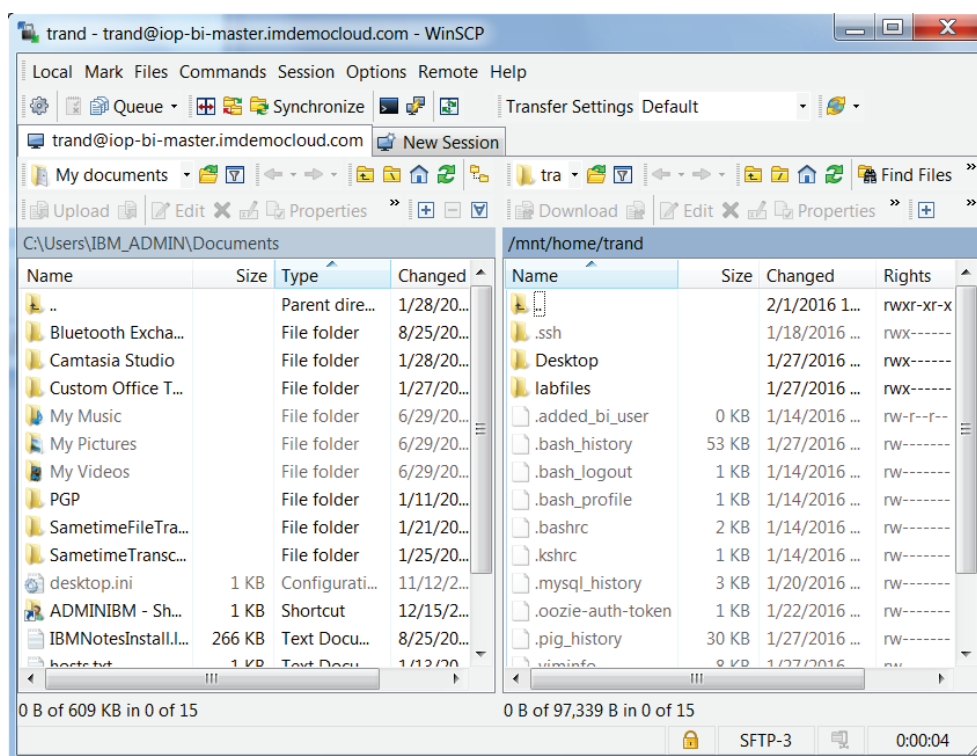
4. Afterwards, WinSCP should return to the start screen, but this time have the session you just created. Select that session and press Login.



5. Input your password into the prompt (same password used to create your account).



6. After Login, the screen should look like this. The left-hand side is your computers' directories, while the right-hand side is the Clouds' directory. Files can be transferred by dragging and dropping them from the left-hand side to the right-hand side and vice-versa.



## 1.5 Summary

Congratulations! You have completed the lab setup, you are now ready to start the labs.

You may move on to the next unit.

## NOTES

[illegible]

## NOTES

[illegible]



---

© Copyright IBM Corporation 2016.

The information contained in these materials is provided for informational purposes only, and is provided AS IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, these materials. Nothing contained in these materials is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software. References in these materials to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates. This information is based on current IBM product plans and strategy, which are subject to change by IBM without notice. Product release dates and/or capabilities referenced in these materials may change at any time at IBM's sole discretion based on market opportunities or other factors, and are not intended to be a commitment to future product or feature availability in any way.

IBM, the IBM logo and [ibm.com](http://ibm.com) are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).



Please Recycle

---