Using Windows Azure Virtual Machines

Overview

In this hands-on lab you will create a Windows Virtual Machine. You can install software you need on that virtual machine and use it as a workstation environment. Windows Azure also provides VMDepot where you can find a lot of community images which you can use directly. You will also add additional disks to your machine for your storage and copy files from your laptop/workstation to your VM machines on Windows Azure.

Objectives

In this hands-on lab, you will learn how to:

- Create a virtual machine with Windows Server 2012 R2 and run R job.
- Create a virtual machine from VMDepot and run ipython job.
- Create a new disk and mount the disk to Windows and Linux VMs.

Prerequisites

The following is required to complete this hands-on lab:

- A Windows Azure subscription - sign up for a free trial
- If you are a Mac Computer user, please install Windows Remote desktop 8.0 from the App Store.
- You **must** use one of the following **browsers**: latest version of **Firefox or Chrome, IE 9, 10, 11.** Browsers like Safari, 360 may have issues with IPython or RDP download.

Exercises

This hands-on lab includes the following exercises:

1. Exercise 1: Create a machine with Windows Server 2012 R2 and run R job.
2. Exercise 2: Create a machine from VMDepot and run ipython job.
3. Exercise 3: Create a new disk and mount the disk to Windows and Linux.

Estimated time to complete this lab: **60** minutes.

Exercise 1: Create a machine with Windows Server 2012 R2 and run R job.

1. You need to login on [Windows Azure Management Portal] (http://manage.windowsazure.com) to get start.

Sign in to Windows Azure

Don't have a Microsoft account? Sign up now ›

Sign in with your organizational account ›

Sign in

Microsoft account What's this?
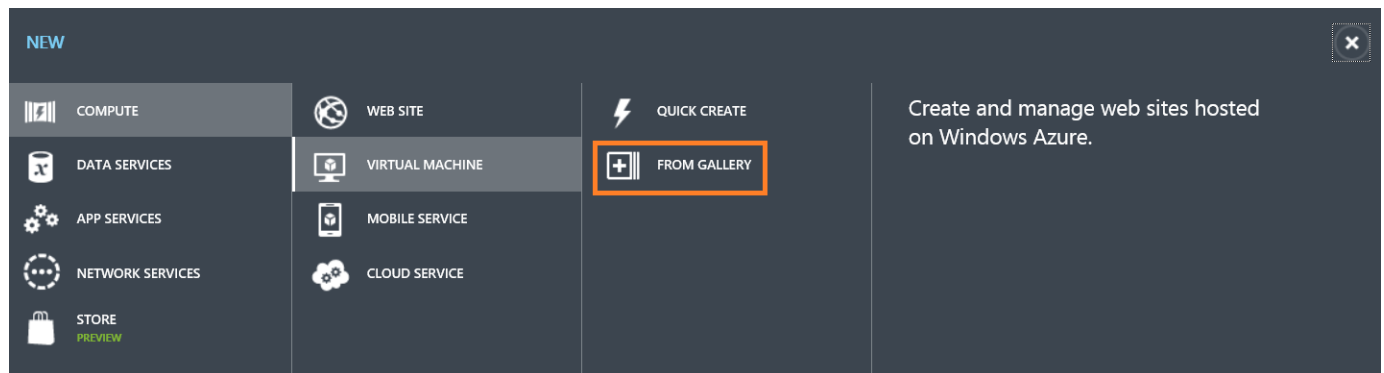
someone@example.com

Password

☐ Keep me signed in

Sign in

Can't access your account?

Sign in with a single-use code

Windows Azure Management Portal

2. On the main page, click **New** -> **Compute** -> **Virtual Machine** -> **From Gallery** to create a new Virtual Machine on Windows Azure.

NEW

COMPUTE	WEB SITE	QUICK CREATE	Create and manage web sites hosted on Windows Azure.
DATA SERVICES	VIRTUAL MACHINE	FROM GALLERY	
APP SERVICES	MOBILE SERVICE		
NETWORK SERVICES	CLOUD SERVICE		
STORE PREVIEW			

Create New Virtual Machine From Gallery

3. On the next page, you will see many different images. Click on **Windows Server 2012 R2** then click on the **Next** arrow at the lower right corner. Please do not use the Windows Server 2012 image.

创建虚拟机

选择映像

| 全部 |
| MICROSOFT |
| WINDOWS SERVER |
| SHAREPOINT |
| SQL SERVER |
| BIZTALK SERVER |
| UBUNTU |
| CENTOS |
| SUSE |
| ORACLE |

我的映像
我的磁盘

☐ MSDN ❓
☐ Supported

特色 ⌄

Windows Server 2012 Datacenter

Windows Server 2012 R2 Datacenter

Windows Server Essentials Experience
Windows Server 2012 R2

Windows Server 2008 R2 SP1

SharePoint Server 2013 Trial

SQL Server 2014 CTP2 Evaluation for Data
Windows Server 2012

SQL Server 2014 CTP2 Evaluation
Windows Server 2012

Windows Server 2012 R2 Datacenter

At the heart of the Microsoft Cloud OS vision, Windows Server 2012 R2 brings Microsoft's experience delivering global-scale cloud services into your infrastructure. It offers enterprise-class performance, flexibility for your applications and excellent economics for your datacenter and hybrid cloud environment.

发布服务器	Microsoft Windows Server Group
操作系统系列	Windows
位置	East Asia;Southeast Asia;North Europe;West Europe;Central US;East US;East US 2;West US

定价信息
定价根据您选择用于设置虚拟机的订阅而异。

→

Select Windows Server 2012 R2 Datacenter

4. Select the **Version Release Date** and **Size** first, then set the **Virtual Machine Name**, **New User Name** and **Password**, then click **Next**. It is really **important** that you **write down** the password you have chosen here.

Virtual machine configuration

VERSION RELEASE DATE (?)

| 9/16/2013 | ∨ |

VIRTUAL MACHINE NAME (?)

[▓▓▓▓▓▓▓▓]

SIZE

| Small (1 core, 1.75 GB memory) | ∨ |

NEW USER NAME

azureuser

NEW PASSWORD   **CONFIRM**

| ●●●●●●●● ✓ | | ●●●●●●● |

Windows Server 2012...

Windows Server 2012 incorporates
Microsoft's experience building and
operating public clouds, resulting in a
dynamic, highly available server platform. It
offers a scalable, dynamic and multi-tenant-
aware infrastructure that helps securely
connect across premises.

PUBLISHER	Microsoft Windows Server Group
OS FAMILY	Windows
LOCATION	East Asia;Southeast Asia;North Europe;West Europe;East US;West US

PRICING INFORMATION

Pricing varies based on the subscription you
select to provision your virtual machine.

← →

Set Machine Information

5. Next, we need to setup the cloud service information. Each virtual machine belongs to one cloud service. You can create a new cloud service or add the virtual machine to an existing cloud service. Here we select *Create a new cloud service* and set the **Cloud Service DNS Name** and set the **Region** to the region closest to where you are. As for the storage account, you can either choose *Use an automatically generated storage account* or select an existing storage account.

> Note: Since the training subscription only has 1 storage account, so make sure you select the existing one.

Virtual machine configuration

CLOUD SERVICE ⑦

Create a new cloud service

CLOUD SERVICE DNS NAME

████████████ ✓ .cloudapp.net

REGION/AFFINITY GROUP/VIRTUAL NETWORK ⑦

East Asia

STORAGE ACCOUNT

████████████

AVAILABILITY SET ⑦

(None)

Windows Server 2012...

Windows Server 2012 incorporates Microsoft's experience building and operating public clouds, resulting in a dynamic, highly available server platform. It offers a scalable, dynamic and multi-tenant-aware infrastructure that helps securely connect across premises.

PUBLISHER	Microsoft Windows Server Group
OS FAMILY	Windows
LOCATION	East Asia;Southeast Asia;North Europe;West Europe;East US;West US

PRICING INFORMATION

Pricing varies based on the subscription you select to provision your virtual machine.

← →

Set Cloud Service Information

6. The next page is to setup the virtual machine's endpoint. The TCP Remote desktop 3389 port and PowerShell 5986 port are added by default. You can add additional ports as needed.

Virtual machine configuration

ENDPOINTS ⑦

NAME	PROTOCOL	PUBLIC PORT	PRIVATE PORT
Remote Desktop	TCP	AUTO	3389
PowerShell	TCP	5986	5986
ENTER OR SELECT A VALUE ⌄			

Windows Server 2012...

Windows Server 2012 incorporates Microsoft's experience building and operating public clouds, resulting in a dynamic, highly available server platform. It offers a scalable, dynamic and multi-tenant-aware infrastructure that helps securely connect across premises.

PUBLISHER   Microsoft Windows Server Group

OS FAMILY   Windows

LOCATION   East Asia;Southeast Asia;North Europe;West Europe;East US;West US

PRICING INFORMATION
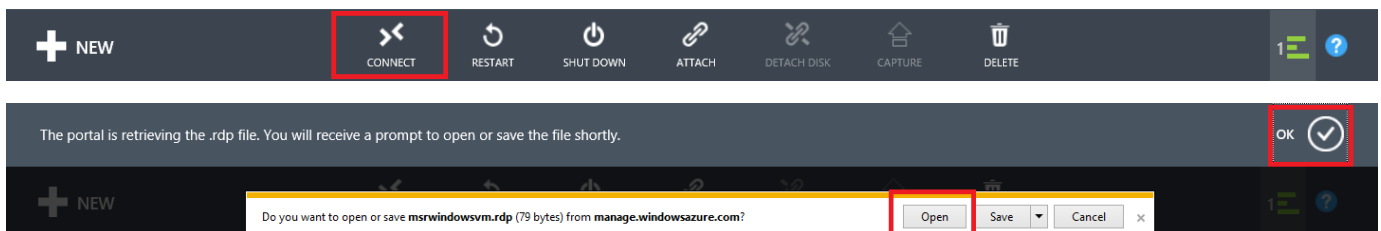Pricing varies based on the subscription you select to provision your virtual machine.

Set Endpoints

7. After clicking the **Finish** button on the right corner, Windows Azure will create the windows virtual machine for you. Wait a few minutes for the provisioning process, we can access that machine via *Remote Desktop*.

← Successfully created virtual machine msrwindowsvm.   OK ✓

✓ Successfully created virtual machine msrwindowsvm.
✓ Successfully provisioned the virtual machine msrwindowsvm.

Successfully Create Windows Virtual Machine

8. Click the **Connect** button, you will download an rdp file. Open the rdp file on your windows machine and connect to the VM we just created. If you are using a **Mac computer**, please make sure you have Windows Remote Desktop Version 8 or higher installed from the App Store. Please also note that you should be using IE 9, Chrome, or firefox to download the RDP session correctly.

| ➕ NEW | ✕ CONNECT | ↺ RESTART | ⏻ SHUT DOWN | 🔗 ATTACH | ✂ DETACH DISK | ⬆ CAPTURE | 🗑 DELETE | 1 ❓ |

The portal is retrieving the .rdp file. You will receive a prompt to open or save the file shortly.   OK ✓

➕ NEW

Do you want to open or save **msrwindowsvm.rdp** (79 bytes) from **manage.windowsazure.com**?   Open   Save ⌄   Cancel   ✕

Connect Windows Virtual Machine

9. Use the *username* and *password* we set in step 4 and log into the Windows virtual machine remotely.
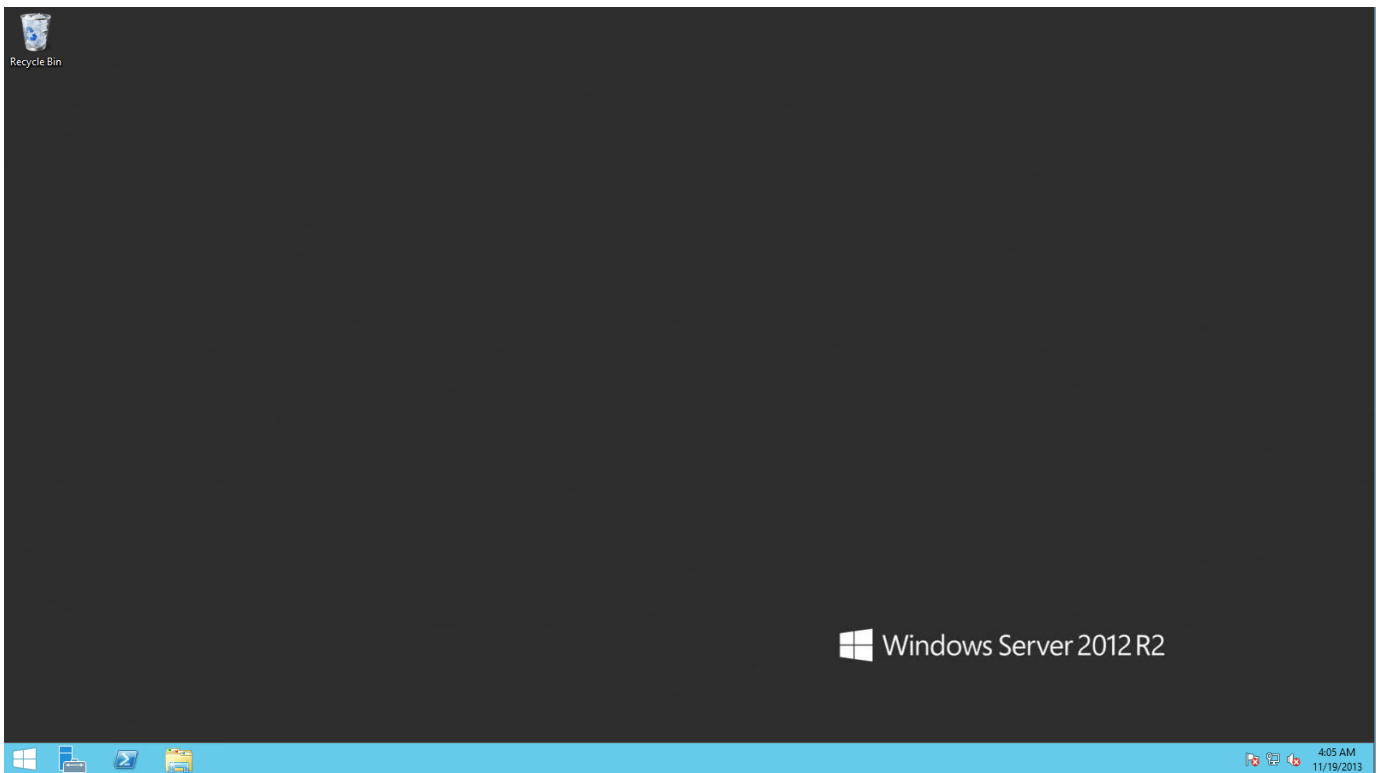
Remote Desktop Window

10. Next, let's install R and R Studio on the Windows Machine and run a few examples. First, open **Internet Explorer** on that machine and download the R install package from R Website. Click the CRAN mirror link from the Getting Started area, then select the nearest mirror to start your download of R for Windows. Please note that the download is from the internet to your VM in the cloud, you will notice that it is much faster than downloading files to your local machine. VMs sitting in Windows Azure typically gives you a much faster interconnect to the rest of the internet than any local machines could achieve. This could potentially speedup your research work.
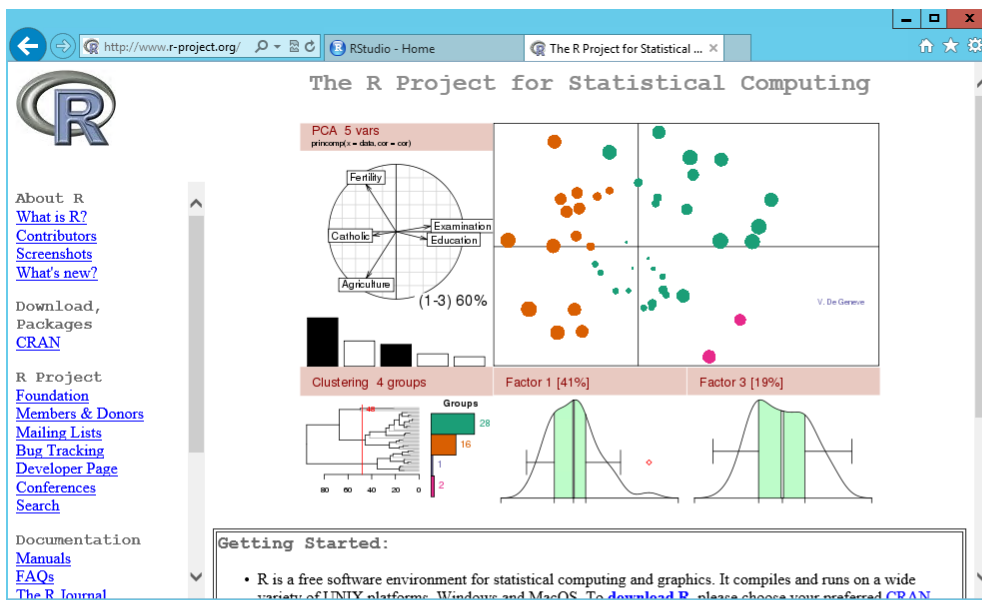
Download R

Note: You might need to disable the **Internet Explorer Enhanced Security Configuration** to avoid adding every website into your security list. In order to disable IES, you can open the **Server Manager**, click **Local Server** -> **IE Enhanced Security Configuration** and set it to Off.
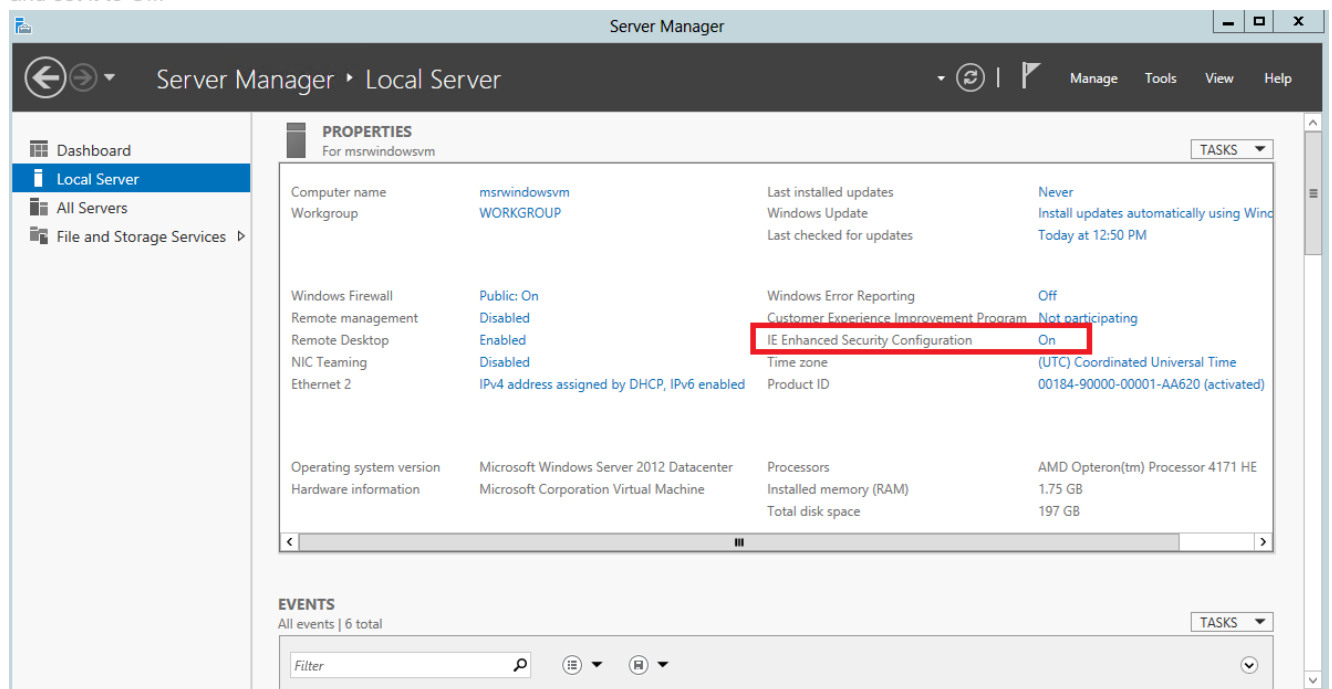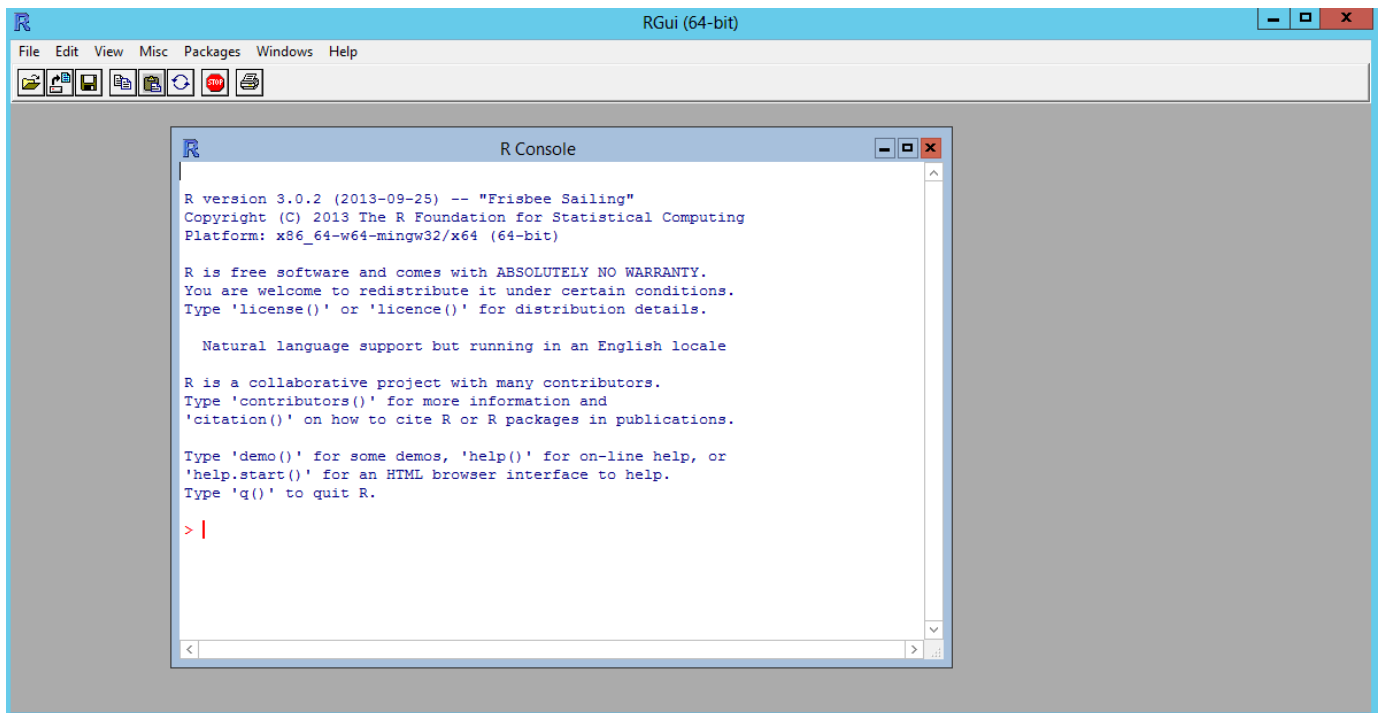

Turn IE Enhanced Security Off

11. Launch/Run the download R-3.0.2-win.exe and install R. After the installation completes, run RGui which presents an R Console window.
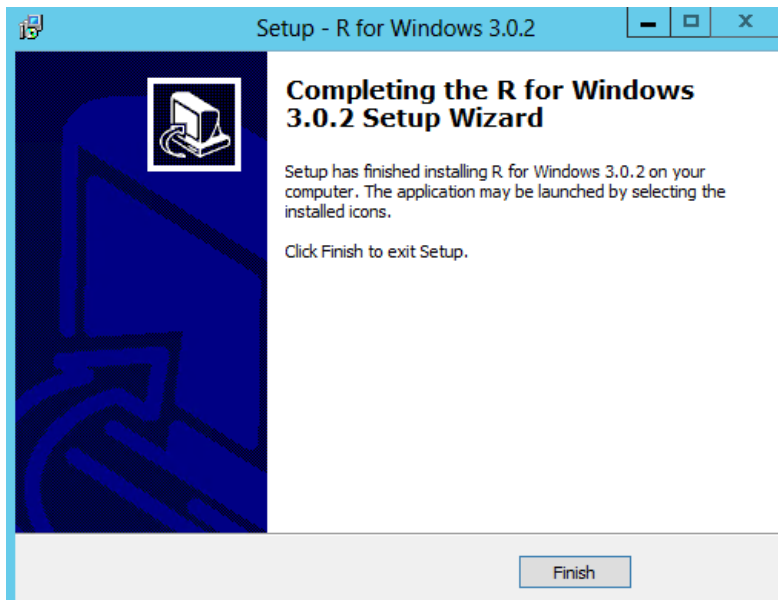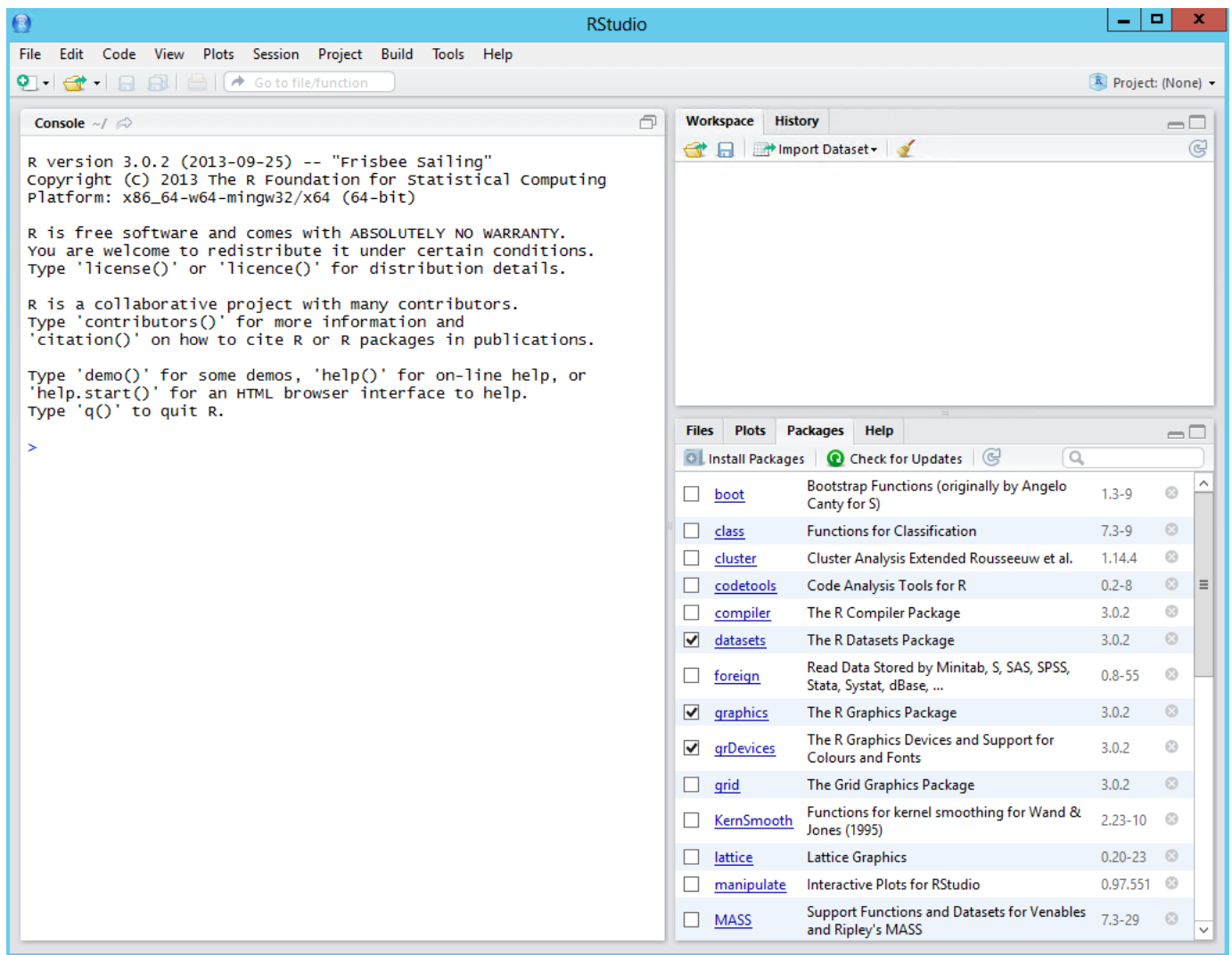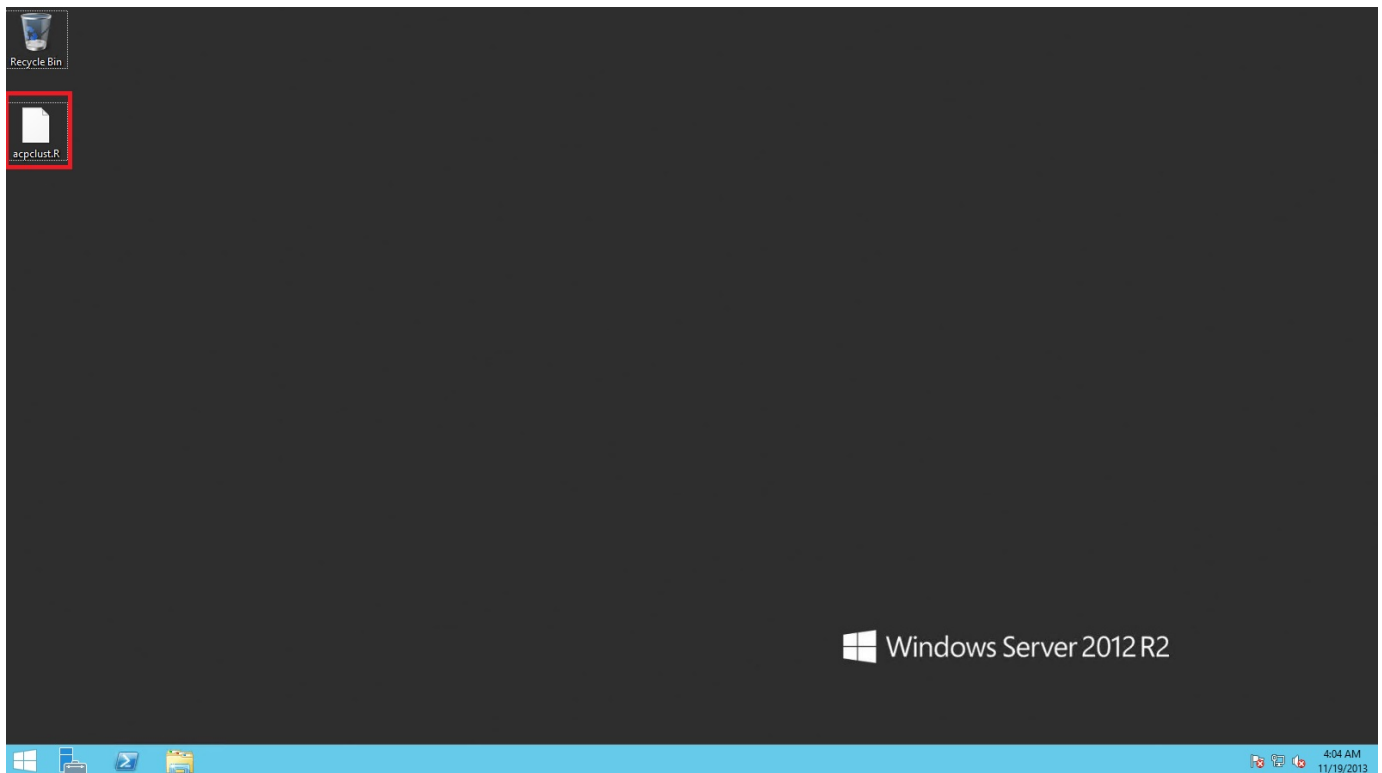
Install R

Launch R

12. Use the same steps to install R Studio and run it. After installation, you can find R Studio by clicking on the Start Menu and then type in: rstudio to find the program. Press return to launch R Studio.

Launch RStudio

13. Next, let's execute an R job. First we need to move an R file on the remote machine on Windows Azure. You can find an file **acpclust.R** under the folder **Source\Exercise1**. Since Windows Remote Desktop supports Copy/Paste from a local machine to a remote desktop machine, you can use this to get the file onto the Windows Azure machine. To do this, right-click the file **acpclust.R**, click **Copy**, then navigate to the Remote Desktop window's desktop and right-click the desktop, click **Paste**. After a few seconds, you will find the file is copied to the remote machine's desktop.

> Note: If you are using a **Mac Computer**, or Linux machine, please create a new text file on the remote machine and paste the content text of the file into notepad or equivalent. You may have to rename the file from acpclust.R.txt to acpclust.R. Simply open a command prompt (start menu, type cmd.exe, enter) and run: rename acpclust.R.txt acpclust.R

Copy File to Remote Desktop

14. To run the **acpclust.R** file, we have to install some additional packages of R. In the RStudio's right corner panel, click **Packages** tab, and click **Install Packages** button.
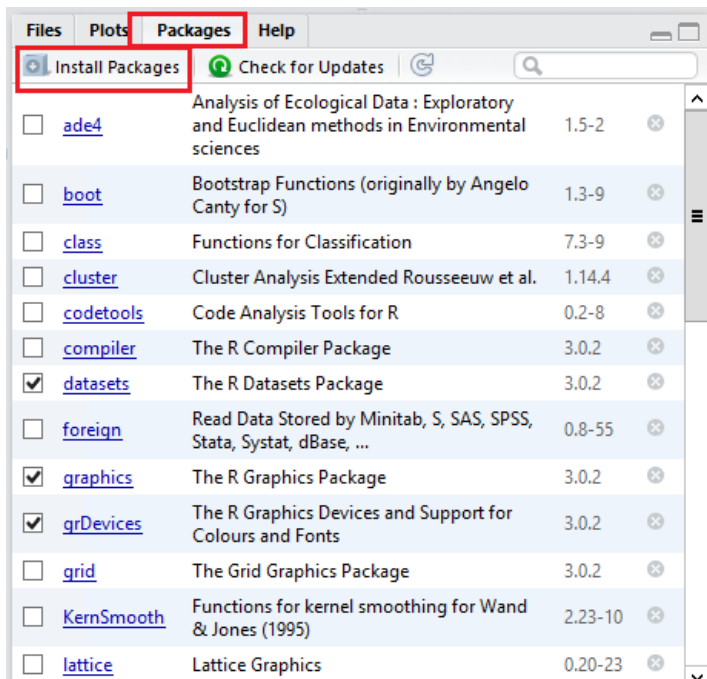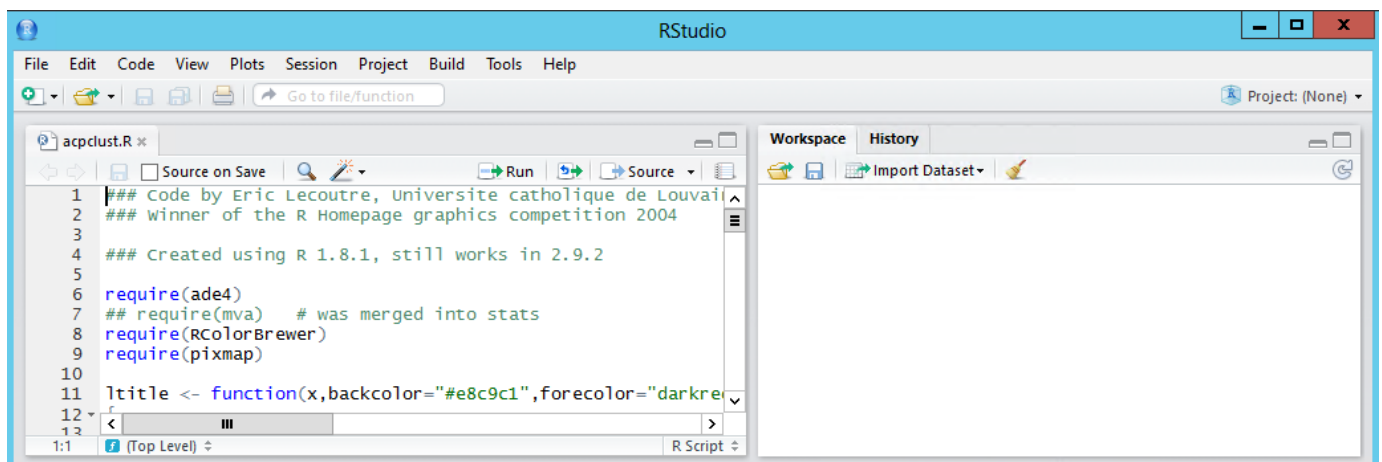

Install R Packages

15. In the **Install Packages** form, input the following command into the Packages line:

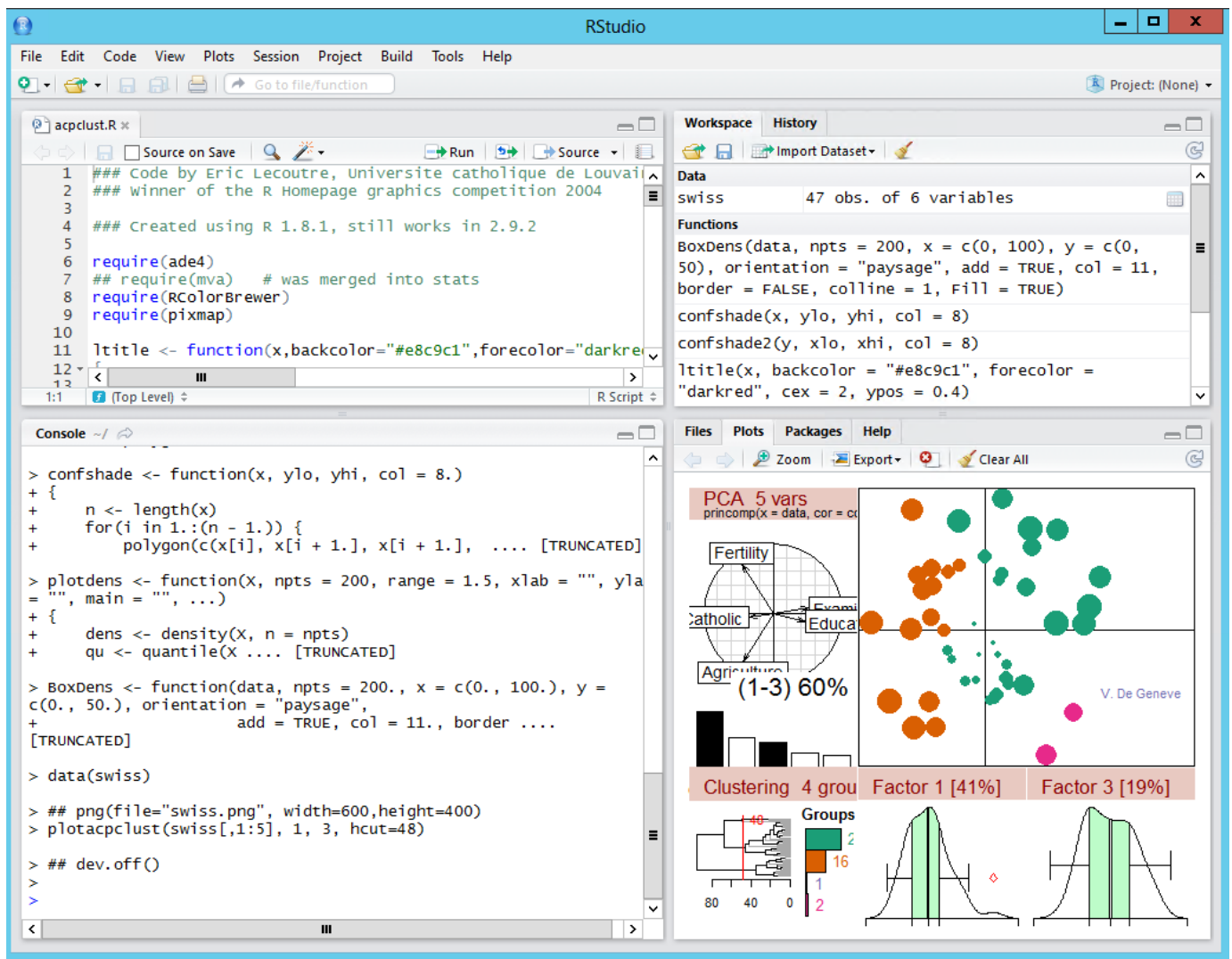`ade4, RColorBrewer pixmap` click **Install** to begin download and installation.
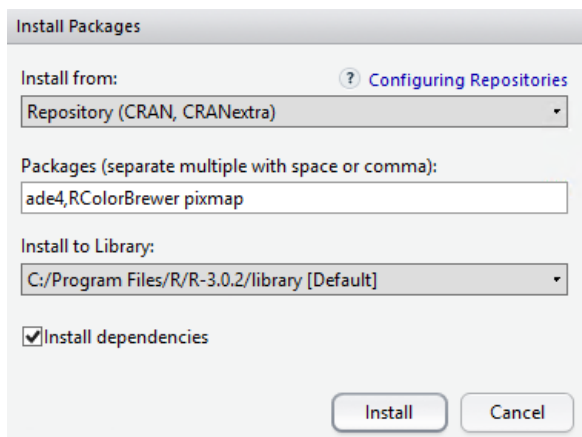
Set R Packages

16. Click Menu **File** -> **Open File** and select the **acpclust.R** file on the desktop to open the file in RStudio.

Open an R File

17. Click the menu **Code** -> **Run Region** -> **Run All**. RStudio will execute the R code and generate the result on the right bottom corner.

Run R File

In this exercise, we created a Windows Server 2012 R2 virtual machine (VM) on Windows Azure, installed R and R studio on that machine. Then we copied a local file to the remote machine used that file as input for running an R job on our VM.

Exercise 2: Create a machine from VM Depot and run IPython examples.

VM Depot is a community-driven catalog of user contributed images containing operating systems, applications, and development environments captured in a virtual disk image that can easily be deployed on Windows Azure. In this exercise, we will create another virtual machine, but this time we will use a Linux image from VM Depot. We will also run some IPython examples on that machine.

1. First we need to copy a community image from VM Depot. In the Windows Azure Management Portal, navigate to **Virtual Machines**, and click **Image** tab. Then you can click **Browse VM Depot**. Please note that VM depot is accessible via the **Images** Tab only.

Virtual Machine Image
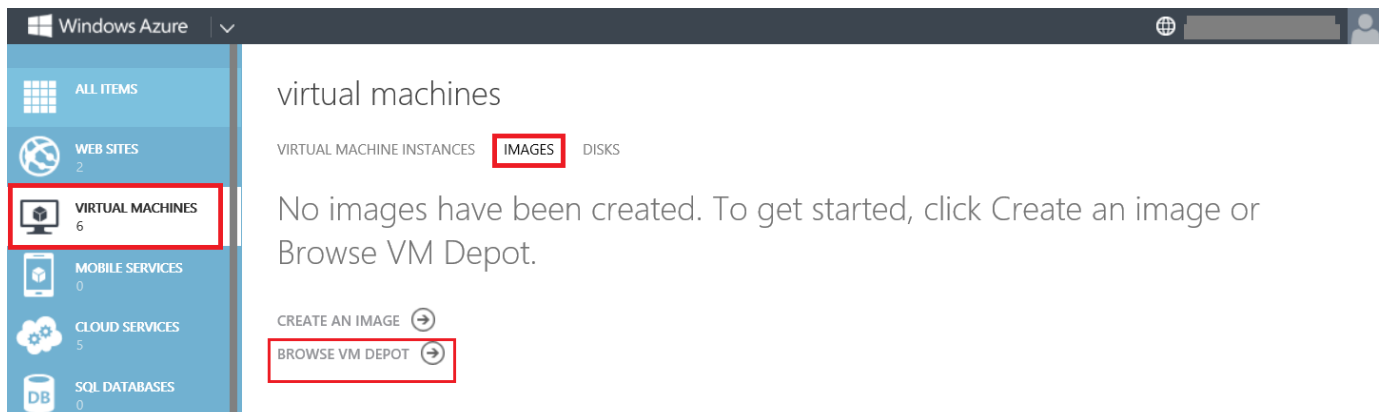
2. On the Browse VM Depot page, click **Ubuntu**, then scroll down to select **Azure Data Analysis** image, then click Next.

Azure Data Analysis Image

> Notes:If you wouldl like to understand more information about the image, we can click **More** link on the right and check more details of the image.

← **Back**

▶ DEPLOYMENT SCRIPT

⬆ PUBLISH

⬘ SHARE

? HELP

📇 Browse Images

👤 My Account

Azure Data Analysis

★★★★★

0 comments

Includes R, openmpi and loads of packages. Include Spark 0.8, Shark 0.7 and Storm and Kafka. Also includes iPython notebook + a whole heap of libraries.

Image name	Azure Data Analysis
Publisher	elastacloud
Platform	ubuntu
Packages	azure sdk git hive ipython ipython-notebook kafka nodejs npm numpy openjdk openjre pandas python scala scikit-learn scipy shark spark storm tornado
Tags	azure sdk git hive ipython ipython-notebook kafka nodejs npm numpy openjdk openjre pandas python scala scikit-learn scipy shark spark storm tornado
Region(s)	All
EndPoints	Public Port / Local Port: 443 / 8888 80 / 80 4040 / 4040

Azure Data Analysis Image Details

3. In the **Choose a storage account** page, we can set the **Image Region** and **Storage Account in image region**. You can also create a new storage account or choose an existing one. You should make sure you store it in the same region that you will be creating your virtual machines.

Choose a storage account

Select a storage account to copy the image to

IMAGE REGION

East Asia ▾ ❓

STORAGE ACCOUNT IN IMAGE REGION

▰▰▰▰▰ ▾ ❓

ATTENTION

Microsoft is not responsible for images provided & licensed by the community members and does not screen for security, compatibility or performance. Users should perform their own due diligence before using the community images. Support for images is provided by the community and not by Windows Azure customer support. By using unsupported images, you might forfeit Windows Azure availability SLA.

1

Alfresco 4.2.c-2 (Ubu...
★★★★★ (0)

Alfresco powered by BitNami is a pre-configured, ready to run image for running Alfresco on Windows Azure. Alfresco is an Enterprise Content Management (ECM) sy...
More

PUBLISHER	BitNami
PUBLISH DATE	3/25/2013
PACKAGES	alfresco, java, jdbc, mysql, tomcat
SIZE	30 GB

← ✓

Set Image Storage Account

4. Then Windows Azure will begin to copy the image from VM Deport to your storage account. It may take 15-30 minutes to finish. Once it completes, you will see the image is inside the image tab.

Copy Image from VM Depot

5. You will also need to register the image before you can create a new virtual machine from it. Click the Register button on the bottom and set the **Name** and **VHD URL** for the image.

Register Image

The image registration process is fast, but make sure to wait until it completes before starting the next step.

1. Next, we will create a new virtual machine from the image we just copied from VM Depot and registered. Click **New** -> **Compute** -> **Virtual Machine** -> **From Gallery** to locate our new image. This time we choose **My Image** and you will see an image called *Azure-Data-Analysis*

there.

Virtual machine image selection

ALL		Azure-Data-Analysis	**Azure-Data-Analysis**
PLATFORM IMAGES			
MY IMAGES			Includes R, openmpi and loads of packages. Include Spark 0.8, Shark 0.7 and Storm and Kafka. Also includes iPython notebook + a whole heap of libraries. More
MY DISKS			

PUBLISHER   elastacloud (VM Depot)

OS FAMILY   Linux

SUBSCRIPTION   Azpas300A0U8655

LOCATION   East Asia

PRICING INFORMATION
Pricing varies based on the subscription you
select to provision your virtual machine.

→

Create VM from My Images

2. Select the image and click *Next*, we will go through similar steps as in Exercise 1 to setup machine's information. Since it is a Linux machine, the information we need to enter is a little different. On the first page, we set **Virtual Machine Name**, **Size**, **User Name** just the same. We can either upload an SSH key or just provide a password for the user. Again, **write down** your password before proceeding.

Virtual machine configuration

VIRTUAL MACHINE NAME ❓

    msrlinuxvm

SIZE

    Small (1 core, 1.75 GB memory) ▾

NEW USER NAME

    azureuser

AUTHENTICATION ❓

☐ UPLOAD COMPATIBLE SSH KEY FOR AUTHENTICATION

☑ PROVIDE A PASSWORD

NEW PASSWORD   **CONFIRM**

    •••••••• ✓   ••••••••

Azure-Data-Analysis

Includes R, openmpi and loads of packages.
Include Spark 0.8, Shark 0.7 and Storm and
Kafka. Also includes iPython notebook + a
whole heap of libraries. More

PUBLISHER	elastacloud (VM Depot)
OS FAMILY	Linux
SUBSCRIPTION	Azpas300A0U8655
LOCATION	East Asia

PRICING INFORMATION
Pricing varies based on the subscription you
select to provision your virtual machine.

← →

Set Linux Machine Information

3. Now, setup the cloud service information, as a reminder a cloud service is simply a container for your VMs and their settings.

Virtual machine configuration

CLOUD SERVICE ❓

Create a new cloud service ▾

CLOUD SERVICE DNS NAME

▒▒▒▒▒▒▒▒▒ ✅ .cloudapp.net

REGION/AFFINITY GROUP/VIRTUAL NETWORK ❓

East Asia ▾

AVAILABILITY SET ❓

(None) ▾

Azure-Data-Analysis

Includes R, openmpi and loads of packages. Include Spark 0.8, Shark 0.7 and Storm and Kafka. Also includes iPython notebook + a whole heap of libraries. More

PUBLISHER	elastacloud (VM Depot)
OS FAMILY	Linux
SUBSCRIPTION	Azpas300A0U8655
LOCATION	East Asia

PRICING INFORMATION
Pricing varies based on the subscription you select to provision your virtual machine.

← →

Set Linux Cloud Service Information

4. We will also need to setup the endpoints information. According to the image description, we will know that the image's publisher, platform, packages and the required endpoints. Add public port 443 mapped from the VM's private port 8888 and then do a same port mapping for port 22 (ssh), 80 (web), and 4040 (SHARK).

Virtual machine configuration

ENDPOINTS ❓

NAME	PROTOCOL	PUBLIC PORT	PRIVATE PORT
SSH	TCP	22	22
HTTPS	TCP	443	8888
HTTP	TCP	80	80
Other	TCP	4040	4040
ENTER OR SELECT A VALUE ⌄			

Azure-Data-Analysis

elastacloud

Includes R, openmpi and loads of packages. Include Spark 0.8, Shark 0.7 and Storm and Kafka. Also includes iPython notebook + a whole heap of libraries. More

PUBLISHER	elastacloud (VM Depot)
OS FAMILY	Linux
SUBSCRIPTION	Azpas300A0U8655
LOCATION	East Asia

LEGAL TERMS

By clicking the Submit button, I acknowledge that I am getting this software from elastacloud and that elastacloud's legal terms apply to it. Microsoft does not provide rights for third-party software.

NOTE

This is an unsupported image supplied by a third party. ❓

PRICING INFORMATION

Pricing varies based on the subscription you select to provision your virtual machine.

Configure Linux Endpoints

5. After several minutes, the Linux virtual machine will be created too. This time we will use PuTTY, a SSH client to connect the that machine. Input the **DNS Name** of the linux machine into the Host Name, click **OK** to connect.

6. Input the username and password, then you will see the welcome screen from that machine upon successful log in.

Login

The machine has already installed a lot of useful packages including IPython, STORM or SPARK/SHARK. We will only use the IPython Notebook on that machine to run a very simple algorithm. First we need to configure IPython as a notebook configuration.

Configure IPython Notebook

1. Now you can start IPython Notebook. You should see that the server has started:

```
sudo ipython notebook --profile=nbserver
```


Start IPython Notebook

2. Navigate your browser (from your own machine, not the VM) to https://\<vm-name>.cloudapp.net. Make sure that you use https and not http. You will see a warning that the certificate is not signed. Since this is your own certificate, you can safely ignore this warning. After ignoring it, you should see the login screen:

IPython Notebook Main Page

Note: If you have trouble reaching the IPython Notebook URL, check the following: - Ensure you typed the *sudo ipython notebook --profile=nbserver* command correctly - Ensure you are using https (not http) - Ensure you are accessing the correct URL (double-check the cloud service name from the Windows Azure Portal)

3. Login with the password. The default password is **Elastacloud123**.

IPython Notebook Main Page

4. **Optional: Change Password** If you want to configure your own password, you can follow the following steps. Execute the following command:

```
python -c "import IPython;print IPython.lib.passwd()"
```


Create IPython Password

5. **Optional Change Password continued:** Then we use nano to edit the configuration and update *c.NotebookApp.password*. Since the file is read only, we can still modify it if we run *nano* through *sudo*. With ctrl+X, you can save the file. You have to press "Y" to confirm the save operation and then press enter.

```
sudo nano /usr/.ipython/profile_nbserver/ipython_notebook_config.py
```

Modify Password Configuration

# Clustering Example with Pandas and Scikit-learn

The following example (cluster-titanic.py) clusters passengers of the titanic based on several attributes. It is taken from www.kaggle.com and has been pre-processed with Excel and then exported in CSV format.

The script will result in an output as:

```
        labelled_data_x[survived[i]][label].append(scaled_coordinates[i][0])
        labelled_data_y[survived[i]][label].append(scaled_coordinates[i][1])


#######################################
# PLOTTING
colors = cm.rainbow(np.linspace(0, 1, NUM_CLUSTERS))
markers = ['x', '^']
for i in kmeans_labels_unique:
    for j in [0, 1]:
        plt.scatter(labelled_data_x[j][i], labelled_data_y[j][i], color=colors[i], marker=markers[j], s=40)

plt.show()
```
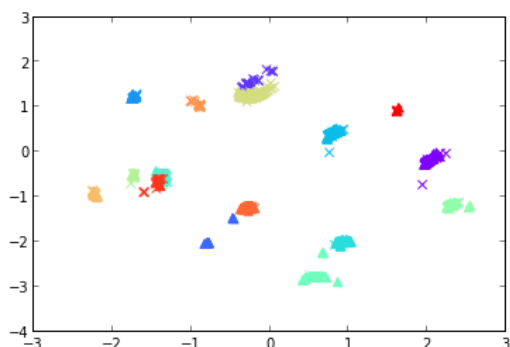

Clustering Example Sample

1. Click **New Notebook** to create a new IPython notebook.

File   Edit   View   Insert   Cell   Kernel   Help

🖫 ✂ ⎘ ⎗   ↑ ↓   ◎ ◎ ▶ ■ Code ⌄ Cell Toolbar: None ⌄

```
In [ ]:
   |
```

Clustering New Notebook

2. In the beginning, the script defines the azure credentials as well as the desired number of clusters for the algorithm to find. The data is loaded from internet and stored in titanic_data.csv. We just upload the csv file to a public windows azure storage account and download it by HTTPs directly.

```
from sklearn.cluster import KMeans
import urllib
import numpy as np
import pandas
import matplotlib
import matplotlib.pyplot as plt
import matplotlib.cm as cm

from sklearn.manifold import MDS
NUM_CLUSTERS = 16
###################################
# download titanic csv data from github
f = urllib.urlopen("http://wrfstorage2.blob.core.windows.net/trainingkit/titanic-data.csv") # YOU MIGHT NEED TO CHANGE THE URL
titanic_csv = f.read()
with open("titanic.csv", "w") as tmp:
  tmp.write(titanic_csv)
```

3. In the next step, the data set is loaded with pandas. Pandas is a data analysis library that makes working with data tables such as CSV data easy. As the "names" and "survived" groups are not needed for the clustering, they are removed from the data frame:

```
# Load data as pandas dataframe
data = pandas.io.parsers.read_csv('titanic.csv', sep=";")
# Remove name and survived dimension to learn
names = data.pop('name')
survived = data.pop('survived')
```

4. In the next code segment, the KMeans clustering operation is initialized, the algorithm is trained and the results (labels for each data set, cluster centers and the set of labels used) are stored in the appropriate variables:

```
# CLUSTERING
# Create KMeans
kmeans = KMeans(n_clusters=NUM_CLUSTERS, init='k-means++', n_init=10, max_iter=300, tol=0.0001, precompute_distances=True, ver
# Train KMeans
kmeans.fit(data)
# Get the results
kmeans_labels = kmeans.labels_
kmeans_cluster_centers = kmeans.cluster_centers_
kmeans_labels_unique = np.unique(kmeans_labels)
```

5. After clustering the data, the multi-dimensional data is reduced to two dimensions for plotting:

```
# PLOT PREPARATION
```

```
# Reduce to two dimensions for plotting
mds = MDS(n_components=2)
mds.fit(data)
scaled_coordinates = mds.embedding_
# PLOT ON TWO DIMENSIONS
labelled_data_x = (dict(), dict())
labelled_data_y = (dict(), dict())
for label in kmeans_labels_unique:
 labelled_data_x[0][label] = []
 labelled_data_y[0][label] = []
 labelled_data_x[1][label] = []
 labelled_data_y[1][label] = []
for i in range(0, len(names)):
 label = kmeans_labels[i]
 labelled_data_x[survived[i]][label].append(scaled_coordinates[i][0])
 labelled_data_y[survived[i]][label].append(scaled_coordinates[i][1])
```

6. The script prepares the data in order to be plotted in multiple colors (depending on their cluster) as well as the status of whether the passenger has survived or not. Surviving passengers and those who did not survive are assigned different markers and, in the end, the plot is shown:

```
# PLOTTING
colors = cm.rainbow(np.linspace(0, 1, NUM_CLUSTERS))
markers = ['x', '^']
for i in kmeans_labels_unique:
 for j in [0, 1]:
  plt.scatter(labelled_data_x[j][i], labelled_data_y[j][i], color=colors[i], marker=markers[j], s=40)
plt.show()
```

7. The result will show:
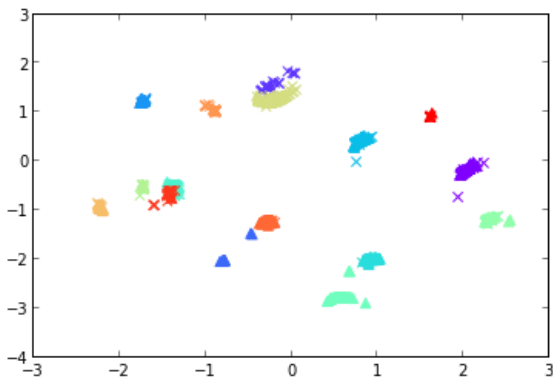
Clustering Result

The full source code can be found in **Source\Exercise2\clustering_sample.py**. If you are having issues, simply copy and paste the content of the source code into IPython notebook and click on the Run (play) button on the menu bar.

Exercise 3: Create a new disk and mount the disk to Windows and Linux.

This exercise will show you how to attach new disks to both Linux and Windows virtual machines.

Attach Empty Disk to Windows

1. You can attach a data disk to a virtual machine to store application data. A data disk is a Virtual Hard Disk (VHD) that you can create either locally with your own computer or in the cloud with Windows Azure. You manage data disks in the virtual machine the same way you do on a server in your office.

Go to the Azure management portal at https://manage.windowsazure.com, select "Virtual Machines" from the bar on the left, click on the VM you want to add the disk to and then go to "Dashboard" at the top bar. In the bar at the bottom, select "Attach" and then "Attach Empty Disk":
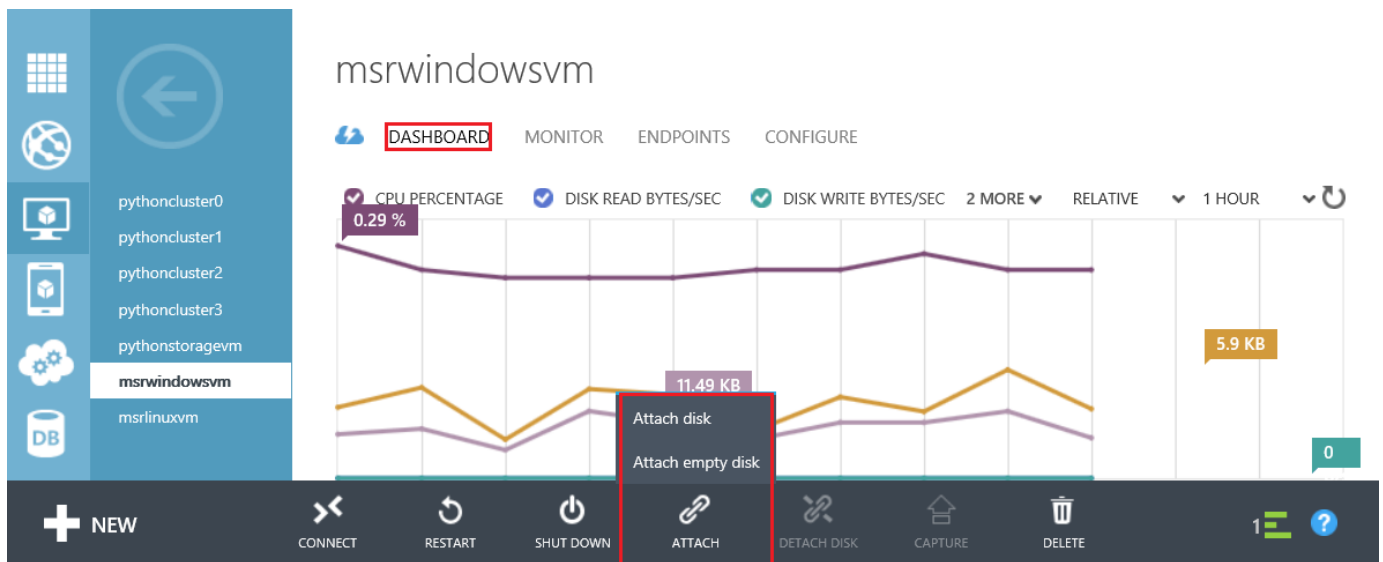
Attach Empty Disk

2. A wizard will open asking you to configure the empty disk. Select a storage location, the file name and the size in GB. In addition, you can choose among the following caching modes:

   ○ Read Only: Reads and writes are cached for future reads but writes are persisted directly to storage
   ○ Read/Write: Reads and writes are cached for future reads. Non-write-through writes are persisted to the local cache first, then lazily flushed to the Windows Azure Blob service. For SQL Server, writes are always persisted to Windows Azure Storage because it uses write-through.
   ○ None (disabled): Requests bypass the cache completely.

The best option to use depends on your intended usage. Read/Write offers the best performance in general, but depending on the type of service you want to use (SQL Server, Apache Cassandra), caching might be counter-productive.

In this example, select Read/Write. You can change this setting later if desired:

Attach an empty disk to the virtual machine

VIRTUAL MACHINE NAME

msrwindowsvm

STORAGE LOCATION

http://labvmstorage.blob.core.windows.net/vhds/

FILE NAME

msrwindowsvm-msrwindowsvm-1007-1

SIZE (GB)

1

HOST CACHE PREFERENCE

| NONE | READ ONLY | READ/WRITE |

Set Disk Property

3. The operation might take a moments. After that, you should see your disk attached on the VM Dashboard in the portal (note that the disk count might need a reload to update):

disks

DISK	TYPE	HOST CACHE	VHD	
msrwindowsvm-msrwindow...	OS disk	Read/Write	http://labvmstorage.blob.cor	
msrwindowsvm-msrwindow...	Data disk	Read/Write	https://labvmstorage.blob.co	

Attached New Disk

4. Then we use Remote Desktop to connect to the machine. Start **Server Manager** from the taskbar. Click *File and Storage Services* on the left panel.
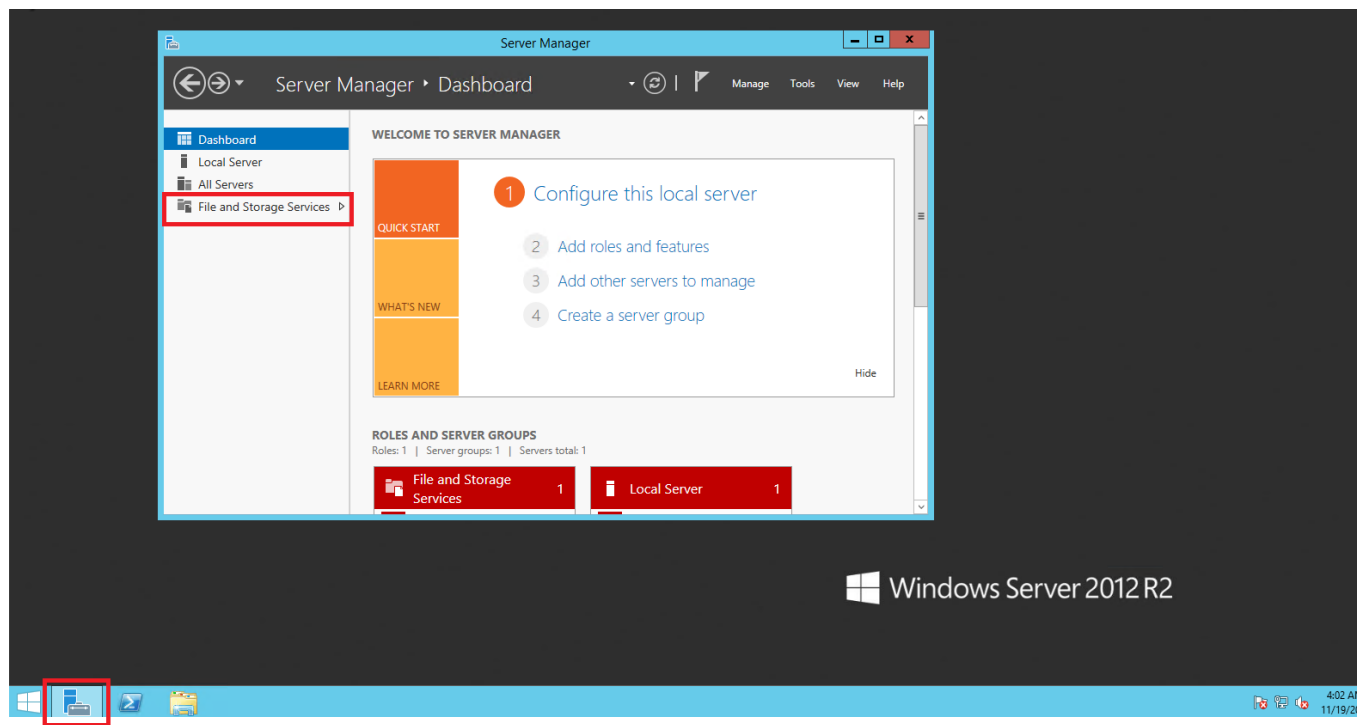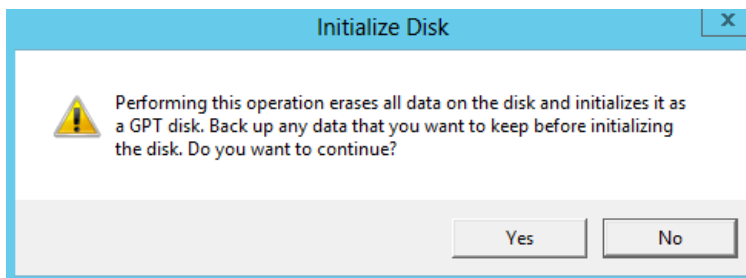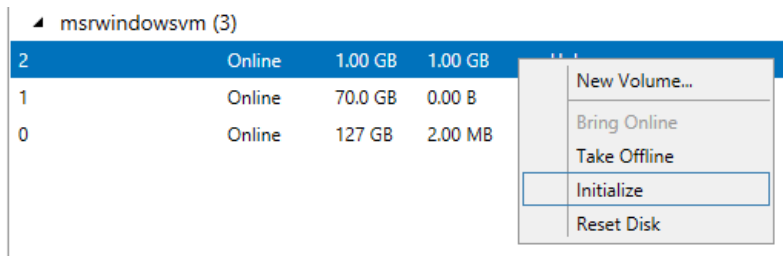
Server Manager
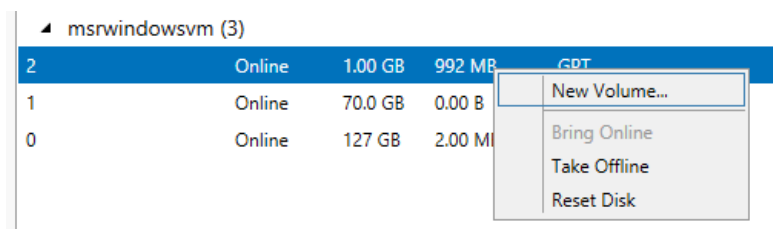
5. Click **Disks** and locate the new virtual disk.
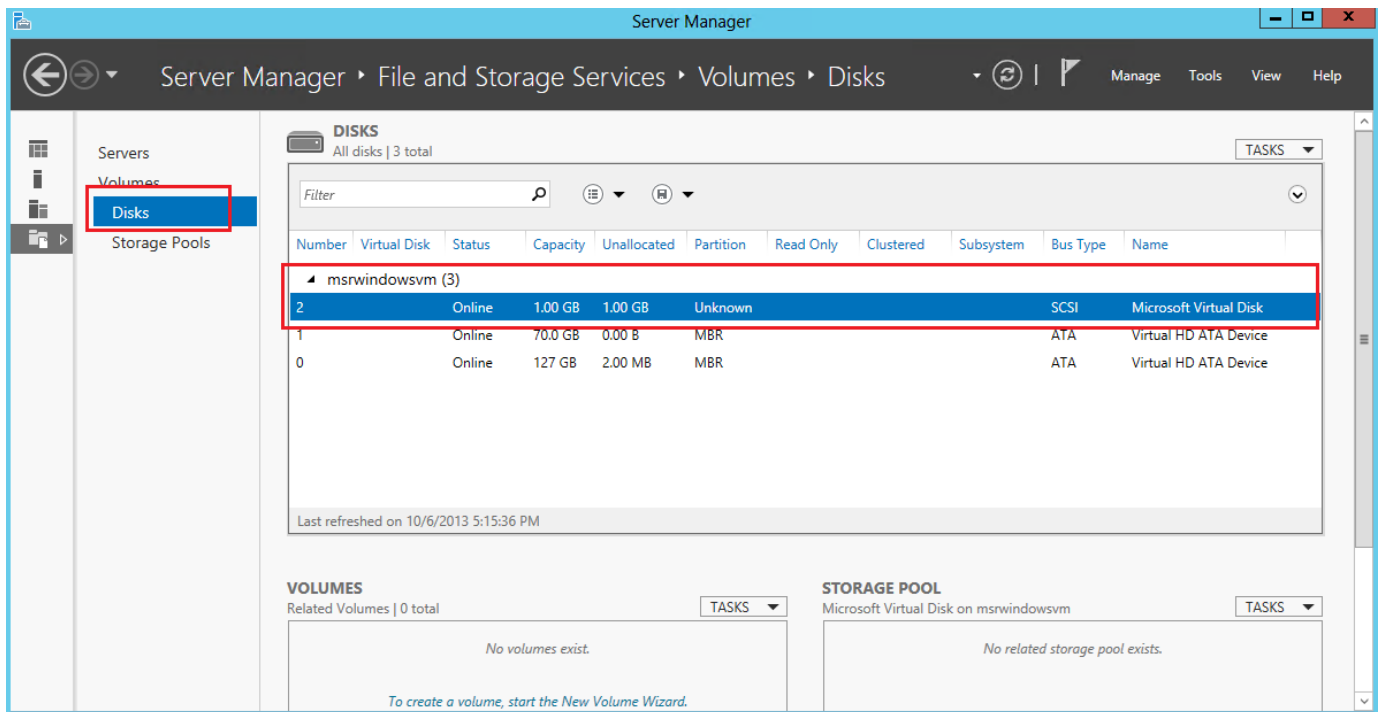
Find New Disk

6. Right click the disk and click **Initialize**. A warning will popup. You can safely ignore it as this disk is empty.
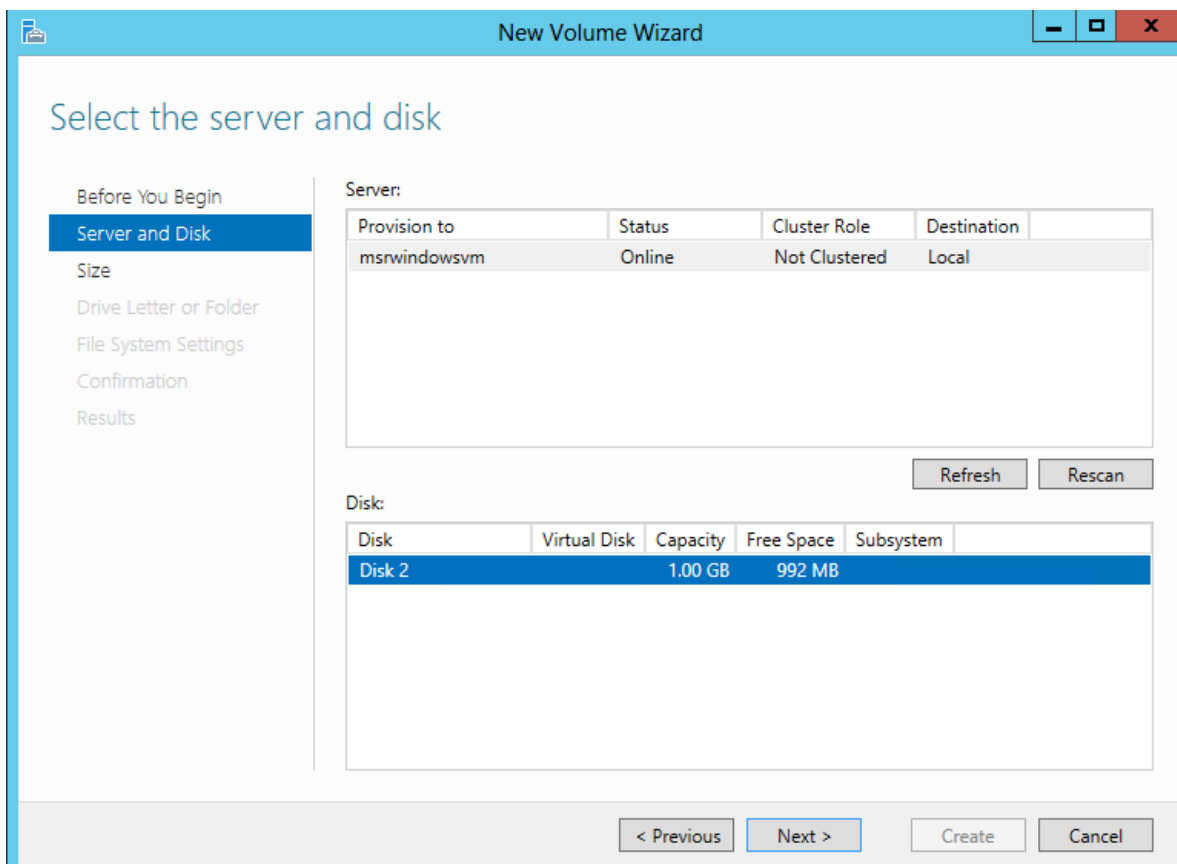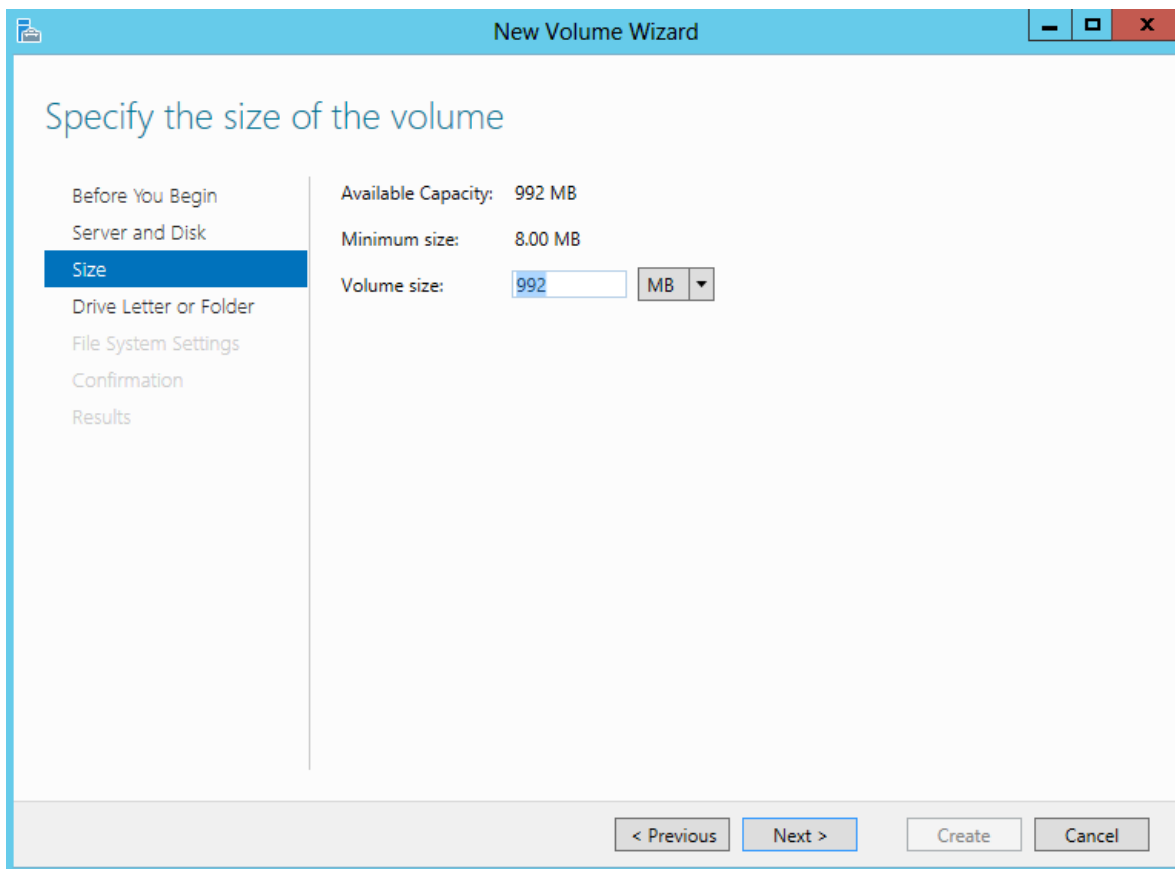
Initialize

7. The initialization process should finish rather quickly. Afterwards, right click on the now initialized disk and select "New volume":

8. A wizard will start. Skip the first section "Before you begin". On the second screen, make sure that your new disk is select and click on "Next":
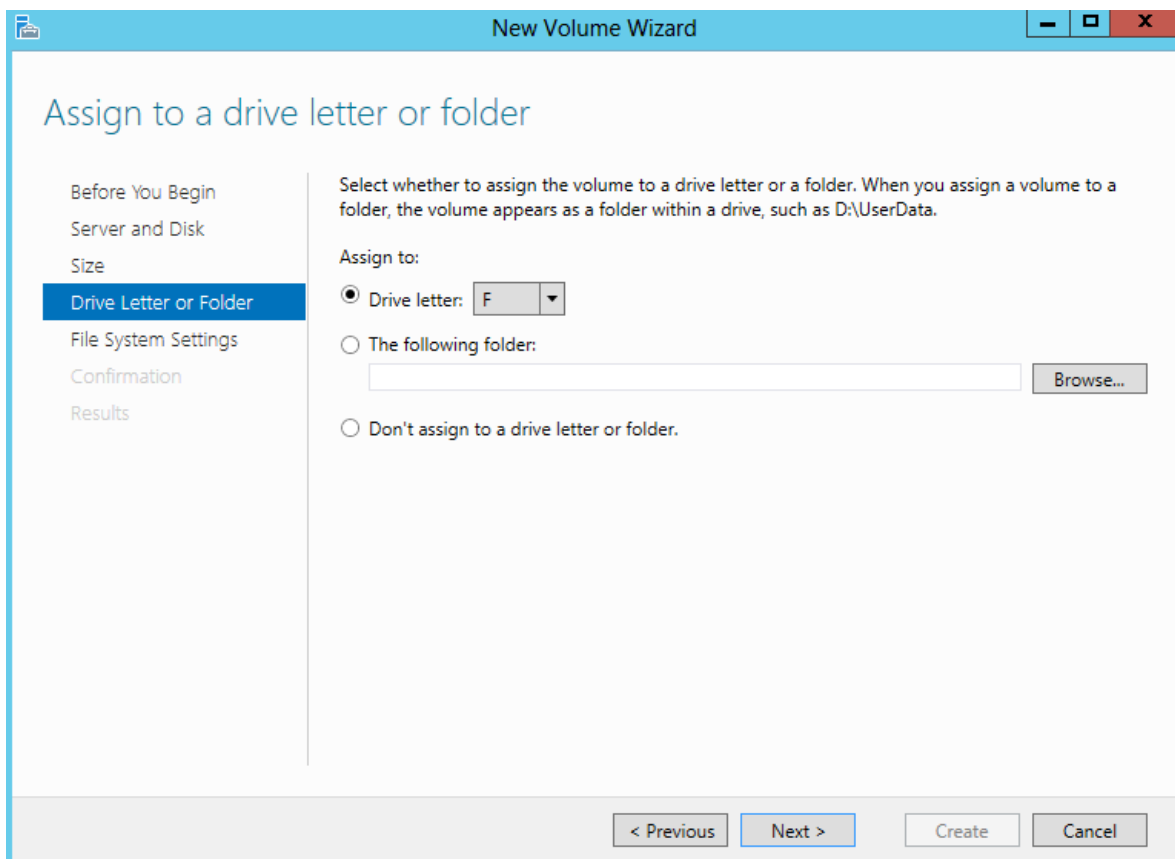

9. On the next page, you can define the size of the volume. The maximum size is selected as default:

Specify the size of the volume

Before You Begin
Server and Disk
Size
Drive Letter or Folder
File System Settings
Confirmation
Results

Available Capacity: 992 MB

Minimum size: 8.00 MB

Volume size: [992]   [MB ▼]

[< Previous]   [Next >]   [Create]   [Cancel]

10. On the next screen, you can assign a drive letter for the new volume, or you can mount it in a particular folder. Choose a drive letter and select "Next":

Assign to a drive letter or folder

Before You Begin
Server and Disk
Size
Drive Letter or Folder
File System Settings
Confirmation
Results

Select whether to assign the volume to a drive letter or a folder. When you assign a volume to a folder, the volume appears as a folder within a drive, such as D:\UserData.

Assign to:

◉ Drive letter: [F ▼]

○ The following folder:

[   ]   [Browse...]

○ Don't assign to a drive letter or folder.

[< Previous]   [Next >]   [Create]   [Cancel]

11. On the next screen, you can select the file system and name the new volume:

New Volume Wizard

Select file system settings

- Before You Begin
- Server and Disk
- Size
- Drive Letter or Folder
- **File System Settings**
- Confirmation
- Results

File system: NTFS

Allocation unit size: Default

Volume label: New Volume

☐ Generate short file names (not recommended)

Short file names (8 characters with 3-character extensions) are required for some 16-bit applications running on client computers, but make file operations slower.

< Previous | Next > | Create | Cancel

12. Confirm your selection on the Confirmation screen, then select "Create". The new volume will be created and mounted:

New Volume Wizard

Completion

- Before You Begin
- Server and Disk
- Size
- Drive Letter or Folder
- File System Settings
- Confirmation
- **Results**

You have successfully completed the New Volume Wizard.

Task	Progress	Status
Gather information		Completed
Create new partition		Completed
Format volume		Completed
Add access path		Completed
Update cache		Completed

< Previous | Next > | Close | Cancel

13. Click on **Close**. If you select your new disk again, you should see the new volume in the "Volumes" window:

Attach Empty Disk to Linux

For linux, the steps to add empty disk is exactly the same. The different is the operation on Linux. After redoing step 1 to step 3, use PuTTY to connect to the linux machine.

1. First we run the following command to find new disk:

```
ls /dev/sd*
```

This shows you all the disks attached to the virtual machine. The new disk is attached at /dev/sdc by default:

2. You can also check the mounted disks with the command:

```
df -h
```

You will see that /dev/sdc is not yet mounted as it is not present in the listing:

3. You need to format the disk to use it. Execute the following command:

sudo fdisk /dev/

In our example:

```
sudo fdisk /dev/sdc
```

When prompted, first enter "n" (new partition), then "p" (primary partition). You can leave the rest of the values at default. This will create a partition over the whole disk. At the end, enter "w" to write the changes to disk:

```
azureuser@msrlinuxvm:~$ sudo fdisk /dev/sdc
[sudo] password for azureuser:
Device contains neither a valid DOS partition table, nor Sun, SGI or OSF disklabel
Building a new DOS disklabel with disk identifier 0x949129bb.
Changes will remain in memory only, until you decide to write them.
After that, of course, the previous content won't be recoverable.

Warning: invalid flag 0x0000 of partition table 4 will be corrected by w(rite)

Command (m for help): n
Partition type:
   p   primary (0 primary, 0 extended, 4 free)
   e   extended
Select (default p): p
Partition number (1-4, default 1): 1
First sector (2048-2097151, default 2048):
Using default value 2048
Last sector, +sectors or +size{K,M,G} (2048-2097151, default 2097151):
Using default value 2097151

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.
Syncing disks.
```

4. Now list the devices again:

```
ls /dev/sd*
```

You should see that a new directory "/dev/sdc1" was added. This directory represents the newly created partition 1 on disk sdc:

```
azureuser@msrlinuxvm:~$ ls /dev/sd*
/dev/sda   /dev/sda1   /dev/sdb   /dev/sdb1   /dev/sdc   /dev/sdc1
```

5. Next, you need to create a file system. In this example, you will use "ext4" as filesystem:

```
sudo mkfs -t ext4 /dev/sdc1
```

```
azureuser@msrlinuxvm:~$ sudo mkfs -t ext4 /dev/sdc1
mke2fs 1.42.5 (29-Jul-2012)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
65536 inodes, 261888 blocks
13094 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=268435456
8 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
        32768, 98304, 163840, 229376

Allocating group tables: done
Writing inode tables: done
Creating journal (4096 blocks): done
Writing superblocks and filesystem accounting information: done
```

6. In the last step, you need to mount the disk. On Linux system, disks are mounted into a directory of your choice, meaning that everything in this directory or in its sub directories are stored on the particular disk/partition.

A common location to mount disks is to use a subdirectory of "/mnt". For this example, create the directory "/mnt/data" as place to store the data:

```
sudo mkdir /mnt/data
```

7. Mount Disk (Temporarily)

You can mount the disk with the following command:

```
sudo mount /dev/sdc1 /mnt/data
```

This will mount the device /dev/sdc1 to /mnt/data. Using the command:

```
df -h
```

again, you should see the mounted disk:

```
azureuser@msrlinuxvm:~$ df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1        29G  4.2G   24G  16% /
none            4.0K     0  4.0K   0% /sys/fs/cgroup
udev            829M   12K  829M   1% /dev
tmpfs           168M  268K  168M   1% /run
none            5.0M     0  5.0M   0% /run/lock
none            840M     0  840M   0% /run/shm
none            100M     0  100M   0% /run/user
/dev/sdb1        69G   52M   66G   1% /mnt/resource
/dev/sdc1       991M  1.3M  939M   1% /mnt/data
```

If you view the files on the disk, you will see only a single folder added by default, lost+found.

```
ls /mnt/data
```

This mounting approach has one drawback: If you reboot the virtual machine, you have to manually mount the disk again. In the next step, you will make the mount configuration persistent.

8. If you want to mount disk persistently, you need to mount the disk in fstab.

    Unmount the disk you mounted in the previous step:

    ```
    sudo umount /mnt/data
    ```

    Next, you need to open an editor to edit /etc/fstab, the file which holds the partitions to mount during startup:

    ```
    sudo nano /etc/fstab
    ```
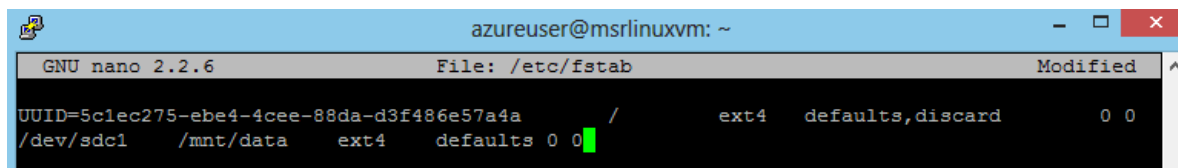
    The format in /etc/fstab is:

    ```
    [Device] [Mount Point] [File System Type] [Options] [Dump] [Pass]
    ```

    Paste the following configuration into a new line in the editor:

    ```
    /dev/sdc1    /mnt/data    ext4    defaults 0 0
    ```

    Then press ctrl+x to exit. Confirm saving with "y" and then enter.

```
                         azureuser@msrlinuxvm: ~                        _  □  ×
  GNU nano 2.2.6                 File: /etc/fstab                    Modified  ^

UUID=5c1ec275-ebe4-4cee-88da-d3f486e57a4a      /       ext4   defaults,discard    0 0
/dev/sdc1     /mnt/data    ext4    defaults 0 0
```

    Now you should try to mount the disks in fstab. This can be done with:

    ```
    sudo mount -a
    ```

    There should be no output if successful. Again, you can check the new partition using:

    ```
    df -h
    ```

```
azureuser@msrlinuxvm:~$ sudo mount -a
azureuser@msrlinuxvm:~$ df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1        29G  4.2G   24G  16% /
none            4.0K     0  4.0K   0% /sys/fs/cgroup
udev            829M   12K  829M   1% /dev
tmpfs           168M  268K  168M   1% /run
none            5.0M     0  5.0M   0% /run/lock
none            840M     0  840M   0% /run/shm
none            100M     0  100M   0% /run/user
/dev/sdb1        69G   52M   66G   1% /mnt/resource
/dev/sdc1       991M  1.3M  939M   1% /mnt/data
```

Summary

By completing this hands-on lab you learned the following:

- Create a machine with Windows Server 2012 R2 and run R job.
- Create a machine from VMDepot and run ipython job.
- Create a new disk and mount the disk to Windows and Linux.