Windows Azure Platform

CSCI 4180 Qin Chuan

Overview

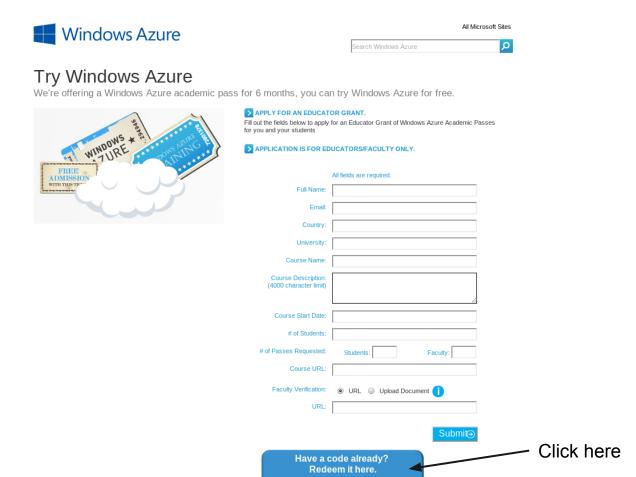
- This year we use Windows Azure Platform to do the assignments
- Get the 12 character code before you start
- Redeem your Windows Azure at
 - www.WindowsAzurePass.com/azureu
- Ensure email won't be blocked
 - admin@windowsazurepassadmin.com

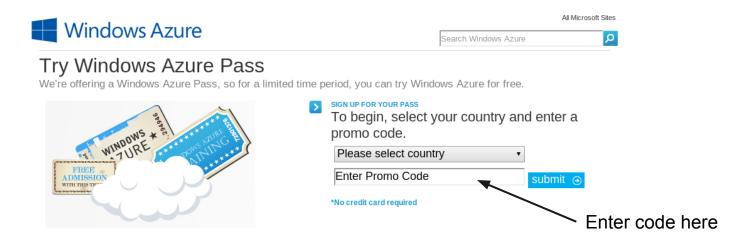
Overview

- Each account includes following setting
- Windows Azure
 - 2 small compute instances
 - 70GB of storage
 - 50,000,000 storage transactions
 - 10 Shared WebSites
 - 10 Shared Mobile Services

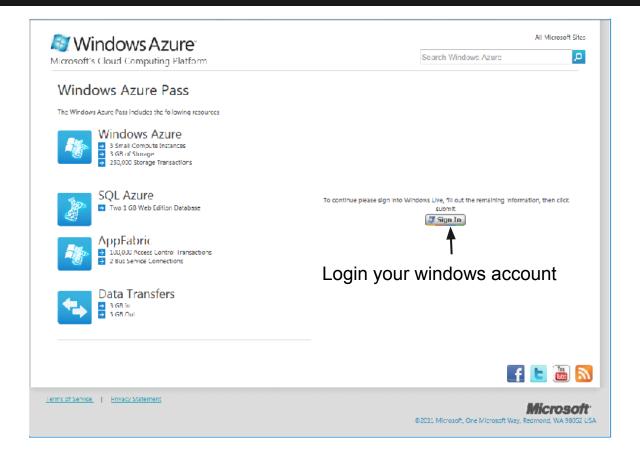
Overview

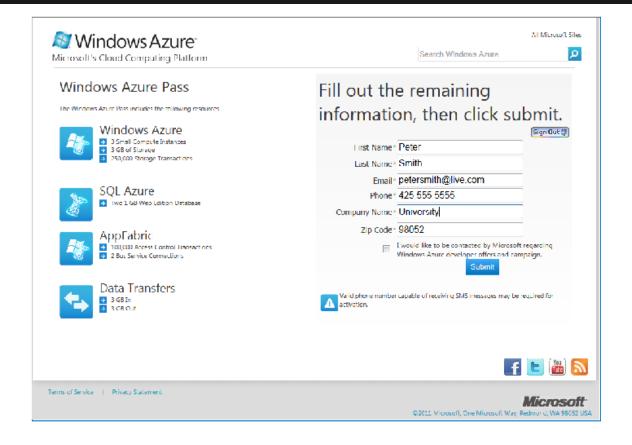
- SQL Azure
 - Two 1GB Web Edition database
- AppFabric
 - 1500 Service Bus Relay hours
 - 500,000 Service Bus Messages
- Data Transfers
 - 12 GB in
 - o 12 GB out

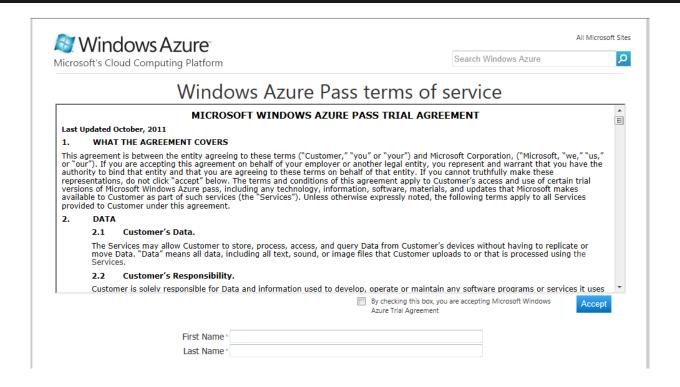




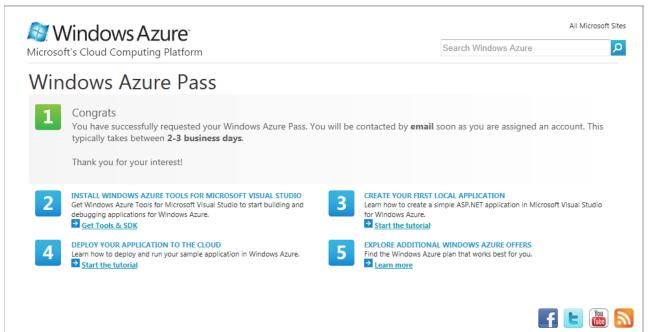
- Followed with some register information
- To redeem the pass, you also need a windows live account







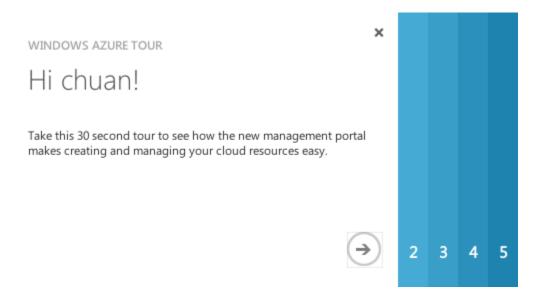
Accept the term of use agreement



 After register, your account will be assigned by register email in 2-3 days

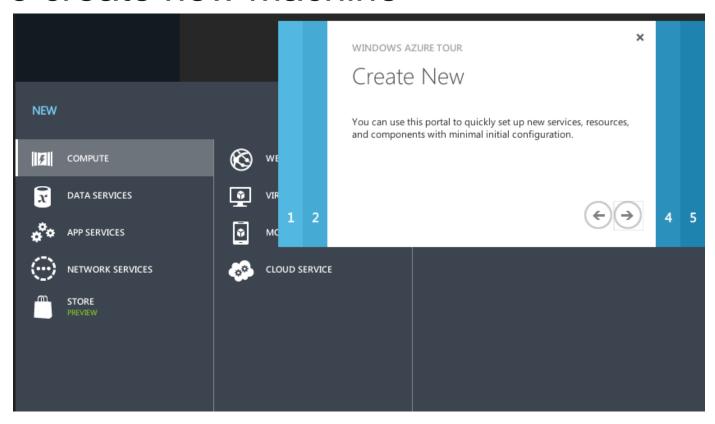
Access Azure

 Once you receive the email, you access the link inside the email



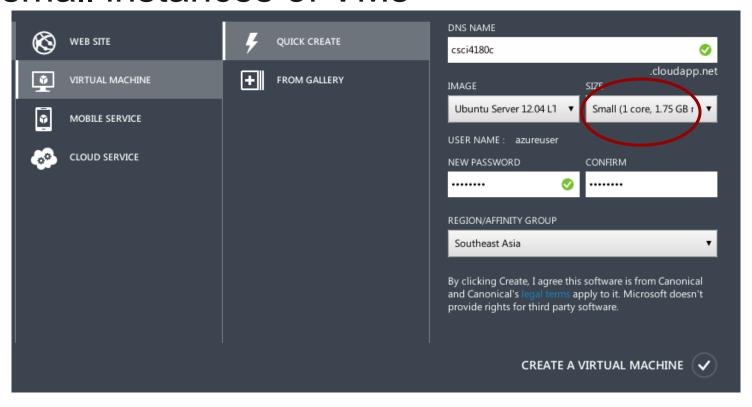
Access Azure

To create new machine



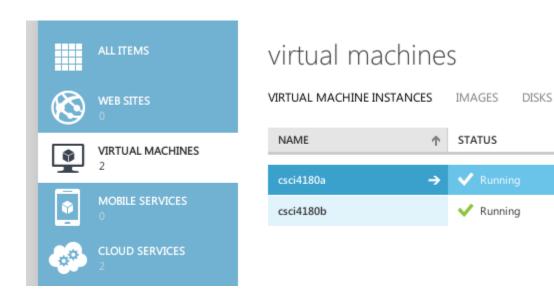
Create VMs

 The student grand only has the quota of two small instances of VMs



Create VMs

When succeeded, you can browse the VMs



Other Services

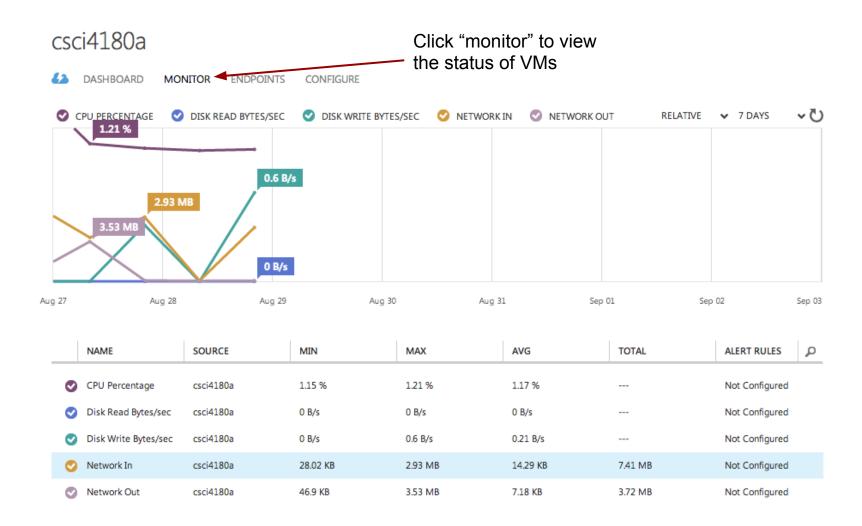
There are also several interesting services on Azure

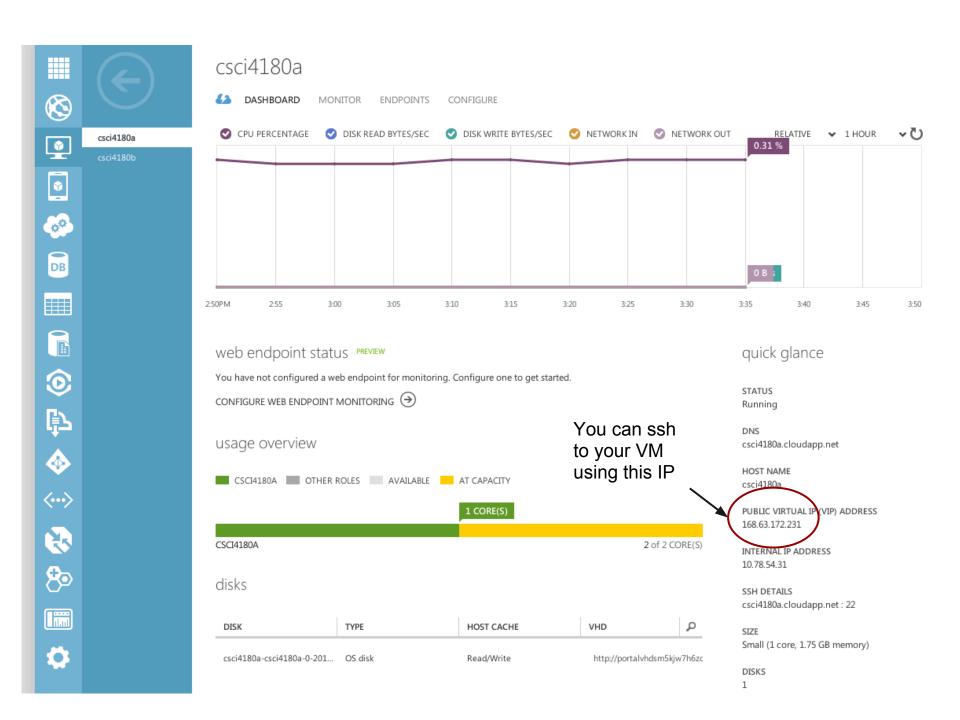
 You might have a try on those cool stuffs

Do remember the limit of academic grand



Browse VM status





Access VMs

In your terminal

```
chintran@Axima ~ $ ssh azureuser@168.63.172.231
azureuser@168.63.172.231's password:
Welcome to Ubuntu 12.04.2 LTS (GNU/Linux 3.2.0-48-virtual x86 64)
 * Documentation: https://help.ubuntu.com/
  System information as of Tue Aug 20 08:04:24 UTC 2013
  System load: 0.0
                                 Processes:
                                                      81
  Usage of /: 5.0% of 28.82GB Users logged in:
                              IP address for eth0: 10.78.54.31
  Memory usage: 12%
  Swap usage: 0%
  Graph this data and manage this system at https://landscape.canonical.com/
  Get cloud support with Ubuntu Advantage Cloud Guest:
    http://www.ubuntu.com/business/services/cloud
  Use Juju to deploy your cloud instances and workloads:
    https://juju.ubuntu.com/#cloud-precise
55 packages can be updated.
31 updates are security updates.
Last login: Thu Aug  1 06:17:53 2013 from sepc463.se.cuhk.edu.hk
azureuser@csci4180a:~$
```

First install Java 6

- ~\$ sudo add-apt-repository ppa:webupd8team/java
- ~\$ sudo apt-get update
- ~\$ sudo apt-get install oracle-java6-installer

Create hadoop user and add key

- ~\$ sudo addgroup hadoop
- ~\$ sudo adduser --ingroup hadoop hduser
- ~\$ su hduser
- ~\$ ssh-keygen -t rsa -P ""
- ~\$ cat \$HOME/.ssh/id_rsa.pub >> \$HOME/.ssh/authorized_keys

- Copy ssh keys
 - only need to be done on namenode
 - \$ ssh-copy-id -i \$HOME/.ssh/id_rsa.pub username@192.168.0.1
 - \$ ssh-copy-id -i \$HOME/.ssh/id rsa.pub username@192.168.0.2
 - \$ ssh-copy-id -i \$HOME/.ssh/id_rsa.pub username@192.168.0.3
- Check whether ssh needs password or not
 - master node should be able to access slave nodes without password after this

Disable IPv6

- ~\$ sudo gedit /etc/sysctl.conf
 - net.ipv6.conf.all.disable_ipv6 = 1
 - net.ipv6.conf.default.disable_ipv6 = 1
 - o net.ipv6.conf.lo.disable_ipv6 = 1
- ~\$ sudo sysctl -p

Download Hadoop

- ~\$ wget http://www.eng.lsu.edu/mirrors/apache/hadoop/core/hadoop-0.22.0/hadoop-0.22.0.tar.gz
- ~\$ cd /home/hduser
- ~\$ tar xzf hadoop-0.22.2.tar.gz
- ~\$ mv hadoop-0.22.2 hadoop

- Set hadoop environment
 - In hadoop/conf/hadoop-env.sh append following
 - export JAVA_HOME=/usr/lib/jvm/java-6-sun
 - #depends where you put the jvm
 - export HADOOP_OPTS=- Djava.net.preferIPv4Stack=true
 - #disable IPv6

- Set hadoop core
 - In hadoop/conf/core-site.xml append following
 - <property>
 - <name>hadoop.tmp.dir</name>
 - <value>/home/hduser/hadoop/tmp</value>
 - /property>
 - <property>
 - <name>fs.default.name</name>
 - <value>hdfs://192.168.0.1:54310</value>
 - /property>

- Set hadoop mapred-site
 - In hadoop/conf/mapred-site.xml append following
 - <property>
 - <name>mapred.job.tracker</name>
 - <value>192.168.0.1:54311</value>
 - </property>
- Set hadoop hdfs
 - In hadoop/conf/hdfs-site.xml append following
 - <property>
 - <name>dfs.replication</name>
 - <value>3</value>
 - </property>

- Set hadoop master
 - In hadoop/conf/masters Add hostname which is supposed to run JobTracker and NameNode
 - o eg. 192.168.0.1
- Set hadoop slave
 - In hadoop/conf/slaves Add hostname which is supposed to run TaskTracker and DataNode
 - eg. 192.168.0.1<mark></mark>反
 - o eg. 192.168.0.2
 - o eg. 192.168.0.3

- Format Hadoop Namenode
 - only on master
 - ~\$ hadoop namenode -format

Important!

- Start hadoop on master
 - ~\$ start-all.sh
- you can type "jps" to see whether the startup is successful

```
5660 TaskTracker
7176 Jps
5436 JobTracker
5141 DataNode
5352 SecondaryNameNode
4923 NameNode
```

- Office SHB 118
- Facebook https://www.facebook.com/groups/200985016728843/

Thank you