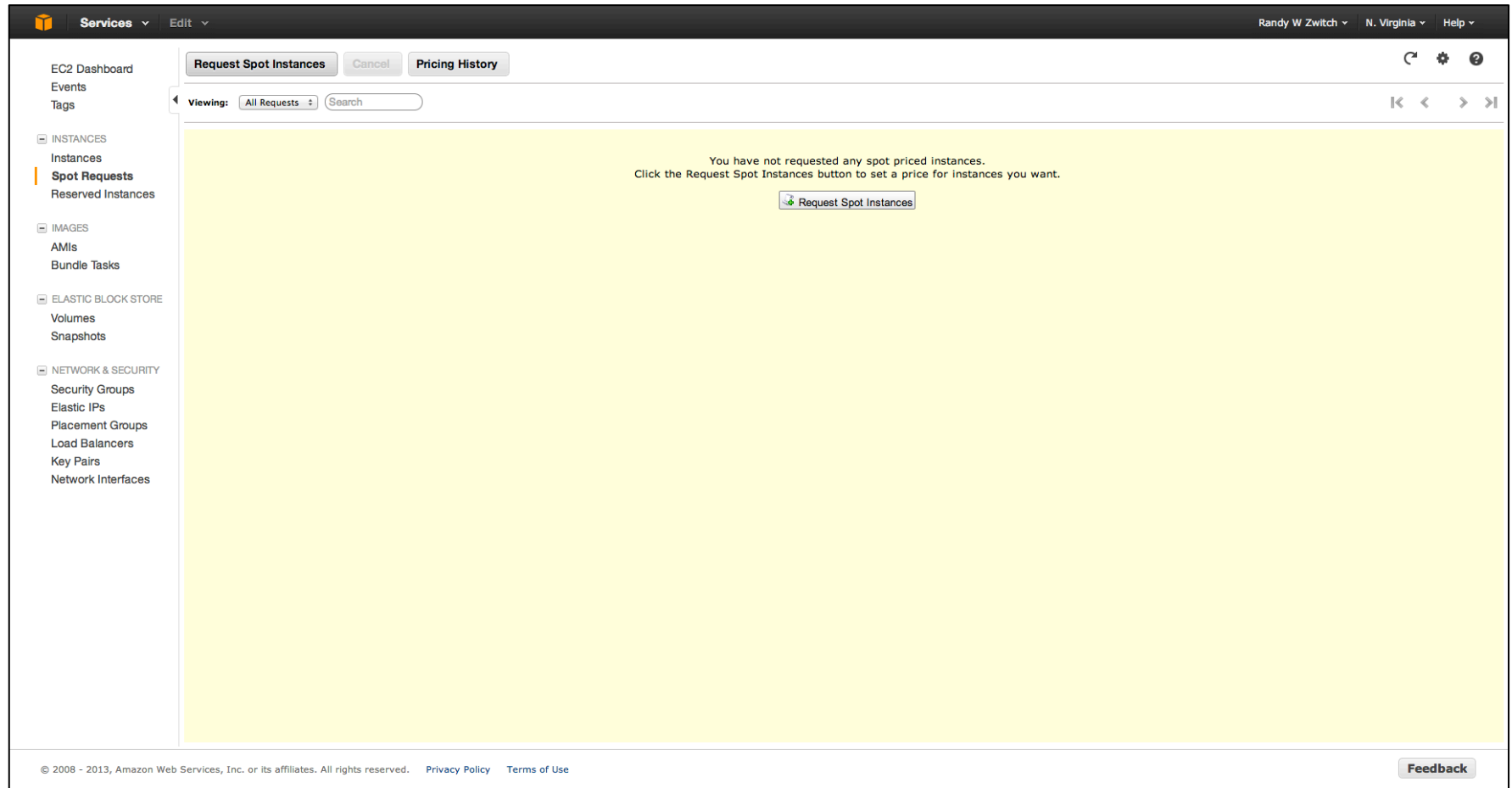


Cluster Computing Using IPython Notebook with Amazon EC2

1. Launching EC2 Cluster Instance

1. Launch Spot Instance from Spot Instance Menu



<http://randywitch.com>

1b. Launch Spot Instance - Ubuntu Server 12.04LTS for HVM Instances

The screenshot displays the AWS Management Console interface for launching an instance. The top navigation bar shows the user is logged in as 'Randy W Zwitch' in the 'N. Virginia' region. The main content area is titled 'Step 1: Choose an Amazon Machine Image (AMI)'. It features a list of available AMIs, each with a logo, name, ID, description, root device type, and virtualization type. The 'Ubuntu Server 12.04.3 LTS for HVM Instances' AMI is highlighted with a blue border. Other visible AMIs include Amazon Linux, Red Hat Enterprise Linux, SUSE Linux, and Microsoft Windows Server. A 'Cancel and Exit' button is located in the top right corner of the AMI selection area.

Logo	AMI Name	AMI ID	Description	Root device type	Virtualization type	Architecture	Action
Amazon Linux	Amazon Linux AMI (HVM) 2013.09.1	ami-d1bfe4b8	The Amazon Linux AMI is an EBS-backed, HVM image. It includes Linux 3.4, AWS tools, and repository access to multiple versions of MySQL, PostgreSQL, Python, Ruby, and Tomcat.	ebs	paravirtual	64-bit	Select
Red Hat	Red Hat Enterprise Linux 6.4 for Cluster Instances	ami-3218595b	Red Hat Enterprise Linux version 6.4 is an EBS-backed, HVM image for use with Amazon EC2 Cluster Instances.	ebs	hvm	64-bit	Select
SUSE Linux	Cluster Instances HVM SUSE Linux Enterprise 11	ami-b6c146df	SUSE Linux Enterprise Server 11 Service Pack 2, 64-bit architecture, and HVM based virtualization for use with Amazon EC2 Cluster Compute and Cluster GPU instances. Nvidia driver installs automatically during startup.	ebs	hvm	64-bit	Select
Ubuntu	Ubuntu Server 12.04.3 LTS for HVM Instances	ami-b93264d0	Ubuntu Server 12.04.3 LTS, with support available from Canonical (http://www.ubuntu.com/cloud/services). For use with Cluster and other HVM instances	ebs	hvm	64-bit	Select
Ubuntu	Ubuntu Server 13.10 for HVM Instances	ami-a1184ac8	Ubuntu Server version 13.10, with support available from Canonical (http://www.ubuntu.com/cloud/services). For use with Cluster and other HVM instances	ebs	hvm	64-bit	Select
Windows	Microsoft Windows Server 2012 Base	ami-7527031c	Microsoft Windows 2012 Standard edition with 64-bit architecture. [English]	ebs	hvm	64-bit	Select
Windows	Microsoft Windows Server 2012 with SQL Server Express	ami-7325011a	Microsoft Windows Server 2012 Standard edition, 64-bit architecture, Microsoft SQL Server 2012 Express edition. [English]	ebs	hvm	64-bit	Select
Windows	Microsoft Windows Server 2012 with SQL Server Web	ami-1d200474					Select

© 2008 - 2013, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#) [Feedback](#)

<http://randytwitch.com>

1c. Launch Spot Instance – Choose cc2.8x1 instance

Services

Edit

Randy W Zwitch

N. Virginia

Help

1. Choose AMI

2. Choose Instance Type

3. Configure Instance

4. Add Storage

5. Tag Instance

6. Configure Security Group

7. Review

Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. [Learn more](#) about instance types and how they can meet your computing needs.

Currently selected: cc2.8xlarge (88 ECUs, 32 vCPUs, 60.5 GiB memory, 4 x 840 GiB Storage Capacity)

All instance types

General purpose

Storage optimized

Compute optimized

GPU instances

Memory optimized

Compute optimized

Compute optimized instances have a higher ratio of vCPUs to memory than other families, and the lowest cost per vCPU among all Amazon EC2 instance types. We recommend compute optimized instances for running CPU-bound scale out applications. Examples of such applications include high traffic front end fleets, on-demand batch processing, distributed analytics, web servers, batch processing, and high performance science and engineering applications

Size	ECUs	vCPUs	Memory (GiB)	Instance Storage (GiB)	EBS-Optimized Available	Network Performance
c3.large	7	2	3.75	2 x 16 (SSD)	-	Moderate
c3.xlarge	14	4	7.5	2 x 40 (SSD)	Yes	High
c3.2xlarge	28	8	15	2 x 80 (SSD)	Yes	High
c3.4xlarge	55	16	30	2 x 160 (SSD)	Yes	High
c3.8xlarge	108	32	60	2 x 320 (SSD)	-	High
cc1.4xlarge	33.5	16	23	2 x 840	-	10 Gigabit
cc2.8xlarge	88	32	60.5	4 x 840	-	10 Gigabit

C3 instances are the latest generation of Compute-optimized instances. For C3 instances, each vCPU is a hardware hyperthread from Intel Xeon E5-2680v2 processors.
C1 instances are based on Intel Xeon processors.
For CC2 instances, each vCPU is a hardware hyperthread from Intel Xeon E5-2670 processors.
Instances launched into the same cluster placement group are placed into a non-blocking 10 Gigabit ethernet network.

Cancel

Previous

Next: Configure Instance Details

© 2008 - 2013, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Privacy Policy

Terms of Use

Feedback

<http://randyzwitch.com>

1d. Launch Spot Instance - Set bid price

For max bid, set price that you're comfortable paying to keep instance running. Cost has been pretty stable at \$0.27/hr for a while

The screenshot shows the AWS Management Console interface for launching a Spot Instance. The top navigation bar includes the AWS logo, 'Services', 'Edit', and user information 'Randy W Zwitch', 'N. Virginia', and 'Help'. Below the navigation bar, a series of tabs indicates the steps: '1. Choose AMI', '2. Choose Instance Type', '3. Configure Instance', '4. Add Storage', '5. Tag Instance', '6. Configure Security Group', and '7. Review'. The current step is 'Step 3: Configure Instance Details'.

The configuration options are as follows:

- Instance Type:** A dropdown menu showing 'us-east-1c 0.2701' and 'us-east-1d 0.27'.
- Maximum price:** A text input field with '\$ 2.50'.
- Launch group:** A text input field with '(Optional)'.
- Availability Zone group:** A text input field with '(Optional)'.
- Request valid from:** A text input field with 'Any time' and an 'Edit' link.
- Request valid to:** A text input field with 'Any time' and an 'Edit' link.
- Persistent request:** A checkbox labeled 'Persistent request' which is currently unchecked.
- Network:** A dropdown menu with 'Launch into EC2-Classic' and a 'Create new VPC' button.
- Availability Zone:** A dropdown menu with 'us-east-1a'.
- Placement group:** A dropdown menu with 'No placement group'.
- IAM role:** A dropdown menu with 'None'.
- Monitoring:** A checkbox labeled 'Enable CloudWatch detailed monitoring' which is unchecked, with a note 'Additional charges apply.'
- Advanced Details:** A section with a dropdown for 'Kernel ID' (set to 'Use default'), a dropdown for 'RAM disk ID' (set to 'Use default'), and radio buttons for 'User data' (set to 'As text', with 'As file' and 'Input is already base64 encoded' as other options). Below these is a text area for user data with '(Optional)' as a placeholder.

At the bottom right, there are four buttons: 'Cancel', 'Previous', 'Review and Launch' (highlighted in blue), and 'Next: Add Storage'. At the bottom left, there is a copyright notice: '© 2008 - 2013, Amazon Web Services, Inc. or its affiliates. All rights reserved.' followed by links for 'Privacy Policy' and 'Terms of Use'. At the bottom right, there is a 'Feedback' button.

<http://randyzwitch.com>

1e. Launch Spot Instance - Set security

I generally open all ports and only allow my IP address as a simplistic security protocol, since this is a spot instance that I use for a few hours

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group: ☒ Create a new security group ☐ Select an existing security group

Security group name:

Description:

Protocol	Type	Port Range (Code)	Source
All TCP	TCP	0 - 65535	My IP 98.114.207.62/32

[Add Rule](#)

[Cancel](#) [Previous](#) [Review and Launch](#)

© 2008 - 2013, Amazon Web Services, Inc. or its affiliates. All rights reserved. [Privacy Policy](#) [Terms of Use](#) [Feedback](#)

<http://randytwitch.com>

1f. Launch Spot Instance - Launch

Services

Edit

Randy W Zwitch

N. Virginia

Help

1. Choose AMI

2. Choose Instance Type

3. Configure Instance

4. Add Storage


5. Tag Instance

6. Configure Security Group

7. Review

Step 7: Review Spot Instance Request

Please review your instance launch details. You can go back to edit changes for each section. Click **Launch** to assign a key pair to your instance and complete the launch process.




Your instance configuration is not eligible for the free usage tier
To launch an instance that's eligible for the free usage tier, check your AMI selection, instance type, configuration options, or storage devices. Learn more about [free usage tier](#) eligibility and usage restrictions.
[Don't show me this again](#)

x

AMI Details

Edit AMI



Ubuntu Server 12.04.3 LTS for HVM Instances - ami-b93264d0
Ubuntu Server 12.04.3 LTS, with support available from Canonical (<http://www.ubuntu.com/cloud/services>). For use with Cluster and other HVM Instances
Root Device Type: ebs Virtualization type: hvm

Instance Type

Edit instance type

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GiB)	EBS-Optimized Available	Network Performance
cc2.8xlarge	88	32	60.5	4 x 840	-	10 Gigabit

Security Groups





Edit security groups

Security group name

launch-wizard-1

Description

launch-wizard-1 created on Thursday, November 21, 2013 1:20:24 PM UTC-5

Protocol 	Type 	Port Range (Code) 	Source 
Custom TCP Rule	TCP	0 - 65535	98.114.207.62/32

Instance Details

Edit instance details

Storage

Edit storage

Tags

Edit tags

Cancel

Previous

Launch

© 2008 - 2013, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Privacy Policy

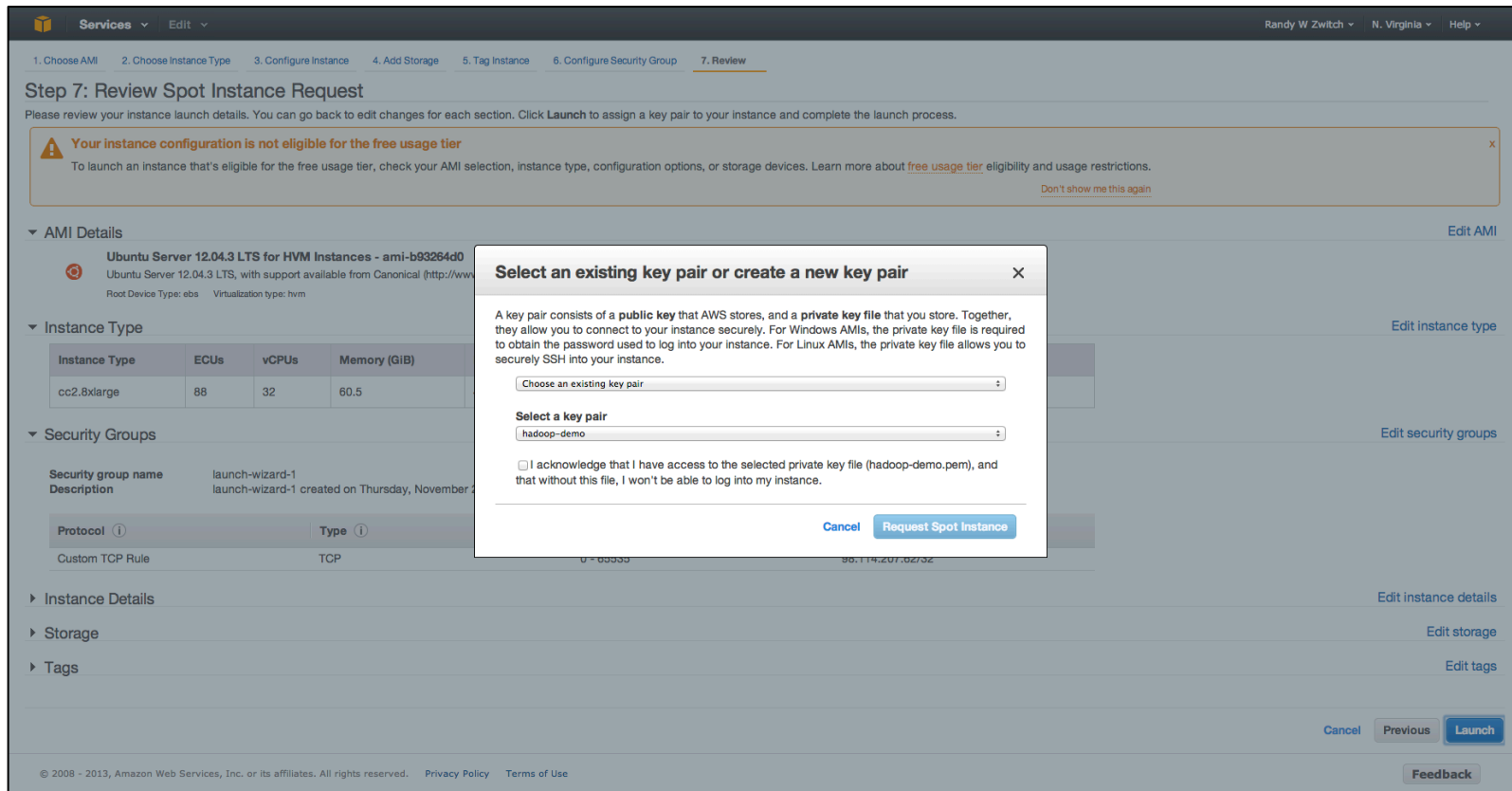
Terms of Use

Feedback

<http://randywitch.com>

1g. Launch Spot Instance - Pick .pem keys

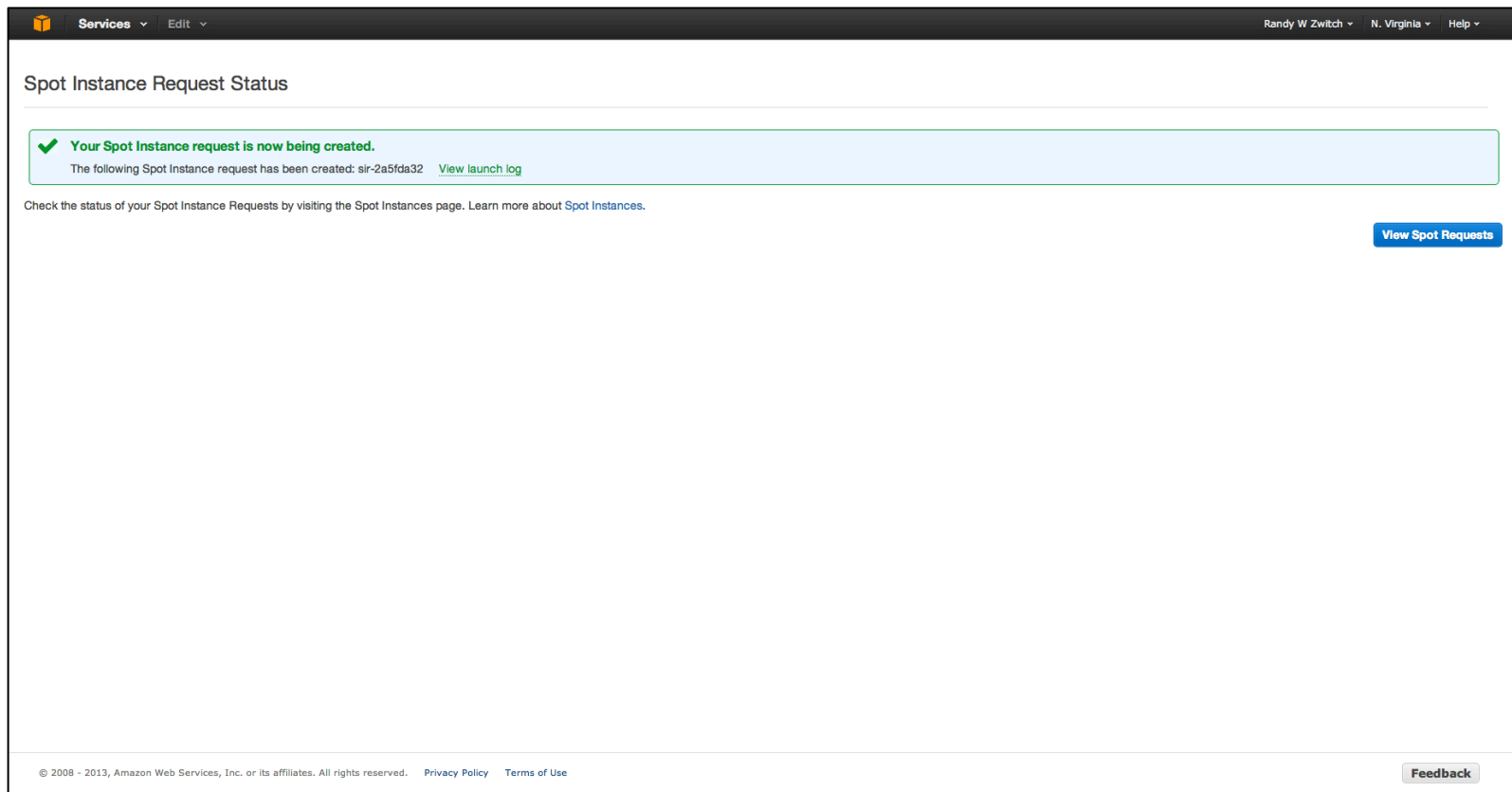
If you don't specify a key pair, you can't login to the instance!



<http://randyzwitch.com>

1g. Launch Spot Instance - Wait for fulfillment

If your spot request is fulfilled, it will take about 5-10 minutes to launch



<http://randyZWitch.com>

2. Installing & Configuring Python/IPython Using Anaconda

FULL INSTRUCTIONS:

[HTTP://IPYTHON.ORG/IPYTHON-DOC/DEV/INTERACTIVE/PUBLIC_SERVER.HTML#NOTEBOOK-PUBLIC-SERVER](http://ipython.org/ipython-doc/dev/interactive/public_server.html#notebook-public-server)

2a. Installing IPython - SSH into EC2 Instance

SSH into EC2 instance, create /temp directory, then download Anaconda (64-bit, Linux).
<http://continuum.io/downloads>

Run script after downloading to install Anaconda: `bash Anaconda-1.8.0-Linux-x86_64.sh`

```
personal_ec2_keys — ubuntu@ip-10-178-132-244: ~/temp — ssh — 156x49
ubuntu@ip-10-178-132-244:~$ mkdir temp
ubuntu@ip-10-178-132-244:~$ cd temp
ubuntu@ip-10-178-132-244:~/temp$ ls
ubuntu@ip-10-178-132-244:~/temp$ wget http://09c8d0b2229f813c1b93-c95ac804525aac4b6dba79b00b39d1d3.r79.cf1.rackcdn.com/Anaconda-1.8.0-Linux-x86_64.sh
--2013-11-21 18:46:16-- http://09c8d0b2229f813c1b93-c95ac804525aac4b6dba79b00b39d1d3.r79.cf1.rackcdn.com/Anaconda-1.8.0-Linux-x86_64.sh
Resolving 09c8d0b2229f813c1b93-c95ac804525aac4b6dba79b00b39d1d3.r79.cf1.rackcdn.com (09c8d0b2229f813c1b93-c95ac804525aac4b6dba79b00b39d1d3.r79.cf1.rackcdn.com)... 165.254.119.18, 165.254.119.33, 2001:688:1:8::500c:6050, ...
Connecting to 09c8d0b2229f813c1b93-c95ac804525aac4b6dba79b00b39d1d3.r79.cf1.rackcdn.com (09c8d0b2229f813c1b93-c95ac804525aac4b6dba79b00b39d1d3.r79.cf1.rackcdn.com)|165.254.119.18|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 488287119 (466M) [application/x-sh]
Saving to: `Anaconda-1.8.0-Linux-x86_64.sh'

100%[=====>] 488,287,119 10.6M/s in 52s

2013-11-21 18:47:08 (9.02 MB/s) - `Anaconda-1.8.0-Linux-x86_64.sh' saved [488287119/488287119]

ubuntu@ip-10-178-132-244:~/temp$ ls
Anaconda-1.8.0-Linux-x86_64.sh
ubuntu@ip-10-178-132-244:~/temp$ bash Anaconda-1.8.0-Linux-x86_64.sh

Welcome to Anaconda 1.8.0 (by Continuum Analytics, Inc.)

In order to continue the installation process, please review the license
agreement.
Please, press ENTER to continue
>>>
=====
Anaconda END USER LICENSE AGREEMENT
=====
```

2b. Installing IPython - Generate Password

In IPython REPL, use the `IPython.lib passwd()` feature to create a password. Copy password to a text editor for later use.

(No, this is not a real password to use on my EC2 instance!)

```
ubuntu@ip-10-178-132-244:~$ ipython
Python 2.7.5 |Anaconda 1.8.0 (64-bit)| (default, Nov  4 2013, 15:30:26)
Type "copyright", "credits" or "license" for more information.

IPython 1.1.0 -- An enhanced Interactive Python.
?                -> Introduction and overview of IPython's features.
%quickref        -> Quick reference.
help             -> Python's own help system.
object?         -> Details about 'object', use 'object??' for extra details.

In [1]: from IPython.lib import passwd

In [2]: passwd()
Enter password:
Verify password:
Passwords do not match.
Enter password:
Verify password:
Out[2]: 'sha1:694bff492f56:e06ae3083855a56269446d2a386a557fb538ff6f'

In [3]:
```

2c. Installing IPython - Create nbserver profile

Create an IPython profile called 'nbserver', which we will use as our profile to create the public Notebook server

```
ubuntu@ip-10-178-132-244:~$ ipython profile create nbserver
[ProfileCreate] Generating default config file: u'/home/ubuntu/.ipython/profile_nbserver/ipython_config.py'
[ProfileCreate] Generating default config file: u'/home/ubuntu/.ipython/profile_nbserver/ipython_notebook_config.py'
[ProfileCreate] Generating default config file: u'/home/ubuntu/.ipython/profile_nbserver/ipython_nbconvert_config.py'
ubuntu@ip-10-178-132-244:~$
```

2d. Installing IPython - Generate SSL certificate

Create a self-signed SSL certificate so that we can use HTTPS on the IPython Notebook

```
ubuntu@ip-10-178-132-244:~$ mkdir certificates
ubuntu@ip-10-178-132-244:~$ cd certificates/
ubuntu@ip-10-178-132-244:~/certificates$ openssl req -x509 -nodes -days 365 -newkey rsa:1024 -keyout mycert.pem -out mycert.pem
Generating a 1024 bit RSA private key
.....+++++
.....+++++
writing new private key to 'mycert.pem'
-----
You are about to be asked to enter information that will be incorporated
into your certificate request.
What you are about to enter is what is called a Distinguished Name or a DN.
There are quite a few fields but you can leave some blank
For some fields there will be a default value,
If you enter '.', the field will be left blank.
-----
Country Name (2 letter code) [AU]:US
State or Province Name (full name) [Some-State]:PA
Locality Name (eg, city) []:Ambler
Organization Name (eg, company) [Internet Widgits Pty Ltd]:COMPANY
Organizational Unit Name (eg, section) []:COMPANY
Common Name (e.g. server FQDN or YOUR name) []:COMPANY
Email Address []:COMPANY@COMPANY.COM
ubuntu@ip-10-178-132-244:~/certificates$ ls
mycert.pem
```

2e. Installing IPython - Modify nbserver profile

Navigate to the `profile_nbserver` directory, then modify the `ipython_notebook_config.py` file with your certificate location and password.

Place these commands at the top of the file; you don't need to uncomment any of the lines generated when nbserver profile was created.

A terminal window titled 'personal_ec2_keys — ubuntu@ip-10-178-132-244: ~/.ipython/profile_nbserver — ssh — 156x49'. The terminal shows the following commands and output:

```
ubuntu@ip-10-178-132-244:~$ cd .ipython
ubuntu@ip-10-178-132-244:~/.ipython$ ls
profile_default  profile_nbserver  README
ubuntu@ip-10-178-132-244:~/.ipython$ cd profile_nbserver/
ubuntu@ip-10-178-132-244:~/.ipython/profile_nbserver$ ls
ipython_config.py  ipython_nbconvert_config.py  ipython_notebook_config.py  log  pid  security  startup
ubuntu@ip-10-178-132-244:~/.ipython/profile_nbserver$ vim ipython_notebook_config.py
```

The contents of the `ipython_notebook_config.py` file are displayed, showing configuration for the IPython kernel and notebook. The configuration includes the kernel name, the location of the certificate file, the IP address, whether to open a browser, the password, and the port number.

```
# Configuration file for ipython-notebook.

c = get_config()

# Kernel config
c.IPKernelApp.pylab = 'inline' # if you want plotting support always

# Notebook config
c.NotebookApp.certfile = u'/home/ubuntu/certificates/mycert.pem'
c.NotebookApp.ip = '*'
c.NotebookApp.open_browser = False
c.NotebookApp.password = u'sha1:694bff492f56:e06ae3083855a56269446d2a386a557fb538ff6f'
# It is a good idea to put it on a known, fixed port
c.NotebookApp.port = 8888

#-----
```


2f. Installing IPython - Launch IPython Notebook

Launch IPython Notebook with the nbserver profile. At this point, we can now access IPython Notebook from our local browser!

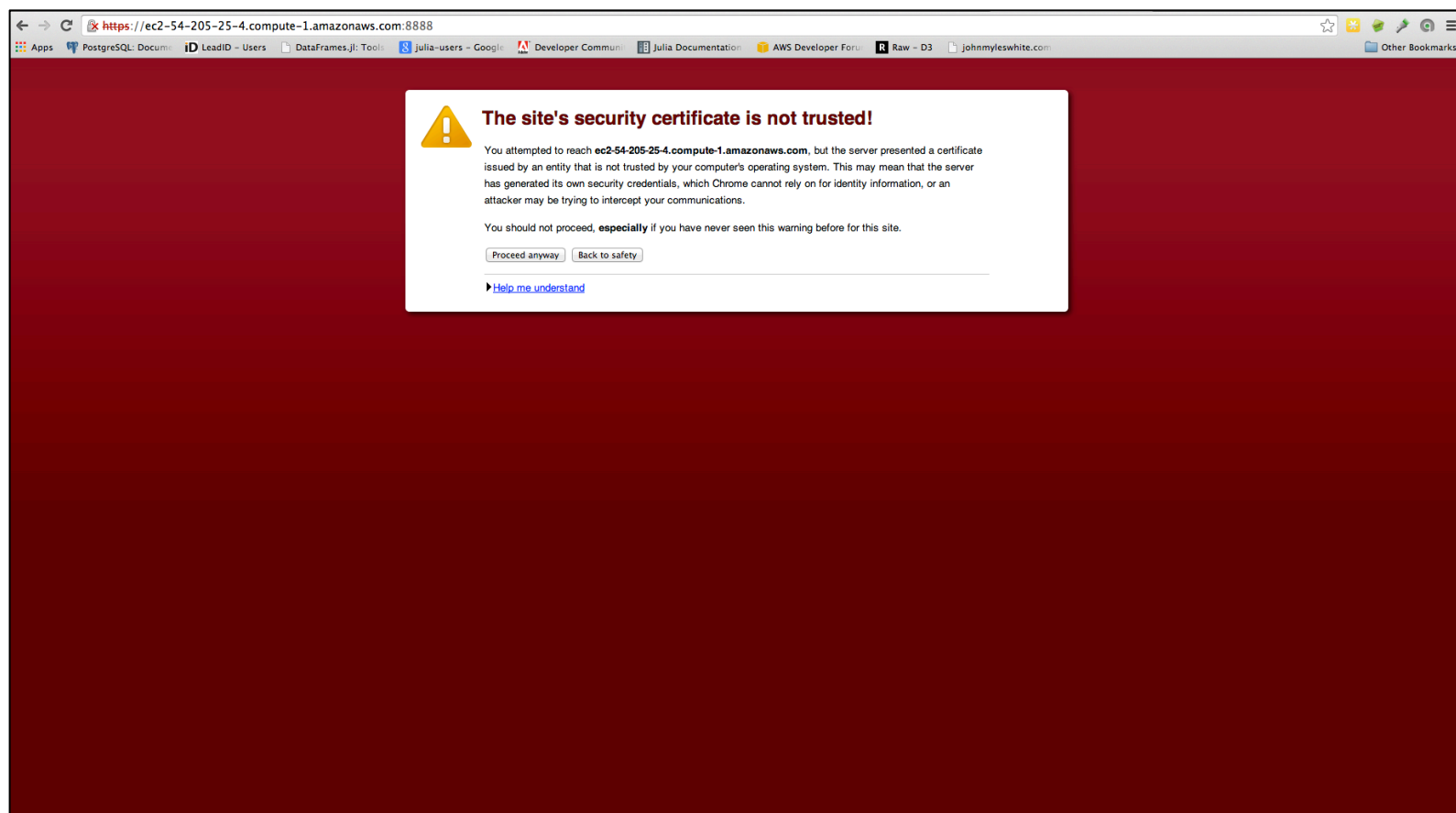
```
ubuntu@ip-10-178-132-244:~/ipython/profile_nbserver$ ipython notebook --profile=nbserver
2013-11-21 19:22:51.167 [NotebookApp] Using existing profile dir: u'/home/ubuntu/.ipython/profile_nbserver'
2013-11-21 19:22:51.173 [NotebookApp] Using MathJax from CDN: https://c328740.ssl.cf1.rackcdn.com/mathjax/latest/MathJax.js
2013-11-21 19:22:51.181 [NotebookApp] Serving notebooks from local directory: /home/ubuntu/.ipython/profile_nbserver
2013-11-21 19:22:51.181 [NotebookApp] The IPython Notebook is running at: https://[all ip addresses on your system]:8888/
2013-11-21 19:22:51.181 [NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
```

3. Using IPython Notebook from Local Browser

3a. Accessing IPython Notebook - SSL Warning

Use any modern browser to access the public DNS of your EC2 image. It is expected to see a warning, as we're using a self-signed SSL certificate

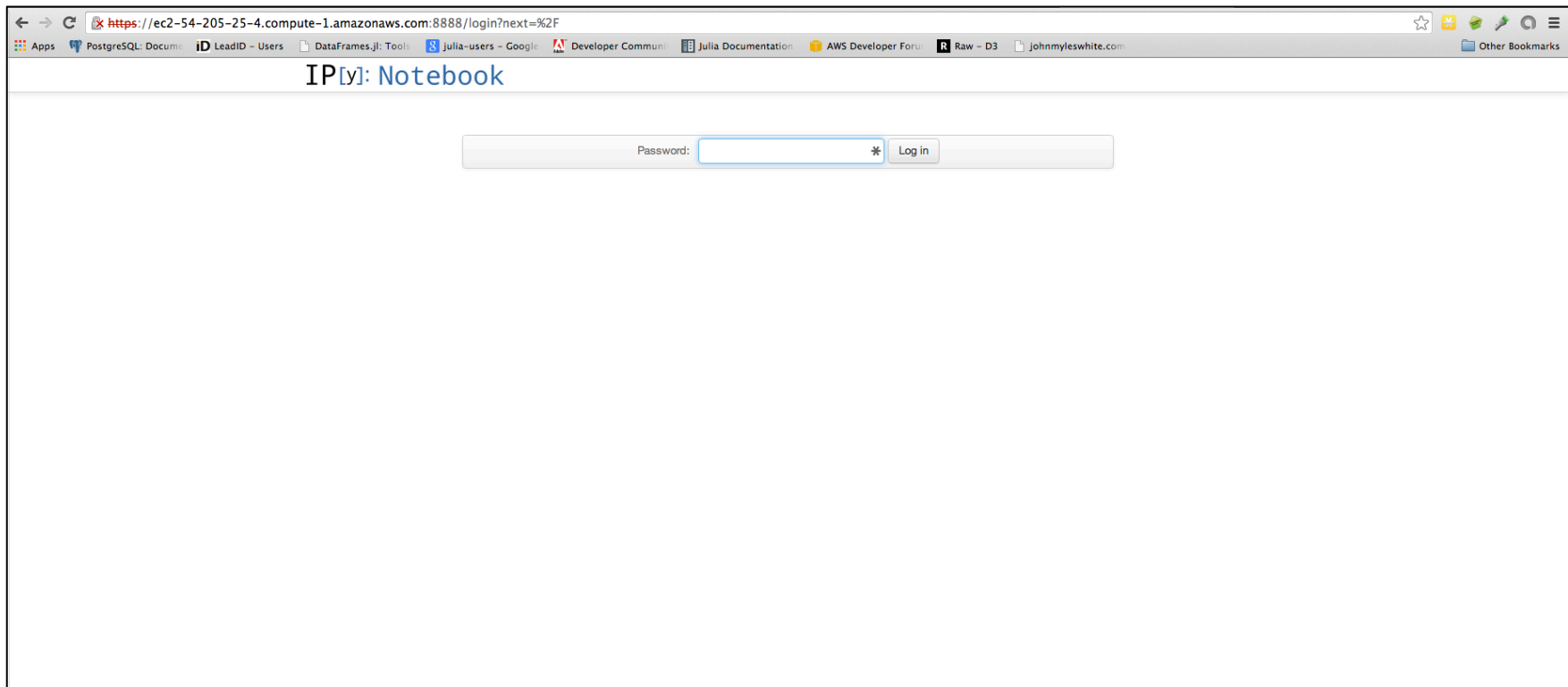
Ex: `https://ec2-54-205-25-4.compute-1.amazonaws.com:8888`



`http://randyzwitch.com`

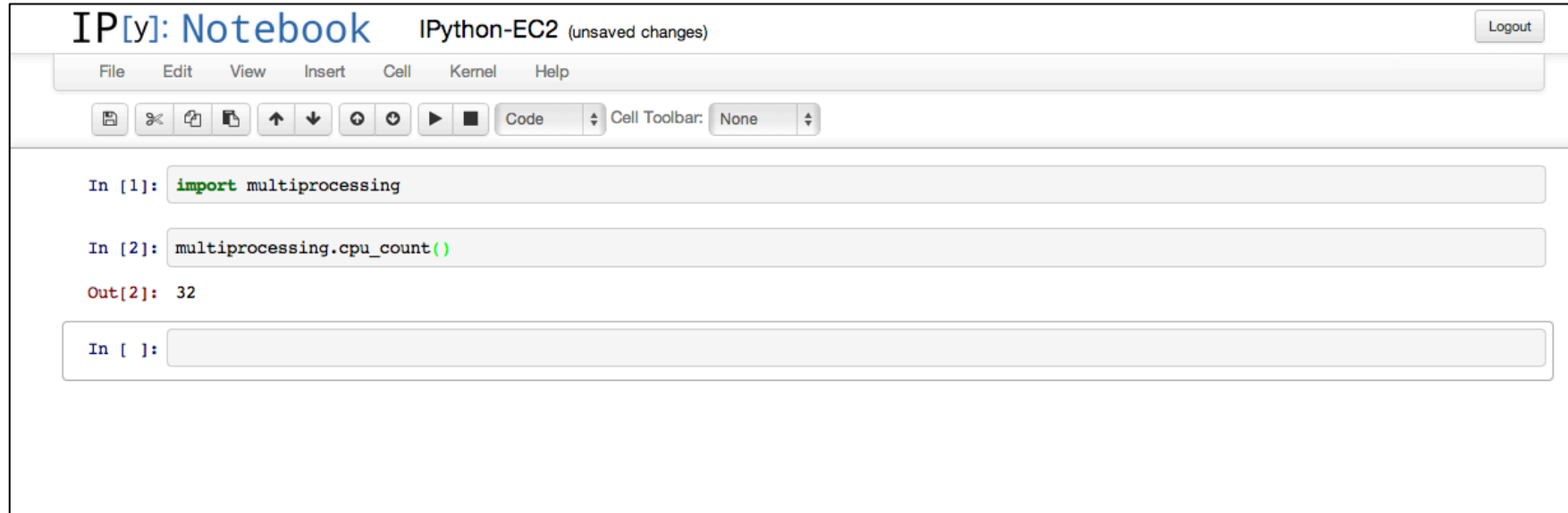
3b. Accessing IPython Notebook - Enter Password

Sign in using password that you set during the prior step (the actual password, not the SHA1 version)



3c. Accessing IPython Notebook – Success!

At this point, you've got a fully functional Python cluster environment running on EC2, which you are accessing from your local browser



3c. Accessing IPython Notebook – Use 32 cores for ML

Running a toy example from Scikit-Learn, we can specify use of 32 cores for the ExtraTreesClassifier

IP[y]: Notebook IPython-EC2 Last Checkpoint: Nov 21 14:42 (autosaved) Logout

File Edit View Insert Cell Kernel Help

Code Cell Toolbar: None

```
In [7]: #Example from http://scikit-learn.org/stable/auto_examples/ensemble/plot_forest_importances_faces.html
print(__doc__)

from time import time
import pylab as pl

from sklearn.datasets import fetch_olivetti_faces
from sklearn.ensemble import ExtraTreesClassifier

# Number of cores to use to perform parallel fitting of the forest model
# A value of -1 represents all cores
n_jobs = 32

# Load the faces dataset
data = fetch_olivetti_faces()
X = data.images.reshape((len(data.images), -1))
y = data.target

mask = y < 5 # Limit to 5 classes
X = X[mask]
y = y[mask]

# Build a forest and compute the pixel importances
print("Fitting ExtraTreesClassifier on faces data with %d cores..." % n_jobs)
t0 = time()
forest = ExtraTreesClassifier(n_estimators=1000,
                             max_features=128,
                             n_jobs=n_jobs,
                             random_state=0)

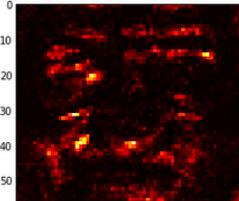
forest.fit(X, y)
print("done in %0.3fs" % (time() - t0))
importances = forest.feature_importances_
importances = importances.reshape(data.images[0].shape)

# Plot pixel importances
pl.matshow(importances, cmap=pl.cm.hot)
pl.title("Pixel importances with forests of trees")
pl.show()
```

Automatically created module for IPython interactive environment

```
Fitting ExtraTreesClassifier on faces data with 32 cores...
done in 1.097s
```

Pixel importances with forests of trees



<http://randyZWitch.com>