Setting Starcluster

Installing necessary dependency

Install starcluster by typing to terminal.

Terminal

```
$sudo apt-get install python-setuptools

$easy_install starcluster
```

Setting up config file

Next step is to create a config file.

On terminal type

$starcluster help

will prompt something like on the next

press 2 to create config file to that location

Terminal

```
$ starcluster help
StarCluster - (http://star.mit.edu/cluster)
Software Tools for Academics and Researchers (STAR)
Please submit bug reports to starcluster@mit.edu

cli.py:87 - ERROR - config file /home/user/.
starcluster/config does not exist

Options:
--------
[1] Show the StarCluster config template
[2] Write config template to /home/user/.starcluster/config
[q] Quit

Please enter your selection:
```

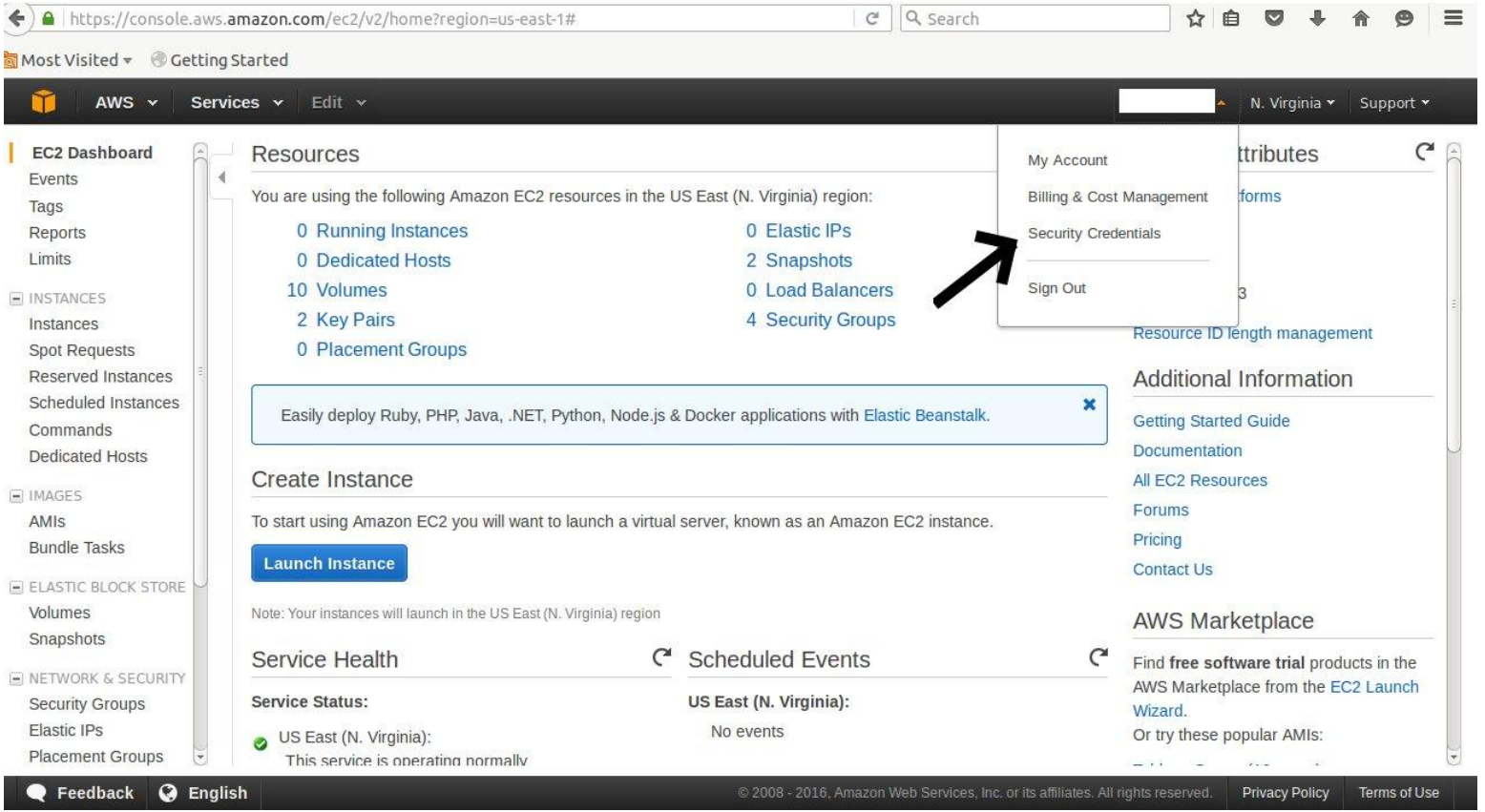
Setting up Config file

Next up is to get AWS account info into the config so program can access it.

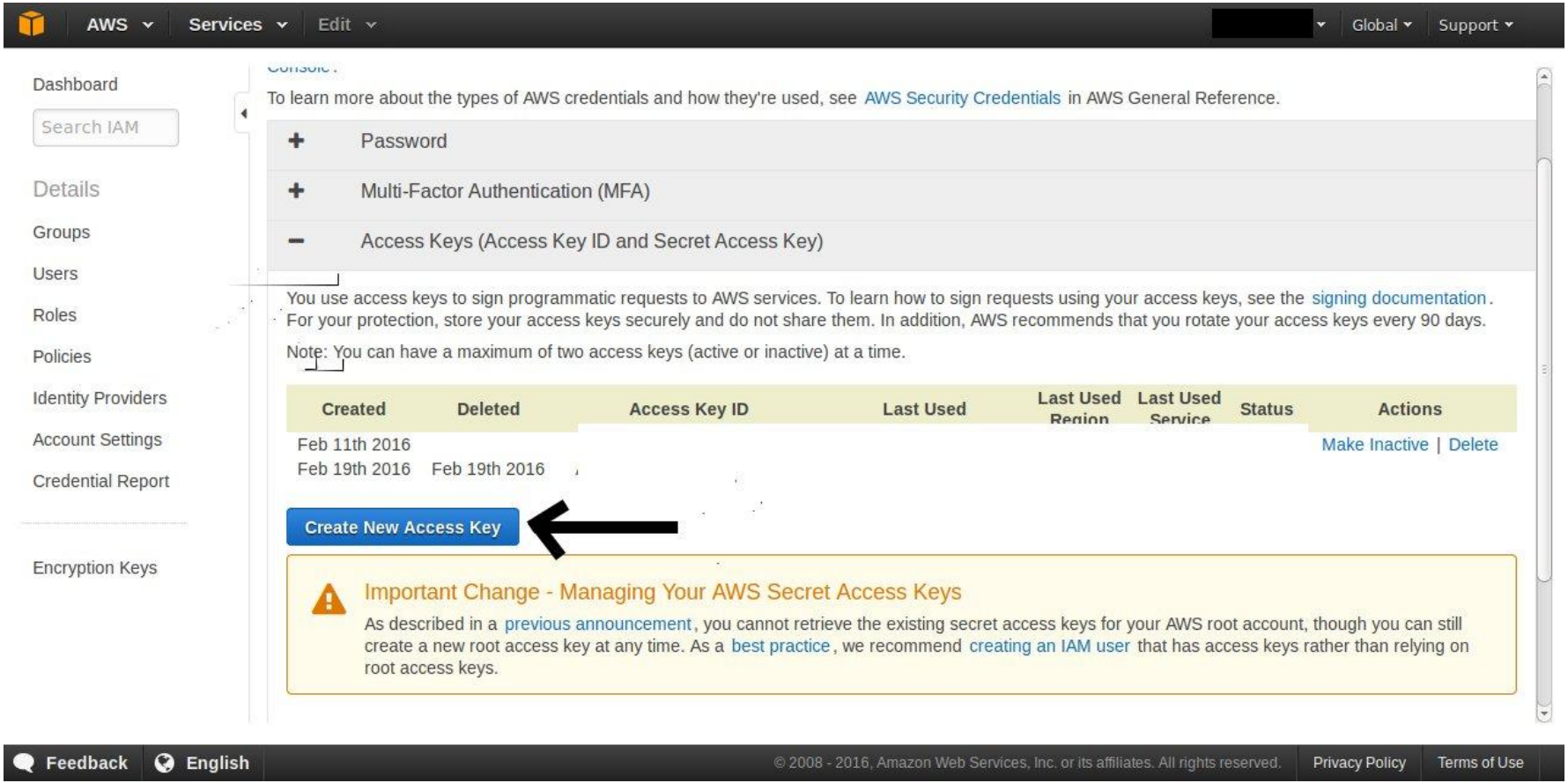
If you haven’t already, please create an account at

http://aws.amazon.com/

Setting Config file - getting AWS info



Setting Config file-getting AWS info



Setting Config - getting AWS info

Pressing create key would generate

AWS_ACCESS_KEY_ID and AWS_ACCESS_SECRET_KEY.

with access to all AWS service.

But this is limited to two key.

If account is shared, you can create a user by:

clicking user tab on left

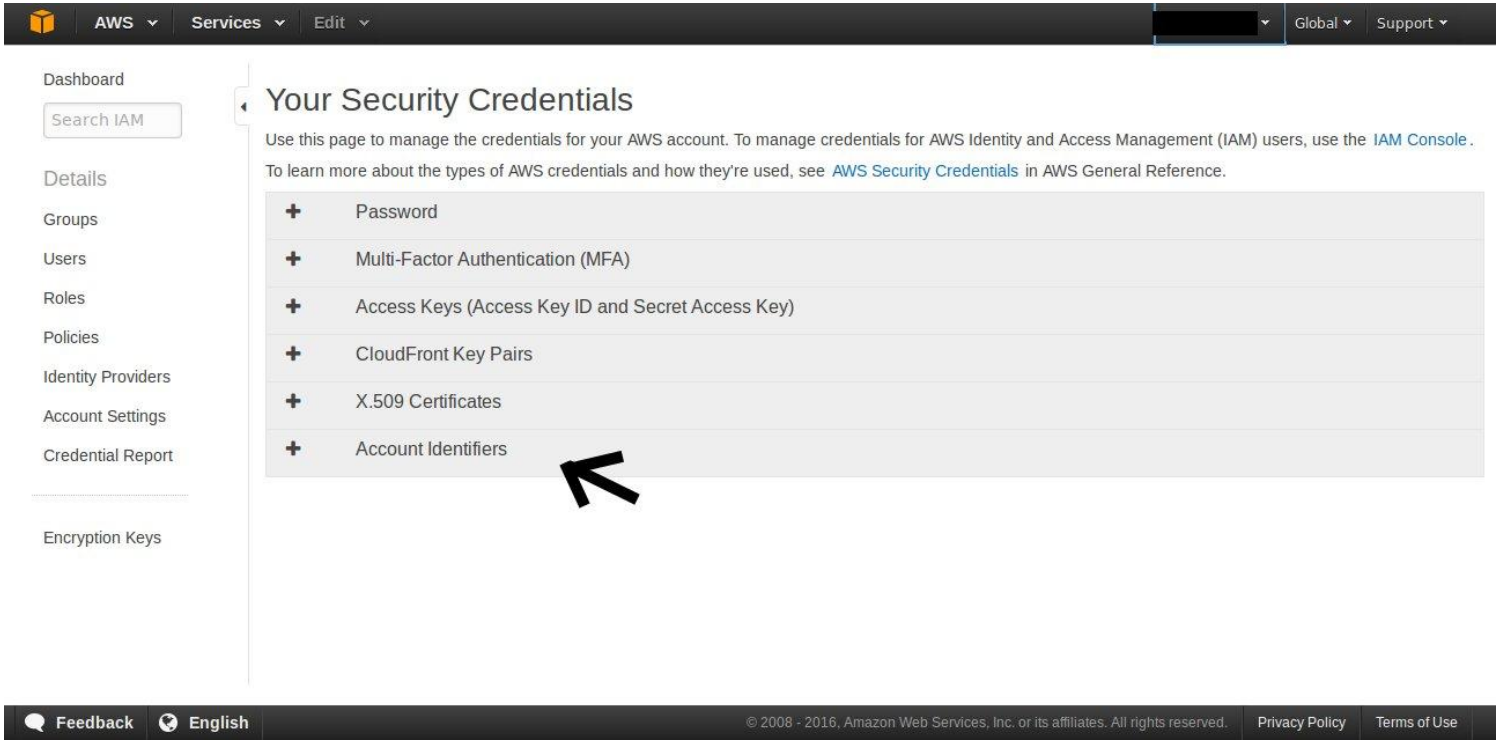
create user and filling in details

under security credential, create key to generate AWS_ACCESS_KEY_ID and AWS_ACCESS_SECRET_KEY. for the user

grant permissions for that user under permission tab (access to EC2, etc)

Setting Config - getting AWS info

ACCOUNT_ID can be found under account identifiers



Setting Config - getting AWS info

with the obtained info, edit $HOME/.starcluster/config file

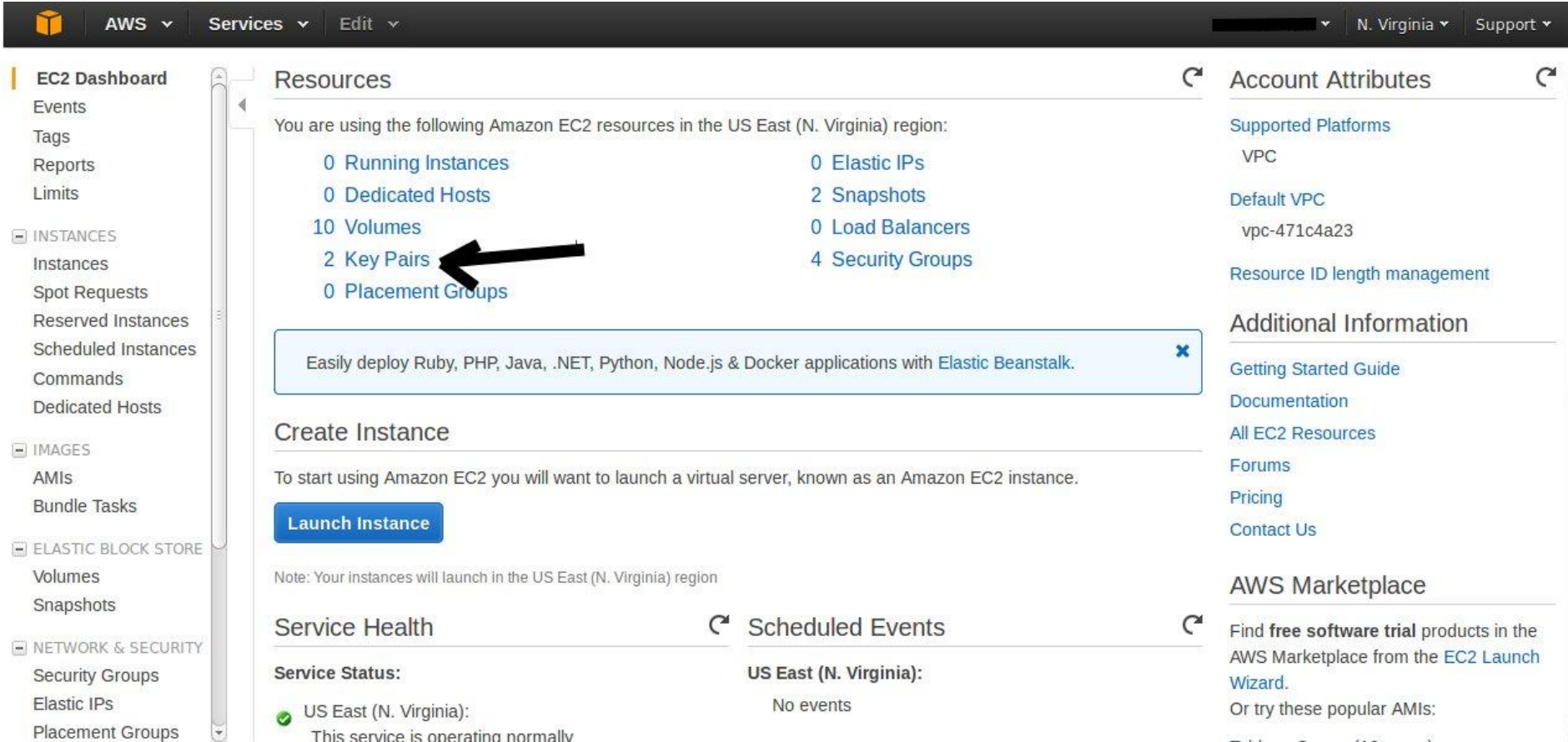
(or wherever config file was saved to).

fill in the information from AWS into appropriate location

[aws info]
aws_access_key_id = #your aws access key id here
aws_secret_access_key = #your secret aws access key here
aws_user_id = #your 12-digit aws user id here

Setting Config getting AWS info

Under EC2 dash board



Setting Config - getting AWS info

Create Key pair and save at some location.

edit $HOME/.starcluster/config or wherever config file was saved

[key mykey]
key_location = path_to_key

Setting config - getting AWS info

You can have multiple keypairs and their paths in config file as this is what is used to verify you have access to the cluster/images/etc made with that keypairs.

format follows the same

[key name_key]
key_location = path_to_key

be sure name_key matches actual key pair name as it may cause an error

Setting Config - cluster setting

cluster format looks like to the right in config file

KEYNAME is keypair that cluster uses
CLUSTER_SIZE is number of the nodes in cluster (it may be added later)
CLUSTER_USER is non-root user account
NODE_IMAGE_IDis image to use.
(you can only use images provided by starcluster to start unless you edit the image to work with starcluster)
NODE_INSTANCE_TYPE is instance of node to use. Please refer https://aws.amazon.com/ec2/instance-types/
PLUGINS is what to use at the starting the cluster

```
[cluster defaultcluster]
KEYNAME = mykey
CLUSTER_SIZE = 2
CLUSTER_USER = sgeadmin
NODE_IMAGE_ID = ami-6b211202
NODE_INSTANCE_TYPE = t2.micro
PLUGINS = ipcluster
```

Setting Config - Plugins

Plugins are called (usually) at starting the cluster to setup the cluster.

Example plugins is built-in ipcluster plugins which generates .json file and sets up the cluster so it could be easily used with IPython.parallel library

Package installer installs python packages indicated to all nodes.

#Sample plugin in config

[plugin ipcluster]

SETUP_CLASS = starcluster.plugins.ipcluster.IPCluster

[plugin package-installer]

SETUP_CLASS = starcluster.plugins.pypkginstaller.
PyPkgInstaller

PACKAGES = scikit-learn #, other packages

Starting cluster

If setup, you can start cluster by typing start command. Here is useful options you can send

-n instance image

-s number of cluster

-u cluster user

-o create-only (do not run starcluster add-ons etc)

-c cluster template to use

$starcluster start testcluster

Stop/terminating cluster

stop will just stop instances.

$starcluster stop clustername

terminate will delete the instances and their storage (EBS) so you are no longer charged for EBS as well

$starcluster terminate clustername