

Amazon Machine Image Setup

Brian M. Brost

2/9/2017

Launch an AMI via the Web Console

Follow these steps to launch a “stock” Amazon Machine Image (AMI) using the Amazon Web Services web console:

1. Login to AWS Console.
2. Select *EC2* (Elastic Compute Cloud) from the available services.
3. Select *Launch Instance* to enter launch wizard.
 - i. Choose an Amazon Machine Image — select *Ubuntu Server 16.04 LTS (HVM), SSD Volume Type* - *ami-7c803d1c* and proceed to next step at bottom.
 - ii. Choose an Instance Type — select *t2.micro (Free tier eligible)* and proceed to next step at bottom.
 - iii. Configure Instance Details — leave default values and proceed to next step at bottom.
 - iv. Add Storage — leave default values and proceed to next step at bottom.
 - v. Add Tags — leave default values and proceed to next step at bottom.
 - vi. Configure Security Group — in addition to SSH (22), consider adding the following rules (port range):
 - Custom TCP Rule (8787) for RStudio Server
 - HTTP (80)
 - HTTPS (443)Proceed to *Review and Launch* at bottom. *Note:* Does changing the source to “Anywhere” compromise security? Is “My IP” better, and if so, are IP addresses static?
4. Launch EC2 instance by selecting *Launch* at bottom.
5. Selecting an existing key pair or create a new key pair — either select an existing key pair or create a new key pair. Agree to the terms and select *Launch Instances* at bottom.

Connect to an EC2 Instance

Establish a remote connecting to an EC2 instance by first navigating to EC2 Instance Manager in the AWS Console (i.e., select *Instances* in left bar). Identify the EC2 instance in the list and select *Connect* at top. Access the EC2 instance using a standalone SSH client with the instructions provided or by following the sequence of steps outlined below:

1. Open SSH client (e.g., Mac OSX Terminal).
2. Modify the security key so it’s not publicly viewable:

```
chmod 400 /Users/brian.brost/Documents/sandbox/aws/testing.pem
```

3. Connect to the EC2 instance using Public DNS:

```
ssh -i "/Users/brian.brost/Documents/sandbox/aws/testing.pem"
ubuntu@ec2-35-161-94-108.us-west-2.compute.amazonaws.com
```

Note: The public DNS will need to be updated to reflect your specific instance.

Add a New User Account to the System

Follow these steps for adding a new user to the AMI and setting permissions to allow remote access:

1. Add a new user to the system:

```
sudo adduser noaa-usr
```

Input and confirm a password for the new user. Note that RStudio Server requires a username and password for web-based access using port 8787. To remove a user from the system:

```
sudo userdel -r noaa-usr
```

2. Establishing remote access to the new user account requires creating a .ssh directory, copying the public key to this directory, changing permissions for each, and changing the ownership of the new folder:

```
sudo mkdir /home/noaa-usr/.ssh/  
sudo chmod 700 /home/noaa-usr/.ssh/  
sudo cp /home/ubuntu/.ssh/authorized_keys /home/noaa-usr/.ssh/  
sudo chmod 600 /home/noaa-usr/.ssh/authorized_keys  
sudo chown -R noaa-usr /home/noaa-usr/
```

3. Update and upgrade package lists:

```
sudo apt-get update  
sudo apt-get upgrade
```

4. Transfer a file to AWS to confirm permissions are set correctly:

```
scp -i /Users/brian.brost/Documents/sandbox/aws/testing.pem  
/Users/brian.brost/Documents/sandbox/aws/testing.R  
noaa-usr@ec2-52-37-94-54.us-west-2.compute.amazonaws.com:
```

This will upload a file to /home/noaa-usr/ on the AMI. Can upload to other directories by inserting that location after the colon at the end of the linux command.

For additional details concerning user accounts:

- <http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/managing-users.html>
- <http://superuser.com/questions/286831/how-do-i-copy-files-into-var-www-with-winscp>
- others?

Install R and OpenBlas Library on AMI

First, update the repository sources to include CRAN:

1. Open the sources.list file:

```
sudo nano /etc/apt/sources.list.d/sources.list
```

2. Add the R repository location to the end of the sources.list file:

```
deb http://cran.rstudio.com/bin/linux/ubuntu trusty/
```

3. Save the updated sources.list file (ctrl-O and enter) and exit (ctrl-X).
4. Update the system to reflect the addition of CRAN:

```
sudo apt-get update
```

For additional details about adding the CRAN repository:

- <https://cran.rstudio.com/bin/linux/ubuntu/README.html>
- <https://help.ubuntu.com/community/Repositories/CommandLine>
- others?

Follow these steps to install R and the OpenBlas library:

1. Install latest version of R:

```
sudo apt-get install r-base
```

2. Install multi-threaded OpenBlas library to optimize linear algebra operations.

```
sudo apt-get install libopenblas-base
```

3. Test performance of optimized linear algebra library. Open R:

```
R
```

Run benchmark test in R:

```
source("http://r.research.att.com/benchmarks/R-benchmark-25.R")
```

Install RStudio Server (Optional)

1. Install gdebi to allow installation of local deb packages:

```
sudo apt-get install gdebi-core
```

2. Optional(?) package:

```
sudo apt-get install libapparmor1
```

3. Change to writeable directory:

```
cd /tmp
```

4. Download the RStudio Server deb package (check for latest version):

```
wget download2.rstudio.org/rstudio-server-1.0.136-amd64.deb
```

Note: that the “https://” portion of the web address provided on the rstudio website has been removed here.

5. Install RStudio Server:

```
sudo gdebi rstudio-server-1.0.136-amd64.deb
```

6. Remove installation files:

```
sudo rm /tmp/rstudio-server-1.0.136-amd64.deb
```

7. Open RStudio Server by opening a web browser and navigating to the Public DNS of the AMI on port 8787, e.g.,

```
ec2-52-38-46-71.us-west-2.compute.amazonaws.com:8787
```

Note that the Public DNS used here will need to be updated to reflect the DNS obtained in Step 4 of Connect to EC2 Instance above. Enter the username and password established during the New User Account setup to proceed.

For additional details about adding RStudio Server:

- <https://www.rstudio.com/products/rstudio/download-server/>

Install spatial libraries

1. Install dev package for Proj.4:

```
sudo apt-get install libproj-dev proj-data proj-bin
```

2. Install dev package for Geos:

```
sudo apt-get install libgeos-dev
```

3. Install GDAL. Three methods for installing GDAL are provided below: (1) binary package, (2) UbuntuGIS PPA, and (3) from source code. Installing from source code may be the best option for ensuring installation of the latest version of GDAL.

- i. GDAL binary package — the version of GDAL in the repository is probably out-of-date. Therefore, instructions on the GRASS-wiki to use the pre-compiled dev package, i.e.,

```
sudo apt-get install libgdal-dev
sudo apt-get install python-gdal gdal-bin
```

may install an old version of GDAL.

- ii. UbuntuGIS PPA binary package — the GDAL binary from UbuntuGIS PPA may be up-to-date. If so, add the UbuntuGIS PPA to the system sources, update the package lists, and install the GDAL binary:

```
sudo add-apt-repository ppa:ubuntugis/ppa
sudo apt-get update
sudo apt-get install gdal-bin
```

See <http://www.sarasafavi.com/installing-gdal-ogr-on-ubuntu.html> for additional details.

- iii. Compile from source — This may be the only sure option for obtaining the latest version of GDAL. First, download the source code:

```
cd /tmp
sudo wget http://download.osgeo.org/gdal/2.1.3/gdal-2.1.3.tar.gz
sudo tar xvfz gdal-2.1.3.tar.gz
cd gdal-2.1.3
```

Next, configure, build, and install the software:

```
sudo ./configure
sudo make
sudo make install

sudo ldconfig
```

Finally, delete the install files:

```
sudo rm -r /tmp/gdal-2.1.3
sudo rm /tmp/gdal-2.1.3.tar.gz
```

For additional details about adding the CRAN repository:

- <http://gis.stackexchange.com/questions/21728/installing-gdal-1-9-0-on-linux-ubuntu-server-10-04/21739#21739>
- <http://trac.osgeo.org/gdal/wiki/BuildingOnUnix>
- https://grasswiki.osgeo.org/wiki/Compile_and_Install_Ubuntu#PROJ.4

Also note that the Python dependencies provided in the web links may not be required and are thus omitted above.

4. Verify GDAL installation:

```
gdal-config
```

```
--version 2.1.3
```

```
gdal-config
```

```
--libs -L/usr/local/lib -lgdal
```

```
gdal-config --cflags
```

```
--I/usr/local/include
```

Install R packages

Install R packages from the command line:

```
wget http://cran.r-project.org/src/contrib/Archive/sp/sp_0.9-37.tar.gz  
sudo CMD INSTALL sp_0.9-37.tar.gz
```

Install R packages from the R prompt:

```
sudo R
```

```
chooseCRANmirror(ind=53)  # Mirror in OR  
install.packages("rgdal")  
install.packages("rgeos")  
install.packages("raster")  
install.packages("sp")  
install.packages("gdistance")  
install.packages("devtools")  
devtools::install_github("edzer/sfr")  # Error on install
```