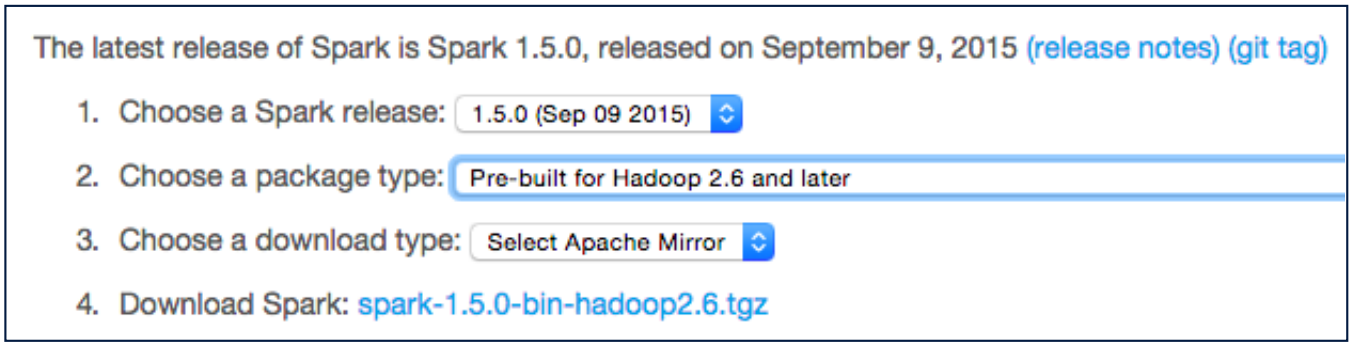
Here we will download **Spark** binaries on our AWS instance, fire up some **Spark** clusters, and try out **RStudio Server**.

- Installing Spark

Installing Spark

This instance will be used to launch our Master and Dependent clusters. Let's get the download URL path for **Spark** (from the downloads page for the latest version):



We'll use `wget` to download Spark directly into our instance, paste the following commands into the terminal:

```
# download spark
sudo wget http://d3kbcqa49mib13.cloudfront.net/spark-1.5.0-
bin-hadoop2.6.tgz

# unpack spark tgz file
sudo tar zxvf  spark-1.5.0-bin-hadoop2.6.tgz
```

Now we need to download our `PEM` key-pair file over to the instance as well. Exit out of the EC2 session and use the `scp command` (Secure Copy):

```
# get out of current session and back to local machine
exit

# upload pem file in same directory on the EC2 instance (swap
the dashed IP with your instance's)
scp -i  udemy.pem udemy.pem ec2-user@ec2-52-88-115-167.us-
west-2.compute.amazonaws.com:/home/ec2-user/spark-1.5.0-bin-
hadoop2.6/ec2
```

Now we go back to the EC2 box:

```
ssh -i "udemy.pem" ec2-user@52.88.115.167

# navigate to the ec2 folder:
cd spark-1.5.0-bin-hadoop2.6/ec2

# change file permissions
sudo chmod 400 udemy.pem
```

We also need to export our access keys:

```
Set access keys:
export AWS_ACCESS_KEY_ID=...yours here...
export AWS_SECRET_ACCESS_KEY=...yours here...
```