Installing Apache Zeppelin on UCB W205 AMIs

We'd like to have a friendly, notebook interface on our EC2 instances. To do this, we'll use Apache Zeppelin, just as we do in the Vagrant environment. To set up Zeppelin, follow these steps.

Preliminaries

1. Start your EC2 instance, attach your EBS volume
2. **Make sure your security group has port 8080 open**
3. Log into your EC2 instance
4. As root, mount the EBS volume under /data

Getting Maven and Zeppelin

We need to use Apache Maven to build Zeppelin. We want to do this as our named user, rather than root. In general, builds should happen as a named user, but installs may happen as root.

First, we need to get Apache Maven. Follow these steps

```
1. su - <named user>
2. cd /data
3. wget http://www.trieuvan.com/apache/maven/maven-
   3/3.3.3/binaries/apache-maven-3.3.3-bin.tar.gz

4. tar xvzf apache-maven-3.3.3-bin.tar.gz
5. export MAVEN_VERSION=3.3.3
6. export MAVEN_HOME=/data/apache-maven-$MAVEN_VERSION
```

Now we have Maven installed and an environment variable pointing to it. Next we need to get the latest version of Apache Zeppelin.

```
1. cd /data
2. git clone https://github.com/apache/incubator-zeppelin.git
   zeppelin
3. cd zeppelin
4. ${MAVEN_HOME}/bin/mvn clean package -Pspark-1.4 -
   Dhadoop.version=2.6.0 -Phadoop-2.6 –DskipTests

5. cp conf/zeppelin-env.sh.template conf/zeppelin-env.sh
```

We need to edit conf/zeppelin.sh to provide Zeppelin more memory. To do this, open conf/zeppelin.sh in a text editor and change the following:
```
# export ZEPPELIN_MEM
```

to
```
export ZEPPELIN_MEM="-Xmx2048m"
```

We can now start zeppelin. In the future, when you boot your instance and mount /data, you can also start zeppelin.

1. `cd /data/zeppelin`
2. `bin/zeppelin.sh`

Use your browser to go to: your-ec2-address.com:8080