Resume Summary (Cloud Project)
Team ID - 19

The document provides details on how to setup Hadoop and web2py which is required to run our project successfully.

Hadoop Installation

Hadoop is supported by GNU/Linux platform and its flavors. Therefore, we need a Linux operating system for setting up Hadoop environment.

- **Pre-Installation step**
  - Creating User: It is recommended to create a separate user for Hadoop to isolate Hadoop file system from Unix file system.
  - Follow the steps given below to create a user. Login as root
    - ```
      # useradd hadoop
      # passwd hadoop
         New passwd:
         Retype new passwd
      ```

- **Installing SSH server: Hadoop uses SSH to login to localhost**
  - Hadoop uses SSH to login to localhost
    - ```
      sudo apt-get install openssh-server
      ```
  - It is required to provide public/private key pair for a Hadoop user and share it with different users. The following commands are used for generating a key value pair using SSH.
    - ```
      $ ssh-keygen -t rsa -P ""
      ```
    - ```
      $ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
      ```

- **Installing Java**

  - Java is the main prerequisite for Hadoop. First step is verify java version
    - ```
      $ java -version
      ```

  - If everything is in order, it will give you the following output.
      ```

```
java version "1.7.0_71"
Java(TM) SE Runtime Environment (build 1.7.0_71-b13)
Java HotSpot(TM) Client VM (build 25.0-b02, mixed mode)
```

If java not installed follow given steps:

- Download Java and extract the tar using following commands
  - ```
    $ cd Downloads/
    $ ls
    jdk-7u71-linux-x64.gz
    $ tar zxf jdk-7u71-linux-x64.gz
    $ ls
    jdk1.7.0_71   jdk-7u71-linux-x64.gz
    ```
- Make java available to all users
  - ```
    mv jdk1.7.0_71 /usr/local/
    ```
- Set up the PATH and JAVA_HOME variables by adding following commands in ~/.bashrc file
  - ```
    export JAVA_HOME=/usr/local/jdk1.7.0_71
    export PATH=$PATH:$JAVA_HOME/bin
    ```
- Apply changes in the current running system
  - ```
    $ source ~/.bashrc
    ```
- Configure java alternatives
  - ```
    # alternatives --install /usr/bin/java java usr/local/java/bin/java 2
    # alternatives --install /usr/bin/javac javac usr/local/java/bin/javac 2
    # alternatives --install /usr/bin/jar jar usr/local/java/bin/jar 2
    # alternatives --set java usr/local/java/bin/java
    # alternatives --set javac usr/local/java/bin/javac
    # alternatives --set jar usr/local/java/bin/jar
    ```

Verify the java -version command from terminal as explained above

Installing Hadoop

➢ Downloading Hadoop
    ```

- wget
  http://apache.claz.org/hadoop/common/hadoop-2.4.1/hadoop-2.4.1.tar.gz
- tar xzf hadoop-2.4.1.tar.gz
- mv hadoop-2.4.1/* to hadoop/

➢ Setting up Hadoop

- ```
  export HADOOP_HOME=/usr/local/hadoop
  export HADOOP_MAPRED_HOME=$HADOOP_HOME
  export HADOOP_COMMON_HOME=$HADOOP_HOME
  export HADOOP_HDFS_HOME=$HADOOP_HOME
  export YARN_HOME=$HADOOP_HOME
  export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
  export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
  export HADOOP_INSTALL=$HADOOP_HOME
  source ~/.bashrc
  ```

➢ Hadoop Configuration

- ```
  $ cd $HADOOP_HOME/etc/hadoop
  export JAVA_HOME=/usr/local/jdk1.7.0_71
  ```

- Updating core-site.xml
  - ```
    <configuration>
       <property>
          <name>fs.default.name </name>
          <value> hdfs://localhost:9000 </value>
       </property>
    ```

- Updating mapred-site.xml. This file is used to specify which MapReduce framework we are using
  - ```
    <configuration>
       <property>
         <name>mapreduce.framework.name</name>
         <value>yarn</value>
       </property>
    </configuration>
    ```

- Updating mapred-site.xml. The hdfs-site.xml file contains information such as the value of replication data, namenode path, and datanode paths of your local file systems. Let us assume following data
  - ```
    <configuration>
       <property>
         <name>dfs.replication</name>
         <value>1</value>
         <description>
           Default block replication. The actual number of
    ```
    ```

```
            replications can be specified when the file is created.
            the default is used if replication is not specified in
            create time.
        </description>
    </property>
</configuration>
```

- Formatting namenode and start hadoop
  - `hadoop namenode -format`
  - `start-dfs.sh`
  - `start-mapred.sh`

- Check if hadoop is successfully installed by running sample example
  - `hadoop jar hadoop-examples-1.1.2.jar pi 3 10`

➢ Verifying whether hadoop is working fine

- `$ hadoop version`

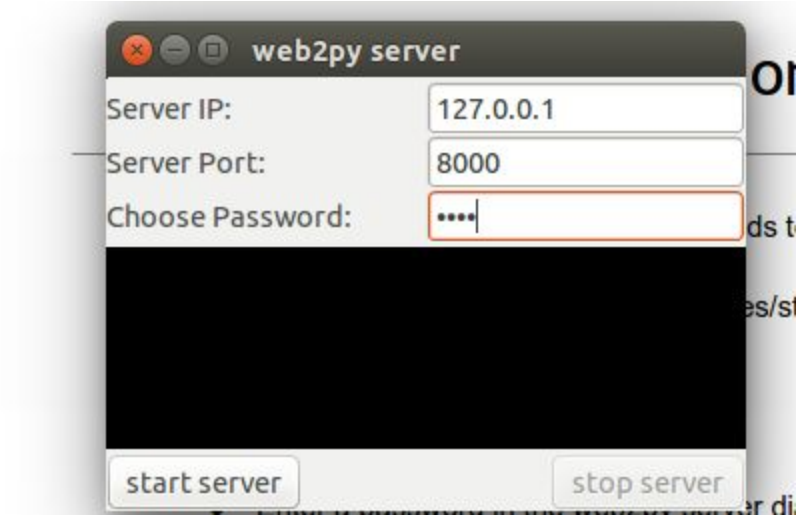
Web2py Installation and Setup

➢ Installing web2py: Run the following commands to install web2py on your local machine.

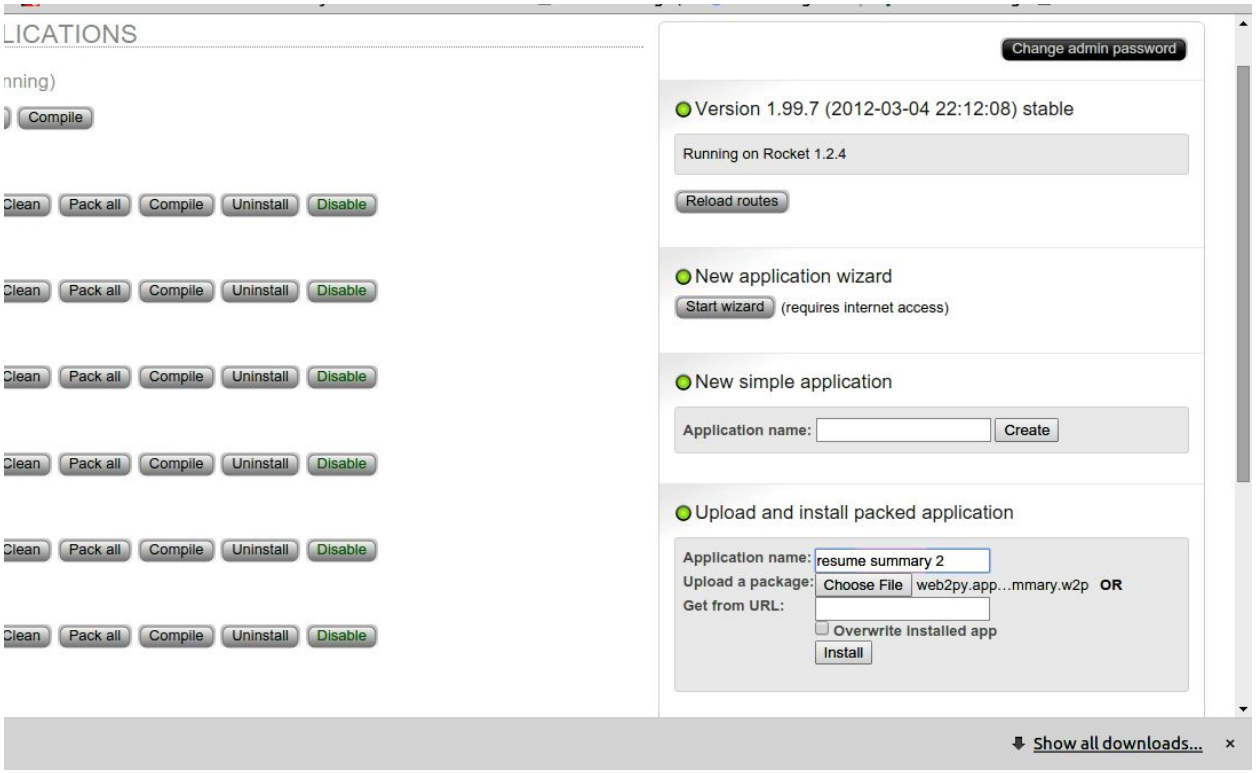
- wget http://www.web2py.com/examples/static/web2py_src.zip
- unzip web2py_src.zip
- cd web2py
- python web2py.py

➢ Deploying on the server:

- Enter a password in the web2py server dialog box.



- Upload the web package(web2py.app.resume_summary.w2p) and provide a suitable name for the application. Click on install button. Your application will be installed and will be ready to use.


LICATIONS
nning)
Compile
Clean Pack all Compile Uninstall Disable
Clean Pack all Compile Uninstall Disable
Clean Pack all Compile Uninstall Disable
Clean Pack all Compile Uninstall Disable
Clean Pack all Compile Uninstall Disable
Clean Pack all Compile Uninstall Disable
Change admin password
Version 1.99.7 (2012-03-04 22:12:08) stable
Running on Rocket 1.2.4
Reload routes
New application wizard
Start wizard (requires internet access)
New simple application
Application name: Create
Upload and install packed application
Application name: resume summary 2
Upload a package: Choose File web2py.app…mmary.w2p OR
Get from URL:
Overwrite installed app
Install
Show all downloads...
×