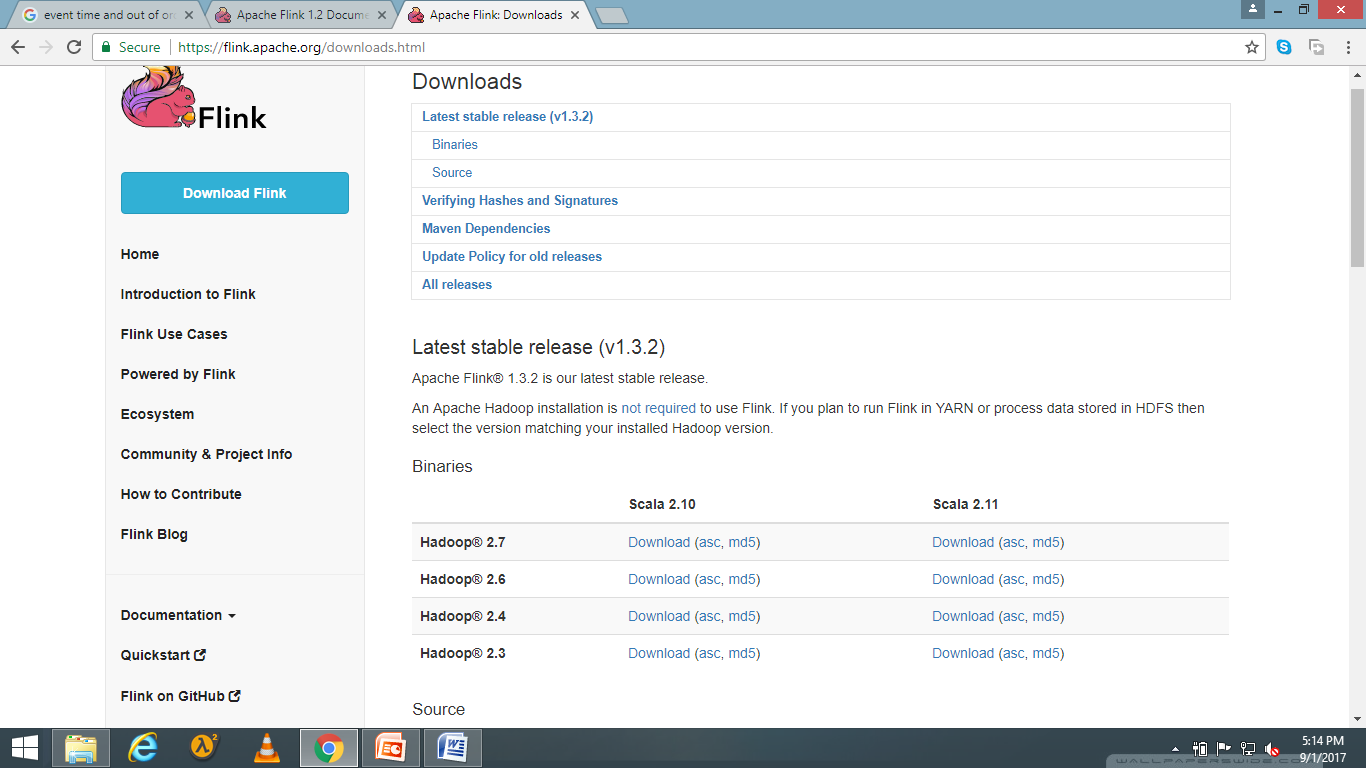
**How to install Apache Flink on your system**

* Download apache flink from the website according to your scala and hadoop version.



* After downloading, untar the file using command

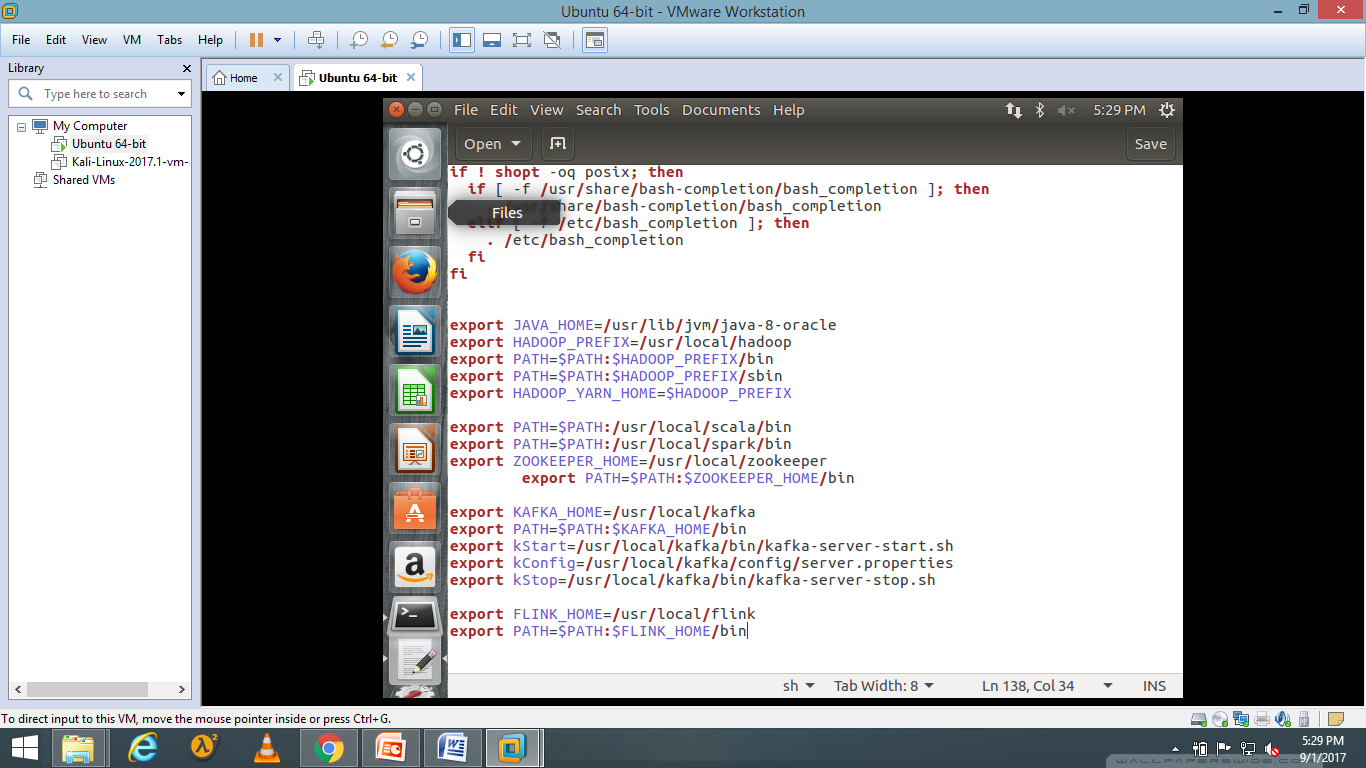
**tar –xzf flink.1.3.2-bin-hadoop27-scala\_2.11.tgz**

* Now rename it using **sudo mv flink-1.3.2 flink**
* Now move the file from the current folder to user local using

**sudo mv flink /usr/local/**

* Now provide a path in the bashrc so that you don’t need to visit it everytime you run Flink

**gedit .bashrc**



**source .bashrc**

* Now to check the commands type

**cd /usr/local/flink/bin;ls**

* Start it on localmode using

**start-local.sh**

**Deloying Apache Flink on Cluster**

**Configure SSH**

**SSH means secured shell which is used for the remote login. We can login to a remote machine using SSH. Now we need to configure passwordless SSH. Passwordless SSH means without a password we can login to a remote machine. Password less SSH setup is required for remote script invocation. Automatically remotely master will start the demons on slaves.**

**a. Install Open SSH Server-Client**

|  |  |
| --- | --- |
| **1** | **$ sudo apt-get install openssh-server openssh-client** |

**b. Generate Key Pairs**

|  |  |
| --- | --- |
| **1** | **$ ssh-keygen -t rsa -P ""** |

**It will ask “Enter the name of file in which to save the key (/home/dataflair/.ssh/id\_rsa):” let it be default, don’t specify any path just press “Enter”. Now it will be available in the default path i.e. “.ssh”. To check the default path use command “$ls .ssh/” and you will see that two files are created “id\_rsa” which is a private key and “id\_rsa.pub” which is a public key.**

**c. Configure password-less SSH**

**Copy the contents of “id\_rsa.pub” of master into the “authorized\_keys” files of all the slaves and master**

|  |  |
| --- | --- |
| **1** | **$ cat $HOME/.ssh/id\_rsa.pub >> $HOME/.ssh/authorized\_keys** |

**d. Check by SSH to all the hosts**

|  |  |
| --- | --- |
| **1** | **$ ssh localhost** |
| **1** | **$ ssh <SLAVE-IP>** |

**It should not ask for any password and you can easily get logged into remote machine since we have configured passwordless SSH**

**2.3. Install Apache Flink in Cluster Mode**

**I. Install Flink on Master**

**a. Download the Flink Setup**

**Download the Flink Setup from its official website**[**http://flink.apache.org/downloads.html**](http://flink.apache.org/downloads.html)

**b. Untar the file**

**In order to extract all the contents of compressed Apache Flink file package, use the below command:**

|  |  |
| --- | --- |
| **1** | **dataflair@ubuntu:~$ tar xzf flink-1.1.3-bin-hadoop26-scala\_2.11.tgz** |

**c. Rename the directory**

|  |  |
| --- | --- |
| **1** | **dataflair@ubuntu:~$ mv flink-1.1.3/ flink** |

**d. Setup Configuration**

**i. Go to Flink conf directory**

|  |  |
| --- | --- |
| **1** | **dataflair@ubuntu:~$  cd  flink** |
| **1** | **dataflair@ubuntu:~/flink$  cd  conf** | |

**ii. Add the entry of Master**

**Choose a master node (JobManager) and set the jobmanager.rpc.address in conf/flink-conf.yaml to its IP or hostname. Make sure that all nodes in your cluster have the same jobmanager.rpc.address configured.**

|  |  |
| --- | --- |
| **1**  **2** | **dataflair@ubuntu:~/flink/conf$  nano flink-conf.yaml**  **Add this line: jobmanager.rpc.address: 192.168.1.3** |

**iii. Add the entry of all the Slaves**

**Add the IPs or hostnames (one per line) of all worker nodes (TaskManager) to the slaves files in conf/slaves. To configure file use the following command.**

|  |  |
| --- | --- |
| **1** | **dataflair@ubuntu:~/flink/conf$  nano slaves** |

**Enter ip addresses like this –  
192.168.1.4  
192.168.1.5**

**II. Install Flink on Slaves**

**a. Copy configured setup from master to all the slaves**

**We will create a tar of configured Flink setup and copy it on all the slaves.**

**i. Create tar-ball of configured setup:**

|  |  |
| --- | --- |
| **1** | **$ tar czf flink.tar.gz flink** |

***NOTE: Run this command on Master***

**ii. Copy the configured tar-ball on all the slaves**

|  |  |
| --- | --- |
| **1**  **2** | **$ scp flink.tar.gz 192.168.1.4:~**  **$ scp flink.tar.gz 192.168.1.5:~** |

***NOTE: Run this command on Master***

**iii. Un-tar configured hadoop setup on all the slaves**

|  |  |
| --- | --- |
| **1** | **$ tar xzf flink.tar.gz** |

***NOTE: Run this command on all the slaves***

**2.4. Start that Flink Cluster**

**I. Start the cluster**

**To start the cluster run below script, it will start all the daemons running on master as well as slaves.**

|  |  |
| --- | --- |
| **1** | **start-cluster.sh** |

***NOTE: Run this command on Master***

**II. Check whether services have been started**

**a. Check daemons on Master**

|  |  |
| --- | --- |
| **1**  **2** | **$ jps**  **JobManager** |

**b. Check daemons on Slaves**

|  |  |
| --- | --- |
| **1**  **2** | **$ jps**  **TaskManager** |