### Spark部署

### Master 上安装 scala

下载scala-2.11.1

解压 tar zxvf scala-2.11.1.tar.gz -C /home/hadoop/

分发到各个节点

scp -r /home/hadoop/scala-2.11.1 hadoop@slave1:~/

scp -r /home/hadoop/scala-2.11.1 hadoop@slave2:~/

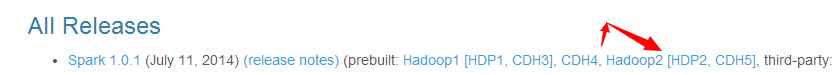
每个节点上配置 vim /etc/profile

export SCALA\_HOME=/home/hadoop/scala-2.11.1  
export PATH=.:$PATH:$SCALA\_HOME/bin

### master安装spark

下载spark 1.0.1

http://spark.apache.org/downloads.html



选择正确的版本。 [Hadoop2 [HDP2, CDH5]](http://d3kbcqa49mib13.cloudfront.net/spark-1.0.1-bin-hadoop2.tgz)

解压到home/hadoop/目录下

添加计算节点

Vim spark目录/conf/slaves

slave1

slave2

修改spark-env.sh

cp spark-env.sh.template spark-env.sh

vi spark-env.sh

添加如下信息

export SCALA\_HOME=/home/hadoop/scala-2.11.1

export JAVA\_HOME=/usr/lib/jvm/jdk1.7.0\_45

export HADOOP\_HOME=/usr/local/hadoop

export SPARK\_MASTER\_IP=172.18.4.70

分发给slave节点

scp -r /home/hadoop/scala-2.11.1 hadoop@slave1:~/

scp -r /home/hadoop/scala-2.11.1 hadoop@slave2:~/

vim /etc/profile

export SPARK\_HOME=/home/haduser/ spark-1.0.1-bin-2.2.0/  
export PATH=.:$PATH:$SPARK\_HOME/bin:$SPARK\_HOME:sbin

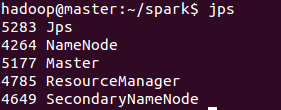
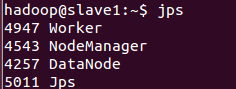
启动spark

cd /home/hadoop/spark-1.0.1-bin-2.2.0/sbin

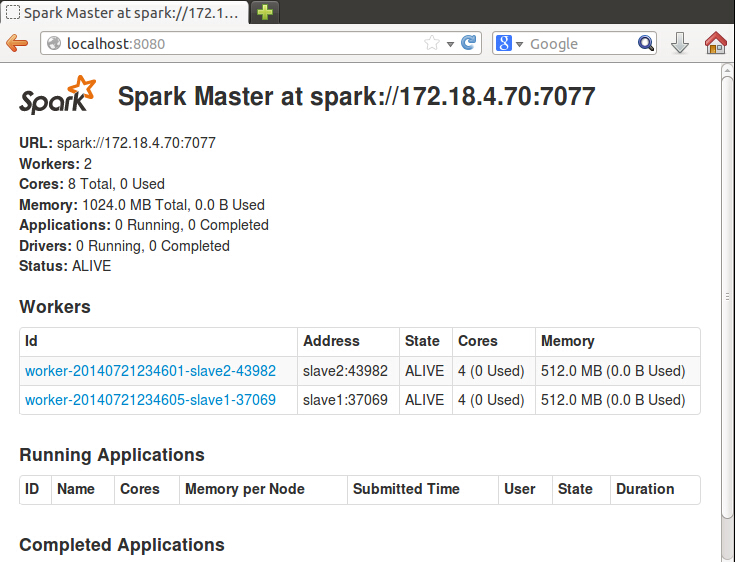
./start-all.sh

检测进程是否启动

$ jps

浏览master的web UI(默认http://localhost:8080). 这是你应该可以看到所有的word节点，以及他们的CPU个数和内存等信息



运行SparkPi例子

$ cd ~/spark-1.0.1-bin-2.2.0/bin

$ ./run-example SparkPi

### Eclipse上安装eclipse

解压eclipse到/opt下

cd /opt

sudo tar zxvf eclipse-jee-luna-R-linux-gtk-x86\_64.tar

然后双击eclipse运行图标，eclipse启动起来了。

创建桌面链接

$ sudo su

# vim /usr/share/applications/eclipse.desktop

写入以下内容

[Desktop Entry]

Name=Eclipse

Comment=Eclipse SDK

Encoding=UTF-8

Exec=/opt/eclipse

Icon=/opt/icon.xpm

Terminal=false

Type=Application

Categories=Application;Development;

完成后可以在Application-Programming下看到eclipse图标。

### 在 eclipse 运行 WordCount

导入第三方jar 包

Spark目录/lib/spark-assembly-1.0。1-hadoop2.2.0.jar

导入成功后：

