**JStorm安装配置**

Jstorm是阿里在storm的基础上开发的一个流计算系统，提供比storm更多的功能和更好的稳定性。

vi详解

http://wenku.baidu.com/link?url=lwNWeoLEG0UnzoMhdw7X2Vek0sNkIqAaRVk0Uh7CkB51yAbXXD8ZarQuIZjuKfGRtEoRXBwn5Yq5zF74EcA7SHIVs84ocI57Vpnid7ruFQ7

说明：

所有软件以及环境依据此文档为准，并且需要用root超级用户来完成，另外需要关闭防火墙。不保证其他环境能成功

**环境：**

系统环境：Ubuntu

集群：两台服务器软件安装完全一致，配置需要修改zookeeper目录下zoo.cfg文件。

需要安装包：

需要安装的软件

* [Python](https://www.python.org/downloads/)和[Java](http://www.oracle.com/technetwork/java/javase/downloads/index.html" \t "_blank)相信大家都安装过了，可以在终端运行python和java -version命令确认一下，注意python要用2.6或2.7版本。
* [Jstorm](https://github.com/alibaba/jstorm/wiki/Downloads)
* [ZooKeeper](http://zookeeper.apache.org/releases.html#download)
* [Tomcat](http://tomcat.apache.org/)

**需要特别注意的是，Github的官方教程中提到了安装zeroMQ和jzmq，这两个其实不是必需的，消息队列我们可以根据自己的需求来配置。**

另外强烈推荐安装 pdsh 和 pdcp ，这两个命令对管理大规模的分布式集群非常有用，可以减少大量工作量。

**安装python**

下载地址：<https://www.python.org/downloads>

# tar –jxvf Python-2.7.3.tar.bz2

# cd Python-2.7.3

# ./configure

# make

# make install

在这里我用apt-get安装或者是在安装系统的时候自带的。

apt-get install python

在终端输入python，如果进入了python的shell编程则说明安装成功：

很多linux 版本自带了python，如果自带的python版本低于安装版本，执行python命令发现版本是旧的，解决如下：

# cd /usr/bin

# ll |grep python

# rm -rf python

# ln -s Python\_HOME/Python-2.7.3/python.exe ./python

# python

**安装JAVA**

1）解压JDK

tar xzvf jdk-7u67-linux-x64.tar.gz

mv jdk1.7.0\_67/ /usr/local/java

2）配置环境变量

vi ~/.bashrc

export JAVA\_HOME=/usr/local/java

export JRE\_HOME=/usr/local/java/jre

export CLASSPATH=.:$JAVA\_HOME/lib:$JRE\_HOME/lib:$CLASSPATH

export PATH=$JAVA\_HOME/bin:$JRE\_HOME/bin:$PATH

使环境变量生效

source ~/.bashrc

执行以下命令

update-alternatives --install /usr/bin/java java /usr/local/java/bin/java 300

update-alternatives --install /usr/bin/javac javac /usr/local/java/bin/javac 300

测试java

java –version

**安装Zookeeper**

下载地址：http://zookeeper.apache.org/releases.html#download

ZooKeeper用来管理和协调整个集群，但不必在所有jstorm节点上都安装。

1. 解压

tar xzvf zookeeper-3.4.6.tar.gz

mv zookeeper-3.4.6 /usr/local/zookeeper

1. 配置环境变量，依旧配置~/.bashrc，加入以下内容：

export ZOOKEEPER\_HOME=/usr/local/zookeeper

export PATH=$ZOOKEEPER\_HOME/bin:$JAVA\_HOME/bin:$JRE\_HOME/bin:$PATH

export CLASSPATH=$ZOOKEEPER\_HOME/lib:.:$JAVA\_HOME/lib:$JRE\_HOME/lib:$CLASSPATH

然后执行

source ~/.bashrc

1. 配置zookeeper参数，配置文件在ZOOKEEPER\_HOME/conf/目录，将zoo\_sample.cfg重命名为zoo.cfg

mv zoo\_sample.cfg zoo.cfg

这个文件中是默认的配置参数，主要有如下参数：

tickTime - 这个时间是作为 Zookeeper 服务器之间或客户端与服务器之间维持心跳的时间间隔，也就是每个 tickTime 时间就会发送一个心跳。

dataDir - 顾名思义就是 Zookeeper 保存数据的目录，默认情况下，Zookeeper 将写数据的日志文件也保存在这个目录里。

clientPort - 这个端口就是客户端连接 Zookeeper 服务器的端口，Zookeeper 会监听这个端口，接受客户端的访问请求。

server.x - 设置第x个ZooKeeper服务器的ip地址和端口信息。

注意修改其中dataDir参数：

dataDir=$ZOOKEEPER\_HOME/data

server.1=10.172.44.208:2888:3888

server.2=10.172.44.252:2888:3888

第一个端口号表示的是这个服务器与集群中的 Leader 服务器交换信息的端口；

第二个端口号表示的是万一集群中的 Leader 服务器挂了，用来执行选举新 Leader 时服务器相互通信的端口。

注意红色字是根据实际地址来填写的。

配置完成后将此配置文件同步到所有ZooKeeper节点相应位置。

1. 接着在每台服务配置dataDir文件中创建一个myid文件，里面内容为一个数字，用来标识当前主机，conf/zoo.cfg文件中配置的server.X中X是什么数字，则myid文件中就输入这个数字。

第二台服务器

vi /usr/local/zookeeper/data/myid

2

cat /usr/local/zookeeper/data/myid

2

1. 至此，ZooKeeper就配置完成了。启动ZooKeeper的脚本在$ZOOKEEPER\_HOME/bin下，启动命令如下：

# zkServer.sh start

提示启动成功后可以执行jps查看是否有ZooKeeper的进程，也可以执行zkServer status查看详细信息。

**安装Jstorm**

下载地址：https://github.com/alibaba/jstorm/wiki/Downloads

1. 解压安装

wget http://42.121.19.155/jstorm/jstorm-0.9.3.zip

unzip jstorm-0.9.3.zip

mv jstorm /usr/local/jstorm

1. 配置环境变量，加入：

vi ~/.bashrc

export JSTORM\_HOME=/usr/local/jstorm

export PATH=$JSTORM\_HOME/bin:$ZOOKEEPER\_HOME/bin:$JAVA\_HOME/bin:$JRE\_HOME/bin:$PATH

source ~/.bashrc

配置$JSTORM\_HOME/conf/storm.yaml

1. 配置jstorm,配置文件为$JSTORM\_HOME/conf/storm.yaml

修改配置文件

nimbus.host: "10.0.0.4"

supervisor.slots.ports:

- 6800

- 6801

- 6802

- 6803

配置项：

storm.zookeeper.servers: 表示zookeeper 的地址，

storm.zookeeper.root: 表示jstorm在zookeeper中的根目录，当多个JStorm共享一个ZOOKEEPER时，需要设置该选项，默认即为“/jstorm”

storm.local.dir: 表示jstorm临时数据存放目录，需要保证jstorm程序对该目录有写权限

java.library.path: zeromq 和java zeromq library的安装目录，默认"/usr/local/lib:/opt/local/lib:/usr/lib"

supervisor.slots.ports: 表示supervisor 提供的端口slot列表，注意不要和其他端口发生冲突，默认是68xx，而storm的是67xx

supervisor.disk.slot： 表示提供数据目录，当一台机器有多块磁盘时，可以提供磁盘读写slot，方便有重IO操作的应用。

topology.enable.classloader: false, 默认关闭classloader，如果应用的jar与jstorm的依赖的jar发生冲突，比如应用使用thrift9，但jstorm使用thrift7时，就需要打开classloader

nimbus.groupfile.path: 如果需要做资源隔离，比如数据仓库使用多少资源，技术部使用多少资源，无线部门使用多少资源时，就需要打开分组功能， 设置一个配置文件的绝对路径，改配置文件如源码中group\_file.ini所示

storm.local.dir: jstorm使用的本地临时目录，如果一台机器同时运行storm和jstorm的话， 则不要共用一个目录，必须将二者分离开

nimbus.host: Storm集群Nimbus机器地址，各个Supervisor工作节点需要知道哪个机器是Nimbus，以便下载Topologies的jars、confs等文件，如：

第二台服务器上配置Slave

nimbus.host: "111.222.333.444"

1. 然后在提交Jar的节点和计划安装WebUI的节点上：

在提交jar的节点上执行:

#mkdir ~/.jstorm

#cp -f $JSTORM\_HOME/conf/storm.yaml ~/.jstorm

**安装Jstorm Web UI（Tomcat）**

下载地址：http://tomcat.apache.org/

必须使用tomcat 7.0 或以上版本， web ui 可以和nimbus不在同一个节点。

1. 解压安装tomcat：

tar xzvf apache-tomcat-7.0.54.tar.gz

mv apache-tomcat-7.0.54 /usr/local/tomcat

1. 修改tomcat/bin/catalina.sh文件添加JAVA路径

export JAVA\_HOME=/usr/local/java

1. 将服务器设置为jstorm的webUI：

cd /usr/local/tomcat/webapps

cp $JSTORM\_HOME/jstorm-ui-0.9.3.war ./

mv ROOT ROOT.old

ln -s jstorm-ui-0.9.3 ROOT

1. 启动tomcat：

cd ../bin

./startup.sh