

# 1 准备

## 1.1 准备linux服务器

本文搭建的是3节点的集群，需要3台linux服务器，我这里使用的是centos7版本的linux虚拟机，虚拟机网络配置如下：

主节点：master 192.168.92.90

从节点：slave1 192.168.92.91

slave2 192.168.92.92

**注：**

*每台linux服务器的时间保持一致*

*虚拟机搭建详见《centos虚拟机搭建与网络配置》*

## 1.2 下载storm

本文使用 apache-storm-1.0.0.tar.gz

下载地址：

<http://archive.apache.org/dist/storm/apache-storm-1.0.0>



# Index of /dist/storm/apache-storm-1.0

Name	Last modified	Size	Description
 <a href="#">Parent Directory</a>		-	
 <a href="#">apache-storm-1.0.0-src.tar.gz</a>	2016-04-05 18:32	7.8M	
 <a href="#">apache-storm-1.0.0-src.tar.gz.asc</a>	2016-04-05 18:32	496	
 <a href="#">apache-storm-1.0.0-src.tar.gz.md5</a>	2016-04-05 18:32	80	
 <a href="#">apache-storm-1.0.0-src.tar.gz.sha</a>	2016-04-05 18:32	268	
 <a href="#">apache-storm-1.0.0-src.zip</a>	2016-04-05 18:32	9.6M	
 <a href="#">apache-storm-1.0.0-src.zip.asc</a>	2016-04-05 18:32	496	
 <a href="#">apache-storm-1.0.0-src.zip.md5</a>	2016-04-05 18:32	77	
 <a href="#">apache-storm-1.0.0-src.zip.sha</a>	2016-04-05 18:32	256	
 <a href="#">apache-storm-1.0.0.tar.gz</a>	2016-04-05 18:32	146M	
 <a href="#">apache-storm-1.0.0.tar.gz.asc</a>	2016-04-05 18:32	496	
 <a href="#">apache-storm-1.0.0.tar.gz.md5</a>	2016-04-05 18:32	76	
 <a href="#">apache-storm-1.0.0.tar.gz.sha</a>	2016-04-05 18:32	225	
 <a href="#">apache-storm-1.0.0.zip</a>	2016-04-05 18:32	146M	
 <a href="#">apache-storm-1.0.0.zip.asc</a>	2016-04-05 18:32	496	
 <a href="#">apache-storm-1.0.0.zip.md5</a>	2016-04-05 18:32	73	
 <a href="#">apache-storm-1.0.0.zip.sha</a>	2016-04-05 18:32	216	

## 1.3 安装zookeeper

详见《ZooKeeper服务器的安装与启动》

## 1.4 安装storm的其他依赖

### 1) 安装JKD

详见《linux下安装jdk》

### 2) 安装Python

我这里linux版本为centos7.0, 自带Python环境

```
Last login: wed Apr 11 23:42:16 2016 from 192.168.92.1
[1by@master ~]$ python
Python 2.6.6 (r266:84292, Nov 22 2013, 12:16:22)
[GCC 4.4.7 20120313 (Red Hat 4.4.7-4)] on linux2
Type "help", "copyright", "credits" or "license" for more information.
>>> █
```

## 2 storm的分布式模式安装

### 2.1 上传apache-storm-1.0.0.tar.gz到服务器

在当前用户的/opt目录下创建software目录

命令：mkdir software

把apache-storm-1.0.0.tar.gz分别上传到3台虚拟机，我这里是上传到以下目录：

/opt/software/

```
anaconda-ks.cfg
apache-storm-1.0.0
apache-storm-1.0.0.tar.gz
data
hadoop-2.6.0-cdh5.14.0
hadoop-2.6.0-cdh5.14.0.tar.gz
idea-IC-181.4445.78
initial-setup-ks.cfg
jdk1.8
jdk-8u162-linux-x64.tar.gz
mydir
mysql-5.6.26-1.linux_glibc2.5.x86_64.rpm-bundle.tar
MySQL-client-5.6.26-1.linux_glibc2.5.x86_64.rpm
mysql-community-release-el7-5.noarch.rpm
mysql-community-release-el7-5.noarch.rpm.1
MySQL-devel-5.6.26-1.linux_glibc2.5.x86_64.rpm
MySQL-embedded-5.6.26-1.linux_glibc2.5.x86_64.rpm
MySQL-server-5.6.26-1.linux_glibc2.5.x86_64.rpm
MySQL-shared-compat-5.6.26-1.
MySQL-test-5.6.26-1.linux_gli
nohup.out
spark-2.2.0-bin-hadoop2.6
spark-2.2.0-bin-hadoop2.6.tgz
workspace
zookeeper-3.4.5-cdh5.14.0
zookeeper-3.4.5-cdh5.14.0.tar
zookeeper.out
公共
模板
视频
图片
文档
下载
音乐
桌面
```

### 2.2 解压apache-storm-1.0.0.tar.gz

解压命令：

tar -zxvf apache-storm-1.0.0.tar.gz

```
[lby@master ~]$ tar -zxvf apache-storm-0.9.6.tar.gz
```

**其他几台服务器同样解压。**

### 2.3 配置storm

#### 2.3.1 进入apache-storm-1.0.0/conf目录

命令：

cd apache-storm-1.0.0/conf

#### 2.3.2 编辑storm.yaml文件

命令：

vim storm.yaml

删除storm.yaml文件的所有注释，在storm.yaml文件中加入如下配置：

---

```
#####
```

```
#
```

#注意：所以设置的开头行都要有空格

```
#
#####
#运行zookeeper的设置，可以使用hostname和IP地址两种方式，我这里的zookeeper的
hostname分别是master、slave1、slave2
storm.zookeeper.servers:
    - "master"
    - "slave1"
    - "slave2"
#指定nimbus的节点机器，使用服务器的hostname，我这里指定的是3台，分别是
master、slave1、slave2
nimbus.seeds: ["master", "slave1", "slave2"]
#指定storm web UI的启动端口，我这里是9999，防止与其他应用端口冲突
ui.port: 9999
#保存storm数据的路径，我这里是/tmp/storm
storm.local.dir: "/tmp/storm"
#指定work（spout和bolt）的运行端口,也是一个supervisor能并行的work的最大数量
（能够开启的最大进程数），每个work占用1个端口，我这里每台supervisor设置的是可以
并行4个work，端口可以任意指定不重复的端口，我这里是6700、6701、6702、6703
supervisor.slots.ports:
    - 6700
    - 6701
    - 6702
    - 6703
```

---

```
#####
#
#注意：所以设置的开头行都要有空格
#
#####
#运行zookeeper的设置，可以使用hostname和IP地址两种方式，我这里的zookeeper的hostname分别是master、slave1、slave2
storm.zookeeper.servers:
    - "master"
    - "slave1"
    - "slave2"
#指定nimbus的节点机器
nimbus.host: "master"
#指定storm web UI的启动端口，我这里是9999，防止与其他应用端口冲突
ui.port: 9999
#保存storm数据的路径
storm.local.dir: "/home/lby/data/storm"
#指定work（spout和bolt）的运行端口,也是一个supervisor能并行的work的最大数量（能够开启的最大进程数），每个work占
用1个端口，我这里每台supervisor设置的是可以并行4个work
supervisor.slots.ports:
    - 6700
    - 6701
    - 6702
    - 6703
#####
```

保存storm.yaml文件并退出

### 2.3.3 创建保存storm数据的文件夹

创建文件夹的路径是storm.yaml文件的storm.local.dir配置项，我这里  
是/home/lby/data/storm，命令：

```
mkdir -p /home/lby/data/storm
```

### 2.3.4 配置环境变量

1) 在用户目录下（我这里是/home/lby），编辑bash\_profile，命令：

```
vim ~/.bash_profile
```

在bash\_profile中加入以下配置：

```
export STORM_HOME=storm安装路径
export PATH=.:$STORM_HOME/bin:$PATH
```

**注：**

*我这里storm路径是/home/lby/apache-storm-1.0.0*

```
export HADOOP_HOME=/opt/software/hadoop-2.6.0-cdh5.14.0
export ZOOKEEPER_HOME=/opt/software/zookeeper-3.4.5-cdh5.14.0
export JAVA_HOME=/opt/software/jdk1.8
export JRE_HOME=${JAVA_HOME}/jre
export CLASSPATH=.:${JAVA_HOME}/lib:${JRE_HOME}/lib
export STORM_HOME=/opt/software/apache-storm-1.0.0
export SPARK_HOME=/opt/software/spark-2.2.0-bin-hadoop2.6
export PATH=.:$SPARK_HOME/bin:~:$STORM_HOME/bin:$ZOOKEEPER_HOME/bin:$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$JAVA_HOME/bin:$PATH
```

2) 保存退出

3) 使配置的环境变量生效，命令：

```
source ~/.bash_profile
```

**所有服务器的storm配置相同。**

## 3 storm的启动与停止

### 3.1 启动

#### 3.1.1 在nimbus主节点执行以下启动命令

我这里master是nimbus主节点，slave1和slave2是备用节点

1) 启动nimbus守护进程，命令：

```
storm nimbus < /dev/null 2<&1 &
```

jps查看进程，出现nimbus进程

```
[hadoop@master storm]$ jps
2374 QuorumPeerMain
6585 nimbus
6874 Jps
3343 AzkabanWebServer
2813 ResourceManager
3401 AzkabanExecutorServer
6356 core
2515 NameNode
2671 SecondaryNameNode
```

2) 启动web UI守护进程，命令：

```
storm ui < /dev/null 2<&1 &
```

jps查看进程，出现core进程

```
[hadoop@master storm]$ jps
2374 QuorumPeerMain
6585 nimbus
6874 Jps
3343 AzkabanWebServer
2813 ResourceManager
3401 AzkabanExecutorServer
6356 core
2515 NameNode
2671 SecondaryNameNode
```

### 3.1.2 在从节点执行以下启动命令

我这里从节点是slave1、slave2

启动supervisor守护进程，在从节点分别执行命令：

```
storm supervisor < /dev/null 2<&1 &
```

jps查看进程，出现supervisor进程

```
[hadoop@slave1 storm]$ jps
2421 NodeManager
2342 DataNode
4136 supervisor
4279 Jps
2274 QuorumPeerMain
```

## 3.2 停止

jps查看进程，执行kill命令杀死进程

例如，杀死nimbus进程：

jps查看进程

```
[hy@mc ~]$ jps
928 NameNode
16368 Jps
1077 SecondaryNameNode
1255 ResourceManager
16206 nimbus
[hy@mc ~]$
```

上图所示，nimbus进程号是16206，执行命令：

```
kill -9 16206
```

jps查看进程，发现nimbus已经被杀死