**Storm 1.1.2 集群安装**

时间：2018-03-12 作者：陈少钦

安装环境如下：   
操作系统：CentOS 6.7 64位   
ZooKeeper版本：3.4.5   
Storm版本：1.1.2   
JDK版本：1.8.0\_151-b12 64位   
python版本：2.6.6   
集群情况：一个主控节点（master）和两个工作节点（slave1，slave2）

**1. 搭建Zookeeper集群**

略，已安装，详见《05-ZooKeeper 安装.txt》。

安装参考：[CentOS下ZooKeeper单机模式、集群模式安装](http://blog.csdn.net/lnho2015/article/details/51142572)

**2. 在Nimbus和Worker机器上安装依赖包**

略，已安装，详见《03-Linux 安装 JDK.docx》。

Java 6   
Python 2.6.6

以上的版本是官方说已经有测试可以和Strom正常运行的版本，本文采用Java 8进行安装。   
参考[CentOS 6使用rpm方式安装JDK8](http://blog.csdn.net/lnho2015/article/details/51093944)。   
由于CentOS 6已经自带Python 2.6.6，这边就不介绍安装方法了。可以采用python --version来检查服务器是否安装python，以及对应的版本。

**消息队列说明：**

ZeroMQ：Storm0.9.0之前的版本中需要安装ZeroMQ，实现Strom Worker之间的通信；Storm0.9.0版本之后采用Netty替代ZeroMQ使Strom部署分布式环境更方便。

Netty：JBOSS提供的一个java开源框架。Netty提供异步的、事件驱动的网络应用程序框架和工具，用以快速开发高性能、高可靠性的网络服务器和客户端程序。

**3. 下载解压最新版的Storm**

打开官网下载页面：<http://storm.apache.org/downloads.html>   
选择适合的版本进行下载。

cd /tmp/pkg

wget " http://mirrors.hust.edu.cn/apache/storm/apache-storm-1.1.2/apache-storm-1.1.2.tar.gz"

tar -zxvf apache-storm-1.1.2.tar.gz -C /opt/bigdata/

cd /opt/bigdata/

mv apache-storm-1.1.2/ storm

**4. 修改配置文件**

打开核心配置文件：storm.yaml

cd /opt/bigdata/storm/conf

vi storm.yaml

至少完成以下3处的配置，笔者的配置文件：

storm.zookeeper.servers:

- "master"

- "slave1"

- "slave2"

storm.local.dir: " /opt/bigdata/storm/data"

nimbus.seeds: ["master"]

其中storm.zookeeper.servers，表示Zookeeper集群地址。如果Zookeeper集群使用的不是默认端口，那么还需要配置storm.zookeeper.port。storm.local.dir用于配置Storm存储少量文件的路径。nimbus.seeds用于配置主控节点的地址，可以配置多个。

**5. 拷贝文件到其余主控节点和工作节点**

scp -r /opt/bigdata/storm hadoop@slave1:/opt/bigdata

scp -r /opt/bigdata/storm hadoop@slave2:/opt/bigdata

**6. 加入环境变量（三台机器）**

sudo vi /etc/profile

export STORM\_HOME=/opt/bigdata/storm

export PATH=$STORM\_HOME/bin:$PATH

. /etc/profile

**7. 启动Storm**

Storm集群中包含两类节点：主控节点（Master Node）和工作节点（Work Node）。其分别对应的角色如下：   
主控节点（Master Node）上运行一个被称为Nimbus的后台程序，它负责在Storm集群内分发代码，分配任务给工作机器，并且负责监控集群运行状态。   
每个工作节点（Work Node）上运行一个被称为Supervisor的后台程序。Supervisor负责监听从Nimbus分配给它执行的任务，据此启动或停止执行任务的工作进程。每一个工作进程执行一个Topology的子集；一个运行中的Topology由分布在不同工作节点上的多个工作进程组成。

**7.1 启动主控节点**

nohup /opt/bigdata/storm/bin/storm nimbus 1>/dev/null 2>&1 &

**7.2 启动工作节点**

nohup /opt/bigdata/storm/bin/storm supervisor 1>/dev/null 2>&1 &

**7.3 启动管理页面**

在主控节点上运行

nohup /opt/bigdata/storm/bin/storm ui 1>/dev/null 2>&1 &

说明：需要分别到每个节点启动服务。管理界面将运行在启动它的服务器的8080端口，日志文件在程序目录中logs文件夹下。上面的启动方式是在前台运行的，如果需要在后台运行，可以使用类似nohup /opt/apache-storm-1.0.0/bin/storm nimbus &的方式运行。

**7.4 查看进程和页面**



