[ZooKeeper-3.4.6分布式安装指南](http://blog.chinaunix.net/uid-20682147-id-4220311.html) 2014-04-22 11:49:33

分类： HADOOP

[[http://blog.chinaunix.net/image/default/fj.png](http://blog.chinaunix.net/blog/downLoad/fileid/12126.html)ZooKeeper-3.4.6分布式安装指南.pdf](http://blog.chinaunix.net/blog/downLoad/fileid/12126.html)

**目录**

[目录 1](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc20081)

[1. 前言 1](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc16828)

[2. 约定 1](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc16667)

[3. 安装步骤 2](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc10153)

[3.1. 配置/etc/hosts 2](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc13485)

[3.2. 设置myid 2](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc16268)

[3.3. 修改conf/zoo.cfg 2](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc27997)

[3.4. 修改/bin/zkEnv.sh 3](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc3088)

[3.5. 修改conf/log4j.properties 3](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc15668)

[3.6. 修改conf/java.env 3](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc12541)

[3.7. 设置PATH 4](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc29496)

[4. 启动ZooKeeper集群 4](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc31283)

[5. 安装验证 4](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc18671)

[6. 基本命令 5](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc2978)

[7. 只启一台zookeeper会如何？ 6](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc26886)

[8. 如何扩容zookeeper？ 7](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc2519)

[9. 相关文档 7](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc16411)

[10. 结束语 7](http://blog.chinaunix.net/uid-20682147-id-4220311.html#_Toc18454)

**1. 前言**

介绍ZooKeeper-3.4.6版本的分布式安装，力求细致，提供精确的安装指导。本文的安装环境是64位的SuSE 10.1 Linux，也适用于ZooKeeper-3.4.8。

ZooKeeper节点间不需要互登录，因此不用配置免密码登录。

**2. 约定**

将ZooKeeper安装在/data/hadoop/zookeeper目录，其中/data/hadoop/zookeeper是指向/data/hadoop/zookeeper-3.4.6的软链接。

ZooKeeper的数据目录设置为/data/hadoop/zookeeper/data。

本文使用了3台机器部署ZooKeeper集群，IP和主机名对应关系如下：

|  |  |
| --- | --- |
| **IP** | **主机名** |
| 10.12.154.77 | DEVNET-154-77 |
| 10.12.154.78 | DEVNET-154-70 |
| 10.12.154.79 | DEVNET-154-79 |

**3. 安装步骤**

Zookeeper机器间不需要设置免密码登录，其它hadoop也可以不设置，只要不使用hadoop-daemons.sh来启动、停止进程，注意不是hadoop-daemon.sh，而是带“s”的那个，带“s”的会读取hadoop的salves文件，远程ssh启动DataNode和备NameNode等。

**3.1. 配置/etc/hosts**

将3台机器的IP和主机名映射关系，在3台机器上都配置一下，亦即在3台机器的/etc/hosts文件中，均增加以下内容（可以先配置好一台，然后通过scp等命令复制到其它机器上，注意主机名不能包含任何下划线）：

|  |
| --- |
| 10.12.154.77    DEVNET-154-77  10.12.154.78    DEVNET-154-70  10.12.154.79    DEVNET-154-79 |

**3.2. 设置myid**

在dataDir指定的数据目录（/data/hadoop/zookeeper/data）下，创建文件myid，文件内容为一个正整数值，用来唯一标识当前机器，因此不同机器的数值不能相同，建议从1开始递增标识，以方便记忆和管理。本文约定如下：

|  |  |
| --- | --- |
| **IP** | **标识数值（取值范围为1~255）** |
| 10.12.154.77 | **1** |
| 10.12.154.78 | **2** |
| 10.12.154.79 | **3** |

可以使用echo命令直接写进去，如：

|  |
| --- |
| **echo**1 > myid或**echo**1 > **/data/hadoop/zookeeper/data**/**myid** |

**3.3. 修改conf/zoo.cfg**

按下述内容，修改ZooKeeper配置文件（3台机器都需要，可以先配置好一台，然后通过scp等命令复制到其它机器上）：

|  |
| --- |
| tickTime=2000  dataDir=**/data/hadoop/zookeeper/data**  **dataLogDir**=**/data/hadoop/zookeeper/datalog**  clientPort=**2181**  initLimit=5  syncLimit=2  autopurge.**snapRetainCount**=10  autopurge.**purgeInterval**=1  server.1=DEVNET-154-77:**2888**:**3888**  server.2=DEVNET-154-78:**2888**:**3888**  server.3=DEVNET-154-79:**2888**:**3888** |

server.N中的N值必须和myid保持相同。

以上内容的配置，参照了ZooKeeper的官方文档：[zookeeperStarted.html](http://zookeeper.apache.org/doc/trunk/zookeeperStarted.html)。server.X用来配置ZooKeeper集群中的各节点，并建议X的值和myid保持一致。

端口2181用于监听客户端的连接，端口2888用于Leader监听Follower的连接，而3888则用于Leader选举。

dataDir用来指定快照文件存放目录，dataLogDir用来指定事务日志文件存放目录。如果只指定了dataDir，而没指定dataLogDir，则事务日志文件也存放于dataDir指定的目录。

默认，autopurge的两个参数是被注释的。

“autopurge.purgeInterval”指定自动清理快照文件和事务日志文件的时间，单位为小时，默认为0表示不自动清理，这个时候可以使用脚本zkCleanup.sh手动清理。不清理的结果是占用的磁盘空间越来越大。

“autopurge.snapRetainCount”用于指定保留快照文件和事务日志文件的个数，默认为3。

**3.4. 修改/bin/zkEnv.sh**

设置环境变量ZOO\_LOG\_DIR为zookeeper的日志存放目录，如：

export ZOO\_LOG\_DIR=/data/hadoop/zookeeper/log

可将这一行放在ZOOKEEPER\_PREFIX之后。

再修改下ZOO\_LOG4J\_PROP，以让日志不是输出到zookeeper.out，而是写入到日志文件，将：ZOO\_LOG4J\_PROP="INFO,CONSOLE"，改成：

ZOO\_LOG4J\_PROP="INFO,ROLLINGFILE"

**3.5. 修改conf/log4j.properties**

log4j.appender.ROLLINGFILE.MaxFileSize=100MB

log4j.appender.ROLLINGFILE.MaxBackupIndex=10

设置每个日志文件大小为100M，滚动10个。

**3.6. 修改conf/java.env**

打开文件bin/zkEnv.sh，可以看到它有引用conf/java.env。可以在java.env中设置JAVA\_HOME这个环境变量，如：export JAVA\_HOME=/data/jdk。

仅在/etc/profile等处设置JAVA\_HOME，可能并不管用。如果不管用，查看log/zookeeper.out，可以看到错误“nohup: failed to run command 'java': No such file or directory”。

**3.7. 设置PATH**

这步不是必须的，但是推荐将zookeeper/bin加入到环境变量PATH，这样操作zookeeper会方便些，如：export PATH=/data/zookeeper/bin:$PATH。

如果计划设置，修改当前用户的.profile或.bashrc即可，而不是/etc/profile等。

**4. 启动ZooKeeper集群**

登录所有ZooKeeper节点，使用bin目录下的脚本zkServer.sh用来启动ZooKeeper集群，但需要带一个start参数，如：

|  |
| --- |
| cd /data/hadoop/zookeeper/bin&&./**zkServer.sh** **start** |

停止参数为stop，重启参数为restart，如：./zkServer.sh restart。

由于启动时，每个节点都会试图去连接其它节点，因此先启动的刚开始会连接不上其它的，导致日志中会包含错误信息，在未全启动之前，这个属正常现象。

**5. 安装验证**

脚本zkServer.sh不但可以用来启动ZooKeeper，还可以用来查看状态。使用方式为带一个status参数，如：

|  |
| --- |
| **./zkServer.sh status** |

如遇到以下错误，请稍安勿躁，过会再试，可能是因为还未完全起来：

|  |
| --- |
| [hadoop@DEVNET-154-77 ~/zookeeper/bin]$**./zkServer.sh status**  JMX enabled by default  Using config: /data/hadoop/zookeeper/bin/../conf/zoo.cfg  Error contacting service. It is probably not running. |

集群成功启动后，将有且只会有一个成为leader，其它是follower：

|  |
| --- |
| [hadoop@**DEVNET-154-79** ~/zookeeper/bin]$ **./zkServer.sh status**  JMX enabled by default  Using config: /data/hadoop/zookeeper/bin/../conf/zoo.cfg  Mode: **leader**    [hadoop@**DEVNET-154-70** ~/zookeeper/bin]$**./zkServer.sh status**  JMX enabled by default  Using config: /data/hadoop/zookeeper/bin/../conf/zoo.cfg  Mode: **follower**    [hadoop@**DEVNET-154-77** ~/zookeeper/bin]$ **./zkServer.sh status**  JMX enabled by default  Using config: /data/hadoop/zookeeper/bin/../conf/zoo.cfg  Mode: **follower** |

**6. 基本命令**

进入ZooKeeper的bin目录，执行zkCli.sh进入ZooKeeper的命令行操作界面。

|  |
| --- |
| ./**zkCli**.sh -**server 10.12.154.78**:**2181** |

参数“-server”中只含一个“-”，用以指定被连接的ZooKeeper节点，可以为Leader，也可以为Follower，“10.12.154.78”为Leader或Follower的IP或主机名，“2181”为ZooKeerp提供的客户端服务端口。

进入ZooKeeper命令行操作界面后，输入help然后回车，可以看到ZooKeeper支持的命令列表：

|  |
| --- |
| [zk: DEVNET-154-79:2181(CONNECTED) 0] **help**  ZooKeeper -server host:port cmd args  **stat**path [watch]  **set**path data [version]  **ls**path [watch]  **delquota**[-n|-b] path  **ls2**path [watch]  **setAcl**path acl  **setquota**-n|-b val path  **history**  **redo**cmdno  **printwatches**on|off  **delete**path [version]  **sync**path  **listquota**path  **rmr**path  **get**path [watch]  **create**[-s] [-e] path data acl  **addauth**scheme auth  **quit**  **getAcl**path  **close**  **connect**host:port |

“ls”有点像HDFS中的“ls”，随便测试下：

|  |
| --- |
| [zk: DEVNET-154-79:2181(CONNECTED) 1] **ls /**  [zookeeper, hbase] |

再测试下：

|  |
| --- |
| [zk: DEVNET-154-79:2181(CONNECTED) 2]**ls /hbase**  [replication, meta-region-server, rs, splitWAL, backup-masters, table-lock, region-in-transition, online-snapshot, master, running, recovering-regions, draining, namespace, hbaseid, table] |

再测试一下：

|  |
| --- |
| [zk: DEVNET-154-79:2181(CONNECTED) 4]**ls /hbase/backup-masters**  [VM-40-171-sles10-64,60000,1398162436871] |

测试创建一个目录：

|  |
| --- |
| [zk: DEVNET-154-79:2181(CONNECTED) 2]**create /test test**  Created /test    [zk: DEVNET-154-79:2181(CONNECTED) 2]**ls /**  [zookeeper, test] |

上面可以看到有HBase，是因为基于它部署了一套HBase，若未基于它安装HBase，则不会有这个条目。

**7. 只启一台zookeeper会如何？**

如果只启动一台，执行zkServer.sh status查看状态时，报如下错误：

|  |
| --- |
| $ zkServer.sh status  ZooKeeper JMX enabled by default  Using config: /data/zookeeper/bin/../conf/zoo.cfg  Error contacting service. It is probably not running. |

如果再启一台zookeeper，会发现一个成了leader，另一个成了follower（经测试两台zookeeper即能正常工作）：

|  |
| --- |
| $ zkServer.sh status  ZooKeeper JMX enabled by default  Using config: /data/zookeeper/bin/../conf/zoo.cfg  Mode: **follower**    $ zkServer.sh status  ZooKeeper JMX enabled by default  Using config: /data/zookeeper/bin/../conf/zoo.cfg  Mode: **leader** |

**8.****如何扩容zookeeper？**

只需要将已有的zookeeper打包复制到新的机器上，然后修改myid文件并设置好，然后启动zookeeper即可。

**9. 相关文档**

《HBase-0.98.0分布式安装指南》

《Hive 0.12.0安装指南》

《ZooKeeper-3.4.6分布式安装指南》

《Hadoop 2.3.0源码反向工程》

《在Linux上编译Hadoop-2.4.0》

《Accumulo-1.5.1安装指南》

《Drill 1.0.0安装指南》

《Shark 0.9.1安装指南》

更多，敬请关注技术博客：[http://aquester.cublog.cn](http://blog.chinaunix.net/uid/20682147.html)。

**10. 结束语**

至此，ZooKeeper分布式安装大告成功！更多细节，请浏览官方文档：<http://zookeeper.apache.org/doc/trunk/zookeeperStarted.html>。