Apache ZooKeeper学习笔记

日期：2016年6月8日

01课时 ZooKeeper简介

ZooKeeper（Hadoop生态系统的管理员）：分布式系统协同服务

ZooKeeper是基于内存的服务，所有安装ZooKeeper的都是ZKServer，在集群内的所有的节点都要运行ZooKeeper。

其中ZKServer集群中的角色：有一个leader，其他的被称为follower。所有节点之间均需要互相通信。

zkClient访问集群中ZKServer Leader的任何一台机器，获得的内容是相同的，ZKServer内存数据结构是完全相同的。

ZooKeeper可以理解为一个文件系统，每一个文件夹对应唯一的路径。

ZooKeeper Server Leader宕机之后，有一个Election机制，选举新的Leader

在分布式应用上有多重配置服务：time同步服务、感知服务（online）等，这些由ZooKeeper完成。

ZooKeeper is for highly reliable distributed coordination: 分布式的开源应用程序协同服务，是Google Chubby的开源实现，是Hadoop HBase的重要组件。

ZooKeeper提供的服务：配置维护、名字服务、分布式同步、组服务。

Zookeeper：封装复杂、易错的关键服务，将简单易用的接口和性能高效、功能稳定的系统提供给用户。ZooKeeper支持Java 和C。

02课时 ZooKeeper下载与安装

1. ZNode的知识

ZNode 数据是有版本的，每次更新都会修改时间戳

ZNode读写操作是基于原子性的，读完整的数据，写入数据是完全覆盖，ZNode的ACL存储用户操作权限；ZNode中的临时节点和Session相关，session结束，节点删除。

2. ZooKeeper安装

1)install java jdk 1.6+

2) download the zookeeper( zookeeper.apache.org) then upzip the file to the /usr/cloud

3)create the soft link “ln -s zookeeper3.4.8 zookeeper”

4)update the /etc/profile, run it by command “source /etc/profile”

#zookeeper environment

export ZOOKEEPER\_HOME=/usr/cloud/zookeeper

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

5)set the configuration file which named “zoo.cfg” for Standalone

tickTime=2000 # heart beat recycle

dataDir=/opt/zookeeper # zookeeper data dir

clientPort=2181 # server port for client connection

03课时 ZooKeeper分布式集群配置与Leader宕机测试

配置方式：

1. Standalone

1)解压文件到/usr/cloud,并且创建zookeeper软链

2)添加环境变量 /etc/profile, source /etc/profile生效 or restart

3)修改配置文件zoo.cfg,注意修改dataDir，不要使用/tmp,同时注意dataDir目录的权限775

tickTime=2000

initLimit=10

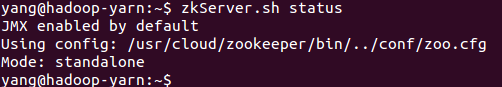
syncLimit=5

dataDir=/opt/zookeeper

dataLogDir=/opt/zookeeper/log

clientPort=2181

4)启动 zkServer.sh start



5)测试 zkCli.sh -server hadoop-yarn:2181(可以服务器上运行，也可以在其他的节点或者是客户端运行)，测试成功 ls 发现zookeeper

2. Pseudo\_destributed(different port)

避免端口一致实现伪分布式配置，生成多个zoo.cfg, 分别使用不同的端口

3. Fully\_distributed

Zookeeper声明集群中的每一个节点：

tickTime=2000 # heart beat recycle

dataDir=/tmp/zookeeper # zookeeper data dir

clientPort=2181 # server port for client connection

# server.n=hostname:port1:port2, the n must be the file myid value, the myid file is in the dataDir folder;

# port1 is leader port, when the host acts as leader, this port is for followers connect;

# port2 is electin port, when electing leader, the port2 is for other followers connect

server.1=hostname:port1:port2 [2888:3888]

server.2=hostname:port1:port2

server.3=hostname:port:port2

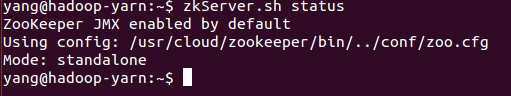
what’s more, the zookeeper cluster handles fault algorithm is 2n+1, so you’d better set the zookeeper cluster server number is odd number(奇数结点)，半数服务器节点挂掉，集群就会挂掉，因此数量最好是奇数

###########################################################

**touch the myid file in the dataDir,** 内容只有一个数字就是n 对应的是每一个节点上面的server.n.

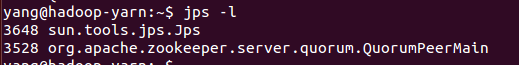
启动Zookeeper服务

zkServer.sh start| stop| status



因为我们仅仅配置了一个主机，所以是Mode：standalone

查看Java进程信息，QuorumPeerMain就是ZooKeeper的进程。



测试完全分布式ZooKeeper集群示例：

配置文件容易出现错误，因为cfg文件对于注释比较严格 # 后面要加一个空格

在单机模式下添加如下配置信息

server.60=hadoop-yarn:2888:3888

server.70=hadoop-yarn-170:2888:3888

server.71=hadoop-yarn-171:2888:3888

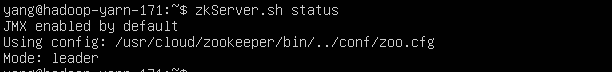
dataDir=/tmp/zookeeper 注意该目录的权限是775 或者是777，不然会出现权限错误

然后将同样的配置文件复制到不同的主机下，分别启动每一台服务器。

zkServer.sh start

zkServer.sh status







只有当ZooKeeper集群中大于一半的节点正常的时候，集群才是正常的，因此我们在开始启动的时候，第一台不报错，之后等到三台全部启动之后，就会正常了。

宕机测试（Leader）：

将Leader主机手动停机或者是关掉服务，因为我们是三台主机，其他的两台follower就会选举一个Leader

04课时 ZooKeeper分布式集群配置与NC linux命令的网络通讯利器

NC： netcat，瑞士军刀

nc -l <port> #让端口处于监听状态

nc hostname port #连接服务器端口，之后两者就可以发送消息

这样便可以将数据定向到指定的文件中

nc -l 1234 > log.txt

nc hoatname port < my.log

端口扫描：

nc ip -z -v port1-port2 扫描处于开放状态的端口

05课时 ZooKeeper分布式集群配置与zk四字符客户端指令

echo stat | nc hostname port

echo conf | nc hostname port

echo dump| nc hostname port

echo envi| nc hostname port

echo reqs| nc hostname port

echo ruok| nc hostname port

echo wchs| nc hostname port

echo wchp| nc hostname port

使用四字节指令夹NC工具，进行服务器之间信息的交互

06课时 ZooKeeper分布式集群配置与zkCli.sh客户端程序使用

zkCli.sh -server <servername:port>

ls <path>

create /yang hellozk

ls /yang

get /yang

set /yang new data

delete /yang

ZooKeeper集群中的数据是一致的，因此无论连接那一台主机，数据都是一致的

07课时 ZooKeeper分布式集群协同客户端API调用

String connectString="hadoop-yarn:2181,hadoop-yarn-170:2181,hadoop-yarn-171:2181";

**int** sessionTimeout = 2000;

ZooKeeper zk = **new** ZooKeeper(connectString, sessionTimeout,**null**);

Stat stat = **new** Stat();

**byte**[] res = zk.getData("/yang",**null**,stat );

System.***out***.println(**new** String(res));

08课时ZooKeeper分布式集群协同客户端API创建和删除PATH

09课时ZooKeeper分布式集群协同客户端API观察者模式与事件回调处理

Zk事件记过是one time trigger (一次触发) 这一部分内容没有学习明白

参考内容：

1. 徐培成ZooKeeper学习视频

2.