

极客学院
jikexueyuan.com

安装HBase的集群

安装HBase的集群 — 课程概要

- 课程体系介绍
- 安装HBase集群之前的准备
- 安装HBase的集群

课程体系介绍

课程体系介绍 — 课程体系介绍

- 如何安装HBase集群
- 数据的导入
- 如何使用HBase管理工具
- HBase的备份和恢复
- HBase监控与诊断
- HBase运维与安全管理
- 问题诊断
- 性能调优
- 高级配置与调优

课程体系介绍 — 学员要求

- 学员要求
 - 可以是HBase零基础
 - 最好懂一些HBase的基本概念和架构
 - 一定要有一些linux的基本知识
- 讲课方式
 - 以实践为目的
 - 全程演示

课程体系介绍 — 本课程要完成的工作

- 搭建一个集群的Hbase，为后面的实验准备实验环境

安装HBase集群之前的准备

■ 安装HBase集群之前的准备 — HBase的安装和运行模式

- 单节点模式 (Standalone Mode)
- 分布式模式 (Distributed Mode)
 - 伪分布式模式 (Pseudo-distributed Mode)
 - 全分布式模式 (*Fully-distributed*)

■ 安装HBase集群之前的准备 — 安装介质的准备

- HBase: hbase-1.0.1.1-bin.tar.gz
- 操作系统: ubuntu14_04.02
- Java JDK: jdk-7u79-linux-x64.tar.gz
- ZooKeeper: zookeeper-3.4.6.tar.gz
- Hadoop: Hadoop-2.6.x

安装HBase集群之前的准备 — **安装主机的准备**

- 准备三台虚拟机或者物理主机
- 配置三台主机的主机名和IP地址
- 配置ftp工具

■ 安装HBase集群之前的准备 — 创建hadoop组与hadoop用户

- 注意：在所有的主机上进行操作
- 创建组hadoop
 - `addgroup hadoop`
- 创建用户hadoop
 - `adduser --ingroup hadoop hadoop`
- 检查用户
 - `id hadoop`

■ 安装HBase集群之前的准备 — 配置安装路径

- 创建安装路径
 - 注意：路径权限的更改
- 也可以采用link的方式
- 我们使用的路径
 - /hadoop是根路径
 - /hadoop根路径下放置子路径，比如hbase的安装路径是/hadoop/hbase



安装HBase的集群

安装HBase的集群

安装HBase的集群 — 配置SSH

- 注意：如果没有安装ssh，需要提前进行安装
- 为了允许master node可以以无密码的方式登录到其他主机，需要配置hadoop用户的公钥
 - `hadoop@master$ ssh-keygen -t rsa -N ""`
- 在所有的从节点上，添加master节点hadoop用户的公钥信息
 - `hadoop@slave1$ mkdir ~/.ssh`
 - `hadoop@slave1$ cat >> ~/.ssh/authorized_keys`
- 检查master节点是否可以以无密码的方式登录到从节点

安装HBase的集群 — 安装Java 7

- 在主节点上安装Java 7
- 将安装好的Java同步到其他两个从节点
 - 例子: `rsync -avz /hadoop/java7/ slave1:/hadoop/java7/`
- 配置三个节点Hadoop用户环境变量.profile
- 检查java的版本
 - `java -version`

■ 安装HBase的集群 — 安装Hadoop

- 在主节点上安装Hadoop
- 修改hadoop-2.6.0/etc/hadoop/hadoop-env.sh，添加JDK支持：
 - `export JAVA_HOME=/usr/java/jdk1.8.0_25`

安装HBase的集群 — 安装Hadoop

- 修改hadoop-2.6.0/etc/hadoop/core-site.xml

注意：必须加在<configuration></configuration>节点内

```
<configuration>
```

```
<property>
```

```
<name>hadoop.tmp.dir</name>
```

```
<value>/tmp</value>
```

```
<description>Abase for other temporary directories.</description>
```

```
</property>
```

```
<property>
```

```
<name>fs.default.name</name>
```

```
<value>hdfs://master:9000</value>
```

```
</property>
```

```
</configuration>
```

■ 安装HBase的集群 — 安装Hadoop

- 修改hadoop-2.6.0/etc/hadoop/hdfs-site.xml

```
<property> <name>dfs.name.dir</name>
  <value>/home/hadoop/hadoop-2.6.0/dfs/name</value>
  <description>Path on the local filesystem where the NameNode stores the
namespace and transactions logs persistently.</description>
</property>
<property> <name>dfs.data.dir</name>
  <value>/home/hadoop/hadoop-2.6.0/dfs/data</value>
  <description>Comma separated list of paths on the local filesystem of a
DataNode where it should store its blocks.</description>
</property>
<property>
  <name>dfs.replication</name>
  <value>1</value>
</property>
```

■ 安装HBase的集群 — 安装Hadoop

- 可选：修改hadoop-2.6.0/etc/hadoop/mapred-site.xml

```
<property>  
  <name>mapred.job.tracker</name>  
  <value>master:9001</value>  
  <description>Host or IP and port of JobTracker.</description>  
</property>
```

■ 安装HBase的集群 — 安装Hadoop

- 修改hadoop-2.6.0/etc/hadoop/slaves

slave1
slave2

■ 安装HBase的集群 — **安装Hadoop**

- 将Hadoop分发到其他两个从节点

— 示例: `rsync -avz /hadoop/hadoop2.6/ slave1:/hadoop/hadoop2.6/`

■ 安装HBase的集群 — 安装Hadoop

- 格式化HDFS文件系统的namenode
 - `bin/hdfs namenode -format`
- 启动hdfs命令如下：
 - `sbin/start-dfs.sh`
- 成功的话输入jps，检查进程
 - jps(Java Virtual Machine Process Status Tool)是JDK 1.5提供的一个显示当前所有java进程pid的命令
- 或者通过浏览器访问<http://master:50070/>
- 停止hdfs命令如下：
 - `sbin/stop-dfs.sh`

安装HBase的集群 — 安装ZooKeeper

- 在主节点上安装ZooKeeper
- 将conf目录下面的 zoo_sample.cfg修改为zoo.cfg
- 修改zoo.cfg
 - dataDir=/usr/local/ZooKeeper/var/data
- 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，可以选择 1,2,3
- 在zoo.cfg中加入如下参数
 - server.1=master:2888:3888
 - server.2=slave1:2888:3888
 - server.3=slave2:2888:3888

■ 安装HBase的集群 — 安装ZooKeeper

- 启动各个节点的ZooKeeper服务器进程
 - `bin/zkServer.sh start`
- 检查安装是否成功
 - `bin/zkCli.sh -server master:2181`
 - `ls /`
 - `quit`
- 停止ZooKeeper
 - `bin/zkServer.sh stop`

安装HBase的集群 — 安装HBase

- 确保HDFS处于启动状态
- 在主节点解压缩Hbase
- 修改HBase的环境变量配置文件hbase-env.sh
 - `export JAVA_HOME=/usr/java/jdk1.6.0/`
- 修改hbase-site.xml

■ 安装HBase的集群 — 安装HBase

- 修改hbase-site.xml

```
<property> <name>hbase.rootdir</name>
<value>hdfs://master.Hadoop:9000/hbase</value>
</property>
<property><name>hbase.zookeeper.quorum</name>
<value>master,slave1,slave2</value>
</property>
<property><name>hbase.cluster.distributed</name>
<value>true</value>
</property>
<property>
<name>hbase.zookeeper.property.dataDir</name>
<value>/home/hadoop/data</value>
</property>
```

■ 安装HBase的集群 — 安装HBase

- 修改regionservers

```
slave1  
slave2
```

■ 安装HBase的集群 — 安装HBase

- 同步两个从节点
- 启动Hbase
 - `start-hbase.sh`
- 检查是否启动成功
 - `jps`

安装HBase的集群 — 测试HBase

- 利用客户端连接到HBase
 - `./hbase shell`
- 创建测试表、查看、插入数据
 - `create 'test', 'cf1'`
 - `list`
 - `put 'test', 'row1', 'cf1:a', 'value1'`
 - `put 'test', 'row1', 'cf1:b', 'value2'`
 - `scan 'test'`
 - `disable 'test'`
 - `drop 'test'`

安装HBase的集群

本套课程中我们详细学习了如何安装HBase的集群，你应当掌握了以下知识：

- 了解了HBase的几种安装方式
- 如何配置SSH的密钥
- 如何安装JDK
- 如何安装和配置Hadoop
- 如何安装和配置ZooKeeper
- 如何安装HBase
- 如何验证HBase是否安装成功

你可以利用本课程所学习的知识来完成HBase集群的安装，如果你想继续提高，欢迎继续关注极客学院HBase系统管理员的其他课程。

极客学院

jikexueyuan.com

中国最大的IT职业在线教育平台

