大规模分布式系统作业四（二）

MAC Hadoop2.8.5上安装hbase2.0.5

16302010002 李云帆

1. 前言：HBase是Hadoop的数据库, 而Hive数据库的管理工具, HBase具有分布式, 可扩展及面向列存储的特点(基于谷歌BigTable). HBase可以使用本地文件系统和HDFS文件存储系统, 存储的是松散的数据(key-value的映射关系).
2. Hbase安装前提&依赖:

* 匹配HBase的Hadoop版本
* Java JDK 1.6+
* SSH（否则需要输入三次密码）

1. 下载hbase binary文件并解压

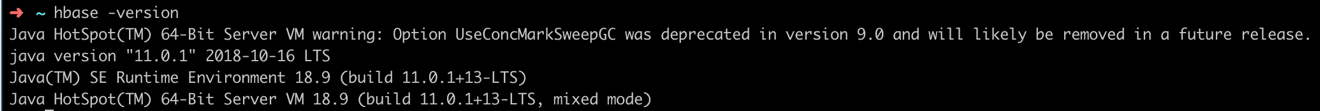
（略）

1. 配置conf/hbase-env.sh

加入export JAVA\_HOME="/XXX/XXX/XXX"

1. 配置conf/hbase-site.xml
2. **<configuration>**
3. **<property>**
4. **<name>**hbase.rootdir**</name>**
5. //这里设置让HBase存储文件的地方
6. **<value>**file:///XXXX/XXXX/hbase**</value>**
7. **</property>**
8. **<property>**
9. **<name>**hbase.zookeeper.property.dataDir**</name>**
10. **<value>false</value>**
11. **</property>**
12. **</configuration>**

1. bin/start-hbase.sh启动HBase，验证是否安装成功
2. $ jps
3. 3440 Jps
4. 3362 HMaster # 有HMaster则说明安装成功
5. 截图：



大规模分布式系统作业四（一）

回答三个问题

16302010002 李云帆

1. Hbase如何处理超过一台服务器存储memory的办法

HBase的服务器体系结构遵循简单的主从服务器架构，它由HRegion服务器（HRegion Server）群和HBase Master服务器（HBase Master Server)构成。HBase Master服务器负责管理所有的HRegion服务器，而HBase中所有的服务器都是通过ZooKeeper来进行协调，并处理HBase服务器运行期间可能遇到的错误。HBase Master Server本身不存储HBase中的任何数据，HBase逻辑上的表可能会被划分为多个HRegion，然后存储到HRegion Server群中，HBase Master Server中存储的是从数据到HRegion Server中的映射。

（数据利用tablet服务器来存储；用三层的、类似B+树的结构在Master服务器存储Tablet的位置信息；）

1. Hbase如何对数据读写进行操作

Hbase写数据流程

a)  Client发起了一个HTable.put(Put)请求给HRegionServer

b)  HRegionServer会将请求匹配到某个具体的HRegion上面

c)  决定是否写WAL log。WAL log文件是一个标准的Hadoop SequenceFile，文件中存储了HLogKey，这些Keys包含了和实际数据对应的序列号，主要用于崩溃恢复。

d) Put数据保存到MemStore中，同时检查MemStore状态，如果满了，则触发Flush to Disk请求。

e)  HRegionServer处理Flush to Disk的请求，将数据写成HFile文件并存到HDFS上，并且存储最后写入的数据序列号，这样就可以知道哪些数据已经存入了永久存储的HDFS中。

由于不同的列族会共享region，所以有可能出现，一个列族已经有1000万行，而另外一个才100行。当一个要求region分割的时候，会导致100行的列会同样分布到多个region中。所以，一般建议不要设置多个列族。

Hbase读数据流程

1. 客户端首先会根据配置文件中zookeeper地址连接zookeeper，并读取/<hbase-rootdir>/meta-region-server节点信息，该节点信息存储HBase元数据（hbase:meta）表所在的RegionServer地址以及访问端口等信息。用户可以通过zookeeper命令(get /<hbase-rootdir>/meta-region-server)查看该节点信息。
2. 根据hbase:meta所在RegionServer的访问信息，客户端会将该元数据表加载到本地并进行缓存。然后在表中确定待检索rowkey所在的RegionServer信息。
3. 根据数据所在RegionServer的访问信息，客户端会向该RegionServer发送真正的数据读取请求。
4. RegionServer接收到客户端的get/scan请求之后，先后做了两件事情：
   1. 构建scanner体系（实际上就是做一些scan前的准备工作）
   2. 在此体系基础上一行一行检索。
5. Hbase是列数据库吗？行，列存储的区别

是的

优缺点：

|  |  |  |
| --- | --- | --- |
|  | **行式存储** | **列式存储** |
| **优点** | * 数据被保存在一起 * INSERT/UPDATE容易 | * 查询时只有涉及到的列会被读取 * 投影(projection)很高效 * 列存储可以通过字典表压缩数据 * 任何列都能作为索引 |
| **缺点** | * 选择(Selection)时即使只涉及某几列，所有数据也都会被读取 | * 选择完成时，被选择的列要重新组装 * INSERT/UPDATE比较麻烦 |

