

第 1 章 HBase 简介

1.1 什么是 HBase

HBASE 是一个高可靠性、高性能、面向列、可伸缩的分布式存储系统，利用 HBASE 技术可在廉价 PC Server 上搭建起大规模结构化存储集群。

HBASE 的目标是存储并处理大型的数据，更具体来说是仅需使用普通的硬件配置，就能够处理由成千上万的行和列所组成的大型数据。

HBASE 是 Google Bigtable 的开源实现，但是也有很多不同之处。比如：Google Bigtable 利用 GFS 作为其文件存储系统，HBASE 利用 Hadoop HDFS 作为其文件存储系统；Google 运行 MAPREDUCE 来处理 Bigtable 中的海量数据，HBASE 同样利用 Hadoop MapReduce 来处理 HBASE 中的海量数据；Google Bigtable 利用 Chubby 作为协同服务，HBASE 利用 Zookeeper 作为对应。

1.2 HBase 中的角色

1.2.1 HMaster

功能：

- 1) 监控 RegionServer
- 2) 处理 RegionServer 故障转移
- 3) 处理元数据的变更
- 4) 处理 region 的分配或移除
- 5) 在空闲时间进行数据的负载均衡
- 6) 通过 Zookeeper 发布自己的位置给客户端

1.2.2 RegionServer

功能：

- 1) 负责存储 HBase 的实际数据
- 2) 处理分配给它的 Region
- 3) 刷新缓存到 HDFS
- 4) 维护 HLog
- 5) 执行压缩

6) 负责处理 Region 分片

1.2.3 其他组件:

1) Write-Ahead logs

HBase 的修改记录, 当对 HBase 读写数据的时候, 数据不是直接写进磁盘, 它会在内存中保留一段时间(时间以及数据量阈值可以设定)。但把数据保存在内存中可能有更高的概率引起数据丢失, 为了解决这个问题, 数据会先写在一个叫做 Write-Ahead logfile 的文件中, 然后再写入内存中。所以在系统出现故障的时候, 数据可以通过这个日志文件重建。

2) HFile

这是在磁盘上保存原始数据的实际的物理文件, 是实际的存储文件。

3) Store

HFile 存储在 Store 中, 一个 Store 对应 HBase 表中的一个列族。

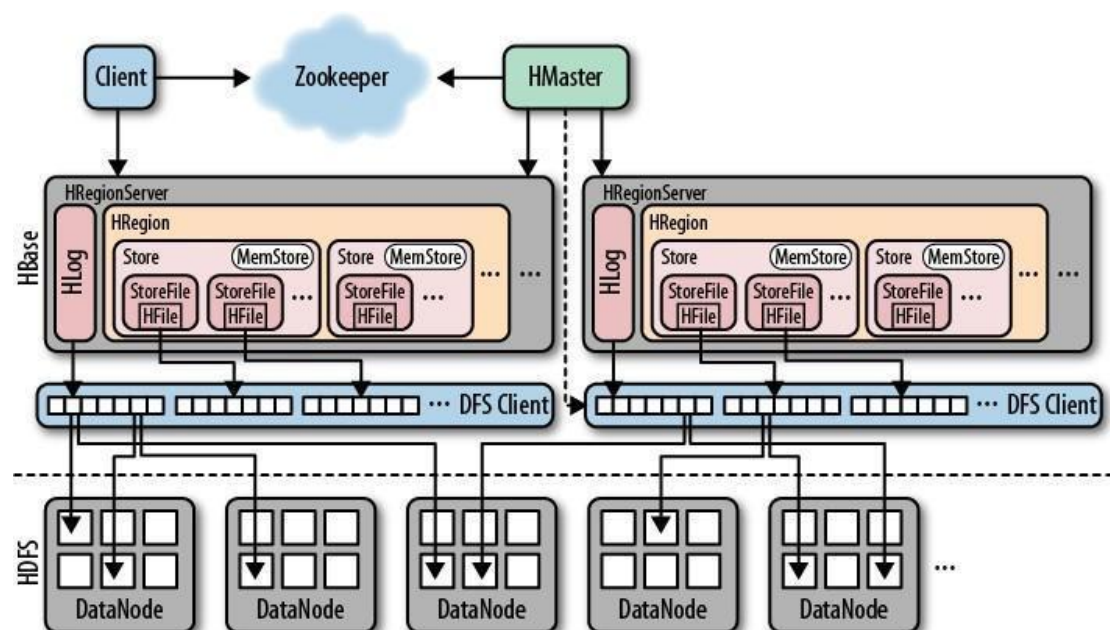
4) MemStore

顾名思义, 就是内存存储, 位于内存中, 用来保存当前的数据操作, 所以当数据保存在 WAL 中之后, RegionServer 会在内存中存储键值对。

5) Region

Hbase 表的分片, HBase 表会根据 RowKey 值被切分成不同的 region 存储在 RegionServer 中, 在一个 RegionServer 中可以有多个不同的 region。

1.3 HBase 架构



HBase 架构图