Hbase数据存储

a.数据分片存储

在HBase中，一张表的数据会被分成几份，每一份数据为一个region；每个region内存储的key是连续范围内的，不同region存储的key范围不重合；这些region可能被存储在同一台机器上，也可能存储在不同的机器上。HBase作为一个分布式数据库，对数据进行分片，可以提升吞吐量。(将数据根据rowkey发送到不同region)

###### b. HLog：Write-Ahead-Log，写操作先写日志

HLog的作用是，当一台regionServer crash了，可以利用HLog来恢复内存中未持久化到硬盘中的数据。需要注意的是，同一台server上的所有region共用一个HLog实例，因为假如每个region拥有一个独立的HLog，服务器会花费很多时间在磁盘寻道上

c. MemStore：写缓存，每个store拥有独立的写缓存

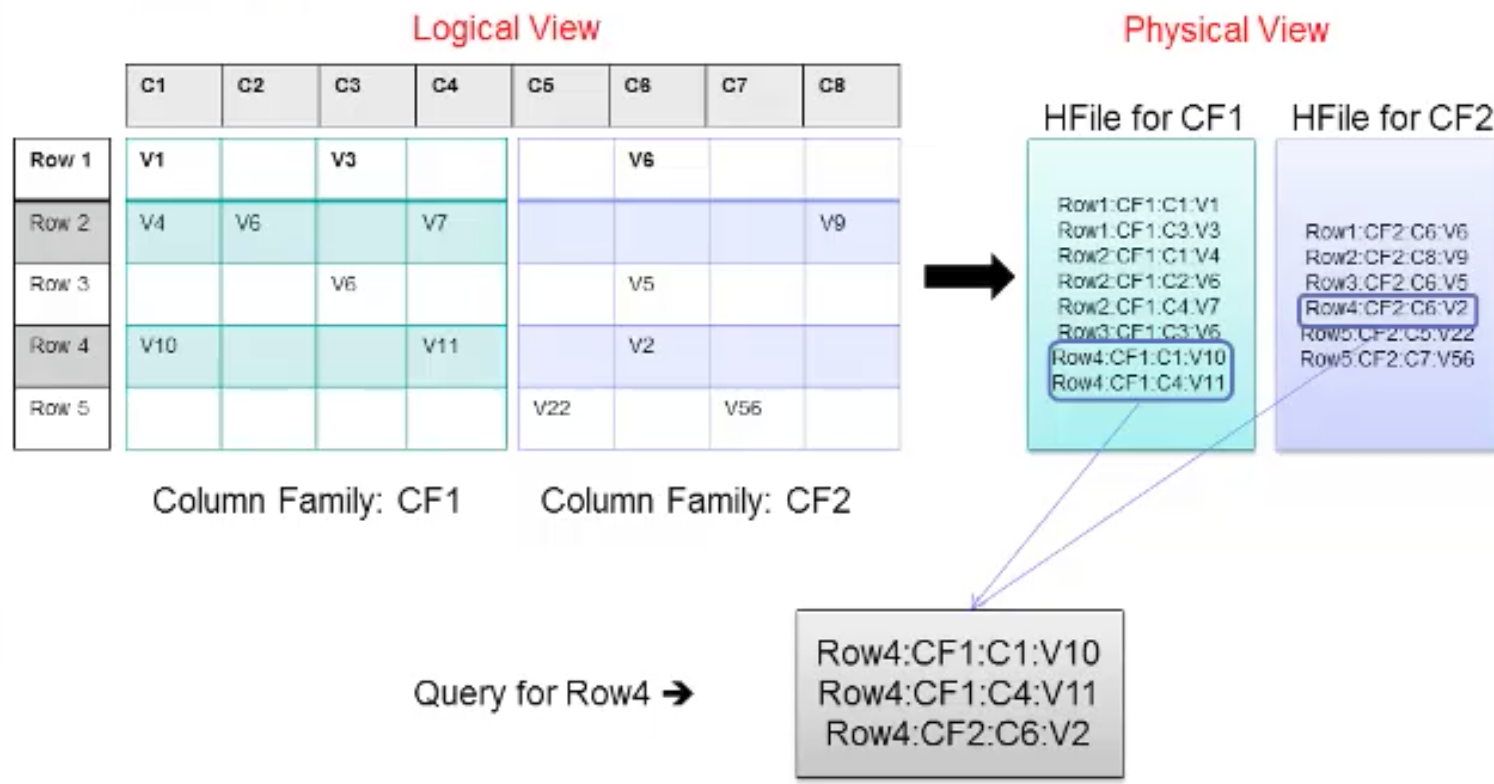
在HBase中，所有的写操作全部写到内存中，当写缓存（MemStore）写满，再刷写（flush）到磁盘中[2]，形成一个新的文件。这样做的目的，是为了高速响应那些写请求

d. HFile：磁盘文件

在存储上，HBase完全依赖HDFS，磁盘操作是直接调用HDFS的API（HDFS在维持data locality这一点上足够智能）。另外，之前提过HBase定义列族的一个原因是为了方便存储，事实上，同一列族的数据会被写到同一文件，因为存储特性本来就是按照列族定义的。HBase的数据在底层文件中时以KeyValue键值对的形式存储的，HBase没有数据类型，HFile中存储的是字节，这些字节按字典序排列。

e. 读缓存：同一server上所有region共用

既然HBase有写缓存，相对应的应该有读缓存。与写缓存不同的是，HBase的读缓存是同一server上的所有region共用的。当HBase读取磁盘上某一条数据时，HBase会将整个HFile block[3]读到cache中[4]。因此，当client请求临近的数据时，因为临近数据已经被缓存到内存中，HBase的响应会更快，也就是说，HBase鼓励将那些相似的，会被一起查找的数据存放在一起。另外，当我们在做全表扫描时，为了不刷走读缓存中的热数据，千万记得关闭读缓存的功能

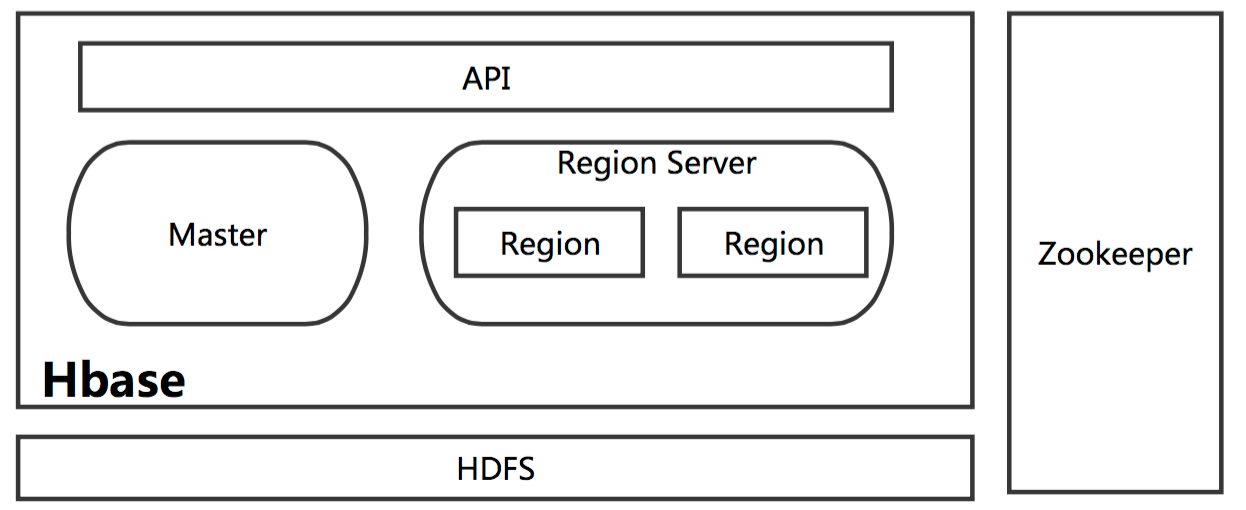


#### 上图：Hbase 中逻辑上数据的排布与物理上排布的关联

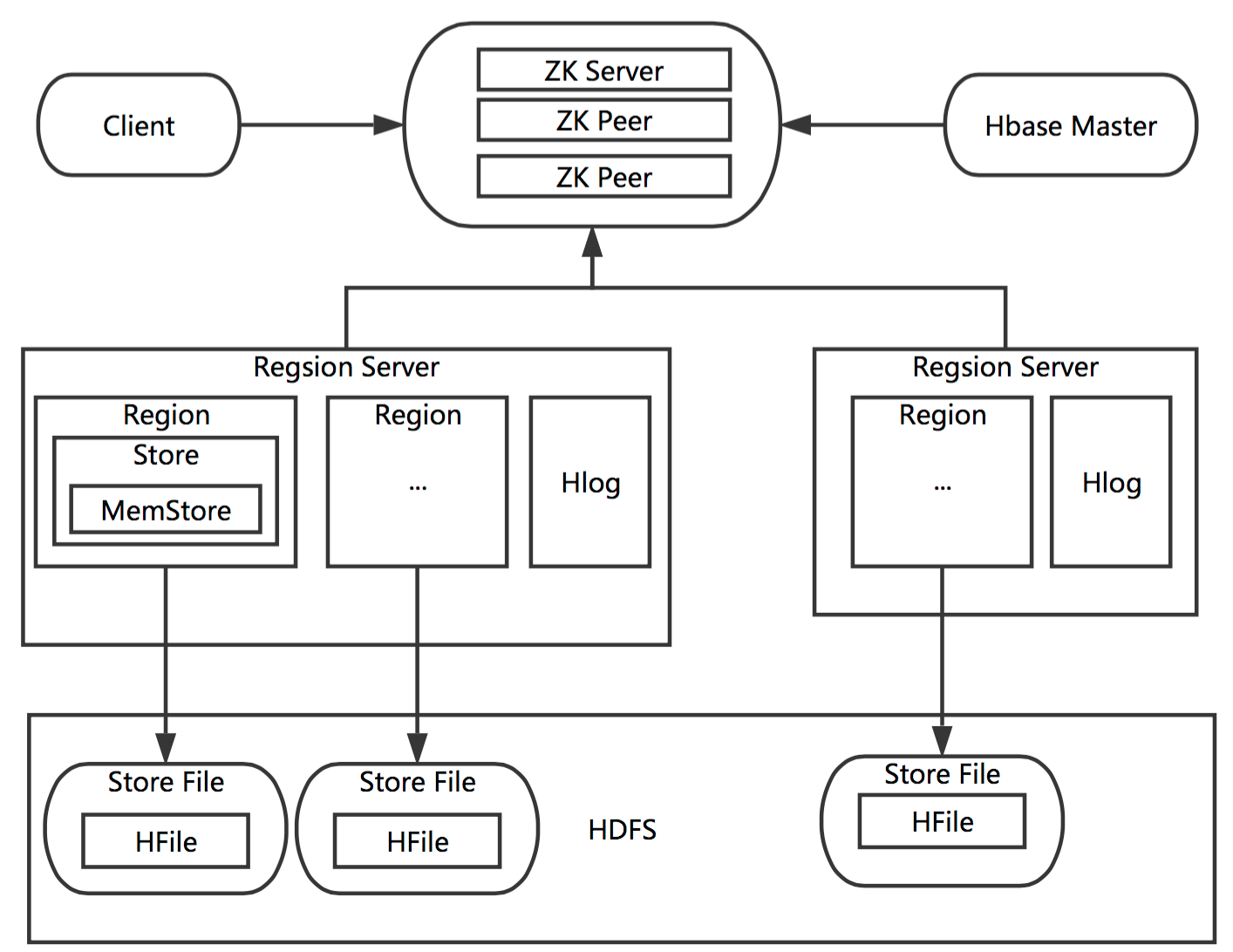
从上图我们看到 Row1 到 Row5 的数据分布在两个 CF 中，并且每个 CF 对应一个 HFile。并且逻辑上每一行中的一个单元格数据，对应于 HFile 中的一行，然后当用户按照 Row-key 查询数据的时候，HBase 会遍历两个 HFile，通过相同的 Row-Key 标识，将相关的单元格组织成行返回，这样便有了逻辑上的行数据。讲解到这，我们就大致了解 HBase 中的数据排布格式，以及与 RDBMS 的一些区别

逻辑上类似表的存储格式，其实看起来就是好多的空值组成，稀疏矩阵类，

而实际的物理存储上，根据CF来分割存储，然后CF在根据rowkey分割region存储到不同的server上。



#### 上图：HBase 的相关模块



当一个 Client 需要访问 HBase 集群时，Client 需要先和 Zookeeper 来通信，然后才会找到对应的 Region Server。每一个 Region Server 管理着很多个 Region。对于 HBase 来说，Region 是 HBase 并行化的基本单元。因此，数据也都存储在 Region 中。这里我们需要特别注意，每一个 Region 都只存储一个 Column Family 的数据，并且是该 CF 中的一段（按 Row 的区间分成多个 Region）。

Region 所能存储的数据大小是有上限的，当达到该上限时（Threshold），Region 会进行分裂，数据也会分裂到多个 Region 中，这样便可以提高数据的并行化，以及提高数据的容量。每个 Region 包含着多个 Store 对象。每个 Store 包含一个 MemStore，和一个或多个 HFile。MemStore 便是数据在内存中的实体，并且一般都是有序的。当数据向 Region 写入的时候，会先写入 MemStore。当 MemStore 中的数据需要向底层文件系统倾倒（Dump）时（例如 MemStore 中的数据体积到达 MemStore 配置的最大值），Store 便会创建 StoreFile，而 StoreFile 就是对 HFile 一层封装。所以 MemStore 中的数据会最终写入到 HFile 中，也就是磁盘 IO。由于 HBase 底层依靠 HDFS，因此 HFile 都存储在 HDFS 之中。这便是整个 HBase 工作的原理简述。

<https://www.ibm.com/developerworks/cn/analytics/library/ba-cn-bigdata-hbase/index.html>

hbase是一种专门为半结构化数据和水平可扩展设计的数据库，hbase表里的数据按照一个四维坐标系统来组织：行键、列族、列限定符和时间版本.hbase是无模式数据库，只需要提前定义列族，他也是无类型数据库，把所有数据不加解释的按照自己诶数组存储。