## HBase-Index

本项目中我主要负责的内容有：

**调研部分：**

Hbase二级索引常用实现方案

对比不同索引实现方案的优劣

对比全局索引

项目介绍ppt

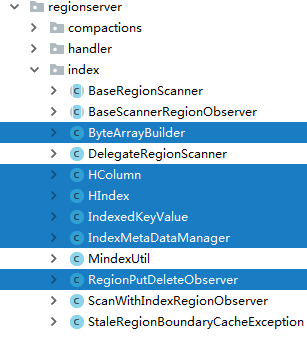
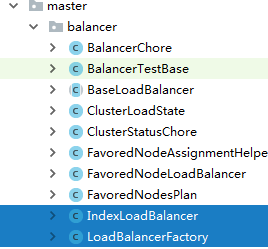
以上内容都包括在了项目说明ppt里。

**开发部分：**

移植MarsBase项目的Balancer代码及Split代码到本项目中。（这部分参见MarsBase那个文档）

索引一致性保证相关代码改进。

**代码位置：**

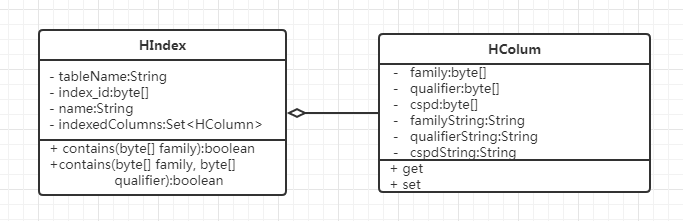
 

hbase-server项目中的org.apache.hadoop.hbase.regionserver.index包下

Hbase-server项目中的org.apache.hadoop.hbase.master.balancer 包下

**一、添加了的类：**

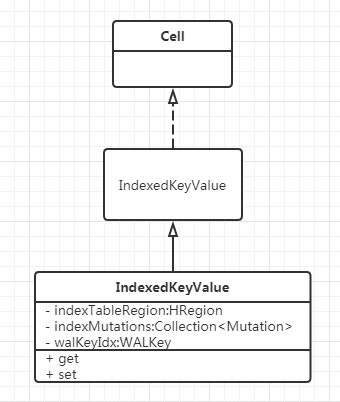
HIndex和HColum分别封装了索引信息和索引列的信息



IndexedKeyValue 继承了KeyValue类，用于封装索引region、索引操作、索引walKey。

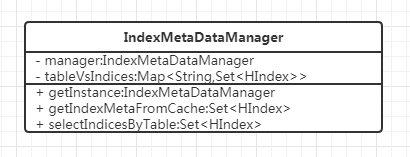
可以作为一个cell添加到主表的WalEdit，传递到主表主流程中来构建索引的WAL内容。

也可以传递到协处理器的post方法中。



IndexMetaDataManager

索引元数据管理类，提供索引元数据的查询方法

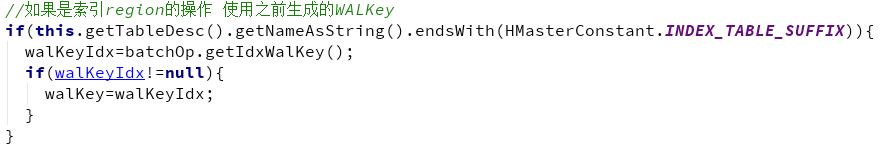


1. **索引数据实时写入流程（对HRegion类的修改）：**

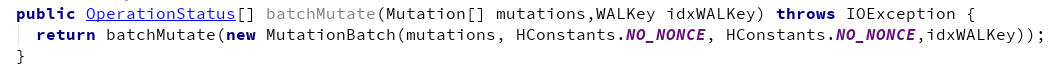
修改目标是实现索引的WAL和主表的WAL可以同时写入和落地，保证索引的最终一致性1. STEP 4. Append the final edit to WAL. Do not sync wal. 添加：



1. 索引写流程中，由于设置属性SKIP\_WAL,所以walKey==null， wal会append的一个空的walEdit，然后更新这次空walEdit的mvccNum。应该防止这样的情况发生，将在主表写流程中生成的索引walKey传递到本次流程中，是其更新的是此walKey的mvccNum.



1. 新添加了一个方法，使索引的walKey能从主表写流程，传递到索引写流程



同时修改了BatchOperationInProgress类



流程图：

