

07 从零实现 KV 存储—数据删除流程



删除数据的流程其实和写数据差不多，根据 bitcask 论文的描述，删除数据实际上也是向数据文件中追加一条记录，只是需要标识其类型是被删除的，也就是类似的一个墓碑值。

在我们的 bitcask KV 存储引擎实现中，也是如此，和 Put 数据一样，我们仍然先判断用户传入的需要删除的 key 是否有效，如果传入的是一个空的 key 值，则直接返回一个报错信息。

然后我们再根据这个 Key 去内存索引中查找，如果没找到的话，说明这个 key 是根本不存在的，我们直接忽略掉这次操作即可，方法直接返回。

需要加上这个判断的原因是，如果用户一直调用 Delete 方法去删除一个不存在的 key，在没有任何判断的情况下，仍然会追加一条记录到数据文件中，这样可能会导致数据文件的膨胀。

但是这样做并没有什么意义，反而占据了多余的磁盘空间，并且在合并数据文件的时候，也会扫描多余的数据，导致效率比较低下，所以我们加上这个判断，如果用户对不存在的 key 执行删除，则什么都不做，直接跳过，这样可以避免后续的很多无用功，更加的简洁。

都通过前面的判断之后，表示我们需要删除的是一个有效的 key，我们需要构造一个 LogRecord，只不过这个结构中不需要存储具体的 value，只是标识其类型是被删除的，如下所示：

```
1 // 这里标识 Entry 的类型是删除
2 record := &data.LogRecord{Key: key, Type: data.LogRecordDeleted}

1 let mut log_record = LogRecord {
2   key: key.to_vec(),
3   value: Default::default(),
4   rec_type: LogRecordType::DELETED,
5 };
```

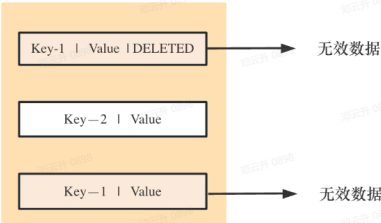
构造好之后，和 Put 一样，调用追加写数据文件的 appendLogRecord 方法即可。

写完之后，如果没有什么错误的话，则更新内存索引。只需要调用内存索引的 Delete 方法，将 key 对应的数据删除掉。

如果删除的时候发生了错误，也就是说更新内存失败，也需要返回一个错误信息。

被删除的数据何时清除

被删除的数据，在磁盘上仍然存在，例如下面的例子：



依次向文件中添加了 key-1 和 key-2，然后我们删除了 key-1，这时候向文件中新增了一条记录，表示 key-1 被删除了，这时候磁盘上的 key-1 数据就是无效的了，我们在后续的 Merge 过程中，会将这样的数据清理掉。

启动时对删除的处理

前面处理完之后，其实还有一个地方需要注意，因为在重启数据库的时候，如果我们不对已删除类型的数据进行处理的话，内存索引是不知道的，那么被删除的数据其对应的索引仍然存在，会导致已经被删除的数据又存在了，数据会发生不一致。

所以在启动 bitcask 实例，从数据文件中加载索引的时候，需要对已删除的记录进行处理，如果判断到当前处理的记录是已删除的，则根据对应的 key 将内存索引中的数据删除。

逻辑类似下面这样：

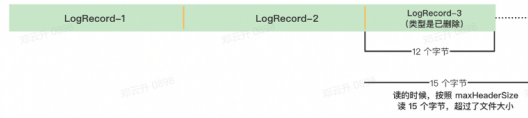
```
1 // 构造内存索引(信息并存储
2 pos := &data.LogRecordPos{Fid: uint32(fileId), Offset: offset}
3 // 如果是删除类型的数据，则删除对应的内存索引信息
4 if logRecord.Type == data.LogRecordDeleted {
5   db.Index.Delete(logRecord.Key)
6 } else { // 否则插入或者更新内存索引
7   db.Index.Put(logRecord.Key, pos)
```

一个特殊 case

搞懂这个 case，需要把后面的对 LogRecord 的编码部分看完之后，才会更加理解这里的描述。

当删除数据时，我们会追加一条记录到文件末尾，如果这条数据恰好是数据文件的最后一条记录，并且这条记录的大小没有超过 maxHeaderSize，但是我们在读取 header 的时候，是按照 maxHeaderSize 来读的，所以有可能会造成 EOF 错误。

大致如下图所示：



所以针对这种情况，我们需要判断读取的时候，是否超过了文件的大小，否则的话，只读取到文件的末尾即可。

以上处理完成后，数据删除的完整流程就基本实现了。



5 人点赞



全文评论



石钟 4月26日 12:44

go 语言中的 Delete105 行，err != nil 的条件下，应该返回 err 吧，为啥是返回 nil 呢



rosesduan 4月26日 18:27

@石钟 能截图发群里吗

输入评论

