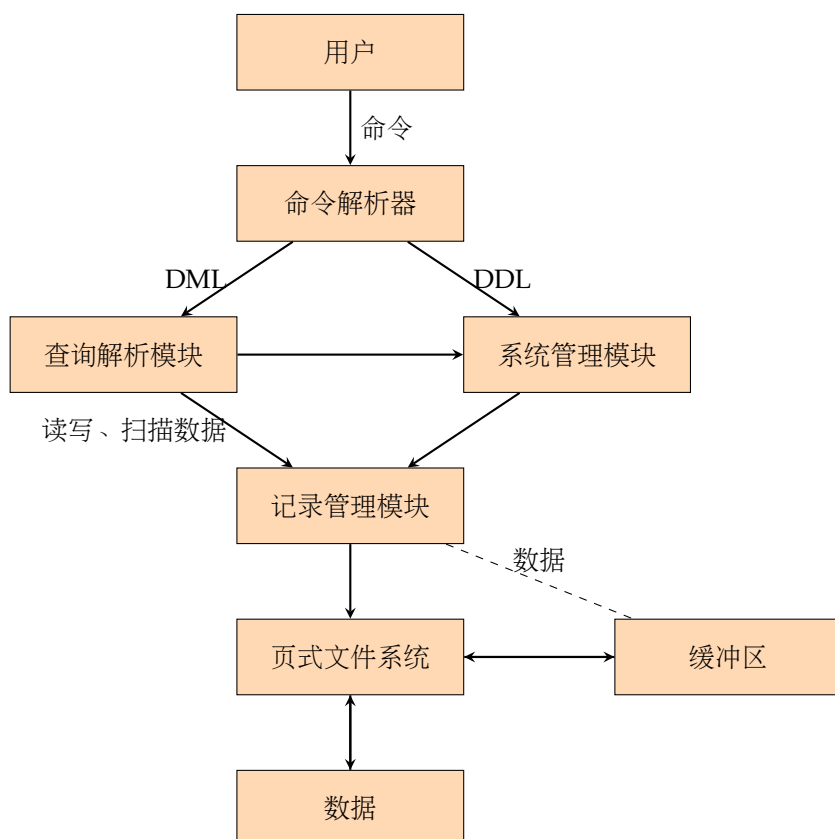


数据库系统概论项目期末报告

徐炜杰 黄欢

January 1, 2016

1 系统结构设计



2 系统功能

2.1 记录管理模块

1. 新建文件、删除文件、打开文件、关闭文件
2. 插入记录、删除记录、更新记录、获取属性值满足特定条件的记录

2.2 系统管理模块

1. 创建数据库、删除数据库、切换数据库、列出当前数据库包含的所有表
2. 创建表、删除表、列出表的模式信息

2.3 查询解析模块

1. 命令解析器
2. 数据类型判断
3. 支持空值操作
4. 连接两个表

2.4 拓展模块

1. 模糊匹配
2. 三个表以上的连接
3. 聚集查询
4. 分组聚集查询

3 主要模块设计原理

3.1 记录管理模块

3.1.1 文件操作

对文件夹、文件的新建、删除、打开、关闭等操作均使用Unix系统下提供的接口实现。

3.1.2 记录操作

对于一个表而言，第零页存储表的元数据，包括列名、数据类型、主键信息等。第一页开始存储表的数据，每一页的开始记录该页的空槽数，每一个槽中维护当前槽中的记录信息和一个指向下一个空槽的指针，其中记录信息包括每个列项是否为空的标记位和每个列项的数据。记录是定长存储的，整数类型数据均采用长整形存储，变长字符串类型数据均采用定长字符串存储。

在进行记录相关的操作之前，需要先获取表的元数据，以明确每个数据项的类型和范围。对于插入操作，依次寻找每一页，找到第一个空槽进行插入，并更新槽指针。对于删除、更新和查询操作，首先需要遍历表中所有记录，并得到所有符合Where子句条件的记录。对于删除操作，需要将记录对应的槽清空，并更新槽指针。对于更新操作，需要将记录的对应键值修改为给定表达式的值。对于查询操作，需要计算出每个表达式的值，并进行输出操作。

记录管理模块仅实现上述操作中有关物理存储的功能，即提供记录的遍历、删除和更新操作，其余功能由查询解析模块完成。

对记录进行面向对象的类封装，有利于记录类和缓存之间的相互转换。

3.2 系统管理模块

3.2.1 文件操作

一个数据库对应于一个文件夹，一个表对应于一个文件，对文件夹、文件的操作可以调用记录管理模块提供的接口。

3.2.2 数据库相关操作

执行创建数据库命令时，创建新的文件夹。执行删除数据库命令时，删除指定文件夹及所有文件。执行切换数据库命令时切换到指定文件夹下。

3.2.3 表相关操作

创建表、删除表时，在当前数据库目录下进行创建文件、删除文件操作。

创建表操作时，需要将表的元数据写入文件第一页中。对元数据进行面向对象的类封装，有利于元数据和缓存之间的相互转换，以及记录相关的操作。

3.3 查询解析模块

3.3.1 命令解析

使用Flex和Bison工具对输入的命令进行词法解析和语法解析，Github开源项目链接详见参考文献。

命令解析完成后，调用记录管理模块和系统管理模块的对应接口来实现对应操作。为了提高效率，记录管理模块和系统管理模块的接口均支持插入、删除、更新和查询操作多条记录批处理的功能。

3.3.2 条件操作

删除记录、更新记录和查询记录都会用到Where子句，通过调用记录管理模块的接口，遍历表中的所有记录，对每一条记录依次计算条件表达式的值，给出符合条件的记录的序号。

3.3.3 表达式操作

更新记录和查询记录都会用到表达式，对于符合条件的记录，分别计算更新语句或查询语句的一系列表达式的值。对于更新记录操作而言，将一系列表达式的值分别赋予对应的列项即可。对于查询记录操作而言，将表达式的值输出到控制台即可。

3.4 拓展模块

3.4.1 模糊匹配

C++11 支持正则表达式匹配，因此直接调用 `regex` 头文件中的接口即可完成模糊匹配。

3.4.2 三个表以上的连接

原理和两个表的连接相同，依次遍历每一个表的每一条记录，并进行多条记录的条件计算和表达式计算即可。

3.4.3 聚集查询

把聚集函数视为表达式操作中的一个项，在递归进行表达式计算过程中，计算到该项时，对相应的数据进行求和、求数量和比较大小的操作即可。

3.4.4 分组聚集查询

在聚集查询的基础上，对分组的键值建立 `map` 映射，根据键值来进行求和、求数量和比较大小的操作即可。

4 实验结果

实现了上文提到的所有功能，并通过了给定的样例指令以及部分自己设计的指令的测试，用Select语句输出指令结果，结果均正确。

性能方面，运行给定的 `book.sql` 样例，需要 15 秒，连接两个表需要 30 - 40 分钟。

5 小组人员分工

徐炜杰	记录管理模块、系统管理模块和查询解析模块的总体设计
黄欢	实现命令解析器、扩展模块和数据库总体测试

6 参考文献

1. 冯建华,周立柱,郝晓龙.数据库系统设计与原理 [M] . 北京:清华大学出版社,2007
2. 命令解析器参考链接: <https://github.com/junkumar/redbase>