|  |
| --- |
| http://www.wpi.edu/~vernescu/logo.png |
| Assignment 3 |
| CS 542: Database Management Systems |
|  |
| **Tyler Carroll, Tom Strott and James Silvia** |
| **3/31/2015** |

Contents

[Table of Figures 1](#_Toc415661984)

[1. Introduction 2](#_Toc415661985)

[2. Assumptions and Limitations 2](#_Toc415661986)

[3. Client/Server Model 3](#_Toc415661987)

[4. Relation Class 4](#_Toc415661988)

[4.1 Open 5](#_Toc415661989)

[4.2 GetNext 6](#_Toc415661990)

[4.3 Close 6](#_Toc415661991)

[5. Isolation Manager 6](#_Toc415661992)

[5.1 Handling the Join 7](#_Toc415661993)

[6. Memory Manager Class 7](#_Toc415661994)

[6.1 Database Table 7](#_Toc415661995)

[6.2 Memory Manager 8](#_Toc415661996)

[6.3 Memory Map 9](#_Toc415661997)

[7. B+ Trees 9](#_Toc415661998)

[8. Test 10](#_Toc415661999)

# Table of Figures

[Figure 1: Database Server Flow 4](#_Toc415662000)

# 1. Introduction

For project 3 our team has implemented a query executioner on top of last project’s value store database with B+ tree indexing. Source code for this project can be found at <https://github.com/jamesesilvia/CS542>. The query executioner consists of an interface that accepts an integer value as a percentage.

1. Query(int percentage);

The query execution has been hard-coded to perform a join of two tables, a city and country table, and perform analysis based on the percentage provided. The query will return a list of cities and corresponding countries if the population of the city makes up at least the given percentage of its country. The example below shows the equivalent SQL query with 40%.

The SQL query is equivalent to:

Select Code, dbo.city.Name, dbo.city.Population from dbo.country

INNER JOIN dbo.city on dbo.country.code = dbo.city.CountryCode

where dbo.city.population >= (dbo.country.population \* 0.4);

Most of the work in Project 3 has been improving the Relation class to handle multiple tables and handling the join of these tables on a query request. Additionally, a database loader was implemented to import both city and country tables from the SQL database to our application.

# 2. Assumptions and Limitations

The schema used for this project was from the world project that is used for MySQL training. We ported this file to SQL before getting started but no data was impacted by the port. The changes mostly consisted of removing quotes around database objects and converting the MySQL enums to a SQL CHECK statement with a default value.

Our application will load all data entries included in the city and country table in to two separate databases on initialization. The data remains persistent on application exit. We do not allow users to add, remove, or edit any of the data stored in the database. We also expect that the schema files used for initialization and the database files are not tampered with during execution. Our code assumes that each entry in the city table has a country code, and each country code only appears in one country table entry.

We provided the ability to execute the query given varying percentages to show the modularity and power of our application.

# 3. Client/Server Model

Our value store database runs in a client/server model that use socket communication to transfer information. The client is a command line interface for a user that accepts commands “query”, and “quit”, where “query” will prompt for a percentage value (0-100).

The database represents the server and is made up of several threads. This server is in charge of handling users requests, managing the information in the database, all while maintaining the integrity of the data stored in the database. Databases often receive multiple requests in an extremely short period of time and must prevent these accesses from overwriting or damaging any information.

The first server thread initializes the server and waits for new client connections. On each client connection a new thread will be spawned in the server to handle the user’s requests to put, get, or remove information from the database. These “client handlers” are responsible for determining the intent of the client and forwarding that information to another piece of the server. These threads can be linked to the database through a Relation. The client handler server requests have a 1:1 mapping to the information in the database (one Relation for one Table).

The manager that handles all requests from the client handlers is referred to as the Isolation Manager and is a single thread on the server. It protects the integrity of the database by acting as a traffic cop of user requests. Once the traffic cop allows information to be accessed it initiates a request to the database manager.

This can be summed up in the diagram below. Further details will be explained shortly. Note in this project, much like the last one, we only have one table in our database so there is only one isolation manager and one relation.

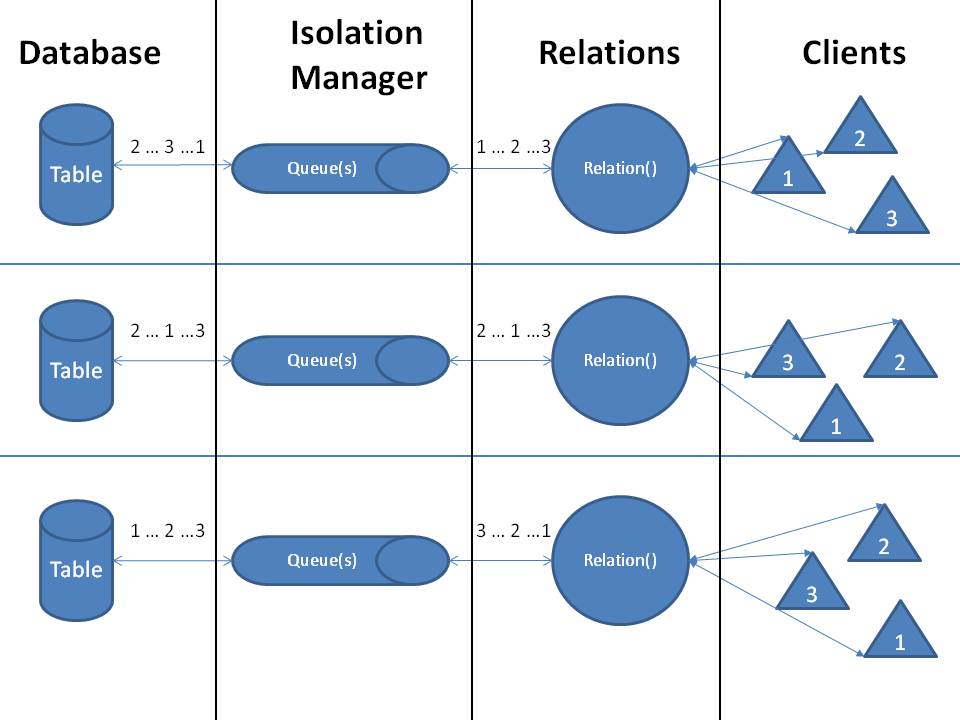


Figure 1: Database Server Flow

## 3.2 Database Initialization

On first execution of the database server two “.csv” files are parsed that contain the information from the MySQL database. Each table entry is stored in the database and indexed in our B-Tree through the Relation class. This information is persistent on each server shutdown/startup; but if either database file is removed; the initialization will take place again. It can be lengthy because we are loading thousands of rows in to our database.

If the database has already been loaded, the only initialization required on startup is re-building the B+ Trees to gain access to the information stored in the database. The initialization makes use of Open, and Close in the Relation. More details in the sections that follow…

# 4. Relation Class

The relation class contains the interfaces used to access the database. There is a single relation for each table that is represented in the database server. For this specific project, there are two instances of a Relation.

The Relation class contains two queues, service\_queue and done\_queue, and is shared across all client handling threads. The queues allow proper request handling when the server is heavily loaded. The Relation class is in charge of placing new requests on to the service queue, removing the same request once complete from the done queue, and forwarding the server response to the correct client. This abstracted block can exclusively put on to the service queue and remove from the done queue.

For project 3 we altered how the Relation class works and how it interacts with the Memory Manager class. In previous projects the Memory Manager class had a single global instance that could be accessed from anywhere because there was only one table in the database. Now that we have multiple tables and there exists both a Relation object and Memory Manager object for each table, we decided to get rid of the global Memory Manager object. We added a Memory Manager object to the Relation class so each Relation object would have its own Memory Manager object that it can use to interact directly with its own table. This also allowed us to let the Relation class have more control over the Memory Manager. We added three new functions to the Relation class for this project, open, getNext and close, which utilize the extra control that the Memory Manager object provides. These functions will be explained in the next sections.

## 4.1 Open

Open is a new function added to the Relation class in order to allow the Relation to choose when the table is mapped into memory. Previously the entire database was mapped into memory at all times on server initialization and kept in memory until the server was shut down. Now that a Memory Manager object is a member of the Relation class, the Relation has control when to map and unmap the table it has control over. The open function allows the Relation to load the table into memory before any accesses occur in order to save memory.

## 4.2 GetNext

In previous projects all accesses to a table were done to access a specific tuple. All reads either asked for a tuple at a specific index or a tuple with a specific indexed value. In order to perform a query and a join we needed a way to be able to get every tuple one at a time. The getNext function allows the Relation object to retrieve one tuple from its table at a time until it runs out of tuples. This functionality allowed us to perform a join by retrieving tuples one at a time and performing the necessary comparisons to see if they should be joined or not. Before getNext can be used the table needs to be loaded into memory using open. Once the join completes the query can use close to unload the table from memory.

## 4.3 Close

The close function performs the opposite action as the open function. Close allows the relation to unmap the table from memory whenever an access has completed. The addition of open and close give the database control of which tables exist in memory at a given time and ensure that tables that are currently not being used do not exist in memory. These functions decrease the average memory use of the server.

# 5. Isolation Manager

The isolation manager is its own thread and is included in the Relation class. However, it can be abstracted out for this document. The isolation manager removes requests to be serviced, determines the action, and forwards it on to the database manager. Once the database manager has handled the request, the isolation manager puts the request on the done queue for the Relation to return the information to the client.

This abstracted block can exclusively remove from the service queue, and put on to the done queue.

## 5.1 Handling the Join

The join is a one-pass algorithm taking place within the isolation manager. It resides in a while loop until there is no next record to get from the city table, iterating through all records of the city table exactly once. GetNext will return a city record to be processed by the join. With the country code of that city table record, we can use an indexed lookup on the country table to efficiently find the country record that corresponds with the city record. From there, we take the user’s percentage input to scale down the country population and compare that with the city population. If the city population is equal or greater, that record is added to the list of entries to be returned to the requesting client. This is why we mandate that all country codes are distinct in the country table and that all cities must have a country code.

# 6. Memory Manager Class

The Memory Manager class is in charge of managing the location of information in the database. The manager has several functionalities, each of which will be described below. The functionalities it has are the database table, memory manager and the memory map. All of these functions are performed by the Memory Manager class. The database manager has one Memory Manager object per table. In previous projects there was only one table so a single global Memory Manager object existed. In project 3 we moved the Memory Manager object into the Relation class for greater database control.

## 6.1 Database Table

The database table keeps track of all the indices in the database. The database table is a list that is ordered by ID, where the ID is unique. The ID cannot be reused unless a table entry has been removed from the database entirely. The unique ID links a table entry to a specific, physical storage location in the table’s database.

## 6.2 Memory Manager

All data that is stored in the database fits inside of a predetermined container. For this project, the container is made up of the unique ID discussed above and the corresponding value and string. For this application, it can be described as an ID, a population value, a location string, and a country code. The container will not shrink or grow based on what it is stored because it is a \_\_packed\_\_ structure and the sizes are limited on input. This simplifies the relationship between the unique ID and its corresponding physical location in storage.

The memory manager is responsible for physically placing the above container in to the database, therefore securing that the data can be recovered. If for any reason the database needs to be expanded (size increase) it handles that as well. The isolation manager will request to the memory manager the request of storing, retrieving, or removing information stored in the database.

The unique ID (UID) that information is stored keyed as links the physical storage location simply by the size of the container previously discussed. A quick example…

* Table Entry UID: 0. Storage Offset: 0
* Table Entry UID: 1. Storage Offset: 113
* Table Entry UID: 2. Storage Offset: 226

If a request is received to place more information in to the database, the memory manager will determine the next available UID, and store the information in that location. If a request to remove is received, the entry is cleared from the database table; which now opens the physical location and UID for a new value. Similarly on a get, the memory manager will extract the relevant information.

B+ Tree’s have been implemented as a searching algorithm to determine what information is stored inside of the database. Without giving too much information, they maintain lists of each table in a specific order for improved database access. They direct the memory manager to the correct UID. More information is available in Section 7.

## 6.3 Memory Map

Our database will be contained within several data files. Each table in the database will have its own file to simplify the memory management. In project 3 there are now two tables so there will be two separate files, one for each table. On the server we will use mmap() to map the entire file into memory. That way we can access individual offsets based on UID’s in the file Each write done to the mapped memory will be immediately written down to the actual file.

The database starts at a size of 1MB and expands in 1kB chunks until it reaches a maximum size of 2GB.

Instead of mapping the entire file into memory on server initialization, the Memory Manager gives the Relation class the functionality to map and unmap the table to and from memory whenever it wishes.

# 7. B+ Trees

B+ trees were implemented to maintain the physical storage location of a given index based on the data that is contained at the location. They are extremely efficient at storing data in block levels (our container). Due to the level of complexity associated with B+ Trees, a licensed, open sourced example has been modified to fit our application1. The modifications are as follows:

1. Conversion to C++ class from C code
2. Multiple data entries per key support
3. Creation of a separate class of B+ Tree that uses strings as the key and creates an alphabetically sorted B+ tree.

Within our application, there is a 1:1 relation between the number of B+ Tree’s and the data that we want to index. In Project 3 the group used B+ Tree’s to index the country code attribute of each tuple. By indexing this attribute we were able to perform our join much more efficiently. The query involved checking each tuple in the City table and comparing the tuple to the matching tuple in the Country table. The country code is an attribute in both tables. This means for each tuple in the City table we were able to easily find the matching tuple in the Country table by searching for the particular country code using our B+ Tree instead of iterating the entire Country table.

# 8. Test

To prove the functionality of our query command, the team performed queries in 10% increments starting at 10% through 100% and compared our application’s results to the corresponding SQL commands results. Included in our delivery are the outputs of our application’s queries (test/Output\_X.results) as well as an excel sheet that contains the SQL results. These SQL results were obtained by using the same query given in Section 1 with a varying constant (for example, 0.4, 0.5, and so on).

The team also verified that a query with 0% population returned an inner join without a where clause of both tables by using the length of tables as a verifier along with country spot-checks to ensure correct permutations.

References

1. GPL Licensed B+ Tree Source Code. Accessed: 3/1/2015.

<http://www.amittai.com/prose/bpt.c>