

CS307 Project 1: The performance of DBMS

- It is a personal project. Each student should do it separately and submit one report for the project.
- You should submit a report before the deadline. The report submitted after the deadline will be zero score.
- DO NOT copy ANY words, figures and others from Internet and your classmates.

DBMS can help us to manage data conveniently, and also can improve the efficiency of data retrieval greatly. Please compare data retrieval implemented by yourself using C/C++, Java or any other language with DBMS to find the differences.

Suggested Steps:

Find some data for experiments. You can use all films with titles, directors, actors and other related information. Some other data such as 'Infected cases by COVID-19', 'Stock data', 'Library data' etc. is also fine. But the size should be reasonably large.

Store the data into a database table. Then use select in SQL to find the films with word "XXX" in titles, and record the execution time.

Store the data into a file, and then load it into RAM. The data in RAM can be any format you preferred. Design an algorithm to search "XXX" in titles. Record the execution time of your algorithm.

Some other comparisons and experiments you would like to do. Such as you can reorganize data into some other format for faster retrieval. You are recommended to study the mechanism of DBMS for storing and retrieval

SUGGESTED structure of your report

1. Title

Please give your report a title that can reflect the core point of your work.

2. Introduction

To introduce what you want to prove in the **experiments**. **DO NOT copy any text from Internet.**

3. Experimental design

3.1 Experimental data and environments.

Introduce the data, DBMS and the programming language you choose. How you organize the data.

3.2 Experiments

Describe your experiments here. To make sure the experiments should be reasonable and the comparisons should be fare.

3.3 Experimental results

Show your results in tables and figures, and analysis the differences of different methods.

4. Conclusions

Give your conclusions here. You can also present you options on DBMS.

Several website can give you some ideas to research, but those are only theory description:

[1. Advantage of database management system over file system](#)

[2. Advantages of Database Management System](#)

[3. Characteristics and benefits of a database](#)

.....

Basic Requirement

1. Clearly and easy understanding design.
2. You need to show that you really do the experiments by some **experimental details** and reasonable analysis.

Bonus

If you can do some of the following functions (but not limit to the following ones), you will win bonus.

1. High concurrency and transaction management
2. User privileges management
3. Database index and file IO
4. Large data sets (the data is large enough and cannot be stored into RAM)
5. Compare performance of multiple databases with file system over different operation systems

Submit

Submit a pdf file into sakai website before March. 29th.

