

# **Monitoring Power Consumption by Changing Query Plan for Single Database System**

by

Arif Hazman Bin Shahir

Dissertation submitted in partial fulfillment of  
the requirements for the  
Bachelor of Technology (Hons)  
(Business Information Systems)

MAY 2014

Universiti Teknologi PETRONAS  
Bandar Seri Iskandar  
31750 Tronoh  
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

**Monitoring Power Consumption by Changing  
Query Plan for Single Database System**

by

Arif Hazman Bin Shahir

A project dissertation submitted to the  
Business Information System Programme  
Universiti Teknologi PETRONAS  
in partial fulfillment of the requirement for the  
BACHELOR OF TECHNOLOGY (Hons)  
(BUSINESS INFORMATION SYSTEMS)

Approved by,

---

(Dr. Rohiza Binti Ahmad)

Department of Computer Information Science  
Universiti Teknologi PETRONAS

UNIVERSITI TEKNOLOGI PETRONAS  
TRONOH, PERAK

May 2014

## CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

---

ARIF HAZMAN BIN SHAHIR  
BUSINESS INFORMATION SYSTEMS  
UNIVERSITI TEKNOLOGI PETRONAS

## **ABSTRACT**

Green computing or Green IT is the environmentally responsible use of computers and related resources as well as the reduction of power resource consumption. In supporting green computing and Green IT, this project is being carried out with the main purpose of creating a draft or template where it can calculate the power consumption of the computer or server when the database query is being process by implementing different type of query plan. The power consumption can be measured by using an energy saving power meter. A consolidated sample of management system will be developed as to indicate the context of the project in testing the query plan from a centralized database system. This project emphasizes on the amount of power usage (in watts) that the query plan consumed in assembling the data. It is more on focusing on the simulation of query plan in consuming power and not on developing a real-time system. Hence making the functionality and reliability of the system is not the main focus and will not be discussed in this paper work. By the end of this project, database developers would be able to see and analyze the various query plan applied in retrieving the same output under a specified condition that gives dissimilar power reading which indicates the different amount of power consumption by each query plan.

## **ACKNOWLEDGEMENT**

Peace be upon everyone,

First of all, I would like to express my deepest gratitude to Allah S.W.T, God The Mightiest for His blessing and guidance for me in completing my 8 months period of this Final Year Project I & Final Year Project II course. My honest appreciation goes to my supervisor, Dr. Rohiza Ahmad and the course coordinator, Pn. Mazlina Mehat (Final Year Project I) as well as Ms Penny Goh Kim Nee (Final Year Project II) for the effort of supervising me and guiding me towards being a responsible and excellent student. The experience and knowledge that I have gained when working with them is priceless compared to the effort I have shown. I would also like to express my deepest and heartfelt appreciation to Pn. Rozana Kasbon, for believing in me by giving me her project and giving me the chance and opportunity to proceed with this project. Thousands of appreciations are also extended to all lecturers of the Department of Computer Information Science for their full support, kindness, assistance, guidance and critics during these 8 months periods. They deserve my respect for their fair assessment and intensive assistant throughout this 8 months period. Last but no least, to anyone who has assisted me directly or indirectly in making my Final Year Project I & Final Year Project II course a success, thank you very much. May God repay your kindness.

Thank you.

## TABLE OF CONTENTS

<b>CERTIFICATION OF APPROVAL</b>	.	.	.	.	.	<b>2</b>
<b>CERTIFICATION OF ORIGINALITY</b>	.	.	.	.	.	<b>3</b>
<b>ABSTRACT</b>	.	.	.	.	.	<b>4</b>
<b>ACKNOWLEDGMENT</b>	.	.	.	.	.	<b>5</b>
<b>TABLE OF CONTENTS</b>	.	.	.	.	.	<b>6</b>
<b>LIST OF FIGURES, TABLES &amp; ABBREVIATIONS</b>	.	.	.	.	.	<b>7</b>
<b>CHAPTER 1: INTRODUCTION</b>	.	.	.	.	.	<b>9</b>
1.1 Background of Study	.	.	.	.	.	9
1.1.1 Significant and Relevancy	.	.	.	.	.	9
1.1.2 Timeframe	.	.	.	.	.	10
1.1.3 Feasibility	.	.	.	.	.	10
1.2 Problem Statement	.	.	.	.	.	11
1.3 Objectives	.	.	.	.	.	12
1.4 Scope of Study	.	.	.	.	.	13
<b>CHAPTER 2: LITERATURE REVIEW</b>	.	.	.	.	.	<b>14</b>
2.1 Green Computing	.	.	.	.	.	14
2.2 Energy Consumption	.	.	.	.	.	14
2.3 Energy Efficient Query Planning	.	.	.	.	.	15
2.4 Data Repository	.	.	.	.	.	16
<b>CHAPTER 3: METHODOLOGY</b>	.	.	.	.	.	<b>18</b>
3.1 Methods for Data Collection	.	.	.	.	.	18
3.2 Key Milestone	.	.	.	.	.	19
3.3 Gantt Chart	.	.	.	.	.	20
3.4 Project Activities	.	.	.	.	.	21
3.5 Data Capturing Procedure	.	.	.	.	.	22
3.6 Tools Utilization	.	.	.	.	.	23
<b>CHAPTER 4: RESULTS AND FINDINGS</b>	.	.	.	.	.	<b>25</b>
4.1 Class Diagram	.	.	.	.	.	25
4.2 Formula Used	.	.	.	.	.	26
4.3 Testing Setup and Environment	.	.	.	.	.	27
4.2 Query Strategy	.	.	.	.	.	28
4.3 Query Execution	.	.	.	.	.	29
<b>CHAPTER 5: CONCLUSION AND RECOMMENDATIONS</b>	.	.	.	.	.	<b>42</b>
5.1 Conclusion	.	.	.	.	.	42
5.2 Recommendation	.	.	.	.	.	43
<b>REFERENCES</b>	.	.	.	.	.	<b>44</b>
<b>APPENDICES</b>	.	.	.	.	.	<b>46</b>

## **LIST OF FIGURES**

- Figure 1.4.1: Centralized System Architecture
- Figure 1.4.2: System Architecture Used for This Project
- Figure 3.2.1: Final Year Project I
- Figure 3.2.2: Final Year Project II
- Figure 3.4.1: Waterfall Methodology
- Figure 3.5.1: Data Capturing Procedure Model
- Figure 4.1.1: Class Diagram for local Restaurant
- Figure 4.1.2: Database Created in Oracle
- Figure 4.3.1: Project Testing Set Up
- Figure 4.5.1: Power Consumption for Projection Query Strategy Condition 1
- Figure 4.5.2: Power Consumption for Projection Query Strategy Condition 2
- Figure 4.5.3: Power Consumption for Projection Query Strategy Condition 3
- Figure 4.5.4: Power Consumption for Projection Query Strategy Condition 4
- Figure 4.5.5: Power Consumption for Selection Query Strategy Condition 1
- Figure 4.5.6: Power Consumption for Selection Query Strategy Condition 2
- Figure 4.5.7: Power Consumption for Selection Query Strategy Condition 3
- Figure 4.5.8: Power Consumption for Join Query Strategy Condition 1
- Figure 4.5.9: Power Consumption for Set Query Strategy Condition 1

## **LIST OF TABLES**

- Table 3.3.1: Project Gantt Chart
- Table 3.5.1: Draft Template for Data Documenting
- Table 3.6.1: Project Tools Utilization
- Table 4.5.1: Projection Query Strategy Condition 1
- Table 4.5.2: Projection Query Strategy Condition 2
- Table 4.5.3: Projection Query Strategy Condition 3
- Table 4.5.4: Projection Query Strategy Condition 4
- Table 4.5.5: Selection Query Strategy Condition 1
- Table 4.5.6: Selection Query Strategy Condition 2
- Table 4.5.7: Selection Query Strategy Condition 3
- Table 4.5.8: Join Query Strategy Condition 1
- Table 4.5.9: Set Query Strategy Condition 1

## **ABBREVIATIONS AND NOMENCLATURES**

CIS	Computer and Information Science
CPU	Computer Processing Unit
FYP	Final Year Project
PC	Personal Computer
SQL	Structured Query Language
IT	Information Technology



# **CHAPTER 1**

## **INTRODUCTION**

The project was proposed in supporting green computing and Green IT with the main purpose of creating a draft or template where it can calculate the power consumption of the computer or server when the database query is being process by implementing different type of query plan. To be more specific, this project is basically concentrating on determining the type of query plan for efficient data executing activities in monitoring the power and energy in a single database system.

### **1.1 Background of Study**

Computer for instances, uses a high level of energy in processing many kinds of applications besides handling large volume of data repository, depending on the specification and features. Many of us are less aware that simple data retrieval from a database consumes energy and contribute to energy released to the environment. Thus in supporting green computing and Green IT, this project is being carried out to see on the power and energy consumed by different query plan when data is retrieved from the single database system.

#### **1.1.1 Significant and Relevancy**

In supporting to green computing, this project is conducted to study the amount of power used by different query plan in query processing activity from a central database system besides creating the awareness and educating the database developers and users in general. By the end of this project, a draft or template will be created where it can calculate the power consumption of the computer or the server when the database query is being process by implementing different type of query which can be used by the database administrator and developers to manage strategies whether to use which type of query plan in executing activities to minimize the consumption of power and energy of the computer. The energy consumed therefore can be

reduced because this tool relatively helps in calculating the computer activity energy. In conclusion, the electricity cost can be reduced and minimized hence preserving the green environment.

### **1.1.2 Timeframe**

This whole project is said to be estimated to complete in approximately 8 months. The first 4 months are used to find the related reading materials for literature review phase. The task in to collect and review the research paper and paper works that have been done by previous experts to help in making the project more focused and transparent. The findings is then be presented and submitted to the supervisor in a formal report. The designing process will carried out in the early 5<sup>th</sup> month which including designing the user interface and develop the data repository. This development phase shall be completed by the end of the 6<sup>th</sup> month. Lastly, the 7<sup>th</sup> and 8<sup>th</sup> month will be used to test the query processing and executing activities and calculating the energy used while the processes take place. Throughout the whole project development and testing phase, meeting with the supervisor is conducted on regular basis to seek for advice on the query plan and system programming to ensure accurate results. At the same time, each progress of work is presented occasionally following to the CIS FYP schedule.

### **1.1.3 Feasibility**

The system developed is referring to a sample centralized management system where it acts as the main branch database and could access to various data from other branches and locations. But the query plan testing is only focused on the local host basis and does not involve any network connection. This project is a comparative research based project that merely emphasizes more on the power consumption in query processing. Therefore the functionality and reliability of the system is not the main focus. Tools used in this project are divided into two, hardware and software. For hardware, the equipment includes personal computer (laptop) and a plug meter to capture the energy and power released. Besides that, for software tools, the Oracle

Database Application Express 10g which is to develop the database and test the data query plan. The features and specification for each tools is discussed further in Chapter 3: Methodology under Tools Utilization section (Table 3.6.1).

For the project feasibility within scope and timeframe, the project is to be delivered with its functionality by the 27<sup>th</sup> August 2014. The scope of the project is aimed to be completed as the deadline. More details are provided below in Project Schedule and Gantt chart.

## **1.2 Problem Statement**

In most organizations these days, the IT department has been one of the most important departments in maintaining the operational activity of the company. By installing huge data repositories, these companies are actually dealing with a large amount of data on daily basis which is also engaging with high electricity cost. Based on previous research, the amount of electricity used for data repositories only are relatively high and may increase if no preventive measure is taken. This electricity usage gives out a huge amount of energy to the environment which contributed to the high electricity cost and can cause the increase of temperature of global warming. This is where green computing comes into this project. Green IT stands for Green Information Technology. This Green IT or green computing is composed of two things:

1. Minimizing the negative impact of information technology use on the environment.
2. Using information technology to help solve environmental issues.

Green computing has now been increasingly attentive in many organizations in dealing with computer activity and preserving the environment while sustaining an efficient business process.

Therefore, this project is supporting the Green IT first objective which is to minimizing the negative impact (such as global warming) of information technology use on the environment. In order to have a better quality control of the environment

in context of conserving energy from computer, one especially the database administrator or database developers should know the amount of energy consumed by a computer and application activities that the computer processes.

Although the world is promoting to have green computing, most users are not observant to the personal usage of their own computers and system activities. Hence this project is conducted to specifically to create a draft or template where it can calculate the power consumption of the computer or server when the database query is being process by implementing different type of query plan. This draft can be very useful for database administrator or developers to monitor and minimize the amount of energy used in every activity of data capturing from the data storage. Besides observing on the CPU energy consumption as a whole, it also helps to educate the users on how a simple energy conserved from data retrieval process can contribute in preserving the energy released. In addition, this study will further analyses the different query plans to be tested whether it has links to the energy produced upon retrieving the data from the data storage.

Based on previous researches, the author found that there is no research which does the comparison in terms of the power consumed by different query plans for solving similar query task. On top of that, there is also no comparative study on power consumption done in terms of similar query planning but different data sizes. Furthermore, the author also found that there is no draft or template available to do the comparative study. All of these problems are making this project significant and reliable to solve these problems.

### **1.3 Objectives**

The main objectives of this project are:

1. To identify the basic query types which make up query plan for relational database.
2. To develop a draft outline that can be used to make comparative study on power consumption of different query plans.
3. To conduct comparative study of some query plans using the draft developed.

## 1.4 Scope of Study

Basically, this project is concentrating on determining the best type of query plan for efficient data executing activities in monitoring the power and energy in a centralized database system. The system of the project is based on a client server and centralized database management model as shown below. But the query plan testing is only focused on the local host basis and does not involve any network connection. Therefore, in this project, only Site A and the Database will be used for testing. Testing with computer network might as well be done but requires more time. Thus in the future, the author seeks to continue this research.

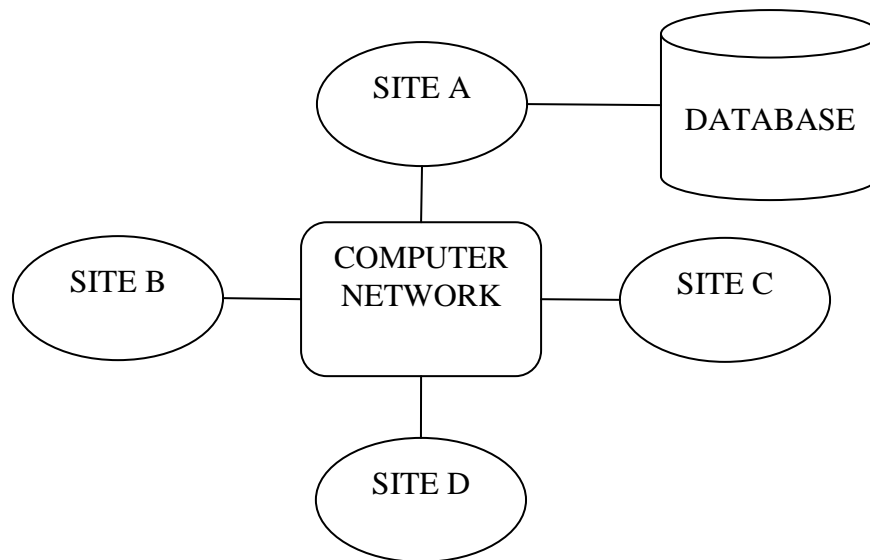


Figure 1.4.1: Centralized System Architecture

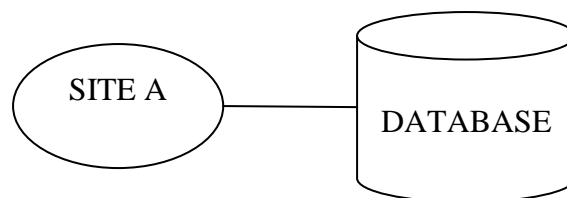


Figure 1.4.2: System Architecture Used for This Project

## **CHAPTER 2**

### **LITERATURE REVIEW**

#### **2.1 Green Computing**

“Green computing and green technology refers to the environmentally responsible use of computers and any other technology related resources. Green computing includes the implementation of best practices, such as energy efficiency central processing units or the CPUs, peripherals and servers. In addition green technology aims to reduce resource consumption and improve the disposal of electronic waste or e-waste (“Green Computing”; 2013)”.

“Size of the Data Centers, with their high energy costs and increasingly negative impact on the environment, are the driving force behind the green computing movement. (Harmon R., Auseklis N.;2009)”.

As said by Patrick Kurp (2008), as the commitment to reduce environmental impact and power consumption are becoming increasingly important objectives for organizations, as architecture leaders are now proactively considering environmental resource constraints along with more traditional IT business goals.

Based on the three quotes above, it is said that supporting Green Computing or Green IT would really helps in improving the negative environmental effect such as global warming. The electronic waste are also said to be increasing and by supporting the Green Computing surely can contribute more or less on this matter.

#### **2.2 Energy Consumption**

According to one study, today, a 1U server costs roughly \$4,000 and typically consumes 500 watts while the 3-year energy cost for operating this 1U server that includes both electricity consumed by the server as well as energy needed to remove heat dissipated from the server would be \$3,942 in the U.S. based on \$0.08 per

kilowatt-hour and \$7,884 in Japan. The total cost if the cost for data center infrastructure is included will be even higher. (Essary D., Amer A.;2008)”.

According to David Wang (2007) a research from ARI (Aperture Research Institute, a division of software company Aperture Technologies) found that 38% of data centers are already running at dangerous levels of power consumption, meaning they are at risk of failing due to a power shortfall. According to Gartner 50% of data centers will have insufficient power or cooling capacity to meet demand by 2008, thus seriously limiting business growth. Bill Clifford, CEO of Aperture, said: "Many users are simply unaware of the dangers they are introducing to their data centers. There is an impending crisis".

Most of big organizations in this world, such as PETRONAS, surely will have big data repositories. These data repositories will keep growing and getting bigger as respectively to the company that is also growing. Therefore, their power consumption is also getting higher and higher each and every year just like what Essary D. and Amer A. said in the above quote. My project is just the draft or template for the small tool of utility to check the power consumption and compare between the different query plans. My project, in the future with further study and research, might as well help these big companies for them to monitor their power consumption.

### **2.3 Energy Efficient Query Planning**

“Achieving higher energy efficiency can be seen from two different vantage points. We can either reduce the time consumed for executing a query which is what exactly DBMS vendors are doing by increasing the performance of their systems or we can reduce the energy consumed while the process is running (Flach T.;2010).”

As said by Tobias Flach (2010) also, if the CPU show a higher power consumption than the HDD, a memory-bound query plan which needs a large time fraction for disk operations and a small time fraction for processor operations will provide a better energy efficiency than a CPU-bound query plan which executes in the same

time and needs a small time fraction for disk operations and a large time fraction for processor operations.

According to Roland Alonso and Solas Ganguly (1992), techniques such as hardware power-model control and query rescheduling have been proposed to save energy in the systems.

The above quote shows that for large database or data repository, having a better query optimization would certainly reduce the power consumption thus helps in energy saving. Most of the database developers are currently saving the energy consume by optimizing the query. In this project, the differences between choosing different query plan are shown and can be used as the reference for the database developers to monitor their power consumption of the CPU.

## **2.4 Data Repository**

“Power consumption at server level is increasing along with performance increase. At the same time, business needs have grown even faster. Higher performance CPU’s along are insufficient to meet ever growing and new business needs; most business demands high performance CPU’s and more of them that are packaged into servers. For example, Google has nearly half a million servers considering Google is barely a few years old. (Wang D.;2007)”.

As mentioned by David Wang (2007), energy consumption can be reduced in smaller area such as having data repository as data storage can be considered another area where performance improvement is achieved with less power consumption.

“Predictive data grouping has advantages in its applicability compared to both prior approaches to reducing latencies and to reducing energy usage (Essary D., Amer A.; 2008)”.

It is said that the power consumption is somehow related to the size of the data repository. For example, Google, as said in the quote above, have half a million of servers. The action to package and divide their server into group is to avoid having high energy consumption at one big server. The tendency of a really big server to



perform perfectly is affected by the CPU performance. If the CPU can not stand very high energy consumption, then the tendency for the server to get damage is really big. Therefore the action to have many servers and package them took by Google is actually because Google is afraid of damaging their server due to the very high energy consumption.

## **CHAPTER 3**

### **METHODOLOGY**

The research methodology chapter represents the strategies and approaches consist of collecting and analyzing the data collected in order for meaningful comparative analysis and interpretations of the research finding to be present. This section focuses on giving the insights on how the research will be carried out. This includes the mode of data collection, how the data is analyzed and the research tool design.

#### **3.1 Methods for Data Collection**

##### **3.1.1 Observation**

This method is used in monitoring the power and electricity reading on the plug meter and the tool while the data execution process takes place. Though the method may seems easy, a careful observation needs to be done to get as close and as accurate result.

##### **3.1.2 Random Sampling Method**

Several sets of random data queries are selected to test on the power and energy consumption. The queries may consist of different amount of attributes or records from several tables and applying various strategies approach. This method is used to test the system and the reliability of the power plug meter.

##### **3.1.3 Systematic Random Sampling Method**

This method is used to test the power and energy consumed of the similar set of several random data queries under similar condition in a specified period of time. The data queries are tested by using different query plan. It is to see which query uses the least energy given time period.

### 3.1.4 Judgment Sampling Method

For this method, it involves tacit and know-how knowledge from experts in a particular domain based on his or her experience. In this project, any relevant and useful information is captured during the formal and informal regular meeting with the supervisor. All the information is helpful in assisting to provide data based on the expert's personal opinion.

## 3.2 Key Milestone

### Final Year Project I

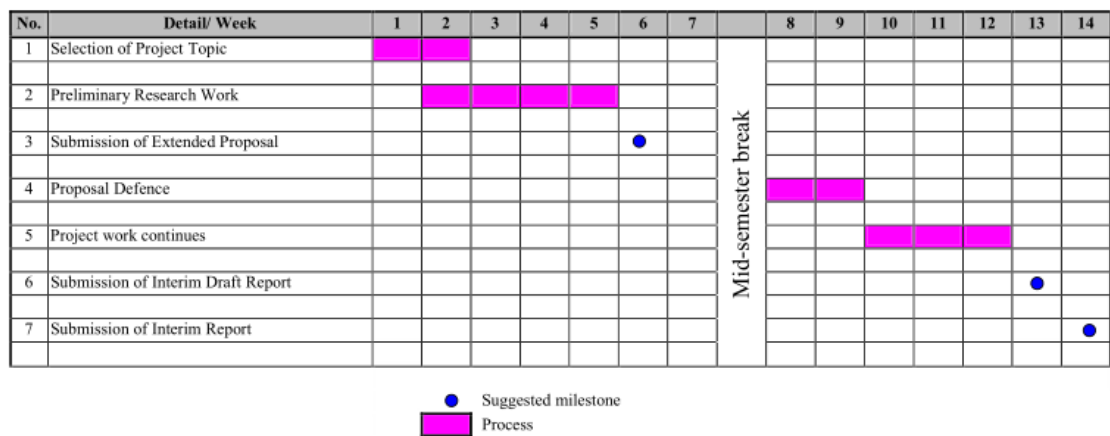


Figure 3.2.1: Final Year Project I

### Final Year Project II

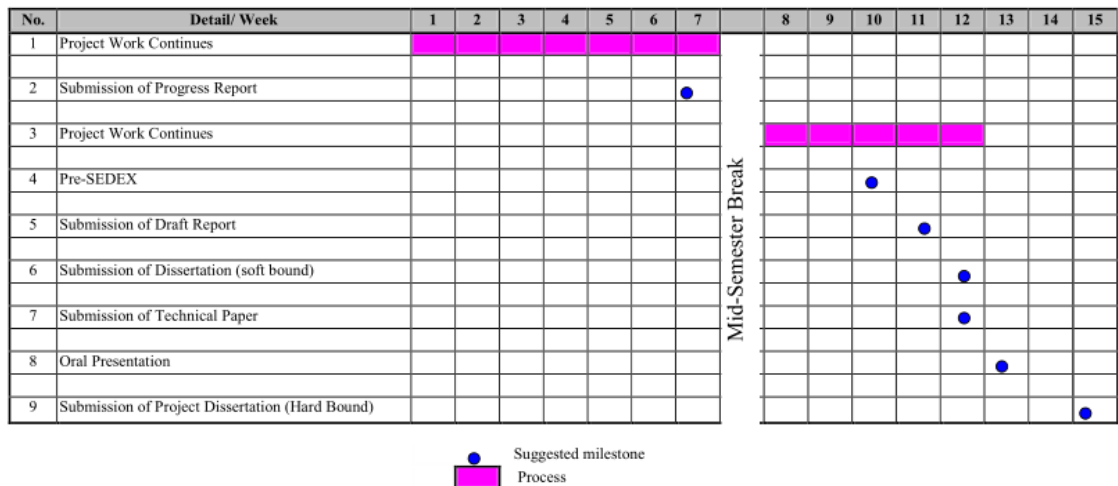


Figure 3.2.2: Final Year Project II

The above shows the key milestone of the Final Year Project as a whole which is focusing on the required deliverables and presentation that final year students need to follow as a guidelines as part of the assessment in completing the course.

### 3.3 Gantt Chart

Task/Month	1	2	3	4	5	6	7	8
<b>1. Define requirements/analysis</b>								
1.1 Determine scope, time and expertise available								
1.2 Collection of reading materials (literature review)								
1.3 Conduct meeting with supervisor								
<b>2. Design</b>								
2.1 Design the system architecture								
2.2 Design the prototype								
2.3 Design the data capturing procedure								
<b>3. Develop/Implementation</b>								
3.1 Develop the prototype (Draft or Template to do the Comparative Study)								
3.2 Linking the prototype with the database and plug meter								
<b>4. Testing/Final</b>								
4.1 Test the (Draft or Template to do the Comparative Study) with query plan								
4.2 Analyze the results								
4.3 Finalize the prototype and analysis								

Table 3.3.1: Project Gantt Chart

The table above shows the Gantt chart of this project from the beginning until it is final stage which will take approximately 8 months or 28 weeks to be accomplished. The activity done in each phase adhered to the project methodology in a whole which is further discussed in the Project Activities section.

### 3.4 Project Activities

For project activities as a whole, the waterfall methodology model is chosen due to its sequential process and its fitness to do the project. This model consists of several stages which every stages has numerous activities and it is distinctively important towards the project progress.

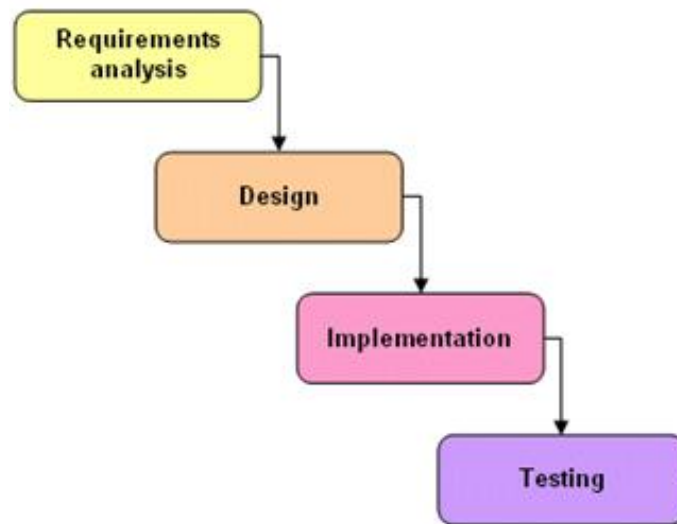


Figure 3.4.1: Waterfall Methodology

The project started by planning and understanding the requirements needed. Before the project can begin, the timeframe in completing the project is wisely planned to give guidelines in meeting the datelines for each task. In synthesizing, the requirements of this project, all initial components such as problem statement, objectives and background of study are analyzed under this beginning phase. This also includes collecting all the research papers related to the topic in exploring the study that has been done by previous experts.

The design phase is started by designing the system architecture of the whole project which gives a clearer view on how the project will run. It also involves in designing the sample database and the drafting or sketching the prototype based on the collective requirements. The data capturing procedure is also designed to describe the procedural activity involved in collecting the result.

The next stage which is the implementation is about making what has been planned in the earlier phase into a real working system. It involves a lot of coding activity whereby the database and the draft or template of power consumption will be developed. Both the database and the prototype will then need to be connected with each other and with the plug meter.

The last phase which is the testing phase is the most important stage for the whole project. Besides testing the functionality and reliability of the whole project itself, this stage is focusing on testing the prototype to retrieve and compare between the power consumption of the using different query plan. This activity is done repetitively to get as accurate result as possible in order to detect any malfunction or improvement that can be made to the prototype.

### 3.5 Data Capturing Procedure

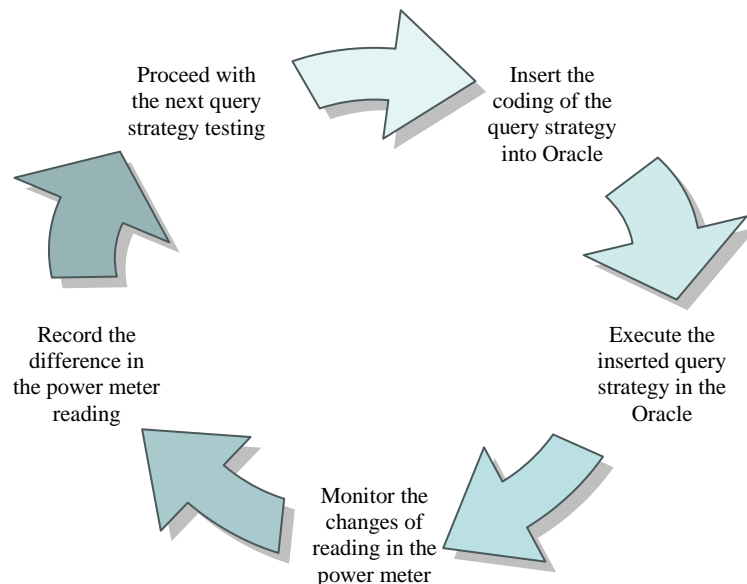


Figure 3.5.1: Data Capturing Procedure Model

The above figure describes the whole procedure of the data capturing method used in this project. This data capturing can only be done after the sample database with the data are generated. The data capturing will be focusing on the power meter

reading differences between each query strategy. However, each query strategy in one similar condition should have the same output of result.

The changes in the power meter reading will be carefully monitored and recorded. The data collected will then be documented in a table form. The table below shows the draft template of the table used to document the data captured.

Query	Strategy	Table Involved	Description	Efficiency Measurement (W)		
				Power Reading (W)	Total Reading (W)	Average (W)

Table 3.5.1: Draft or Template for Data Documenting

### 3.6 Tools Utilization

The below table shows the project tools features and specifications.

Category	Client	Remarks
Hardware	Dell XPS 15 (Laptop) 1) Intel(R) Core(TM) i7 CPU Q740 @ 1.73GHz 2) 8GB DDR3 SD RAM 3) 750 GB Hard Drive Storage 4) 64-bit Windows 7 Professional 5) 1GB NVIDIA GeForce GT435M	Personal computer with relative specification and features. Most of the project phase especially during the implement and testing phase that are needed to be done with this hardware.
	Energinie Power Saving Meter: 1) Voltage Rating: 240V ac 50Hz 2) Current Rating: 13A 3) Maximum Load: 3120W 4) Power Consumption: <0.5W 5) Metering: Voltage 200 to 276 V AC, +/- 1.5% 6) Frequency Range: 45 – 65 Hz	This hardware is the device that will be used to calculate the electric current or power and energy consumption used during the query data processing. This device will be linked to the prototype in the implement phase.

Software	Oracle Application Express 10g	This software is the software that will be used to develop the sample database management system that this project will be using
----------	--------------------------------	--

Table 3.6.1: Project Tools Utilization



## CHAPTER 4

### RESULTS AND FINDINGS

#### 4.1 Class Diagram

Figure below shows the class diagram of the centralized database system that is used by a local restaurant. This database system will be used for this project as a sample database system. There are a total of nine tables of entity with defined attributes involved in the database including three bridge entities which are DepartmentSupplier and OrderMenu. The relationship and cardinality were shown in the figure below.

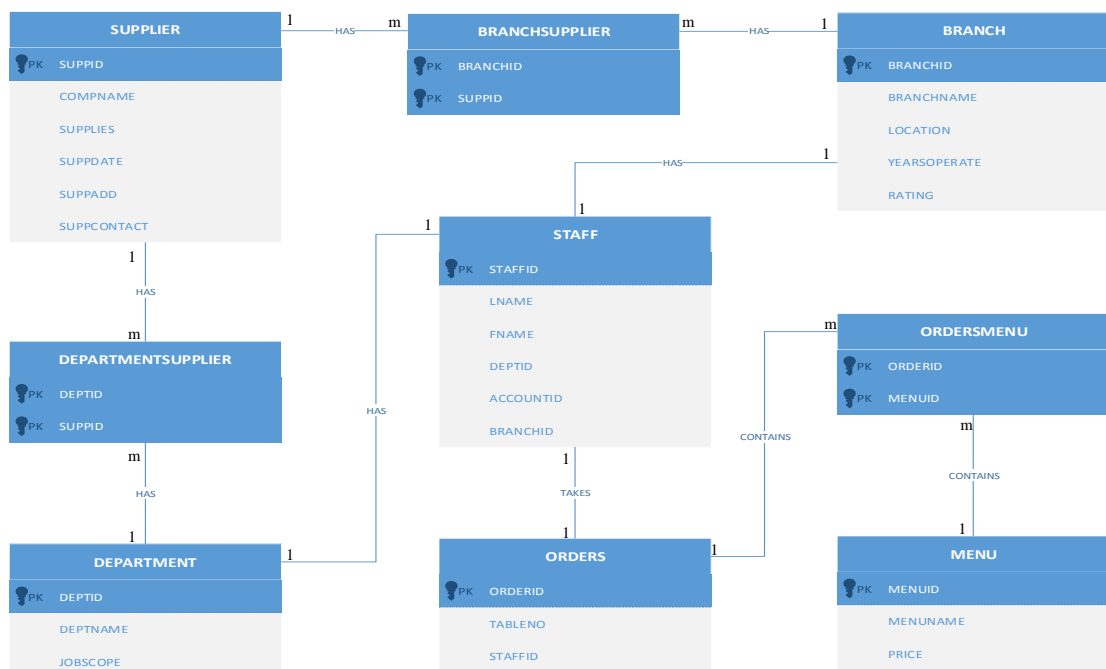


Figure 4.1.1: Class Diagram for local Restaurant

Based on the class diagram, a database is developed by using the Oracle Database Application Express 10g.

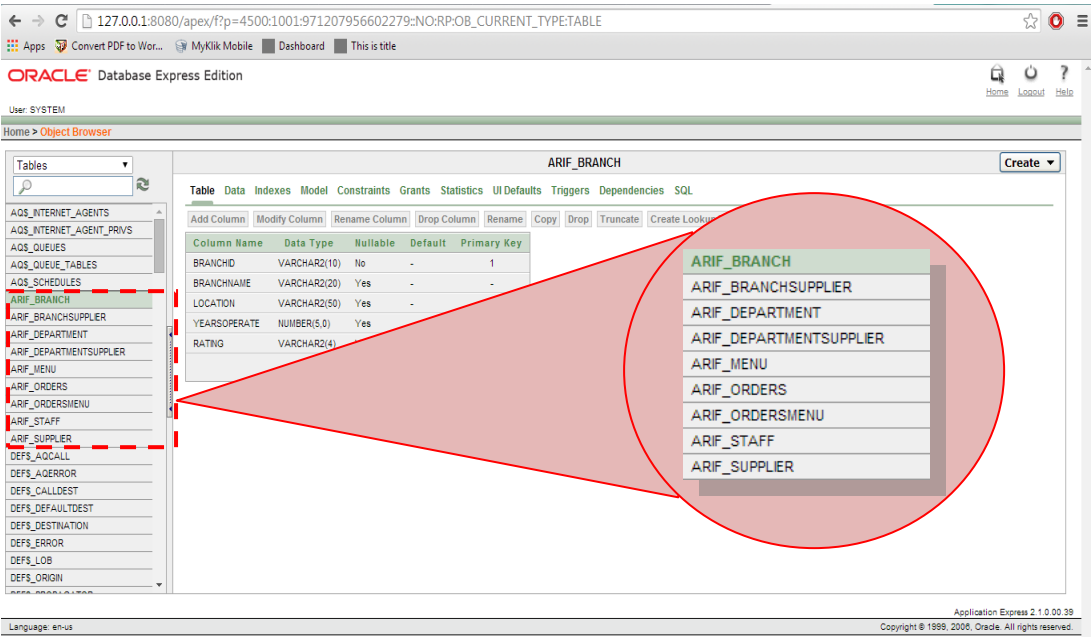


Figure 4.1.2: Database Created in Oracle

The figure above shows the interface of the created database in the Oracle for the project. As generating a huge amount of data which is need in this project is too time-consuming, the data in the database will be generated automatically by a website named generatedata.com. This website allow the author to optimize the time to complete the project as it provides an enormous amount of random data based on the selected criteria such as name, address, state and so on. It also generates SQL statement which helps in monitoring the time taken at the data entry level. Each table has different number of records depending on reasonableness of the attributes in the tables.

## 4.2 Formula Used

Of course through several researches, there are a lot of ways in calculating the energy consumed by the query processing. However, in this project, the testing process will be conducted five times and all the result will be recorded. Then the average of the result will then be calculated as the final result of the power consumption of the query plan. The formula can be derived as follows.

$$\text{Power Consumption, } W = \frac{\sum n}{k=5} ; \text{ where } \begin{array}{l} n = \text{number of power reading} \\ k = \text{loops of testing process} \end{array}$$

### 4.3 Testing Set Up and Environment

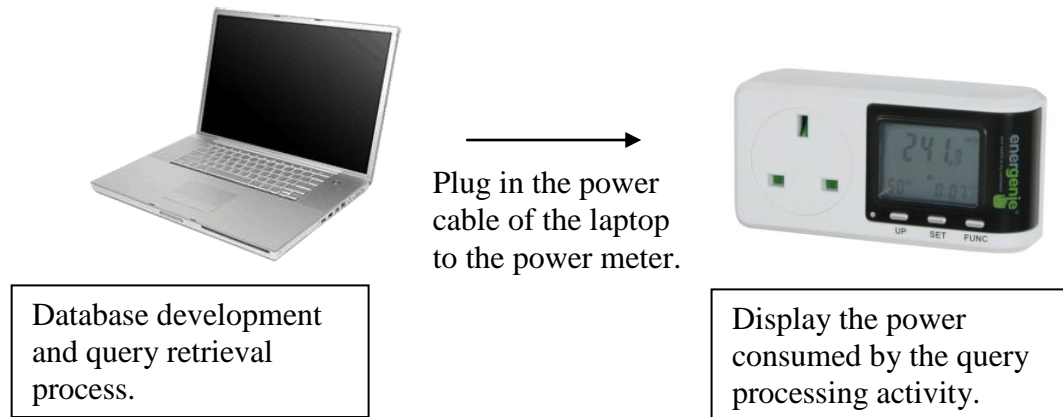


Figure 4.3.1: Project Testing Set Up

Figure above shows the set-up of how the project will look literally. The laptop is used to develop the central database of the local restaurant management system by using Oracle Application Express 11g. Besides that, the retrieval process also takes place in the laptop. The initial value, which is the power consumed by the laptop at the current state before the data is retrieved from the database, is setup on the energy power reading before the query is being executed. Once the query is being executed, any immediate changes on the power consumption will be captured by the power meter and recorded.

There are also certain conditions be made in getting the favorable result on the power consumption. All of the conditions were made to minimize the error during the process of taking the power consumed by the query and to make sure the power meter is being held carefully to get more accurate result. The site of the project would be mainly takes place in the author hostel room which is in Universiti Teknologi Petronas and the result gathered will be based on the specification of tools and equipments used in this project (refer to 3.5 Tools Utilization). Other than that, no other application of software is running on the system background along with the Oracle to enhance the accuracy of the result. Lastly, network connection will be

unplugged during the whole procedure of the project as the project is only comprises single database system and it is local host basis.

#### **4.4 Query Strategy**

A query plan or query strategy can be defined as the ordered set of steps that will be used to access the data in a SQL relational database management system. Therefore, since SQL is declarative, there are many kinds of alternative strategy to execute a certain query which however executed the same output or data. The query strategy will be divided into four which are the:

i. **Projection Query**

For this strategy, the author basically wants to compare between each query and to identify whether the position of column is affecting the power consumed or not.

ii. **Selection Query**

Conditions are commonly used in the query strategy to specify the output. For this strategy, the author wants to compare between conditions query such as OR, LIKE and NOT. The author will try different style of query to determine whether it will affect the power consumption or not.

iii. **Join Query**

Generally joining tables is the most common way to perform a query. In this strategy, the author wants to compare between each different types of join such as EQUI JOIN and NATURAL JOIN.

iv. **Set Query**

Combining the results sets can be done by using certain SQL statement such as UNION and INTERSECT. The author uses this query strategy to compare the power consumed by each different set query.

## 4.5 Query Execution

Based on the query discussed before, a number of conditions are specified with different strategies are being applied to test on the query power consumption. Below are the conditions that will be used for this project.

### Projection Query

#### Condition 1: To project single column from different position

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_SUPPLIER	To select the FIRST column of the table	<u>Q1</u> SELECT SUPPID FROM ARIF_SUPPLIER;	5	20.9	4.18
			4		
			4.9		
			2.8		
			4.2		
	To select the THIRD column of the table	<u>Q2</u> SELECT SUPPLIES FROM ARIF_SUPPLIER;	5.6	25.7	5.14
			5.2		
			5.4		
			5.2		
			4.3		
	To select the LAST column of the table	<u>Q3</u> SELECT SUPPCONTACT FROM ARIF_SUPPLIER;	3.7	18.4	3.68
			3.2		
			4.9		
			3.1		
			3.5		

Table 4.5.1: Projection Query Strategy Condition 1

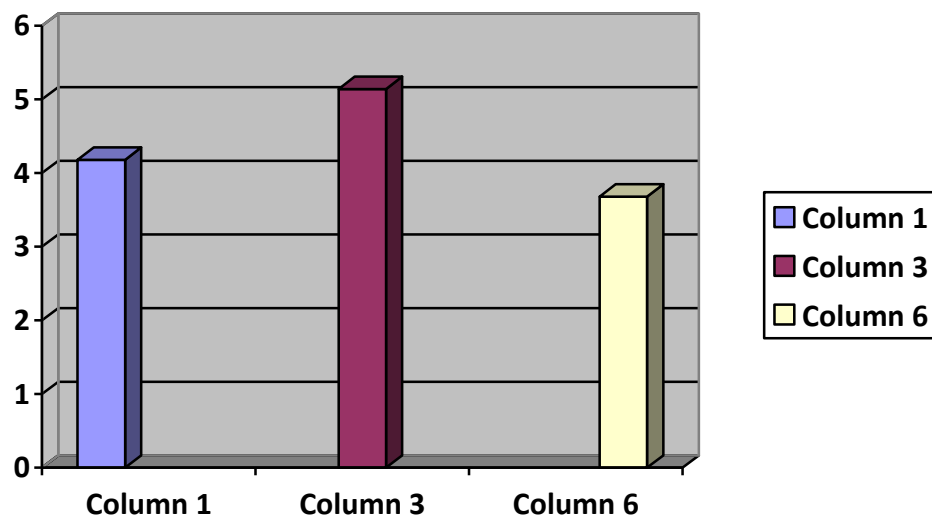


Figure 4.5.1: Power Consumption for Projection Query Strategy Condition 1

**Condition 2: To project multiple columns from different or adjacent position.**

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_SUPPLIER	To select the FIRST and SECOND column of the table	<u>Q1</u> SELECT SUPPID, COMPNAME FROM ARIF_SUPPLIER;	4.2	16.9	3.38
			3.1		
			3.8		
			3.1		
			2.7		
	To select the FIRST and LAST column of the table	<u>Q2</u> SELECT SUPPID, SUPPCONTACT FROM ARIF_SUPPLIER;	3.9	16.6	3.32
			3.3		
			2.9		
			2.5		
			4.0		

Table 4.5.2: Projection Query Strategy Condition 2

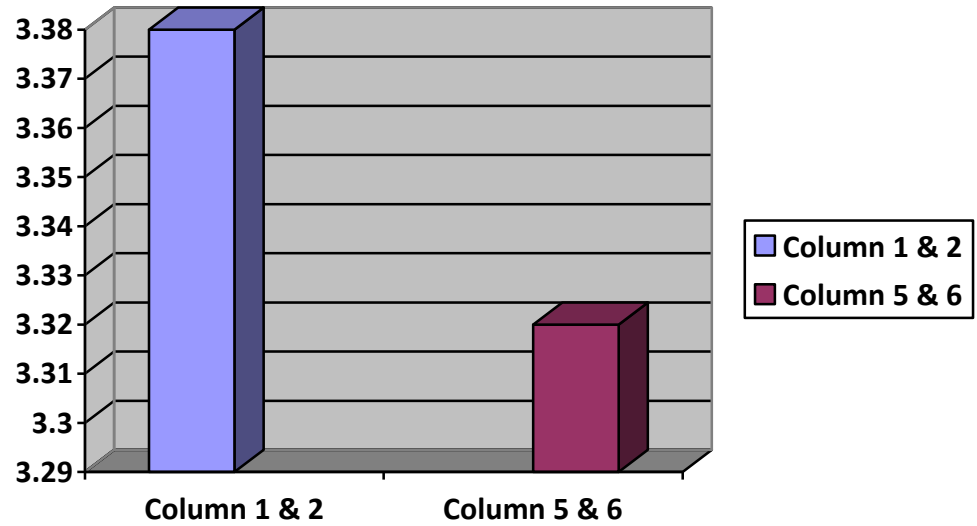


Figure 4.5.2: Power Consumption for Projection Query Strategy Condition 2

**Condition 3: To project all columns using different strategy.**

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_BRANCH	To select all columns by using *	<u>Q1</u> SELECT * FROM ARIF_BRANCH;	4.6	16.3	3.26
			2.2		
			3.9		
			2.5		
			3.1		
	To select all columns by listing one by one	<u>Q2</u> SELECT BRANCHID, BRANCHNAME, LOCATION, YEARSOPERATE, RATIONG FROM ARIF_BRANCH;	3	14.2	2.84
			3.2		
			1.8		
			3		
			3.2		

Table 4.5.3: Projection Query Strategy Condition 3

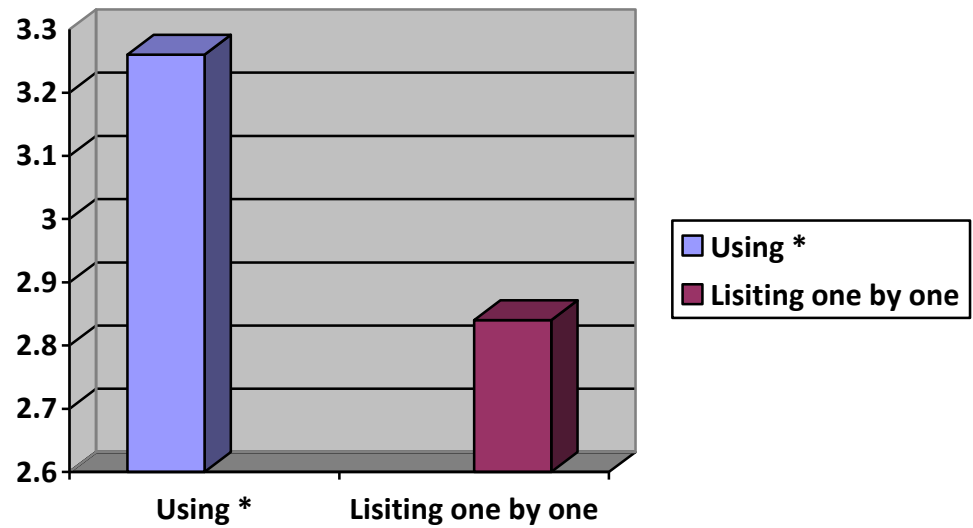


Figure 4.5.3: Power Consumption for Projection Query Strategy Condition 3

**Condition 4: To project first column from different table.**

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_MENU	To select the FIRST column of the table	<u>Q1</u> SELECT MENUID FROM ARIF_MENU;	4.3	18.3	3.66
			3.3		
			4		
			3		
			3.7		
ARIF_DEPARTMENT	To select the FIRST column of the table	<u>Q2</u> SELECT DEPTID, FROM ARIF_DEPARTMENT;	3.9	20.5	4.1
			3.2		
			3.7		
			4.9		
			4.8		

Table 4.5.4: Projection Query Strategy Condition 4



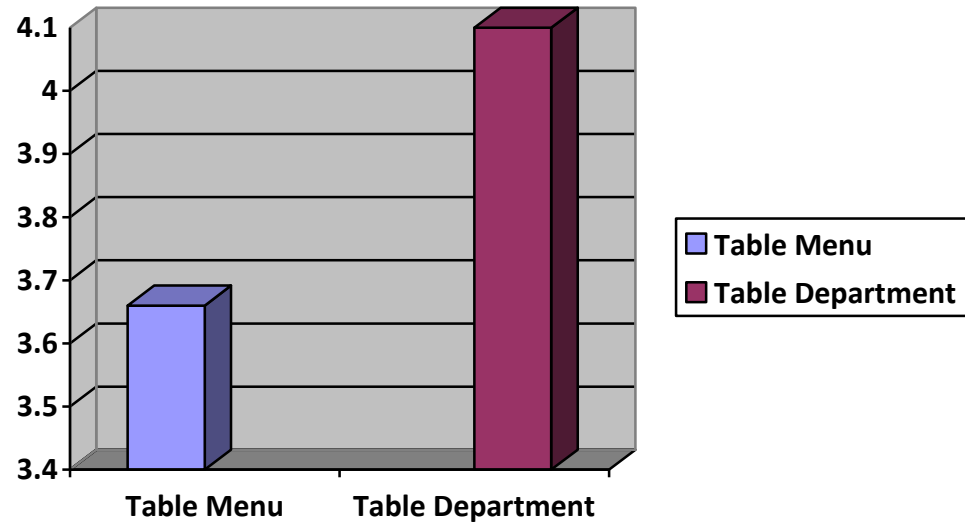


Figure 4.5.4: Power Consumption for Projection Query Strategy Condition 4

### **Discussion for Projection Query**

Figure 4.5.1 shows that column 1 and 6 give the lowest power consumption compared to column 3. This could prove that the screening process of the database starts from the first or last columns. Column 1 and column 6 have almost the same power consumption therefore making it hard to identify if the screening process of the database starts from the first or last column. However the amount of data in that particular column might also have affected the power consumption.

Figure 4.5.2 shows that power consumption for selecting column 1 & 2 and selecting column 5 & 6 are almost the same with 0.06W in difference. Therefore in this case, it is hard to conclude the testing. However, having column 5 & 6 the least power consumed could prove that the screening process of the database might start from the last column.

Figure 4.5.3 shows that selecting the entire columns in that particular table Branch using \* uses more energy compared to listing all of them one by one. These two queries give the same output of results however consumed different power while it is being executed. It can somehow be concluded that listing all the columns one by one could save more energy.

Figure 4.5.4 shows that selecting the first column from Table Department uses more energy compared to selecting the first column from Table Menu. Table Department

has 12 rows of data while table Menu consists of 20 rows of data. This power consumed however maybe influenced by the data type as the data type for the first row of table Menu is integer while for table Department is varchar.

### Selection Query

#### **Condition 1: To select using OR statement**

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_STAFF	To select using OR	<b>Q1</b> SELECT FNAME, DEPTID FROM ARIF_STAFF WHERE FNAME = 'Clark' OR FName = 'Rooney';	3.4	14.4	2.88
			3		
			3.3		
			3		
			1.7		
	To select using IN	<b>Q2</b> SELECT FNAME, DEPTID FROM ARIF_STAFF WHERE FNAME IN ('Clark', 'Rooney');	3.1	14.6	2.92
			2.1		
			3.7		
			2.8		
			2.9		

Table 4.5.5: Selection Query Strategy Condition 1

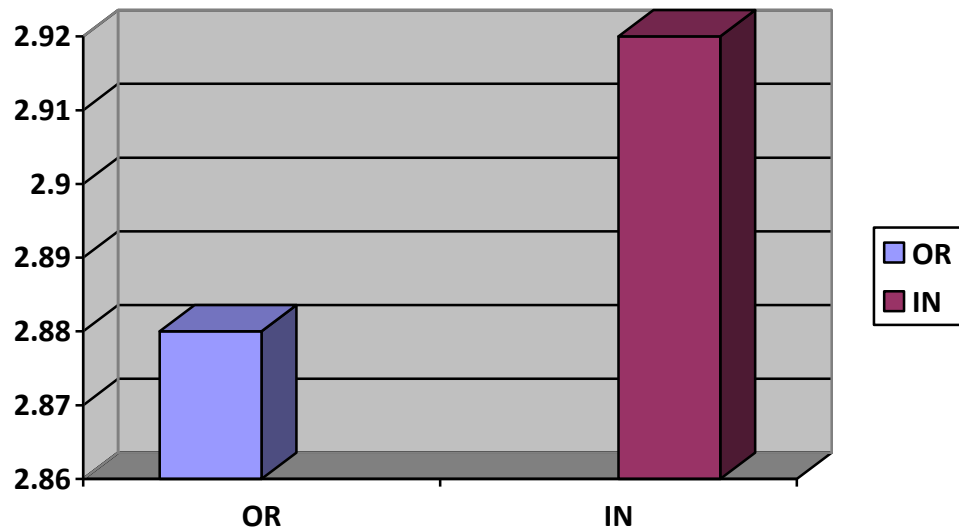


Figure 4.5.5: Power Consumption for Selection Query Strategy Condition 1

**Condition 2: To select using LIKE statement**

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_STAFF	To select using LIKE	<u>Q1</u> SELECT FNAME, DEPTID FROM ARIF_STAFF WHERE FNAME LIKE 'Rooney';	4.6	13.5	2.7
			1.5		
			3.4		
			1.5		
			2.5		
	To select using =	<u>Q2</u> SELECT FNAME, DEPTID FROM ARIF_STAFF WHERE FNAME = 'Rooney';	2.5	12.5	2.5
			2		
			2.5		
			3.4		
			2.1		
	To select using IN	<u>Q3</u> SELECT FNAME, DEPTID FROM ARIF_STAFF WHERE FNAME IN ('Rooney');	2.5	18.8	3.76
			4.4		
			5.9		
			2.2		
			3.8		

Table 4.5.6: Selection Query Strategy Condition 2

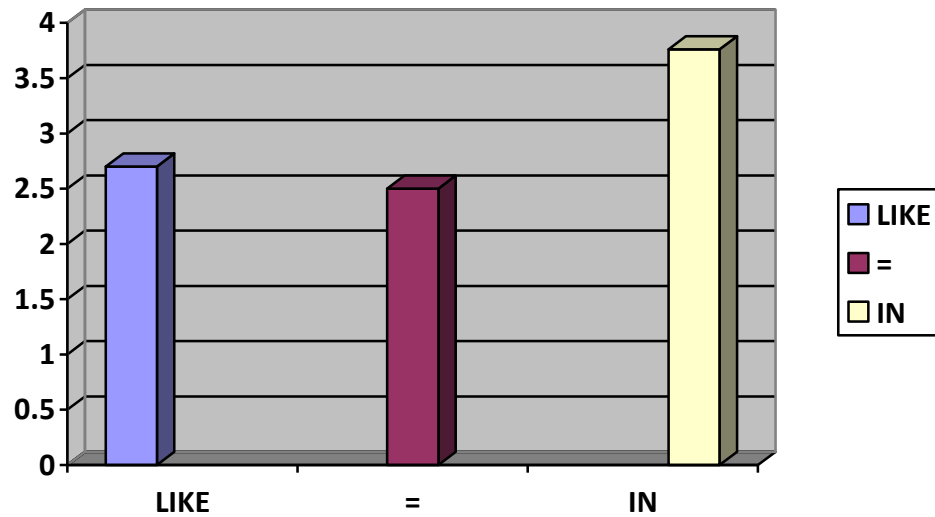


Figure 4.5.6: Power Consumption for Selection Query Strategy Condition 2

**Condition 3: To select using NOT statement**

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_MENU	To select using NOT LIKE	<u>Q1</u> SELECT MENUNAME FROM ARIF_MENU WHERE PRICE NOT LIKE 10;	4.9	22.9	4.58
			4.9		
			5		
			4.9		
			3.2		
	To select using !=	<u>Q2</u> SELECT MENUNAME FROM ARIF_MENU WHERE PRICE != 10;	3.8	15.9	3.18
			2.4		
			3.6		
			2.2		
			3.9		

Table 4.5.7: Selection Query Strategy Condition 3

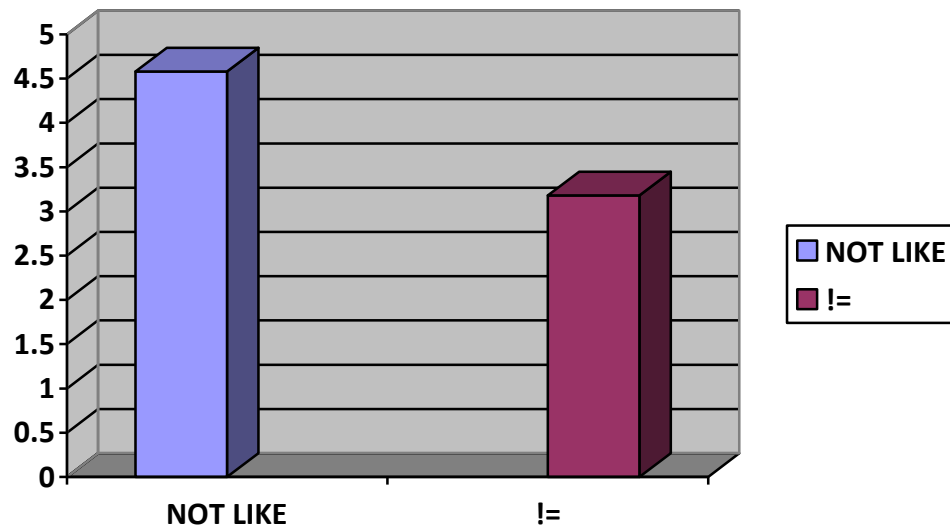


Figure 4.5.7: Power Consumption for Selection Query Strategy Condition 3

### **Discussion for Selection Query**

Figure 4.5.5 shows OR statement used in the query could save energy compared to using IN statement in the query strategy. As both queries is giving out the same output, the amount of data and data type in this condition will not be affecting the power consumed.

Figure 4.5.6 shows using algorithm save more energy compared to other statement such as LIKE and IN. This figure again shows that IN statement uses the most energy, as in figure 4.5.5 also shows that IN statement uses the most energy. This means that IN statement somehow make the database system to use more energy during execution. The algorithm = uses the least energy while LIKE statement having the difference of 0.2W.

Figure 4.5.7 proof that using algorithm will certainly save more energy. As from the result, using algorithm != saves more energy compared to using NOT LIKE statement. This support the results in Figure 4.5.6 which also shows that using algorithm save more energy by having it using the least amount of energy compared to other statement.

## Join Query

### Condition 1: To join 2 tables

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_STAFF & ARIF_ORDERS	To join 2 tables by ON	<u>Q1</u> SELECT S.STAFFID,S.FNAME,O.ORDERID FROM ARIF_STAFF S JOIN ARIF_ORDERS O ON S.STAFFID = O.STAFFID ORDER BY O.ORDERID;	3.6	16.8	3.36
			3		
			3.3		
			3.7		
			3.2		
	To join 2 tables by EQUI JOIN	<u>Q2</u> SELECT S.STAFFID,S.FNAME,O.ORDERID FROM ARIF_STAFF S, ARIF_ORDERS O WHERE S.STAFFID = O.STAFFID ORDER BY O.ORDERID;	4.4	23.6	4.72
			5.9		
			4		
			4.5		
			4.8		
	To join 2 tables by USING	<u>Q3</u> SELECT STAFFID,FNAME,ORDERID FROM ARIF_STAFF JOIN ARIF_ORDERS USING (STAFFID) ORDER BY ORDERID;	2	14.1	2.82
			3.9		
			2.6		
			2.6		
			3		
	To join 2 tables by NATURAL JOIN	<u>Q4</u> SELECT STAFFID,FNAME,ORDERID FROM ARIF_STAFF NATURAL JOIN ARIF_ORDERS ORDER BY ORDERID;	3.5	17.8	3.56
			2.2		
			4.4		
			4.4		
			3.3		

Table 4.5.8: Join Query Strategy Condition 1

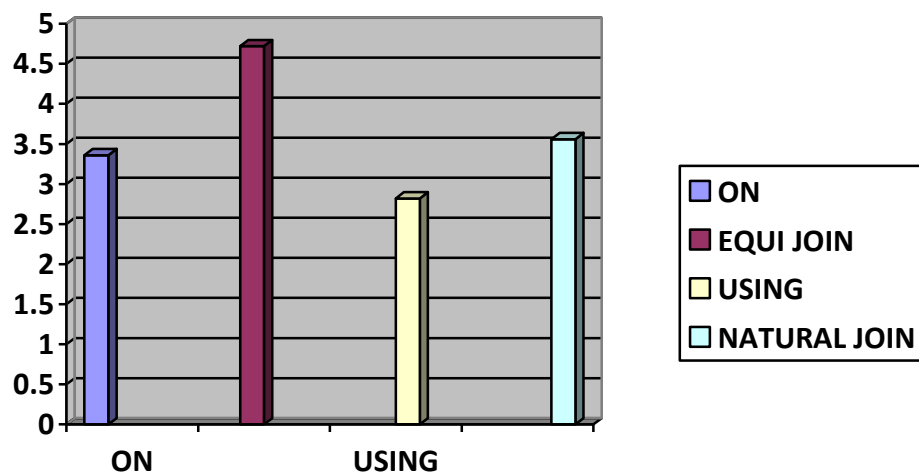


Figure 4.5.8: Power Consumption for Join Query Strategy Condition 1

#### **Discussion for Join Query**

Figure 4.5.8 shows the result of the Join Query done. The author joins two tables of Staff and Orders with both are having 100 data. The query with USING clause uses the least energy while query with EQUI JOIN clause is using the most power consumed. Query with ON clause uses the energy as almost the same as the query with NATURAL JOIN clause. The amount of data and data type of the table is not affecting the power consumption as both queries is using the same tables thus handling the same amount of data and the same type of data type. Therefore, setting up query strategy with USING clause could save more energy compared to other clauses or statements.

## Set Query

### Condition 1: To return 2 sets of select queries

Table Involved	Query Strategy	Query Coding	Power Consumption Measurement (W)		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
ARIF_ORDERS MENU	To return 2 sets of select queries by INTERSECT	<u>Q1</u> SELECT ORDERID FROM ARIF_ORDERSMENU WHERE MENUID = 'M12' INTERSECT SELECT ORDERID FROM ARIF_ORDERSMENU WHERE MENUID = 'M5';	2.9	17.3	3.46
			4.1		
			3.2		
			2.3		
			4.8		
	To return 2 sets of select queries by SUBQUERY	<u>Q2</u> SELECT ORDERID FROM ARIF_ORDERSMENU WHERE ORDERID IN (SELECT ORDERID FROM ARIF_ORDERSMENU WHERE MENUID = 'M12') AND MENUID = 'M5';	4.7	15.9	3.18
			2.6		
			3.6		
			2.6		
			2.4		

Table 4.5.9: Set Query Strategy Condition 1



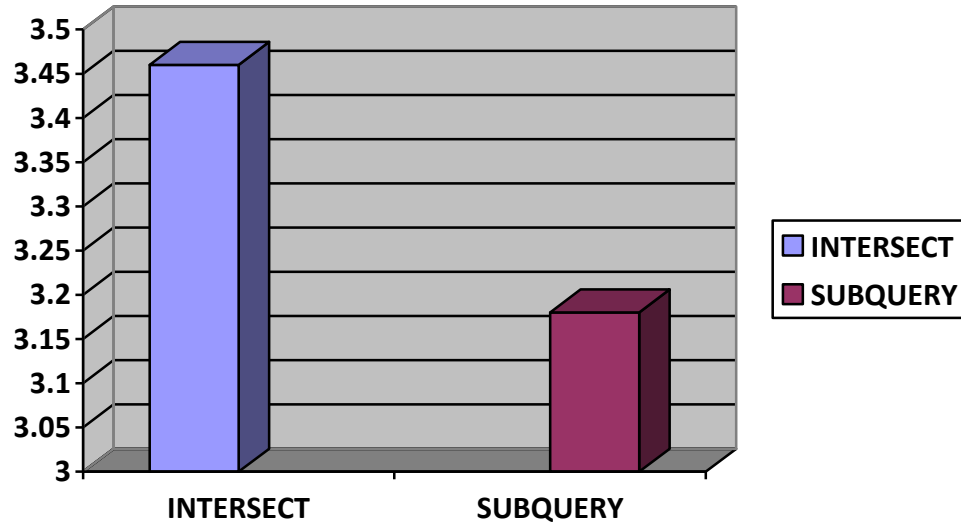


Figure 4.5.9: Power Consumption for Set Query Strategy Condition 1

#### **Discussion for Join Query**

Figure 4.5.9 shows INTERSECT query consumed more power compared to the query using SUBQUERY. As both of the queries is handling the same amount of data and data type, therefore the data and data type of the column are not affecting the power consumption. The platform seems to be using more energy when executing the INTERSECT statement. Therefore, database developers should be using SUBQUERY instead of INTERSECT whenever possible to support Green IT.

## **CHAPTER 5**

### **CONCLUSION AND RECOMMENDATIONS**

#### **5.1 Conclusion**

This project managed to be completed within 8 months thorough finding and relevant results. As completing the project, there are several strengths and weaknesses identified. Among the strength of this project are the sufficient related resources that the author can find easily as there are many related project been done in UTP itself before the author's project. These resources had help in making the project more precise and accurate. Other than that, UTP have been provided sufficient tools required with the help from group of experts in both technology and power domain. However, the main concern of this project is the power meter reading, where only small changes that is almost undistinguishable between power used by different set of queries. This is because this project is done within the small scale of database and which the author believes, in larger scale, the differences can be easily seen.

Based on the result gathered and captured on all the conditions, it is said that for each condition, the amount of power consumed differs and the least and most power consumed by the query strategy also not similar. Hence, it can be concluded that various query strategies and approaches do consume different amount of power under the same query condition. Hence, the objective of the project is achieved.

## **5.2 Recommendations**

In recommendation to this project, selecting the best query strategy that consume the least power to be used in query processing from the centralized database system would be a good alternative after analyzing the power usage of each query dataset. As much as it can save cost in electricity power and the operational cost in a bigger scale view for an organization, it also helps in reducing the release of heat and energy to environment which contributed to more favorable environment.

For future research recommendation, it would be best if the powers consumed by the data queries are being measured also with the time access required for the query retrieval activity and focuses more on reasoning factor why the queries use such power consumption. This project and research is a good study in having better performance in computer activity and in IT field in general which help in preserving green environment. It would contribute much to the benefits of technology advances in supporting the green computing.

## REFERENCES

- [1] Alonso R., Ganguly S. (1992) *Energy Efficient Query Optimization*. Technical Report, Matsuhita Infor Tech Lab.
- [2] Amer. A., Essary. D., (2008). *Predictive Data Grouping : Defining the bounds of energy and latency reduction through predictive data grouping and replication*. ACM Transaction on Storage (TOS), 4(1) ; 1-23. Retrieved June 22, 2014 from <http://dl.acm.org/citation.cfm?id=1353454>
- [3] Auseklis. N., Harmon. R., (2009). *Sustainable IT Services : Assesing the impact of green computing practices*. Management of Engineering; 1707-1717. Retrieved June 22, 2014 from <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=05261969>
- [4] Barroso L.A., Holzle U. (2007) *The case for Energy-Proportional Computing*. Computer 40(12):33-37.
- [5] *Green computing energy efficiency guide*. (2013). Retrieved June 21, 2014, from <http://www.computerweekly.com/guides/Using-green-computing-for-improving-energy-efficiency>
- [6] *Institute for Energy Efficiency*. (2008). UC Santa Barbara, Retrieved June 21, 2014, from <http://www.iee.ussb.edu/greenscale>
- [7] Kumar R. (2010) *Data center power, cooling and space: A worrisome outlook for the next two years*. Retrieved June 22, 2014 from [http://www.napatech.com/sites/default/files/dn-0431\\_multi-appliance\\_v02\\_us\\_letter\\_online.pdf](http://www.napatech.com/sites/default/files/dn-0431_multi-appliance_v02_us_letter_online.pdf)
- [8] Kurp. P., (2008). *Green Computing*. Communication of the ACM. Retrieved June 21, 2014 from <http://dl.acm.org/citation.cfm?id=1400186>

[9] Flach T., (2010). *Optimizing query execution to improve the energy efficiency of database management systems.*, 16-40. Hasso Plattner Institut.

[10] Wang. D., (2008). *Meeting Green Computing Challenges*. Proceedings of HDP'07 (858) , 1 – 5. Retrieved November 21, 2013 from <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04283590>

## APPENDICES



Project Set Up Environment



Power Meter Used