Analyzing Power Consumption in Query Processing
for Centralized Database System

By
Muhammad Nasrulhaq bin Mohd Bohari

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

SEPTEMBER 2012

UniversitiTeknologi PETRONAS
Bandar Seri Iskandar,
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

Analyzing Power Consumption in Query Processing for Centralized Database System

By

Muhammad Nasrulhaq bin Mohd Bohari

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

Approved by,

____

(Rozana bt Kasbon)

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.

MUHAMMAD NASRULHAQ BIN MOHD BOHARI

12753

ABSTRACT

Green computing is the environmentally responsible use of computers and related resources as well as reduced power resource consumption. In supporting to green computing, this project is being carried out with the main purpose to analyze the efficient performance of query processing activity in consuming power from a centralized database by implementing different query strategies in data grouping. In achieving the goal, each data query strategy retrieved from the database is measured based on its power consumption by using an energy saving power meter. A consolidated hotel management system is developed as to indicate the context of the project in testing on the query power usage from a centralized database system. This project emphasizes on the amount of power (in watts) that the query consumed based on the size of the query and the strategy used in assembling the data. It is more focusing on the simulation of query dataset in consuming power and not on developing a real-time system. Hence, 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, readers would be able to see and analyze that various query strategies applied in retrieving the same output under a specified condition gives dissimilar power reading which indicates the different amount of power consumption by each data query.

ACKNOWLEDGEMENT

Alhamdulillah, praise be to Allah S.W.T, the Most Gracious and Most Merciful for giving me the strength and courage in completing this project. I would like to express my greatest gratitude and appreciation to my supervisor, Puan Rozana Kasbon for her understanding and helpful guidance in order for me to complete the development of this project. Her critics and positive comments have been the milestone for me to make this project a huge success.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND NOMENCLATURES

CIS	Computer and Information Science
CPU	Computer Processing Unit
FYP	Final Year Project
PC	Personal Computer

TABLE OF CONTENT

CHAPTER 1
INTRODUCTION

1.1 Problem Statement

In most organizations these days, IT infrastructure has been the fundamental component in maintaining the operational activity. By installing huge data repositories, they are dealing with a large amount of data on daily basis which also engaging with high electricity cost. According to [1], the U.S. Environmental Protection Agency estimates that centralized computing infrastructures also known as data centers currently use 61 billion kilowatt hours of electricity used in the year of 2010. The statistics tells 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.

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. In order to have a better quality control of the environment in context of conserving energy from computer, one should know the amount of energy consumed by a PC and application activities that the PC processes. Although the world is promoting to have green computing, most users are not alarm to the personal usage of their computers and system activities. Hence, this project is conducted to specifically monitor the amount of energy used in data captured from 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 be further analysed by constructing different strategies of data queries to test whether it has links to the energy produced upon retrieving the data from the data storage.

1.2 Background of Study

Computer for instance, 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 surrounding. Thus, this project is being carried out to see on the power and energy consumed by different query strategies when data is retrieved from the database.

1.2.1 Significant and Relevancy

In supporting to green computing, this project is conducted to study the amount of power used by different query strategies in query processing activity from a central database system besides create awareness and educate the database developers and users in general. By reordering and applying good technique on data grouping, the energy consumed can be reduced and relatively helps in reducing computer activity energy while reducing the electricity cost and preserve the green environment.

1.2.2 Timeframe

This whole project is estimated to complete in approximately 8 months. The first 4 months are to find the related reading materials for literature review phase. The task is to collect and review the research 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 is carried out in the early of 5^{th} month which include designing the user interface and develop the data repository. This development phase shall be completed by the end of 6^{th} month. The remaining 7^{th} and 8^{th} month is used to test the query processing and executing activities and calculating the energy and power used while the processes take place. Throughout the development and testing phase, meeting with supervisor is conducted on regular basis to advise on the query structuring activity and ensuring accurate result. At the same time, each progress work is presented occasionally following to the CIS FYP schedule. By the end of the project, a complete hard bound paper work is submitted to the supervisor.

1.2.3 Technical Feasibility

- The system developed is referring to the centralized Hotel Management System where it acts as the main branch database and could access to various data from other hotels and from different branch and location. But the query testing process is only focused on the local host basis and does not involve any network connection.

- This project is a research based project which merely emphasizes more on research element on the amount of energy consumed in query processing. The functionality and reliability of the system is not the main focus.

- Tools used in this project are divided into two categories; hardware and software. For hardware, the equipment includes PC (personal notebook) and a plug meter to capture the energy and power released. While for software involves Oracle Database Application Express 11g to and develop the database and test the data query. The feature and specification for each tools is discussed further in Chapter 3 : Methodology under Tools Utilization section (Table 2).

1.3 Objectives

The main objective of this project is to analyze the efficient performance of the query processing in consuming power by applying various query processing strategies in a centralized database system. It is done by measuring the amount of power used for each query executed at a fixed time period. Among other relevant objectives include :

- To look upon on the different strategies used in constructing data query that can produce different amount of energy.
- To identify the contributing factors that affect to the energy consumed by the data query based on the following hypothesis :

Hypothesis

Larger volume of data contained in the database will give higher electricity reading.

1.4 Scope of Study

To be more specific, this project is concentrated on determining how efficient data executing activities in consuming power and energy in a centralized database system. The system architecture of the project is based on client-server and centralized database management model. See Figure 1.1 below.

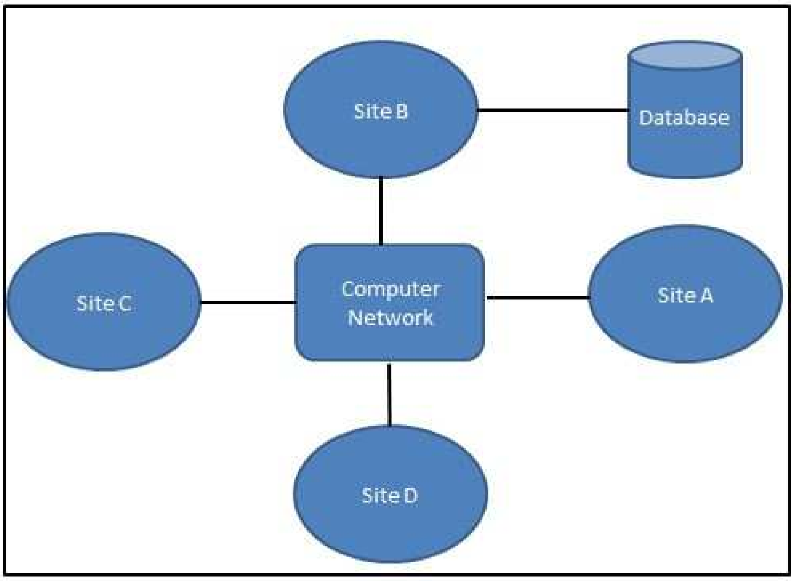

Figure 1.1: Project System Architecture

CHAPTER 2
LITERATURE REVIEW

2.1 Green Computing as Better Alternative

Along the fast pace changing of technology, people are getting more concern on having a green and safe environment. Every day, more and more users are taking into account on their surrounding climatic change and find ways to help in conserving the energy released. In IT context, the energy produced by data repository does contribute to the world by releasing heat and increases the surrounding temperature. In supporting green computing, we will see on how much the energy consumed in a single and multiple data retrieval from a database.

As the commitment to reduce environmental impact and power consumption are becoming increasingly important objectives for organizations, architecture leaders are now proactively considering environmental resource constraints along with more traditional IT business goals [1]. Day by day, more organizations are realizing on the importance of having a controlled amount of power consumption. Provided with high maintenance of large volume data storage (mainly servers) and power consumed by business operations, firms are now taking precautionary and wise actions in placing environmental resource parallel to IT business goals. According to [2], data centers are the driving force behind the green computing environment.

2.2 Data Storage Reduces Power Consumption

Many people think that energy consumed can only be done in big activity scale such as shutting down the PC, put the PC under sleep or hibernate and so on. But in large companies, this task is nearly hard to do especially during the working hours as it may interrupt the business process. Excessive restarts can be detrimental to system reliability, incurring delay, and always consuming power which we would prefer to conserve if possible [3]. Energy consumption can also be reduced in smaller area such as having data repository. Data storage is another area where performance improvement is achieved with less power consumption, not more [4] . As compared to the traditional daily business process where it requires less data storage capacity, today's modern technology has certainly contribute in reducing the energy consumed.

2.3 Energy Consumption in Data Accessing

As much as data storage helps in lessen the usage of energy, the process of accessing the data from the database is still power consuming. Nearly 59 % of overall energy in a typical query execution is spent in the main memory (excluding I/O devices) [5]. Though it releases small percentage of energy, if it is involved in repetitive tasks from large number of units, the energy produced will relatively become larger. Plus, as the volume of data and query size increases, the access time will also increases [6] , which is going to affect the efficient performance of the data execution process. A study has been done by taking a disk as an example of data storage to illustrate the energy consumed upon retrieving data from its storage. As a result, it suggests that predictive data grouping can be a better alternative in reducing energy consumption for accessing data. According to [3], for energy, he shows how predictive grouping can even save energy use for an individual disk that is never idle. Apart of other factors such as data access latency, this approach has subsequently helping in lessen the energy consumption of data retrieval from PC.

CHAPTER 3
METHODOLOGY

The research methodology chapter represents the strategies and approaches that consist of collecting and analyzing data collected in order for meaningful analysis and interpretations of the research findings 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 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

Under this 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. Apart from testing the system itself and the reliability of the plug meter, the reason of having this method is simply to test on the energy reading of the query retrieved from the database.

3.1.3 Systematic Random Sampling Method

This method is used to test on 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 strategies. It is conducted to see which query strategies use the least energy in a given time period correspond to the query complexity and volume of data contained. The data is then recorded.

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 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 Methods for Data Presentation

The data collected is being presented in a table form. Table below shows the draft template of the table used to organize the data captured.

Table 3.1: Draft Template for Recorded Result in Data Representation

Query	Strategies	Table Involved	Description	Efficiency Measurement		
				Power Reading (w)	Total Power Reading (w)	Average Power Consumed (w)

3.3 Key MileStone

3.3.1 Final Year Project I

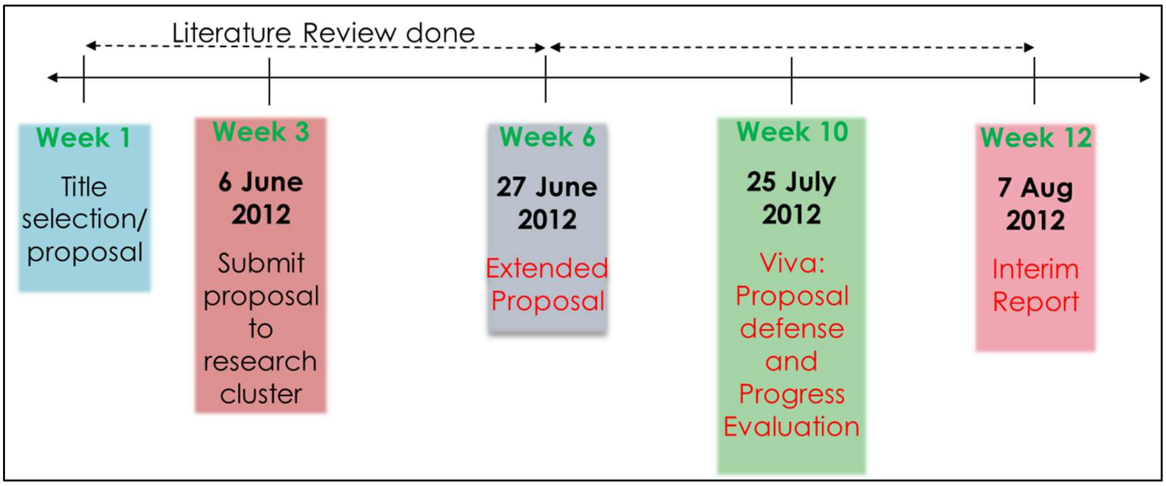

Figure 3.1: Final Year Project I Key Milestone

3.3.2 Final Year Project II

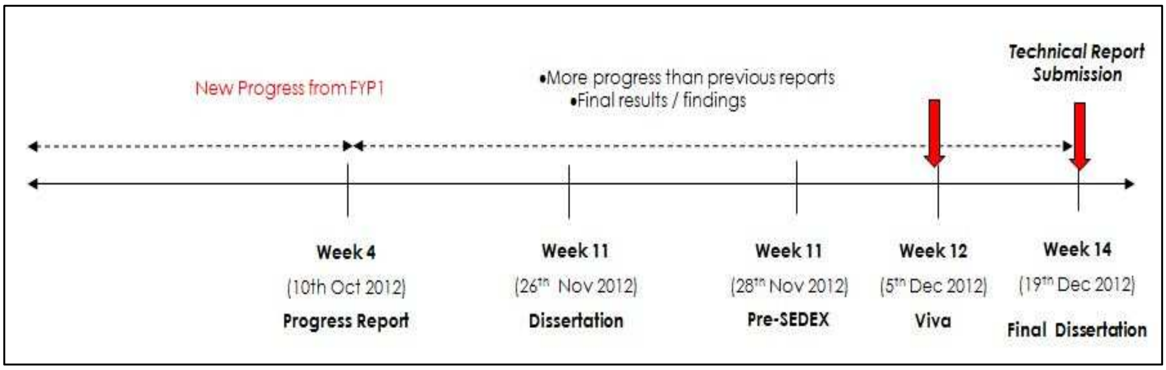

Figure 3.2: Final Year II Key Milestone

Both Figure 3.1 and 3.2 above show the key milestone of the Final Year Project as a whole. It highlights more on the formal report and paper work activity with submission of deliverables to the respective panels such as supervisor, coordinator and external examiner. Also, it is focused on the required presentations that final year students need to follow as guidelines as part of assessment in completing the course.

3.4 Gantt Chart

No	Phase/Month	1	2	3	4	5	6	7	8
1	Requirement / Analysis								
2	Make project timeframe plan	■							
3	Set topic / problem statement / objective and analyze project background	■	■						
4	Collect reading materials / researched paper		■	■					
5	Design								
6	Design the system architecture of the project			■	■				
7	Design the Class Diagram				■	■			
8	Design the Data Capturing Procedure				■	■			
9	Implement / Code								
10	Develop database in Oracle					■	■		
11	Structuring the query strategies					■	■	■	
12	Testing								
13	Test each query strategies on power consumption							■	■
14	Tabulate the result							■	■

Figure 3.3: Project Gantt Chart

Figure 3.5 above shows the Gantt chart of the project from the beginning until it is solely completed which takes about 8 months to be accomplished. The activity done in each phases adhered to the project methodology in a whole which is further discussed in the Project Activities section.

3.5 Project Activities

3.5.1 Waterfall Methodology Model

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.

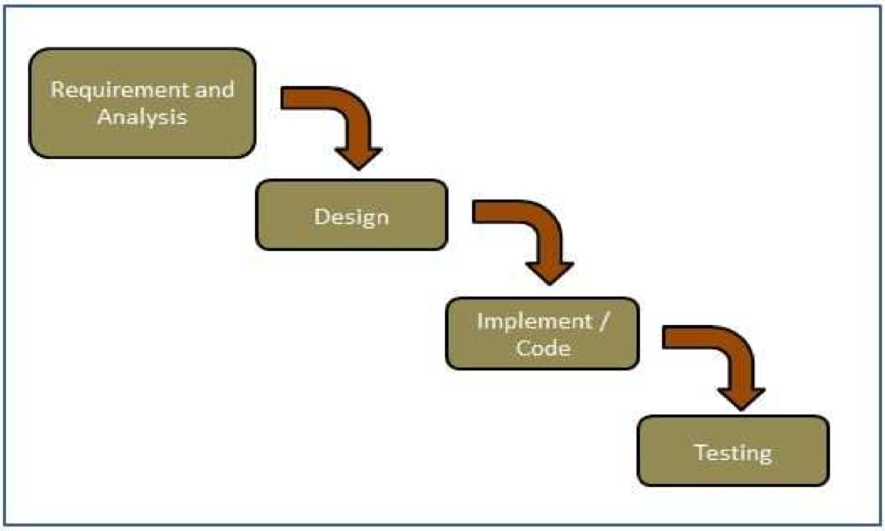

Figure 3.4: Project Activity Model: Waterfall Methodology

- Requirement and Analysis

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, objective and background of study are analyzed under this beginning phase. This also includes collecting all the research papers on related topic in exploring the study that has been done previously by the experts.

- Design

This phase is started by designing the system architecture of the whole project which gives clearer view on how the project system runs. It also involves designing the class diagram and the possible query structure based on the collective requirements. The data capturing procedure is also designed to describe the procedural activity involved in collecting the result.

- Implement / Code

This stage is about implementing what has been planned in the earlier phases into a working system. It involves a lot of coding activity whereby the database is developed and the possible query strategy is constructed.

- Testing

This phase is the most important phase for the whole project. Besides testing on the functionality and reliability of the whole system itself, this phase also focuses on testing the query processing while measuring the power consumption. This activity is done repetitively to get as accurate result as possible in order to produce precise outcome of the project. All the information is recorded to be as the project result.

3.5.2 Data Capturing Procedure

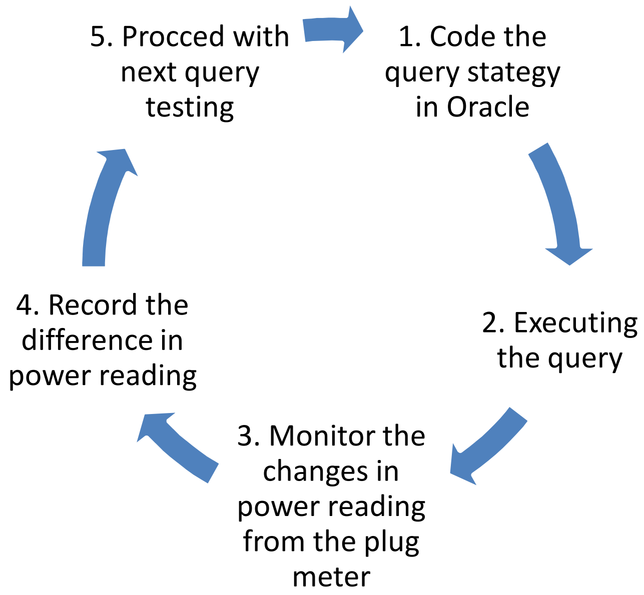

Figure 3.5: Data Capturing Procedure

Figure 3.6 above describes the procedure used in capturing the data required for the project. To begin, the database is developed by using Oracle with the number of 10 tables involved. After that, the query strategy is constructed by joining several tables depending on the data required for the specified condition. Once the query has been coded, it is then being executed to see on the result or output. All query strategies in one similar condition should have the same output of result. Any changes in the power reading on the energy plug meter is carefully monitored which then be recorded for further analysis. The steps are then repeated for the next query strategies onwards.

3.6 Tools Utilization

Table 3.2: Project Tools Feature and Specification

No	Category	Client	Remark(s)
1	Hardware	Mobile Device (PC) Intel® Core™ 2 Due CPU 2.00GHz, 2000Mhz, 2Core(s) 3.00 GB RAM	Personal computer with relative specification and features. It will be used to develop the system and to run the usability testing
		Energy Saving Power Meter • Voltage Rating: 240V ac 50Hz • Current rating: 13A • Maximum load: 3120W • Power Consumption: <0.5W • Metering: • Voltage 200 to 276V AC,. +/- 1.5% • Frequency range 45 to 65Hz	This device will be used to calculate the electric current and energy used for data accessing process.
2	Software	Oracle Application Express 11g	This application will be used to develop a database
3	Operating system	32-bit Windows 7 Home Premium	The OS used to run and test the entire project

CHAPTER 4
RESULT AND FINDINGS

4.1 Class Diagram

Figure 4.1 below shows the class diagram of centralized hotel management system that is used for this project. There are a total of ten tables with defined attributes involved in the database including two bridge entities, Booking and Room_Promotion_Rate. The relationships and cardinality are also mentioned.

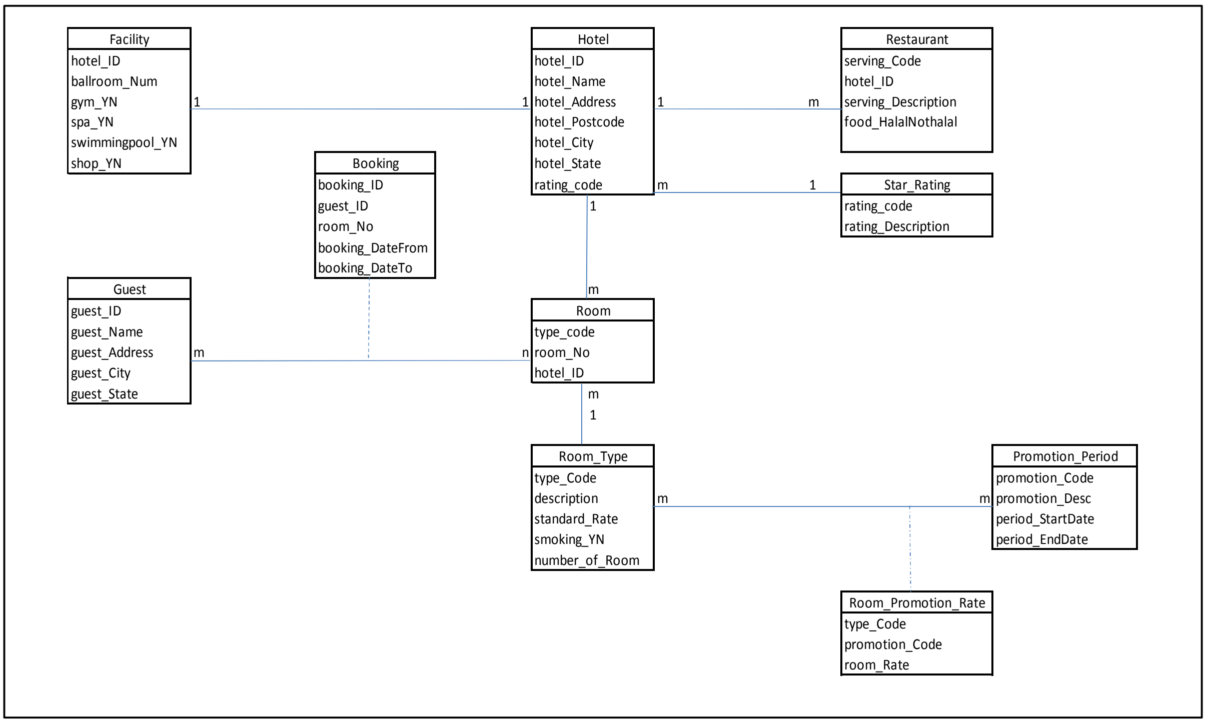

Figure 4.1 : Class Diagram on Centralized Hotel Management System

Based on the class diagram, a database is developed by using Oracle Application Express 11g.

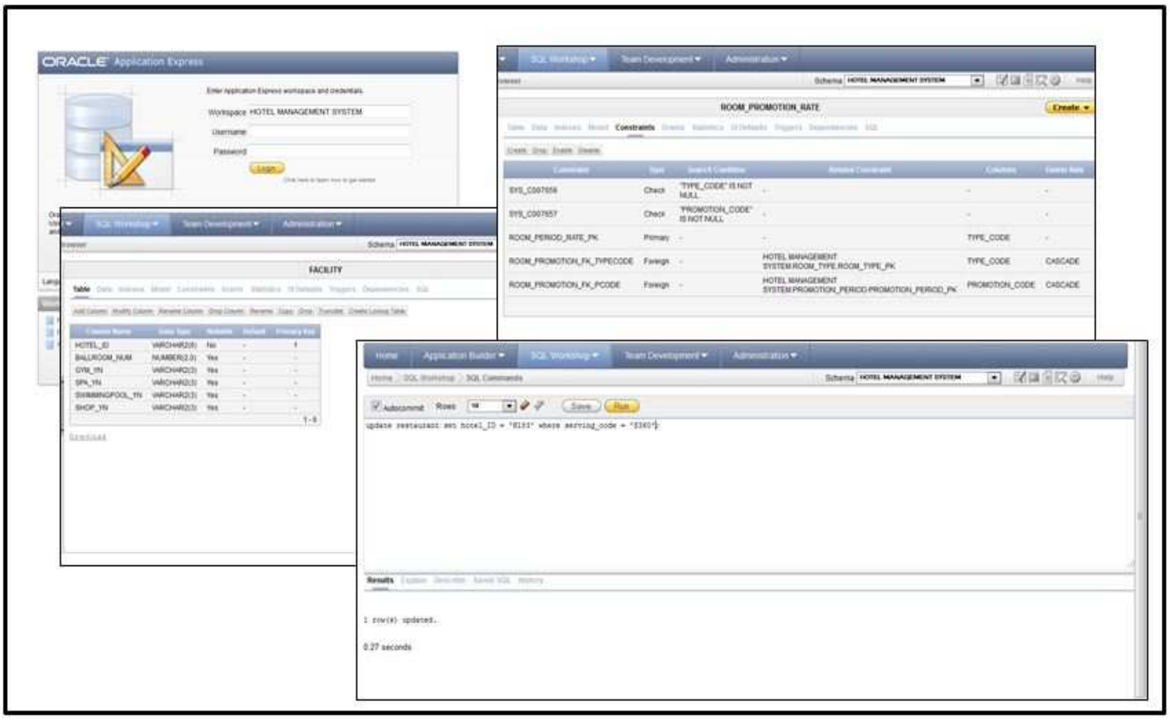

Figure 4.2: System Interface in Oracle

Figure above shows the system interface in Oracle for the project. Generating a huge amount of data may be time consuming. Thus, to optimize the time to complete the project, the data in the database is generated from a website named generatedata.com. This website provides an enormous amount of random data based on selected criteria such as name, address, state and so on. It also generates SQL statement which helps in minimizing the time taken at the data entry level. Each table have different number of records depending on reasonableness of the attributes in the tables.

4.2 Type of JOIN query and Other Forms of Query

In creating the query, many approaches are used to develop various kind of queries which at the same time executed the same data. For a single database system, generally joining the tables is the most common way to perform a query and it is focused on which there are many ways to perform. Below are the strategies to construct the query :

4.2.1 JOIN query

I. EQUI JOIN query

This join is the most common in SQL and said to be the default join type where it uses only equality comparison in combining column values of two tables based on the join-predicate. The query compares each row in both tables to find all pairs of rows which satisfy the join-predicate. When the join-predicate is satisfied, column values for each matched pair of rows are combined into a result row. EQUI JOIN divides into two types as below:

I. OUTER JOIN

This join does not require the column in the two joined tables to have similar records. On the other hand, the joined tables remain the records even if there is no matching record exists.

II. INNER JOIN

This type of EQUI JOIN returns all rows from tables where the key record of one table is equal to the key records of another table.

II. NATURAL JOIN query

A type of equi-join and is structured in such a way that, columns with same name of associate tables will appear once only.

4.2.2 Other Forms of Query

I. UNION query

The SQL UNION query allows combining the result sets of 2 or more SQL SELECT statements. It removes duplicate rows between the various SELECT statements. Each SQL SELECT statement within the UNION query must have the same number of fields in the result sets with similar data types.

II. INTERSECT query

The SQL INTERSECT query allows returning the results of 2 or more "select" queries. However, it only returns the rows selected by all queries. If a record exists in one query and not in the other, it will be omitted from the INTERSECT results.

III. MINUS query

The SQL MINUS query returns all records in the first SQL SELECT statement that are not returned in the second SQL SELECT statement. Each SQL SELECT statement must have the same number of fields in the result sets with similar data types.

Besides mentioned above, there are other functions and clauses that are used in the making of query which also used in this project. Such examples include ON condition, IN condition, USING clause and much more.

4.3 Formula in Calculating Power Consumed

By thorough research, there are many ways in calculating the energy consumed by the query processing. In this project , the testing to find the power reading for a query is conducted for five times and the average of total power reading is then be recorded to be the result of the power consumed by the query set.

The formula is as follow :

$$\textbf{Power Consumption, W} = \frac{\sum n}{k=5}$$

Where n = number of power reading and k = number of times the test is conducted

4.4 Testing Set up and Environment

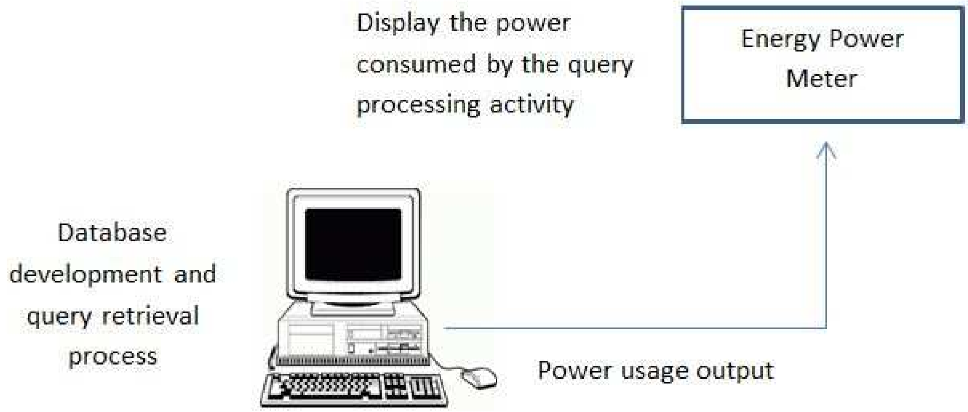

Figure 4.3: Project Testing Setup

Figure 4.3 above shows the testing set up of the project. The PC is used to develop the central database of the hotel management system by using Oracle Application Express 11g. Also, the retrieval process takes place in the PC. The initial value is set up on the energy power reading before the query is being executed. This initial value indicates the highest power consumed by the PC at the moment before the data is retrieved from the database. Once the query has run, any immediate changes on the power reading is captured by the power meter and recorded.

In getting the favorable result on the power consumption, certain conditions need to be set up as testing environment to minimize the error while recording the power reading on the power meter and it is being held carefully to get more accurate result. The testing procedure takes place at the room in the hostel and the result gathered is based on the specification of the tools and equipment used in this project (refer to table 3.2). In enhancing the accuracy of the result, no other application is running on the system background along with the Oracle. Besides that, no network connection is involved during the procedure is performed since it only comprises single database system and it is local host basis.

4.5 Query Strategies and the Power Consumed

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. Some of the conditions are as follow :

Condition 1

To display hotel name and its star rating from different state

Condition 2

To display the Guest booking information

Condition 3

To identify the number of hotels that have gym facility and halal food restaurant

Condition 4

To identify the 5 star hotels that offers promotion rate for Luxury room

Condition 5

To display guest name list from 5-star Hotel

The results of all the power readings for the different query for the conditions are gathered and calculated in the following tables.

4.5.1 Condition 1

To display hotel name and its star rating from different state

Table 4.1: Query Strategies and Power Consumption for Query Condition 1

Table Involved	Strategy	Query	Power Consumption Measurement		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
HOTEL and STAR_RATING	To join 2 tables by using ON clause	**Q1** SELECT H.HOTEL_NAME, H.HOTEL_STATE, SR.STAR_NUMBER FROM HOTEL H JOIN STAR_RATING SR ON H.RATING_CODE = SR.RATING_CODE ORDER BY SR.STAR_NUMBER;	1.9 2.3 4.0 1.1 2.0	11.3	2.26
	To join 2 tables by using EQUI JOIN approach	**Q2** SELECT H.HOTEL_NAME, H.HOTEL_STATE, SR.STAR_NUMBER FROM HOTEL H, STAR_RATING SR WHERE H.RATING_CODE = SR.RATING_CODE ORDER BY SR.STAR_NUMBER;	0.4 2.9 2.4 3.2 2.0	10.9	2.18
	To join 2 tables by using USING clause	**Q3** SELECT HOTEL_NAME, HOTEL_STATE, STAR_NUMBER FROM HOTEL JOIN STAR_RATING USING (RATING_CODE) ORDER BY STAR_NUMBER;	0.0 1.3 3.0 1.9 1.9	8.1	1.62
	To join 2 tables by using NATURAL JOIN approach	**Q4** SELECT HOTEL_NAME, HOTEL_STATE, STAR_NUMBER FROM HOTEL NATURAL JOIN STAR_RATING ORDER BY STAR_NUMBER;	2.2 1.2 0.0 1.1 2.5	7.0	1.40

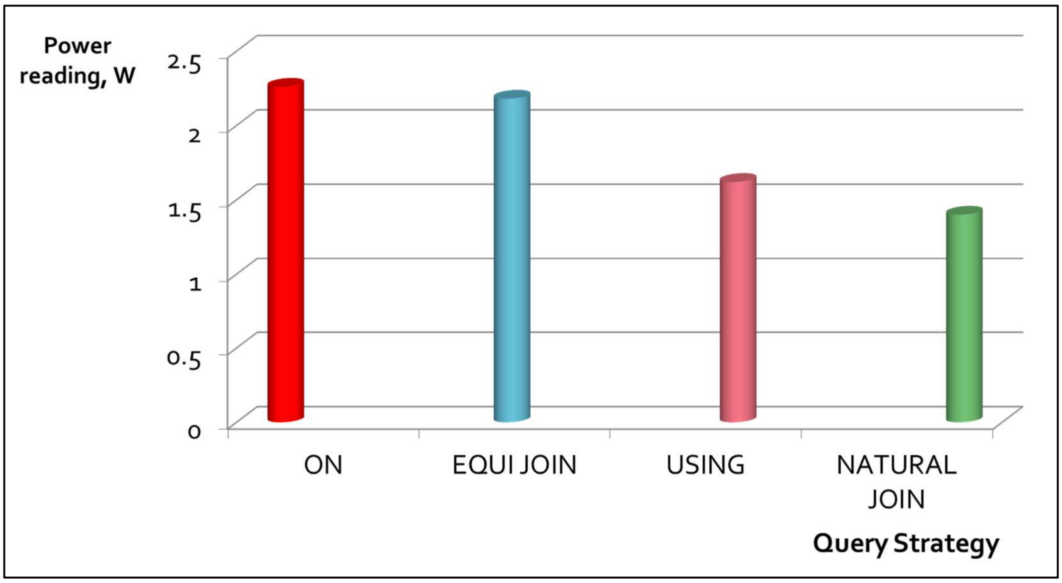

Figure 4.4: Power Consumption by Different Query Strategies for Condition 1

Figure 4.4 above shows the power consumed by each query strategies for condition 1. The number of tables involve are two, Hotel which has 200 records and Star_Rating for 50 records. Based on the condition, the output returns 200 number of records. Based on the recorded data, all strategies use different amount of power and the query strategy that uses NATURAL JOIN approach has the lowest power consumption while the query that use ON clause has highest power reading.

4.5.2 Condition 2

To display the Guest booking information

Table 4.2: Query Strategies and Power Consumption for Query Condition 2

Table Involved	Strategy	Query	Power Consumption Measurement		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
GUEST and BOOKING	To join 2 tables by using ON condition	**Q1** SELECT G.GUEST_NAME, B.BOOKING_ID, B.BOOKING_DATEFROM, B.BOOKING_DATETO, B.ROOM_NO FROM GUEST G JOIN BOOKING B ON G.GUEST_ID = B.GUEST_ID AND ROOM_NO IS NOT NULL ORDER BY BOOKING_ID;	2.0 1.1 1.9 2.0 1.9	8.1	1.62
	To join 2 tables by using EQUI JOIN approach	**Q2** SELECT G.GUEST_NAME, B.BOOKING_ID, B.BOOKING_DATEFROM, B.BOOKING_DATETO,B.ROOM_NO FROM BOOKING B, GUEST G WHERE G.GUEST_ID = B.GUEST_ID AND B.ROOM_NO IS NOT NULL ORDER BY B.BOOKING_ID ;	1.1 0.7 1.9 0.6 2.4	6.7	1.34
	To join 2 tables by using USING clause	**Q3** SELECT GUEST_NAME, BOOKING_ID, BOOKING_DATEFROM, BOOKING_DATETO, ROOM_NO FROM GUEST JOIN BOOKING USING (GUEST_ID) WHERE ROOM_NO IS NOT NULL ORDER BY BOOKING_ID;	0.8 1.0 0.8 1.9 1.9	6.4	1.28
	To join 2 tables by using NATURAL JOIN approach	**Q4** SELECT GUEST_NAME, BOOKING_ID, BOOKING_DATEFROM, BOOKING_DATETO, ROOM_NO FROM GUEST NATURAL JOIN BOOKING WHERE ROOM_NO IS NOT NULL ORDER BY BOOKING_ID;	1.7 0.9 1.5 1.1 0.7	5.9	1.18

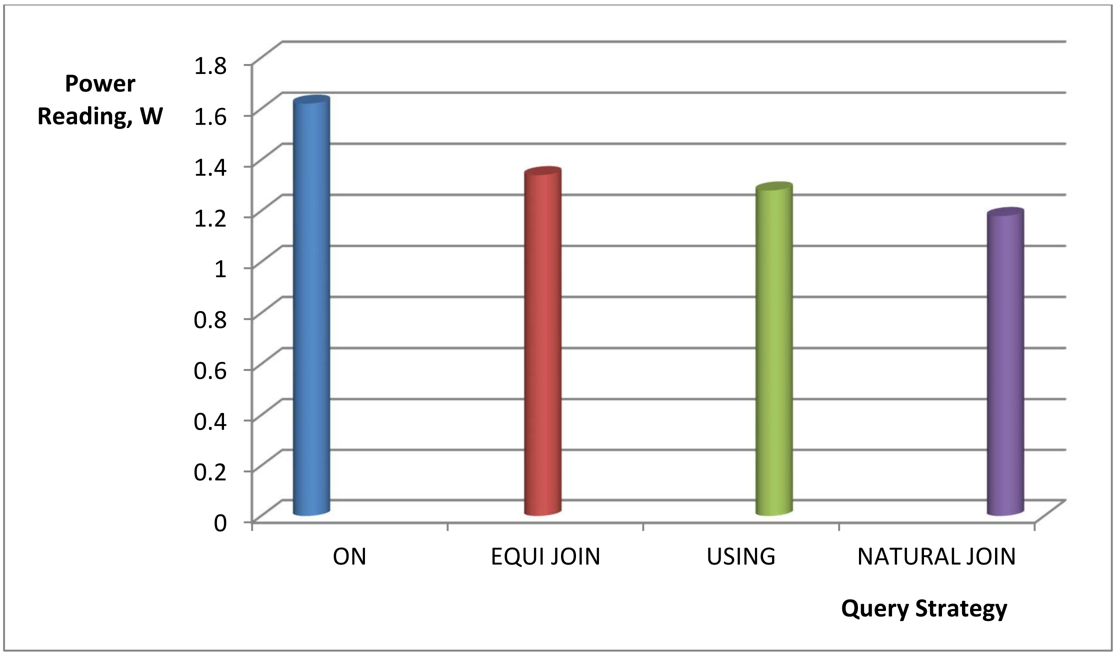

Figure 4.5: Power Consumption by Different Query Strategies for Condition 2

Figure 4.5 above shows the power consumed by each query strategies for condition 2. There are 2 tables involve under this condition, Guest and Booking. Each table has over 1000 records and 500 records respectively. Since it requires to display the information of guest, the output of the query is 500 records. Based on the recorded data, different query strategies implied to get the output consume different amount of power and the query that uses NATURAL JOIN approach has the lowest power consumption while the query that use ON clause has highest power reading.

4.5.3 Condition 3

To identify the number of hotels that has gym facility and halal food restaurant

Table 4.3: Query Strategies and Power Consumption for Query Condition 3

Table Involved	Strategy	Query	Power Consumption Measurement		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
HOTEL, FACILITY and RESTAURANT	To join 3 tables by using EQUI JOIN approach	**Q1** SELECT COUNT (DISTICNT H.HOTEL_ID) AS HOTEL_ID FROM HOTEL H,FACILITY F, RESTAURANT R WHERE H.HOTEL_ID = F.HOTEL_ID AND F.GYM_YN = 'Yes' AND H.HOTEL_ID = R.HOTEL_ID AND R.FOODHALALNOTHALAL= 'Halal';	0.0	4.1	0.82
			1.1		
			2.2		
			0.9		
			0.0		
	To join 3 tables by using INTERSECT clause	**Q2** SELECT (HOTEL_ID) FROM HOTEL INTERSECT SELECT (HOTEL_ID) FROM FACILITY WHERE GYM_YN = 'Yes' INTERSECT SELECT (HOTEL_ID) FROM RESTAURANT WHERE FOOD_HALALNOTHALAL = 'Halal';	2.1	9.8	1.96
			2.7		
			2.4		
			0.6		
			2.0		
	To join 3 tables by using IN clause	**Q3** SELECT COUNT (HOTEL_ID) FROM HOTEL WHERE HOTEL_ID IN (SELECT HOTEL_ID FROM FACILITY WHERE GYM_YN = 'Yes' AND HOTEL_ID IN (SELECT HOTEL_ID FROM RESTAURANT WHERE FOOD_HALALNOTHALAL = 'Halal'));	0.0	8.1	1.62
			1.3		
			3.0		
			1.9		
			1.9		

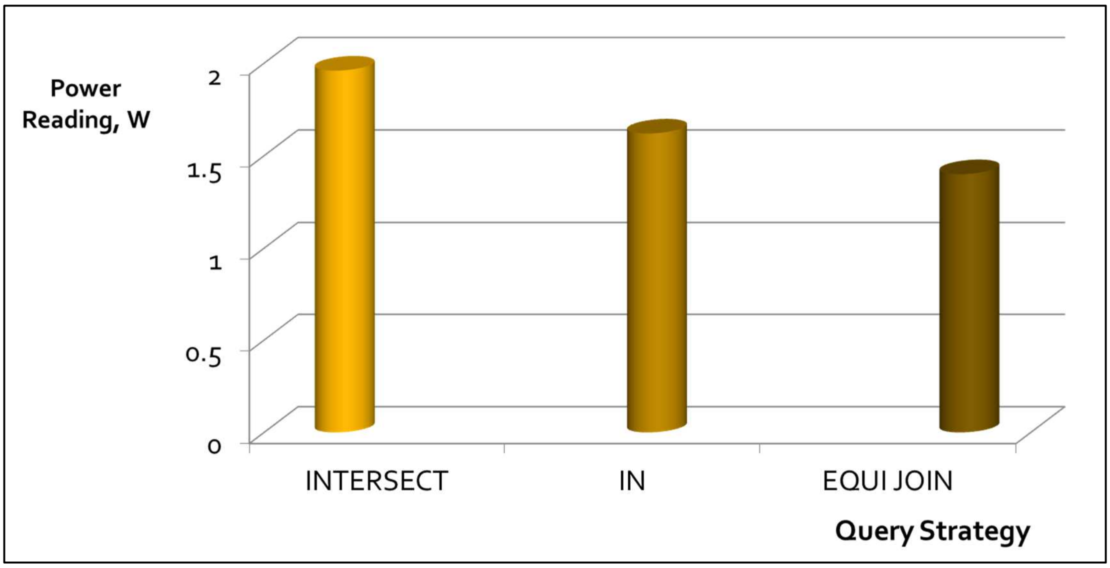

Figure 4.6: Power Consumption by Different Query Strategies for Condition 3

Figure 4.6 above illustrates the query strategies used condition 3 and distinguish power usage by each query strategies. Three tables are involved for this condition, Hotel with 500 records, Facility and Restaurant with 200 records each. According to the condition, there are only 7 hotels that have gym and provide halal food restaurant, which means the query only returns 7 records. Three possible queries are used to get the output which is EQUI JOIN approach, INTERSECT-clause approach and IN-clause approach. Based on the three strategies, the query that use EQUI JOIN approach has the lowest energy reading and INTERSECT-clause strategy has the highest power consumption.

4.5.4 Condition 4

To identify the 5 star hotels that offers promotion rate for Luxury room

Table 4.4: Query Strategies and Power Consumption for Query Condition 4

<table>
<tr><th rowspan="2">Table Involved</th><th rowspan="2">Strategy</th><th rowspan="2">Query</th><th colspan="3">Power Consumption Measurement</th></tr>
<tr><th>Power Reading (W)</th><th>Total Power Reading (W)</th><th>Average Power Consumed (W)</th></tr>
<tr><td rowspan="15">HOTEL, STAR_RATING, ROOM and ROOM_ PROMOTION_ RATE</td><td rowspan="5">To join 4 tables by using EQUI JOIN approach</td><td rowspan="5"><u>Q1</u>
SELECT H.HOTEL_NAME , H.HOTEL_STATE FROM HOTEL H, STAR_RATING SR, ROOM R, ROOM_PROMOTION_RATE RP WHERE H.RATING_CODE = SR.RATING_CODE AND SR.RATING_NUMBER = 5 AND H.HOTEL_ID = R.HOTEL_ID AND R.TYPE_CODE = 'TC0050' AND R.TYPE_CODE = RP.TYPE_CODE;</td><td>0.7</td><td rowspan="5">7.6</td><td rowspan="5">1.52</td></tr>
<tr><td>0.0</td></tr>
<tr><td>1.8</td></tr>
<tr><td>2.9</td></tr>
<tr><td>2.2</td></tr>
<tr><td rowspan="5">To join 4 tables by using IN clause approach</td><td rowspan="5"><u>Q2</u>
SELECT HOTEL_NAME, HOTEL_STATE FROM HOTEL WHERE RATING_CODE IN (SELECT RATING_CODE FROM STAR_RATING WHERE STAR_NUMBER = 5) AND HOTEL_ID IN (SELECT HOTEL_ID FROM ROOM WHERE TYPE_CODE = 'TC0050'AND TYPE_CODE IN (SELECT TYPE_CODE FROM ROOM_PROMOTION_RATE));</td><td>0.8</td><td rowspan="5">8.4</td><td rowspan="5">1.68</td></tr>
<tr><td>2.8</td></tr>
<tr><td>2.1</td></tr>
<tr><td>2.0</td></tr>
<tr><td>0.7</td></tr>
<tr><td rowspan="5">To join 4 tables by using ON and IN clause approach</td><td rowspan="5"><u>Q3</u>
SELECT H.HOTEL_NAME, H.HOTEL_STATE FROM HOTEL H JOIN STAR_RATING SR ON H.RATING_CODE = SR.RATING_CODE WHERE SR.RATING_NUMBER = 5 AND H.HOTEL_ID IN (SELECT R.HOTEL_ID FROM ROOM R JOIN ROOM_PROMOTION_RATE RP ON R.TYPE_CODE = 'TC0050' AND R.TYPE_CODE = RP.TYPE_CODE);</td><td>1.1</td><td rowspan="5">7.3</td><td rowspan="5">1.46</td></tr>
<tr><td>1.1</td></tr>
<tr><td>2.2</td></tr>
<tr><td>1.2</td></tr>
<tr><td>1.7</td></tr>
</table>

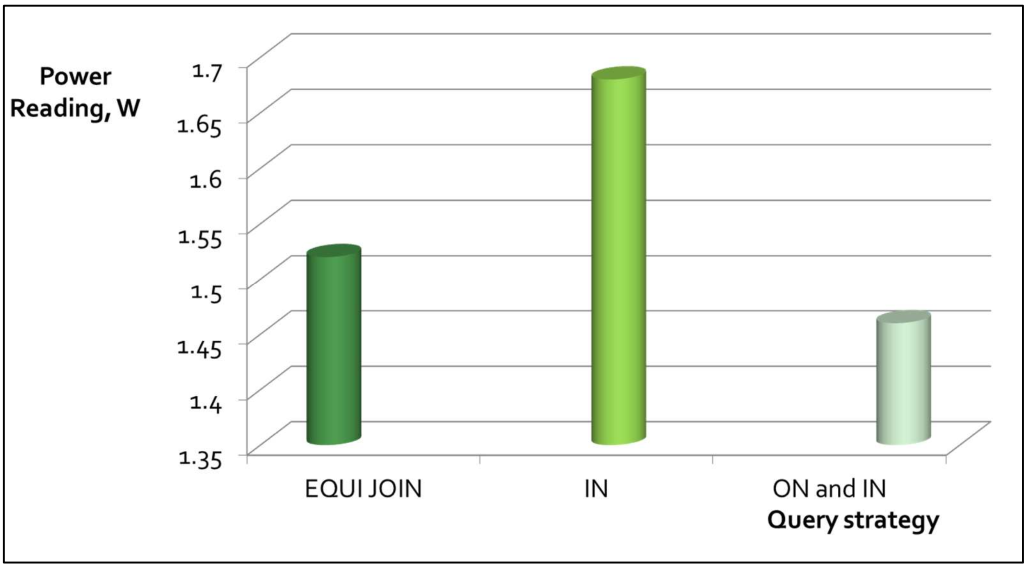

Figure 4.7: Power Consumption by Different Query Strategies for Condition 4

Figure 4.7 above illustrates the analysis under condition 4 on power usage by different query strategies applied. There are 4 tables involved for this condition, Hotel, Star_Rating, Room and Room_Promotion_Rate. Each table has different number of records which as follows; Hotel: 500 records, Star_Rating: 50 records, Room: 564 records and Room_Promotion_Rate: 50 records. Based on the data retrieved from the database, there are three 5 star hotels that offers promotion rate for Luxury room. To execute the output, three possible queries are applied which are EQUI JOIN approach, IN-clause approach and ON-and-IN-clause approach. Based on the data gathered, it shows that IN query strategy has relatively high power consumed as compared to other query strategies and hence the highest among all three strategies, while ON-and-IN approach has the least power consumption in retrieving the data from the database.

4.5.5 Condition 5

To display guest name list from 5-star Hotel

Table 4.5: Query Strategies and Power Consumption for Query Condition 5

Table Involved	Strategy	Query	Power Consumption Measurement		
			Power Reading (W)	Total Power Reading (W)	Average Power Consumed (W)
HOTEL, STAR_RATING ROOM, BOOKING, and GUEST	To join 5 tables by using EQUI JOIN approach	**Q1** SELECT G.GUEST_NAME FROM GUEST G, BOOKING B, ROOM R, HOTEL H, STAR_RATING SR WHERE H.RATING_CODE = SR.RATING_CODE AND SR.STAR_NUMBER = 5 AND H.HOTEL_ID = R.HOTEL_ID AND R.ROOM_NO = B.ROOM_NO AND B.GUEST_ID = G.GUEST_ID;	1.0 1.5 1.4 1.1 1.3	6.3	1.26
	To join 5 tables by using IN clause approach	**Q2** SELECT GUEST_NAME FROM GUEST WHERE GUEST_ID IN (SELECT GUEST_ID FROM BOOKING WHERE ROOM_NO IN (SELECT ROOM_NO FROM ROOM WHERE HOTEL_ID IN (SELECT HOTEL_ID FROM HOTEL WHERE RATING_CODE IN (SELECT RATING_CODE FROM STAR_RATING WHERE STAR_NUMBER = 5))));	2.5 1.3 1.4 0.7 2.3	8.2	1.64
	To join 5 tables by using ON and IN clause approach	**Q3** SELECT G.GUEST_NAME FROM GUEST G JOIN BOOKING B ON G.GUEST_ID = B.GUEST_ID AND B.ROOM_NO IN (SELECT R.ROOM_NO FROM ROOM R JOIN HOTEL H ON R.HOTEL_ID = H.HOTEL_ID AND H.RATING_CODE IN (SELECT RATING_CODE FROM STAR_RATING WHERE STAR_NUMBER = 5));	2.2 1.7 1.3 0.5 1.8	7.5	1.50

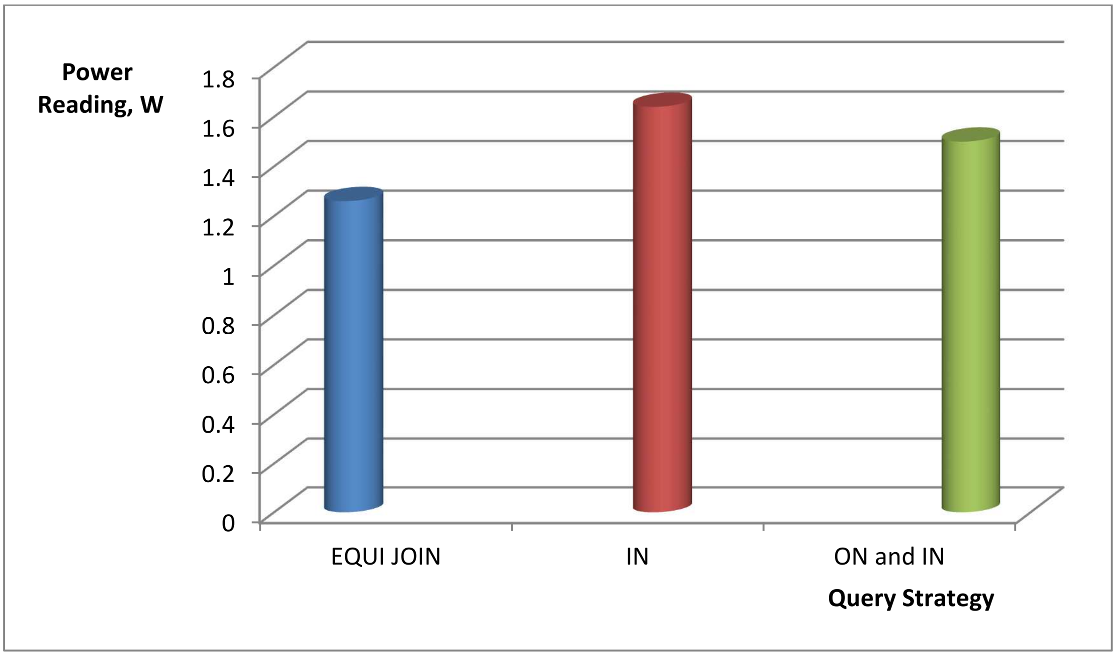

Figure 4.8: Power Consumption by Different Query Strategies for Condition 5

Figure 4.8 above shows the power consumption by three different query strategies applied under condition 5. This condition is involving 5 tables, Hotel, Star_Rating, Room and Booking and Guest. Hotel has 200 records, Star_Rating has 50 records, Room has over 500 records, Booking has 500 and Guest has over 1000 records accordingly. The output of this condition is 36 guests that are staying in 5-star hotel, which indicates 36 number of records. Three query strategies are applied for the data retrieval activity which enlisted by EQUI JOIN approach, IN-clause approach and ON-and-IN-clause approach. Based on the data captured and calculated, the number of power reading of all strategies vary in which IN-clause approach has the highest consumption power among all the three strategies and EQUI JOIN query strategy has the lowest power reading.

CHAPTER 5
CONCLUSION AND RECOMMENDATION

5.1 Conclusion

This project is being conducted for 8 months to complete with thorough findings and relevant results. Upon completing this project, several strengths and weaknesses are identified. Among strengths of the project includes sufficient resources to do the project. There have been many research been done before on this area and it helps in making the project more precise. Plus, the tools needed are sufficient and assisted by group of experts in both technology and power domain. However, the main concern in this project is the energy reading, where the small changes in the reading that is almost undistinguishable between energy used by different set of queries since it is being conducted in a small scale of database.

Based on the result gathered and captured on all the five 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 Recommendation

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] Kurp, P. (2008). Green computing. Communications of the ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1400186

[2] Harmon, R. (2009). Sustainable IT services: Assessing the impact of green computing practices. Management of Engineering &, 1707–1717. doi:10.1109/PICMET.2009.5261969

[3] Essary, D. (2008). Predictive data grouping: Defining the bounds of energy and latency reduction through predictive data grouping and replication. ACM Transactions on Storage (TOS), 4(1), 1–23. Retrieved from http://dl.acm.org/citation.cfm?id=1353454

[4] Wang, D., & Ph, D. (2008). Meeting Green Computing Challenges, (858), 121–126.

[5] Pisharath, J., & Choudhary, A. (2004). Reducing energy consumption of queries in memory-resident database systems. for embedded systems, 35. doi:10.1145/1023833.1023840

[6] Stockinger, K., Berkeley, L., & Shoshani, A. (2002). Strategies for Processing ad hoc Queries on Large Data, 72–79.