Data Processing of an Energy Management System Using the Google Cloud Platform

by

Dina Ghovind
17957

Dissertation submitted in partial fulfilment of
the requirement for the
Bachelor of Engineering (Hons)
(Electrical and Electronic)

JANUARY 2017

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

Data Processing of an Energy Management System Using the Google Cloud Platform

by

Dina Ghovind
17957

A project dissertation submitted to the
Electrical and Electronic Engineering Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirement for the
Bachelor of Engineering (Hons)
(Electrical and Electronic)

Approved by,

____

(IR DR PERUMAL NALLAGOWNDEN)
Project Supervisor

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
JANUARY 2017

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.

DINA GHOVIND

ABSTRACT

The energy consumption of most industries has been growing year after year and this has called for an efficient method of managing energy wisely. However, endeavours to monitor and manage energy have been neglected due to the lack of ability of current energy management systems to acquire and process energy data accordingly. The advent of the Internet of Things, cloud computing capabilities as well has Big Data has brought about a new means of collecting and processing data; all acquired in real-time, subsequently making data analysis possible. This paper, as such presents an Efficient Energy Management System model which leverages on the technology of the Internet of Things (IoT) and cloud computing for developing a real-time system with data processing capabilities using the Google Cloud Platform. This paper also discusses the drawbacks of current energy management systems as well as the associated technologies and blueprint which are required to enable the development of a real-time energy management system. The data processing model will be examined and processed using the Dataflow Engine and the Storage Bucket in the Google Cloud Platform environment.

ACKNOWLEDGEMENTS

I would like to extend my utmost gratitude to my supervisor, Ir Dr. Perumal Nallagownden for his guidance and support throughout the completion of my Final Year Project. His experience being an engineer, coupled with the years he spent lecturing has enabled him to offer me advice as to the understanding of my project and in providing solutions to any problems I have faced during the course of this project.

At the same time, I would like to extend my sincerest gratitude to all the staff and lecturers of the Electrical and Electronic Engineering Department for their assistance and guidance during the course of this project. I would also like to thank Universiti Teknologi PETRONAS for guiding me and providing the resources which have helped me to complete this project.

Last but never the least, I would also like thank my parents for their continuous moral support and love and in the assistance they have provided in the course of completing this project. Their loving words and encouragement motivated me to pursue this project with excellence.

Thank you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

Over the last few years, developing an effective Energy Management System (EMS) has been a prime focus in the industry. This was to tackle the impending issue of high energy consumption and to move towards a more automated approach of managing energy. The Internet of Things (IoT) has also become an omnipresent term, being an automated technology thus making the connected device more independent which leads to energy efficiency according to Zanella [2]. Already an expansion of the concept of "anytime" and "anyplace" which connects to "anything", IoT has become a technological revolution in both the technology and engineering sector.

Essentially, the need for proper energy management has to be developed without compromising on production. The advent of the Internet of Things (IoT) has enabled us to utilise tools such as cloud computing to be more energy efficient in handling energy loads from devices [3]. From sensors to microcontrollers, these technological marvels have allowed us to work simultaneously with online systems to tackle energy efficiency and reintroduce a more integrated Energy Management System (EMS). The Internet of Things (IoT), being a novel idea in energy management, has a huge potential in solving the issue of high energy consumption.

From home automation to industrial application, the Internet of Things (IoT), will allow for a seamless connection between physical devices across an active network. In short, this tool allows for a real-time data collection, a data analysis medium and a platform for us to leverage on. We would in turn achieve better results with an increase in energy efficiency and productivity by having real-time access to relevant information according to industrial needs. Such data can then be processed to even make real-time decisions and predictive analysis models to prevent failure or breakdown.

This study will as such focus on presenting an Energy Management System (EMS) which will be an integrated system of smaller sub-systems. The sub-systems or modules are divided into the input system, cloud computing and the data analytics component which all make up the entirety of the Energy Management System (EMS) as a whole. The cloud computing system, particularly the data processing mechanism will be demonstrated in the Google Cloud Platforms.

1.2 Problem Statement

Current Energy Management System (EMS) models are either very costly or proprietary. There is no cost-effective system or platform for industrial use and this has caused industries to either neglect or focus on reducing costs in other sectors of their operations. With strict regulations from energy suppliers, organisations now are left with no choice but to put in time and money to research and develop energy management systems. This only gives more reason for the conduction of this study as it addresses a key concern among industry giants and at the same time, presents methods to develop a cost-effective system.

Furthermore, present energy management systems function solely as an energy monitoring system, without providing means of reducing energy consumption. The advent of the Internet of Things (IoT) has enabled us leverage on its smart data collection capabilities, harnessing the availability of cloud computing as well. The unavailability of this technology in energy management systems has in turn caused mass usage of energy. Thus, inefficient management of energy has called for new technologies like IoT to be implemented alongside an energy management system for better energy efficiency.

Leveraging on the technology of the Internet of Things (IoT), industries will be empowered to manage their energy loads more efficiently to prevent energy loss. In addition to that, an Energy Management System (EMS) would enable industries to use the capabilities of the Internet of Things to be more energy efficient in maintaining the power consumption of power plants.

1.3 Objectives

The objectives of the project are as follows:

a) To study the various data collection models and propose a suitable model using the Internet of Things (IoT) and cloud computing.
b) To develop a data processing model for an efficient Energy Management System (EMS) with the Google Cloud Platform.

1.4 Scope of Study

The scope of study is essential to define clear parameters from the objectives which will cover the spectrum of this research. As such, this research will cover the the layered architecture of the Internet of Things (IoT) which is the main component in the data input sub-system, cloud computing and the data analytics sub-system. The cloud computing subsystem will be thoroughly analysed, particularly in data processing mechanisms. Leveraging on the Google Cloud Platform, the cloud computing processes will utilise the Dataflow Engine and the Storage Bucket to process and store data.

This research will in the end, provide a model of data processing which will be showcased through the Google Cloud Platform environment using the Dataflow Engine.

CHAPTER 2

LITERATURE REVIEW

The huge usage of energy processes has caused immense energy consumption in production industries [1]. This in turn, has created an enormous potential for energy saving methods in industries. In an effort to not only showcase environmental responsibility, Energy Management Systems (EMS) are paramount in reducing production costs and increasing efficiency. Correspondingly, energy efficiency can only be achieved with improvement in productivity, meaning the optimization of production equipment by reducing energy consumption when not in use. The challenge is in ensuring energy efficiency is achieved while still maintaining production output.

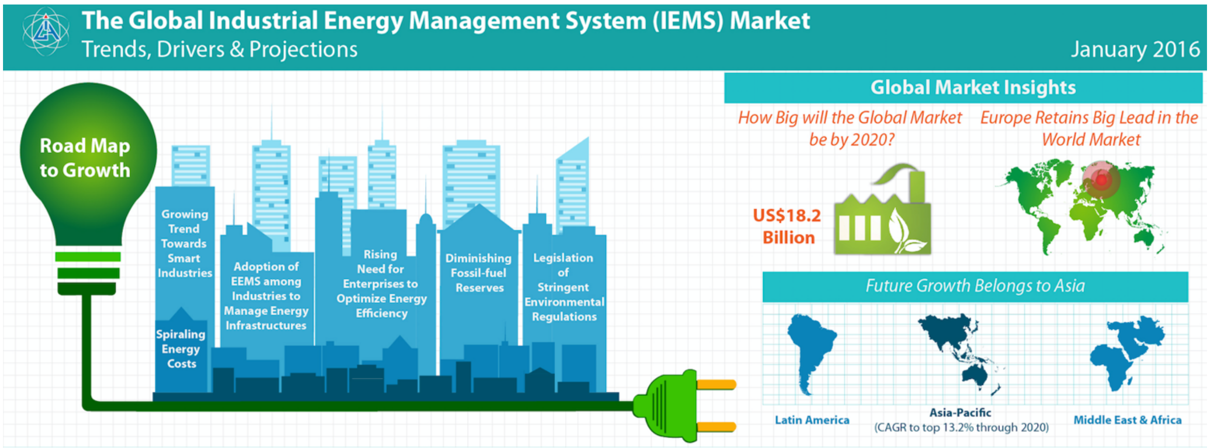

Figure 1: Road to Energy Management System Growth & Global Ems Market

The road map of the Energy Management System growth per Figure 1 illustrates the industry trend in adopting new energy management alternatives to be more cost-effective. In essence, spiralling energy costs was the main reason behind the development of energy management systems, thus leading to a growing trend towards smart industries. The soaring costs of energy consumption in industries has resulted in optimised energy efficient systems to not only automate the data collection process but also to build on a predictive analysis model which will result in energy efficiency. To top it all off, the introduction of legislation of stringent environmental regulations has caused organisations to re-think their strategies to be more energy efficient.

There are various barriers for energy efficiency in non-energy intensive and energy intensive industries respectively as discussed by Rodhin and Thollander [9] and Trianni et al [10]. There required level of control in energy usage is also absent as suggested by Garetti and Taisch [11]. Research into methods of estimating energy consumption patterns have been led to because of these shortcomings. Estimation of energy consumption are classified into two categories of forecasting models which are driving factors and historical data. The estimated energy consumption readings however, can be inaccurate and available at large time intervals. The use of an energy monitoring system is therefore essential to assess energy performance with improved accuracy in real-time.

In the industrial sector, energy savings has always called for much of the "policy" issue that estimates the magnitude of cost-effective management systems. The implementation of all the efficiency improvements that can be cost justified has been asserted from one side is that industry is "economically efficient". Energy-efficient and cost-justified opportunities that significantly documents engineering analyses are found when focused on the other side. Both statements will be proven correct depending on how the question is framed. More complex issues can be lead to raising questions of how to evaluate energy efficiency within the industrial economic process. There is also a doubt whether this is this particular question is truly what policy makers really intend to question.

There is an extensive literature on the "technical" potential for industrial energy efficiency improvements. Both analytical reviews and documented case studies showcases potential while at the same time there are costs for improvements on the extensive literature. However, there are conclusions that are less than satisfactory in regard to efficiency improvements that are actually implemented but not in regard to a systematic review of those studies.

The manufacturing processes required to produce the various products is directly involved in much of the energy consumed by the industry. Electricity accounts for about a third of the primary energy consumed by the industries while oil and gas resources such as coal, petroleum and natural gas takes up about 7, 26 and 28 percent respectively Traditionally thought of energy sources for non-fuel purposes are also made use by industries.

Crucial strategies of the manufacturing companies today are both energy efficiency and lean manufacturing. Lean manufacturing is the way of placing the entire value stream for specific products relentlessly in the foreground and rethinking every aspect of jobs, careers, functions, and firms [1]. One of the most widely used tools in lean manufacturing is value stream maps. It allows the observation of the flow of material and information as a product or service as it makes it way through the value chain. The state value stream maps set for the future are prepared to support new plans and strategies when the current state is displayed in the value stream map.

In terms of energy growth of a country, Europe dominates the world market as of 2016. However, it is projected that the future energy growth belongs to Asia, with Latin America and the Middle East and Africa not far behind. The current energy market is booming and is huge due to interests from big industries. Rising energy consumption has deemed it necessary to find a viable solution at the crux of the matter, which is energy itself. By 2020, it has bee projected that the global energy market will be US$ 18.2 billion. Raking in over eighteen billion dollars in an industry which is not at its peak, makes research at this stage extremely crucial to present an energy management system which is independent and automated.

Most organisations forgo energy saving methods because often times, these systems can prove to be very expensive and complicated to manage. The additional manpower needed to operate these systems and the costs to install new smart meters and devices have caused organisations to simply neglect energy efficiency [7]. In recent times however, energy suppliers have come down hard on organisations who do not implement any energy saving methods or systems. Some organisations are even penalised for their lack of adherence to newly imposed regulations. The upside behind these regulations are organisations are beginning to rethink their approach to energy savings and are starting to implement measures to address these issues with the introduction of cost-effective Energy Management Systems (EMS) [2].

Computing is starting to be transformed to models that consists of services where that has delivery and commodity services that is in a way slightly similar to traditional utilities such as water, electricity, gas, and telephony. Users access their services based on what they require not in regard to where those services are hosted or delivered. Most recently cloud computing, grid computing and cluster computing has shown promise to showcase this utility computing vision of several computing paradigms.

Presently, it is generally common to independently access various content across the internet with no reference to the underlying hosting infrastructure. Content providers monitor and maintain around the clock this infrastructure which consists of data centres. Business applications have their capabilities exposed as sophisticated services that will be able to be accessed over a network is cloud computing which is an extension of this paradigm. For accessing these services, profits made by the charging consumers are given as incentives to the cloud service providers. To help reduce or eliminate all costs in regard with “in-house" provision for these services, this opportunity serves to attract consumers. Since cloud applications is essential and crucial to core business operations, consumers need to be guaranteed by their providers on service delivery.

Various data centres for hosting Cloud computing applications in various locations across the world have been established by providers such as Google, Salesforce, Amazon, IBM, Microsoft, and Sun Microsystems to prove to be redundant and also to have consistent reliability in case the site collapses. While keeping users isolated from the underlying infrastructure, service providers have to maintain and ensure their flexibility in their service delivery as the user requirements for their cloud services are sometimes varied. Due to recent progressions in software and microprocessor technology, commodity hardware to run and use applications with Virtual Machines

(VMs) has been increasing on an efficient scale. To suit the consumers, these Virtual Machines (VMs) allows the consumer to customize their platforms and also allows the isolation of the applications from the general underlying hardware. By allowing consumers to install their own applications, providers are allowed to expose their applications running within VMs and provide access to VMs as a service (e.g. Amazon Elastic Compute Cloud). The service providers are challenged with the VMs as there is a large demand from the users competing resources due to the intelligent allocation of the VMs.

To save time and money, service consumers with global operations distributed the workloads to multiple locations at the same time through various cloud that enabled a faster response time. Also, to enable dynamic interconnection and cloud provisioning from multiple domains within various and multiple enterprises, this creates the need for establishing a computing atmosphere. Creating clouds and cloud interconnections involve huge challenges [7].

As such, by leveraging on the technological advancements of the Internet of Things (IoT), energy efficiency can be achieved by a developing a comprehensive Energy Management System (EMS). The Internet of Things (IoT) enables us to connect anywhere via a network with the implementation of a physical device. The network is an active, with real-time exchange of information of different properties. The real-time transmission of data simply provides a means of access to the the physical world which allows for smart decision making, thus leading to an increase in energy efficiency [4]. Therefore, harnessing the power of the Internet of Things (IoT) and incorporating it within the Energy Management System (EMS) will result in an effective system which would solve the issue of high energy consumption.

Storage systems, supercomputers, data sources as well as specialized devices, property of various organizations are among a wide range of geographically distributed resources that is meant to solve large-scale resource-intensive problems in engineering, science, commerce, and grid computing [3] . Grid computing was first motivated when large-scale, resource (computational and data)-intensive scientific applications required more resources that a single computer may have been able to provide in an administrative domain that was driven by the electrical power grid's pervasiveness, ease of use, and reliability [10]. Since then, after the internet and the World Wide Web, grid computing is said to be revolutionary. This is based on its ability to create an impact for the 21st century generation as compared to the 20th century when the electric power was said to be revolutionary as well.

Peer-to-Peer (P2P) computing peer nodes (computers) is used to share content and information in a direct manner to another in a decentralized manner. Since there is no notion of clients or servers, all peer nodes are considered to be equal and simultaneously serve as the client and server in pure P2P computing. Next-generation data centres that are built on storage technologies and virtualized computers to deliver cloud computing that is said to be one of the latest paradigm [7] which shows promises of fully reliable services.

The cloud will be accessible to consumers wherever they are in the world and will be able to have access to applications and data on demand. The users are assured of the availability of the cloud infrastructure to be accessed anywhere at any time as the infrastructure is also very robust and said to be of high reliability, autonomic and scalable to support global access and have dynamic discovery are some of computing services which are essential. The required levels of service that are indexed through the Quality of Service (QoS) parameters, are noted within the SLAs establishment with the providers are indicated by consumers in particular. The paradigm that has shown a lot of promise to leverage and build on the developments from other paradigms is the recently emerged cloud computing paradigm.

Production and utility make up the energy consumed in manufacturing systems. The entire process and operation involved in the technicalities of making the actual goods using manufacturing facilities is directly related to consumption attributed to production. The infrastructure that indirectly supports the production process is due to the factors such as heat, electricity, and air.

Productivity improvement is correlated with energy efficiency for production. While reducing the task time also enhances the energy efficiency by decreasing the energy usage while producing the same amount of goods, the utilization rate of production equipment is improved by reducing unnecessary energy consumption during standby time. Energy per productive output ratio in the production system has been the focus of most of the studies that have considered the topic of energy efficiency [4]. Coordinated by information networks, the research on increasing production efficiency has accelerated through the development of intelligent manufacturing systems, known Smart Factory [5]. This is because the production of excellent value-added products is the purpose of the manufacturing system. Industry 4.0, presents a smart factory that supports more intelligent and agile manufacturing to meet the rapid changes in consumers' requirements which is recently trending in the industry. In view of cost effectiveness as well as social requirements, energy efficiency is one of the important objectives of the smart factory. In organizing a smart factory, green manufacturing is also an important factor to consider. When considering the energy efficiency problem in manufacturing, the security of the energy supply should be satisfied without reducing the productivity.

A wide range of environmental and sustainability issues including resource material selection, transportation, manufacturing process and pollution cover green manufacturing [6]. There have been some efforts to reduce the total energy consumption in the manufacturing process with the introduction of Industry 4.0 which makes energy efficiency issues for manufacturing more practical, and [11].

The main electricity bill components, which are charge and energy charge, should be considered to gain the green manufacturing achievement in terms of usage of electricity. Reducing energy consumption is an obvious issue because high demand in electricity leads to charge a large amount of social energy costs so it would be smart and well-meaning to reduce the charge said to be in demand. Among many, one of the common energy forms required in the energy industry is heat. Thus, combined heat pumps are mainly introduced along with electricity [9]. In achieving improved total energy efficiency, a research has long been conducted as its importance for the optimal concurrent operation of heating and electricity and vitality has been increasing gradually.

The high energy consumption density from the industry compared to residential and commercial energy consumption has caused a global warming pressure on industry to increase. The wide and garnered development and instalment of photovoltaic panels on large surface areas such as parking lots and factory roofs have been driven because of such trends. The amount of energy that has been reduced due to energy charges and CO2 emissions is due to renewable energy technologies [3]. However, the output characteristics that are the source of the unstable regional power have been difficult to predict because these resources have variations. This is because, mostly, the generation matches grid peak time and the reduction of demand charges may be due to photovoltaics, in no way can it guarantee the contribution. Essentially high demand charges can be determined by one day of cloudy weather. One of the potential solutions is electric energy storage (EES).

The high initial capital cost should be worked on and overcome by potential means. to meet industrial needs, the developed Energy Management System (EMS) has to be cost-effective as well. It should be cost viable in the long run because an expensive system will only deter organisations from using it. To prevent machine failures or breakdown, the system should be able to make smart decisions in real-time as well. While still delivering a fully functioning smart system, all these factors have to considered in developing an Energy Management System (EMS), which should address all concerns and issues.
The providers took charge with peak cutting as a step to be more green but that’s not the only purpose. It is also to reduce energy cost that the individual manufactures have to bear. This is all part of a development for greener technology. In most manufacturing industries, high demand charges are usually due to the consumption of high short-term energy. A most suitable solution for peak cutting is EES. The recent improvement of EES technology enables a fast response is an excellent compensation

for peak reduction. The adoption of EES as an energy shifting or demand response is slow because of the high cost of EES. However, in relative to the small amounts of EES, intermittent peak cutting can be implemented; thus enabling the acquisition of economic feasibility. There are many studies in this area especially in the appropriate sizing of the EES which is important approach.

There are many production-related factors that are considered by the manufacturing industries before executing their plans that mainly focuses on resource, facility, efficiency reliability during the production scheduling. The main concern isn't the planning process but however it's the manufacturing execution. It's whether the execution plan has an effect on the electricity usage patterns. With a focus on energy savings in the industry it is highly suggested and to combine or co-optimize manufacturing execution planning with an energy management plan [10, 11]. In the long term when managements consider time-varying electricity prices and execution planning, energy consumption considerations will be taken more seriously with respect to management.

The process efficiency is directly related the unit facility efficiency improvement although it has a very small effect on the impact range. However, the utility control and combined heat and generation operation is indirectly affecting the manufacturing process during which it has a large range of impact. The co-development of manufacturing and energy management is one of the optimum goals of green manufacturing into a single organic system.

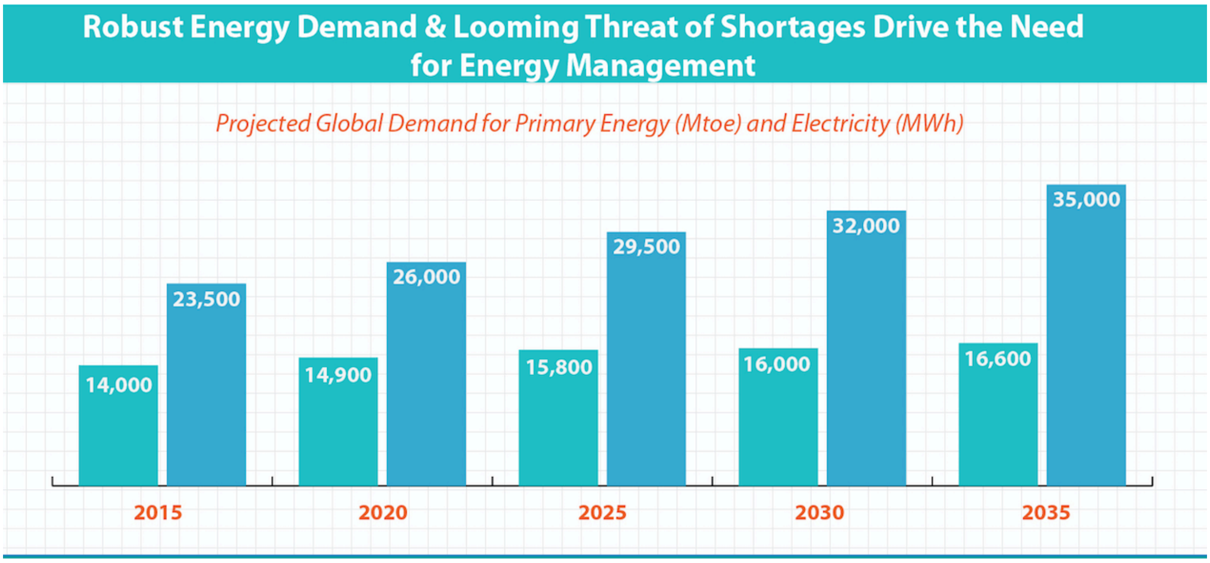

Figure 2: Projected Energy Global Demand for The Next Two Decades

Figure 2 depicts the projected energy global demand for the next two decades in terms of primary energy (Mtoe) and Electricity (MWh). Both mediums of energy show an exponential increase in consumption. The need for energy management stems from robust energy demand and the looming threat of an energy shortage. With strict legislations from environmental regulations, if the energy consumption of an industry exceeds the energy supplied, a fine will be imposed as well as additional costs for the extra energy supply. This energy shortage in the industry has called for a well-developed Energy Management System, which will result in better management of energy as well as better energy efficiency.

The proposed Energy Management System (EMS) will integrate the Internet of Things (IoT) with cloud computing to deliver data analytics on energy consumption. The layered structure of the IoT will comprise of four layers, which are primarily the hardware layer (edge technology layer), data layer (access gateway layer), middleware layer and application layer. The hardware layer can consist of any hardware element which provide information storage, control, actuation and information processing [6]. The data layer is responsible for data handling and is the first stage of message routing and communication. The middleware layer which is one of the most important layers, behaves as a data access layer between the application layer and the hardware layer. It operates like a central processing unit by handling all critical functions such as device management and data aggregation. Meanwhile, the application layer communicates the IoT device to the the common user. Figures 1 shows the architecture of the Internet of Things (IoT).

Each of these layers are crucial in communicating the IoT device to the cloud platform which are the servers. The sensors which are usually embedded systems or soft sensors communicate with the servers via an active Internet network to transmit data. The data is analysed for storage, predictive analysis and to make real-time decisions. All these processes take place in real-time leveraging on IoT capabilities. The stored data is then sent to the cloud computing environment to be cleaned and aggregated to generate daily or weekly energy consumption reports.

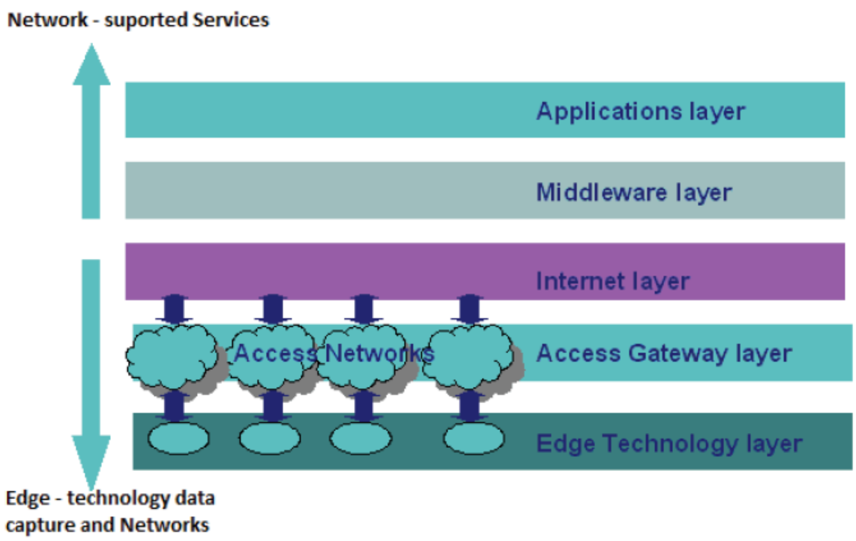

FIGURE 3: LAYERED ARCHITECTURE OF THE INTERNET OF THINGS (IoT)

The Internet of Things (IoT) paradigm allows for real-time data acquisition ad transfer of data from various sources. Due to security concerns related to transferring of critical data through the internet and sharing of corporate network for internet access, the wide scale adoption of IoT-based systems has been constrained. Having the ease of implementation over an existing infrastructure, the use if Wi-Fi for the communication network provides data transfer speed of about 10 Mbps.

While data processing capabilities are embedded deep within remote energy meters or smart devices, computers in energy management systems have been traditionally used to monitor and logging of data. The required processing power for data analysis is provided by a centralised system which is operated on demands while the energy meters as terminal units are utilised with minimum processing power. With respect to this, cloud computing plays an essential role in which uses a network of computers on the internet for processing the data.

Another essential component is data storage; particularly the medium in which it is stored. The Google Cloud Platform is one the most popular and sought after storage platforms for it extensive scalability and ease of use. Based on the analysis of acquired real-time data, the storage of data is the key component in the predicting failures in a system. This predictive analysis is usually done in the cloud computing environment. Big Data has enabled us to store and process raw data on remote servers, without any complexities. Cloud computing has enabled us to establish data servers with Big Data capabilities and all of this can be done in the cloud. Latency is also reduced accessing the data for analysis with the implementation of data storage and analysis on the cloud platform.

CHAPTER 3

METHODOLOGY

The design of a real-time energy management system (EMS) would enable better insights into the energy industries and the use of the Google Cloud Platform would enable a faster and more efficient data processing model. This would in turn lead to faster processing time and quick decision making capabilities for the user.

The research methodology of this project will be carried out as shown in the flowchart below. It briefly details the steps which will be carried out for the duration of the project and the research phases as well as the sub-system developments of the Energy Management System (EMS).

3.1 Project Methodology (Project Flow)

Figure 4 depicts the project methodology flowchart or the project flow to develop an Energy Management System (EMS). In the first stages of the project life-cycle, various research was done on current energy management systems. This was then followed by a detailed blueprint of the proposed energy management system. The subsystems were then broken down into individual components which are the input system, cloud computing system and data analytics. Once the blueprint was identified, the data processing of the cloud computing system was developed on the the Google Cloud Platform using the Dataflow Engine. This subsystem is developed utilising the capabilities of the Internet of Things (IoT) and cloud computing. This model of the energy management system is to showcase how data is collected, parsed, stored and processed and consequently generated to the end-user. The “job” is then run on the cloud platform to determine if it was successful or not. If it was not, the code is debugged and re-run until the desired results are achieved.

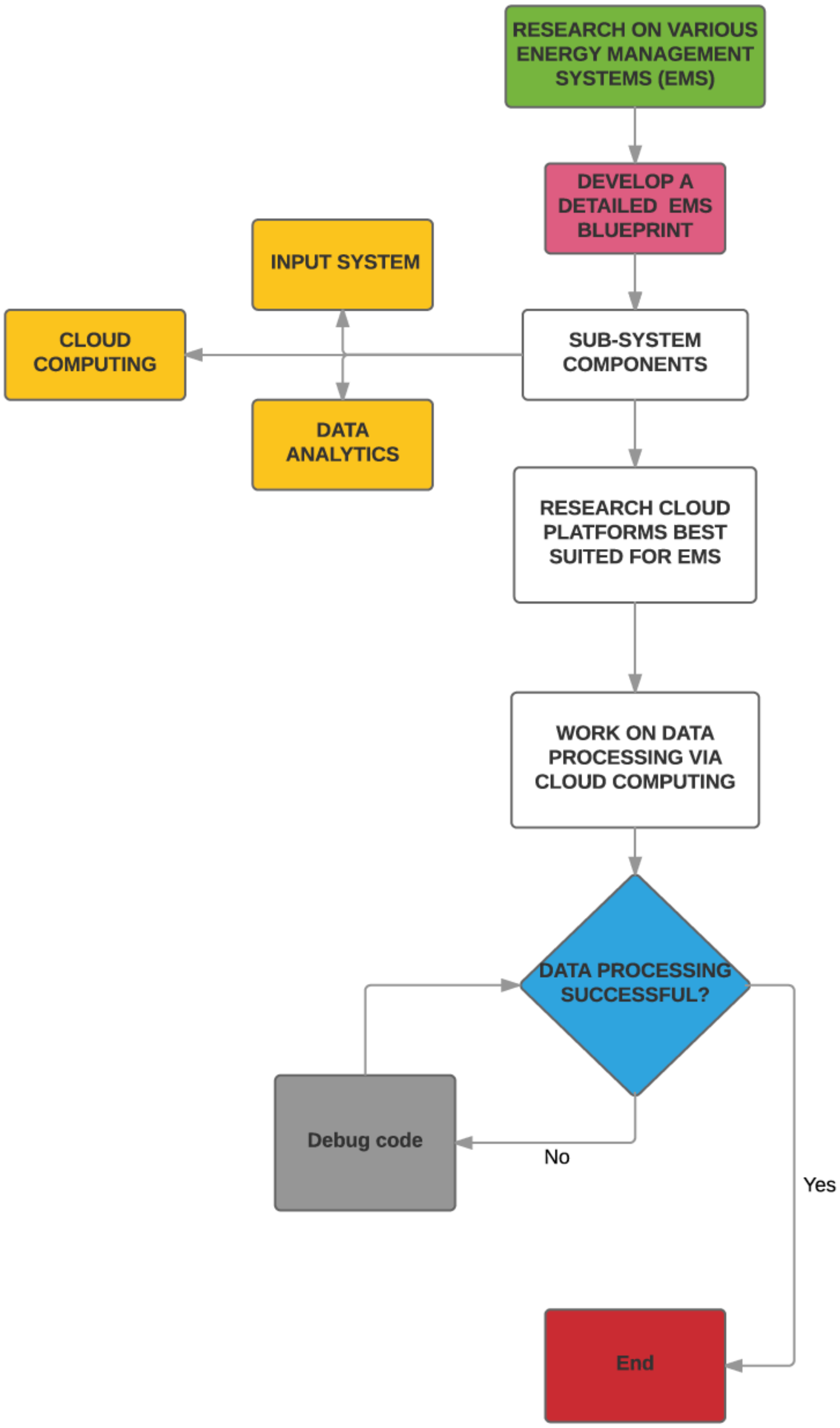

Figure 4: Flowchart of EMS Design

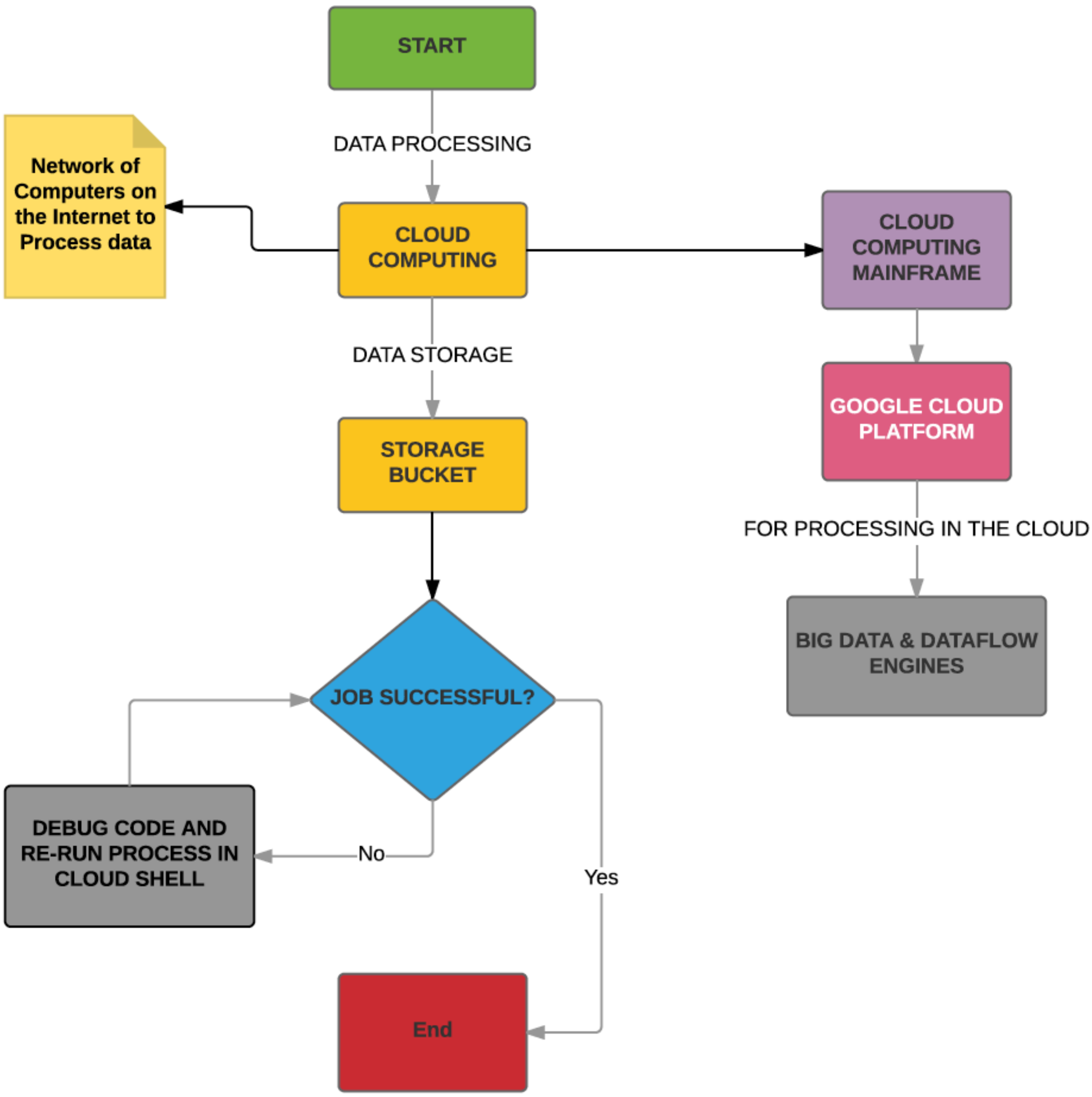

Figure 5: Flowchart of Project Methodology (Project Flow)

3.2 Ems System Design

Figure 5 depicts the flowchart of the energy management system detailing the subsystems and infrastructures. First of all, raw data is sent to the Google Cloud Platform to be stored. This data which is stored is not ready for the end-user as it has not been processed and requires the data to be filtered. This entire process of data collection and processing works off of the Internet of Things (IoT) paradigm.

Cloud computing is essentially a network of computers on the internet which processes data. The data is accessed via the SQL instance in the Google Cloud Platform. Using the Maven architecture, which is a new architecture of processing data clusters, the data is subsequently run through a “job” utilising the Dataflow Engine. Built on the Google Compute Engine, Dataflow is a unified programming model and managed service for developing and executing a wide range of data processing capabilities using batch execution of pipelines.

In the Dataflow Engine, the Google Cloud Shell is used to interact with the cloud platform. It is simply a command based terminal to configure and input instructions to run processes within the Google Cloud Platform. Using this model, various energy data is run through the Dataflow engine (pipelining process) and outputs the data accordingly to the user. Once the data has been through the cloud computing mainframe, it is then stored in the Storage Bucket mainframe within the Google Cloud Platform.

Then, the run “job” status will determine if it was successful or not. In the event that the status shows that it is “unsuccessful”, the Java code will be debugged for errors or any other problems in the functions. The “job” is then run again to ensure it is error free and successful.

3.3 Structured Project Plan- Key Milestones

The Final Year Project should achieve certain milestones to meet the required objectives which are as follows:

- Development of an Energy Management System (EMS) Design
- Working on Dataflow Engine of the Google Cloud Platform
- Developing the Data Processing Model on the Google Cloud Platform in Java
- Working on the Maven Architecture within the Dataflow Engine
- Running the scripts on the Google Cloud Shell.
- Testing of the Run Job and its corresponding results.

Table 1: Key Milestones

NO.	TASKS	DURATION	ESTIMATED START DATE	ESTIMATED END DATE
1.	Development of Data Processing Model on Google Cloud Platform	84 Days	24/01/2017	07/03/2017
2.	Progress Report Submission	5 Days	20/03/2017	24/03/2017
3.	Pre-SEDEX Preparation	7 Days	20/03/2017	22/03/2017
4.	Pre-SEDEX Presentation	1 Day	-	22/03/2017
5.	Final Report (Draft)	1 Day	-	10/04/2017
6.	Final Report Submission	1 Day	-	17/04/2017
7.	Technical Paper Submission	1 Day	-	17/04/2017
8.	FYP 2 VIVA	1 Day	-	25/04/2017

CHAPTER 4

RESULTS & DISCUSSION

4.1 Subsystem Design

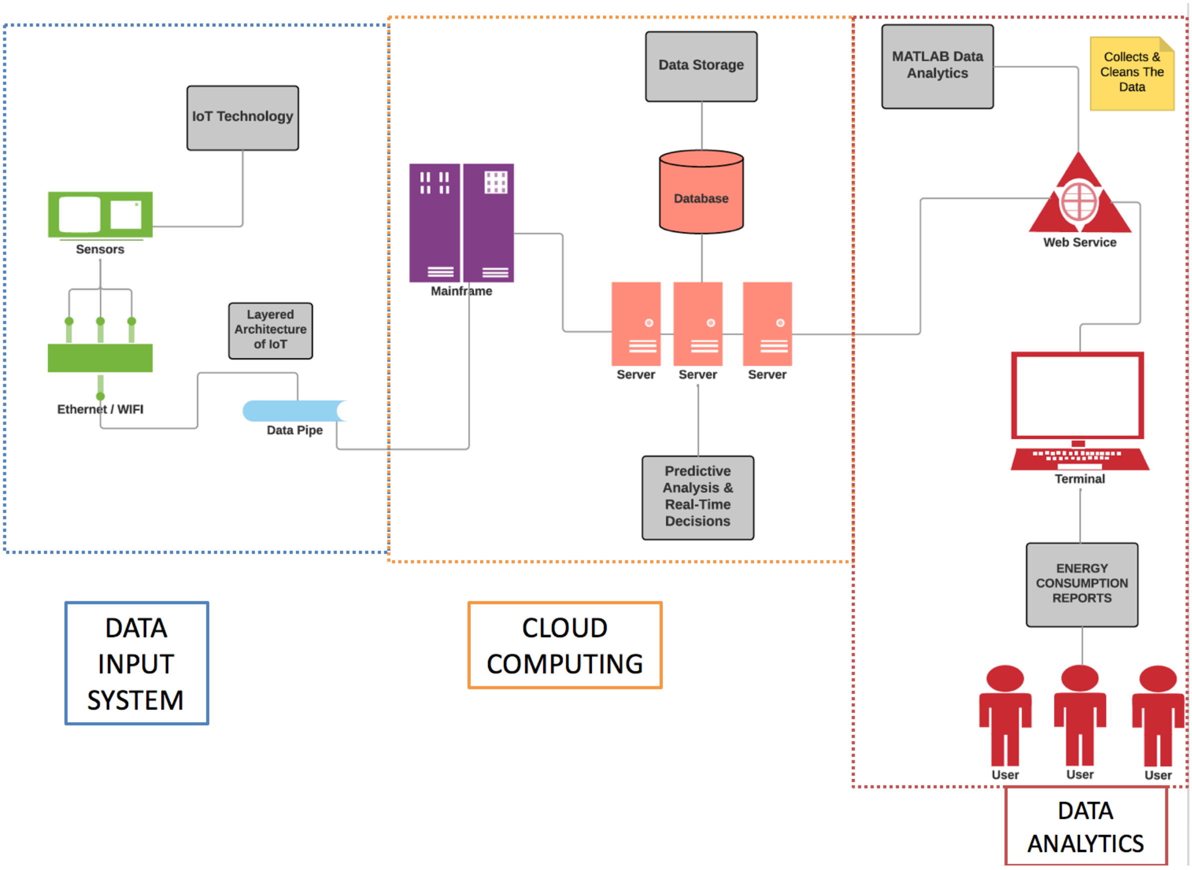

FIGURE 6: SUBSYTEM DESIGN OF EMS

Figure 6 shows the blueprint of the Energy Management System (EMS) which comprises of three individual subsystems. These subsystems are the data input system, cloud computing mainframe and data analytics system. These systems are developed individually on separate platforms but work hand-in-hand and make up the entire energy management system.

For this project, the focus would be on cloud computing, particularly the data processing model within the Google Cloud Platform.

4.2 Cloud Computing

Figure 7: Google Cloud Platform

The Google Cloud Platform is utilised for data processing of the energy management system for its speed and scalability capabilities. The Dataflow Engine as well as the Storage Bucket is used to process data and store it accordingly. Using the SQL instance in the Google Cloud Platform, raw data is uploaded onto the cloud using the cloud computing paradigm. A Storage Bucket is created on the cloud with the energy values.

The data is then processed in the Dataflow Engine within the Google Cloud Platform. The raw data is processed through scripts in the Google Cloud Shell and outputs a txt. l form of data representation. This data shows the energy values and average energy values throughout a 24-hour period. The Java code Figure 11 below shows the pipelining process of storing the data onto the Dataflow engine in the Google Cloud platform. The pipelining process allows for a seamless flow of data from the SQL engine to the Dataflow engine using batch executions for a fast processing of data.

```
energy-data-162109 ×  +
Welcome to Cloud Shell! Type "help" to get started.
dinaghovind@energy-data-162109:~$ gcloud auth list
Credentialed Accounts:
 - dinaghovind@gmail.com ACTIVE
To set the active account, run:
    $ gcloud config set account `ACCOUNT`

dinaghovind@energy-data-162109:~$ gcloud config list project
Your active configuration is: [cloudshell-15484]

[core]
project = energy-data-162109
dinaghovind@energy-data-162109:~$
```

Figure 8: Google Cloud Shell Script

Figure 9 shows the Google Cloud Shell in which all the data processing commands are keyed in. It essentially is the interface between the user the cloud platform and enables the user to manipulate the environment and commands per the user's requirement. Most of the commands to activate the console and run processes happen using this environment. It requires a working Internet connection for data uploading and data processing. The Maven architecture per Figure 9 is also activated via the Google Cloud Shell.

```
mvn archetype:generate \
    -DarchetypeArtifactId=google-cloud-dataflow-java-energy-data \
    -DarchetypeGroupId=com.google.cloud.dataflow \
    -DarchetypeVersion=1.9.0 \
    -DgroupId=com.data \
    -DartifactId=first-dataflow \
    -Dversion="0.1" \
    -DinteractiveMode=false \
    -Dpackage=com.data
```

Figure 9: Maven Architecture

The Maven architecture is a data management tool to handle big amounts of data. Maven makes the user's life easier by downloading the Java libraries from the Google App Engine SDK, thus not requiring the user to do it. Maven also builds the data clusters and therefore managed the entire project. It behaves like a brain which is responsible in ensuring the data is aggregated and processes accordingly per the Java Dataflow Code.

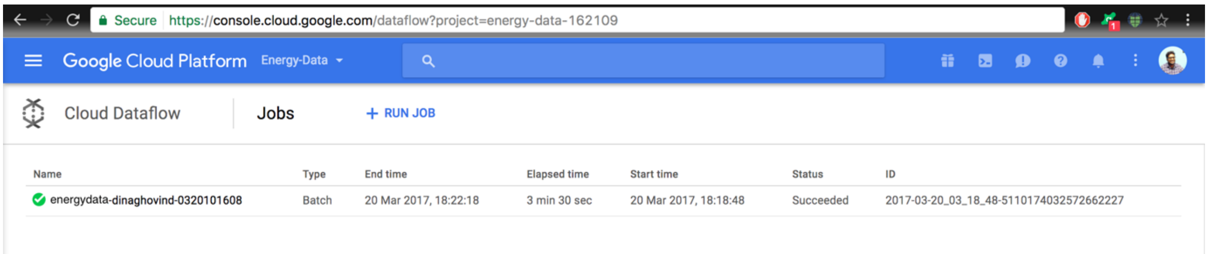

Figure 10: Dataflow Job Status

Figure 10 shows the dataflow job status which details the time taken, status and project ID of the project. The Google Cloud Platform console manages the active projects and all associated instances such as Storage, Stackdrivers and Big Data Engine. The Dataflow engine is part of the Big Data engine and is the most ideal engine in processing energy data using its unified programming model.

```java
package com.google.cloud.dataflow.examples;

import com.google.cloud.dataflow.sdk.Pipeline;
import com.google.cloud.dataflow.sdk.io.TextIO;
import com.google.cloud.dataflow.sdk.options.DataflowPipelineOptions;
import com.google.cloud.dataflow.sdk.options.PipelineOptionsFactory;
import com.google.cloud.dataflow.sdk.runners.BlockingDataflowPipelineRunner;
import com.google.cloud.dataflow.sdk.transforms.Count;
import com.google.cloud.dataflow.sdk.transforms.DoFn;
import com.google.cloud.dataflow.sdk.transforms.MapElements;
import com.google.cloud.dataflow.sdk.transforms.ParDo;
import com.google.cloud.dataflow.sdk.transforms.SimpleFunction;
import com.google.cloud.dataflow.sdk.values.KV;


public class EnergyData {

    public static void main(String[] args) {
        // Create a DataflowPipelineOptions object. This object lets us set various execution
        // options for our pipeline, such as the associated Cloud Platform project and the location
        // in Google Cloud Storage to stage files.
        DataflowPipelineOptions options = PipelineOptionsFactory.create()
        .as(DataflowPipelineOptions.class);
        options.setRunner(BlockingDataflowPipelineRunner.class);
        // CHANGE 1/3: Your project ID is required in order to run your pipeline on the Google Cloud.
        options.setProject("energydata");
        // CHANGE 2/3: Your Google Cloud Storage path is required for staging local files.
        options.setStagingLocation("gs://energy_kwh/energy");

        // Create the Pipeline object with the options we defined above.
        Pipeline p = Pipeline.create(options);

        // Apply the pipeline's transforms.

        p.apply(TextIO.Read.from("gs://dataflow-energydata/energy/*"))

        .apply(ParDo.named("ExtractReadings").of(new DoFn<String, int>() {
            @Override
            public void processElement(ProcessContext c) {
                for (int i : c.element().split("[^a-zA-Z']+")) {
                    if (!data.isEmpty()) {
                        c.output(data);
                    }
                }
            }
        }))
        .apply(Count.<int>perElement())
        // Apply a MapElements transform that formats our PCollection of word counts into a printable
        // string, suitable for writing to an output file.
        .apply("FormatResults", MapElements.via(new SimpleFunction<KV<int, i>, int>() {
            @Override
            public int apply(KV<int, i> input) {
                return input.getKey() + ": " + input.getValue();
            }


        }))
        // Concept #4: Apply a write transform, TextIO.Write, at the end of the pipeline.
        // TextIO.Write writes the contents of a PCollection (in this case, our PCollection of
        // formatted strings) to a series of text files in Google Cloud Storage.
        // CHANGE 3/3: The Google Cloud Storage path is required for outputting the results to.
        .apply(TextIO.Write.to("gs://energydata/energy"));

        // Run the pipeline.
        p.run();
    }
}
```

Figure 11: Dataflow Java Code

4.3 Data Output

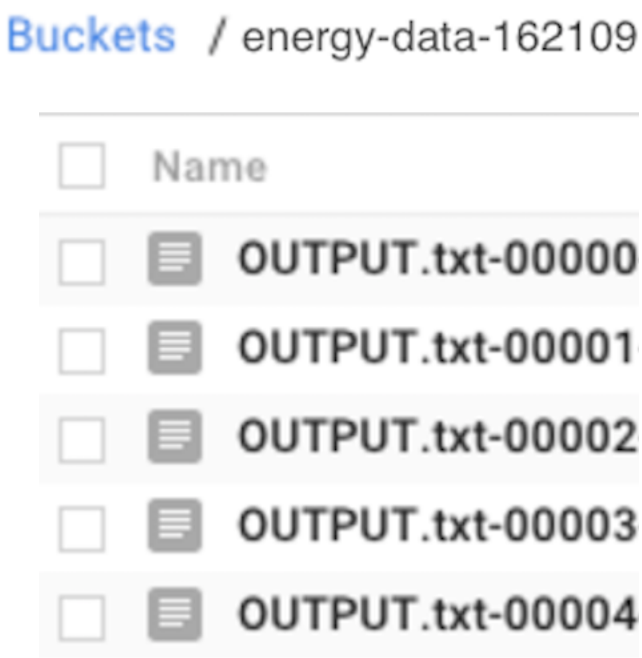

Figure 12: Data Processing Output

The data format in Figure 12 details the output energy data in the txt. format. It is split into several files because of the pipelining batch execution process which allows for a much faster processing mechanism. The energy values vary from KWh, power ratios, voltage values to heat, humidity and temperatures of machines in the factory plants.

The files are as small as 5KB which also means the compression algorithm behind the Google Cloud Platform is also great. The average processing time of the data on the cloud is roughly between four to seven minutes but this varies according to the strength of the Internet connection as well.

```
0:00    0.075,  0.18,   0.089,
1:00    0.082,  0.195,  0.105,
2:00    0.039,  0.085,  0.105,
3:00    0.085,  0.084,  0.068,
4:00    0.039,  0.102,  0.067,
5:00    0.059,  0.177,  0.066,
6:00    0.115,  0.233,  0.1,
7:00    0.213,  0.185,  0.11,
8:00    0.072,  0.19,   0.112,
9:00    0.038,  0.074,  0.108,
10:00   0.164,  0.042,  0.108,
11:00   0.098,  0.035,  0.124,
12:00   0.112,  0.024,  0.073,
13:00   0.131,  0.055,  0.081,
14:00   0.106,  0.024,  0.116,
15:00   0.131,  0.035,  0.281,
16:00   0.132,  0.045,  0.214,
17:00   0.195,  0.022,  0.172,
18:00   0.173,  0.038,  0.455,
19:00   0.194,  0.029,  0.881,
20:00   0.127,  0.234,  0.295,
21:00   0.194,  0.187,  0.197,
22:00   0.11,   0.167,  0.172,
23:00   0.109,  0.205,  0.146,
24:00   0.162,  0.153,  0.141,
```

Figure 13: Output KWh Data in Txt. Format

The data format shown in Figure 13 details the time stamp per half an hour as shown in the first column, followed by the energy readings of three different machines in KWh. The data is then used by the user to make smart decisions as to which machines take up the most power. By knowing this, the user can then manage the energy efficiency of the power plant much better. Likewise, other data values such as humidity, power ratio, KW, voltage levels and temperature readings can be processed in the similar manner.

4.3 Discussion

The Energy Management System (EMS) is broken down into several subsystems for easier development. As a whole, the Energy Management System brings about not just financial gain to businesses, but also a strategized carbon controlled and sustainable tool which solves issues related to pollution and climate change. The processed data which will be available to the user instantaneously will allow the user to make smart decisions as to which machines take up the most power and and therefore make calculated decisions in saving costs and maximising energy efficiency.

The web-based nature of the Energy Management System allows for the user to access the data anywhere. Apart from being user-friendly, the system supplies data which allows the user to make better decisions in terms of energy usage. The flexibility of the system also allows the system to be scaled up or scaled down depending on the size of the business.

The Google Cloud Platform's processing speed using its pipelining features allows data to be made available to the user in about five minutes. It will greatly supplement the Energy Management System in serving at its cloud computing base for data processing.

CHAPTER 5

CONCLUSION & RECOMMENDATIONS

5.1 Conclusion

This paper presented a new method of dealing with energy efficiency with the implementation of an Energy Management System (EMS) with the utilisation of the Internet of Things (IoT). It also presented the methods of developing the system as a whole via the sub-systems and processes involved within the Energy Management System (EMS). The cloud computing subsystem was developed using the Google Cloud Platform for its scalability and speed factors in data processing. The Dataflow Engine was used as the data processing engine and the Storage Bucket was used to store the output data. The data processing mechanism was fast because of the capabilities of the Maven architecture and the pipelining execution in the Dataflow Engine.

The Energy Management System (EMS) is aided with various technologies such as IoT which includes the concept of Big Data and Storage Bucket, all of which makes energy data processing in real-time an actual reality. The data processing in the Google Cloud Platform is much faster and allows the Energy Management System (EMS) to be scaled according to the user’s needs.

5.1 Recommendations

The availability of real-time energy data to be processed automatically on the cloud computing platform for every hour to save energy and towards proactive maintenance is possible. In the near future, the availability of real-time energy consumption data will pave the way for the development of systems for predictive analysis based on Artificial Intelligence and Machine learning. This will further increase energy efficiency to benefit industries in terms of proactive and preventive maintenance. Deep learning algorithms that will make use of predictive analysis of real-time data will pave the path for ultimate energy efficiency within a system.

REFERENCES

[1] A. Beloglazov *et al*., "Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing", Future Generation Computer Systems, 28(5), 755-768, 2012.

[2] A. Zanella et al., "Internet of Things for Smart Cities", IEEE Internet of Things, 2014.

[3] F. Satoh et al., "Total Energy Management System for Cloud Computing", 233-240, 2013.

[4] L. Skorin-Kapov et al., "Energy Efficient and Quality-Driven Continuous Sensor Management for Mobile IoT Applications", 2014.

[5] M. Condoluci *et al*., "Enabling the IoT Machine Age with 5G: Machine-Type Multicast Services for Innovative Real-Time Applications", IEEE Access, 2016.

[6] M. Primorac et al., "Microgrid Energy Management System in a Public Building", IEEE Xplore, 2016.

[7] M. Stamenić et al., "Energy Management System for Energy Efficiency Improvement in the Industrial Sector of the Republic of Serbia", IEEE Xplore, 2016.

[8] S. Ezhilarasan et al., "Design and Development of Energy Management System for DG Source Allocation in a Micro Grid with Energy Storage System", Indian Journal of Science and Technology, Vol 8(13), 2015.

[9] Rohdin, P., & Thollander, P. (2006). Barriers to and driving forces for energy efficiency in the non-energy intensive manufacturing industry in Sweden. Energy, 31(12), 1836-1844.

[10] Trianni, A., Cagno, E., Thollander, P., & Blacklund, S. (2013). Barriers to industrial energy efficiency in foundries.: a European comparison. Journal of Cleaner Production, 40, 161-176.

[11] Garetti, M., & Taisch, M. (2012). Sustainable manufacturing trends and research challenges. Production Planning & Control, 23(2-3), 83-104.

APPENDIX I: PROJECT GANTT CHART

Task \ Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
DEVELOPMENT OF DATA PROCESSING MODEL ON GOOGLE CLOUD PLATFORM	■	■	■	■	■	■	■	■	■	■	■	■			
PROGRESS REPORT SUBMISSION								■							
PRE-SEDEX PREPARATION PRESENTATION									■						
PRE-SEDEX POSTER PRESENTATION										■					
FINAL REPORT (DRAFT)													■		
FINAL REPORT SUBMISSION														■	
TECHNICAL PAPER														■	
FYP2 VIVA															■