Intelligent Cloud Based Decision Support Framework for Smart Grid

Walid Mohamed Aly
College of Computing and Information Technology
Arab Academy for Science, Technology & Maritime
Alexandria, Egypt
Email: walid.ali [AT] aast.edu

Mohamed Kholief
College of Computing and Information Technology
Arab Academy for Science, Technology & Maritime
Alexandria, Egypt

Eslam Adel Kamel
College of Computing and Information Technology
Arab Academy for Science, Technology & Maritime
Alexandria, Egypt

Abstract **—— a smart grid is an electricity grid that uses digital information and communication technology, it allows grid operators to monitor, analyze, control and manage a network so that more informed and efficient decisions can be adopted. A smart grid system typically includes efficient database management which allow data storage, transaction processing and querying. This paper introduces an Intelligent Cloud Based Decision Support framework for Smart Grid, this framework deals with the subject of a cloud-based Smart Grid for Big-Data management and analysis, and it identifies trends and forecasts demand.**

The proposed formation of smart grid will deal with Big Data set which will include, the power usage patterns of customers, weather data, the current demand and supply details. The grid will operate on the data being fetched from the cloud storage. This paper also focuses on the important role of NoSQL systems in solving challenges the progress of smart grid .

A case study is included to present the efficiency of the proposed framework.

Keywords -Smart Grid, Cloud Computing, Hadoop, Big Data, MongoDB, Predictive analytics

I. INTRODUCTION

Electric power systems have developed through the years, and currently reached a new system called Smart Grid. The development of this trend is regarded as the most significant technological Innovation of the electric power system in the 21st century. A Smart Grid is an electrical network which depends on digital technology that is used to provide electricity to consumers through two-way digital communication [1].

This system is provided to be used in monitoring, analysis, control and communication within the supply chain to improve efficiency, as well as reducing consumption of energy and cost.

The smart grid system is neither a mere smart meter nor a digital meter. The Smart Grid is a highly intelligent monitoring system that contains several sensing devices.

The Smart Grid uses an intelligent monitoring system tocontinuously track all the digital sensing devices to obtain power usage data. Furthermore, it may work with renewable energy systems and power generation systems.

The purpose of the use of Smart Grid is to overcome the short comings of the traditional electrical networks which use smart meters, but need advanced analytics over millions of data streams for decisions support and forecasts demand [2]. The databases are also increasing, and growing tremendously. As for the historic database, they are helpful in the future analysis and can be taken as a factor for forecasts demand and decision making [3]. Big Data is a high volume, high velocity and high variety information asset that demand cost-effective, innovative forms of information processing to reach an enhanced insight and decision making.[4] This is a Big Data challenge that requires advanced informatics techniques and infrastructure. Energy uses events streaming from millions of smart meters, sampled at 15 minutes, which need to be collected and correlated with a consumer's historical profile [5]. Data mining and pattern matching are two necessary techniques to Detect critical situations,

And to correct them with low latency to ensure Grid stability. Analytical and computational models can help predict the power supply and its demand within a service

area to take preemptive actions to curtail this demand by notifying and incentivizing consumers[5].

Smart Grid produces massive amounts of data which need to be managed and processed. However, running analytics on this data using traditional data warehouses is very complex, therefore traditional data warehouses are not capable of managing and processing big data in a fast and effective way. Distributed file system framework like Map-Reduce system can process big data quickly, which can be readily used for Analytics purpose [6].

Nowadays, one of many important trends in information management is the management of tasks through cloud computing, which has been widely regarded as the next-generation of computing and storage. The concept of cloud computing is based on an enormous data centers with massive computation and Storage capacities operated by cloud providers, which makes computing and storage services indispensable utilities. Managing information through Cloud computing, and using the Smart Grid simultaneously makes sense for the following four reasons [7]:

First, highly scalable computing and storage services, provided by cloud providers, match with the requirement of the information processing in the smart grid.

Second, the level of information integration in the smart grid can be effectively enhanced by leveraging cloud information sharing.

Third, due to the complication of the Smart Grid, it may lead to a highly complex information management system.

Fourth, outsourcing information management to the clouds allows these new players to focus and concentrate on their innovation rather than focusing on building data centers to achieve scalable goals.

The paper aims in giving a proposal to form a cloud-based Smart Grid for Big-Data management and analysis which will act as an intelligent system to identify trends and forecasts demand.

The remainder of this paper is organized as follows. Section 2 we review the Cloud computing in the field of cloud data management and Hadoop implementations and non-relational databases, while Section 3 presents related work in the field of big data analytics and cloud data management, we describe the system model in Section 4 the proposed intelligent framework for a cloud-based for smart grid. Section 5 the case study that is used to demonstrate the proposed framework is presented, Finally Section 6 concludes the paper.

II. BACKGROUND

A. *Cloud computing*

Cloud computing is an emerging paradigm, with a growing popularity and it's also being widely adopted. Cloud providers hold shared servers, and deliver computing, storage, processing data, and software to end consumers as a service.

Performance, resiliency, and power is noted to be a key requirement. This includes provisioning, scheduling, monitoring, controlling of failure management, performance optimization, and energy management. Traditional monitoring and managing of technologies were developed for enterprise environments, and are typically catered to meet the needs of single customer environments deployed in the range of thousands of servers, serving millions of users [9].

B. *NoSql*

Within the last few years, non-relational databases have dramatically risen in popularity. These databases are commonly and widely known as NoSQL databases, clearly marking them different from the traditional SQL databases. Most of these are based on storing simple key-value pairs on the premise that simplicity leads to speed.

Beyond simple values, but also have the ability to store objects. The objects are in some serialized form such as XML, JSON, and BSON (binary encoded JSON). [10]

NoSQL is a non-relational database management system, database provides a mechanism for store and process huge data sets, like Facebook, Google and Amazon.com, NoSQL databases based on the partition the data and store the partitions via different server nodes, these partitions can also be replicated in multiple servers so that the data is still available even in case of servers' failures. Use

NoSQL with cloud computing these characteristics make NoSQL data stores especially suitable for use as cloud data management systems. Horizontal scalability, readily achievable for NoSQL solutions, fits incredibly well with cloud computing.

C. *Hadoop*

The Apache Hadoop is an open-source software framework which helps in distributed storage and processing of large data sets. Hadoop can be scaled up from one server to thousands of servers with local computation. Hadoop uses the concept of Map Reduce. The Map Reduce framework consists of single master/slave node per cluster [3].

The proposed framework to form a cloud-based Smart Grid for Big-Data management and analysis which will act as an intelligent system to identify trends and forecasts demand via using Hadoopand database NoSQL.

Big data

Big Data is high volume, high velocity and high variety of data that demand cost-effective, storage or compute capacity for information processing and decision making.

D. Big Data Analytics

"Big Data Analytics is the process of examining large amounts of data of a variety of types (big data) to uncover Hidden patterns, unknown correlations and other useful information."[11].Since start of the increase in the volume of data, we have become a new term called "Big Data" are difficult to manage and analyzed via traditional data analysis methods. so we are now in need of some new technologies that can be applied to handling the big data like parallel processing, distributed file system, Hadoop and NOSQL Database, cloud computing.

III. RELATED WORK

Many researchers have contributed to the field of data mining in the smart grid. In this section, an overview on a few representative works is presented.

M.Mayilvaganan and M.Sajitha[3] gave a comprehensive case study about the smart grid the main purpose formation of smart grid will deal with Big Data set which will contain the data regarding the power usage patterns of customers, historic weather data of the location, the current demand and supply details. The paper has suggested architecture to big data analytics in the cloud, but this study need to implementation of the architecture with smart grids, storing, big data analytics to calculate the user demand and supply.

Kendall E. Nygard, Steve Bou Ghosn, Md. Minhaz and Prakash Ranganathan [2] designed a software for directly supporting decision model independence, in their research, they discussed the advantages of an agent-based framework as a methodology for fully automating the electrical grid. The concluded that by supporting decision models separate from the monitoring and action agents, alternative models can be easily evaluated. And that adjustable and situated agent autonomy adds further depth and power to the design.

Xi Fang, Dejun Yang, and GuoliangXue[2] studied the optimization problem of leveraging the cloud domain to reduce the cost of information management in the smart grid. They proposes a cloud-based smart grid information management and presented a cloud and network resource optimization framework to solve the cost reduction problem in cloud-based smart grid information storage.

the precedent studies depends on explaining and clarifying the advantages of using cloud computing with smart grid and using cloud database for big data analytics, but it lacks practical application on realistic data and especially the user model to predict the expected process . (Long term and short term).

IV. INTELLIGENT FRAMEWORK FOR SMART GRID

A. The Proposer Framework

The architecture consists four partitions Data source, Dataset, Cloud computing and demand forecast.

The data are stored in cloud based Mongo Database. To manage the storage and retrieval of the data the system uses the Hadoop.

The processing is done using the Map-Reduce. The smart grid in the proposed architecture uses a prediction algorithm 'regression model ' to predict the demand and supply.

The proposed architecture of implementing big data analytics will be occasion for the smart grid environment.

The proposed framework apply Regression Model to forecasting by machine learning approaches (Regression) acquire hidden knowledge by learning from a smart grid data set via use Hadoop in the cloud, The framework also the smart grid used cloud services with the big data analytics to efficiently balance the load in power production and supply.

The advantages cloud computing, big data analytics, and smart grid are together to produce an improved version from smart grid to perform the group from functionalities:

- Analysis of consumer behavior Patterns to forecast demand in advance.
- Analyze the weather data and forecasting the energy production.
- Balance the demand/supply.
- Decision support of switching between the high/low priority demands.
- Smart grid framework consists of four partitions Data source, Dataset, Cloud computing and demand forecast.

B. Data Source

Data Source are the sources which are responsible for collecting all the data which affect anything that has to do with prediction. Below are some examples:

- (Data weather) Seasons, weather forecast, Temperature, etc.
- Smart meter is one of the most important resources which is used to collect data and store it every 15 to 30 minutes to the grid.
- •Data about the kind of residence (the kind of consumption (house - hospital - company) address.
- Data on about the day (weekend -seasonal vacation - working day - half working day).

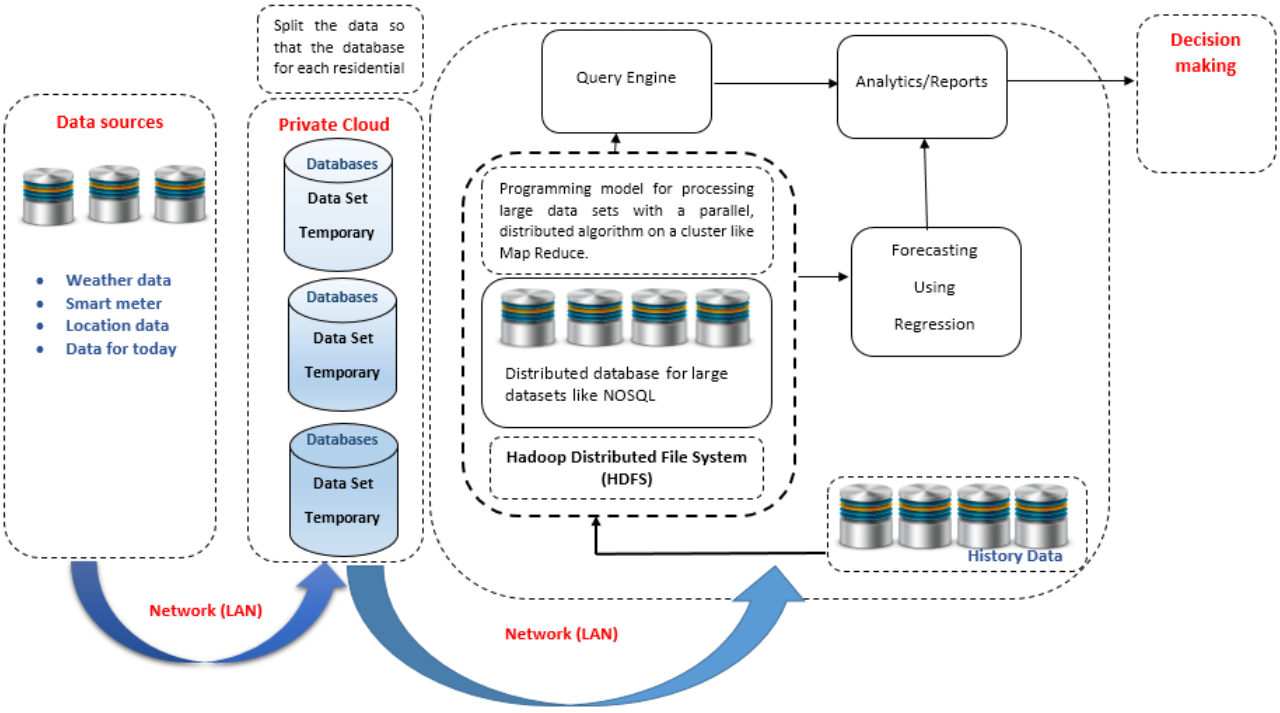
REFERENCE	TITLE OF PAPER	OBJECTIVES	PUBLISHER
[13]	"A cloud-based architecture for Big-Data Analytics in Smart Grid: A Proposal."	The paper has suggested architecture to implement big data analytics and cloud architecture in the smart grids to make them even better.	IEEE 2013
[14]	"Smart Grid Data Analytics for Decision Support."	This paper focuses on identifying the variables of interest that are important in the electric grid which will help decision support process for system operators. Use M5 model and J 48 decision tree machine learning technique is investigated using the real electric grid data.	IEEE 2011
[04]	"SOFT GRID - Big Data Analytics for Smart Grid."	This paper presents a survey on the future Smart Grid - Soft Grid; and how this technology addresses the Indian Power crisis.	VESIT2014
[15]	"Emphasizing Big Data Engineering and Software Architecture on Cloud Computing"	In this paper its availability the requirements to use in efficient way between the three concepts which is big data, efficient software architecture and clod computing, how to use big data on cloud computing.	IJCSIT 2015

TABLE CVII. SEVERAL RELATED STUDIES THAT DEAL WITH BIG DATA AND SMART GRID VIA THE USE OF CLOUD

FIG.107. FRAMEWORK BLOCK DIAG

C. Data Set Temporary

Split the data so that the database for each residential area separately use .It's a temporary data base, the storage of data is done through it from the smart meters and the sensor



machines, then sending it to the Cloud Data Base each 24 hours, connecting the network traffic and this reduces the problem Cloud Computing, In addition to all data set belonging to a particular area through which you can also predict the process for a particular area.

D. Big data

Cloud computing provides users to process distributed queries via multiple data sets and Return resultant sets in a timely manner, also provides the underlying engine through the use of Hadoop, shown in Fig.1 Big data are stored in the cloud distributed database and processed via a programing Model for large datasets with a parallel distributed algorithm in a cluster.

Map Reduce is an appropriate example of big data processing in a cloud environment, it allows for the processing of large amounts of data stored in parallel in the cluster.

V.CASE STUDY

The Electricity Data used in this case study is obtained from The Ministry of Electricity in Egypt.

This Case study implementation on a cloud-based architecture for Big Data Analytics in Smart Grid Using Hadoop, In Egypt is not found smart meter until now, so we will implementation forecasting on the Traditional data, Forecasting Energy Consumed for client in the first half year of 2015.

A. Data Set

The consumption data is collected from the ministry of electricity during the period from2009 to 2014 to form a data set known as (Data Smart Grid). Data smart grid includes 80Cases Study, each Case includes 72 record has12 attributes. Not all the attributes will be used in the data mining process, some of the attributes in the data set such as the Client ID, Client Name, Address, or Home Number present personal information that do not expand anyknowledge for the data set under processing. The selected attributes are shown in table 2.

Table 108. Dataset

Customers ID	SM_ID	Address	Date	Time	Temperature	Consumption	Cost	Type
ABN0001205	002154	Alexandria	22/5/2015	02.50	28	25	50	Home
ABN0001205	002154	Alexandria	22/5/2015	02.55	28.2	27	54	Home
10 Bytes	6 Bytes	10 Bytes	9 Bytes	5 Bytes	4 Bytes	2 Bytes	2 Bytes	2 Bytes

Customers ID	Name	Phone	Mobile	Email	Notes
ABN0001205	Eslam Adel	035404427	012345678	Eng.eslamAdel@Hotmail.com	

Storage capacity per client per year (In case found Smart Meter)

- One Row = 10+6+10+9+5+4+2+2+2 = 50 Bytes
- New row Every 5 seconds.
- 20 rows in one minute.
- 1200 rows in one hour. 20 row * 60 minutes
- 28800 rows in one day. 1200 row * 60 hours
- 864000 rows in one month. 28800 row * 30days
- 10368000 rows in one year. 864000 row * 12 months
- 10368000 Rows * 50 Bytes = 518400000 Bytes
- 518400000 Bytes / 1024 = 506250 KB
- 506250 KB / 1024 = 494.38 MB for One Client in one year.

We need to big servers of information processing and storage huge data size this is big cost.

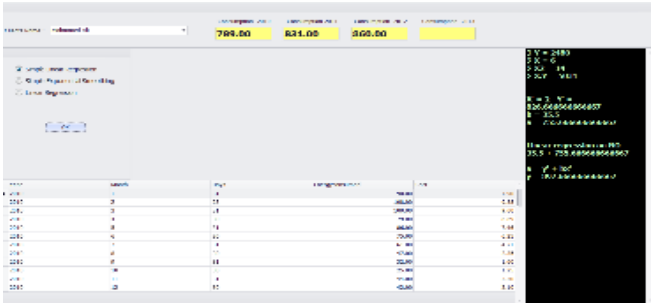
The best solution to this problem is the use of cloud technology.

Traditional data: is data collected from the Traditional meter one row in one month, 72 record about 72 months. Request forecasting Energy Consumed for client in the first half year of 2015.

No.	ClientID	Month	Count Days	UtilityType	Energyconsumed	Cost	years	Temperature
1	1	1	31	Home	80	7.9	2009	16
2	1	2	28	Home	98	9.88	2009	20
3	1	3	31	Home	90	9	2009	23
4	1	4	30	Home	70	6.69	2009	24
5	1	5	31	Home	86	7.46	2009	25
6	1	6	30	Home	75	6.25	2009	25
7	1	7	31	Home	64	4.71	2009	29
8	1	8	30	Home	48	2.35	2009	30
9	1	9	31	Home	35	1.6	2009	27
10	1	10	30	Home	26	1.25	2009	26
11	1	11	31	Home	45	2.2	2009	25
12	1	12	30	Home	42	2.1	2009	19
13	1	1	31	Home	90	7.9	2009	16
14	1	2	28	Home	108	9.88	2010	20

B. SGSoft for Forecasting Using Hadoop

SG-Soft is a new software tool designed for forecasting using by Hadoop, also it is scalable Software Platform for



forecasting using Cloud Technologies and The Apache

Hadoop. SG-Soft is a software tool written in C#.Net programming language developed and mongo database in cloud Environment and using Map Reduce.

C. Results for Forecasting Using Regression Model Reduce.

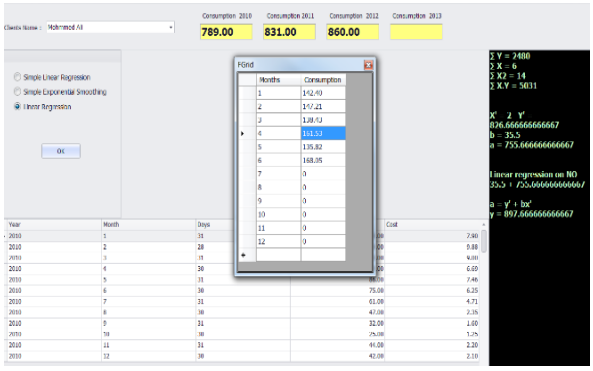
$$Y = a + b + e$$

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$a = y - bx$$

$$r = \frac{\sum xy}{\sqrt{\sum x^2 + \sum y^2}}$$

Forecasting energy consumed for one client in the first 6 months year of 2015.Comparison between real data and the resulting prediction



=== Summary ===

Correlation coefficient	**0.9654**
Mean absolute error	**5.2835**
Root mean squared error	**7.426**
Relative absolute error	**22.5362 %**
Root relative squared error	**26.0747 %**
Total Number of Instances	**72**

TABLE 109 RESULT SYSTEM

VI.CONCLUSION

In this paper, we have dealt with the efficient - Planning of power production and Hadoop Distributed File System. The implementation and usage of smart grids have increased now days. The paper has suggested architecture to implement big data analytics and cloud architecture in the smart grids to make them even better. The proposed

Real data	**RESULT PREDICTION**
140 KW	142.42 KW
140 KW	147.21 KW
141 KW	138.43 KW
183 KW	161.52 KW
120 KW	135.82 KW
160 KW	168.05 KW

Architecture is a proposal and the implementation of the smart grid is in process and forecasting.

REFERENCE

1. "Application Research of Power Supply Company Decision Support System under Smart Grid", 2010 China International Conference on Electricity Distribution.
2. "Managing Smart Grid Information in the Cloud Opportunities Model and Applications", IEEE 2012, ISSN 0890-8044.
3. "A cloud-based architecture for Big-Data
4. Analytics in Smart Grid: A Proposal" P.S.G College of Arts and Science, Coimbatore, India 2013.
5. "SOFT GRID - Big Data Analytics for Smart Grid" VESIT, International Technological Conference-2014 (I-TechCON), Jan. 03 – 04, 2014.
6. "Cloud-Based Software Platform for Big Data Analytics in Smart Grids", University of Southern California.
7. Managing Smart Grid Information in the Cloud:Opportunities, Model, and Applications.
8. Evolving Smart Grid Information Management Cloudward: A Cloud Optimization Perspective Xi Fang, Dejun Yang, and Guoliang Xue, Fellow, IEEE 2013
9. "Implementation of Smart Power Management and Service System on Cloud Computing" Department of Computer Science, Tunghai University, Taichung City 40704, Taiwan RO. 2012
10. "Cloud Management: Challenges and Opportunities" Tim Forell, Dejan Milojicic, Vanish Talwar Hewlett-Packard.
11. "A performance comparison of SQL and NoSQL databases" Yishan Li and Sathiamoorthy Manoharan Department of Computer Science University of Auckland New Zealand.2011.
12. BigData Analytics Definition ,http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics Util
13. "Emphasizing Big Data Engineering and Software Architecture on Cloud Computing", Mohammed Faisal et al, / (IJCSIT) International

Journal of Computer Science and Information Technologies, Vol. 6 (3) , 2015, 2852-2855 .

14. "A cloud-based architecture for Big-Data Analytics in Smart Grid: A Proposal ", M.Mayilvaganan, M.Sabitha Department of Computer Science, P.S.G College of Arts and Science, Coimbatore, India. mayilvaganan.psg@gmail.com,sabithamuthusamy.psg@gmail.com ,978-1-4799-1597-2/13/$31.00 ©2013 IEEE

15. "Smart Grid Data Analytics for Decision Support" ,Prakash Ranganathan, Kendall Nygard Department of Electrical Engineering, University of North Dakota, Grand Forks, ND Department of Computer Science, North Dakota State University, Fargo,ND ,2011 IEEE Electrical Power and Energy Conference .

16. "Emphasizing Big Data Engineering and Software Architecture on Cloud Computing".

17. Mohammed Faisal,Tamil Selvi Madeswaran, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (3) , 2015, 2852-2855 .