A Survey on Realtime Analytics Framework for Smart Grid Energy Management

K.Sornalakshmi[1], G.Vadivu[2]

Assistant Professor, Department of Information Technology, SRM University, Tamil Nadu, India[1]

Professor and Head, Department of Information Technology, SRM University, Tamil Nadu, India[2]

ABSTRACT: Smart grids are modernized electricity grids with information technology support. Smart Grids are the most promising development in the energy and utilities market.Smart grids are being installed in many countries and it is expected to have multi-fold benefits in efficient energy management. The Smart Grids receive real time meter data with high velocity and volume. In such scenario, near real time efficient analytics of streaming smart meter data and quick decision making is significant. In this paper, we survey the existing methodologies and means for real time energy data management in smart grids.

I.   INTRODUCTION

Smart grids are electronic grids that enable two way communications. Smart grids have many interconnected components [6] like sensors, Phasor measurement units (PMUs), smart meters, micro grids, Pluggable Hybrid Electric Vehicles (PHEVs) etc.Fig 1 shows the basic components of smart grid.All these components have to communicate with the grid. These components also monitor or generate data that has to be analysed for grid operations. Therefore smart grids are different from traditional grids in the sense that they include automation, telecommunication and information technology support.

Data available in smart grids are of various formats namely real time AMI (Advanced Metering Infrastructure) data, static offline consumer data, generation data etc. Such data has the dimensions of velocity, variety, volume and veracity of big data.Technologies like Automated Meter Reading (AMR) only allow a single way communication from meters to grid whereasAMI in smart grids also offer two way communications [4]. Notifications from grid on dynamic pricing and plans will be sent to customers. Distributed energy generated via solar panels and Pluggable Electric Vehicles (PEVs) should be fed into the grid. Such functionality not only requires latest hardware infrastructure, but also requires more efficient software framework that works in parallel.

Smart Grid uses AMI, which is an integrated system of smart meters, sensors, Intelligent Electronic Devices (IEDs) and data management systems. AMI requires smart meters to be installed on consumer premises for usage monitoring.Smart grids are now being deployed in many countries due to many advantages like remote control, energy savings and personalized consumption [4].Modern smart grid is a Cyber Physical System (CPS), which requires the state of the art computing infrastructure to support its operations.This paper surveys the real time data analytics platforms and technologies available for smart grids. The remaining section of this paper is organized as follows. First we discuss the current technologies used in smart grids, then the key challenges in implementing the software platform for smart grids.
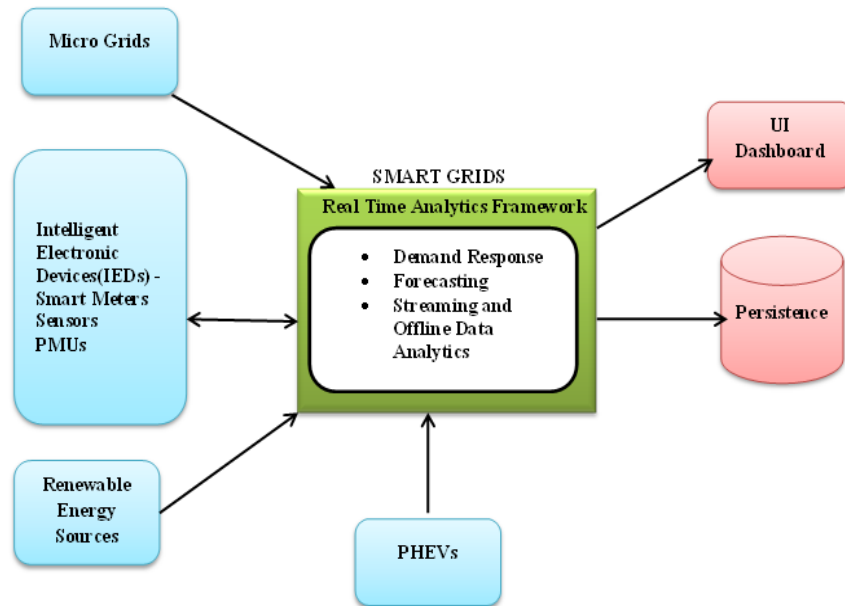
Figure 1 – Smart Grid

II.   TECHNOLOGY IN SMART GRIDS

A.   STREAMING DATA ANALYTICS

Smart meters send data once in every 15 to 30 minutes to the grid.To make the grid really smart and enable timely in place decisions that benefit the customers, traditional static processing of stored data is not efficient. Earlier predictions like planned power shut down ortime based pricingalone are not sufficient. Real time short term accurate prediction on consumption or outage is very much essential.Close to real time streaming data from the smart meters have to be immediately analysed and predictions done for supply demand mismatch, tariff recommendations etc.Such type of real time data is called as data in motion as compared to the traditional data at rest.

[1] addresses dynamic demand forecasting using the real time data from smart meters ingested using an efficient automated data flow pipeline, Floe into the software analytics platform on top of hybrid cloud.According to [4], there are two types of data generated in smart grids – event data and usage data. Both of these data are arriving as fast changing streams to the substation's control centre. Advanced analytics as compared to the normal statistics based prediction in conventional SCADA is needed in data driven decision making.

For efficient analytics, characteristics of streaming data have to be considered [5] like one pass for each record, no storage before processing, limited CPU and memory. Millions of reading per second would be reported at substation levels, which should be compared against variety of data like consumer profile, historical data, weather data and predictions made in time

Many distributed stream computing platforms are available for effective real time computation on big data streams [8] – like Yahoo S4 [15], Apache Storm [16], Apache Spark [9]. Spark introduces in memory partitions and computing, thereby reducing frequent hard disk reads and writes, which improves the response time which is the key characteristic of stream computing. Discretized Streams [11]are tuples of Resilient Distributed Data Sets(RDDs) [10], whichprocess streams as short, deterministic tasks which are also stateless. RDDs reconstruct themselves through lineage information, thereby achieving fault tolerance [12].

B.   FORECASTING

Load forecasting is not only based on streaming AMI data, but other data such as consumer profile, weather forecast, consumer locality and day of prediction say weekend or weekday or festivals/special days etc.In [1], demand forecasting is done using ARIMA (Auto Regressive Moving Average) and regression tree.Advantages of both the methods are discussed. ARIMA is good at following trends and the Regression tree model has low error rate. It has also concentrated on scalability for assembling abundant ensemble trees.Clustering consumers according to their profiles using k-means as consumer groups are useful when dynamic price recommendations and plans have to be proposed to the consumers [4].Prediction of renewable resources like wind is discussed in [5]. Such renewable energy generation prediction is useful since the load is also predicted for short term. Wind power generation is predicted based on wind speed, locality and direction of wind mill based on the hybrid approach, say physical and statistical approaches combined. Generation and the corresponding usage are forecasted in the given time horizon.[5] also discusses on how to decide on the number of historical points or the forecasting window to be fed into thealgorithm. Static window and dynamic window length where only similar recent historical points are considered are compared. Third option splits the duration to multiple windows, then performing prediction on each of these windows.

C.   CLOUD

Usage of cloud computing for smart grids is inevitable due to its characteristics like cost effective, reliability, scalability, availability, elasticity etc.[1] discusses the advantages of using private and public clouds. Public clouds are cost efficient whereas private clouds offer more security over sensitive data like customer usage patterns. Hence a hybrid cloud infrastructure is followed for the components of the grid, depending on the component's characteristics.According to [5], the huge amount of data generated from smart meters varies with time and is not constant. Also, apart from the streaming smart meter data, offline date like consumer profile, weather, tariff plans have to be referenced. Hence highly reliable platform is required for real time computation and storage at nominal cost. Cloud Computing has the answer.

D.   DEMAND RESPONSE

There are huge demand response prospects in residential sectors and in industries [2]Through proper demand response and customized tariff programs for domestic and industrial usage, if it is possible to move considerable load off peak, then there would be significant cost saving. Moving to distributed generation and resources like micro grids, Pluggable Hybrid Electric Vehicles (PHEVs), solar energy generation via panels at homes etc, fully utilize the potential of smart grids. Incorporating this feature in smart grids require precise demand response programs.[1] proposes the Dynamic Demand Response Program to help in intelligent decision making via demand forecasting and curtailment forecasting models.These models accurately predict the near real time areas of peak and mismatch and also identify the buildings or customers for curtailment.

E.   STATIC ANALYTICS

Data from sub stations must be integrated for midterm and long term forecasting.Persistence before storage will cost high latency in streaming analytics. Hence after analytics the data from smart meters could be stored for further analysis. Midterm (MTLF) and Long term load Forecasting (LTLF) could use such stored data. Also, various options for general tariff recommendations and personalization at the substation level are possible by analysing such data. It has been shown in [3] thateach time the system predictsdemandit is stored in any database for big data like Cassandra[13] and manage storage and retrieval using HDFS [14].Static offline analytics also reveals energy theft or meter tampering patterns.

III.   KEY CHALLENGES

Apart from the various advantages of smart grids, a real time analytics framework also has major challenges.

Privacy, Ethics and Security: The software platform is analysing more sensitive data like consumer energy/device usage pattern in home as well as industry. Such data, is accessed maliciously can lead to the understanding of behaviour of customers. Also, authentication of smart meters and frequent outlier analysis or divergence analysis is required by utilities for fraud or anomaly detection.

Scalability: Pilot projects in smart grids are currently being tested with scalability in mind. When millions of homes are sending high frequent data, which has to be compared analysed and economic incentives recommended with latency in seconds, scalability of software platform has to be considered.

Localized Services: Though the analytics framework is general for the entire smart grid, some locality based service option could be necessary as in heavily concentrated industrial areas, academic areas, tourist spots, local festivals etc where activity pattern might differ.

Service Desk Operations: When demand response and customer economic incentives are significant features of smart grid operations, a service desk for customers that operates24/7 is essential. Data and feedback from such desks should be incorporated into the analytics or recommender platforms for improved customer satisfaction.

Response Time: The demand response has to quickly respond and come up with energy saving options for customers. The customers should be able to view the options at least 3-4 hours in advance to better plan and use the incentives proposed.

Outages: Renewable energy generating stations often undergo equipment repairs, outages or damages, which some time require more than a day to be fixed. In such cases the utilities or grid should be able to still control the supply demand mismatch.

Relevancy of historical data: With the use of electronic items increasing multifold, reference of historical data for prediction is a concern. A domestic customer profile five years back need not match his usage pattern today. His usage of grid energy would have increased, based on new equipment added to his home or would have decreased, in case his house is equipped with solar panels. Such cases should be considered when deployed in real world.

IV.   CONCLUSION

This paper has surveyed the current software analytics frameworks and technologies in smart grids.We infer that streaming analytics scalable platform in cloud is required for smart grid big data analytics. Real time solutions on a very high scale are required for smart grids compared to pilot projects. The key challenges are also discussed. Our future work will involve on building a real time analytics architecture for smart grids.

REFERENCES

1. Simmhan, Y,Aman, S, Kumbhare, A., Rongyang Liu , Stevens, S. , Qunzhi Zhou and Prasanna, V, "*Cloud-Based Software Platform for Big Data Analytics in Smart Grids*", Computing in Science & Engineering,Vol:15,Issue: 4,DOI: 10.1109/MCSE.2013.39 ,2013
2. Zhong Fan, Qipeng Chen, Georgios Kalogridis, Siok Tan, and Dritan Kaleshi, "*The Power of Data: Data Analytics for M2M and Smart Grid*", 3rd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe), 2012
3. M.Mayilvaganan and M.Sabitha, "*A cloud-based architecture for Big-Data Analytics in Smart Grid: A Proposal*",IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) ,2013
4. Alahakoon and D. Xinghuo Yu,"*Advanced Analytics for Harnessing the Power of Smart Meter Big Data*", IEEE International Workshop on Intelligent Energy Systems (IWIES), 2013
5. Couceiro, M. , Ferrando, R. , Manzano, D. and Lafuente, L., "*Stream analytics for utilities. Predicting power supply and demand in a smart grid*", 3rd International Workshop on Cognitive Information Processing (CIP), 2013
6. Arup Sinha, S.Neogi, R.N.Lahiri,, S.Chowdhury, , S.P.Chowdhury and N.Chakraborty, "*Smart Grid Initiative for Power Distribution Utility in India*" ,IEEE Power and Energy Society General Meeting, 2011
7. Bitzer, B. and Gebretsadik, E.S., "*Cloud Computing Framework for Smart Grid Applications*" ,48th International Universities' Power Engineering Conference (UPEC), 2013
8. Osman, A. , El-Refaey, M. and ElNaggar, A. "*Towards Real-time Analytics in the Cloud*", 2013 IEEE Ninth World Congress on Services
9. M. Zaharia, M. Chowdhury, M. Franklin, S. Shenker, and I. Stoica,"*Spark: cluster computing with working sets*," in Proceedings of the 2nd USENIX conference on Hot topics in cloud computing, pp. 10–10, USENIX Association, 2010.
10. M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M.McCauley,M. Franklin, S. Shenker, and I. Stoica, "*Resilient distributed datasets:A fault-tolerant abstraction for in-memory cluster computing*," in Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation, 2011.
11. M. Zaharia, T. Das, H. Li, S. Shenker, and I. Stoica, "*Discretized streams: an efficient and fault-tolerant model for stream processing on large clusters*," in Proceedings of the 4th USENIX conference on Hot Topics in Cloud Ccomputing, pp. 10–10, USENIX Association, 2012.

International Journal of Innovative Research in Science, Engineering and Technology

12. M. Zaharia, T. Das, H. Li, T. Hunter, S. Shenker, and I. Stoica, "*Discretized streams: A fault-tolerant model for scalable stream processing*," UC Berkeley Technical Report UCB/EECS-2012-259, 2012
13. Apache Foundation Hadoop HDFS Architecture, [Online]: http://hadoop.apache.org/docs/stable/hdfs_design.htm
14. Apache Cassandra, [Online]: http://cassandra.apache.org/
15. Yahoo S4, [Online]: http://incubator.apache.org/s4/
16. Apache Storm, [Online]: https://storm.apache.org/