A Secure Cloud Architecture for Data Generated in the Energy Sector

Michael Pham-Hung, Pirathayini Srikantha(✉), and Deepa Kundur

University of Toronto, Toronto, ON, Canada
{michael.pham.hung,pirathayini.srikantha,dkundur}@mail.utoronto.ca

Abstract. In urban cities, intelligent devices such as smart meters are deployed extensively and these provide comprehensive real-time snapshots of energy consumption patterns to the electric power utility (EPU). This data can be leveraged by the EPU, other third-party solution providers or consumers themselves to make informed and smart decisions that enable sustainable power consumption. However, managing vast amounts of meter data generated at a regular basis in a secure and economical manner is not a trivial task. We propose a cloud-based architectural model that leverages recent developments in standardized data access and encryption techniques to enable flexible, secure and economical management of energy data. We also present a prototype of a mobile application that performs analytics on meter data to provide information to consumers that will incentivize more sustainable power consumption patterns.

Keywords: Cloud services · Green button standard · Security · Homomorphic encryption

1 Introduction

Today's grid is a cyber-physical system consisting of many interconnected intelligent devices that perform measurement and actuation in real-time. Information and tools made available by these devices allow both the power supplier and consumer to make informed decisions on sustainable energy management. The advanced metering infrastructure (AMI) is a critical component of the electric power utility (EPU) composed of smart meters that are equipped with bi-directional communication capability [9]. These transmit local energy usage measurements of thousands of consumers to the EPU at a daily basis. This data is not only convenient for billing consumers, it contains interesting insights on energy consumption patterns that can be leveraged by the EPU, third party solution providers and the consumers themselves to promote sustainable power consumption behaviour. In order to glean these insights, this data must be effectively stored, managed and processed. This represents a significant challenge for the EPU as metered data is vast and is continually generated at a fast pace. Significant resources are necessary to store and manage data at this scale and using

A. Leon-Garcia et al. (Eds.): Smart City 2015, LNICST 166, pp. 374–383, 2016.
DOI: 10.1007/978-3-319-33681-7_31

in-house methods for this purpose is not economical. A very viable alternative for the EPU is the cloud.

The cloud provides on-demand access to storage and computational resources without prior commitment. Moreover, pricing models of cloud services are extremely economical. Although the cloud consists of many features that address most needs of the EPU for handling metered data, security is a primary concern. As the cloud providers do not reveal details on the internal workings of their infrastructure, the EPU cannot ascertain the security level of the system (i.e. unpatched vulnerabilities, etc.). Security is an important consideration for the EPU as metered data can be analyzed to reveal confidential information of consumers. We propose a cloud-based architectural model with a central focus on security. Main components of our proposed model include additional processing of metered data via homomorphic encryption prior to cloud storage and the integration of the Green Button standard that provides interfaces for accessing authorized data with privacy considerations in tact. In order to demonstrate how our proposed architecture provides means for secure data management and processing, we present a mobile application prototype that can be provisioned by the EPU to encourage sustainable energy consumption based on analytics performed on metered data managed by our cloud model.

2 Background

In this section, we provide a brief background on the Green Button standard, homomorphic encryption and cloud services.

2.1 Green Button Standard

The Green Button initiative provides interfaces that can be used to access available metered data in a standardized manner by solution providers and consumers. Privacy is incorporated into this standard as any data obtained via Green Button excludes Personal Identifiable Information (PII) [1]. Only meter measurements obtained within a specified time interval is provided without including information of its origins. Utility providers are already using smart meter data for billing purposes. However, proprietary protocols have been used to access this data. This makes data sharing extremely challenging. The green button standard provides a universal format that allows consumers to gain access to their data through the use of the *Download my Data* (DmD) interface. Apart from being able to view their own data, consumers are able to authorize access to their data for usage by third party solution providers via *Connect my Data* (CmD) interface. These features promote privacy and security. The proposed architectural model revolves around the Green Button standard as the cloud will be able to gain access to standardized formatted data from EPU companies through these interfaces and this data can then be processed by third parties to infer interesting estimations and models without impinging on consumer privacy.

2.2 Homomorphic Encryption

Homomorphic encryption represents a group of semantically secure encryption functions that allow certain algebraic operations on the plaintext to be performed directly on the ciphertext. Mathematically, given a homomorphic encryption function E( ), and two messages $x, y \in N$, we are able to compute $E_k(x \star y) = E_{k1}(x) \circ E_{k2}(y)$, without knowing the plaintext x, y or the private key [8]. Ideally, a fully homomorphic encryption allows any function E( ) to be computed on ciphertext, however these schemes, such as Gentry, are hugely inefficient [15]. More practical schemes, such as Pailler's scheme, are partially homomorphic [14]. These schemes can only compute a limited amount of functions on ciphertext, however are much more efficient and therefore practical.

2.3 Cloud Services

According to the National Institute of Standards and Technology (NIST), cloud computing systems enable convenient, on-demand network access to a shared pool of configurable computing resources and services (e.g. networks, servers, storage, applications) that can be rapidly provisioned and released with minimal management effort or service provider interaction [2]. The attractiveness of cloud computing, specifically the *Infrastructure as a Service* (IaaS) model, stems from its scalability and economic viability. There is potential to obtain access to a vast amount of computing services according to the needs of clients without up-front commitment [3]. In this paper, we are focusing on the IaaS cloud service which is typically provided by a third party cloud provider.

With the vast and increasing amount of data that the energy sector generates and processes, an economically viable option for utilities is to outsource computation and storage to cloud providers. However, there are a few concerns listed in the following that must be addressed by the utilities prior to fully engaging in cloud services.

Forensic Accessibility: Metered data contains highly revealing private information of consumers. For example, energy signatures can be applied to metered data to determine the occupancy and activities of a home [11]. When handling sensitive consumer data, utilities must be prepared to provide logs and forensic records for regulatory compliance purposes. It will be impossible for cloud providers to provide this information without exposing their internal architecture and algorithms. This lack of transparency renders significant difficulties for the power EPU sector as it is accountable for the protection and integrity of the data it stores [4].

Multi-tenancy: The method by which data is distributed across the cloud is completely under the control of the cloud provider. In a multi-tenant public cloud system, data is typically spread across multiple physical servers that may be shared with other users. Any unresolved vulnerabilities in the virtual environment that divides physical resources can be exploited by other tenants to access data in an unauthorized manner [5].

Resource Location: End-users use the services provided by the cloud providers without knowing exactly where the resources for such services are located. The physical systems supporting the cloud can possibly reside in other legislative domains. As local laws apply, exposure to possible issues that affect the integrity and privacy of data maybe inevitable.

Authentication and Trust of Acquired Information: Since data is stored within a third party infrastructure, it possible for information to be altered without the owner's consent. Authenticity and integrity of data must be guaranteed [4].

These security issues can be overcome by integrating standard interfaces for data access and incorporating additional layers of security to data prior to cloud deployment in a practical manner. The proposed architecture will preserve privacy and data integrity.

3 Related Work

Many industries have recently begun adopting cloud computing to share data in a quick and cost-effective manner.

In the health care industry, data sharing is essential for health problem detection, solution identification and medical resource allocation. Reference [6] proposes a solution for preserving the privacy of medical records stored within the cloud. Their approach involves vertical partitioning of their data into plaintext, anonymized and encrypted sections. The data owner can then authorize the merging of the partitioned data to improve medical search or analysis. As an added level of security, their proposal also allows for a hybrid method to check the integrity of data due to different requirements between data owners and recipients. Such a comprehensive solution for metered data adds significant complexity and can introduce difficulties maintaining a standard method for data access.

In the energy sector, research is plentiful in exploring how the smart grid can be combined with the cloud. One solution proposes the use of a cloud architecture that provides a platform in which third parties are allowed access to Green Button Data and development tools to provide further analysis of energy usage data [7]. These tools allow users to gain valuable information from the usage data with analysis and results aiming to reduce costs for both the users and the suppliers themselves. This paper, in contrast, proposes a security based cloud architecture that leverages standardized protocols to access sensitive data while preserving privacy and integrity.

In terms of security, reference [8] provides an excellent example of an encryption model, however it is currently not integrated with green button services. Through the use of homomorphic encryption, they suggest encrypting data starting from the smart meter to maintain anonymity. Homomorphic encryption allows data to be encrypted and have plaintext algebraic functions performed on the cipher text. This ensures that all smart meters participate in the aggregation while simultaneously maintaining user privacy. Within this paper, such a model is assumed to exist and this paper incorporates the encryption scheme to improve the security of our proposed architectural model.

4 Proposed Architecture

The main focus of our proposed smart grid cloud architecture is to provide a secure solution to store and organize EPU data. Revolving around the Green Button standard format, the cloud will be able to utilize existing DmD and CmD interfaces along with added security features to improve the integrity and usability of data.

As clouds provide resources that otherwise would have been too costly to manage internally, access to ample resources such as storage on the fly becomes trivial. With the use of the cloud, EPU companies certified by Green Button will be able to store information within a database residing in the cloud. Utilities, consumers and third parties, with authorization, will then be able to access this data from the past or present. Green Button data is stored in Extensible Markup Language (XML) which can be easily accommodated by databases.

The cloud is also able to provide access to powerful computational resources. One of the main activities that the utilities and third parties accessing the data will perform is executing various computations for further analysis. Therefore, within the proposed architecture, a separate computational component in the cloud is available for this purpose. As comprehensive powerful features such as these are provisioned by the cloud, all major operations can be executed on the cloud.

Before storing data, our model encapsulates this data with an added security layer via homomorphic encryption. This addresses many of the security and privacy issues raised earlier. As all stored data is stored as encrypted cipher-text, data will not be subjected to privacy issues due to leakage. The major attractiveness of homomorphic encryption is that mathematical operations can be directly performed on the cipher-text (i.e. there is no need to decrypt data). Since all processing of metered data involve computations, these can be performed on the cloud in the encrypted state. In other words, during the lifetime of the data, third parties hosting the cloud will never be able to gain access to the content of the data as data is never decrypted on the cloud. Figure 1 illustrates the important components of our proposed architecture.

Each process in Fig. 1 is labelled and described in detail in the following:

1. Since 2010, every home in Ontario has been installed with a smart meter [9]. Not only do smart meters measure energy and gas usage, but they are able to store and send measurement data to EPU companies.
2. Assuming that the model referenced in [8] exists, homomorphic encryption is applied to all information gathered from the smart meters. The secure data transfer is visualized in orange.
3. Utility companies (Toronto Hydro etc.) convert data received from the meters into the Green Button Standard format [1].
4. The cloud will actively gather data by using Green Button's "Download my Data" and "Connect my Data" interfaces and possibly a third interface to gather the completely anonymous data.
5. The web server will act as the wrapper and the communicating point between the cloud, users and EPU companies.

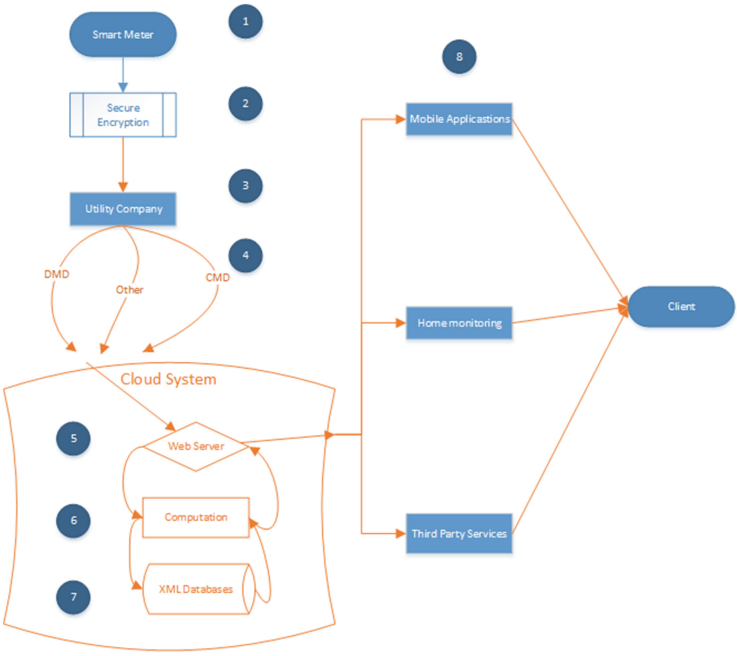

Fig. 1. Proposed architecture model

6. Computations on data before or after storage are executed by computational applications installed in the cloud.
7. The cloud will store data within an XML database. Every set of data will be tagged/labelled with parameters such as location for different types of computation and aggregation.
8. This data can be accessed by third party entities which can perform necessary operations on data in the cloud and then use these in applications that for example promote sustainable use of energy.

We are in the process of implementing a testbed modelling this cloud architecture at a smaller scale. We have used only open-source software to implement the entire testbed. Usage data is generated from Green Button's Sandbox which is populated with sample personal data [1]. The cloud itself is managed by the open source software Cloudstack [13]. The database and querying component is supported by the Basex XML Database engine.

5 Mobile Application

5.1 A Proof-of-Concept

In order to demonstrate the practicality of the proposed cloud-based smart grid data management architecture, we propose a mobile application that accesses data according to the architecture illustrated in Fig. 1.

The main motivation of this application is to encourage sustainable behaviour with respect to energy consumption. The application also serves as a proof-of-concept that illustrates the possibilities of useful applications which can access data in a manner that maintains security and privacy. This application will be deployed via the Android platform to ensure portability and convenience. It will be developed on the Android environment with the use of Android Studio. The end-product will then be tested along with the test-bed cloud which will simulate the interactions between the two entities.

The application will collect the aggregate consumption data of the user's neighbourhood, compute the average consumption rate and then compare this with the user's usage via a simple interface. The aggregation will be able to operate on anonymous data using the proposed cloud architecture at a time scale that matches the rate at which data is produced (e.g. hourly). Users will receive a real-time rendering that compares their consumption with the average consumption in their neighborhood. Following is a detailed overview of the processes required for typical interactions with the mobile application:

1. Users will supply their log-in credentials into the single time log-in screen of their mobile application. This will implicitly evoke the DmD and CmD interfaces.
2. User supplied information will then be delivered to the cloud's web server. Data sets authorized for the user according to the postal code associated with the user's residence is fetched via DmD from the XML database in the cloud. This information is still encrypted. In the case of this application, Paillier homomorphic encryption suffices due to the low complexity of the computational requirements of this application [14].
3. Using the computation component, aggregation operations are performed on the rendered data sets to obtain average consumption at the geographical location corresponding to the user.
4. The geographical average consumption and the user's consumption data are then sent via the web server to the user's device.
5. The user's device then decrypts and plots both sets of data, highlighting the differences between these. To ensure users are enticed to alter their behaviour, the graph will include a dynamic vertical scaling in order to ensure that the differences are displayed prominently to the user. Green and red trend lines will show the user the time periods at which their consumption rate is above or below average.
6. Also displayed within the application is an estimate of the amount of savings the user has been able to achieve from his or her typical energy consumption.

In Fig. 2, there are two illustrations of the Graphical User Interface (GUI) of the mobile application. The first is the primary screen the user sees after logging in. This screen is labeled as the "SAVE" screen. The "SAVE" screen provides users with a quick summary of their progress towards more sustainable power consumption behaviour: a positive value indicates that they are well on track of achieving their goals. Additionally, the "SAVE" screen presents advice/tips

to the user on how they can lower their energy consumption rates. This advice is dynamic and is based on energy signatures that can be used to detect and suggest the conservation steps that can be taken by the user. Also, this advice is tailored to various periods in a day (i.e. expensive during peak and inexpensive off-peak hours).

The second screen, labelled as "COMPARE", provides a graphical rendering that compares the differences between the user's usage and the average power consumption in the system. The user will also be able to view other data within Green Button. The user's usage trend is colored in the illustration provided in Fig. 2. This color can change depending on whether the user's trends are above or below the geographical average depicted in black. In Fig. 2, the user's consumption is illustrated in red where the application informs the user that their consumption rate is greater than the average. When the user consumption rate is green, this indicates that the user's consumption is well below the average and this is desirable. The colored cues provide an intuitive visual for users to interpret their data. The graph rendering is therefore dynamic and interactive. The user is able to pan in and zoom out for further analysis of their consumption.

On both screens, two buttons are present to show the user which screen they are on. The screens are labelled respectively with the depressed and bold button showing the current screen.

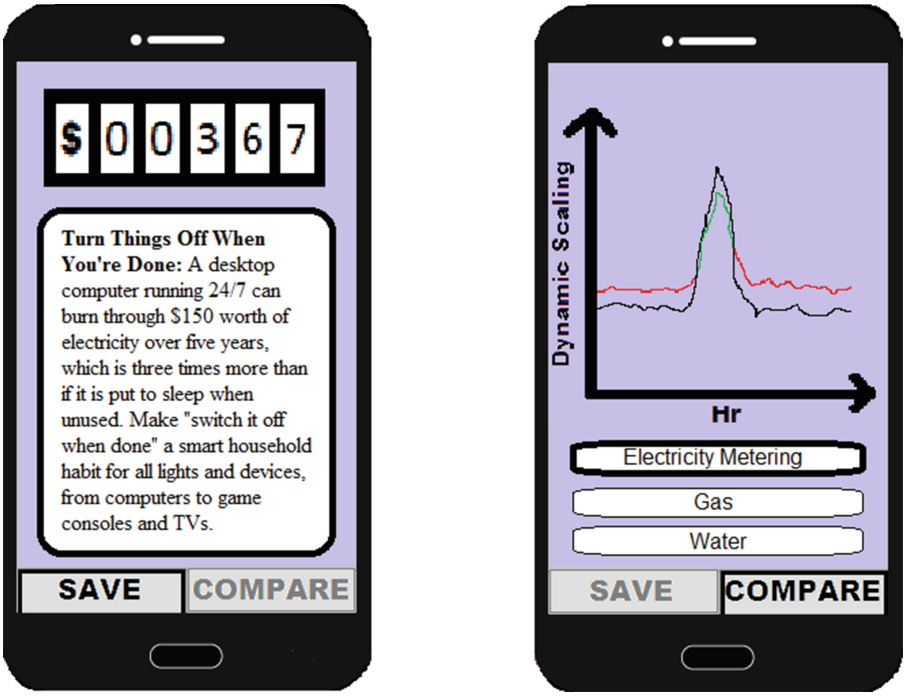

Fig. 2. GUI of mobile application prototype

Applications similar to the one proposed in this paper can be shown to encourage purely sustainable behaviour in the long-run with appropriate incentives. For instance, work in [12] evokes evolutionary game theory to prove that with appropriate incentive mechanisms in place, consumers can be induced to select more profitable actions (i.e. sustainable behaviour) with a probability of one. As future work, we intend to deploy this mobile application to the public and analyze whether this is indeed true in a realistic setting.

5.2 Our Contributions

The proposed cloud architecture provides a secure access point to data custodians, retail customers and third parties, revolutionizing the Green Button standard. The main purpose of the mobile application is to test the architecture and evoke each process flow. The mobile application provides an example of the possibilities the architecture provides for analysis and comparison tools without compromising personal information.

6 Conclusions

In this paper, we have proposed a comprehensive cloud-based architecture that enables secure big data management for the EPUs. We have utilized existing standards to enable universal accessibility for authorized clients. We have also proposed additional processing of data with homomorphic encryption to securely store data while also capitalizing on the extensive computational resources provided by the cloud. Our mobile application prototype demonstrates how various flows induced by interactions between the cloud, the EPU and the client are efficient and secure. As future work, we intend to deploy our testbed and mobile application to real consumers and analyze whether incentives displayed by the application effects any impact on power consumption patterns.

References

1. Green Button Data, An Overview of the Green Button Initiative, 25 June 2015. http://www.greenbuttondata.org/learn/. Accessed 14 July 2015
2. Mell, P., Grance, T.: The NIST Definition of Cloud Computing. National Institue of Standards and Technology (NIST), Gaithersburg (2011)
3. Ugale, B.A., Soni, P., Pema, T., Patil, A.: Role of cloud computing for smart grid of india and its cyber security. In: International Conference on Current Trends in Technology, pp. 1–5 (2011)
4. Rong, C., Nguyen, S.T., Jaatun, M.G.: Beyond lightning: a survey on security challenges in cloud computing. Comput. Electr. Eng. **39**(1), 47–54 (2013)
5. Simmhan, Y., Kumbhare, A.G., Cao, B., Prasanna, V.: An analysis of security and privacy issues in smart grid software architectures on clouds. In: IEEE 4th International Conference on Cloud Computing, pp. 1–8 (2011)
6. Yang, J.-J., Li, J.Q., Niu, Y.: A hybrid solution for privacy preserving medical data sharing in the cloud environment. Future Gener. Comput. Syst. **43**, 74–86 (2015)

7. Ballijepalli, V.M., Khaparde, S.A.: Smart grid standards conformed cloud based demand side management tools. Int. J. Eng. Res. Technol. (IJERT) **1**(5), 7 (2012)
8. Li, F., Luo, B., Liu, P.: Secure information aggregation for smart grids using homomorphic encryption. In: First IEEE International Conference on Smart Grid Communications, pp. 327–332 (2010)
9. Hydro One, Smart Meter, Hydro One. http://www.hydroone.com/myhome/myaccount/mymeter/pages/smartmeters.aspx. Accessed 21 July 2015
10. Gens, F.: IDC eXchange, 2 October 2008. http://blogs.idc.com/ie/?p=210. Accessed 20 July 2015
11. Ruzzelli, A., Nicolas, C., Schoofs, A., O'Hare, G.: Real-time recognition and profiling of appliances through a single electricity sensor. In: Seventh IEEE Communications Society Conference on Sensor Mesh and Ad Hoc Communications and Networks, pp. 1–9 (2010)
12. Ramchurn, S.D., Vytelingum, P., Rogers, A., Jennings, N.: Agent-based control for decentralised demand side management in the smart grid. In: 10th International Conference on Autonomous Agents and Multiagent Systems (2011)
13. A. S. Foundation, Apache Cloudstack. https://cloudstack.apache.org/index.html. Accessed 29 July 2015
14. Paillier, P.: Public-key cryptosystems based on composite degree residuosity classes. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS, vol. 1592, pp. 223–238. Springer, Heidelberg (1999)
15. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: STOC, vol. 9, pp. 169–178 (2009)