Scalable Cloud-Based Data Storage Platform for Smart Grid

Hnin Yu Shwe[1(✉)], Soong Boon Hee[1], and Peter Han Joo Chong[2]

[1] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore
hninyushwe@ntu.edu.sg

[2] Department of Electrical and Electronic Engineering, Auckland University of Technology, Auckland, New Zealand

Abstract. Real-time conditional monitoring and instinctive fault diagnosis plays an important role in the smart grid systems. Conventional data storage and management systems are facing the difficulties with the disseminated, heterogeneous and large volume of data. Conventional data management infrastructure utilizes centralized approach which comprises of large-scale server, disk array data storage hardware and relational database management system (RDBMS) for database services. This type of infrastructure may result in poor performance to support the requirement of smart grid applications for very large database with high data request rates at very low latency. Precise, fast, vulnerable and cooperative information system is the basis requirement for the smart grid applications. In this paper, we proposed a distributed data storage and management platform which guarantees high scalability and reliability data storage and management with the huge amount of information in the smart grid application. In order to evaluate and validate our scalable cloud-based data storage platform for smart grid, we deployed our proposed infrastructure in our cyber-physical test-bed which is a facility to test innovative technologies for building efficiency and urban sustainability.

Keywords: Mesh · Cloud · Internet of Things · Data storage · Data management · Smart-grid

1 Introduction

The smart grid is the progression of existing electrical grid, using new technologies such as two-way information communication and computational intelligence in order to optimize the conservation and delivery of power [1]. Transforming the conventional electric delivery system to a smart grid incurs incorporation of real-time monitoring and control, distributed power sources, bi-directional energy flow between the sources and the customers. Integration of intelligent communication and control across the existing electric energy system can maximize the system interoperability and flexibility and minimize the carbon footprint.

E.T. Lau et al. (Eds.): SmartGIFT 2017, LNICST 203, pp. 259–265, 2017.
DOI: 10.1007/978-3-319-61813-5_26

In the age of IoT, a number of physical and virtual things are interconnected to become smarter, intelligent, informative and enable people to enjoy a better quality of life [2]. Smart grid is one of the application domains of urban IoT in which grid components are considered as IoT objects. Smart grid consists of several components such as smart appliances, smart meters, ecological vehicles, home and office buildings, power plants and renewable energy resources. From IoT perspective, smart grid requires an architecture that supports all information and resources to be accessible from everywhere, obtain relevant grid operational information in efficient time manner, understand data in actual time to allow for further improvements in reliability or power quality [3].

The above illustrated requirements of the smart grid can be achieved by taking advantage of the cloud computing model. Cloud computing practices distributed resources and, as a consequence, it has the profits of greater reliability, flexibility, extensibility and powerful computing capability. To obtain a reliable infrastructure for smart grid, cloud computing solutions and services must be integrated with the conventional scheme. In this work, we proposed a scalable cloud-based data storage platform for smart grid which offers better architecture than conventional centralized scheme to accommodate massive data, analyse data in real time and make rapid decisions to improve reliability and power quality. Considering the key requirements of smart grid architecture and utilizing the modern advanced technologies, the basic idea of this work is to embed control, computational power and monitoring capabilities into the cloud.

The paper is organized as follows. In the next section, we discuss about the other related works in the literature. In Sect. 3, we describe our proposed scalable cloud-based data storage platform for smart grid system followed by its services. Finally, we provide future challenge and conclude our paper.

2 Background

Smart grid is an active field of research, and renewable energy production, real-time monitoring and prognosis health system, remote and local control over communication system play important research areas of interest [4]. The components of smart grid are distributively located in the grid, ranging from residential and office buildings, power sources to utility data and control stations. As described in the introduction, each grid component has the ability to access and exchange data via different communication protocols. Data storage and management system provides facility to store the data in a systematic way and enable the data to extract, compute and analyse either promptly or subsequently. In a conventional system, centralized magmatic hard disk drives are used to store the data.

Smart grids can be assumed as a big data challenge which involves massive amounts of data and analytic. The huge amount of data from smart grid components need to be adequately managed to achieve the maximum reliability and sustainability of the smart grid. Nowadays, cloud computing and cloud database become attractive and many people emphasize on how to obtain profits from cloud technologies in the smart grid environment [5]. A cloud database can either be a virtual or physical database. A virtual cloud database is a database which is installed on a virtual machine instance which can be purchased for a limited time. Alternatively, a physical cloud database is a database as a service in which the service provider installs and maintains the database, and application owners pay according to their usage.

Different approaches towards data center location could be used by a cloud storage provider. In a centralized cloud data storage architecture, the data are being kept in a central data center which is physically located in one geographic location and users can able to access the data via internet connection. Popular cloud data storage provider, such as Dropbox [6], follows the centralized topology and it requires the clients to access the same data center which is located in a specific location [7]. In a distributed cloud data storage architecture, in contrast, the data are being kept in several distributed data centers. The clients can access the different data center which is closest to its location geographically. Eg., Google Drive [8] utilizes the distributed topology and it uses multiple cloud storage data centers which are spread over a geographical area [9,10].

In addition, cloud data storage can either be public or private. If the infrastructure and services are provided off-site by a third-party provider over the Internet, it is called public cloud. In private cloud, on the other hand, infrastructure and services are maintained on a private network. In spite of both cloud storage services dominate conventional data storage in terms of scalability and performance, the private cloud offers higher level of control and security over public cloud. Additionally, private cloud offers the great ability to customize the storage, computation and networking elements to optimal suit with the requirements of the specific entities.

Integrating the two modern technologies; cloud services and wireless mesh network, in this work, we introduce a cloud-based data storage platform which satisfies the requirements of smart grid applications. Proposed cloud-based data storage platform ensures users to utilize the data storage center more effectively and efficiently in terms of data storage, data accessibility, data analysis, data processing with optimal performance.

3 Proposed System

3.1 Cloud-Based Data Storage System Architecture

High level overview of our proposed cloud-based data storage architecture is presented in Fig. 1. The system is constituted by two-tier network and it mainly supports for multi-level decision making and hierarchical scheduling.

In first-tier, regional wireless mesh access points (APs) are installed in each grid point to form a mesh distributed cloud data storage network. The mesh APs are installed regionally in the grid to distributively store real-time or time-critical data. Central data storage center is considered as second-tier, in which historical and long-term data of the grid system are stored.

In wireless mesh backbone network, all mesh APs are designed with local storage drive to store the data from the grid components and organized as a

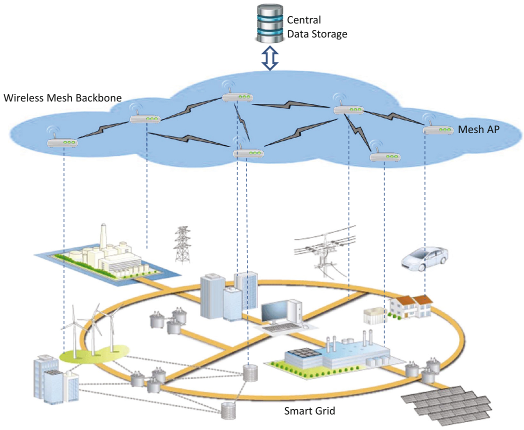

Fig. 1. High level architecture of cloud-based data storage platform.

private cloud distributed data storage network. Every node in the network collaborates for data distribution and data is available throughout the network seamlessly. Gateway node collects raw data from the grid point, store in its local drive and performs first-level data processing or analysis. The proposed cloud-based distributed data storage platform supports large scale data interaction and information exchange among different systems (regions) in the smart grid. Data interaction is available in real-time and, hence, a mesh client can able to retrieve time critical information of the other regions in the grid with near-zero latency. In order to test the performance of our proposed framework, we set up the wireless mesh network by using four-radio dual-band wireless mesh APs. Figure 2 shows the MeshRanger MN4300 AP which is compatible in both outdoor and indoor environment.

Fig. 2. A wireless mesh AP: MeshRanger MN4300.

Cloud platforms are essential to establish a software architecture that meets the requirements of smart grid applications. In the smart grid system, it is very important to constantly monitor and analyse the continuous data streams from multiple sources with near-zero latency. Taking the advantage of cloud architecture over centralized approach, our proposed distributed cloud storage architecture provides the facility to process huge, persistent streams of data in the most efficient way and transparent to the user.

3.2 Data Storing and Processing

One of the facilities of our proposed architecture is to support the user with multi-level decision making and hierarchical scheduling. To achieve this facility, our data storage architecture is composed with two-tier network and it offers data to the users hierarchically. In our data storage architecture, data are generally classified into two types: Type-A and Type-B.

We define time-critical and real-time data as Type-A and which are stored in the first tier, the wireless mesh network. Type-B data comprises historical and long-term data and these data are stored in the second tier, the central data storage. Information exchange can be taken placed among mesh APs over the wireless mesh backbone network. Efficient decision making is supported by providing data access to regional data in each mesh AP. In this way, data can be retrieved in efficient time manner and make the decision locally and perform the control timely, reducing the auxiliary effort to connect the central decision server for every decision making process.

Moreover, our proposed cloud-based data storage architecture takes the power of the cloud and offers it to the users. Proposed system performs first-level data processing with the use of cloud data computing and can execute simple data processing such as data aggregation and data extraction. Another feature of the cloud platform is that the capacity of the network storage can be efficiently and effortlessly scaled up as the system requires. In addition, the proposed architecture offers a semantic platform to make actual information available to each grid component, including connectivity to variety of devices and interoperability of multiple systems.

With the use of proposed cloud data storage platform, the grid operators can obtain the real-time information from the grid and able to understand the current situation of the grid and perform the necessary action to optimize the conservation and delivery of power. By integrating the proposed architecture with decision making service, the proposed platform compromises more value and power to all the users in the smart grid system.

3.3 Services

This section will discuss the services offered by our proposed system.

Database as a Service (DaaS): It supports the grid operators to constantly monitor the performance and condition of the grid and make the control decisions promptly and precisely. Moreover, grid operators are granted to obtain the grid data in real-time in order to immediately detect grid performance issues such as power outages and shortages. Additionally, grid operators can also able to access the long-term grid data to perform trend analyses, prediction, power distribution planning and other necessary arrangements.

Dynamic Cloud Topology: To minimize the resource request time, it supports dynamic cloud topology. The network topology varies dynamically depend upon the user's request.

High Throughput: It offers higher throughput by avoiding network congestion at central data storage. Since the users are able to obtain certain data from the regional mesh APs, it can avoid to access the central stand-alone data center for every process which may prone to heavy traffic congestion and bottlenecks at the central data center.

Data Transparency: Data transparency provides the ability to easily access the data from any point within the grid. Therefore,the users can transparently use the data storage, analyse and exchange the data among the grid components located in the different grid regions.

Low Data Communication Cost: One of the challenging factors of data communication network is data communication cost. Our proposed architecture offers lower data communication cost with the use of distributed data storage network. It minimizes the data communication by providing local data access to the regionally distributed mesh APs.

On-the-go Scaling: Since grid data are increasing rapidly from time to time, scaling plays a key role in the data storage architecture. Our cloud data storage architecture offers the users not to worry about running out of the storage space. The users can scale up the current storage space any-time, easily and seamlessly.

Easy-Access-to-Information: Another advantage of our proposed platform is easy access to the information which are geographically distributed throughout the grid. The users are not necessary to worry about their geographic locations and they can conveniently access the grid information via wireless mesh backbone network.

Massive System: This service is very important for the smart grid which is composed with several different sub-systems. Proposed framework stands as a unified data collector for assorted data types of distinct sub-systems and provides them to work under a consolidated management engine for better achievement.

4 Conclusion

This paper presented a scalable cloud-based data store platform for smart grids. The main objective of this paper is to propose a data storage framework which can handle huge amount of data in smart grid applications. From implementation and testing in our cyber-physical test-bed, we can validate the proposed framework was able to extract, store, process and query the huge amount of grid data with lowest latency. In addition, we reach our target of addressing two main critical issues, data scalability and data efficiency. From the testing in the test-bed, we could see our proposed framework achieves greater data scalability, transparency and efficiency along with an efficient and effective data storage service. We believe this framework can easily scale-up to the deployment of large-scale systems. In near future, we are going to deploy in large-scale and real environment and will improve the system by taken into accounts other performance factors.

Acknowledgement. This research is funded by the Republic of Singapore's National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore.

References

1. Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M.: Internet of Things for smart cities. IEEE Internet Things J. **1**(1), 22–32 (2014)
2. Miorandi, D., Sicari, S., Pellegrini, F.D., Chlamtac, I.: Internet of Things: vision, applications and research challenges. Ad-Hoc Networks **10**(7), 1497–1516 (2012)
3. Sadiku, M.N.O., Musa, S.M., Momoh, O.D.: Cloud computing: opportunities and challenges. IEEE Potentials **33**(1), 34–36 (2014)
4. Tao, F., Cheng, Y., Xu, L.D., Zhang, L., Li, B.H.: CCIoT-CMfg: cloud computing and Internet of Things-based cloud manufacturing service system. IEEE Trans. Ind. Inf. **10**(2), 1435–1442 (2014)
5. Hu, W., Yang, T., Matthews, J.N.: The good, the bad and the ugly of consumer cloud storage. ACM SIGOPS Oper. Syst. Rev. **44**(3), 110–115 (2010)
6. https://www.dropbox.com/
7. Drago, I., Mellia, M., Munafo, M.M., Sperotto, A., Sadre, R., Pras, A.: Inside dropbox: understanding personal cloud storage services. In: Proceedings of the 2012 ACM Conference on Internet Measurement Conference, Boston, Massachusetts, USA, pp. 481-494 (2012)
8. https://www.google.com/drive/
9. Bocchi, E., Drago, I., Mellia, M.: Personal cloud storage: usage, performance and impact of terminals. In: 2015 IEEE 4th International Conference on Cloud Networking (CloudNet), Niagara Falls, ON, pp. 106-111 (2015)
10. Drago, I., Bocchi, E., Mellia, M., Slatman, H., Pras, A.: Benchmarking personal cloud storage. In: Proceedings of the 2013 Conference on Internet Measurement Conference (IMC 2013), pp. 205-212 (2013)