DATA COLLECTING AND PROCESSING METHOD IN DISTRIBUTION SYSTEM USING EDGE COMPUTING TECHNOLOGY

Haizhu WANG, Wenxin GUO
Electric Power Dispatching Control Center of Guangdong Power Grid Co., Ltd– China
20801244@qq.com, guowenxin@gddd.csg.cn

Caishan GUO, Yuyan SUN
School of Electric Power, South China University of Technology – China
{gcs7601, jess894597048}@163.com

Jiangang LU, Ruifeng ZHAO, Yang LIU
Electric Power Dispatching Control Center of Guangdong Power Grid Co., Ltd – China
{lujiangang, zhaoruifeng, liuyang}@gddd.csg.cn

ABSTRACT

With the improvement of automation system and the proliferation of Internet of Things (IoT) in the distribution system (DS), massive heterogeneous data generated at the edge of DS is utilized for various applications. Edge computing can be a good solution to manage those data while guarantee diverse performance requirements, especially limited latency. In this paper, we propose a cloud-edge architecture for DS, and specifically illustrate the logic design of the edge computing (EC) device for DS (DSEC). In order to manage data at the edge of DS, a DSEC-based data collecting method, including a protocol identification and conversion method, a naming method and a data management method, is proposed. Lastly, a DSEC-based data processing method considering local processing and cloud offloading is designed to improve the utilization of cloud & edge resources. The proposed methods can ensure rapid data analysis requirements and better data service in DS.

I. INTRODUCTION

With the penetration of renewable energy generators, energy storage systems and electric vehicles to the distribution system (DS), as well as the growing demand side management requests, the operation and control requirements for the DS have become more complicated [1]. Moreover, with the improvement of automation level and the proliferation of Internet of Things (IoT) in the DS, the data generated at the edge of DS shows big data characteristics, like volume, variety, velocity, etc [2]. The above changes challenge the current data collecting and processing platforms in the DS, which is difficult to cope with the storage and data analysis issues due to such massive and heterogeneous data. Therefore a new data management architecture that can provide fast, reliable and economy data services is of great importance.

Cloud computing, which has gain widespread attention in different areas, is a new computing method to tackle the above problems by dynamically providing elastic and massive computing resources [3]. There are also various studies that attempt to introduce cloud computing in the distribution system and have verified its efficiency and benefits [4]-[6]. However, as a centralized computing method, cloud computing may fail to support some limited latency applications especially when workloads are extremely heavy or generated far from cloud computing platform [7].

A feasible and effective way is to push data services from the centralized center to the edge of the network, the proximity of data sources [8]. In this context, a new paradigm called cloud-edge computing is proposed to provide cloud computing capabilities at the edge of pervasive networks in close proximity to data producers. In consequence, the computing workloads of the data center and data transmission delay are decreased. Another benefit of edge computing is that, similar to cloud computing, virtual machine (VM) mechanism is utilized to provide flexible computing resource. Admittedly, VM resource can be allocated between cloud and edge computing, which can further improve the computing resource and minimize the service delay at the same time [9],[10]. Due to the tension between resource-hungry applications and resource-constrained edge computing terminals, the resource allocation mechanism between edge computing terminals and cloud center based on virtualized technology also seems promising and helpful.

Considering all those benefits of edge computing and its practical applications, such as managing massive data and latency guarantee applications, this paper aims at introducing edge computing in the DS to support data collecting and processing applications. The main work of this paper includes: 1) a cloud-edge architecture for the DS, as well as the software & hardware deployment scheme, 2) a data collecting method, and a data processing method for the cloud-edge-architecture-based DS.

II. OVERVIEW OF DSEC

Implementation scheme

The demand-side data is generated from multi-type and widely-distributed power distribution devices during the power distribution and utilization. The great value contained can satisfy the demand of the current traditional power distribution service development. On the other hand, it plays an essential and basic role in the development of distribution system operator (DSO) business. For the data value exploitation and meet the real-time and reliability requirements of different services, this paper proposes a cloud-edge architecture for DS (DSEC) with cloud center, DSEC based on edge computing and DS devices, as shown in Figure 1.

The bottom of cloud-edge architecture is close to users in DS. It is the source of massive, heterogeneous, multi-type data generated from devices in DS. The devices include renewable power sources (such as photovoltaic, wind power, etc.), energy storage devices, regional energy management units (such as smart buildings, smart homes, etc.), controllable load devices and intelligent data acquisition and adjustment units. These devices can upstream the raw data and also accept the regional and centralized control for controllable load, controllable power supply and controllable energy storage if the users agree.

The middle one is DSEC, serving as a hub for the devices and cloud center. Situated in the neighbour of data sources, DSEC meets the requirements of unified access, unified management and unified delivery for data in DS. DSEC is a distributed open platform integrated with network, computing and storage. With the unified interface, users and managers could get intelligent services and send command through DSEC. This is important for real-time, reliability and accuracy of the service.

On the top of cloud-edge architecture is cloud center, integrating the global data in DS. Cloud center is the brain for monitoring and decision-making of DS. First, it receives the data uploading by DSEC and performs basic management such as data analysis, protocol conversion and data quality verification. Second, cloud center deploys APPs for differentiated services on the standardized platform, where APP process the data in DS. Combined with the analysis results with the global data in DS, cloud center can make a centralized control for DS.

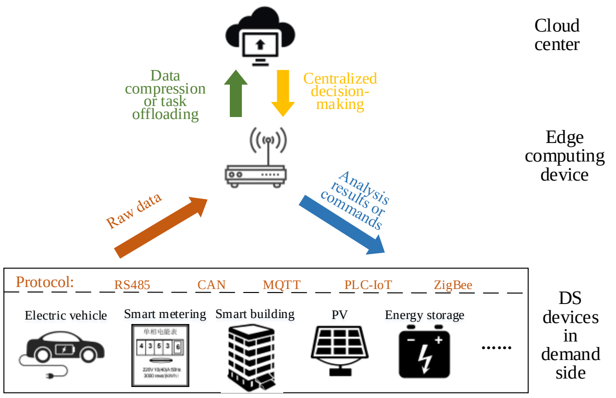

Fig.1. The cloud-edge architecture scheme for DS

Logic design

In order to meet the requirements of accessing massive, heterogeneous, multi-sources data in DS and regional distribution control, this paper proposes a logic design for DSEC based on the design concept of overall layered and modular logical architecture, as shown in figure 2. The vertical layers contains three parts: information access layer, edge computing layer and standardized transmission layer.

A. *Information access layer*

DSEC collect data from a large number of devices in DS through multiple protocols such as RS485, CAN, MOTT, PLC-IoT, ZigBee, etc. To dispose of the integrated data reported by different sources, we come up with several method in this layer including protocol identification and conversion method, Naming, data management.

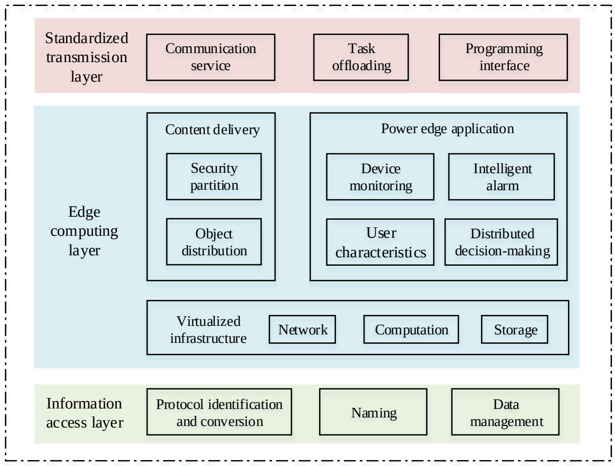

Fig.2. The logical design of DSEC

In the part of protocol identification and conversion, DSEC firstly provides available identification services for the added data, assessing the validity or otherwise of the added data according to its protocol. When the added data pass the identification, the data ought to be converted to a standardized protocol for integrated application of massive data in DS. Naming is an important issue that help to communicate and locate the data source, describing the basic attributes of the data including type, ID, time, contents, etc. Meanwhile, to ensure that the DSEC benefits from massive data, it is of great importance to manage data in logical and organized fashion. In this case, data management is responsible for monitoring the data quality, storing the data in DS in a proper structure and simplifying data processing.

B. *Edge computing layer*

As we discussed above in implementation scheme, edge computing service in DSEC should be provided for the devices and cloud center for the reduction of data transmission delay and the computing load in cloud center. There are three components in the edge computing layer. First, for the concern of efficiency and power consumption, we proposed a virtualized infrastructure, which is consist of network, computation and storage. Second, it is quite common that different data generated by devices means various degree of value, so we deliver data into different isolated containers according to its contents and security degrees. Last, power edge application is data processing at the edge of DS. It contains device monitoring, intelligent alarm, user characteristics and distributed decision-making.

C. *Standardized transmission layer*

In this layer, we form a bridge for DSEC and cloud center to fulfil the demand of communication and task cooperation. Standardized transmission layer has three integral parts, such as communication service, task offloading and programming interface. In the part of communication service, DSEC are required to communicate with cloud center in a transmission standard. Tasks from users in DS are always constrained for the consideration of limited resources in DSEC and demand of QoS for customers. So DSEC should offload some tasks under certain constraints, which called task offloading. The last part is programming interface, which is responsible for adding, eliminating and configuring the various application programming interface (API) of users and managers in DS.

III. DATA COLLECTING METHOD FOR DS BASED ON EC

In this section, we focus on the data collecting and management methods for DSEC from three aspects: naming, protocol identification and data management.

Data variety

To begin with, we discuss about data variety issue in DSEC. The more DSO business arise, the more data variety grows in number of quantity. Here we list the main data variety according to different applications such as distribution automation, micro-grids, demand response (DR) and regional energy management units. However, it is not limited to these.

A. *Distribution automation*

Distribution automation system play an essential role in operation monitoring and controlling. There are data varieties including telemetry, signal and remote control collected by feeder terminal unit (FTU), distribution terminal unit (DTU), transformer terminal unit (TTU) and remote terminal unit (RTU), respectively. What's more, measuring the performance of a power distribution room, where the above terminals work, is another important challenge. It is required to describe the performance by another quantity of metrics such as temporary, humidity, smoke, etc. Also, there are data varieties to monitoring the security, including access control and videos.

B. *Micro-grids*

The components in micro-grids are consist of distributed renewable power supply, energy storage and controllable load, etc. For the concern of operation of these components, DSEC need to obtain electrical measurement data and state quantity data among these components.

C. *Demand response*

Demand response (DR) refers to the users changing the normal power consumption pattern with response to the prices and incentive signals. To perform better demand response to power grid, power utilization data should be totally mined and exploited. Also, there is some controllable load for dispatch department and energy management and control systems (EMCS). The demand responses caused by different states and prices should also be collected.

D. *Regional energy management unit*

With the help of IoT, regional energy management unit obtain data from every aspect of a home, a building, a community and so on, collected by low-cost, chip-type and powerful sensors.

Naming method

Naming is a foundation for data collecting. With more and more devices access, DSEC should be aware of the added device automatically and effectively. Naming is a service that help to know about the specific information of devices and their data.

DSEC names each device with description of the following information: ID (who), data description (what), IP address/ Mac address (where), sampling time (when) and communication protocol (how). Each name is unique and it benefits data management and processing. Also, this naming mechanism makes protocol identification much convenient.

Protocol identification and conversion method

In order to achieve more concise data management, a uniform protocol within cloud-edge computing paradigm is expected. As discussed above, we mainly consider a piece of information about "who, what, where, when and how". Therefore, those messages carried in different protocols should be abstracted into such uniform one. In this manner, DSEC is able to support different communication protocols, such as IEC 60870 of distribution automation system, IEEE 1547 of EV and LoRa of edge devices, etc. Considering the widespread usage of IEEE 61850 in smart grid, here we recommend to select IEEE 61850 as the basic protocol in DSEC and cloud center of DS, which means that all other protocols existed in DS should be translated to IEEE 61850. Due to the length limitation, this is not discussed in detailed in this paper.

Data management method

With abundant data upstream, the proper data management should be considered as an important issue. We discuss about how to perform various procedures to keep data in managed and useful form from three aspects: data quality verification, data storage and data abstraction.

A. *Data quality verification*

DSEC acquired data from multiple devices that are in different states, such as normal, abnormal or broken. The quality of the acquired data depends on the state of the devices. If incomplete and incorrect data mix up in the database, data-based application will provide wrong information to the users so as the decision made by them. Therefore, the data quality verification model could detect abnormal data according to the historical data record and data logical relationship. What's more, it could further analyse the reason for missing data or incorrect data, which also could be useful for users to detect device failure and even attack outside.

B. Data storage

Data storage is one of the fundamental part of data management. Based on information technology such as HDFS, data could be store in the form of flat file, text file, spreadsheet, or table. With data structures that reflect the relations between indexes, key data entities and so on, the complicated relational database could be available to process data.

C. Data abstraction

Data is stored in different containers according to its security and privacy. Applications running on DSEC should be isolated from this containers, which means they should abstract raw data from the containers with permission. It is the first step that DSEC knows what the applications need or what they are interested in. Then the called service for abstraction should perform validation process. Without this validation service, it is probable that outsiders can get the details of sensed data from DSEC which means no security and privacy. The abstracted data is ought to be usable and reliable so that the applications could learn enough knowledge from DSEC.

IV. DATA PROCESSING METHOD FOR DS BASED ON EC

Virtualized infrastructure allocation

To ensure isolation of applications and healthy utilization of physical resources, such as computation, storage and networking, cloud providers typically make use of virtualization technology to create virtual machines (VM), which is also deployed in our proposed DSEC.

Both DSEC and cloud center are physical machines (PM) except that the DSEC is one PM while DC manages several. A PM is able to accommodate more than one VM as long as the total resources of all VMs are within its resource ability. Each VM contains its own operation system, computation, storage and network resources which can all be dynamically configured by a VM manager (VMM). Afterwards, certain applications in DS like DR can be processed in the VM. Since all VM can operate independently, data can be analysed insolently and applications can be processed simultaneously. After finishing the application, VM can be shut down and return its resource back to the resource pool in PM, waiting for the next resource request.

As shown in Fig.3, initially, edge applications with different requirements of quality of service (QoS) and service level agreement (SLA) are mapped into different VMs. QoS represents an application's real-time, reliability and safety constraints while SLA means the agreement between service provider and its customers. Then the VM resource allocation would output a possible scheme after analysing the PM resource application characteristics. Lastly, VMs with certain sizes are activated to process corresponding applications.

A well-allocated VM method can efficiently deliver appropriate resources for applications without violating the SLAs, reduce process delays and increase resource utilization [11]. Therefore, an optimization method to allocate virtualized infrastructures in DSEC is of great importance.

Content delivery

As mentioned above, data can be isolated among VMs and edge computing can prevent privacy leaky. Here we propose a content delivery method that be utilized in DSEC to achieve better resource utilization.

For some applications, their data collection frequency and correspond application process frequency may not be the same. For instance, EV charger status are collected periodically, but the EV charging strategy application is activated only when a driver' request is arrived. Therefore, we divide the VM into two categories, the first one is processing VM discussed in the last subsection, and the second one is storage VM. Considering the real-time and safety properties of different data, a four dimension content delivery method is illustrated in Fig.4.

The first dimension relates to the real-time property which includes real-time (RT) and non-real-time (RT). The vertical axis denotes the safety property in terms of high safety (HS) and low safety (LS) requirement. Any data that collects into the DSEC would be delivered into different storage VM based on its real-time and safety constraints. When an application is activated, a processing VM is then generated and retrieve necessary data from different storage VM. Similar to the power grid dispatching network security management, different categories of storage VM in DSEC requires different data manage methods.

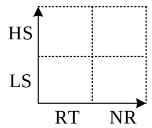

Fig.4 A 4-dimension content delivery method for DSEC

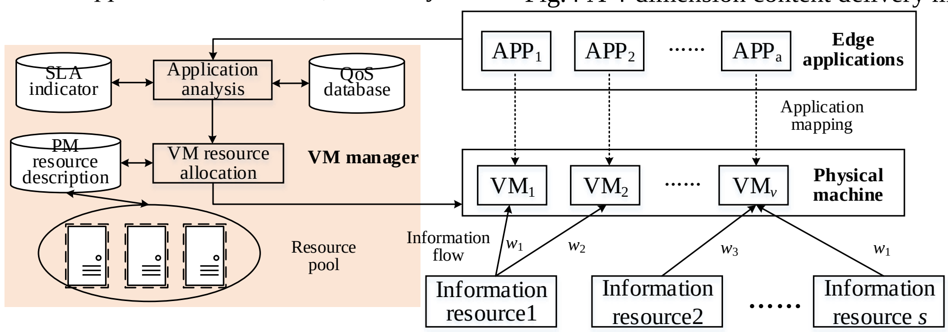

Fig.3. The cloud-edge architecture scheme for DS

Task offloading

Edge computing exhibits a better trade-off between delay-sensitivity and computing-intensive applications by offloading some applications to the remote cloud. The DSEC has the ability to select certain tasks to be offloaded and processed on cloud center. The rest of tasks are supposed to be processed locally. The main design issues are where, when, what and how to offload applications from DSEC to cloud center to meet SLA [12].

One critical factor of cloud-edge computing is the computation load profile characteristics of computing applications. In the DS, any operation and control applications can be divided into three categories: 1) constant and continuous applications, such as equipment monitoring, network analysis, etc; 2) continuous but fluctuating applications like EV charging or demand response management; 3) fluctuating and shiftable applications, for example, equipment fault diagnosis.

Considering the current computing resource and the characteristics of computation applications that need to be processed at time *t*, an optimal task offloading is then required to perform customized resource allocation, which should include the optimal processing VM provision and the exact applications that should be processed in edge VMs or cloud VMs.

Typically, reserved processing VMs can be provided for those constant computation applications (category 1 that mentioned above). For those fluctuating applications in category 2 and 3, on-demand VMs strategy is needed.

V. CONCLUSION

To manage massive and heterogeneous data generated in the distribution system, this paper proposes a cloud-edge architecture and design an edge computing device (DSEC) for DS. The data collecting and processing methods of DSEC are also illustrated. In order to fully exploit the benefits of edge computing in DS, we discuss a computing resource allocation scheme by introducing virtualized machine (VM). Data is stored in different VMs based on a 4-dimension content delivery method we proposed. Operation and control applications in DS are computed in reserved or on-demand VMs according to their computation load characteristics and available computing resources. The proposed data collecting method is practical and necessary when facing complicated data access requests in DS. And the proposed data processing method based on virtualized technology can reduce service time and increase resource utilization.

Acknowledgments

This research is supported by the Science and Technology of China Southern Power Grid. The authors cooperated with Guangzhou Suihua Energy Technology Co.,LTD to develop an edge computing terminal prototype for distribution system.

REFERENCES

[1] LASSETER R H, 2011, "Smart distribution: coupled microgrids", *Proceedings of the IEEE*, vol.99, 1074-1082.

[2] Z. Guan, J. Li, L. Wu, et al, 2017, "Achieving efficient and secure data acquisition for cloud-supported internet of things in smart grid", *IEEE Internet Things J.*, vol.4, 1934-1944

[3] Hayes, B, 2008, "Cloud computing", *Communi-cations of the ACM*, vol.51, 9-11.

[4] N.D. Popovic, D.S. Popovic and I. Seskar, 2018, "A novel cloud-based advanced distribution manage-ment system solution", *IEEE Trans. Indust. Info.*, vol.14, 3469-3476.

[5] Y. Melike, V. C. Gungor and B. Sel ҫuk, 2014, "Cloud computing for smart grid applications" *Computer Networks*, vol.70, 312-329.

[6] B. Samaresh, S. Misra and J. Rodrigues, 2015, "cloud computing applications for smart grid: a survey." *IEEE Trans. Parallel & Distributed Syst.*, vol.26, 1477-1494.

[7] N. Abbas, Y. Zhang, A. Taherkordi, et al, 2018, " Mobile edge computing: a survey " *IEEE Internet Things J.*, vol.5, 450-465.

[8] W. Shi, J. Cao, Q. Zhang, et al, 2016, "Edge computing: vision and challenges", *IEEE Internet Things J.*, vol.3, 637-646.

[9] T. G. Rodrigues, K. Suto, H. Nishiyama, et al, 2017, "Hybrid method for minimizing service delay in edge cloud computing through VM migration and transmission power control", *IEEE Trans. Computers*, vol.66, 810-819.

[10] C. You, K. Huang, H. Chae, et al, 2017, "Energy-efficient resource allocation for mobile-edge computation offloading", *IEEE Trans. Wireless Communi.*, vol.16, 1397-1411.

[11] C. Guerrero, I. Lera, B. Bermejo, et al, 2018, "Multi-objective optimization for virtual machine allocation and replica placement in virtualized Hadoop", *IEEE Transactions on Parallel & Distributed Syst.*, vol.29, 2568-2581.

[12] Y. Hao, M. Chen, L. Hu, et al, 2018, "Energy efficient task caching and offloading for mobile edge computing", *IEEE Access*, vol.6, 11365-11373.