Energy Big Data Application based on Energy Big Data Center

Jia Liu, Jiangbo Sha, Zhenhua Yan, Rui Ma, and Dongge Zhu
Electric Power Research Institute of State Grid Ningxia Electric Power Co.,Ltd. Yinchuan Ningxia, 750002,China
y_l_321@163.com

ABSTRACT
Energy big data has great potential application value in promoting the flow of energy resources on demand and the optimization and transformation of energy structure. Based on the investigation of national energy information system and energy industry enterprise big data platform, the development orientation of energy big data application center is defined, and the architecture design, security protection and data whole process control scheme of "three platforms and two systems" are proposed. This paper introduces the functional design scheme of energy big data application scenario for the government, enterprises and the public.
Keywords: *Energy big data; Architecture design; DATA Governance*

1 INTRODUCTION

With the popularization of information collection system, the amount of data in the energy field is increasing explosively. The rapid development of digital information technology makes it possible to mine and analyze energy big data. Mining the economic and social values hidden behind big energy data, promoting the optimization and transformation of energy structure and the on-demand flow of energy resources have become a powerful starting point for promoting the reform of China's energy industry system[1]. In order to meet the needs of the development of new energy industry, it is urgent to design the functional system and application scheme of energy big data application center, and promote the integration and sharing of data resources in the energy industry. Thanks to the installation of smart meters, the deployment of information systems such as dispatching data acquisition and monitoring system and user power information acquisition system, the digital information level of power grid enterprises is in the forefront of enterprises in the energy industry. In recent years, the mining and analysis of power big data has gradually become a hot spot in the industry[2]. At present, the research on its architecture and function positioning mostly starts from the power grid enterprises themselves, and there is relatively little comprehensive application analysis of cross domain, cross business and cross department. The research on the architecture, operation mode and data acquisition mechanism of energy big data application is still in its infancy.
Energy big data center is an important carrier to promote the deep integration of energy revolution and digital revolution. It is an energy data aggregation center, Innovation Incubation Center and shared service center. Among them, the energy data aggregation center will realize the whole link and chain aggregation access, safe storage and overall management of all kinds of energy data; The Innovation Incubation Center will focus on the needs of the government, enterprises and the public, carry out energy data value mining, and serve the scientific decision-making of the government, the coordinated development of digital economy, the Lean Development of enterprises, intelligent operation and the convenient use of energy by the public; The shared service center will open energy big data resources to the outside world, realize energy product service innovation, core technology innovation and business model innovation, make the achievements benefit more enterprises and people, and build an innovative, open, cooperative and shared energy big data ecology[3]. We studied the functional system design and application scheme of the energy big data application center, proposed the functional system design scheme, including platform design, security system design, data management and application scenario development, and constructed the energy big data diversified analysis application system [4].

2 RELATED WORK

At present, the application of energy big data in electric power, new energy, oil and gas enterprises has been widely carried out, but there is no mature mode at the government and industry level. The whole country and some provinces have carried out the construction of

energy information system in advance. For example, the national energy administration has built a national energy monitoring and early warning and planning management system, which can initially realize the daily monitoring of energy operation and dynamic evaluation of energy planning; Hubei, Hebei, Inner Mongolia, Zhejiang, Jiangxi, Guangdong and other provinces have also built their own energy information systems, whose functions are basically consistent with the national energy system; In view of the characteristics of abundant wind and light resources in Qinghai Province, relying on Qinghai electric power company, Qinghai Province has built a new energy big data innovation platform to provide planning, construction, operation and maintenance services for the government and related enterprises.

Rodolph et al.[1] proposed energy-saving cross-layer optimization of big data transfer based on historical log analysis. Ahmad w a proposed an energyefficient big data workflow scheduling algorithm under budget constraints for heterogeneous cloud environment[2]. Zhu j et al. proposed bigdata assisted energy conversion model for innovative city application[3]. Jeong s y et al. proposed development and application of a big data analysis-based procedure to identify concerns about renewable energy[4]. Pak w et al. proposed proposal of the energy consumption analysis process for the residential houses using big data analytics technique[5]. Jia y et al. proposed construction of energy scheduling model for iron and steel enterprises based on big data[6]. Chen h et al. new energy generation forecasting and dispatching method based on big data[7]. Khare v et al. baredar p proposed assessment and decision-making of biomass energy conversion system by big data and game theory technique - sciencedirect[8]. Zhai y et al. proposed green internet of things and big data application in smart cities development[9]. Monga k et al. proposed a dual mode in-memory computing unit using spin hall assisted mram for data-intensive applications[10].

This paper consists of the following parts. The first part introduces the related background and significance of this paper, the second part is the related work of this paper, and the third part is data analysis. The fourth part is example analy-sis. The fifth part is conclusion.

2.1 General idea

For the government: achieve data aggregation in the energy field, meet the government's demand for energy planning, operation situation and other industry management and supervision, and help energy transformation. Enterprise oriented: promote energy enterprises to improve operational efficiency, achieve lean management, and promote industrial enterprises to improve energy efficiency. Facing the public: meet the public's demand for convenient and efficient energy use, and promote the public to reduce the cost of energy use [6].

2.2 overall structure

The overall architecture of the energy big data application center is shown in Fig 1. It follows the design of "three platforms and two systems" and serves three types of objects: government, enterprises and the public. The three platforms are "software and hardware infrastructure platform, data management platform, and application crowdsourcing platform". The software and hardware infrastructure platform provides basic resource services, storage, resource management, and servers for the energy big data application center by building the energy big data cloud platform; Through public service components and developer community, the data management platform provides decision support, management consultation, convenient energy use and other services for different types of users[7]. The two systems are "security protection system" and "operation management system", in which the security protection system provides four levels of security guarantee, namely, technology, management, credibility and service; Operation management system includes setting up efficient operation organization, mining data value and ensuring sustainable development.

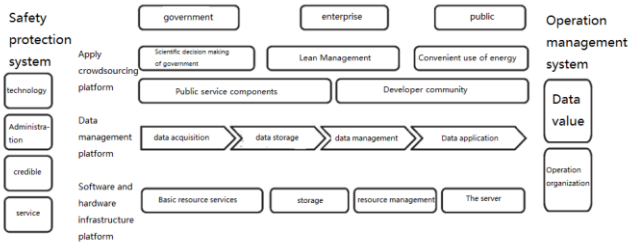

Fig. 1. Overall architecture of energy big data application center

2.3 Physical architecture

The physical architecture of energy big data application center is divided into security management area, energy big data production area (including development and testing area), security access area, communication channel area and data acquisition area. The security management area is mainly used to deploy network management equipment and operation and maintenance equipment to realize network security and operation and maintenance management; The production area is main-

ly used for deploying energy big data application cloud platform, management services, applications and data management components; The safety receiving area ensures the safety of information intranet communication by adding safety protection equipment; In the communication channel area, Internet is used to realize data remote communication; The data collection area mainly realizes the safe collection of power intranet, other energy industry cloud and government data, meteorological data[8].

3 DATA ANALYSIS

The construction of security protection system is the key and difficult point in the construction of energy big data application center. Based on the design of security protection system, the corresponding data security collection, sharing and interaction scheme should be developed.

3.1 Safety protection system

According to the latest requirements of "Regulations on classified protection of network security" and referring to the big data security capability maturity model, the security protection system construction of energy big data application center is carried out from three aspects of "management, technology and operation".
Information security management system includes security management system, emergency plan, etc. According to the relevant management requirements of the country, industry and enterprise, the data security strategy, specification and effectiveness measurement mechanism are established, the big data security management organization is set up, and the data asset security management specification, classification standard, method, process, guide, mechanism and checklist are formulated. In view of service interruption, large-scale virus infection, data leakage and other major security risk scenarios, emergency plans are formulated and relevant training exercises are carried out.
Information security technology system around the network communication layer, platform layer, data layer, application layer. Adopt cloud monitoring, dual link, load balancing and other strategies, deploy firewall, system vulnerability scanning system and security risk analysis management system to improve network security protection ability. Through the virtual network micro isolation protection software, virtual firewall, endpoint detection and response system, the security protection of virtual resources is realized. To develop a unified data security strategy, establish a data access behavior risk monitoring system, deploy data desensitization system and isolation device, and realize risk early warning and risk control for data assets.
Information security operation system includes security services, monitoring services, infrastructure security, etc. Provide data security suite for data tenants and developers, and set up security service team to provide security detection, security assessment and security consulting services. Intelligent monitoring engine is designed to monitor data flow and abnormal events in real time, and cooperate with various security resources to respond quickly. Various protection means such as identity authentication, access control and data encryption are adopted, and multi-level application level intrusion prevention system is deployed.

3.2 Data security protection scheme

Design a data application user full link data security protection scheme, including:
(1) Data asset identification and management: systematically identify the internal data assets and the classification of data assets in the data center, understand the overall distribution of data assets, and formulate a unified data security strategy.
(2) Data behavior monitoring: establish a unified data risk behavior monitoring for the main data access behavior in the data center, including data access through application, interface call, database management, operation and maintenance personnel's access to the database.
(3) Comprehensive data risk control: conduct comprehensive risk analysis combined with data assets and data behavior, and give early warning of risks and take real-time risk control measures.
(4) Data desensitization: take data desensitization measures to ensure the data desensitization security of energy big data application center and meet the demand of data desensitization business. The data desensitization system is deployed in the data storage and computing network area of the energy big data application center, and the deployment mode can reach the network.
Energy enterprises transmit data to the energy big data application center: the data source collects the data to the enterprise front-end processor for data file encryption, and transmits it to the isolation area of the energy big data application center through the Internet. The front-end processor receives the data uniformly, and the front-end processor transmits the received data to the internal network unified buffer of the energy big data application center for data decryption and verification, Finally, it is stored in the target database.
The energy big data application center shares data with energy enterprises: the energy big data application center first carries out data desensitization and file encryp-

tion, and then transmits the data to the enterprise front-end processor through the front-end processor in the isolated area of the energy single data application center for data file decryption and verification, and stores the data in the target database after decryption and verification [11].

4 EXAMPLE ANALYSIS

It is necessary to comb the data directory, determine the data source and access mode, and determine the data security and permission management process in combination with the application R & D requirements.

4.1 Data directory

The data catalog covers the whole process data of production and supply, consumption and investment, resource transfer and utilization efficiency of coal, oil, natural gas, electric power, new energy and other energy resources, as well as cross sectoral and cross domain data of macroeconomic operation, ecological environment, meteorology, geographic information and transportation, mainly including macro level, energy industry and other relevant data. The macro level data include the provincial macro-economic operation, development planning, industrial policy, system reform, market development trend and other data, as well as the world's major countries and regions, the country, advanced provinces and other economic and social, energy development data[12]. The data of energy industry includes the whole process data of resource endowment, mining and processing, transportation and distribution, energy conversion and energy consumption of electricity, coal, oil, gas, new energy and other energy categories. Other relevant data include ecological environment, meteorology, geographic information, transportation, technological innovation, industrial price, etc.

4.2 Data sources

Data sources include government departments, energy companies, the Internet and other channels. Macro level data mainly comes from relevant government authorities or research institutions, official policy documents, Statistical Bulletins, research reports, as well as the regular acquisition of relevant data from international authoritative data statistics platforms establish a fixed information submission mechanism, and regularly report relevant energy information through the information submission system. Other relevant data mainly come from the public data of relevant departments. The docking and collection mechanism of departments and units is established to realize the regular submission of relevant data, as shown in Fig 2.

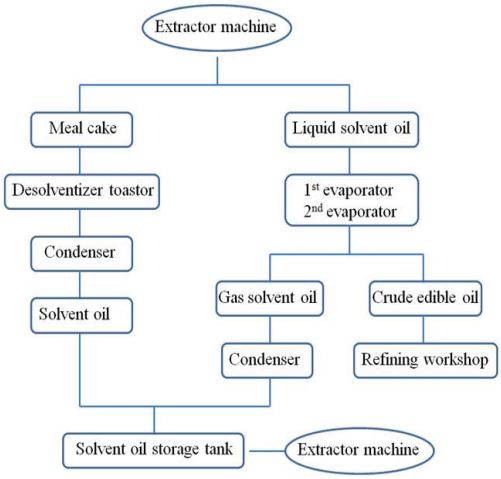

Fig. 2. "Internet plus" service convenience tree

4.3 Data access

According to different data types and sources, the energy data receiving methods are divided into three types.
1) Data submission: for all levels of government departments and energy related enterprises, a fixed information submission mechanism is formed, and energy information is regularly reported online through the information submission system. The submission process adopts the multi-level audit process to ensure the accuracy of the data.

$$\|\Delta x_{k+1}(t)\| \le \left(pk_f + m_2 + m_3\right) \int_0^t \Delta x_{k+1}(\tau)d\tau + \int_0^t (m_1 \|\Delta u_k(\tau)\| + pd)d\tau \quad (1)$$

$$\|\Delta u_{k+1}(t)\| \rho \|\Delta u_k(t)\| + m_4 (\int_0^t e^{(pk_f+m_2+m_3)(t-\tau)} (m_1 \|\Delta u_k(\tau)\| + pd)d\tau) + m_5 d \quad (2)$$

$$\|\Delta u_{k+1}(t)\| e^{-\lambda t} \le \rho \|\Delta u_k(t)\| e^{-\lambda t} + m_5 d e^{-\lambda t} + m_4 (\int_0^t e^{(pk_f+m_2+m_3)(t-\tau)} e^{\lambda(t-\tau)} e^{-\lambda t} (m_1 \|\Delta u_k(\tau)\| + pd)d\tau) \le \rho \|\Delta u_k(t)\| e^{-\lambda t} + m_5 d + m_1 m_4 \int_0^t e^{(pk_f+m_2+m_3-\lambda)(t-\tau)} e^{-\lambda t} \|\Delta u_k(\tau)\| d\tau + pdm_4 \int_0^t e^{(pk_f+m_2+m_3)(t-\tau)} d\tau \quad (3)$$

2) Data capture: for various public information sources published on the Internet, the information crawling, cleaning and entering are completed by means of big data.

$$\|\Delta u_{k+1}(t)\|_\lambda \le \rho\|\Delta u_k(t)\|_\lambda + m_5 d + m_1 m_4 \|\Delta u_k(t)\|_\lambda \frac{1-e^{(pk_f+m_2+m_3-\lambda)t}}{pk_f+m_2+m_3-\lambda} + pdm_4 \frac{1-e^{(pk_f+m_2+m_3)t}}{pk_f+m_2+m_3} \le \left(\rho + m_1 m_4 \frac{1-e^{(pk_f+m_2+m_3-\lambda)t}}{pk_f+m_2+m_3-\lambda}\right)\|\Delta u_k(t)\|_\lambda + \left(m_5 + pm_4 \frac{1-e^{(pk_f+m_2+m_3)t}}{pk_f+m_2+m_3}\right) d \le \left(\rho + m_1 m_4 \frac{1-e^{(pk_f+m_2+m_3-\lambda)t}}{pk_f+m_2+m_3-\lambda}\right)\|\Delta u_k(t)\|_\lambda + m_5 d \quad (4)$$

3) System access: for other information systems running on the Internet or government extranet, data interface shall be established according to the data access mode and content standard agreed by both parties to realize data docking. Data access is shown in Fig 3.

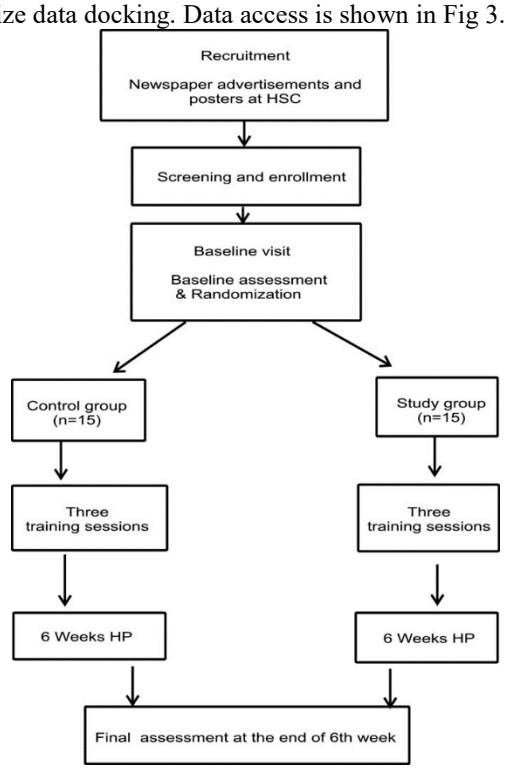

Fig. 3. Data access

5 CONCLUSION

For core data assets, a safe and reliable protection system and corresponding data security collection, sharing and interactive whole process control scheme are set up. We will promote the integration and integration of energy big data in an orderly manner, break the data resource barriers, innovate new regulatory models for the energy industry, improve the timeliness and accuracy of energy statistics, analysis, prediction and other businesses, and foster new forms of energy industry development.

REFERENCES

[1] Rodolph L , Nine S , Tacchio L D , et al. Energy-saving Cross-layer Optimization of Big Data Transfer Based on Historical Log Analysis[J]. 2021.

[2] Ahmad W , Alam B , Atman A . An energyefficient big data workflow scheduling algorithm under budget constraints for heterogeneous cloud environment[J]. The Journal of Supercomputing, 2021.

[3] Zhu J , Vadivel T , Sivaparthipan C B . Bigdata Assisted Energy Conversion Model for Innovative City Application[J]. Journal of Interconnection Networks, 2021.

[4] Jeong S Y , Kim J W , Joo H Y , et al. Development and Application of a Big Data Analysis-Based Procedure to Identify Concerns about Renewable Energy[J]. Energies, 2021, 14.

[5] Pak W , Inhan K , Jungsik C . Proposal of the energy consumption analysis process for the residential houses using big data analytics technique[J]. Journal of Computational Design and Engineering, 2021.

[6] Jia Y , Jiang C , Yang J , et al. Construction of Energy Scheduling Model for Iron and Steel Enterprises Based on Big Data[M]. 2021.

[7] Chen H , Zhang H . New energy generation forecasting and dispatching method based on big data. 2021.

[8] Khare V , Nema S , Baredar P , et al. Assessment and decision-making of biomass energy conversion system by big data and game theory technique - ScienceDirect. 2021.

[9] Zhai Y , Liu J , Faqiri H , et al. Green Internet of Things and Big Data Application in Smart Cities Development[J]. 2021.

[10] Monga K , Chaturvedi N , Gurunarayanan S . A Dual Mode In-Memory Computing Unit using Spin Hall assisted MRAM for Data-Intensive Applications[J]. IEEE Transactions on Magnetics, 2021, PP(99):1-1. He Hongyan, research on Key Technologies of smart grid based on big data [J]. Power technology, 2016,40 (8): 1713-1714.
[11] Hu Yanling. Research on big data transaction status and pricing [J]. Price monthly, 2017 (12): 16-19