Prognostics and health management system for hydropower plant based on fog computing and docker container

Jian Xiao[1,4], Mingqiang Zhang[2], Haiping Tian[1], Bo Huang[1] and Wenlong Fu[3]

[1] Hunan Electric Research Institute of State Grid Corporation, Changsha Hunan, China;
[2] Hunan Engineering Company of State Grid Corporation, Hengyang Hunan, China;
[3] College of Electrical Engineering and New Energy, China Three Gorges University, Yichang, China.

[4] xiaojianhn@163.com

Abstract. In this paper, a novel prognostics and health management system architecture for hydropower plant equipment was proposed based on fog computing and Docker container. We employed the fog node to improve the real-time processing ability of improving the cloud architecture-based prognostics and health management system and overcome the problems of long delay time, network congestion and so on. Then Storm-based stream processing of fog node was present and could calculate the health index in the edge of network. Moreover, the distributed micros-service and Docker container architecture of hydropower plants equipment prognostics and health management was also proposed. Using the micro service architecture proposed in this paper, the hydropower unit can achieve the goal of the business intercommunication and seamless integration of different equipment and different manufacturers. Finally a real application case is given in this paper.

1. Introduction

In the recent years, hydropower, as one of the most important green and renewable energy in the world, has a rapid development in China [1]. A failure in hydropower generating set (HGU) results in not only the catastrophic loss of economic consequence but also lead to safety and environmental problems. Therefore, the requirement that the HGU should work under the premise of high reliability, low environmental risk and personal safety has been rapidly increasing. Operational condition of hydropower equipment has a crucial impact on the economic benefits and stability of hydropower plant. These effective maintenance strategies are more and more being taken into account and as far as possible to minimize the maintenance cost. And the HGU equipment could ultimately be kept with overall healthy and stable operation [2].

Condition-based maintenance (CBM) strategy is new and popular industrial maintenance mechanisms which perform maintenance depend on the observation of the "health condition indicator" of the HGU and its components. The CBM could enable real-time diagnosis of impending failures and prognosis of future equipment health. Prognostics and Health Management (PHM) is one of important aspects in CBM. PHM aim to identify the performance degradation condition which depends primarily on health condition parameters and predict incipient faults and estimated Remaining Useful Life (RUL) of mechanical components using heterogeneous sensor data. Nevertheless, it is a challenging task to apply the PHM technology into the practical application. On one hand, traditionally hydropower plant often uses standalone condition monitoring server and fault diagnosis system to handle all the data

processing and computing and it could not adapt the current changeable environments. To access the remote fault diagnosis server with the fiber, enormous human and material resource could be cost [3]. On other hand, this method could not share the expert experience and failures case.

Therefore, there is an urgent demand on designing a scalable and extensible data processing framework and ensuring reliable transmission of monitoring data under the high-speed acquisition and fluctuation environmental. Although large-scale cloud fault diagnosis system has been partially applied in hydropower plants, and could partly meet the requirements of users to some extent, but it also faces great challenges on the burden of exploding amount of big data [4].The sensor data stream in hydropower plant exhibits a new characteristic: low-latency, which requires that the fault diagnosis could analyses the equipment condition on the spot timely to prevent equipment from unexpected breakdown and failures. Furthermore, numerous mobile devices are increasingly accessing applications and services from the cloud. The drawback of this model is that when many users simultaneously request services from the data server, bandwidth limits may lead to data congestion and even server crashes [5-7].
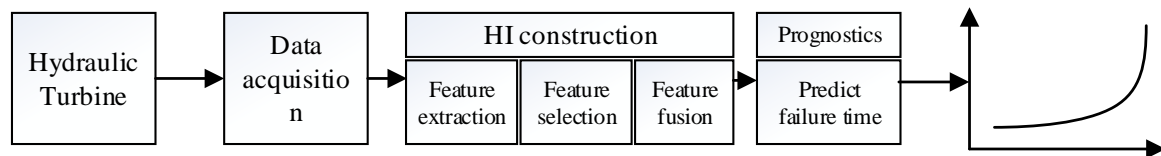
To tackle these problems, we proposed a fog computing architecture-based PHM system for hydropower plant. Contrary to the cloud computing, the system uses fog nodes to perform delay sensitive applications. The fog computing has the advantages that the computing resources are closer to the consumption of resource, and tries to minimize the impact of fog calculated by processing the application and close to the user's data, which has obvious advantages in the following aspects: 1) Connectivity between physical devices and network infrastructure could use. Fog nodes can be adapted to a variety of access methods, such as RS-485, TCP/IP, or hard wiring, etc. 2) Low network latency. Fog computing reduces the network latency by deploying computing devices at geographically closer servers on the edge of the network, and collects and stores sensor data in fog nodes. The hydropower station can collect and analyse a large amount of data on the local edge nodes, and no need to upload all the original data to the server or store. The significant reduction in data transmission can solve the bandwidth bottleneck problem. 3) Ubiquitous and instant remote access to near real-time data without spatial constraints. Users can use desktop computers or mobile devices and other platforms to access the entire server's data. 4) Secure and high volume data storage. The fog computing node can be a large number of local data storage, cloud storage only features, data indexing and call when necessary, greatly improves the data security and storage efficiency.

The main contribution of this paper is firstly proposed, prototypical 4-layer fog-based computing architecture was proposed to demonstrate the effectiveness and the feasibility of the system in the paper. The PHM system comprises three components: terminal device, fog node and public cloud server. Because fog node is the key part of fog diagnosis system, a flow computation framework is proposed to deal with real-time massive state data in fog nodes.

This next part of paper is organized as follows. Section 2 details the Storm architecture and fog node. Section 3 introduced real-time fault diagnosis system with a practical application. Finally, a brief conclusion is offered in Section 4.

2. Basic theory of the PHM
As illustrated in figure 1, the framework of the PHM methods is generally based on the following three steps: (1) data acquisition, (2) construction of health indicators (HIs), and (3) prognostics (i.e., prediction of a failure time). Among these steps, the performance of HIs has a crucial effect on prediction accuracy of RUL. In literature, synthesized HIs have received much attention in recent years. The synthesized HIs are usually constructed through some data fusion techniques, where multidimensional statistical features, e.g. root mean square (RMS), variance, kurtosis, etc., are transformed into an one dimensional HI .
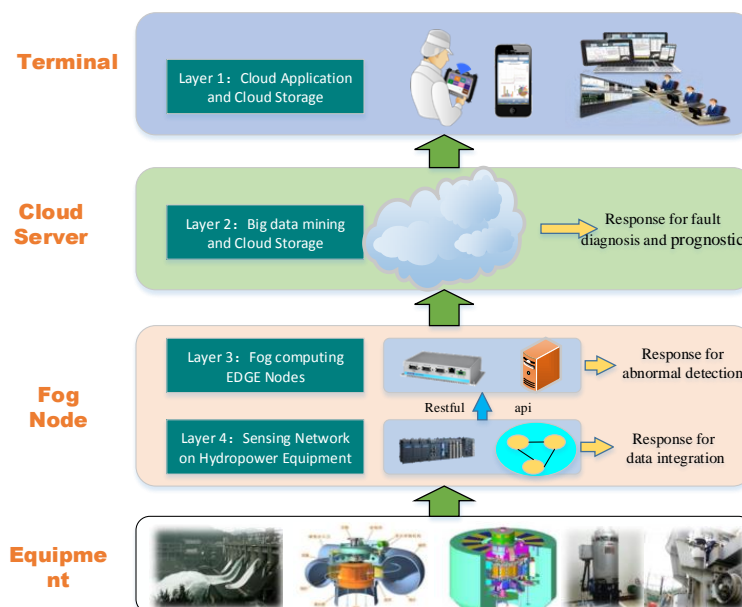
Figure 1. The framework of PHM.

3. The Hierarchical distributed architecture of fog computing platform for hydropower plant

3.1. Hierarchical distributed architecture

Hierarchical distributed architecture based on fog computing is the next generation hydropower equipment condition monitoring and fault diagnosis network. The power plant can obtain diagnostic services through on-demand calculations, mobile applications, and parallel machine learning tools.
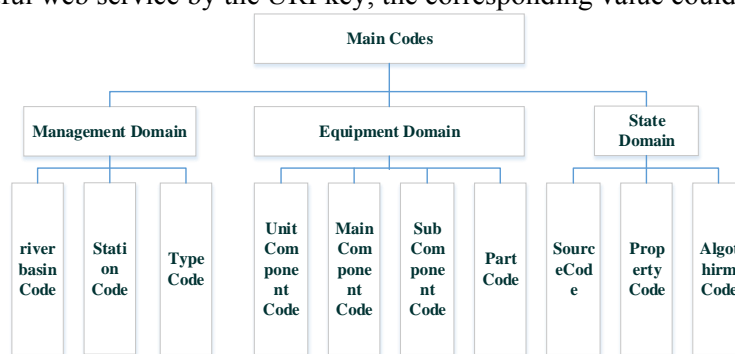
With the extended framework of cloud services, fog computing based PHM systems have superior benefits. From a hardware point of view, fog based computing architecture can solve the problem of persistent lack of affordable sensing technology, which can be easily integrated into different IT application systems. On the one hand, fog based computing architecture can solve the problem of persistent lack of affordable sensing technology, and it can be easily integrated into different IT application systems. On the other hand, the parallel PHM algorithm has sufficient computational power and bandwidth for high speed. Although high performance computing servers can be deployed on the cloud, training and storage of all raw data are very expensive. In view of the above shortcomings, four fog calculation based on PHM system is proposed in this paper. Fog computing architecture allows the system to collect a large number of real-time streaming data and monitor the health status and process of the machine. The architecture of the four layer fog calculation is shown in figure 2.
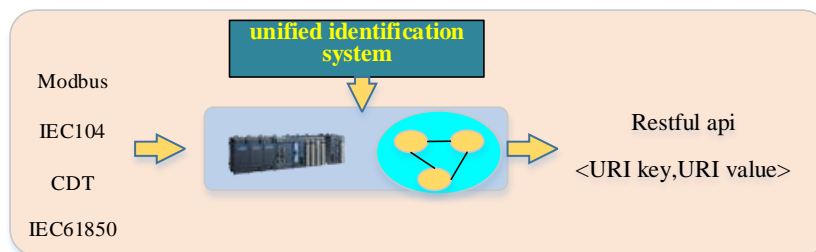

Figure 2. Architecture of system.

At the edge of the network, the fourth layer is a sensor network that contains many sensory nodes. These sensor nodes monitor the state changes of hydropower devices and transfer them through monitoring system and SCADA system. Since all of these heterogeneous data sources come from different devices and protocols, such as IEC104, Modbus, IEC61850, CDT, etc., this paper establishes a unified information coding strategy using RESTful communication protocol with Unified identification system for power plant equipment. Since all these heterogeneous data source from

different equipment and protocol such as IEC104, Modbus, IEC61850, CDT, etc., this paper developed a unified information encoder strategy using the restful communications protocols and unified identification system for power plant equipment. Unified identification system for power plant equipment consists of three parts: Management Domain, Equipment Domain, and State Domain [13]. The Management Domain represents the information of management hierarchy. Equipment Domain could locate the details part of data source. State Domain is often related with the computing method of performance indicator. The unified identification system for power plant equipment is shown in Figure 3. As is shown in the figure 4, once the edge of network receives the raw data, it could parse the real value according to the description of state domain code. Then the edge of network converts the real value into "key-value" format. And all the data could be defined as the restful Uniform Resource Identifier (URI) which consists of a URI key and a URI value. Once the client initiated request with restful web service by the URI key, the corresponding value could be required.

Figure 3. The unified identification system for power plant equipment.

Figure 4. The protocol integration process using the unified identification system and restful API.
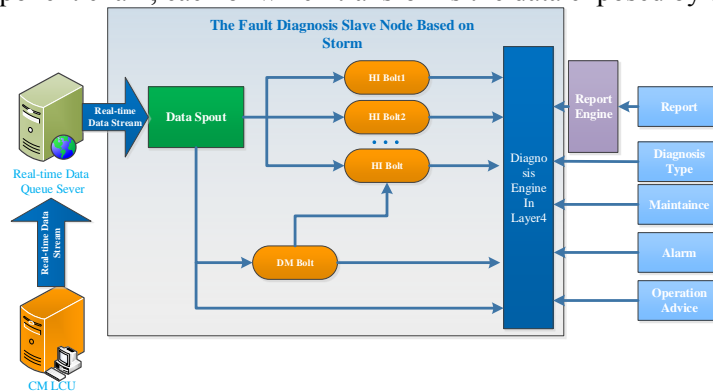
The nodes at the edge forward the condition information with restful **Uniform Resource Identifier (URI)** into the layer 3, which aims to compute the performance indicator based on the Storm-based stream computing. The condition information is feed into the computing unit of Storm to compute the performance indicator. The output of performance indicators has two parts: the first are reports as the fault feature to the fault diagnosis Cloud server at its next upper layer, while the second is simple and quick feedback alarm to failure equipment for preventing from unexpected break down.

The second layer is a cloud server, which is composed of multiple application services. Each application service is related to the performance index of the third layers, and the potential fault reasons are identified by combining historical temporal data and related event instances. At the same time, the rapid response to the fault, such as friction, cracks and other fault detection equipment. In addition, cloud computing and cloud storage in this layer provide complex and long-term big data analysis, such as large-scale event queries, long term pattern recognition and relational modelling.

3.2. The Fog Computing Node based on Storm
In this section, we detail the core implementation of fog computing nodes based on storm framework. As the core framework of Twitter, storm is a distributed, robust, fault-tolerant massive data stream processing system, which shows superior performance in real-time monitoring and computing. Storms have different types of components, each of which is responsible for a simple specific processing topology. There are two kinds of computing units in the storm frame system, called the master and the worker. The work of storms is divided into two categories: spout and bolt, which are responsible for a simple specific fault diagnosis task, such as feature extraction and fault

classification. The task of the bolt module is to process the input stream of the storm cluster. As shown in figure 5, the spout passes through the monitoring data bolt assembly, either adhering to data in certain storage, or through some other health index (HI) bolts. Storm clusters can be viewed as a bolt component chain, each of which transforms the data exposed by the spout.
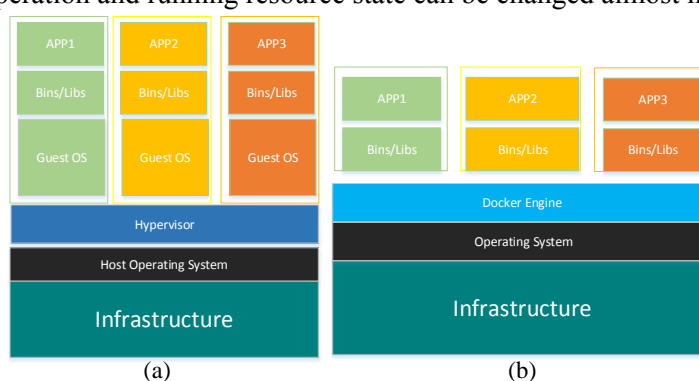

Figure 5. The topology model.

The spout setting should describe the KKS code to acquire the monitoring stream by querying from the "Key-Value" pairs in the restful URI. Then the spout pass the data need into each blot which is embedded with the performance indicator model. All these models could be cascaded and weighted composed.

3.3. Prognostics and Health Management Cloud Service based on Micro Service Architecture and Docker Container

In the Layer2, We developed the Prognostics and Health Management distributed cloud service (PHMDCS) based on the micro-service architecture and Docker container. Micro-services architecture could improve productivity effect through maximizing the automation in all life circles of PHM service. However, micro-services architecture approach also introduces a lot of new complexity and therefore the Docker container technology has been applied to implementing micro-services architecture. Figure 6 shows the main difference between traditional virtual machine technology and Docker technology. Docker which is container based could avoid the resource waste due to guest OS in its operation and running resource state can be changed almost instantly.
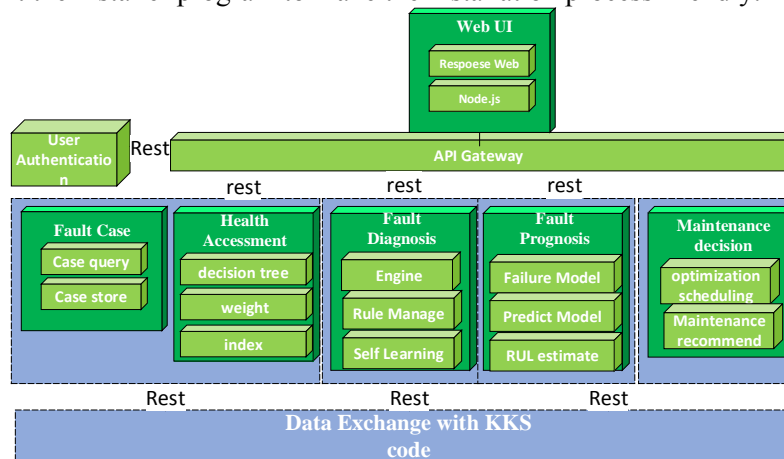

Figure 6. The topology model.

Figure 7 shows the PHMDCS micro service architecture. The PHMDCS client, including user interface and run on the user's Web browser, an application programming interface with PHMDCS server (API) gateway. This gateway covers the complexity of communicating with different services, and provides an integrated API to clients. Micro services include PHMDCS divided into three so-called vertical and related functional group services. In the first vertical phase, we arranged micro services related to simple functional engineering management. Second vertical, we arrange the fault diagnosis related micro service, and third vertical micro service settings and fault prediction and. In the fourth vertical, micro service arrangement, relevant maintenance decisions are utilized by a dock

worker to run each micro service in an isolated container. These containers can be deployed directly to private or public cloud infrastructure. Through the dock window machine, the same container can also run on the desktop computer Mac OS X or Microsoft Windows execution. For these two platforms, we built the installer program to make the installation process friendly.

Figure 7. The micro-service model of PHMDCS.

4. Case study

In this section, we applied the fog computing-based fault diagnosis system into ZhengTouBa hydropower plant, and the master server system is located in the Changsha which linked the slave system by Restful protocol which is shown in figure 8. This equipment health information of ZhengTouBa power plant could be acquired from various sources (e.g. condition monitoring system, SCADA system,etc) and passed into the 3 Layer Fog computing node. Finally, the PHMDCS Server located in ChangSha are computing and identifying the fault type and degree. Then the engineering in power plant could acquire the fault information by mobile phone and PC.

Figure 8. The main UI of PHMDCS

In the PHMDCS system of hydropower plant, the performance health indicators bolt are continuously acquired from the monitoring data spout. When an abnormal alarm is pushed to the diagnosis service and the prognostics service is carried out by the diagnosis bolt. Finally the result message will be sent to the user interface. Once the engineer in the hydropower plant wants to find out

EEEP2017

IOP Conf. Series: Earth and Environmental Science **121** (2018) 042029

IOP Publishing

doi:10.1088/1755-1315/121/4/042029

the deeper of abnormal alarm, the detailed performance health index of each sub-component could be seen in the index analysis module.

5. Conclusions

In this paper, a new fault prognostics and health management system architecture for hydropower plant equipment based on fog calculation and Docker container is proposed. We use fog computing technology to improve the real-time processing capability of Cloud Architecture Based on fault prediction and health management system and overcome the problem of long delay time, network congestion and so on. Then, a storm based data flow processing method is proposed to calculate the health index at the edge of the network. In addition, distributed micro service and Docker container building hydropower plant equipment fault prediction and health management are also proposed. Using the micro service architecture proposed in this paper, the power station can realize the business intercommunication and seamless integration between different equipment and different manufacturers. Finally, the system is applied to the power plant of the ZhenTouBa, and the safety operation of the power plant equipment is given through the system, which can guide the power plant production operation well

References

[1] Lee J, Wu F, Zhao W, Ghaffari M., Liao L and Siegel, D 2014 Prognostics and health management design for rotary machinery systems—reviews, methodology and applications *Mechanical Systems & Signal Processing* pp 313–334

[2] GEBRAEEL N and PAN J 2008 Prognostic degradation models for computing and updating residual life distributions in a time-varying environment *Reliability, IEEE Transactions on* pp 539-550

[3] Tan W P and Xiao J 2016 Real-time degradation assessment system for hydropower equipment based on the streaming data processing *Advances in Power and Energy Engineering: Proceedings of the 8th Asia-Pacific Power and Energy Engineering Conference* pp 173–214

[4] KONDO D, JAVADI B, MALECOT P, CAPPELLO F and ANDERSON D P 2009 Cost-benefit analysis of cloud computing versus desktop grids *Parallel & Distributed Processing IPDPS 2009. IEEE International Symposium on, 2009* pp 1–12

[5] Krishnan Y N, Bhagwat C N and Utpat A P 2015 Fog computing—Network based cloud computing *Electronics and Communication Systems (ICECS), 2015 2nd International Conference on. IEEE* pp 250–251

[6] Yi S Li C and Li Q 2015 A survey of fog computing: concepts, applications and issues *Proceedings of the 2015 Workshop on Mobile Big Data*. pp 37–42

[7] Chiang M and Zhang T 2016 Fog and IoT: An overview of research opportunities[J]. *IEEE Internet of Things Journal* pp 854–864