ITU-T

Technical Report

TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU

(09/2021)

ITU-T Focus Group on Environmental Efficiency for Artificial Intelligence and other Emerging Technologies (FG-AI4EE)

FG-AI4EE D.WG3-03

Data centre energy-saving: Application of Al technology in improving energy efficiency of telecom equipment rooms and Internet data centre infrastructure

Working Group 3 – Implementation Guidelines of Al and Emerging Technologies for Environmental Efficiency

Focus Group Technical Report



FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The procedures for establishment of focus groups are defined in Recommendation ITU-T A.7. ITU-T Study Group 5 set up the ITU-T Focus Group Environmental Efficiency for Artificial Intelligence and other Emerging Technologies (FG-AI4EE) at its meeting in May 2019. ITU-T Study Group 5 is the parent group of FG-AI4EE. Deliverables of focus groups can take the form of technical reports, specifications, etc., and aim to provide material for consideration by the parent group in its standardization activities.

Deliverables of focus groups are not ITU-T Recommendations. For more information about FG-AI4EE and its deliverables, please contact Charlyne Restivo (ITU) at tsbfgai4ee@itu.int.

NOTE

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

© ITU 2021

This work is licensed to the public through a Creative Commons Attribution-Non-Commercial-Share Alike 4.0 International license (CC BY-NC-SA 4.0). For more information visit https://creativecommons.org/licenses/by-nc-sa/4.0/.

Technical Report FG-AI4EE D.WG3-03

Data centre energy-saving: Application of Al technology in improving energy efficiency of telecom equipment rooms and Internet data centre infrastructure

Summary

Telecom equipment rooms and Internet data centre (IDC) infrastructure is a data centre that contains a huge amount of information and communication equipment. In order to keep the equipment running continuously and reliably for a long time, the room is necessarily equipped with air-conditioners to create an environment suitable for equipment operation. Nevertheless, this causes a large amount of energy consumption and carbon emissions. This technical report focuses on the application of artificial intelligence (AI) technology and other emerging technologies such as digital twin technology, to improve the energy efficiency and reduce the carbon emissions of telecom equipment rooms and IDC infrastructures.

Most existing equipment rooms do not have a full ability to identify indoor temperature distribution. Therefore, it is difficult to analyse the power consumption in real-time and make appropriate adjustments in a timely manner. Consequently, this leads to unnecessary consumption of energy. This report will address how AI-based power management can achieve the following capabilities:

- Data collections in telecom equipment rooms and IDC infrastructure;
- Real-time analysis of the historical power consumption and parameters of the target equipment room;
- The ability of training an intelligent model; and
- Making reasonable timely adjustments to the air-conditioning and temperature, so as to achieve energy-saving in the equipment rooms and IDC infrastructure.

Keywords

AI, digital twin, energy-saving, telecom equipment room, IDC infrastructure.

Change Log

This document contains Version 1.0 of the ITU-T Technical Report on "Application of Al technology in improving energy efficiency of telecom equipment rooms and IDC infrastructure" approved at the ITU-T Study Group 5 meeting held on XXX.

Editor:

Shi Ying Tel: +86 20 38639321
China Telecommunications Corp. Fax: +86 20 38639370
China E-mail: shiying@chinatelecom.cn

China E-mail: shiying@chinatelecom.cn

Wu Tong Tel: +86 10-64224092

National Institute of Metrology Fax: +86 10-64204738 China Email: wut@nim.ac.cn

Tan Rumeng Tel: +86 20 38639526
China Telecommunications Corp. Fax: +86 20 38639370
China E-mail: tanrum@chinatelecom.cn

Pan Guanfu Tel: +86 15201457307

National Institute of Metrology, China
China E-mail: pangf@nim.ac.cn

Zheng Pin-Di

Tel: +861851 1976805

Beijing Rainspur Technology Co.,Ltd

Fax: +861068277400

China

F mail: zhennindi@minspur.com

China E-mail: <u>zhenpindi@rainspur.com</u>
Huang Dongmei Tel: +8613601394566

Beijing Rainspur Technology Co.,Ltd Fax: +861068277400
China E-mail: huang@rainspur.com

Table of Contents

1	Scope	<u> </u>					
2	_						
		rences					
3		itions					
	3.1	Terms defined elsewhere					
	3.1	Terms defined in this Technical Report					
4	Abbre	eviations and acronyms					
5	Conv	entions					
6	AI Te	chnology					
	6.1	AI concepts					
	6.2	Machine learning					
	6.3	Deep learning					
	6.4	Digital twin technology					
	6.5	Digital twin and AI					
7	AI En	AI Energy-saving application scenarios					
	7.1	AI Energy saving implement process					
	7.2	Process of energy-saving tuning with AI					
	7.3	Air conditioning terminal system energy saving control by AI					
	7.4	Chilled water system energy saving by AI					
8	Futur	e Prospects					
	8.1	AI for the future problems					
	8.2	AI application prospects					

Technical Report FG-AI4EE D.WG3-03

Data centre energy-saving: Application of Al technology in improving energy efficiency of telecom equipment rooms and Internet data centre infrastructure

1 Scope

This Technical Report identifies the new application of AI technology in improving energy efficiency of telecom equipment rooms and Internet data centre infrastructure. It will address how AI-based power management can achieve the following capabilities:

- Data collections in telecom equipment rooms and IDC infrastructure.
- Real-time analysis of the historical power consumption and parameters of the target equipment room.
- The ability of training an intelligent model and
- Making reasonable timely adjustments to the air-conditioning and temperature, so as to achieve energy-saving in the telecom equipment rooms and IDC infrastructure.

2 References

None.

3 Definitions

3.1 Terms defined elsewhere

This Technical Report uses the following terms defined elsewhere:

- **3.1.1 artificial intelligence** (**AI**) [b-ITU-T F.749.4]: An interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning.
- **3.1.2 digital twin** [b-ISO/TR 24464]: Compound model composed of a physical asset, an avatar and an interface.

3.1 Terms defined in this Technical Report

None.

4 Abbreviations and acronyms

This Technical Report uses the following abbreviations and acronyms:

ACU Air Condition Unit

AI Artificial Intelligence

BA Building Automation

BMS Building Management System

CFD Computational Fluid Dynamics

DC Data Centre

DCIM Data Centre Infrastructure Management

DNN Depth Neural Network

DT Digital Twin

GSMA Global System for Mobile Communications Association

IDC Internet Data Centre

IT Information Technology

KPI Key Performance Indicator

ML Machine Learning

MNO Mobile Network Operator

PUE Power Usage Effectiveness

SDG Sustainable Development Goal

1

SLA Service Level Agreement

5 Conventions

None.

6 AI technology

6.1 AI concepts

Artificial intelligence (AI) is a new technical science that studies and develops theories, methods, technologies and application systems for simulating, extending and expanding human intelligence. At the Dartmouth conference in 1956, John McCarthy suggested that "artificial intelligence is to make machines behave like people behave intelligently", marking the birth of AI. Different from the natural intelligence embodied by human beings and other animals, AI is the general name of the intelligence embodied by machines. The AI goal is to achieve a system of understanding, thinking, learning and behaviour like humans.

6.2 Machine learning

In the 1980s, a way to achieve AI machine learning was developed. Types of machine learning include supervised learning (training data marks correct answers, divided into regression and classification) and unsupervised learning (training data is not marked, specific structures are extracted from data by clustering) and so on. Machine learning is when you give an algorithm as much sample data as possible, that is, training data, which it can use to make predictions or decisions about events in the real world without the need to perform tasks through explicit programming. An often-cited example is that of identifying cats: you enter thousands of pictures of cats for the machine, and it can then recognize the appearance of cats without having to define cats in the system.

The training process of machine learning is as follows: first, the hypothetical model is defined, such as linear classification, linear regression, logical regression, deep neural network and so on; second, the loss function is defined to measure the quality of the trained model; finally, an algorithm, such as the least square method and the gradient descent method, is selected to optimize the hypothesis model and finally obtain the optimal solution. Different hypothetical models use different algorithms, so that for example linear regression is usually solved by least square method, logical regression is solved by gradient descent method, and neural network is solved by reverse derivation. In today's big data era, machine learning is essential. The purpose of collecting, transmitting and storing big data is to "utilize" big data (data pre-processing, data mining, etc.), and without machine learning technology to analyse big data, "utilizing" it will not be possible.

6.3 Deep learning

In the 2000s, a method to realize machine learning deep learning (learning based on deep neural network) developed, which is inspired by the physiological structure of the human brain, using a multi-layer neural network to establish an algorithm. The cognitive behaviour of simulating the interaction between neurons is also an example like that of identifying cats. If a shallow machine learning system is used, it takes time to define the edges of a cat, while deep learning uses a multi-layer learning algorithm to understand the data. The first few layers can calculate the edges of a cat image, and the latter layers can focus on the cat's whiskers and eyes. Compared with shallow machine learning, deep learning usually requires more data, higher computational power, but less manual intervention.

6.4 Digital twin technology

6.4.1 Digital twin concept

Digital twin refers to the creation of a completely corresponding and consistent virtual model for physical entities in the real world in a digital way, which can simulate their behaviour and performance in the real environment in real time.

6.4.2 Data centre digital twin

Because of characteristics including high security requirement, high trial and error cost and high energy consumption, it is very important to predict data centre security and energy consumption management. The whole life cycle of the data centre is divided into layout planning, process planning, rapid delivery, energy consumption optimization, intelligent operation and maintenance. Numerical simulation technology can be used in the planning, design and construction stages, and digital twin technology can be used in the operation and maintenance stage of the data centre.

In the planning and design stage, the performance of the data centre equipment and system can be verified in advance by numerical simulation technology, and the possible quality and function defects of the equipment and system can be captured as soon as possible to ensure that the design requirements are met.

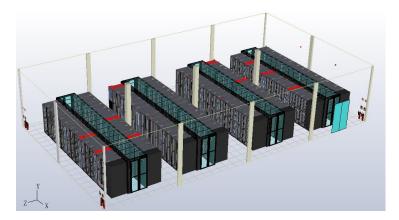


Figure 1 – DC digital model

Figure 1 illustrates an example of the data centre (DC) digital model. The digital model formed in the planning, design and construction phases can be reused in the operation and maintenance phases, and through a digital twin model can obtain the real-time information of the system, air conditioning equipment, power equipment and so on in the actual operation process. It can alsp predict and troubleshoot other scenarios such as equipment changes and scheme management in the future, and perform the management of the real data centre in the digital model.

Digital twin in data centres is often inseparable from computational fluid dynamics (CFD) simulation technology. The 3D data centre model built by CFD software can not only reproduce the present situation of the data centre, but also connect with the physical data centre dynamically and in real time through the interface, so it is called a twin model. Corrected digital twin models have been highly unified with physics (usually more than 90% CFD for the accuracy requirements of model simulation and measured comparison), and their digital attributes are of great significance in predicting the future operation state, energy utilization rate, space utilization rate and operational risk of the data centre.

Therefore, both static numerical simulation technology and dynamic digital twin technology provide an important basis for the safe and energy-saving operation of the data centre and the decision-making of the personnel, which play an important role in the operation of the data centre.

6.5 Digital twin and AI

Digital twin and AI are a pair of companion technologies. In the actual running data centre, tens of thousands of data will be generated in each system and equipment, and each data centre cannot transmit each data in real time because of the limitation of sensor positions, quantity and other factors. The monitoring system's monitoring of the data centre is local and one-sided. Digital twin technology can not only monitor the parameters of each system and equipment, but also predict the running state of the data centre in the future. After "filtering and cleaning", data mining and data analysis of the artificial intelligence algorithm, the data generated by the digital twin can be trained to save energy.

Digital twin has been widely used in energy-saving control of air conditioning terminals and chilled water systems. It is believed that with the continuous evolution of digital twin technology, digital twin technology will be closer to the real world.

7 AI energy-saving application scenarios

7.1 AI energy-saving implement process

The refrigeration system is the provider of cooling capacity, and DCs are a demander of cooling capacity. Take the water-cooling system as an example, the cooling demand model of a computer room can be built by using water temperature difference and water flow rate. As the provider of cooling capacity, the capacity of a refrigeration system to provide cooling capacity is related to environmental parameters (external temperature and humidity), and is also related to adjustable parameters such as freezing pump frequency. When the adjustable parameters are fixed, then the higher the temperature and humidity of the external environment, the weaker the capacity to provide cooling capacity. When the external environment is fixed, improving the operating power of the refrigeration system will also increase cooling capacity, but it will increase power consumption. AI energy-saving needs an algorithm model to learn the relationship between the refrigeration capacity and environmental parameters and adjustable parameters.

At the same time, the energy utilization ratio of the refrigeration system has a relationship with environmental parameters and adjustable parameters, which can be fitted by an algorithm model. In addition, different refrigeration systems have different refrigeration capacities and different energy utilization curves.

Figure 2 shows an implement process of data processing and parameters controlling.

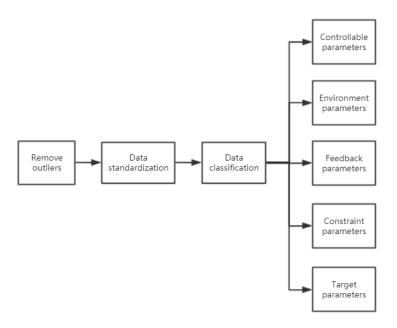


Figure 2 – Implement process of data processing and parameters controlling

Controllable parameters: frequency of cooling pumps, freezing pumps and the pumps used in cooling water tower.

Environmental parameters: air temperature, humidity.

Feedback parameters: water temperature, temperature difference, etc. in each equipment.

Constraint parameters: threshold value or parameter range of equipment operation.

Target parameters: power usage effectiveness (PUE) value, defined as a visual expression of energy efficiency.

Figure 3 shows the optimization method of the PUE parameter. The PUE curve fitting parameter optimization method is illustrated through four parameters, and uses a machine learning and depth learning algorithm and consideration of the influence of time series and related parameters, accurately fitting the optimization target PUE value. Adjustable parameters are optimized to make the PUE value lowest.

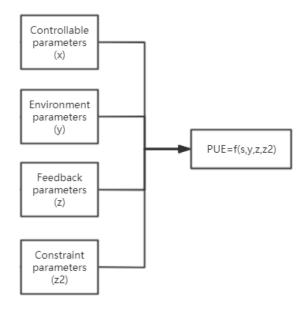


Figure 3 – Optimization method of PUE parameter

A data centre AI energy-saving process includes data acquisition access and storage, an artificial intelligence algorithm platform, a visual interactive interface, a device command value control output, a feedback and monitoring system and other functions, as well as an auxiliary data centre to achieve efficient operation and maintenance.

Key index prediction: using a machine learning and depth learning algorithm to accurately predict PUE and other parameters to provide an accurate PUE perception for operators.

Advice of operation strategy: through a machine learning and decision optimization algorithm, the best equipment operation combination and parameters that can reduce the PUE can be directly linked to the equipment and automatically controlled.

Intelligent operation and maintenance monitoring: 24/7 for daily monitoring and system maintenance services, real-time monitoring of equipment operation status, key performance indicator (KPI) changes and failures and other information.

Closed loop tracking control: the algorithm outputs the operation parameter strategy of the equipment, the system can track the parameter changes and the energy-saving output effect in real time, promote the feedback of the model learning effect, carry on the self-learning upgrade and realize closed-loop control.

Promotion of building automation (BA) upgrade: update and repair existing BA systems through problem and fault analysis, add or upgrade the existing sensors and communication equipment to provide sufficient data support for the algorithm.

7.2 Process of energy-saving tuning with AI

The application of energy-saving tuning by AI is shown in Figure 4. The necessary sensor data can be accessed via a data access layer offering overall data mining ability and an intelligent analysis algorithm for data which contains a common foundation framework, a machine learning algorithm, a deep learning algorithm and a neural network algorithm. The internal relationships and values of sensor data can be identified by automatic selection and processing of features, automatic adjustment and training. The complete solution of data mining applications for the data centre can be obtained by data analysis, integration, testing and evaluation. The systematic ability of operation and maintenance can be assessed and data collected through the integration of the system and a BA system. Data interaction can be accomplished through protocol while the data conversion module can also perform data cleaning and data formatting. The whole process of data mining such as data storage, analysis and processing inside the system can be realised. The forecasting of data, machine

operation, checking, monitoring and the exporting of authority management can be implemented by a data visualization platform.

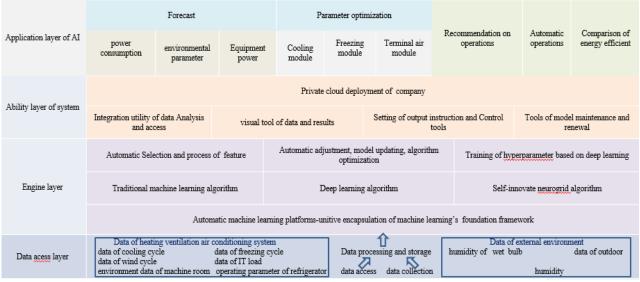


Figure 4 – Overall architecture of AI energy-saving tuning

The modules used in this system are described as follows:

Data acquisition module: This module can get data from a BMS/BA/DCIM (building management system/ building automation/data centre infrastructure management) system, parse the data protocol, transfer it to the data format that the system needs and send it the data-to-data processing module. The main functions of this module include importing the BMS/BA/DCIM historical base and basic information of equipment, receiving the real time performance data of the machine sent from BMS/BA/DCIM and sending of the data-to-data process module via a service interface.

Data process module: The function of this module is receiving data from the data acquisition module, storing the data locally and satisfying the forecasting movement of the algorithm module.

Data storage module: This module is responsible for the forever storing of data, dividing the data into hot and cold data for storing. The cold data can be used for shown of history analysis, while the hot data can be used for real time inquiry and analysis.

Intelligent analysis of arithmetic module: The function of this module is mainly comprised of three parts: model training, model self-learning and closed-loop tracking. The optimized operation composition and parameters for reducing PUE can be provided by machine learning, deep learning and forecasting of PUE, environmental parameters and equipment power.

Device module: The function of this module is to control the equipment based on the controlling instructions from arithmetic training. This module can receive the output results and control instrument then adjust and control the parameters of equipment.

Self-monitoring module: This module can implement the inside monitoring, recognise emergency warnings and common warnings such as failures of data acquisition, equipment control, arithmetic training, and so on. The functions of this module also includes warning alert cooperation with the visualization platform, and giving pre-warning to the head monitoring station such as the controlling system of machine rooms. If there is no problem, the system can self-recover by running a self-triggered procedure.

Data visualization module: This module can provide the data analysis and visualization per visual component. The main functions include the management of pages, views and elements, management of user rights, operation log management and alert and curvilinear figure display.

The different modules and their functions are illustrated in Figure 5.

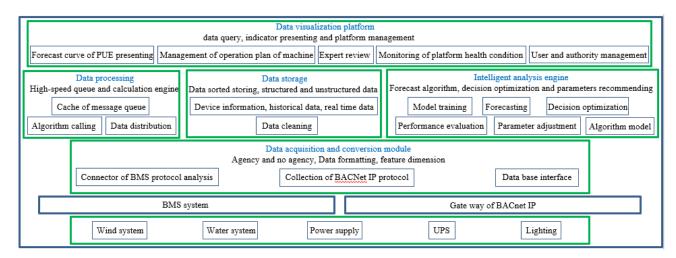


Figure 5 – AI energy-saving tuning modules and their functions

7.3 Air conditioning terminal system energy-saving control by AI

For chilled water or direct expansion air conditioning systems, the energy consuming component at the terminal of the air conditioner in DC rooms is the air condition unit (ACU) fans, so the energy-saving is mainly reflected in the energy consumption of the fan.

The reduction of the fan speed can effectively reduce the energy consumption of the fan. Fan laws has proved this statement in theory.

According to the fan laws, when the rotating speed of the same fan (with the same fan diameter and the same air density) changes, the ratio of the fan shaft power before and after the speed change is directly proportional to the third power of the fan speed. For example, if the fan speed is reduced to 90%, the input power will be reduced to 72.9%.

The minimum air conditioning fans speed can be realized by AI algorithm, that is, using on-site and real-time monitoring data, with AI technology such as supervised learning or reinforcement learning, etc. In this way the accurate evaluation of the minimum terminal fan speed that meets the inlet temperature requirements of IT equipment will be achieved.

To reduce the fan speed, there are two scenarios:

- (1) Reduce the fan speed without adjusting the control temperature.

 The output air flowrate of ACUs greatly exceeds the demand air flowrate of IT equipment, then reduces the speed and reduces the invalid air flowrate output and gets the minimum fan output when the air supply volume is close to the demand air flowrate of IT equipment.
- Adjust the control temperature to reduce the fan speed.
 To improve the air conditioning control temperature, the DC room needs to meet some conditions.

First, determine whether the air distribution and cooling configuration in the machine room are reasonable, and whether the maximum air inlet temperature of the current IT equipment is within the allowable range required by ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers), so as to provide a safety guarantee for improving the control temperature.

Second, using numerical simulation technology to calculate the digital twin model, predict and improve the reasonable value of the control temperature in advance and reduce the potential safety hazards of IT equipment.

The first method realizes the balance between supply airflow and demand airflow. The control temperature can be increased under this condition.

There are two benefits of increasing the control temperature:

- 1 After increasing the control temperature of the ACU, the energy efficiency of the unit can be effectively improved and energy-saving can be realized.
- Improving the air distribution system in the room will improve the ACU supply efficiency, this can further reduce the fan speed and achieve the purpose of energy-saving.

Finally, the fan speed should be sent to on-site. The trained model will automatically adjust the fan control value according to the load change of IT equipment and keep the fan running with minimum energy consumption.

7.4 Chilled water system energy-saving by AI

AI energy efficiency adjustment, that is, using AI technology to connect the data between the IT load to the refrigeration system and the whole external environment to collect, manage, train and reason a large number of historical operations and maintenance data, and to obtain the PUE prediction model.

Realizing the energy-saving of a chilled water system with AI requires the cooperation of software and hardware. For example, the chiller, water pump, valve, terminal and other components have AI accessible interfaces, and AI and equipment data are exchanged and shared. The specific implementation can be divided into four steps described below and shown in Figure 6.

- Data acquisition: through the moving ring management system, collects the freezing station (including outdoor working conditions, freezing station operating state, etc.), terminal air conditioning (including cold and hot aisle temperature, terminal operating state, etc.), IT load (including IT load change, historical PUE, etc.) and other system operating parameters.
- 2) **Data governance**: the original data collected into the data lake on the cloud, desensitization, merging, noise reduction, cleaning, correlation analysis feature engineering and other governance, to provide high quality data for subsequent model training.
- Model training: using managed high-quality data, training depth neural network (DNN), output high-precision PUE prediction model, business prediction model (for SLA business operation guarantees). The PUE model will be updated in real time according to the optimization situation in the early stages of operation, and continuously optimized during stable operation.
- 4) **Reasoning execution**: the trained PUE prediction model is distributed to the reasoning platform. The online reasoning is completed by using relevant algorithms with the current outdoor environment and business forecast model under its load as constraints. The lowest adjustable parameter combination of PUE is obtained. Finally, the group of parameters is issued to a group control system for execution. When the outdoor wet bulb temperature and its load change, the next round of reasoning and optimization will be triggered automatically, and the adjustment cycle can also be set manually according to the business requirements.

In the initial operation of AI tuning, the inference parameters are confirmed by experts and then issued to ensure business security. After a month of manual judgment, the instruction parameters can be automatically issued. Finally, through standardized practice guidance, continuous optimization and regular measurement, the best PUE can be obtained.

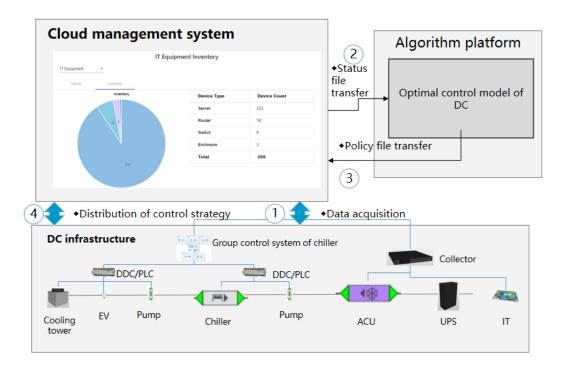


Figure 6 – Implementation steps of AI energy efficiency tuning

AI energy efficiency tuning has the following value characteristics:

- (1) Energy-saving and high efficiency, overall regulation of refrigeration system, and system level energy efficiency optimization.
- The operation is safe. The data reading and command parameter distribution are executed through the group control system without directly controlling the refrigeration equipment. When the AI optimization function is disabled, the building management system (BMS) system will automatically return to the state before optimization; The AI reasoning process follows the operation and maintenance specifications and meets the parameter range constraints of the refrigeration system, and the reasoning results meet the SLA requirements. It can be confirmed by experts before issuance and implementation, so as to enhance the system security.

8 Future prospects

8.1 AI for the future problems

Data security risks: The general AI model and data which are stored in a public cloud may be at risk of being stolen and the operation of the data centre may face unknown security risks. Moreover, the ability of technical personnel will be not enough to handle increasing data in the future. However the AI model will have the ability to control and optimize the operation parameters of the cooling system automatically with the continuous improvement of AI at decision-making level. As a result, engineers will have less opportunity to participate in the operation and this also means that less human errors will exist. Not only that, intelligent firewalls controlled by AI will increase the safety level of public cloud and control systems significantly.

Coupling conflict: The operation system of the data centre contains multiple control subsystems, and every subsystem has an independent control model. This operation system will become more and more complex in the future. If these subsystems work respectively, there may be conflicts between them. AI systems will give an excellent coordinate controlling for this huge and integrated system.

8.2 AI application prospects

At present, although the energy consumption of the optimized data centre has been reduced by 10% via the AI cooling systems made by some vendors, obtaining an optimized energy consumption model requires a lot of data and training time and until now the feedback speed has not been fast enough. The improvement of AI model algorithms which are not limited to data-driven machine learning, but are also integrated with physical model algorithms, optimization calculation and experience accumulation are expected in the future.

AI energy-saving can greatly reduce the energy consumption of data centres that contain IT equipment and a cooling system. The energy consumption models of IT equipment and data centre infrastructures will be used to form a more complex AI model. The AI model can make multiple optimization and control objectives for data centres, such as data capacity, energy efficiency, safety, and other indicators. The AI model will balance these objectives and provide a best solution for control plans.

Bibliography

[b-ITU-T F.749.4] Recommendation ITU-T F.749.4 (2021), *Use cases and requirements for multimedia communication enabled vehicle systems using artificial intelligence.*

[b-ISO/TR 24464] ISO/TR 24464:2020, Automation systems and integration — Industrial data — Visualization elements of digital twins.

[b-Deloitte 2020] Deloitte. 2020. Tech trends 2020. Tech Trends 2020 (deloitte.com)