2020 IEEE 4th Information Technology,Networking,Electronic and Automation Control Conference (ITNEC 2020)

Intelligent Cloud Training System based on Edge Computing and Cloud Computing

Zhijia Chen, Yanqiang Di, Hongli Yuan, Shaochong Feng*
Equipment simulation Training Center, Shijiazhuang Campus of Army Engineering University
Shijiazhuang, China
youshenshui@163.com

***Abstract*—Equipment simulation training based on cloud computing is emerging. However, the latency between cloud center and client is long, and the energy consumption management is difficult, which are influencing the development of cloud training. Intelligent cloud training system based on edge computing and cloud computing is introduced in this paper. Intelligent gateway is introduced, through which the task and resources are scheduled and managed together. The popularity of training resources is analyzed. The management of servers in cloud center and edge is intelligently switched between timing sleep and task-activation. Intelligent training service provisioning is achieved through above measures. The simulation results show that the system and management methods are effective on improving training service quality and lower the energy consumption.**

Keywords—edge computing, cloud training, resource scheduling, energy consumption management.

I. INTRODUCTION

The emerging of cloud computing [1] and cloud simulation [2] proposed new road for large scale associated training and network simulation training. In the cloud training mode, each user is provided with independent virtual hardware via network, to realize allocat[1]ion on demand and full share of computing resources and simulation resources. Multi users visit the datacenter via network to deal with the tasks including data calculation and graphic processing, etc. By cloud training, users can go into equipment operation training, equipment maintenance training, and equipment command training without restriction of geographic location.

During the training procedure, as the long distance connection is restricted by network bandwidth, the transaction and processing of large amount of 3D models, graphics and videos faces huge pressure, which results long time delay when transmitting data between client and cloud center. Besides, whether troops have training tasks, cloud training center is running continuously. The energy consumption can be lowered in despite of resource scheduling technology, however, the problem of high energy consumption cannot be solved completely. Use a new resource processing mode – edge computing as a reference, putting the resources used to be deployed in the cloud center to the edge node and gives the edge node resource processing ability. Cloud center and edge node cooperation method is proposed. Part of the training resources, models, tasks used to be deployed in cloud center are decomposed and migrated to high performance edge nodes. Use this method to reduce the load of cloud center and then achieve the goal of lowering down energy consumption, reducing time delay, improving training efficiency.

II. RELATED WORKS

Lots of researchers have been conducting researches in edge computing, cloud computing and the application to training in depth. Meng[3] et al. proposed a multi-users multi-tasks simulation training conception to solve the equipment simulation training problem in the condition of informationization. The training system could satisfy the requirement of large scale, multi person-time equipment teaching and training. Grid technology, web service technology, HLA/RTI technology based multi users, multi tasks mode is realized and some application problems are solved. Due to the particularity of air arm training, Huang[4] et al. took full advantage of the powerful computing and storage space capabilities of cloud computing in processing complex system, and proposed a new architecture of air arm training simulation system based on cloud computing platform. The main contents which should be processed by cloud computing platform are figured out. Song[5] et al. researched the method of saving computing resources and improving the utilization in the environment of Internet. Virtual container technology is used to assemble series cloud services into PC platform, which forms local assistant computing unit. Through the cooperation work of service management program and Intelligent gateway, the energy save algorithm is designed, the cloud center computing load is lowered, and system energy conservation management is realized. Yang[6] et al. allocated the tasks of cloud center to edge nodes, and analyzed the resources requirements of users, the resource utilization frequency, the restriction of node scales. The problem of resource transmitting cost and time delay between cloud center and clients are abstracted as multi dimension bag optimization problem. Finally, the problem is solved by optimization algorithm and the time delay is reduced. Zhang[7] et al. provided IT and cloud computing capabilities for wireless network and give the wireless network the transmission ability of low time delay and high bandwidth. The cost of operation is reduced in a certain extent.

The researches on edge computing, cloud computing and the application in training produced certain effect. This paper uses the experience of the above researches achievements to

[1] Shaochong Feng*, the corresponding author.

978-1-7281-4390-3/20/$31.00 ©2020 IEEE

improve the service ability and reduce the energy consumption of cloud training.

III. SYSTEM ARCHITECTURE

The edge-cloud computing based intelligent training system is constructed as shown in figure 1. The system is composed of cloud computing servers, edge devices/servers, and Intelligent gateway. In this system, intelligent gateway is the key of system task and resource scheduling. The form could be a training computer or training client device in which the intelligent gateway control program is installed. The intelligent gateway controls the training tasks and is mainly used in training resources analysis, training tasks analysis, training resources and tasks matching, etc.

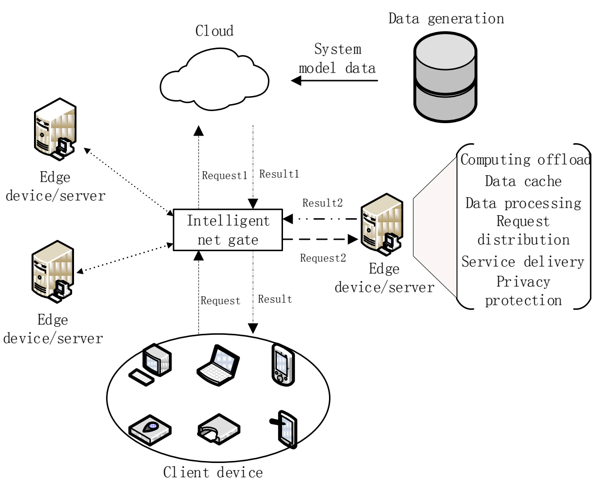

Figure 1 the architecture and the application of edge-cloud intelligent training system

Edge cloud intelligent training resources scheduling architecture is shown in Fig.1. The training system is developed and deployed in cloud center. Users get training services by using client devices. In the cloud-edge intelligent architecture, intelligent gateway is added between cloud center and client to realize data filter and allocation. Intelligent gateway records the services that edge nodes could provide. When client requests arrive, gateway indexes the record database quickly. If edge node could provide the service, then the request (request 2) is forwarded to the edge node. If not, the request is submitted to cloud center (request 1). If more than one edge node could provide the service, gateway provides intelligent decision. If edge node receives the requests forwarded from intelligent gateway, it deals with the tasks and return the results (result 2) to intelligent gateway and the gateway returns the results to client. In the edge side, service node could provide computing service, data cache and storage service, data processing service and service delivery, etc.

IV. SYSTEM MODEL

A. Basic hypothesis

In cloud center, there are m types of training system. The requirement of resources of each type of training system is $R = \{R_1, R_2, \ldots\ldots, R_m\}$. In this array, each element is an array and includes resource requirements sub array of CPU, GPU, memory and storage, which could be represented as $R_i=\{r_{i1},r_{i2},\ldots\ldots,r_{in}\}$. n=5 represents 5 types of resources. The utilization heat $h=\{h_1, h_2, \ldots\ldots, h_n\}$ represents the using frequency of the training systems. It is interrelated with user request frequency and history heat. Edge node is represented as $e=\{E_1, E_2,\ldots\ldots, E_t\}$, in which t represents the total number of edge nodes. The resources in each node could be used is $RE=\{RE_1, RE_2, \ldots\ldots, RE_m\}$. Each element is an array including available resources such as CPU, GPU, memory and storage and could be represented as $RE_i=\{re_{i1}, re_{i2}, \ldots\ldots, re_{in}\}$. n=5 represents 5types of resources.

Restriction condition: a) total resources that each edge node caches could not exceed the whole storage number of all edge nodes. b) Each user could only get resource one time, either from edge node or from cloud center, never repeatedly. c) Each type of resources could exist in multi edge nodes.

B. Training system heat evaluation model

When the center and edge node provide training resources in cooperation, and the resources migrate from center to edge node or from edge node to center, we should first calculate the heat of training system, then calculate the amount of resources to make the training system more close to client and allocating enough resources for training system.

Training system heat is determined by system visit frequency and history heat. Heat attenuation should be considered as well. Training system heat is defined as:

$$p^i = \alpha\cdot\left(C\cdot t/T\right)+\left(1-\alpha\right)\cdot p^{i-1}$$

p^i is training heat to be calculated. p^{i-1} is history training heat. α is attenuation coefficient ranging [0,1] representing the weight relationship between current heat and history heat. C is training system using number of times in this statistic period. t is the average time. T is the period time. We can see that the influences of history heat to data heat is decreasing after the summation heat undergoing multi-time attenuation by (1-α) coefficient. If α is bigger, the decreasing ratio is quicker and the influences of history factor is smaller or vice versa.

C. System energy management model

When cloud training center and edge device provide training service in cooperation, single limitation energy management approach is difficult to satisfy this multi-region and hyper-device energy saving requirement. In traditional cloud center, servers will shutdown or sleep when there is no or few visiting request. In this system, intelligent gateway is used to make unified management of energy. When the coefficient of utilization of servers in cloud training center, some servers could shutdown or sleep, which may cause unavailability of services provided by cloud center. Then the service provision should be switched and provisioned by edge device. Otherwise, the cloud center server will be awakened to provide service continually. If edge device sleeps and the edge service is unavailable, cloud service should be switched and data should be transmitted to cloud center. The key problem is when to sleep or awaken. If the sleep time is too long or is not awakened in time, it may cause service unavailability. However, if sleep time is too short or awakened up too frequently, energy consumption will rise up and device life will decrease. Multi-conditional intelligent sleep-awaken algorithm is proposed based on user task monitor information, edge node processing task scale monitor information, processing period monitor information, center processing task scale monitor information and processing period monitor information. Using this algorithm, servers in edge and center could sleep logically. The goal of providing reliable services and saving energy could be realized.

Step1. Calculating node active degree ξ, which is exponential relationship with time interval from initial time to current time and send-receive frequency.

$$\xi = 1 - e^{-\lambda \cdot \frac{k}{\Delta t}}$$

Δt is statistic time interval. k is the send-receive data number of times in Δt, which determines the attenuation ratio of active degree curve in Δt. If $\frac{k}{\Delta t} \to 0$, $\xi \to 0$. If $\frac{k}{\Delta t} \to \infty$, $\xi \to 1$.

Step2. According to active degree, pre-define sleep period of edge node servers and center servers. Sleep period is inversely proportional to active degree.

Step3. According to the sleep period, the servers awaken and process tasks every other time. When the task is done, servers sleep and next period time count begins.

Step4. This method may cause the edge node and center server both sleep at the same time, and some task could not be processed in time. Therefore, edge node and center server that process the same task need to be separated. The sleep period could be alternative in the timeline, but not overlapped.

Step5. If some edge nodes and center servers could not be totally alternative in the timeline, intelligent gateway is introduced to control this. The intelligent gateway will not sleep and will collect task information continuously. If the task capacity increases and statute for limitation is strict, then first awaken the edge node server and guarantee the task real time processing. If edge node could not handle this totally, then awaken the center server to process the task.

Through these measures, center-edge energy intelligent optimization management with service high availability guaranteed is realized.

D. training service get model

After intelligent gateway receives user requirements, intelligent gateway will calculate the forwarding destination according to the distance (d) between edge node and users, training system heat (p), current task's resources requirement (R) of edge node, net throughput (n) of edge node, processing power (RE) of edge node. Multi-attribute weight algorithm is introduced to calculate the destination.

Step 1. Assign different parameter with different weight, denoted as w_1, w_2, w_3, w_4, w_5.

Step 2. Calculate destination overall target index, $f = w_1 \cdot \frac{1}{d} + w_2 \cdot p + w_3 \cdot \frac{1}{R} + w_4 \cdot n + w_5 \cdot RE$.

Step 3. According to intelligent recording edge servers number m, get the collection of all overall target indexes, denoted as F.

Step 4. Arrange collection F with descending order, and choose edge node to provide the service.

Through above method, intelligent task scheduling and service scheduling is realized and center-edge service hybrid intelligent providing is realized.

V. EXPERIMENT AND ANALYSIS

Compared with cloud center, time delay of edge node is shorter. Local servers and application lies in the same LAN, which guarantee high bandwidth and low time delay. While, cloud center is far from clients, the bandwidth is narrow and time delay is high. Table 1 shows the custom bandwidth and time-delay of edge node and cloud center.

TABLE I. CUSTOM BANDWIDTH AND TIME-DELAY OF EDGE NODE AND CLOUD CENTER

Type	Time delay (ms)	Bandwidth (Mb/s)
Edge node	＜1	＞100
Cloud center	＞10	＜10

Figure 2 shows comparison time of cloud center and edge node processing the same tasks. We can see that when only adopting edge node to process tasks, as the processing capability is comparatively weak, the processing time is long. However, the time delay is short, with the increasing of tasks capacity, the disadvantage of weak processing capability is distinct. When only adopting cloud center to process tasks, benefit from the powerful computing capability, the consuming time is shorter than only by edge node. When adopting intelligent optimizing method, combined with the low time delay of edge node and high performance of center, the choosing of processing node is optimized. The total consuming time is shorter than only adopting edge node or center server.

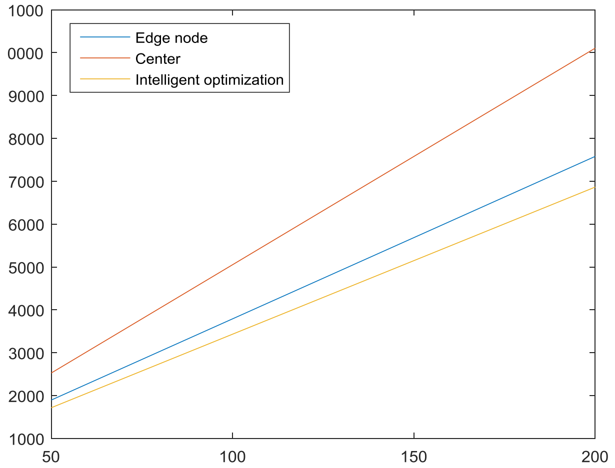

Fig.2 task processing time comparison

Figure 3 shows the energy consumption when processing the same tasks. We can see that when only adopting center servers to process tasks, though some energy saving algorithm is introduced, high performance computing needs lots computing devices and supporting devices, the energy consumption is high in whole. When only adopting edge devices to process tasks, the energy consumption is lower than center servers, but the long processing time and weak processing performance cause energy consumption high. When adopting intelligent optimizing method, combined with the low energy consumption of edge node and high performance of center, the energy consumption is optimized and lower than the other situations.

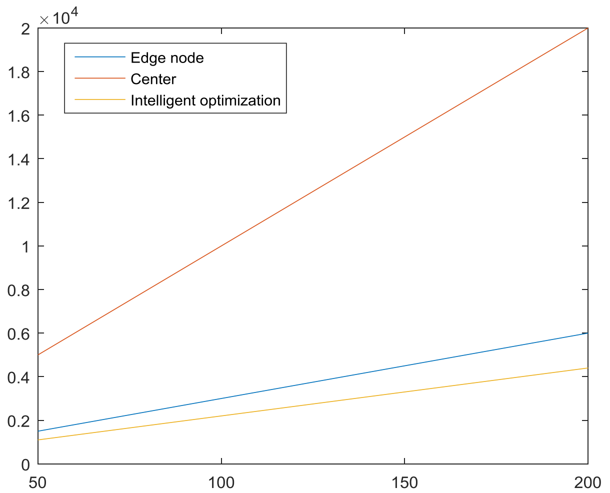

Fig. 3 energy consumption comparison

Conclusions

This paper analyzes cloud training problems and further researches the mode of edge-cloud intelligent providing services. Resources providing and energy management is researched in depth. The experiment results show that the algorithm could improve service performance and lower down the energy consumption. Next step is to apply the system to troop training and improve this training system in application.

References

[1] Ling Qian, Zhiguo Luo, Yujian Du, et al. Cloud Computing: An Overview[J]. Lecture Notes in Computer Science, 2009, 5931(1):626-631.

[2] Chai Xudong, Li Bohu, Hou Baocun, et al. Networked Modeling & Simulation Platform Based onConcept of Cloud Computing – Cloud Simulation Platform [J]. Journal of System Simulation, 2009, 21(17):5292-5299.

[3] Meng Xian-Guo, Di Yan-Qiang, Zhu Yuan-Chang, Et Al. System Design and Implementation:a Web-based Simulation Training System in Grid Environment[J]. Journal of System Simulation, 2009, 21(4):1202-1205.

[4] HUANG An-xiang, FENG Xiao-wen, LI Jin-song, et al. Aviation Simulation Architecture Based on Cloud Computing Platform[J]. Journal of System Simulation, 2011, 23(S1):106-109.

[5] SONG Pengtao, LI Chao, XU Liting, Et Al. Edge Computing System for Smart Home Based on Personal Computer [J]. Computer Engineering, 2017, 43(11):1-7.

[6] Yang Shihao. Resources deployment and allocation problem in mobile edge computing environment [J]. Chinese New , 2017, 21:122-124.

[7] ZHANG Jianmin, XIE Weiliang, YANG Fengyi, et al. Mobile Edge Computing and application intraffic offloading [J]. telecommunication Science, 2016, 7:132-139.