HK.NCCP
International Journal of Applied Mathematics and Soft Computing
Volume 3, Issue 1, June, 2017

Cloud Scheduling Algorithm based on Phase Space Analysis Model

Haogui CHEN
Hunan City University, Yiyang, Hunan, China

Abstract: To solve the problem like calculation of migration frequently existing in the cloud computing system, storage migration, backup computing, storage, backup and node failure handling high coupling between nodes operating characteristics, this paper presents analysis model based on phase space cloud scheduling algorithms. This article makes cloud computing system project to phase-space, defining and analyzing the generalized phase space temperature, generalized normalized entropy, generalized gravity, generalized normalized momentum, generalized thermodynamic parameters etc. Taking advantage of thermodynamic principles to establish a cloud computing system phase space analysis model of the basic theoretical framework, taking advantage of phase space analysis model scheduling algorithm for cloud computing cloud computing can be projected to the parameter phase space and momentum phase space. Experimental results show that: the proposed algorithm can effectively make cloud computing system achieve load balancing in phase space remain thermodynamically non-equilibrium systems.

Keywords: Projection Analysis Model; Phase Space; Generalized Temperature; Normalized Momentum

1. Introduction

Cloud computing is a new computing paradigm. It is a distributed computing, parallel computing and grid computing. For those data who have requirements of a higher consistency and concurrency, Cloud storage layer will have to pay even more dearly, that is read in the cloud to be much less than the cost of the cost of writing, especially large files. The most important feature of cloud computing is to provide the most reliable and secure data storage center; users do not have to worry about data loss, virus attacks and other problems. Because in the "cloud" the other side, there are the world's most professional team to help users manage information, has the world's most advanced data centers to help the user to save data. Meanwhile, the strict rights management strategies can help users safely be shared with the user-specified data. Secondly, cloud computing client devices require a minimum, it is also the most convenient to use. Users only have a computer with Internet access, have their own favorite browser, you can type in the browser URL, and then enjoy the cloud and have fun.

As cloud computing system containing frequent calculation migration, storage migration, backup computing, storage, backup, node failure handling high coupling between nodes makes cloud computing system into a high coupling cluster system, which is cloud computing data Center and traditional data center are very different, there is a low coupling between traditional data center server ,the logical server is put together physically, it is not much meaningful to obtained the overall trend of the data center in this case , while, because of the strong coupling of cloud computing data center and the server is a single pool of resources, which are more concerned with the overall work situation [1]. Massive cloud computing system nodes and high coupling between nodes have brought great difficulties to analysis and modeling of cloud computing system architecture, to find a theoretical model, which can effectively describe the overall situation of the cloud computing system, has an important significance on cloud computing technology development. In recent years, cloud computing system architecture and modeling research has become a hot area of research, Google company put forward a representative Map-Reduce distributed processing model of cloud computing, several other companies and researchers have also put forward their own cloud computing model, but there is still lacking of a system can effectively reflect the massive cloud computing system nodes and high coupling characteristics of cloud computing model to support cloud computing architecture and model [2-5]. In mathematics and physics, phase space is a system which is used to represent all possible states of the space; each possible state of the system has a corresponding phase space points [6]. By using the idea of phase space to express cloud computing system working state, each server node in cloud computing system corresponds to a point on the phase space, the location of point describe the server's current working state, points' movement has described the server parameter changes, phase space projection method provides a new view to cloud computing system simulation, scheduling algorithm design, analysis, so that we can learn from the whole profound observation system working state.

For different applications provide differentiated service quality has always been the core objectives of the network resource scheduling, With the diversity and complexity of network applications increasing, the cloud network resource scheduling faces new challenges. Based on the traditional bandwidth sharing resource allocation strategy aims to achieve higher average network throughput and better resource utilization, however, the aspect in considering users' satisfaction to the quality of services provided by network is deficiencies.

At the same time resource scheduling and load balancing is the traditional topic of cloud computing and cluster system research, various balancing algorithms are endless, but according to different scheduling policies, the existing equalization algorithm can be broadly classified into two categories: static load balancing algorithm and dynamic load balancing algorithm. Static load balancing algorithm has low overhead, simple features, was once the center of attention, but because of lack of consideration of the actual load conditions of each node server, poor flexibility, etc [7-8]. PHASE SPACE ANALYSIS MODEL

Cloud computing system is composed of a large number of high coupling server nodes, if they are directly observed and analyzed for each server, we will be unable to understand the overall operation of the system state, which can not be on the system resources for effective scheduling and analysis. In this section we have conducted research of cloud computing system's parameters phase space projection and momentum phase space projection.

1.1. Cloud Computing System in The Parameter Phase Space Projection Analysis Model

We found that when the number of nodes in the cloud computing server reaches a certain size, a single server parameters (such as CPU utilization and memory usage, etc.) with a large server clusters to form a clear distinction between micro and macro, this difference in the phase space can be more easily described and analyzed, especially the use of thermodynamic models, in order to found a theory description tool in the phase space for the cloud computing system with massive high coupling nodes.

Definition1. By the server's two or more parameters of a generalized coordinate axis formed by the two-dimensional or multidimensional space called cloud computing system parameter phase space.

The use of a phase-space projection m servers cloud computing system to do two parameters can be used when projecting on a two-dimensional phase space point set A to represent:

$$A=\{(x_i,y_i):i\le m, x_i\le 1, y_i\le 1\} \tag{1}$$

x_i , y_i is the normalized server parameters.

If the parameter is projected onto each server to each parameter is a parameter generalized coordinates phase space, the server parameter change process parameters on the phase space into a point on the course of the campaign, this mapping will transform into a cloud computing system dynamics systems research. By contrast can found that in the cloud computing system parameters on the number of nodes in the phase space movement and a lot of thermodynamic systems corresponding to the movement of gas molecules, while the relationship between nodes coupled with thermodynamics of complex interactions between gas molecules force corresponds Using this correspondence can be used in thermodynamics has a clear physical meaning of the macroscopic and microscopic amount of the amount of cloud computing system working state to describe, analyze and control. As cloud computing system is not an actual thermodynamic system, so we generalized thermodynamic parameters are defined as follows.

Definition2. Cloud computing system parameters defined in the phase space projection with a clear physical meaning physical quantity called the generalized thermodynamic parameters. Generalized thermodynamic parameters is an important tool in the parameter phase space analysis of cloud computing system status and effect of scheduling algorithms, through generalized thermodynamic parameters to establish a single server node status parameter phase space linkages with cloud computing system, and can be used to construct generalized thermodynamic Some criteria scheduling algorithm parameters.

Definition3. The moving average projection point distance on phase space the server every time is called the generalized mean free path. Mean free path on the load request, the external load on the overall load larger request, the larger the mean free path of the server in the phase space, the difficulty of a large mean free path of scheduling system resources will become large.

1.2. Calculation of Cloud Computing System Parameters Normalized Phase Space Generalized Entropy

Phase spatial analysis of cloud computing systems needs to use another generalized thermodynamic parameters: the generalized normalized entropy. Generalized normalized entropy can describe the current extent of deviation from the equilibrium state, which is now very similar to the definition of thermodynamics, a better load-balancing cloud computing system whose generalized normalized entropy value should be relatively small. For ease of comparison, the entropy values were normalized to form generalized normalized entropy.

In thermodynamics entropy R is defined as follows:

$$R = yIn\Omega \tag{2}$$

Here, k is the Boltzmann constant, Ω is the number of microstates.

The generalized normalized entropy is defined as the entropy of the current system with the system R_1 state in balancing the ratio of the entropy value R_2.

$$y = \frac{R_1}{R_2} = \frac{yIn\Omega_1}{yIn\Omega_2} = \frac{In\Omega_1}{In\Omega_2} \quad (3)$$

Where, Ω_1 is the number of the current system of microstates, Ω_2 for the system in equilibrium when the number of microstates.

Cloud computing system generalized normalized entropy is approximate calculated as follows: Suppose you want to analyze cloud computing nodes is m servers, we compared the parameter space is divided into a grid of $n \times n$, m server's current working parameter mapping phase space will inevitably fall to the parameter $n \times n$ grid, in the equilibrium state m servers will fall in a random manner a $n \times n$ grid, while far from equilibrium when m servers will fall into a grid, then $1 \le j \le n \times n$, shown in Figure 1, Figure 1 (a) when the server parameters for the equilibrium phase space distribution of the parameter, Figure 1 (b) a non-equilibrium state server parameters in the parameter phase space distribution.

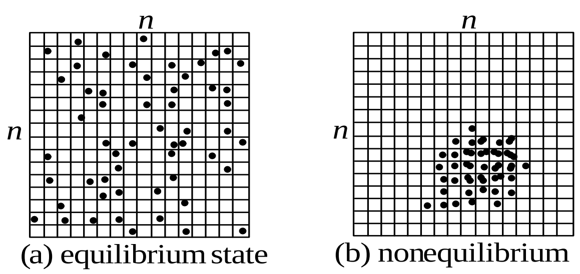

Figure 1. Equilibrium and Non-equilibrium State When the Server Parameters in the Parameter Phase Space Distribution of Contrast

When in equilibrium m-server parameters fall into a grid $n \times n$ number of states $\Omega_2 = (n \times n) * m$; non-equilibrium we approximate that m-server parameters will fall l grids , the states $\Omega_1 = l_m$. So when one has m servers cloud computing server parameters in the parameter phase space l just fell into a grid, the system of generalized entropy can be normalized

$$y = \frac{Inl^m}{In(n \times n)^m} = \frac{Inl}{In(n \times n)} \quad (4)$$

As $1 \le l \le n \times n$, so that the normalized entropy broad range of [0,1]

The actual calculation of the grid can be adjusted to obtain different densities normalized entropy generalized approximation, generally larger number of servers can be used a more dense grid, less the number of servers can be reduced the density of the grid. Entropy is a statistic because we usually make in meshing $m \succ n \times n$, so as to ensure statistical validity, while the number of nodes in the server is too small, such as the case when only a few nodes, generalized normalized reference value of entropy is not big.

In the case of massive node can draw the following referenced conclusions.

Conclusion2. In the massive cloud computing system to node n-dimensional phase space projection parameters, sub-optimal projection point does not appear in the phase space of events for the small probability event.

Proof, According to the previous proof, n-dimensional phase space parameters, through the center of gravity can be divided into $2n$ sub-phase space, then the probability of either a server sub-optimal projection point does not appear in the phase space is:

$$1 - \frac{1}{2^n} \quad (5)$$

The probability of m servers' projection point is not appeared in sub-optimal phase space is:

$$\left(1 - \frac{1}{2^n}\right)^m \quad (6)$$

Taking into account the cloud computing system the number of nodes is very large, so we can set $m \square 1$, and $m \square n$, then there

$$\left(1 - \frac{1}{2^n}\right)^m - \succ 0 \quad (7)$$

The event of M servers' projection point is not appeared in sub-optimal phase space for small probability. For example, when $n = 2$, $m = 1000$, the server does not appear in the projection point sub-optimal phase space probability of only 1. $15 \times 10 - 126$, QED.

Conclusion 2,Show that the mass node scheduling system usually can find the comparatively lightly loaded node with the high probability from the optimal sub-phase space and assign to the external load requests.

The establishment of conclusion provides a theoretical basis for phase space we designed scheduling algorithms only consider the phase space from the sub-optimal allocation of servers. In sub-optimal phase space we are based on the projection of the center of gravity from the server as a sort of basis for assigning priority queue, in sub-optimal phase space nodes farther away from the center of gravity lower consolidated load, the load allocation priority allocation.

2. Phase Space Scheduling Algorithm

As cloud computing phase space analysis model has profound and intuitive features, we propose a cloud computing system phase space scheduling algorithm; this algorithm is also behind the simulation algorithm in comparative examples.

Basic phase space scheduling algorithm is implemented as follows:

(1) Determine the working status of the mapping parameters of the number of parameters that space dimension n, construct the corresponding parameter phase space.
(2) All nodes in cloud computing system project to the parameter phase space.
(3) Calculating the position of center of gravity $G(x_0 ... y_0, z_0)$, which is the projection of current parameters space.
(4) Center of gravity position and system settings on the critical value, if the center of gravity of the system exceeds the threshold, the system begins loading the police, through continuous release of waiting tasks, cancel the alarm again after jump to step 3, otherwise the next step.
(5) The parameter phase space is divided into $2n$ sub-phase space, and mark the best sub-phase space and worst sub-phase space, in order to determine the optimal sub-phase space in all the projection points.
(6) Calculate the phase space in sub-optimal projection point and the center of gravity of all the distance d_i , d_i values at descending node to be assigned priority queues formed load-queue, the d_i node will be preferentially allocated to the load request to use, while the formation of all the tasks to be assigned the task assigned priority queue task queue, the request is arranged in chronological order, FIFO.
(7) The task and nodes in Load-queue and task-queue were paired until load queue been allocated, jump to step 3.

3. Simulation Experiments and Analysis

3.1. Simulation Model and Parameter Description

(1) phase space simulation model cloud computing system of cloud computing systems in the phase space model, the working status of a server can be used to describe a state vector of parameters (parameter 1, parameter 2, parameter 3, ...), such as (CPU occupancy rate of 0.3, memory usage of 0.2, the number of connections occupancy of 0.1, ...), the parameter vector dimension corresponds to the dimensions in the phase space, the server parameter vector projected onto the corresponding phase space, end position of vector in the phase space on behalf of the server's work states, the size of the vector represents the state of the server forming a plurality of integrated work load. End vector of server parameter changes over time in the phase space into particle movement.
(2) the method ,an external load request one, which of simulation model used vector model to describe the phase space projection, provides a convenient for the establishment of an external load simulation, also uses the same external load request load parameter vector to describe, but more a time slice length parameter, this parameter is the load request describe the length of time occupied by the server, the time slot length is the length of a relative time unit, the duration of the absolute time length is not used during the simulation in order to realize a computer simulation in the general experimental a large number of requests per unit time of generation and processing load. Each unit length of time the number of requests issued by the load simulation system pressure concurrent access, through per unit of time, the load is reduced one time slice. Load parameter vector in each component of the value of the range distribution may reflect different external application resource demand characteristics. The CPU occupancy rate, for example, the low occupancy rate of concurrent requests on the Web can be described most of the computation request, the high occupancy rate of concurrent requests can describe some numerical requests. Therefore, through different parameters and component ranges can be achieved with a few adjustments for different situations simulated external requests.
(3) the simulation parameters description in the simulation experiments comparing the proposed phase-space scheduling algorithm, polling algorithm, minimum load-first algorithm and dynamic load balancing algorithm, applied to the feedback of cloud computing scheduling effect after the simulation model, for convenience of the simulation analysis, we only consider two parameters projection (CPU usage and memory usage) of the two-dimensional case, the specific experimental parameters are shown in Table 1, the entropy in table is the generalized normalized entropy, its value can reflect scheduling effect in this cloud system simulation model scheduling algorithm, usually generalized normalized entropy value is smaller load on the system more balanced. Generalized table normalized entropy is 10×10 grid density using equation (2) calculated; sampling time position are 100 time chip length; broad temperature is calculated using the formula (1) is calculated The; single load task takes take a random integer ranging from 1 to 10 chip length; Load occupancy rates ranging from 0 to 0.01 and 0 to 1 are two cases, corresponding to the small load range and large load variations.

3.2. Performance Evaluation

Figure 2 is a section of 1000 points, a single task load range from 0 to 0.01, per unit time task requests chip 20000 times the number of simulation results, this parameter is a small load corresponding to the state of high pressure access, Fig. 3 (a) ~ (d) are scheduling result projection of phase space scheduling algorithm, polling algorithm, minimum load-first algorithm and dynamic load balancing feedback algorithm in the parameter phase space (upper) and momentum phase space (lower). Contrast distribution of the projection point of four algorithms results in the parameter phase space, it is clear that ,under the action of phase space scheduling algorithm ,node in the parameter phase space projection showing a good state of aggregation, suggesting that the current system load more balanced, which states in the thermodynamic sense, is a far from equilibrium state. Simultaneous contrast cloud computing system in the momentum phase space projection found that the phase space of the system scheduling algorithm scheduling phase space momentum gathered in the phase space projection point of a small area of origin, in the momentum phase space projection points closer to the origin indicates the current system dynamic working condition more stable, and load balancing state is also better, because the projection point in the momentum phase space near the origin of the phase space in the parameter corresponding to the server's slower movement and closer to the center of gravity position. In comparison, result of the polling algorithm, minimum load-first algorithm and dynamic load balancing algorithm feedback in the parameter phase space and momentum phase space scheduling are not better than the phase space projection scheduling algorithms, from Table 1, we can see that the phase-space algorithm, polling algorithm, minimum load-first algorithm and dynamic load balancing algorithm feedback at this time generalized normalized entropy were 0.350, 0.502, 0.543, 0.500, phase space algorithm enables the system maintains low entropy, to ensure the system is in good load balancing state. While under the effect of the phase space algorithm for generalized cloud computing system temperature is only 0.0213, lower than the polling algorithm, minimum load-first algorithm generalized temperature, indicating that the phase-space algorithm for job stability at this point is better than polling algorithms and minimum load-first algorithm. Dynamic load balancing algorithm feedback generalized phase space temperature is lower than 0.0173 algorithms, dynamic feedback time job stability of load balancing algorithm is superior to phase-space algorithm, but due to dynamic load balancing algorithm feedback generalized normalized entropy to higher than the phase-space algorithm, which indicates that the load balancing algorithm scheduling dynamic feedback system under load balancing algorithm is not as good as phase space. By operating results projected onto the server parameter phase space and momentum phase space, we can describe operation of the system from the perspective of both static and dynamic, projection point in the phase space distribution of binding different parameter values can be generalized thermodynamic react intuitively the overall state of the system, phase space analysis of massive high coupling cloud computing system has a strong ability to describe, so the phase space analysis is a powerful tool for analysis of cloud computing system.

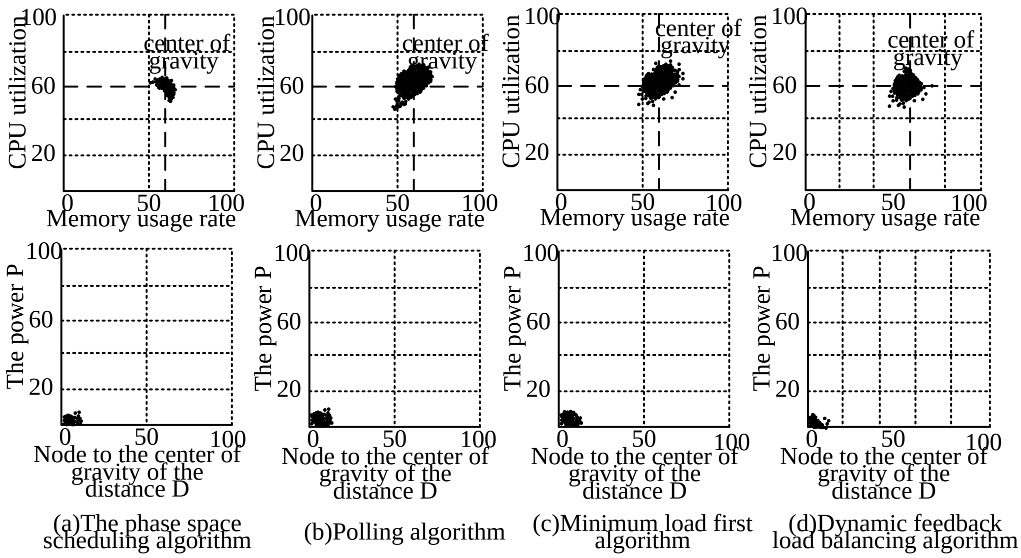

Figure 2. Is 1000 Points; a Single Task Load of 0 to 0.02, the Task Request is 20000 Times Simulation Results

Figure 3 is a section of 100 points, a single task load range from 0 to 0.02, per unit time task requests chip 1600 times the number of simulation results, Fig 4 (a) ~ (d), respectively, the scheduling result projection of the phase spatial scheduling algorithm, Polling algorithsm, minimum load-first algorithm and dynamic load balancing algorithm feedback in the parameter phase space (upper) and momentum phase space (lower). Figure 3 shows the comparison, under the experimental conditions; scheduling effect of this scheduling algorithm is superior to the polling scheduling algorithm, minimum load-first algorithm and dynamic load balancing algorithm feedback. Under the action of the phase space of the system scheduling algorithm generalized normalized entropy of 0.240, 0.0198 generalized temperatures is better than the other three algorithms, and we also draw a visual comparison of algorithm execution effects from distribution of the projection point in parameter phase space and momentum phase space, phase space algorithm under cloud computing system in the parameter space of the center of gravity relative to aggregation and momentum phase space in the aggregation of the origin are better than other three algorithms. The projection of phase space can be described qualitatively intuitive movement system; generalized thermodynamic parameters can be given a quantitative comparison scheduling effect of macroscopic parameters.

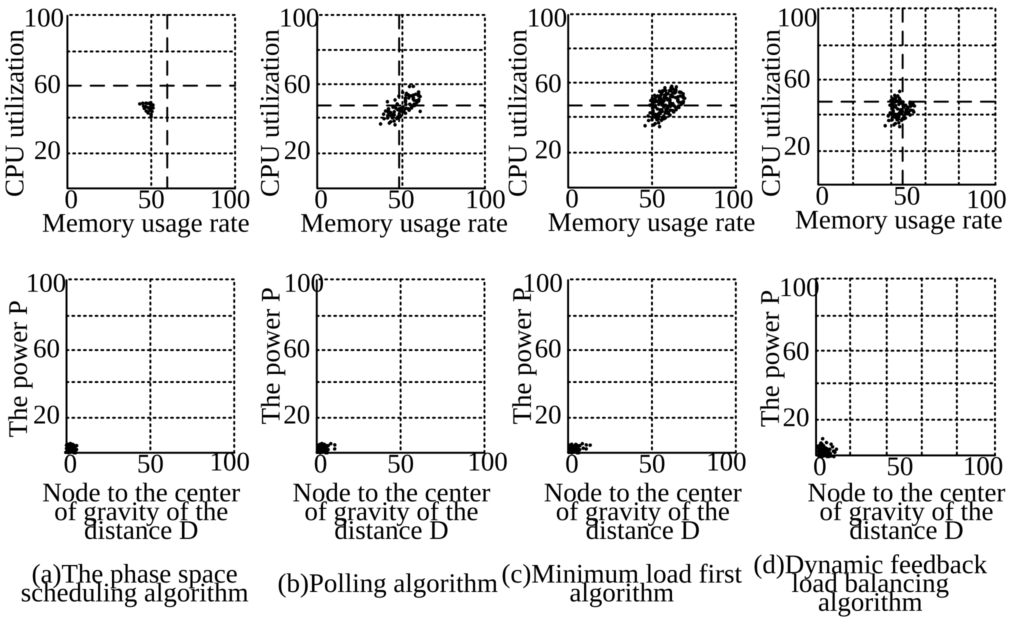

Figure 3. Is 100 Points, a Single Task Load of 0 to 0.02, Task Requests, a Simulation Result for the 1600 Times

Figure 4 is a section of 100 points, a single task load range from 0 to 2, unit time slice task requests is 16 times the simulation results, Figure 4 (a) ~ (d), respectively the projection of scheduling result of phase space scheduling algorithm, polling algorithm, minimum load-first algorithm and dynamic load balancing algorithm feedback in the parameter phase space (upper) and momentum phase space (lower).

This simulation conditions external load requests varied greatly and the load limit of 1, then the parameters of the server in the movement space of the larger mean free path. Compare the previous experiment, in which case the generalized temperature is higher, four algorithms 0.0850,0.1243,0.0850,0.1532 generalized temperatures were, compared to Figure 3, Figure 4 and Figure 5 , a large range of load changes in the outside world when the effects of four algorithms scheduling phase space has emerged in the parameter random distribution of cases, the phase space projection server load balancing algorithm has not been effectively gathered, the system is in thermodynamic equilibrium, the server the load is not balanced, the results are also shown in Table 1 are reflected in the entropy, broad four algorithms

Righteousness normalized entropy were 0.904, 0.850, 0.905, 0.890, these values are very close to the generalized normalized entropy maximum value of 1. The momentum from Figure 5, the phase space diagram, we also see the server in the momentum projection points on the phase space is no longer gathered at the origin, which indicates the current operating status of the system is very unstable, and the system load is uneven.

The effect differences of scheduling algorithm under the mean free path and different temperatures is full compliance with thermodynamic principles, the thermodynamic temperature diffusion of gas molecules has a strong tendency, it is very difficult to maintain the system in a non-equilibrium state. The same principle, cloud computing system generalized phase space very high temperatures scheduling algorithm to maintain the system in the phase space in a non-equilibrium state is very difficult, which appeared in the same generalized scheduling algorithm at different temperatures appears different scheduling effect results. In the generalized system

needs higher temperatures to pay a higher price to pay to keep the system calculate load balancing, scheduling algorithms should be in the phase space of this
Circumstances make the appropriate optimization.
Based on the above analysis, we can get the cloud computing system in the parameter phase space and momentum projection points on the phase space form two qualitative conclusions:

(1) Cloud computing system in the parameter phase space projection points on the aggregation of the better system load balance better.
(2) Cloud computing system in the momentum phase space projection points on the origin of the phase space to gather the better system stability and load balance work better.
(3) Through the above two qualitative conclusions we can directly from the parameters of phase space and momentum phase space readout system on the current operating status of phase space that role cloud computing can be applied to the design of dedicated monitoring system, and the system The more nodes the more obvious advantages of this method, which is precisely the advantage of the thermodynamic approach.

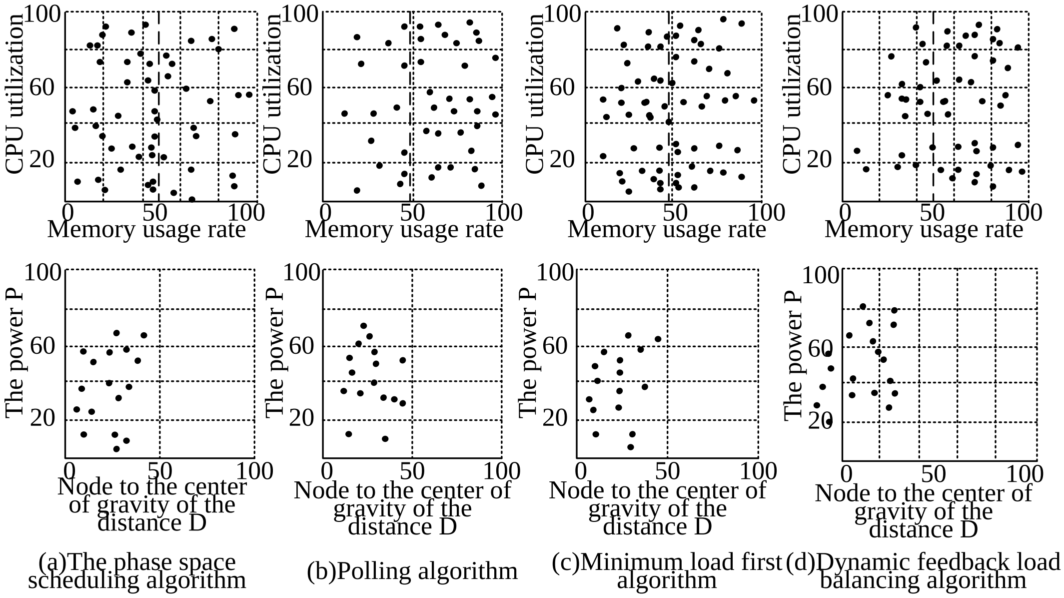

Figure 4. Simulation Result of 100 Points in a Single Task Load Which is 0 to 1 and the Task Request is 16

4. Conclusion

This article makes the cloud computing system project to phase-space and defines and analyzes the generalized phase space temperature, generalized normalized entropy, generalized gravity, generalized normalized momentum, etc. generalized thermodynamic parameters. The thermodynamic principles are used to establish the basic theoretical framework of a cloud computing system phase space analysis model. This theory is suitable for analysis of mass nodes with high coupling cloud computing system. The phase-space analysis are used to make the implementation of cloud computing system project to the parameters phase space and momentum phase space and provide a great convenience for the establishment of cloud computing simulation model, research of scheduling algorithm and analysis of system status. The phase space scheduling algorithms are also proposed, for cloud computing system phase space simulation model based on the phase space analysis model to realize the simulation experiment for multiple scheduling algorithms on the phase space simulation model. The simulation results show that the phase-space scheduling algorithm can effectively makes the cloud computing system in space to maintain the sense of non-equilibrium thermodynamics and achieve load balancing system. And also the results demonstrate that the phase space analysis model has unique advantages in system analysis, simulation and algorithm design.

References

[1] Jian-hua WANG, Yan YU, Madukasi Charles Nnaemeka. A High Throughput MAC Protocol for Wireless Sensor Networks in Surveillance Applications. Journal of Networks, Vol 8, No 9 (2013), 2005-2012

[2] Zhang Weiqiang, Xu Chen, "A Denoising Algorithm of Wavelet Threshold Based on Translation invariance, Modern Electronics Technique, Vol.6, 2003, pp. 29-31.

[3] W. Xing and B. Yuan, "3D part based structural description extracting and modeling", In: Proceedings of the International MultiConference of Engineers and Computer Scientists 2012 (IMECS 2012), 14-16 Mar., 2012, Hong Kong. vol. I, pp.209-212, 2012.

[4] S. Zheng and Y. Qi, "Estimation of coronary arterial dynamics from X-ray angiographic sequences", Computerized Medical Imaging and Graphics, vol. 35, pp. 353-364, July 2011.

[5] J. He, Y. Geng and K. Pahlavan, Modeling Indoor TOA Ranging Error for Body Mounted Sensors, 2012 IEEE 23nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), Sydney, Australia Sep. 2012 (page 682-686)

[6] Li Ya-feng, Wang Bao-bao, "New method for partial discharge signals detection," JOURNAL OF XIDIAN UNIVERSITY, V01. 37, No. 3, pp. 530-533, Jun.2010.

[7] A. Hernàndez-Sabaté, D. Gil, J. Garcia-Barnés, E. Martí, "Image-based cardiac phase retrieval in intravascular ultrasound sequences", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 58, pp.60-72, Jan. 2011. http://dx.doi.org/10.1109/TUFFC.2011.1774 PMid: 21244975

[8] S. Zheng, T. Meiying, S. Jian, "Sequential reconstruction of vessel skeletons from X-ray coronary angiographic sequences", Computerized Medical Imaging and Graphics, vol.34, pp. 333-345, July 2010.