Future Generation Computer Systems 74 (2017) 132–141

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

DataABC: A fast ABC based energy-efficient live VM consolidation policy with data-intensive energy evaluation model[☆]

Jianhua Jiang[a], Yunzhao Feng[a], Jia Zhao[b], Keqin Li[c,*]

[a] School of Management Science and Information Engineering, Jilin University of Finance and Economics, Changchun, PR China
[b] School of Computer Science and Engineering, Changchun University of Technology, Changchun, PR China
[c] Department of Computer Science, State University of New York, New Paltz, NY, USA

HIGHLIGHTS

- Address energy and performance constrained VM scheduling with data-intensive jobs.
- Propose an energy-efficiency evaluation model with two influencing factors: CPU and GPU utilization rates.
- Develop an artificial bee colony algorithm to optimize VM selection and allocation policies.
- Perform extensive simulations to verify the analytical results.

ARTICLE INFO

Article history:
Received 16 September 2015
Received in revised form
6 April 2016
Accepted 7 May 2016
Available online 18 May 2016

Keywords:
Artificial bee colony algorithm
Energy-efficiency
Green data center
Virtual machine consolidation
Virtual machine migration

ABSTRACT

Live Virtual Machine (VM) consolidation is an effective method of improving energy-efficiency level in green data centers. Currently, to evaluate energy consumption in green data centers, energy-efficiency evaluation model with CPU utilization rate has been proposed. However, it is not suitable for data-intensive computing due to great energy consumption by GPU-intensive processing. In this paper, we have proposed a new energy evaluation model with CPU and GPU utilization rates. There are two kinds of policies in live VM consolidation: one for VM selection and the other for VM allocation. Some researchers have proposed their solutions based on VM selection policy or VM allocation policy respectively. However, it will be a better energy-efficiency VM consolidation policy if these two polices are integrated together. Based on these two policies, a fast Artificial Bee Colony (ABC) based energy-efficiency live VM consolidation policy with data-intensive energy model, named as DataABC, is proposed. DataABC adopts the idea of Artificial Bee Colony algorithm to get a fast and global optimized decision of VM consolidation. Compared with two state-of-art policies of PS-ABC and PS-ES, the total energy consumption of DataABC evidently drop by 9.72% and 5.84% respectively. As a result, based on the ESV metric, the DataABC approach has proved that (a) the energy-efficiency evaluation model with data-intensive computing is valid and that (b) DataABC can save energy with a good Quality of Service (QoS) in green data centers.

1. Introduction

A large number of data centers worldwide consume huge energy. A report [1] from China says that energy consumption by data centers accounted for about 1.3% of the global total energy consumption in 2010. Based on the estimation from Gartner Group in 2013, the energy consumption takes about 10% proportion of the total operation cost in the first year, and about 50% in the next few years [2]. It is necessary to find an effective approach to minimize energy consumption with the low violation rate of Service Level Agreements (SLA) in data centers. Live Virtual Machine(VM) consolidation policy is an important method to find its balance between performance and energy consumption in data centers [3,4]. Some heuristic algorithms [5–7] have been proposed to save energy.

☆ This research is supported by the National Natural Science Foundation of China under Grant Nos. 61202306, 61472049, 61402193, 61572225, and Foundation of Jilin University of Finance and Economics under Grant Nos. XJ2012007 and 2013006.

* Corresponding author.
E-mail address: lik@newpaltz.edu (K. Li).

http://dx.doi.org/10.1016/j.future.2016.05.013
0167-739X/

First, energy-efficiency evaluation models for computing-intensive applications have been proposed before. However, it is not reasonable to construct energy-efficiency evaluation models based on the single influencing factor with CPU utilization rate, since there are huge amount of data-intensive jobs in data centers. As we all know, GPU is also an important energy consumption unit in a host with a huge amount of data-intensive jobs. Therefore, GPU should be taken as an important influencing factor in an energy-efficiency evaluation model with data-intensive jobs [8].

Second, different VM selection or allocation polices have been designed. There are two steps in VM consolidation: VM selection and VM allocation [9]. However, the importance of cooperation between these two phases is neglected. Furthermore, it is unreasonable to make only one VM migration decision in Δt. To meet the requirement of data-intensive jobs in data centers, m VM consolidation decisions should be made parallelly in Δt. Therefore, a fast and parallel energy-efficiency VM consolidation policy is required to be proposed to get a good tradeoff between performance and energy consumption in data centers.

In this paper, a fast ABC based energy-efficient live VM consolidation policy with data-intensive energy evaluation model, named as DataABC, is proposed in data centers. DataABC adopts the idea of an artificial bee colony algorithm [10] to get a fast global optimized VM consolidation decision with the features of self-organization and collaboration.

The key contributions of this paper are as follows.

- A novel energy-efficiency evaluation model is proposed based on influencing factors of CPU and GPU utilization rate for data-intensive jobs.
- Different from state-of-art VM consolidation policies, ABC-based idea is applied to both VM selection and VM allocation policy with the characteristics of self-organization, collaboration and rapid convergence.
- Compared with other classical or state-of-art VM consolidation policies in CloudSim3.0, DataABC achieves the goal of a higher energy-efficiency level.

The rest of this paper is organized as follows. In Section 2, we present related work supporting our proposed approach for minimizing energy consumption in green data centers. In Section 3, firstly, the proposed problem is given; secondly, its formulation is presented. In Section 4, some discussions are made to explain reasons for simulation performance of DataABC algorithm and its implementation in detail. In Section 5, performance evaluation of DataABC is evaluated in CloudSim3.0. Conclusion and future work are given in the last section.

2. Related work

First, from the perspective of energy-efficiency evaluation model in data centers, some researchers have proposed their energy consumption models by the single influencing factor of CPU utilization rate [3,4,9].

Fan et al. [11] have found a strong correlation between single server and the CPU utilization rate, i.e., there is a direct linear correlation between the CPU utilization rate and the server energy consumption of servers. However, the energy-efficiency evaluation model will not be good when there are a huge number of data-intensive jobs in data centers.

Beloglazov et al. [6] have proposed an energy consumption evaluation model, which is according to CPU, RAM and network bandwidth utilizations. However, it does not consider the influencing factor of data-intensive jobs in a data center.

In data centers, there are a huge amount of data-intensive jobs, which lead to great energy consumption due to GPU-intensive computing. To get a better energy-efficiency evaluation model, the GPU utilization rate should be adopted as an important influencing factor.

Second, from the perspective of VM consolidation policies, some policies have been proposed from VM selection and VM allocation respectively. The related polices are given as follows.

Nathuji et al. [12] have proposed a novel VM energy controller named as VirtualPower. It can reduce energy by itself on the virtual homogeneous and heterogeneous platforms. However, it does not switch off or hibernate hosts whose workloads are idle or low. In our proposed approach, idle or low workload hosts will be hibernated to save energy, and there will be a decreased probability to wake them up.

Beloglazov et al. [6] have proposed an approximated pack problem algorithm about VM consolidation. Its main theories are firstly ranking the VMs in descending order with the CPU utilization rate, secondly allocating each VM to per host by turns, thirdly calculating the increased energy, and finally choosing the target host with minimum increased energy. In addition, they have proposed the minimal migration algorithm and maximum growth potential algorithm.

Fei et al. [13] have proposed a strategy of distributed VM management. It selects the VM in the host of high CPU utilization rate to migrate by monitoring hosts in local management and selects the appropriate host by global management. However, it just considers computing-intensive jobs rather than data-intensive jobs. The policy of VM selection and VM allocation based on the threshold trigger mechanism, and its value of threshold cannot be dynamically. In our proposed approach, the energy-efficiency evaluation model is conducted with behaviors of data-intensive jobs. Furthermore, live VM migration is triggered by the ABC-based policy, which is much more reasonable.

Jin et al. [14] have proposed a novel energy-efficiency approach in data centers, which adopts the real and dynamic VM redistribution mechanism to migrate VMs for improving the utilization rate of resources in data centers. It ranks hosts by their CPU utilization rate, migrates VMs from minimum utilization rate hosts to the optimal hosts on the principle of minimum incremental power and switches off some idle hosts to save energy. It is unreasonable to set global optimal of the upper and lower threshold of the CPU utilization rate in hosts, because that it is set statically.

Zhao et al. [15] have proposed a novel heuristic approach, called PS-ABC. It combines three ideas into the artificial bee colony (ABC), such as the uniform random initialization, the binary search and Boltzmann selection policy to achieve an improved ABC based approach. It can improve the global optimization capability and the local optimization capability. However, the importance of VM selection is neglected. Migrant VMs are not selected by ABC-based optimization. Zhao et al. [7] have proposed a novel heuristic approach named as PS-ES. It combines the PSO (particle swarm optimization) idea with the SA (simulated annealing) idea to achieve an improved PSO-based approach with the better global searches ability. And it uses the Probability Theory, Mathematical Statistics and the SA idea to deal with the data obtained from the improved PSO-based process to get the final solution. However, the energy-efficiency evaluation model is not redesigned when considering different types of jobs in data centers.

All in all, current live VM consolidation policies have been proposed to save energy. Among them, the energy-efficiency evaluation model is conducted by the single influencing factor of CPU utilization rate without considering the behavior characteristics of data-intensive jobs in data centers. Furthermore, ABC idea is considered as a good idea to get an optimized VM allocation policy. However, ABC based VM selection policy and ABC based VM allocation policy are not considered integrated to achieve a good global optimal.

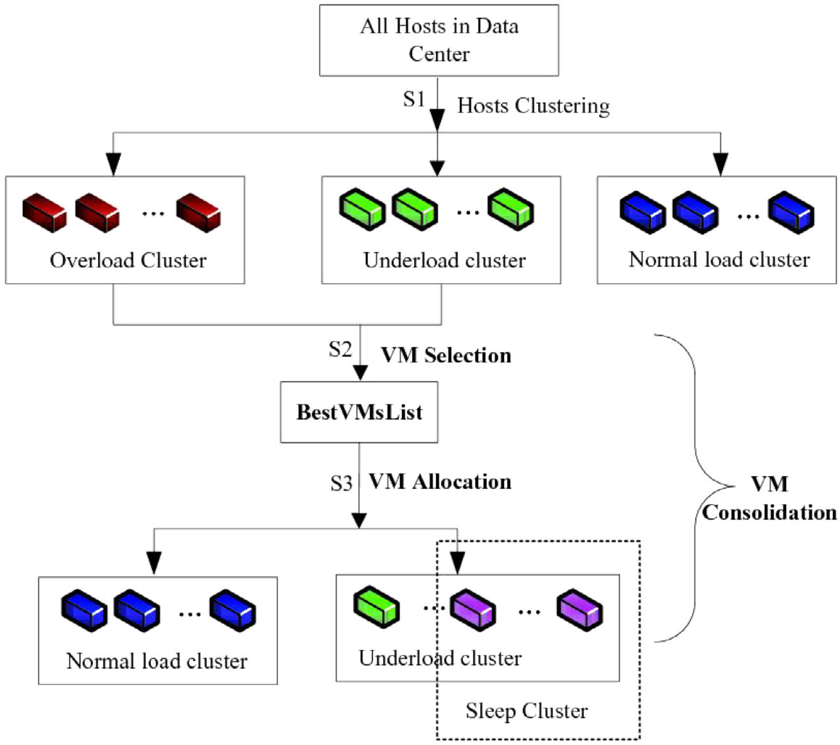


Fig. 1. In the figure of architecture of VM consolidation, VM consolidation has two steps of VM selection and VM allocation. Before making decision of VM consolidation, all hosts should be clustered into overload cluster, underload cluster and normal load cluster in the steps of hosts clustering.

3. Problem definition

3.1. The proposed problem

In data centers, jobs can be classified into computing-intensive jobs and data-intensive jobs based on their different behaviors. In [16], it was mentioned that GPU is as important as CPU in energy consumption. However, current energy-efficiency evaluation models in VM consolidation are based on CPU utilization rate only. Obviously, it is not reasonable. Therefore, the first problem is to design a new energy-efficiency evaluation model with influencing factors of CPU and GPU utilization rates.

VM consolidation includes VM selection and VM allocation. Currently, some novel energy-efficiency VM allocation policies have been proposed. However, there are two problems need to be solved. One is that energy-efficiency VM migration in VM allocation needs a quick heuristic algorithm to allocate m VMs to m target hosts in Δt. The other one is that VM selection and VM allocation policy should be considered together to find a live energy-efficiency VM consolidation policy.

3.2. Problem formulation

As can be seen from Fig. 1, there are n hosts in a green data center. All hosts in the data center can be represented by $H = \{t, h_1, h_2, \ldots, h_k, \ldots, h_n\}$, where t refers to the current timestamp and h refers to each host. It is obvious that RAM utilization rate has a high positive correlation with the GPU utilization rate in the data center. To evaluate the factor of GPU utilization rate, we took RAM utilization rate as a substitute variable in cloudsim3.0 [10]. Therefore, each host h can be represented $h = \{ID, CPU_{rate}, RAM_{rate}\}$. In a green data center, the state of hosts H is varied dynamically in real time in the light of corresponding workloads. We assume the state of hosts will not be varied in a time window Δt. Some other assumptions are also given as below: There are three steps labeled as S1, S2 and S3 in Fig. 1. In step of S1, the *k-means* approach is adopted to divide all hosts into 3 clusters, such as *overload*, *underload* and *normalload* cluster. $C_{overload}$ is a set of x available physical hosts denoted by $C_{overload}(x, t) = \{h_{o1}, h_{o2}, \ldots, h_{ox}\}$ at time t. $C_{underload}$ is a set of y available physical hosts denoted by $C_{underload}(y, t) = \{h_{u1}, h_{u2}, \ldots, h_{uy}\}$ at time t. $C_{normalload}$ is a set of z available physical hosts denoted by $C_{normalload}(z, t) = \{h_{n1}, h_{n2}, \ldots, h_{nz}\}$ at time t, where $x + y + z = n$. The step of S1 can be denoted as $K-means(n, k, t) = \{H, C_{overload}, C_{underload}, C_{normalload}\}$, where $k = 3$.

In step of S2, it implements the operation of VM selection. Its objective is to find m VMs from the *overload* cluster and the *underload* cluster. m is the parameter of VM selection, and it is assumed that m VMs are selected from m hosts respectively. The problem of VM selection is to find the *BestVMsList* with its energy-efficiency evaluation model. A five-tuple $VMSelection = \{C_{overload}, C_{underload}, BestVMsList, E, ABCS\}$ for the proposed problem scenario is defined. *BestVMsList* is a set of m VMs which come from m hosts respectively. It can be denoted by $BestVMsList(m, t) = \{t, hvm_1, hvm_2, \ldots, hvm_m\}$, where t is the start time, hvm is a vector of $\{h, vm\}$ and vm is the best migrant VM in host h. $E(m, t, \Delta t) = \{E_1, E_2, \ldots, E_m\}$ is the energy consumption by the m physical hosts in a resource pool. ABCS refers to the proposed VM selection policy with ABC algorithm.

$$TotalE(t, \Delta t) = \Sigma(E(Host_i, t, \Delta t)) \quad (1)$$

where $E(Host_i, t, \Delta t) = \Sigma(E(VM_j, t, \Delta t))$.

$TotalE(t, \Delta t)$ denotes the energy consumption in a data center within Δt time from the start time t.

The energy consumption of VM in each host can be denoted by $E(VM_j, t, \Delta t)$,

$$E(VM_j, t, \Delta t) = E_{(cpu,j,t,\Delta t)} + E_{(gpu,j,t,\Delta t)} + E_{(punish,j,t,\Delta t)} \quad (2)$$

where $E_{(cpu,j,t,t)}$ refers to energy consumption occurred by CPU utilization in VM_j in Δt time from the start time t. $E_{(gpu,j,t,\Delta t)}$ is caused by GPU processing with a huge number of data-intensive jobs. In CloudSim3.0, we use the RAM utilization rate in place of GPU utilization rate. There are two reasons. One is that it will be more convenient to measure GPU utilization rate in the CloudSim3.0, and the other one is that the correlation between the GPU utilization rate and the RAM utilization rate is highly positive. Therefore, formula (2) can be rewritten with formula (3).

$$E(VM_j, t, \Delta t) = E_{(cpu,j,t,\Delta t)} + E_{(ram,j,t,\Delta t)} + E_{(punish,j,t,\Delta t)} \tag{3}$$

where $E(ram, j, t, \Delta t)$ refers to energy consumption of VM_j caused by GPU processing, $E(punish, j, t, \Delta t)$ refers to extra energy consumption brought by unreasonable VM selection policy.

$$E(punish, j, t, \Delta t) = P(\lambda) \times \Delta E \tag{4}$$

where $P(\lambda)$ refers to the probability of Poisson distribution, and ΔE is a constant value with relative higher energy cost.

In step of S3, VM allocation is executed. It can be formulated as selecting m hosts from the other candidate target hosts. An m-dimensional vector of target hosts is adopted to represent a solution of this proposed problem. The *No.* of each element in m-dimensional vector of target hosts represents the location of each host, and the element is the target host. The step of S3 can be denoted as $VMAllocatin(m, t) = \{C_{normalload}, C_{underload}, C_{sleep}, BestVMsList, E, ABC\ A\}$, where $C_{sleep} = \{h_{s1}, h_{s2}, \ldots, h_{st}\}$, where $t \leq y$, represents the subset of the *underload* cluster that should be hibernated. ABCA refers to a proposed VM allocation policy with ABC idea. Similar to the VM selection policy, it should be optimized with the energy-efficiency objective. Metrics of a VM allocation policy includes performance metrics and energy consumption metrics. As for performance metrics, $SLA(t, \Delta t)$, $SLAV(t, \Delta t)$ and $ASLAV(t, \Delta t)$ can be used to evaluate performance of a data center. A metric SLA violation rate specifies the violation rate of the Service Level Agreement (SLA), which is described in the policy tables, and ASLAV refers to the average of SLA violation rate. As for energy consumption metrics, $TotalE(t, \Delta t)$ is to measure the energy consumption in the data center in a Δt time from the start time t. To evaluate the proposed VM allocation policy of ABCA, we use $ESV(t, \Delta t)$ to measure its energy-efficiency level.

$$ESV(t, \Delta t) = TotalE(t, \Delta t) \times SLAV(t, \Delta t). \tag{5}$$

To achieve an energy-efficiency VM allocation policy, it is necessary to minimize $ESV(t, \Delta t)$.

4. Methods

To get an energy-efficiency live VM consolidation policy in data centers, an ABC-based policy is adopted in this paper. To achieve a high energy-efficiency level, we focus on three sub-problems, such as energy-efficiency evaluation model for data-intensive jobs, ABC-based energy-efficiency live VM selection policy and ABC-based energy-efficiency live VM allocation policy.

4.1. Energy-efficiency evaluation model for data-intensive jobs

There are a huge number of data-intensive jobs in data centers. The paper [16] has found that energy consumption ratio of GPU and CPU is equivalent to each host in a data center. The energy-efficiency evaluation model, based on a single influencing factor of CPU utilization rate, is not suitable in data centers. To evaluate energy consumption in data centers, which have a huge number of data-intensive jobs, energy-efficiency evaluation model should be modeled with both CPU and GPU utilization rate. The energy-efficiency evaluation model is proposed in Section 3.2 with formula (1)–(4).

4.2. ABC-based energy-efficiency VM selection policy

To reduce energy consumption in data centers, selecting a reasonable VM to be migrated out is the key point. It is obvious that the underload host will not be hibernated quickly to save energy if there is a wrong VM selection decision. Similarly, the overload host will not be changed quickly into normal load status if there is an unreasonable VM selection decision. Furthermore, unreasonable VM selection decisions will increase the frequency of VM migrations. Therefore, a fast, convergent and global optimized VM selection policy is necessary. The major features of ABC are fast convergence and global optimization. The idea of ABC can help bees to find the best food source quickly. Similarly, we take the source hosts as a candidate solution to the VM selection policy. The solution can be formulated by selecting m hosts from all candidate hosts in a data center. In this paper, candidate hosts come from underload and overload hosts. To get m best global VMs to be migrated from the candidate hosts in Δt, an ABC-based idea is necessary to be applied to the VM selection policy.

ABC is a colony intelligence algorithm, which focuses on cooperation based on division of bee scouters, bee followers and bee employers. Therefore, it is crucial to the VM selection policy to find the best global optimal for their corporation. A two-dimensional solution space can be generated from CPU and RAM utilization rate of VMs in underload and overload clusters. Each VM is a point in the two-dimensional solution space. k bee scouters can be dispatched randomly to k points to check their predicted incremental energy consumption. After comparison with these k points values, m minimal points are selected to explore m best global optimized points. Some bee followers will be sent to check their predicted incremental energy consumption and to calculate their fitness value. Other bee followers will be dispatched according to their fitness value until m best global optimized points are found.

It is difficult to introduce the idea of ABC into the VM selection policy, because the VM allocation policy and workload balance in a data center should be taken into consideration when finding m best migrant VMs to reduce the frequency of VM migrations. As for workload balance, it is a good decision if VM migrations are occurred from an overload host to a normal load host or from an underload host to a normal load host. As for VM allocation policy, candidate target hosts should be selected in advance to avoid unnecessary VM migrations. Therefore, a good VM selection policy will help to decrease the frequency of VM migrations. However, traditional VM selection policies generate a huge number of VM migrations. Its major reasons are as follows.

- **The dynamic change of workload between source and target host is not considered when making a VM selection policy**. The migrant VMs in overload target host cannot be processed smoothly, if VM migrations occurred from normal load or underload hosts to the target overload hosts. As a result, these migrant VMs will be migrated out to find another target hosts. The ABC based VM selection policy must make sure that the migration route is from overload or underload hosts to normal load hosts. Therefore, the change of workload among different load hosts should be considered when making a VM selection policy.
- **The VM selection policy should be considered together with VM allocation policy to avoid generating some unnecessary VM migrations**. As for data-intensive jobs in migrant VMs, the distribution of their required data or replicas in candidate target hosts can be taken as a precondition to avoid unnecessary VM migrations. Currently, traditional VM selection decision policies neglect the requirement of data-intensive jobs. It is obvious that the target host will be more reasonable if it can access the required data or replicas as soon as possible. In a word, the frequency of VM migrations will be increased, if VM allocation policy is not considered in VM selection policy.

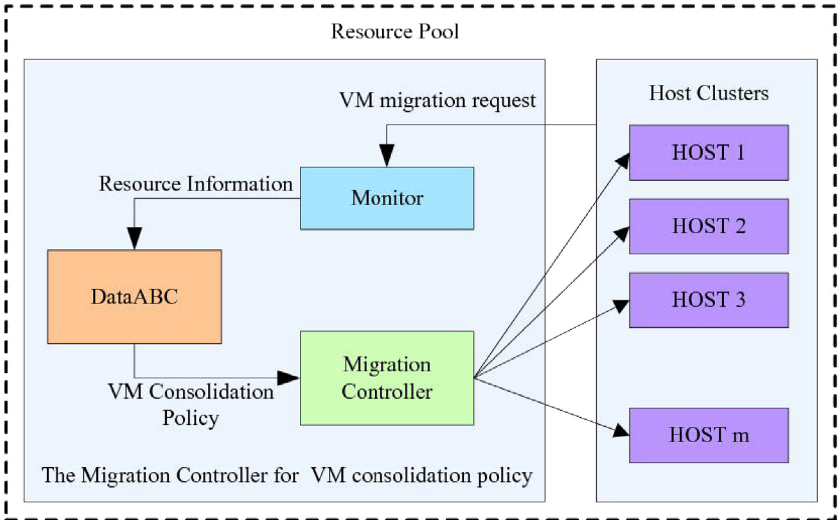


Fig. 2. In this figure of DataABCs architecture, the right part represents the set of available physical hosts, where a huge number of VMs are running. The left part represents the controller of three modulers of Monitor, DataABC and Migration Controller. It implies the proposed architectures processing flow.

4.3. ABC-based energy-efficiency VM allocation policy

VM allocation policy is the decision of selecting m best target hosts from the normal load cluster to receive the m migrant VMs. To reduce the frequency of VM migrations and to save the energy consumption of a data center, a global optimized VM allocation policy should be proposed.

The idea of ABC has the major features that are fast convergence and global optimization. Choosing m best target hosts can be regarded as choosing m best food source in ABC. All combination of selecting m target hosts from N hosts in the normal cluster is the solution space. As for considering characteristics of data-intensive jobs, the candidate target hosts in normal cluster can constitute a two-dimensional solution space which is based on CPU and GPU utilization rates. The ABC based VM allocation policy gets the best global optimized solution by cooperation of bee scouters and followers. In Algorithm 2, it tells us how bee scouters and followers corporate together to get global optimized target hosts. The bee scouters will check some candidate target hosts randomly. Therefore, the bee followers will check their neighbor candidate target hosts to calculate their fitness for fast convergence.

There are two difficult questions needed to be solved in the VM allocation policy. One is how to reduce the frequency of VM migrations, and the other is how to achieve a high energy-efficiency level. It will be a good VM allocation decision that the migrant VM will be allocated to the target host, which has little probability to be migrated out. To decrease the probability of the migrant VM being migrated out, it is a good idea for migrant VMs with data-intensive jobs to be allocated to target hosts with low GPU utilization rate. As illustrated in Fig. 4, the metric of ASLAV is used to measure the average value of SLAV. It shows that the ABC based VM consolidate policy can get a competitive ASLAV value when compared with other traditional polices. Obviously, there is a conflict between saving energy and decreasing the violation rate of SLAV. In VM allocation policy, it is the simplest way of reducing SLAV value, if migrant VMs are migrated to underload target hosts. However, this will increase energy consumption in data centers, because many underload hosts cannot be either slept or hibernated to save energy. As illustrated in Fig. 4, it will be a good ASLAV value when the VM consolidation policy is combined IQR, THR, or MAD based VM allocation with RS or MMT based VM selection. However, these traditional VM consolidation policies would increase energy consumption, which is illustrated in Fig. 3. To solve this conflict, this paper takes behavior characteristics of data-intensive jobs, energy-efficiency evaluation model and SLAV metric into ABC-based VM allocation into consideration. There are two key points as follows:

- The key to reduce SLAV value lies in that the VM allocation decision should be made based on workload balance.
- The key to save energy is making underload hosts to be either slept or hibernated as many as possible and making overload hosts to be normal load hosts with workload level between 60% and 85% to make sure that each VM can be processed fluently.

4.4. The proposed system architecture

In the paper, the system architecture of DataABC is proposed as illustrated in Fig. 2, in which the position of the controller DataABC for VM selection and allocation policies can be seen and its interaction with other entities is clear. Within a Δt time, live VM migration requests will be accumulated by the *Monitor* while the current available amount of computing resource such as CPUs, GPUs, memory, storage and network bandwidth etc. as well as energy consumption is updated. After a Δt time, these information will be transferred to the DataABC *controller*, where the VM selection policy is supposed to be generated by utilizing the proposed DataABC approach to obtain m migrant VMs, and the VM allocation policy is also generated by it to get the m target hosts. Subsequently, the generated migration solution of VM selection and VM allocation will be sent to *Migration Controller*, which is in charge of executing live migration of these VMs. Eventually, these m migrant VMs are migrated into their m target hosts.

4.5. Solution representation

The solution includes VM selection and VM allocation policy. The solution of VM selection is to select a migrant VM queue of BestVMsList from underload cluster and overload cluster within a time window Δt, and the solution of VM allocation is to select a target host queue from clusters of $C_{normalload}$ and $C_{underload}$.

4.6. The main idea of DataABC

Its main idea can be divided into three aspects. Firstly, from the perspective of characteristics of data-intensive jobs in a green cloud data center, a new energy-efficiency evaluation model should be designed by influencing factors of CPU and GPU utilization rate. Secondly, in a workload balanced data center, VM migration is adopted to turn off or hibernate hosts to save energy, which will lead to the increasing of SLAV value and number of VM migrations. However, the QoS of a green cloud data center will be decreased. To solve this problem, VM selection is considered as an important step to save energy and reduce the violation rate of SLA. Lastly, to get an energy-efficiency VM consolidation approach, an ABC-based VM consolidation approach named DataABC is proposed, which proves that it can reduce energy consumption with an acceptable QoS in a green cloud data center.

4.7. The implementation of DataABC

The implementation of DataABC is divided into two steps, which are VM selection and VM allocation. We will introduce DataABC from these two perspectives respectively.

4.7.1. The implementation of VM selection policy

In this section, we describe the specific process of VM selection of DataABC. Details are as follows:

Find the *BestVMsList* from *underload* and *overload* cluster. A five-tuple $VMSelection = \{C_{overload}, C_{underload}, BestVMsList, E, ABCS\}$ for the proposed problem scenario is defined. The ABC-based VM selection policy can be described as Algorithm 1. There are some constrains as follows:

Algorithm 1 ABCS: ABC-based VM selection policy

```
Input: VMs ← Coverload ∪ Cunderload
Output: BestVMsList
CandidateVMsList ← VMs, receive m migration tasks in Δt
for all i such that 1 ≤ i ≤ m do
  repeat
    Send the i scouter based on its getDirection() function, and
    get its predictEnergyCost()
    Send 2 neighbor followers, get their predictEnergyCost()
    Get current optimal migrant VM based on predictEnergy-
    Cost()
    Determine its direction based on getFitness()
  until getFitness() < θ
  Update BestVMsList and CandidateVMsList
end for
Return BestVMsList
```

- Migrant VMs can be selected from all hosts in $C_{overload}$ and $C_{underload}$.
- Only one migrant VM can be selected from a host at time t in Δt.
- The size of BestVMsList is determined by the number of migration requests from a data center at time t in Δt.

4.7.2. The implementation of VM allocation policy

This section mainly introduces the VM allocation policy of DataABC approach. Details are as follows:

Find the m target hosts from the normal load cluster. A six-tuple $VMAllocation(m, t) = \{C_{normalload}, C_{underload}, C_{sleep}, BestVMsList, E, ABCA\}$ for the proposed problem scenario is defined. The ABC based VM allocation policy can be described as Algorithm 2. There are some constrains as follows:

- m migrant VMs should be migrated to m target hosts selected from cluster of $C_{normalload}$ or $C_{underload}$ respectively.
- Each migrant VM is not allowed to be migrated to the host where it comes from.
- The host from $C_{underload}$ will be moved to C_{sleep} if it has been migrated a VM out from time t.
- The host from $C_{underload}$ will continue to receive migrant VMs if it has received a VM from time t.

Algorithm 2 ABCA: ABC-based VM allocation policy

```
Input: BestVMsList, Cunderload, Cnormalload
Output: Csleep
TargetHostsList ← all Hosts, m migrant VMs ← BestVMsList in
Δt
for all i such that 1 ≤ i ≤ m do
  repeat
    Send the i scouter based on its getDirection(), and calculate
    its predictEnergyCost()
    Send 2 neighbor followers, and get their predictEnergy-
    Cost()
    Get current optimal target Host based on their predictEn-
    ergyCost()
    Determine its direction based on its getFitness()
  until getFitness() < θ
  Update BestVMsList, TargetHostsList, Csleep
end for
Return Csleeps
```

5. Performance evaluation

In this section, an array of experiments are designed and conducted to evaluate the performance of the proposed DataABC policy. The new energy-efficiency evaluation model needs to be verified. To verify the importance of VM selection in minimal incremental energy consumption and SLA Violation rate, the proposed policy of VM consolidation is evaluated with different traditional policies and two state-of-art VM consolidation policies such as PS-ABC [15] and PS-ES [7] policies in the CloudSim3.0 platform. It can be manifested from the final experiment results that the DataABC policy proposed can save more energy and has an acceptable execution performance. Notably, DataABC has reduced the frequency of live VM migrations greatly, which reduces the workload of network greatly.

5.1. Experimental scenarios

To evaluate and test the proposed DataABC policy, we have designed and conducted simulation experiment on the CloudSim platform. The simulation experiment is conducted in a single computer. The relevant parameters are shown as follows: Intel(R) Core(TM) i7-3770 CPU @ 3.40 GHz, the OS is Windows 7, RAM is 4.00 GB and its System architecture is 64bit.

As can be seen from Table 1, sample data of 10 days in PlantLab computing environment [17] are selected as experimental data at random. PlantLab Project, started in 2003, consists of many computers all over the world. There are 1160 computers under the control of 547 stations, distributed at 25 nations. The features of the experimental data are large data volume, various data types, low value density and fast processing speed.

As can be seen from Table 2, the simulated data center comprised 800 heterogeneous physical nodes, half of which were HP ProLiant ML110 G4 servers, and the other half consisted of HP ProLiant ML110 G5 servers. The distribution of energy consumption with different workloads of these two servers is depicted in Table 2. The frequencies of the servers CPUs were mapped onto MIPS ratings: 1860 MIPS each core of the HP ProLiant ML110 G4 server and 2660 MIPS each core of the HP ProLiant ML110 G5 server. Each server had 1 GB/s network bandwidth.

With the exception that all VMs were single-core in servers, the number of cores for each VM type divided the amount of RAM: High-CPU Medium Instance (2500MIPS, 0.85 GB), Extra Large Instance (2000MIPS, 3.75 GB), Small Instance (1000 MIPS, 1.7 GB) and Micro Instance (500MIPS, 613 MB).

Table 1
Characteristics of the workload data (CPU utilization rate).

Data	Num of VMs	Mean	St.dev	Quartile 1	Median	Quartile 3
03.03.2011	1052	12.31%	17.09%	2%	6%	15%
06.03.2011	898	11.44%	16.83%	2%	5%	13%
09.03.2011	1061	10.70%	15.57%	2%	4%	13%
22.03.2011	1516	9.26%	12.78%	2%	5%	12%
25.03.2011	1078	10.56%	14.14%	2%	6%	14%
03.04.2011	1463	12.39%	16.55%	2%	6%	17%
09.04.2011	1358	11.12%	15.09%	2%	6%	15%
11.04.2011	1233	11.56%	15.07%	2%	6%	16%
12.04.2011	1054	11.54%	15.15%	2%	6%	16%
20.04.2011	1033	10.43%	15.21%	2%	4%	12%

Table 2
Power consumption by the selected servers at different load levels in Watts.

Server	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
G4	86	89.4	92.6	96	99.5	102	106	108	112	114	117
G5	93.7	97	101	105	110	116	121	125	129	133	135

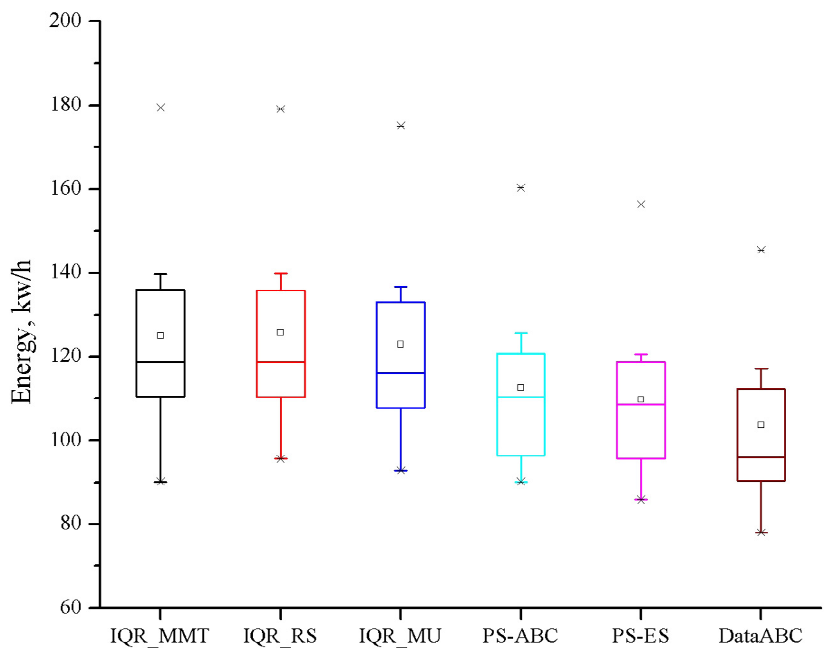


Fig. 3. In the figure of energy consumption, the proposed DataABC is the last one. It shows that it can save energy greatly when compared with other approaches.

5.2. Energy consumption evaluation

As can be seen from Figs. 3–6, each VM consolidation policy consists of a VM selection policy and a VM allocation policy. The proposed DataABC approach is compared with three traditional VM consolidation policies such as IQR_MMT, IQR_RS, IQR_MU and two state-of-art policies of PS-ABC and PS-ES on energy consumption with the proposed energy-efficiency evaluation model in the simulated cloud data center to evaluate its efficiency in the experiment scenario. In this experiment, Δt is 30 s, and safety parameter is 1.5.

In Fig. 3, the proposed DataABC policy is located in the last place. While the other 5 comparable policies are shown in different colors and labeled as IQR_MMT, IQR_RS, IQR_MU, PS-ABC and PS-ES. Noticeably, DataABC can save energy greatly. Based on the average index of *TotalE* in Cloudsim, it can save energy 19.32% to IQR_MMT, 19.25% to IQR_RS, 22.75% to IQR_MU, 9.72% to PS-ABC and 5.84% to PS-ES respectively. The reason is that energy model of these five policies is based on the only influencing factor of CPU utilization rate. However, it is not a good energy model in data-intensive data centers. The proposed DataABC is modeled with CPU utilization rate and GPU utilization rate. Therefore, it is proved that GPU utilization rate should be adopted as an important influencing factor in data-intensive data centers.

5.3. Violation rate of SLA evaluation

SLA is acronym for Service Level Agreement, and SLAV is for Violation rate of SLA. ASLAV is the average of SLAV, and it is adopted to evaluate the performance of service level provided by data centers in Cloudsim3.0. It has a poor performance if its ASLAV value is high. As shown in Fig. 4, DataABC policies are in the final position, and the other five approaches named as IQR_MMT, IQR_RS, IQR_MU, PS-ABC and PS-ES are compared with DataABC. Among them, IQR_MU has the worst performance. Compared with IQR_MMT and IQR_RS, IQR_MU is different with its VM selection approach of MU. MU is acronym for Minimum Utilization policy, which selects the VM as the migrant VM with minimum CPU utilization rate. In Fig. 4, it proves that VM selection policy of MU

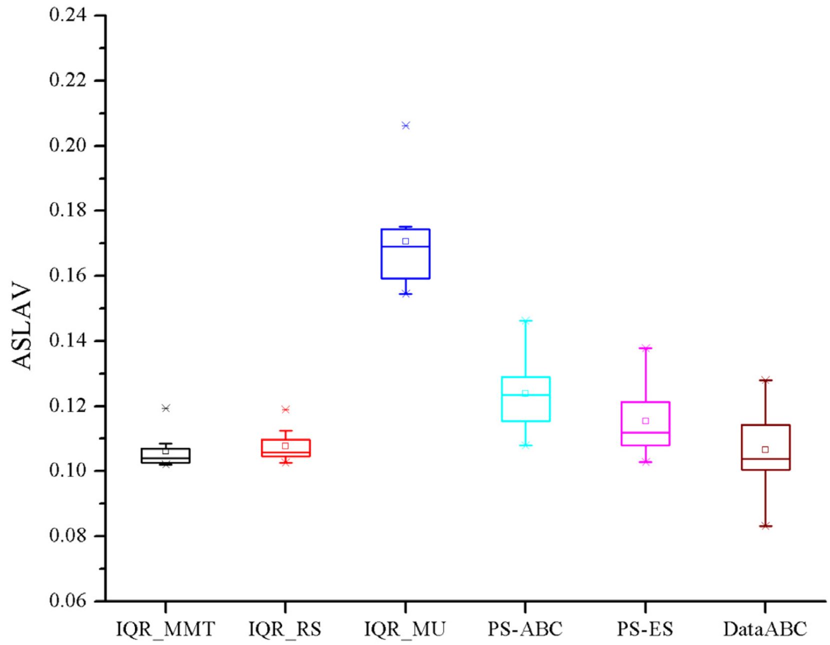


Fig. 4. In the figure of ASLAV metric, the proposed DataABC is the last one. It shows that it has a good ASLAV value when compared with other approaches.

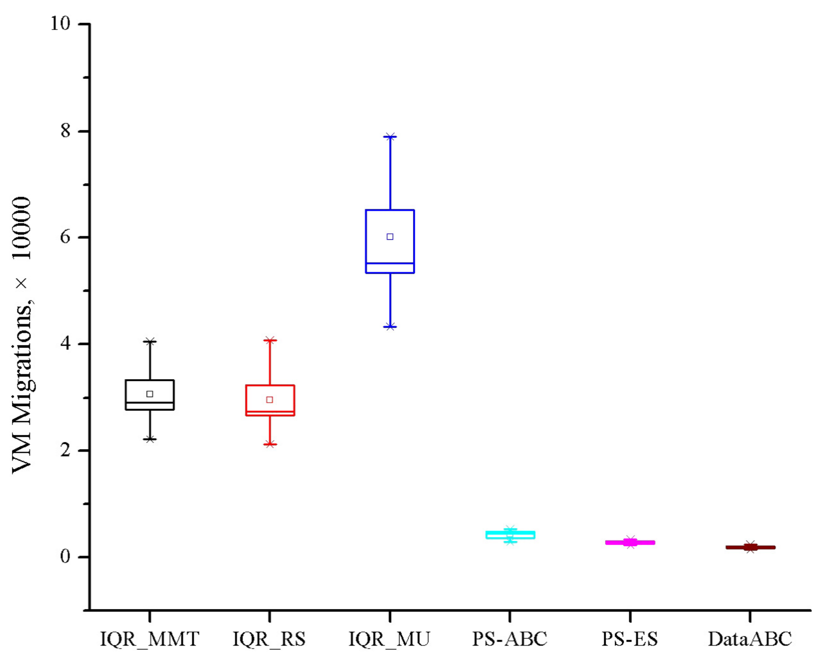


Fig. 5. In the figure of number of VM migrations, the proposed DataABC is the last one. It shows that it can decrease the frequency of VM migrations greatly when compared with three traditional approaches.

makes high ASLAV level. In other words, it also proves that VM selection policy influences ASLAV value in a data center.

In this paper, ABC is the proposed VM selection policy in DataABC. Apparently energy consumption and ASLAV value are negatively correlated. Service level will be highly satisfied if all resources are satisfied. Therefore, VM selection policy should be designed to find a good selection policy to reduce ASLAV. RS is acronym for VM selection policy of Random Selection, and MMT is acronym for VM selection policy of Minimum Migration Time. Observably, the proposed DataABC approach has the similar effect when compared with IQR_MMT, and IQR_RS. Compared with these two state-of-art policies such as PS-ABC and PS-ES, ASLAV of DataABC is 9.1% lower to PS-ABC and 6.2% lower to PS-ES respectively, because its phase of VM selection is redesigned with ABC algorithm. At last, it proves that DataABC can reduce energy consumption with acceptable ASLAV.

5.4. Number of VM migrations evaluation

Observably from the above section, VM selection policy will influence the result of violation rate of SLA directly. VM selection policy of RS and MMT can also lead to a low ASLAV value except for ABC from DataABC. However, it is unreasonable that choosing RS or MMT as the VM selection policy for DataABC. The reason is that number of VM migrations is another metric to evaluate VM selection policy. The incremental energy consumption will be lower if fewer VM migrations occurred.

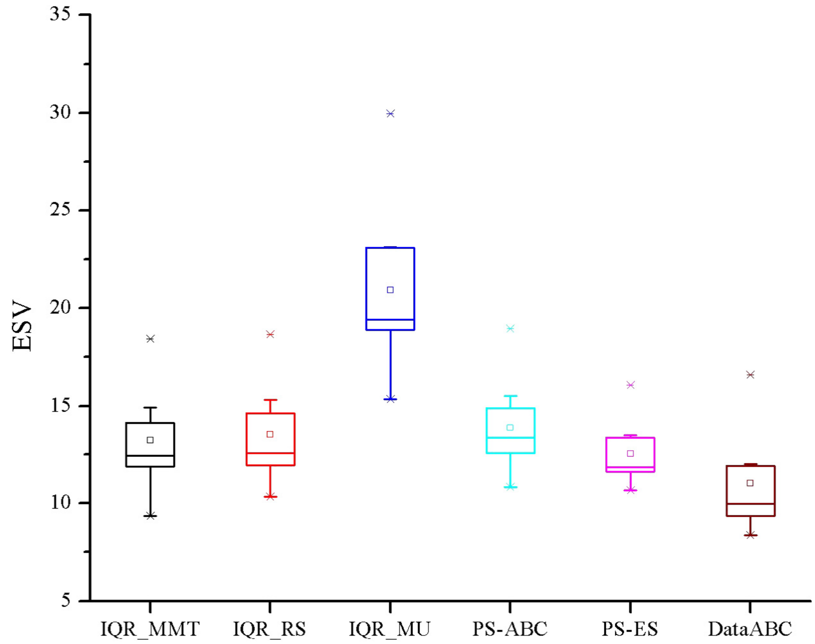


Fig. 6. In the figure of ESV metric, the proposed DataABC is the last one. It shows that it is an energy-efficiency VM consolidation policy when compared with other methods.

As shown in Fig. 5, the DataABC policy is labeled at the last position. From the simulation experiment the number of VM migrations in DataABC policy is decreased exponentially. The number of VM migrations is about 20,000–80,000 times every day by adopting 3 classical policies of IQR_MMT, IQR_RS, and IQR_MU, while about 1000 times every day by our proposed DataABC policy. Compared with IQR_MMT, IQR_RS and IQR_MU, the range of ratio variation of number of VM migrations of DataABC is from 1/80 to 1/20. Even more, compared with two state-of-art policies of PS-ABC and PS-ES, the percentage of number of VM migrations of DataABC are 42.65% and 66.61% respectively.

The purpose of VM Selection is to find best migrant VMs as BestVMsList in Δt, which can reduce energy cost by decreasing the frequency of VM migrations. Explicitly, number of VM migrations is an important statistic to VM selection policy. In data centers, unreasonable VM selection decision will increase the number of VM migrations, improve the error rate of VM allocation, and increase the utilization rate of network bandwidth.

5.5. *Energy-efficiency evaluation*

Based on the above two sections, apparently our proposed approach DataABC can save energy with acceptable ASLAV. In a green data center, ESV is an aggregative statistic to evaluate its energy-efficiency level with energy consumption and SLAV. ESV can be calculated in formula (4).

In a green data center, we know that EC and SLAV are negatively correlated. A good energy-efficiency policy in a data center should get the lowest ESV value when compared with other policies. We have obtained a minimum EC value illustrated in Fig. 3, and obtained a relative low ASLAV value illustrated in Fig. 4. Therefore, it is necessary to evaluate energy-efficiency with the ESV value.

As illustrated in Fig. 6, observably, the proposed DataABC policy has a good performance in the metric of ESV value, while the other five policies have a relatively high value in the metric of ESV value. The primarily reason is that GPU utilization rate should be taken as an important influencing factor in energy-efficiency evaluation model in a green data-intensive data center. The secondary reason is that VM selection is a crucial step to make an energy-efficiency VM consolidation policy. Adopting ABC based VM selection policy can reduce the frequency of VM migrations greatly to decrease the violation rate of SLA. The third reason is that it proves that the proposed ABC based VM consolidation with its characteristics of self-organization, collaboration and rapid convergence is an energy-efficiency policy.

6. Conclusion

In this paper, a novel VM consolidation policy of live VM migration, called DataABC, is proposed to minimize the energy consumption cost and the SLAV value. It adopts the improved ABC based approach to optimize the VM migration decision from VM selection and VM allocation, and redesigns an energy-efficiency evaluation model for data intensive jobs in green data centers. In the improved ABC based approach, we apply a new energy-efficiency evaluation model, VM selection policy, VM allocation policy and the ABC optimized idea to the VM consolidation policy to find an energy-efficiency solution in data-intensive data centers. Firstly, with the observation of the behavior of data-intensive jobs in data centers, we have reconstructed the energy-efficiency evaluation model with two influencing factors of CPU and GPU utilization rate. Henceforth, it is verified that the proposed energy-efficiency evaluation model has its reasonability and effectiveness in CloudSim3.0. Secondly, to save energy and to get an acceptable SLAV value, ABC-based idea is applied to VM consolidation policy. After some simulation experiments in CloudSim3.0, some findings are given as follows:

- A good VM migration decision is made based on both VM selection policy and VM allocation policy.
- The number of VM migrations can be controlled efficiently by ABC based VM consolidation policy.
- The ABC based VM consolidation policy can save 25%–30% energy compared with other classical polices.
- The ABC based VM consolidation policy has the minimal ESV value, which proves that DataABC can get minimized energy consumption based on an acceptable level of SLAV.

In the future, it is necessary to find a global energy-efficiency approach. Some other colony intelligence ideas are necessary to be applied into VM consolidation policy to evaluate their energy efficiency level. Furthermore, the proposed energy-efficiency evaluation model should be verified in computing-intensive jobs.

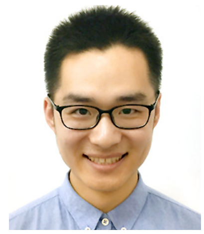
Acknowledgments

The authors would like to thank the editor and the reviewers for their constructive suggestions which have helped us to improve the quality of the paper.

References

[1] M. Armbrust, A. Fox, A view of cloud computing, Commun. ACM 53 (4) (2010) 50–58.
[2] Gartner group, http://www.gartner.com.
[3] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, A. Warfield, Xen and the art of virtualization, Oper. Syst. Rev. 37 (5) (2003) 164–177.
[4] P. Graubner, M. Schmidt, B. Freisleben, Energy-efficient virtual machine consolidation, It Professional 15 (2) (2013) 28–34.
[5] A. Beloglazov, R. Buyya, Energy efficient allocation of virtual machines in cloud data centers, in: M. Parashar, R. Buyya (Eds.), 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, Vol. 1, Institute of Electrical and Electronics Engineers, IEEE Conference Publishing Services, Melbourne, Victoria, Australia, 2010, pp. 577–578.
[6] A. Beloglazov, R. Buyya, Energy efficient resource management in virtualized cloud data centers, in: M. Parashar, R. Buyya (Eds.), 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, Vol. 1, Institute of Electrical and Electronics Engineers, IEEE Conference Publishing Services, Melbourne, Victoria, Australia, 2010, pp. 826–831.
[7] J. Zhao, L. Hu, Y. Ding, G. Xu, M. Hu, A heuristic placement selection of live virtual machine migration for energy-saving in cloud computing environment, PLoS One 9 (9) (2014) e108275–e108275.
[8] A. Beloglazov, R. Buyya, Managing overloaded hosts for dynamic consolidation of virtual machines in cloud data centers under quality of service constraints, IEEE Trans. Parallel Distrib. Syst. 24 (7) (2013) 1366–1379.
[9] J. Broberg, S. Venugopal, R. Buyya, Market-oriented grids and utility computing: The state-of-the-art and future directions, J. Grid Comput. 6 (3) (2008) 255–276.
[10] R.N. Calheiros, R. Ranjan, A. Beloglazov, et al., Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms, Softw.-Pract. Exp. 41 (1) (2011) 23–50.
[11] X. Fan, W. Weber, L. Barroso, Power provisioning for a warehouse-sized computer, ACM Sigarch Comput. Arch. News 35 (2) (2007) 13–23.
[12] R. Nathuji, K. Schwan, Virtualpower: coordinated power management in virtualized enterprise systems, Oper. Syst. Rev. 41 (6) (2007) 265–278.
[13] M. Feia, F. Liu, Z. Liu, Efficiency energy-saving distributed management method of virtual machine in cloud computing, Comput. Eng 38 (11) (2012) 5–7.
[14] X. Liao, L. Hu, H. Jin, Energy optimization schemes in cluster with virtual machines, Cluster Comput. 13 (2) (2010) 113–126.
[15] G. Xu, Y. Ding, J. Zhao, L. Hu, X. Fu, A novel artificial bee colony approach of live virtual machine migration policy using bayes theorem, The scientific world journal 2013 (3) (2013) 215–227.
[16] L. Minas, B. Ellison, Energy Efficiency for Information Technology: How to Reduce Power Consumption in Servers and Data Centers, Intel Press, 2009.
[17] A. Beloglazov, Energy-efficient management of virtual machines in data centers for cloud computing (Ph.D. thesis), The University of Melbourne, 2013.

Jianhua Jiang is an Associate Professor of computer science. His current research interests include energy-efficient computing, cloud computing, supply chain management and internet finance. He has published over 20 papers, and has received 3 best paper awards.



Yunzhao Feng received the bachelor degree in 2009. Currently, he is studying for a master's degree from School of Management Science and Information Engineering, Jilin University of Finance and Economics. His main research interests include energy-efficient computing, cloud computing, supply chain management and internet finance.

Jia Zhao received the master's degree in 2008. He received the Ph.D. degree from the College of Computer Science and Technology, Jilin University in 2013. He began his research on distributed system in 2005. Currently, he is with the College of Computer Science and Engineering, Changchun University of Technology, China. His main research interests include distributed system, cloud computing, network technology. He has participated in several projects.

Keqin Li is a SUNY Distinguished Professor of computer science. His current research interests include parallel computing and high-performance computing, distributed computing, energy-efficient computing and communication, heterogeneous computing systems, cloud computing, big data computing, CPU-GPU hybrid and cooperative computing, multicore computing, storage and file systems, wireless communication networks, sensor networks, peer-to-peer file sharing systems, mobile computing, service computing, Internet of things and cyber-physical systems. He has published over 410 journal articles, book chapters, and refereed conference papers, and has received several best paper awards. He is currently or has served on the editorial boards of *IEEE Transactions on Parallel and Distributed Systems*, *IEEE Transactions on Computers*, *IEEE Transactions on Cloud Computing*, *Journal of Parallel and Distributed Computing*. He is an IEEE Fellow.