International Journal of u- and e- Service, Science and Technology
Vol. 7, No. 6 (2014), pp. 369-380
http://dx.doi.org/10.14257/ijunesst.2014.7.6.32

Towards Energy-aware Web Service Composition

Wei Wang[1], Suichu Zhai[2], and XiaoXia Shi[3,4*]

[1]*School of Digial Information Technology, Zhejiang Technical Institute of Economics Hangzhou, China*
[2] *Hangzhou Power Supply Cooperation, Hangzhou, China*
[3]*Key Laboratory of Complex System Modeling and Simulation, Ministry of Education*
[4] *School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, China*
wangw_me@163.com, sxx19890220@hotmail.com, zhaisuichu@126.com

Abstract

In recent years, the Web service composition technology has developed rapidly, as a result, the energy consumption problem is also increasingly prominent. Faced with this situation, the paper firstly gives out the power model of single service; then derives the model of service composition according to its characteristics, and gives out the corresponding QoS constraints at the same time. After finishing the model, the paper talks about the allocation of feasible path, which is one of the problem in model calculation. The author introduces a graph-plan based searching algorithm, and uses the concept of hierarchical search to solve the problem. At last, the experimental results show that compared with traditional methods, the approach in this paper can not only satisfy the user requests with QoS restricted, but also reduce the energy consumption of service composition effectively.

Keywords: *service composition, energy-aware, QoS constraints, multi-path, composite graph-plan*

1. Introduction

In the recent years, with the popularization of e-commerce, the network based Web service technology has gained rapid development and application, and also the service-oriented computation has become to be an important approach to build the network applications. However, the services released on Internet currently are mostly of single structure and simple function, which make them unable to satisfy the complicated requirements from users. Therefore, the service composition technology arises at the historic moment. The new technology aims at satisfying users` requirements, it tries to achieve the final goal through the composition of different kinds of services.

On the other hand, with the fast development of economy, the performance of computer system and network is increasing, which as a result aggravates the energy consumption of computers, and the situation is becoming worse. In order to solve the problem, the concept of green computing has emerged. This new technology takes low energy consumption of system as the purpose, it provides foundation support platform for the developing information service by building energy efficient computing and running environment. In today`s complex

[*] Corresponding Author: sxx19890220@hotmail.com

ISSN: 2005-4246 IJUNESST
Copyright ⓒ 2014 SERSC

computing environment, the green computing is can use the resource more properly, and can meet the users` various computing requirements in a low energy consumption way.

As the service composition technology has developed rapidly, the most direct result is that the energy consumption produced by services is growing accordingly. This has not only burdened the electricity situation of system, but also influenced the whole performance of service composition. In order to make better use of service composition technology, and at the same time, to increase the utilization rate of resource, the content of green computing needs to be further expanded in the light of the characteristics of service composition. However, the energy consumption research of service composition is very scarce at present, there are still a lot of problems to be solved [1]. On the basis of previous studies, this paper has made some improvements of the service composition, that is, to propose a new Web service composition method, which takes the non-functional attributes QoS as constraint and the energy consumption optimization as purpose. In the current service computing environment, the proposed method can not only reduce the energy consumption of the server, but also improve the efficiency of service composition, thus making it to be better used by us.

Aiming at the shortcomings of current study, the author has made a further discussion of service composition in the perspective of energy awareness, the main contents of the paper are as follows: (1)To search for power model under the framework of service composition, to build the model through the relationship between energy consumption and service parameter; (2)To study the energy-aware service composition method, to generate some feasible paths which can satisfy the QoS constraints based on the power model that already got; (3)To build the relevant simulation and test platform, through which to verify the proposed method.

The paper can be divided into 6 parts, section 2 introduces the research status at home and abroad; Section 3talkl about the model of energy-aware service composition; Section 4 gives the specific algorithm which can solve the searching problem of feasible path in the solution method for model; Section 5 is the experiment of the paper, and Section 6 gives the conclusion and prospect.

2. Related Work

Firstly, the section will introduce Web service composition. In the aspect of classification, on the basis of different division ways, some experts divide it into two models, called business process driven and real-time task solution, others divide it into two methods, i.e. static and dynamic. In the aspect of modeling, the workflow has been introduced, and it has been used to represent and verify the service composition process [2]; a Petri-Network algebraic [3] has been introduced to model Web service control flow. In the aspect of optimization, there are two different methods of Web service optimization, which are local optimization and global optimization [4]. Recently, researchers combines the two methods together [5], taking advantage of both strengths, so that to obtain the best effect of service composition optimization.

As to speak of the energy consumption optimization, the earlier research is mainly engaged in the aspect of hardware. In the non-compute intense activities, the user can reduce the energy consumption by lowering the clock frequency [6]; the researchers also put forward a heterogeneous multi-core architecture of single instruction [7], which can be a mechanism of lowering the energy consumption of processors. In the aspect of system and software, literature [8] tells us the energy-aware design in hard real-time system; when users attempt to minimize the energy consumption in network optimization, they can try the method of studying the routing problem [9]; literature [10] talks about the measurement problem of single Web service energy consumption. In the aspect of modelling and management, the

 Copyright ⓒ 2014 SERSC

DVS(Dynamic Voltage Supply) method has been proposed [11], which is used to lower the usage frequency of servers through the mechanism of adjusting voltage dynamically; an adaptive framework based on service application is also put forward [12], which can reduce energy consumption from the perspective of workload observation.

As to the energy consumption of service composition, literature [13] gives a summary of reconfigurable service composition and classification under energy-aware mobile computing environment, the paper takes the mechanism of integrating different energy-saving methods to achieve the goal of reducing energy consumption.

3. Service Composition Power Model

This section will select multiple time period as the research sample, every time period is called sample interval, denoted by Δt. Suppose during the sample interval $[i\Delta t, (i+1)\Delta t]$, the request arrival rate of server γ_i is fixed, where i means discrete time index. In order to guarantee the non-functional attributes constraint of composite service, here introduces the QoS index, *i.e.* Quality of Service. There are a lot of QoS index, we will choose three of them, which are defined as follows:

(1) Throughput: in a certain time the measurement of service load, denoted by $Q^{(T)}$;

(2) Response Time: including the time of handling user request by services and the waiting time of relevant requests, denoted by $Q^{(R)}$;

(3) Reliability: refers to the rate of successfully executing user requests by services, i.e. success rate, denoted by $Q^{(E)}$.

In view of the energy consumption in service models, the paper will make the following two specifications: (1) environmental/idle energy consumption will be neglected; (2)the energy consumption of network communication will be neglected. The paper will only consider about the energy consumption generated from the sever has received requests.

3.1. Power Model of Single Service

Assuming that the service system has been in a steady state during the interval i. Set the instantaneous power consumption of a single service as p_i, according to literature [14], p_i can be described as:

$$p_i(v) = \mu \times v_i^{\alpha} \, (\alpha > 1) \tag{1}$$

μ 、 α are constants, here we call them the power parameters, v_i means service rate. Particularly, the values of μ 、 α are related to specific service equipment and environment, and they will not be discussed here temporarily. According to the relationship between service load C_i and request arrival rate during the time interval i, the following formula was established:

$$C_i = \gamma_i \times \Delta_t \tag{2}$$

Δt refers to the duration of time interval i. According to the relationship between service load and service rate, the formula can be as follows:

$$v_i = \frac{C_i}{T_{f,i}} \tag{3}$$

Where $T_{f,i}$ refers to effective service time. Combine type (2) with (3.10), we can get the formula:

$$v_i = \frac{C_i}{T_{f,i}} = \frac{\gamma_i \Delta t}{T_{f,i}} \tag{4}$$

According to type (1) and (4), the new expression of instantaneous power consumption p_i can be derived as:

$$p_i(v) = \mu \times (\frac{C_i}{T_{f,i}})^{\alpha} = \mu \times (\frac{\gamma_i \Delta t}{T_{f,i}})^{\alpha}, \ (\alpha > 1) \tag{5}$$

We can see that, there exist certain relationship between instantaneous power consumption p_i and effective service time $T_{f,i}$. In a given period of time, the longer the effective service time, the higher the service efficiency, and the smaller the instantaneous power consumption; while on the contrary, the greater.

According to the relationship between instantaneous power consumption and the total energy consumption, we can obtain the total energy consumption calculation formula of single service in an interval $[t_1, t_2]$ as follows:

$$E_i(t_1, t_2) = p_i \times (t_2 - t_1) = \mu \times (\frac{\gamma_i \Delta t}{T_{f,i}})^{\alpha} \times (t_2 - t_1) \tag{6}$$

Therefore we have got the energy consumption model of single service.

We can see from the above model that apart from self-parameter, the energy consumption of service composition has been influenced mainly by service load and effective service time. Improve the effective service time is one of the methods to reduce energy consumption; while when the service load is fixed, if we want to lower the total energy consumption, the service load must be distributed, that is the multi-path service composition method this paper has adopted, the specific modelling process will be discussed next.

3.2. Energy-aware Service Composition Modeling

Suppose that a request is accomplished by k ($1 \le m \le k$) candidate service sets, and every one of them includes l ($1 \le n \le l$) candidate services which have similar function, we call them the service scale, denoted by (k, l). In consideration of convenient calculation, the concept of sample interval will be ignored in the following modelling process(i.e. leave out the discrete time index i). In a continuous time period, there comes a list of request flow in turn which has a speed of γ. Suppose p_j to be one of a service composition feasible paths which can satisfy users` QoS constraints, $\Phi(p)$ represents the set of feasible paths, here assuming that at least exist one feasible path. Set $|\Phi(p)| = d$, d represents the total number of found feasible paths, apparently we can get: $0 < j \le d$.

According to the optimization goal of the paper, here given the total service composition energy consumption $\min \sum_{m=1}^{k} \sum_{n=1}^{l} P_{mn}$ as the objective function. Every feasible path in set $\Phi(p)$ has a distributed flow, and the sum of all paths` flow can`t exceeds the total request flow. Similarly, the effective service time of every feasible path can be no more than total

 Copyright ⓒ 2014 SERSC

effective service time. From above, the power model of service composition can be drawn as follows:

$$\begin{cases} \min \sum_{m=1}^{k} \sum_{n=1}^{l} P_{mn} \\ subject\ to \quad \sum_{p_j \in \Phi(p)} r_j = \gamma,\ r_j \ge 0,\ 0 < j \le d \\ \qquad \sum_{p_j \in \Phi(p)} z_{mnj} t_{mj} = T_{f,i} \le R_i,\ 1 \le m \le k; 1 \le n \le l \\ \qquad \sum_{n=1}^{l} z_{mnj} = 1,\ z_{mnj} = 0,1;\ 1 \le m \le k;\ p_j \in \Phi(p) \end{cases} \tag{7}$$

r_j represents the request rate distributed in the feasible path p_j, and represents the effective service time of the m th candidate service set in feasible path p_j. z_{mnj} represents a marking variable, where $z_{mnj} = 0$ means p_j did not go through the service node S_{mn}, and on the contrary, $z_{mnj} = 1$ means p_j did go through the service node S_{mn}. The third constraint ensures that for any feasible path p_j,it only and must go through one of the service node in any candidate service set.

After obtaining the power model, next will give the condition formula of user global QoS constraint:

$$\begin{cases} \min \ \Omega(X_1,\ldots,X_m) \\ subject\ to \quad F_T(X_1,\ldots,X_m) \le Q^{(T)} \\ \qquad F_R(X_1,\ldots,X_m) \le Q^{(R)} \\ \qquad F_E(X_1,\ldots,X_m) \le Q^{(E)} \end{cases} \tag{8}$$

$\Omega(X_1,\ldots,X_m)$ represents the total energy consumption function of service composition, $F_T(X_1,\ldots,X_m)$ 、 $F_R(X_1,\ldots,X_m)$ and $F_E(X_1,\ldots,X_m)$ represent the QoS aggregate function of throughput, response time and reliability in a service composition respectively [12], which are used to accomplish the calculation of aggregate QoS in a composite service. According to the relationship between server`s response time R_i and QoS attribute response time $Q^{(R)}$, the following condition was established:

$$R_i \le Q^{(R)} \tag{9}$$

Combining type (9) with the former content, when the relationship between R_i and the service parameter of energy consumption $T_{f,i}$ (i.e. effective service time) is already known, it is easier to obtain the relationship between $T_{f,i}$ and $Q^{(R)}$ indirectly, thus guaranteeing the users` global QoS constraint of web service composition.

3.3. Further Discussion

In model (7), z_{mnj} (marking variable) and r_j (request rate of single feasible path)are the variables which need to be solved. Therefore, the calculation of the model will be divided into two steps, that is the searching of service composition feasible path and the distribution of request rate. Specially, the calculation of the former one is much more complicated, it will be discussed in next section; as to the latter, it can be thought as a nonlinear optimization problem, where the objection function is energy consumption and the independent variable is request rate. The chaos theory can be applied to solve the problem, and then use Matlab to finish the programming calculation.

4. The Search of Service Composition Path

This section will introduce a graph-plan based service composition method to solve the searching problem of feasible path model calculation, which is also called the composite graph-plan method. The algorithm has the ability to limit some operations of problem domain, and can guarantee that at least find one solution of the problem in a linear time complexity. From the perspective of structure, it is a directed acyclic hierarchical graph which includes two kinds of nodes, i.e. service node and parameter node, the two parts appear alternately.

4.1. Main Process

As shown in Figure 1(flow chart), the algorithm can be divided into three steps generally.

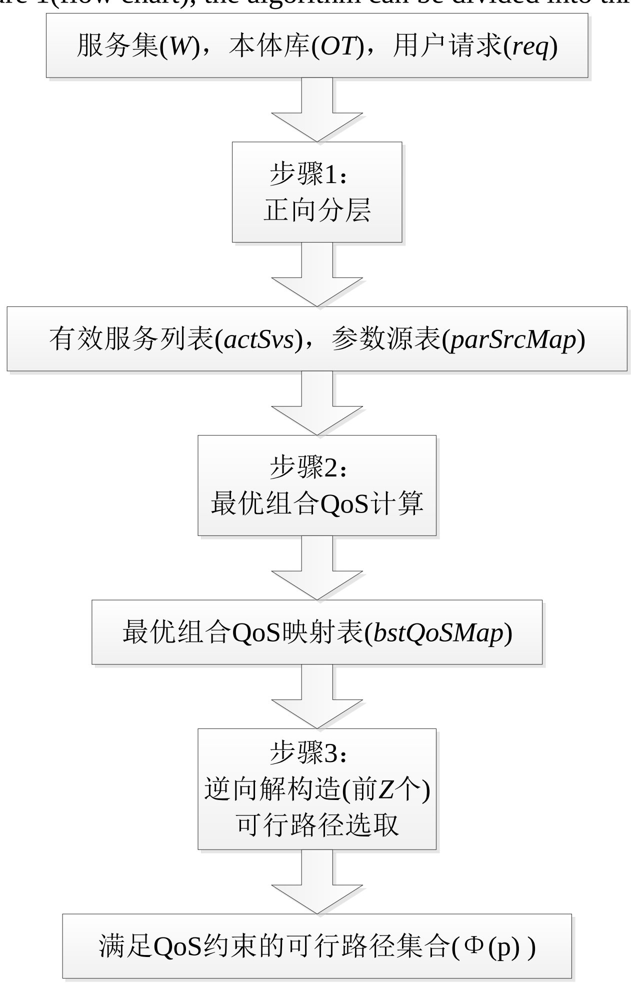

Figure 1. Algorithm Flowchart

Copyright ⓒ 2014 SERSC

The first step of algorithm is hierarchical searching stage, for each service in the to-be investigated service set, the algorithm will decide whether to add it to the graph-plan through filtering. The second step mainly uses the rounds-update calculation method to obtain the composite QoS optimal value of service continuously. The third step firstly makes the solution structure by reverse searching, and by means of QoS threshold value to limit the searching space, then makes the selection of feasible path according to the QoS constraints. Taking the problem of searching space expansion into consideration, here introducing the algorithm parameter *Z*, when the number of found feasible path is more than Z, then the algorithm will only record *Z* paths according to certain filtering rules. At last, the found feasible path set with satisfaction of QoS constraints is the solution we need. Due to space limitations, here only gives the implementation process of the third step of algorithm, as the following algorithm 4.1.

```
Algorithm 4.1 FeasiblePathSelecting
Input：bstQoSMap, parSrcMap, Z
Output：Φ(p)
//search for QoS optimal solution
1. [bstQoSComp, PreSvsMap] = SearchBstQoS(bstQoSMap, parSrcMap, Z);
//determine whether the number of solution is more than Z
2. if(｜bstQoSComp｜≥ Z)    solutions = bstQoSComp;
3. else{
4.      delts = CalDelts(bstQoSMap, parSrcMap, bstQoSComp, PreSvsMap);
5.      solutions = SearchBstZ(bstQoSMap, parSrcMap, delts, Z);
6. }
//search for feasible path with satisfaction of QoS constraints
7. while(solutions != ∅){
8.      AcFlag = false；
//make comparisons of each feasible path in the set solution
9.      for(s∈solutions){
10.     cmpaQoS(s.QoS, rstrQoS, AcFlag);
11.               if(AcFlag)          put(Φ(p) , s);
12.     else  solutions = solutions – {s};
13. }
14. }

15. return Φ(p);
```

4.2. Analysis of Performance

The key stop of this algorithm is the part of selecting feasible path, so here will talk about the time complexity of third step only. In this stage, the most time consuming part is the algorithm *SearchBstZ*. Set the average source group number of each service as *c*, and the total number of service as *n*, we can get the time complexity of searching all solutions is $O(c^n)$, it is an exponential order. However, the algorithm of this paper uses threshold value and algorithm parameter to limit the searching space, which greatly improves the total performance.

5. Simulation experiment

5.1. Experimental Environment and Parameter

In order to verify the energy-saving efficiency of the proposed energy-aware service composition method, this section will give the simulation experiment. The experimental environment is Lenovo Intel(R) Core(TM) i5-2400 CPU、4G RAM、32bit Windows 7 OS. Besides, the experiment also select a QoS based global optimization method (denoted by GO, Global Optimization) [15] as a comparison. At last, the specific parameter configuration of the experiment is shown in Table 1 below.

Table 1. Parameter Configuration of Experiment

parameter names	parameter values
service scale (k, l)	$k \in [20,40]$, $l \in [200,600]$
algorithm parameter Z	$Z \in [10, 30]$
user`s global QoS constraint ($Q^{(R)}$)	randomly generated in $[0.5k, k]$
request arrival rate(γ , reqs/s)	20~100

Before the experiment, firstly we`ll give out two performance index as the specific evaluation criterion, which are defined as follows:

Definition 5-1 (service composition time, *SCT*): the total calculation time algorithm needs to generate a service composition scheme of user`s request, including the generation of service composition and the distribution of request rate, denoted by *SCT*. It is mainly influenced by service scale and algorithm parameter.

Definition 5-2 (energy consumption ratio): the ratio of total energy consumption produced by the method of the paper`s and the compared one. It is influenced mainly by QoS constraint and request arrival rate. When the value of *EPC* is less than 1, it means the paper`s method is more energy efficient than the compared method.

5.2. Experimental Results and analysis

5.2.1. Service composition time *SCT*:

Experiment 1 Relationship between *SCT* and service scale (k, l)

This group of experiments mainly observe the change of *SCT* while selecting different service scale (k, l) , the results are shown in Figure 2 below. We can see from the diagram that when the value of algorithm parameter *Z* is fixed, with the increase of candidate service number *l*, *SCT* is also increasing fast; particularly, when the value of candidate service set k is relatively larger, the increase of *SCT* is more evident.

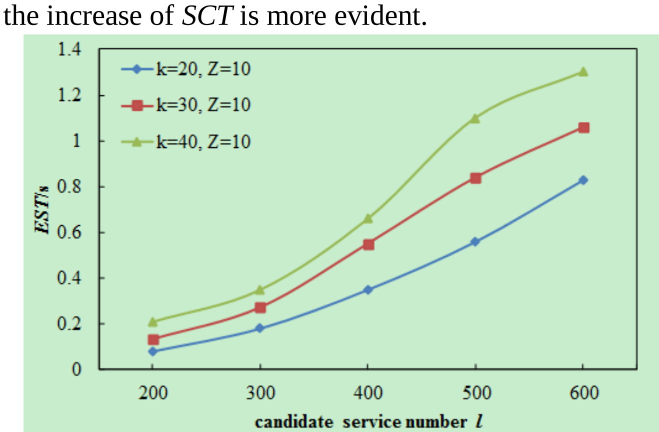

Figure 2. Relationship between Service Composition Time and Service Scale

 Copyright ⓒ 2014 SERSC

Experiment 2 Influence of algorithm parameter *Z* to *SCT*

The experimental results of this group are shown in Figure 3 below. It is observed that the algorithm parameter *Z* has a linear positive correlation with *SCT* generally. Thus we can see that, when Z has a larger value, the performance of algorithm will be improved certainly; however at the same time, the value of *SCT* is also increasing. Therefore, when in the practical application, the value of *Z* should be set reasonably according to the demands; we need to guarantee the control of computation overhead while satisfying the user`s request.

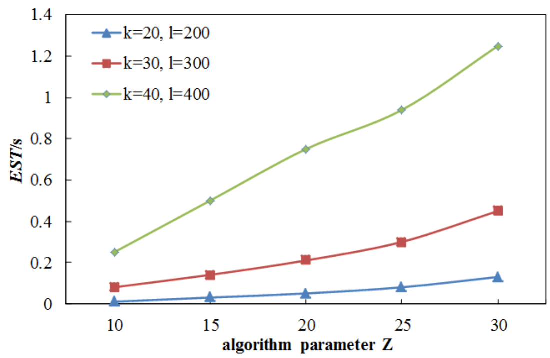

Figure 3. Relationship between Algorithm Parameter and Service Composition Time

5.2.2. Energy Consumption Ratio *EPC*:

Experiment 3 Relationship between *EPC* and request arrival rate γ

This group of experiments mainly observe the influence of request arrival rate to *EPC*, the results are shown in Figure 4 below, where we can see the value of EPC generally varies in the range of 0.3~0.8. The results show that, compared with traditional composition methods, the energy-saving effect of this paper is much more obvious. This is all resulting from the multi-path service composition method applied in the article, which achieves the goal of distributing service load effectively, and thus avoiding the situation of producing too much energy consumption because of service overload.

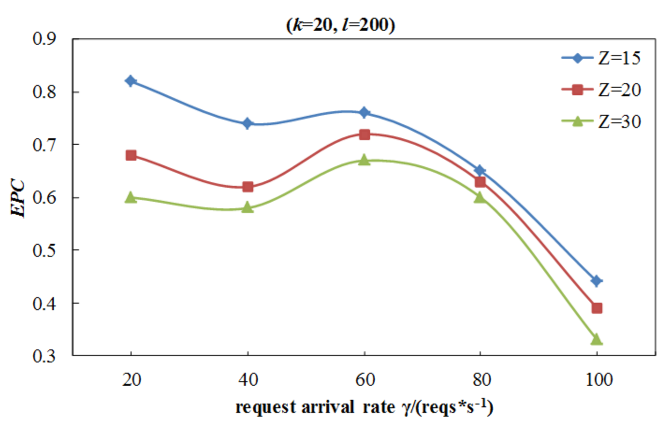

Figure 4. Relationship between Request Arrival Rate and Energy Consumption Ratio

Experiment 4 Influence of QoS constraint to *EPC*

The experimental results of this group are shown in Figure 5 below, and the chosen QoS constraint is response time $Q^{(R)}$. The diagram shows that *EPC* has an obvious negative correlation with response time $Q^{(R)}$. With the increase of $Q^{(R)}$, the value of EPC is decreasing gradually, while we all know that the lager of the $Q^{(R)}$, the lower of the QoS constraint. This shows that the energy-saving effect of the method is more obvious when the QoS constraints are relatively lower, otherwise it will be restrained. So when in practical application, we need to hold the user`s QoS demands in a certain range through ways like negotiation, avoid causing large energy consumption by excessive pursuit of QoS optimization.

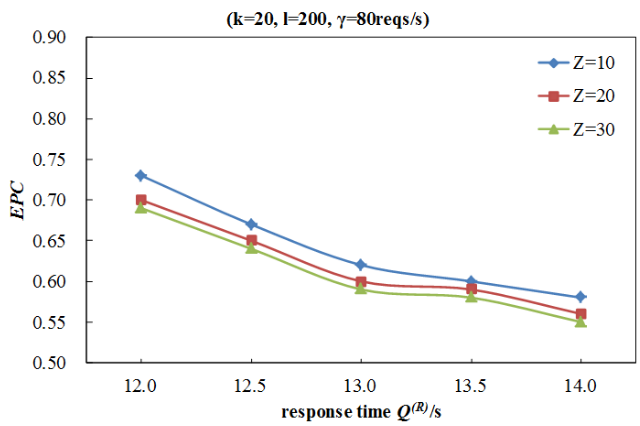

Figure 5. Relationship between User`s Global QoS Constraint and Energy Consumption Ratio

6. Conclusion and Prospect

This paper mainly studied the problem of energy-aware service composition, we have adopted the multi-path service composition method to reduce the energy consumption of service effectively. At the beginning of modelling, we excluded the influence of environment and network communication to the energy consumption. However in practical application, they can't be overlooked usually. So our next work should make a deep research of this problem.

References

[1] D. Perez-Palacin, R. Mirandola and J. Merseguer, "QoS and Energy Management with Petri nets: A Self-adaptive Framework", Journal of Systems and Software, vol. 85, no. 12, **(2012)**, pp. 2796-2811.
[2] W. M. Van Der Aalst, "Verification of workflow nets", Application and Theory of Petri Nets. Springer, **(1997)**, pp. 407-426.
[3] R. Hamadi and B. Benatallah, "A Petri net-based model for web service composition", In Proceedings of the 14th Australasian database conference, vol. 17, **(2003),** pp. 191-200.
[4] D. Ardagna and B. Pernici, "Global and local QoS constraints guarantee in web service selection", IEEE International Conference on Web Services, **(2005)**, pp. 805-806.
[5] M. Alrifai and T. Risse, "Combining global optimization with local selection for efficient QoS-aware service composition", In Proceedings of the 18th International Conference on World Wide Web, **(2009)**, pp. 881-890.

 Copyright ⓒ 2014 SERSC

[6] I. Hewlett-Packard, "Microsoft, Phoenix, and Toshiba, "Advanced configuration and power interface specification", **(2004)**.
[7] R. Kumar, K. Farkas, N. P. Jouppi, R. Parthasarathy and D. M. Tullsen, "Processor power reduction via single-ISA heterogeneous multi-core architectures", Computer Architecture Letters, vol. 2, no. 1, **(2003)**, pp. 2-8.
[8] O. S. Unsal and I. Koren, "System-level power-aware design techniques in real-time systems", Proceedings of the IEEE, vol. 91, no. 7, **(2003)**, pp. 1055-1069.
[9] M. Andrews, A. F. Anta, L. Zhang and Z. Wenbo, "Routing for energy minimization in the speed scaling model", Proceedings of the INFOCOM, IEEE, **(2010)**, pp. 1-9.
[10] P. Bartalos and M. B. Blake, "Engineering energy-aware web services toward dynamically-green computing", Proceedings of the Service-Oriented Computing-ICSOC 2011 Workshops, Springer, **(2012)**.
[11] D. Kusic and N. Kandasamy, "Risk-aware limited look ahead control for dynamic resource provisioning in enterprise computing systems", Cluster Computing, vol. 10, no. 4, **(2007)**, pp. 395-408.
[12] M. C. Jaeger, G. Rojed-Goldmann and G. Muhl, "QoS aggregation for web service composition using workflow patterns", Proceedings of the 8th IEEE International. Enterprise Distributed Object Computing Conference, **(2004)**, pp. 149-159.
[13] E. Park and H. Shin, "Reconfigurable service composition and categorization for power-aware mobile computing", Parallel and Distributed Systems, vol. 19, no. 11, **(2008)**, pp. 1553-1564.
[14] C. Lin, Y. Tian and M. Yao, "Green network and green evaluation: energy-efficient mechanism, model and evaluation", Journal of Computers, vol. 34, no. 4, **(2011)**, pp. 593-612.
[15] L. Zeng, B. Benatallah, A. H. Ngu, M. Dumas, J. Kalagnanam and H. Chang, "QoS-aware middleware for web services composition", Software Engineering, IEEE Transactions, vol. 30, no. 5, **(2004)**, pp. 311-327.

Authors

Wang Wei, lecturer, she work at Zhejiang Technical Institute of Economics the main research field of computer network technology, grid computing.

Suichu Zhai, she received the M.S. degree from School of information engineering Northeast Dianli University, her major is computer application technology. He research interests focus on service computing and cloud computing.

Xiaoxia Shi, she received the M.S. degree from School of Computer on Hangzhou Dianzi University, her major is computer application technology. She studies in the lab of Cloud Computing Technology Research Center, her research interests are service computing and cloud computing, mainly the energy-aware service composition.

Copyright ⓒ 2014 SERSC