

A Cloud-computing-based Resource Allocation Model for University Resource Optimization

Cong Liu

Information Technology Center, Tsinghua, Beijing, China
liuc@tsinghua.edu.cn

Abstract

The resource allocation is the basic guarantee for the effective implementation of the university development strategy, while resource allocation mechanism is an effective control measure to strengthen the university management. So the importance of resource allocation has caused widespread attention and concern of university administrators. Combining with the actual situations and encountered problems of university resource management, we introduced cloud computing ideas and proposed a cloud-computing-based resource allocation model to conduct a preliminary exploration and research on this issue. We first formulated the resource allocation problem, and then proposed the resource allocation concept model and its corresponding application model in universities. Based on the proposed model, we further expanded detailed analyses and descriptions on resource definition and description, resource allocation rules and methods, resource management mode, and a general mathematical model of resource allocation optimization in university. Finally, we gave a cloud-computing-based university resource allocation framework to offer references and consult for further applications of technology based on cloud computing in the university resource construction, aiming to achieve optimal university resource allocation and improve teaching quality and efficiency.

Keywords: *cloud computing, resource allocation concept model, resource allocation application model, resource classification tree, resource optimization*

1. Introduction

With the continuous development of information technology, colleges and universities information technology has become an important part of the university's business, university leadership has been more attention with the teachers and students. Resources are the basis for development and the essential guarantee for higher education quality. However, the actual rate of growth in resources development of universities is far from meeting the demand for resources needed. With the expansion of colleges and universities, will inevitably lead to the relative per capita possession of resources reduced. Obviously, the resource management technology in university is also facing the great challenges and opportunities with the transition from the traditional management techniques to the information resource modernization and sharing [1]. Furthermore, in recent years, the emerging cloud computing technology has brought new opportunities and provided a new way of thinking for the development of the resource management technology in university.

Cloud computing as a new product with the times because of its powerful product features in some areas has been the basic application of [2]. Cloud computing stores the data, applications and services in the cloud, takes full advantage of the powerful computation

capabilities of data center, achieves self-adaptability of user business systems. Based on this idea, the cloud-centric innovative applications continue to produce. We should design following the concept, i.e. Application as a Service (AaaS), and provide all the resources and functions as a service to the cloud clients.

This paper analyses the characteristics and usage of university resources, in accordance with the “Fair, Open and Justice” principle, elaborates the resource allocation model and its application based on cloud computing. The rest of this paper is organized as follows. In Section 2, we propose the resource allocation model including concept model and application model that we research in this paper, and some definitions and descriptions about the resources and allocation procedures are given. Section 3 describes the application of cloud-computing-based resource allocation model in detail. Finally, we conclude our paper in Section 4.

2. Resource Allocation Model

2.1. Problem Formulation

First, we briefly analyze the resource allocation procedure, as shown in Figure 1. In fact, the resource allocation is the result set of mappings from users to resources in accordance with certain rules. Reasonable resource configuration to improve resource utilization is to optimize resource allocation, which to obtain an optimal result set of mapping from user to resource according to certain rules.

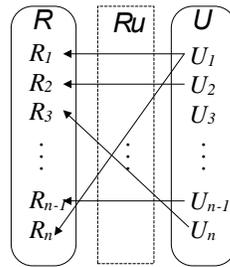


Figure 1. User-Resource Mappings

The same resource defined in the usual sense is different from each other because of its spatial and time such factors. In this way, a vector including spatial property, time property and other basic properties of the resource can truly identify a valid resource, i.e., minimal resource unit. The resource definition is described in Section 2.4.

2.2. Resource Allocation Concept Model

The resource allocation concept model shown in Figure 2 gives a uniform solution to define, explain, supervise and manage throughout the life cycle of the resource allocation and usage in universities.

From the perspective of above model, Resource Manager is a function entity responsible for the control and allocation of users and resources. Rule Service is a function entity to provide corresponding resource rules from the Rule Database which is used to store reservation rules of various types of resources. Authorized Authentication Center is a function entity to issue the identity certificate. Users, i.e., organization or individual user, can reserve resources only through authorized authentication.

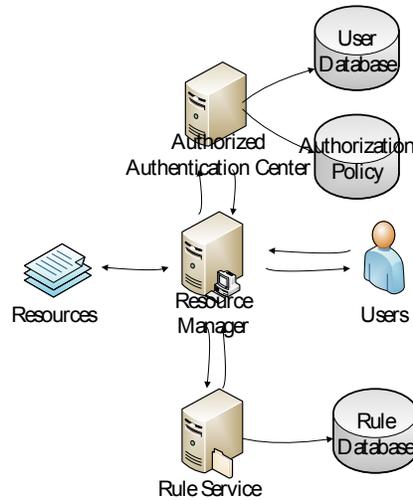


Figure 2. The Resource Allocation Concept Model

2.3. Resource Allocation Application Model

The resource allocation application model is shown in Figure 2. The model can be described by a three-tuple as $M = (U, R, Ru)$, where U is the user set, i.e. $U = \{u_1, u_2, \dots, u_{|U|}\}$, R is the resource set, i.e. $R = \{r_1, r_2, \dots, r_{|R|}\}$, Ru is the constraint set of resource usages defined for a number of different types of resources, $|R|$ is the total number of resources to be allocated to $|U|$ users.

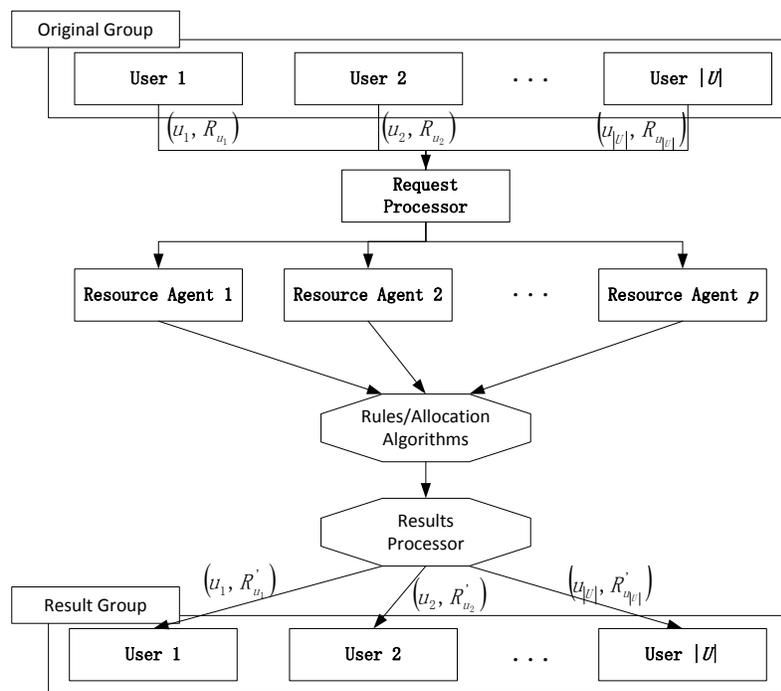


Figure 3. The Resource Allocation Application Model in Universities

The resource allocation application model in universities is mainly composed of five parts according to the function flow. To start a round of resource allocation, first, each user in original group submits his resource request to the Request Processor; second, the Request Processor put the users' requests forward to the appropriate Resource Agent according to the type of resources users attend to apply for; third, the Resource Agent is responsible for resource allocation and integration of the results in accordance with these resources' rules and algorithms; finally, the allocation results are returned to the users in Result Group and the current round of resource allocation ends. The main functions of each part are as follows:

Request Processor: The request processor is responsible for parsing the user requests. It will distribute these requests to the specified resource agent according to certain rules, such as resource type, load balancing and so on. Here, the set of resource requests for each user can be described as $\{(u_1, R_{u_1}), (u_2, R_{u_2}), \dots, (u_{|U|}, R_{u_{|U|}})\}$, where R_{u_i} represents resource set user u_i applies for, $i = 1, 2, \dots, |U|$.

Resource Agent: The resource agent is responsible for the management of publishing and allocating resources. Furthermore, the agent constraints the user's applying behavior according to the rules for each type of resources. The users who cannot be satisfied with requirements for resource usage will not be included in the resource allocation.

Original group: The original group is the set of all the users applying for resources. Users according to their requirements apply for all kinds of resources which are needed.

Rules/Allocation Algorithms: This function module stores usage rules and allocation algorithms for each type of resources.

Result Group: After resource allocation ends, the set of resource allocation for each user can be described as $\{(u_1, R'_{u_1}), (u_2, R'_{u_2}), \dots, (u_{|U|}, R'_{u_{|U|}})\}$, where R'_{u_i} represents resource set assigned to user u_i , $i = 1, 2, \dots, |U|$.

2.4. Resource Description

This section explicates resource description in detail.

Definition 1. Resource unit refers to the minimum allocation item of resources, which is indivisible. For convenience, we call a resource unit as a resource in this paper.

Let an m -dimensional vector express a resource, i.e., $\forall r_i \in R$, $r_i = (r_{i1}, r_{i2}, \dots, r_{im})$, where m represents the vector dimension to describe various attributes of the resource, such as space, time, category, relations with other resources and so on.

Definition 2. Resource classification tree is a finite set of n nodes. In the set, we define a unique relationship, which satisfies the following conditions:

- Root node is an only node with no parent node in the set.
- Except root node, each node has one and only a parent node.
- Except root node, any node n has an only node sequence from the root to this node, i.e., $n_0, n_1, n_2, \dots, n_k$, where n_0 is the root, n_k is the node n and n_i is the children node of the node n_{i-1} . This node sequence is called a path from the root to n and k is the path length.

- Each node represents a specific classification in certain kind of resource categories in university.
- Leaf node is the node with no children nodes, which represents the minimum classified catalogue in the classification tree.

Definition 3. The level of the resource classification tree is the maximum path length from the root to all nodes, *i.e.*, the root node is at level 0, and the nodes whose parent node is root are at level 1, and so on.

The first level of the tree represents first categories of the university resources, the second level of the tree represents second categories of the university resources, and so on, and the n^{th} level of the tree represents n^{th} categories of the university resources. Each node represents some sort of a resource, whose parent node represents the upper classification of the current node and whose children nodes represent the sub-classification of the current node. We encode for each node in accordance with the following rules to generate:

- According to the experiences and statistical data, we determine the maximum classified number of the nodes with the same parent node under each level of the tree and define several digits used to describe them.
- At a certain level in the tree, the nodes with the same parent node are orderly numbered from left to right.

Then, the category code of any node in the tree can be composed by the description digit determined by the tree level and codes of all the nodes in the path. For example, the category codes of node A and node B are 10201201201 and 10299299299 respectively, as shown in Figure 4.

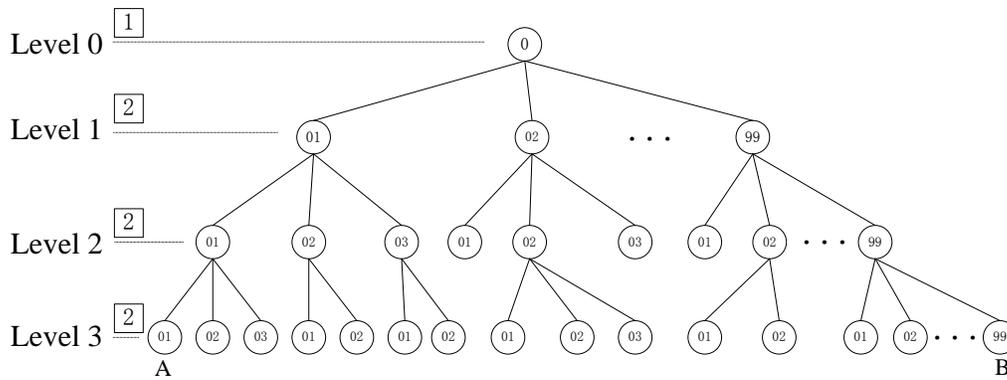


Figure 4. An Example of Resource Classification Tree and its Encoding

According to the definition of the resource classification tree, we have established a multi-layer resource classification tree to form a university resource classification system, facilitating unified maintenance and management of the university resource classification. In the resource classification tree architecture, the university resources will be divided into several categories, each category is divided into many subcategories, and each subcategory can be divided into much smaller subcategory, and so on. Thus, each category has a class number and each kind of resources can be assigned to a certain category. Assume that category code of any resource must be described by the leaf node of the resource classification tree, through which we can trace and analyze the affiliations and various categories of the resource classification at all levels.

2.5. Resource Management and Allocation

The ownership of the university resources belongs to the university, but university administrations or units are authorized to manage these resources. Therefore, we build a resource management tree based on university organization tree to analyze the resource management mode.

According to the management and usage requirements of the university resources, the resource allocation rules can be divided into two categories:

- No user participation. Administrative departments or units directly allocate their responsible resources in accordance with the corresponding resource rules.
- User participation. Users participate in the resource allocation procedure, such as student's selecting courses, activity ground reservations and so on. Administrative departments or units launch resource allocation based on the users' requirements under certain rules.

No matter what kind of allocation rules, we need to consider user permissions, usage requirements, restrictions and other factors for the use of resources.

Due to various allocation methods, there are various allocation policies in the rule constraints, such as first come first served basis, average allocation, random allocation, resource reservation, *etc.*, According to the mathematical modeling, we achieve fair, reasonable and multi-win resource allocation using kinds of algorithms, such as heuristic algorithm, bionic algorithm, intelligent computing and so on.

2.6. The Mathematical Model of Resource Allocation Optimization

The resource allocation issue can be considered as a multi-constraint and multi-objective programming problem, aiming to optimize the user satisfaction, resource utilization and costs, and ensuring to optimize resource allocation for each category at the same time to make all resources allocated to achieve optimization overall. Therefore optimization of resource allocation strategy can be an interesting and challenging task. A classical optimal resource allocation problem can be presented as:

$$\min \text{ or } \max f(x_1, x_2, \dots, x_{|U|}) \quad (1)$$

$$\text{s.t.} \quad \sum_{i=1}^{|U|} x_i \leq |R| \quad (2)$$

$$\sum_{i=1}^{|U|} x_{ij} \leq |R_j| \quad (3)$$

$$\sum_{k=1}^{|C|} x_{ik} = x_i \quad (4)$$

$$|R| \geq 0, |R_j| \geq 0, R_j \subseteq R, j = 1, 2, \dots, |C| \quad (5)$$

where objective value function $f(x_1, x_2, \dots, x_{|U|})$ becomes minimized (*e.g.*, cost) or maximized (*e.g.*, profit) through the comprehensive consideration of many factors. The variable x_{ij} represents the amount of resources with category j allocated to user i and the

variable x_i represents the amount of all resources allocated to user i . $|R_j|$ represents the amount of resources with category j and $|C|$ represents the number of resource categories. In the case of cloud computing, the resources are elastic, *i.e.*, $|R_j|$ and $|R|$ cannot be regarded as a static parameter. Therefore, it is an important research task to investigate optimal resource allocation algorithms, as presented in [3], but not assuming $|R_j|$ or $|R|$ as a fixed value. Thus online monitoring of $|R_j|$ or $|R|$ is necessary to decide the best time for allocation or reallocation towards optimization.

3. The Application of Cloud-Computing-Based Resource Allocation Model

3.1. Cloud Computing Overview

Cloud computing is a development and convergence product of grid computing, distributed computing, virtualization and other computer technology and network technology [4, 5]. A core concept of cloud computing is that by continuously improving the processing capability of cloud, thereby reducing the processing burden on the user terminal, the user terminals eventually reduced to a simple input and output devices, and can enjoy on-demand the powerful computing processing capability of cloud. Cloud computing is a way of computing mode to provide dynamic scalable virtual resources using Internet as a service, which can be expressed as four kinds of deployment models, three kinds of service models and seven characteristics, as shown in Figure 5.

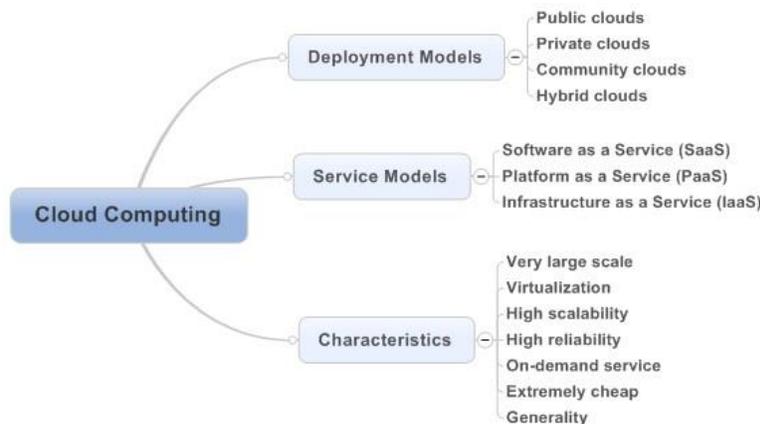


Figure 5. Cloud Computing Manifestations

3.2. The Architecture of Cloud-Computing-Based Resource Allocation Model

The resource allocation model belongs to SaaS research category, and its cloud computing based architecture is shown in Figure 6. The architecture is based on the popular B/S service mode, including service terminal devices and cloud computing platform. Users through such as the mobile terminal, personal computer or temporary offline on-demand access various resource services provided by cloud computing platform. The cloud computing platform is composed of physical layer, virtual layer and application layer.

Physical layer provides computing, network and storage resources to the upper layer in the form of resource pools. Virtual layer through virtualization management platform fully integrates and efficiently utilizes computing, network and storage resources. According to the function, application layer is divided into management, service and data modules, which orients different service objects and provides different application services. Management module provides authentication, billing, security and other management functions of the cloud computing platform. Service module provides management-related specific businesses in the university resource allocation, such as resource information inquiry, resource usage applying, resource allocation results inquiry, resource management, etc., providing service on demand based on the users' needs. Data module is the storage of all the entities in the data model and is part of the basis for management and services. User layer, including teachers and students, staffs in the university as well as possible authorized social users, supports users in different roles through mobile terminals, personal computers and a variety of ways to access or manage the data in the application layer.

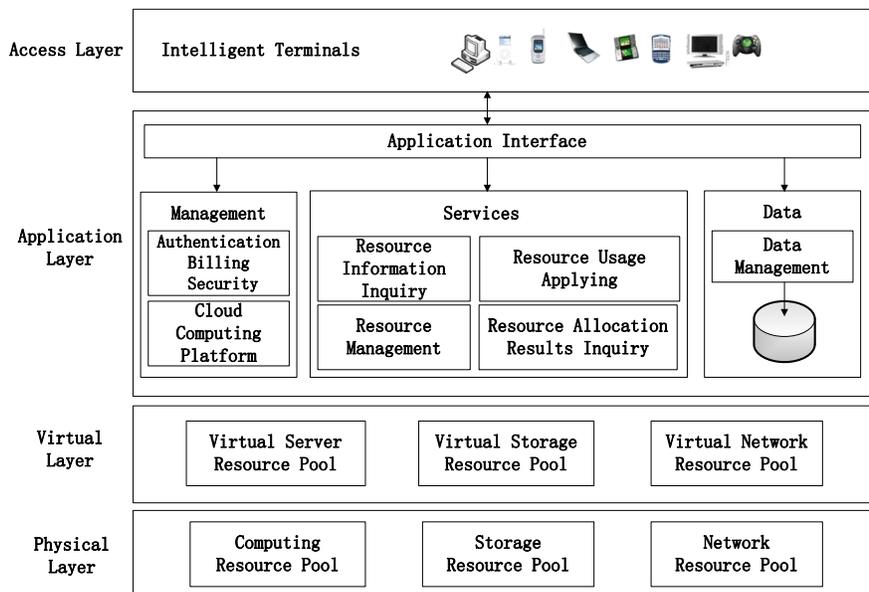


Figure 6. The Architecture of Cloud-Computing-Based Resource Allocation Model

4. Conclusions

Cloud computing is a research hotspot in IT fields, which utilizes virtualization technology, integrates distributed computing, storage, data, applications and other resources, provides collaborative computing model and strong information processing capabilities. A new idea which is applying cloud computing to the resource management in universities is presented in this paper, for solving shortcomings in the current inefficiency resource management, lack of resource information timely communication and resource sharing. At present this model still exist some imperfect places, pending further research, improvement and practice.

Acknowledgements

The author would like to thank Ruiwei Meng for excellent technical support and critically reviewing the manuscript.

References

- [1] J. Dongxing, S. Zongkai, C. Huaichu, L. Qixin, S. Peihua and W. Yingxue, "University resource planning system", Chinese J Tsinghua Univ (Sci & Tech), vol. 4, no. 44, (2004).
- [2] C. Quan and D. Qianni, "Cloud computing and its key techniques", Journal of Computer Applications, vol. 9, no. 29, (2009).
- [3] T. Ibaraki and N. Katoh, "Editors, Resource allocation problems: algorithmic approaches, Foundation of Computing series", The MIT Press, Cambridge, Massachusetts, (1998).
- [4] R. Buyyaa, Y. Cheeshin, V. Srikumar, B. James and B. Ivona, "Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility", Future Generation Computer Systems, vol. 25, (2009).
- [5] L.S. Vesna, H.D. Vesna and B. Zoran, "Editors. Cloud computing-issues, research and implementations", Proceedings of the ITI 2008 30th International Conference on Information Technology Interfaces, (2008) June 23-26, Cavtat, Croatia.

Authors



Cong Liu, is an engineer in Information Technology Center at Tsinghua University. Her research interests are major in university education in formationization including electronic school affairs system and online learning. She holds an MS in computer software and theory from Northeastern University.

