



IGNITED MINDS
Journals

*International Journal of
Information Technology
and Management*

*Vol. VI, Issue No. I,
February-2014, ISSN 2249-
4510*

**EFFECTIVE LOAD BALANCING ALGORITHM
USING ANT COLONY OPTIMIZATION
TECHNIQUE**

AN
INTERNATIONALLY
INDEXED PEER
REVIEWED &
REFEREED JOURNAL

Effective Load Balancing Algorithm Using Ant Colony Optimization Technique

Sai Prasanna Kodam¹, Divya Anusha Gandla², Anusha Janapatla³, Chandu Naik Azmera⁴

¹Student, Department of Computer Science and Engineering, Varadha Reddy College of Engineering, Warangal, India

²Student, Department of Computer Science and Engineering, Varadha Reddy College of Engineering, Warangal, India

³Student, Department of Computer Science and Engineering, Varadha Reddy College of Engineering, Warangal, India

⁴Assistant Professor, Department of Computer Science and Engineering, Varadha Reddy College of Engineering, Warangal, India

Abstract – Load balancing in the cloud computing environment has an important impact on the performance. Good load balancing makes cloud computing more efficient and improves user satisfaction. This article introduces a better load balance model for the public cloud based on the cloud partitioning concept with a switch mechanism to choose different strategies for different situations. The algorithm applies the game theory to the load balancing strategy to improve the efficiency in the public cloud environment. Cloud computing is efficient and scalable but maintaining the stability of processing so many jobs in the cloud computing environment is a very complex problem with load balancing receiving much attention for researchers. Since the job arrival pattern is not predictable and the capacities of each node in the cloud differ, for load balancing problem, workload control is crucial to improve system performance and maintain stability. Load balancing schemes depending on whether the system dynamics are important can be either static or dynamic. Static schemes do not use the system information and are less complex while dynamic schemes will bring additional costs for the system but can change as the system status changes. A dynamic scheme is used here for its flexibility.

Key Terms: Load Balancing, Dynamic Scheme, Cloud

I. INTRODUCTION

Cloud computing is an attracting technology in the field of computer science. In Gartner's report, it says that the cloud will bring changes to the IT industry. The cloud is changing our life by providing users with new types of services. Users get service from a cloud without paying attention to the details. NIST gave a definition of cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. More and more people pay attention to cloud computing. Cloud computing is efficient and scalable but maintaining the stability of processing so many jobs in the cloud computing environment is a very complex problem with load balancing receiving much attention for researchers. Since the job arrival pattern is not predictable and the capacities of each node in the cloud differ, for load balancing problem,

workload control is crucial to improve system performance and maintain stability. Load balancing schemes depending on whether the system dynamics are important can be either static or dynamic. Static schemes do not use the system information and are less complex while dynamic schemes will bring additional costs for the system but can change as the system status changes. A dynamic scheme is used here for its flexibility. The model has a main controller and balancers to gather and analyze the information. Thus, the dynamic control has little influence on the other working nodes. The system status then provides a basis for choosing the right load balancing strategy. The load balancing model given in this article is aimed at the public cloud which has numerous nodes with distributed computing resources in many different geographic locations. Thus, this model divides the public cloud into several cloud partitions. When the environment is very large and complex, these divisions simplify the load balancing. The cloud has a main controller that chooses the suitable partitions for arriving jobs while

the balancer for each cloud partition chooses the best load balancing strategy.

II. RELATED WORK

Cloud computing is efficient and scalable but maintaining the stability of processing so many jobs in the cloud computing environment is a very complex problem with load balancing receiving much attention for researchers. Since the job arrival pattern is not predictable and the capacities of each node in the cloud differ, for load balancing problem, workload control is crucial to improve system performance and maintain stability. Load balancing schemes depending on whether the system dynamics are important can be either static and dynamic. Static schemes do not use the system information and are less complex while dynamic schemes will bring additional costs for the system but can change as the system status changes. A dynamic scheme is used here for its flexibility. Load balancing schemes depending on whether the system dynamics are important can be either static and dynamic. Static schemes do not use the system information and are less complex while dynamic schemes will bring additional costs for the system but can change as the system status changes. A dynamic scheme is used here for its flexibility. The model has a main controller and balancers to gather and analyze the information. Thus, the dynamic control has little influence on the other working nodes. The system status then provides a basis for choosing the right load balancing strategy.

The load balancing model given in this article is aimed at the public cloud which has numerous nodes with distributed computing resources in many different geographic locations. Thus, this model divides the public cloud into several cloud partitions. When the environment is very large and complex, these divisions simplify the load balancing. The cloud has a main controller that chooses the suitable partitions for arriving jobs while the balancer for each cloud partition chooses the best load balancing strategy. Load balancing schemes depending on whether the system dynamics are important can be either static and dynamic. Static scheme does use the system information and are less complex.

III. EXPERIMENTAL RESULTS

A flow of events is a sequence of transactions (or events) performed by the system. They typically contain very detailed information, written in terms of what the system should do, not how the system accomplishes the task. Flow of events are created as separate files or documents in your favorite text editor and then attached or linked to a use case using the Files tab of a model element.

A flow of events should include:

- When and how the use case starts and ends

- Use case/actor interactions
- Data needed by the use case
- Normal sequence of events for the use case
- Alternate or exceptional flows



Construction of Use case diagrams:

Use-case diagrams graphically depict system behavior (use cases). These diagrams present a high level view of how the system is used as viewed from an outsider's (actor's) perspective. A use-case diagram may depict all or some of the use cases of a system.

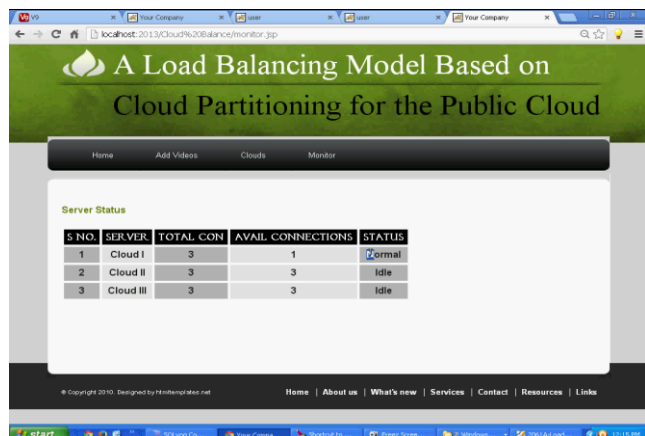
A use-case diagram can contain:

- actors ("things" outside the system)
- use cases (system boundaries identifying what the system should do)
- Interactions or relationships between actors and use cases in the system including the associations, dependencies, and generalizations.

Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus

the objective of input design is to create an input layout that is easy to follow



[6] N. G. Shivaratri, P. Krueger, and M. Singhal, Load distributing for locally distributed systems, Computer, vol. 25, no. 12, pp. 33-44, Dec. 1992.

[7] B. Adler, Load balancing in the cloud: Tools, tips and techniques, <http://www.rightscale.com/info-center/white-papers/Load-Balancing-in-the-Cloud.pdf>, 2012

IV. CONCLUSION

This paper proposed the basic concepts of Cloud Computing and Load balancing. In addition to that, the load balancing technique that is based on Swarm intelligence has been discussed. We have discussed how the mobile agents can balance the load of a cloud using the concept of Ant colony Optimization. The limitation of this technique is that it will be more efficient if we form cluster in our cloud. So, the research work can be proceeded to implement the total solution of load balancing in a complete cloud environment. Our objective for this paper is to develop an effective load balancing algorithm using Ant colony optimization technique to maximize or minimize different performance parameters like CPU load, Memory capacity, Delay or network load for the clouds of different sizes.

V. REFERENCES

[1] R. Hunter, The why of cloud, <http://www.gartner.com/DisplayDocument?doc cd=226469&ref=g noreg>, 2012.

[2] M. D. Dikaiakos, D. Katsaros, P. Mehra, G. Pallis, and A. Vakali, Cloud computing: Distributed internet computing for IT and scientific research, Internet Computing, vol.13, no.5, pp.10-13, Sept.-Oct. 2009.

[3] P. Mell and T. Grance, The NIST definition of cloud computing, <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>, 2012.

[4] Microsoft Academic Research, Cloud computing, <http://libra.msra.cn/Keyword/6051/cloud-computing?query=cloud%20computing>, 2012.

[5] Google Trends, Cloud computing, <http://www.google.com/trends/explore#q=cloud%20computing>, 2012.