TELKOMNIKA, Vol. 11, No. 2, February 2013, pp. 954~959
ISSN: 2302-4046

One Method of Cloud Computing Bandwidth Allocation Based on Fairness

Yiquan Kong
Zhanjiang Normal University Guangdong Zhanjiang 524048 China
Corresponding author, e-mail: freemankyq@163.com

Abstract

In order to solve the bandwidth allocation unfairness problem in the cloud computing network, one method uses fairness congestion control algorithm, access control list (ACL) and traffic policing and traffic shaping in the paper. The method can rationally solve the problem after analyzing the reason of cloud computing bandwidth allocation unfairness. For illustration, one network video conference example was utilized to show the method in solving bandwidth allocation unfairness problem.The experimental *results show network bandwidths are fairly allocated, packet loss ratio and latency is obvious improvement. The method deals with non-adaptive UDP and TCP adaptive flow congestion and provides the end-to-end quality of service over the differentiated services networks, and the bandwidth allocation problem based on fairness in the cloud computing network is solved well.*

Keywords: *cloud computing; bandwidth allocation; fairness*

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

With the development of Internet and multimedia services, the architecture of cloud computing has been more widely used. Cloud computing which has features such as simple and efficient, urgent expansion, and it becomes one core technology for the Internet backbone network. Service differentiation and fairness guarantees are the targets of the cloud computing network to realize. But cloud computing network has the problem of fairness in the bandwidth allocation. It includes the problem of responsive-flow and unresponsive-flow bandwidth allocation, which has influenced the network performance seriously [1].

The fairness means every connection can share the resources of the network fairly when network congestion occurs [2]. The root reason of fairness in the bandwidth allocation is the competition takes place fighting for limited network resources among every data flow as the network has to cause the data packet to lose while happening connectedly [3]. The competitive data flow gets more network resources, but has damaged the interests of other flow.

When the cloud computing network is overloaded, all traffic throughputs are proportional to the target rate of decline in. There is a fair measure of the proportion in standard [4]. The definition of proportional fairness is defined like formula (1).

$$FI = (\sum_{i=1}^{n} x_i)^2 \Big/ n \times \sum_{i=1}^{n} x_i^2 \tag{1}$$

The proportion of subsection I convergence flow which gets the remaining bandwidth to subsection I convergence flow CIR. N is the number of flow convergence. The more equitable index FI close to 1 indicates the better fairness [5].

In order to solve the bandwidth allocation unfairness problem in the cloud computing network, We must deal with non-adaptive UDP and TCP adaptive flow congestion.The paper suggests one method using congestion control algorithm and ACL to fairly realize bandwidth guarantees in the cloud computing network, and provides the end-to-end quality of service over the differentiated services networks.

Received September 17, 2012; Revised January 2, 2012; Accepted January 15, 2013

2. Research Method

Firstly, we should analyse the reason of cloud computing bandwidth allocation unfairness. On the one hand reason, TCP adaptive flow is caused by the different responses when congestion occurs mainly in instructions of congestion. As TCP flows have congestion control mechanism, the source end can reduce the transmission speed on own initiative after receiving the congestion indication. But since there is no end to end congestion control mechanism in UDP flow, UDP cannot reduce the data transmission speed when the congestion happens. As a result, congestion adaptive TCP flow gets less and less resource, non-adaptive UDP congestion gets more and more resources when network congestion happens, at last it leads to the problem of unfair distribution of network resources.

On the other hand reason, micro-flow boundary in the region will be aggregated together for the flow in the cloud computing network. Then PHB processing object area is aggregated traffics rather than micro-flow. The micro-flow with first class aggregated traffics essentially shared reserved resources. Adaptive stream TCP and UDP flow between internal non-adaptability of fairness run mainly by setting up different discarded priority levels. One obvious disadvantage is that the service can not distinguish different types of data streams combined in RED which does not support service differentiation and unable to provide effective equity protection.

2.1. The Congestion Control Algorithm of Cloud Computing Bandwidth Allocation

Our spectral method is constructing matrix by the similar relationship of data points, obtaining the values and eigenvectors of the matrix, choosing the right eigenvectors, and using them to cluster different data points. Spectral clustering algorithm was first used in computer vision, VLSI design and other fields. It has recently begun to apply in machine learning [6]. Without the assumption of the overall structure of data, so spectral clustering is very suitable for the cloud computing network practical problems.

The paper’s method sets up a 2-way objective function NCut based on two sub-graphs A and B divided by graph G:

$$\begin{aligned} Ncut(A,B) &= \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \\ cut(A,B) &= \sum_{u\in A, v\in B} w(u,v) \\ assoc(A,V) &= \sum_{u\in A, t\in V} w(u,t) \end{aligned} \tag{2}$$

Formula (2) can be transformed into:

$$Ncut(A,B) = \frac{1}{1+\dfrac{\sum_{u\in A, v\in A} w(u,v)}{\sum_{u\in A, v\in V} w(u,v)}} + \frac{1}{1+\dfrac{\sum_{u\in B, v\in B} w(u,v)}{\sum_{u\in B, v\in V} w(u,v)}} \tag{3}$$

Seen from formula (3), we can know that the minimum of the objective function NCut meets not only the low similarity of the cloud computing network expression data in the different clusters, but also the high similarity of the cloud computing network expression data in the same cluster.

Given a set of the cloud computing network expression data S= {s1, s2,…,sn} in Rm that the paper wants to cluster it into k subsets. The steps are as following:
Step1 : Constructing the affinity matrix W= (wij):

$$w_{ij} = \begin{cases} \exp\left(-\dfrac{\| s_i - s_j \|^2}{2\sigma^2}\right) & i \neq j \\ 0 & i = j \end{cases} \tag{4}$$

Step2 : Where, si denotes the cloud computing network expression data point, and σ is the parameter proposed first;

Step3 : Constructing the Laplacian matrix L. Defining the diagonal matrix D whose (i, i)-element is the sum of W's i-th row, then;

Step4 : Finding the k largest eigenvectors x1, x2, xk of L, and constructing the matrix X=[x1, x2... xk]? Rnxk by stacking the eigenvectors in columns;

Step5 : Constructing the matrix Y from X by normalizing each of X's rows, i.e.

$$Y_{ij} = X_{ij} / \sqrt{\sum_j x_{ij}^2} \tag{5}$$

Step6 : Treating each row of Y as a point in Rk, cluster them into k clusters via K-means or any other algorithm that attempts to minimize distortion;

Step7 : Assigning the original point si into cluster j if and only if row i of the matrix Y was assigned into cluster j.

The advantage of the spectral clustering method is that it can be used on any shape of sample space and converge in the global optimal [7].

2.2. The Access Control List of Cloud Computing Bandwidth Allocation

The paper divides the network data stream into different levels of priority groups using access control list (ACL) and virtual local area network (VLAN).When the network is congested, they are in different queues by separating TCP and UDP flows. Then we use congestion control algorithm carrying on the rational distribution in between the different rank data stream. It carries on the reasonable dispatch to the network resources, thereby enhances the efficiency of network resources and guarantees the fairness of the data stream. For example, VLAN 200 contains the real-time multimedia data (VOIP and video conferencing) flow all through the UDP protocol transmission, and other VLAN contains TCP protocol transmission flow.

ACL and VLAN for data traffic classification configuration:

```
Router(config-acl)#ip access-list extended 100 permit tcp any any
Router(config-acl)#ip access-list extended 100 permit tcp any any eq ftp
Router(config-acl)#ip access-list extended 100 permit tcp any any eq ftp-data
Router(config-acl)#ip access-list extended 101 permit udp any any range 100 2500!
Router(config)#class-map match-all tcp-class
Router(config)#match access-group 100
Router(config)#class-map match-any udp-class
Router(config)#match access-group 101
Router(config)#match vlan 200
Router(config)#policy-map ACL
Router(config)#class udp-class
Router(config)#police cir 2000000
Router(config)#class tcp-class
Router(config)#police cir 5000000
Router(config-if)#service-policy input ACL
Router(config-if)#switchport trunk encapsulation dot1q
Router(config)#switchport mode trunk
Router(config)#Wrr-queue dscp-based 2 40 46
Router(config-if)#switchport trunk encapsulation dot1q50
Router(config)#mls qos trust dscp
Router(config)#Wrr-queue threshold 1 50 100
```

2.3 Traffic Policing and Traffic Shaping In the Cloud Computing Network

In the cloud computing network, the paper used CAR and GTS to deploy traffic policing and traffic shaping [8]. CAR (Committed Access Rate) limited message traffic through data packet mark or remark to categorize the messages. In the network application, we controlled and improved the network conditions using different flow control, such as controlling FTP the maximum occupied bandwidth. GTS (Generic Traffic Shaping) solved network link interface rate matching problem to make the interface flow to established characteristics.

The traffic policing and traffic shaping results in the experimental was shown in Figure 1. We controlled different flow in the cloud computing network, so that the traffic on an interface consistent with the characteristics set, which was advantageous in guarantees in the network between each service to coordinate mutually [9], thus improved the quality of service fairness. The main method is to restrict the flow of packets, buffer the packets exceeding the traffic convention, and then to send the buffered message out at the appropriate time.

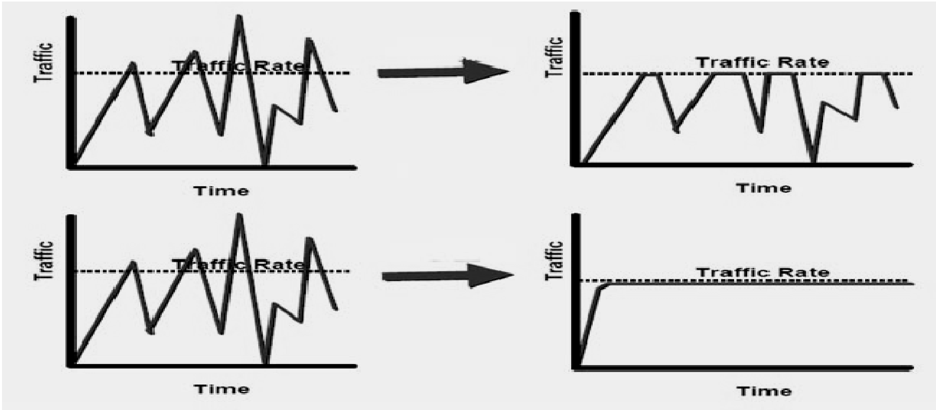

Figure 1. Traffic Policing and Traffic Shaping

3. Results and Analysis

In the simulation experimental, the paper give the preliminary result using ACL and congestion control algorithm to fairly realize bandwidth guarantees in the Cloud Computing Networks. The experiment uses six routers which contained four Cisco2600 and two Cisco3600 and one Catalyst3550 switch supporting routing faction. All equipments are installed Cisco Internetwork Operating System (IOS) software12.2 (15), supporting differentiated services [10]. We monitor the network performance and resource utilization by system service assurance agent (SAA). The Experiment judges the procedure realization application effect through the network test response time. Our network experimental topology is in Figure 2.

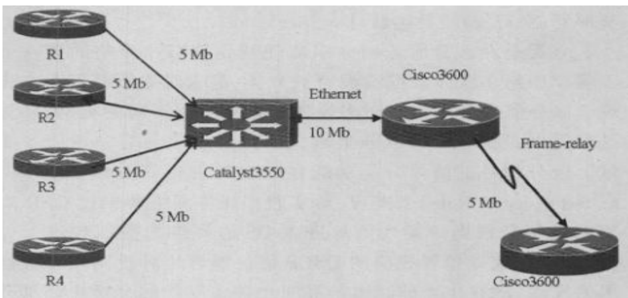

Figure 2. Experimental Topology

Recently, the use of network video conference technology as a means of voice communication is growing rapidly, and the number of VoIP users also increases. In addition, this technology is cheaper, and supports video streaming as well. Overall, the VoIP infrastructure is also available in the cloud computing network [11]. The paper taked the network video conference as the example.The experimental result shown in Figure 3. By using the above-mentioned method for the network conference video frequency when the network delay and network jitter changed, we dedicated bandwidth, managed and avoided network congestion. It reduced packet loss rates and controlled network traffic, thus guaranteed the QoS fairness in the cloud computing network.

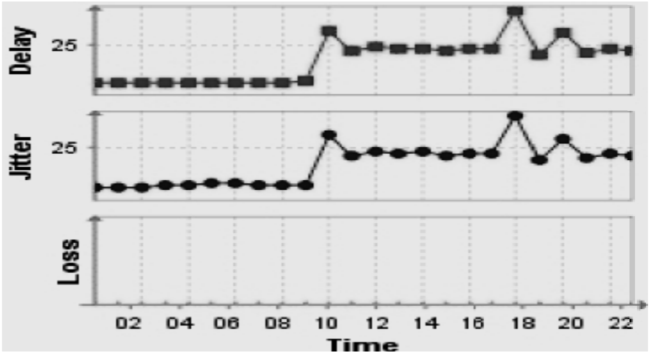

Figure 3. Analysis of Experimental Results

The experimental result in Figure 4 shows the effect of our method dealing with bandwidth allocation in the Cloud computing network. It enhances the efficiency of network resources and guarantees the fairness of the data stream. The real-time multimedia data (VOIP and video conference) flow all through the UDP protocol transmission and other TCP protocol transmission flow are also able to provide effective equity protection. The method allocated network bandwidth fairly and improved network throughput, significantly reduced the packet loss and delay of data packets, differentiated services in order to achieve end to end QoS in the Cloud Computing Networks. In VoIP networks, voice and video streaming can be delivered continuously and simultaneously that require constant connection to traffic Quality of Service(QoS) can be guaranteed.

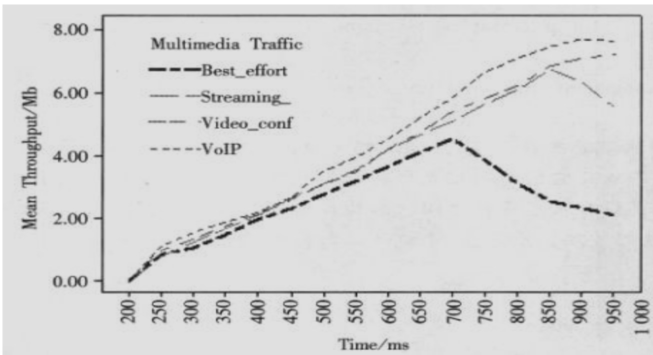

Figure 4. Analysis of Experimental Results

4. Conclusion

To save network resources, multicast transmissions are more and more adopted by the operators when the same information has to reach several destinations in parallel, such as in IPTV services, radio broadcast and video-clip streaming[12]. With the development of the Internet, the application and user of cloud computing are increasing rapidly. Because different applications and users share the network bandwidth, the fairness of bandwidth allocations is becoming more and more important. The paper suggests one method using control list (ACL) and traffic policing and traffic shaping to fairly realize bandwidth guarantees in the cloud computing Networks. The experiment result shows that network bandwidths are fair allocated. The packet loss ratio and latency are obvious improvement, provides the end-to-end quality of service over differentiated services networks. The evolution of cloud computing enables organizations to reduce their expenditure on IT infrastructure and is advantageous to both the serving and served organizations [13]. Further work this paper is on buffering and scheduling bandwidth in the cloud computing network.

References

[1] Gurudatt Kulkarni. Cloud Computing Software as Service. *International Journal of Cloud Computing and Services Science.* 2012; 1(1): 11-16.
[2] Xu Westhead. An Investigation of Multilevel Service Provision for Voice over IP under Catastrophic Congestion. *IEEE Communications Magazine.* 2004; 42(6): 94-100.
[3] Floyd S, Jacobson V. Random early detection gateways for congestion avoidance*. IEEE/ACM Trans on Networking*. 1993; 1(4): 397-413.
[4] AGILENT Technologies Inc. Qos Management Solution for MPLS Virtual Private Network Managed Services. *IEEE Communications Magazine.* 2004; 42(8): 14-14.
[5] Eun-Chan, Park. Feedback-Based Adaptive Packet Marking for Proportional Bandwidth Allocation. *IEEE Transactions on Parallel & Distributed System.* 2007; 18(2): 225-239.
[6] He X, Sun X, Gregor VL. QoS Guided Min-min Heuristiec for grid task scheduling. *Journal of Computer Science and Technology.* 2003; 18(4): 442-451.
[7] Sun Y, He S, Leu J Y. Syndicating Web services: A QoS and user-driven approach. *Decision Support Systems.* 2007; 43(1): 243-255.
[8] Alrifai M, Risse T. *Combining global optimization with local selection for efficient QoS-aware service composition.* Proceedings of WWW 2009. Spain. 2009; 14: 881-890.
[9] Stiliadis D, Varma A. Rateproportional servers: a design methodology for fair queueing algorithms. *IEEE ACM Transon Networking.* 1998; 6(2):164-173.
[10] Menasc DA. Composing Web services. A QoS view. *IEEE Internet Computing.* 2004; 8(6): 88-90.
[11] Suardinata, Kamalrulnizam in Abu Bakar. A Fuzzy Logic Classification of Incoming Packet for VOIP. *TELKOMNIKA.* 2010; 8(2): 165-174.
[12] Augustin WAYAWO. The Transmission Multicast and the Control of QoS for IPV6 Using the Insfrastructure MPLS. *International Journal of Information and Network Security.* 2012; 1(1): 9-27.
[13] Devki Gaurav Pal. A Novel Open Security Framework for Cloud Computing. *International Journal of Cloud Computing and Services Science.* 2012; 1(2): 45-52.