Fair Virtual Bandwidth Allocation Model in Virtual Data Centers

Ying Yuan, Cui-rong Wang, Cong Wang
School of Information Science and Engineering
Northeastern University
Shenyang, China
School of Computer and Communication Engineering
Northeastern University at Qinhuangdao
Qinhuangdao, China
yuanying1121@gmail.com

Journal of Digital Information Management

ABSTRACT: *Network virtualization opens a promising way to support diverse applications over a shared substrate by running multiple virtual networks. Current virtual data centers in cloud computing have flexible mechanisms to partition compute resources but provide little control over how tenants share the network. This paper proposes a utility-maximization model for bandwidth sharing between virtual networks in the same substrate network. The aim is to improve the fairness of virtual bandwidth allocation for multi tenants, especially for virtual links which share the same physical link and carry non-friendly traffic. We first given a basic model and then proved the model can be split into many sub models which can run on each rack-switch port. In the sub model, every physical link is associated with a fairness index constraint in maximize utility calculating. The goal is to limit the differences among the bandwidth allocation of the virtual links which share the same physical link. Experimental results show that the virtual bandwidth allocation between tenants is more reasonable.*

Categories and Subject Descriptors:
C.2.6 **[Internetworking]**: Standards; C.2 **[COMPUTER-COMMUNICATION NETWORKS]** Data Communications

General Terms:
Network Protocols, Network Virtualization, Virtual Bandwidth Allocation

Keywords: Network Virtualization, Multi-tenant, Bandwidth Allocation, Virtual Data Centers

Received: 1 February 2013, Revised 16 March 2013, Accepted 25 March 2013

1. Introduction

Cloud computing delivers infrastructure, platform, and software that are made available as subscription-based services in a pay-as-you-go model to consumers[1]. It aims to power the next generation data centers as the enabling platform for dynamic and flexible application provisioning. This is facilitated by exposing data center's capabilities as a network of virtual services so that users are able to access and deploy applications from anywhere by the demand and QoS (Quality of Service) requirements [2].

In practical applications, data centers which are built using virtualization technology with virtual machines as the basic processing elements are called virtual data centers (VDCs) [3]. Comparing with the traditional data centers, VDCs could provide some significant merits such as server consolidation, high availability and live migration, and provide flexible resource management mechanisms. Therefore, VDCs are widely used as the infrastructure of existing Cloud computing systems [4, 5].

To achieve cost efficiencies and on-demand scaling, VDCs are highly multiplexed shared environments, with VMs and tasks from multiple tenants coexisting in the same cluster. While VDCs provide many mechanisms to schedule local compute, memory, and disk resources [6].

Existing mechanisms for apportioning network resources in current VDCs fall short. End host mechanisms such as TCP congestion control are widely deployed to determine network sharing today via a notion of flow-based fairness. However, TCP does little to isolate tenants from one another: poorly-designed or malicious applications can consume network capacity, to the detriment of other applications which generate non TCP friendly flows. Thus, while resource allocation using TCP is scalable and achieves high network utilization, it does not provide robust performance isolation.

Network virtualization can extenuate the ossifying forces of the current Ethernet and stimulate innovation by enabling diverse network architectures to cohabit on a shared physical substrate [7]. So it can support diverse applications over a shared substrate by running multiple virtual networks, which customized for different performance objectives. Leveraging network virtualization technology, the VDC supervisor can classify virtual machines of same application provider users who rental the VDC resources (CPU, RAM, storage, bandwidth of servers) into same virtual networks so as to achieve more agility that needed by cloud computing and support multi-tenant demand by cloud computing.

In this condition, in order to save energy and improve hardware resource utilization, virtual machines (VMs) which belong to different users may be distributed on a single physical machine. In cloud computing environment, there are competition issues between these users. VMs on same physical machine will compete for network resources, because they all want to occupy enough bandwidth. Even in some extreme cases there may exist malicious competition, some VMs may try to fill the bandwidth to affect the opponent regardless of whether it really need so much network resources.

Therefore the supervisor should to ensure the fairness of bandwidth allocation between those virtual networks first of all. This paper makes the case for virtual bandwidth allocation between virtual links in multi-tenant VDCs to provide more fairness and proactively prevent network congestion. We fist given a basic model and then proved the model can be split into many sub models which can run on each rack-switch port. In the sub model, every physical link is associated with a fairness index constraint in maximize utility calculating. The goal is to limit the differences among the bandwidth allocation of the virtual links which share the same physical link.

The rest of the paper is organized as follows. Section 2 gives a basic utility-maximization model for bandwidth sharing between virtual networks of multi tenants. Section 3 proves the basic model can be split into many sub models which can run on each rack-switch port and introduces the sub model in details. Section 4 gives the conclusion.

2. Basic Model

While VDCs provide many mechanisms to schedule local compute, memory, and disk resources[8], existing mechanisms for apportioning network resources fall short. For example, end host mechanisms such as TCP congestion control are widely deployed to determine network sharing today via a notion of flow-based fairness. However, TCP does little to isolate tenants from one another: poorly-designed or malicious applications can consume network capacity, to the detriment of other applications which generate non TCP friendly flows[9]. Thus, while resource allocation using TCP is scalable and achieves high network utilization, it does not provide robust performance isolation.

There is a trade-off between ensuring isolation and retaining high network utilization. Bandwidth reservations, as realized by a host of mechanisms, are either overly conservative at low load, which can achieve poor network utilization, or overly lenient at high load, which can achieve poor isolation. An ideal bandwidth isolation solution for cloud datacenters has to scale and retain high network utilization. It has to do so without assuming well-behaved or TCP conformant tenants. Since changes to the NICs and switches are expensive, take some time to standardize and deploy, and are hard to customize once deployed, edge or so ware based solutions are preferable.

Network virtualization is a promising way to support diverse applications over a shared substrate by running multiple virtual networks, which customized for different performance objectives. VDC supervisor can classify virtual machines of same application into same virtual networks so as to achieve more agility that needed by cloud computing. In this condition, the fabric owner as the substrate provider to provide the physical hardware, service provider rent virtual machines and virtual networks to run their applications to provide service, therefore can achieve a flexible lease mechanism [10] which is more suitable for modern commercial operations.

Today, network virtualization is moving from fantasy to reality, as many major router vendors start to support both switch virtualization and switch programmability. In the main thought of network virtualization, virtual machines of same applications in data center are partitioned into same virtual networks by network slicing. In this case, for QoS guarantee, the most important thing in VDC network is the bandwidth indemnification in each virtual link, i.e. the virtual bandwidth allocation mechanism must be well considered. We will first give a basic model for bandwidth allocation between virtual networks of multi tenants, and then proof that the global problem can be divided into sub-problems on each rack switch's port.

The virtual bandwidth sharing problem in this section is considered from the overall perspective of the system which contains virtual networks and substrate hardware resource in VDC under multi-tenants market mechanism under cloud computing. Consider a substrate network with a set L of links, and let C_j be the capacity of substrate link $j \in J$. The network is shared by a set N of virtual networks and indexed by i. Define a vector b_{il} which denotes the allocated bandwidth of virtual network i in link j. Let $U_i(b_{il})$ be the utility of virtual network i as a function of his bandwidth b_{il}. Note that the utility $U_i(b_{il})$ should be an increasing, nonnegative and twice continuously differentiable function of b_{il} over the range $b_{il} \geq 0$. The bandwidth control problem can be formulated as the following optimal problem. The bandwidth control problem can be formulated as the following optimal problem $P1$:

$$P1: \quad \text{MAX} \sum_{i=1}^{N} \sum_{l=1}^{L_i} U_i(b_{il}) \quad \text{s.t.} \quad \forall i, \forall l, \quad b_{il} \geq 0, \quad \forall l \in L, \ \sum_{i}^{N} b_{il} \geq C_l, \quad F(b) \leq D$$

The meaning of the constraints are: 1) for each virtual network, its bandwidth must be greater than or equal to 0; 2) the total bandwidth of all virtual networks in physical link *l* does not exceed the maximum bandwidth C_l allowed by the hardware; 3) $F(b) \leq D$ is an addition condition to introduce the fairness constraint which we will discussed in next section and D is a constant. Since the utility functions are strictly concave, and hence continuous, and the feasible solution set is compact then the above optimal problem has a unique optimal solution.

$P1$ is an overall problem with multiple constraints. For each virtual network, it can calculate its profit for the allocated virtual bandwidth b_{il}. The goal is to achieve the maximize sum utility of all virtual networks. However, problem $P1$ for virtual bandwidth allocation in VDC is still a global issue. Obviously, it's impossible to manage vast amounts of virtual nodes in large-scale network as well as in VDC. There is no gain in using the local interpretation unless we can devise a local way to solve the problem.

3. Distributed Sub Model

In practical VDCs, using a distributed manner to address the complex optimization system such as $P1$ is an efficient way. In this subsection, we divided the overall problem into sub-problems which can be solved on the port of rack switches in VDC. The overall optimization problem $P1$ can be split into several distributed problem. Recall the problem in the previous section, each b_{il} in vector $b_{il} = (b_{il}, l \in L)$ denotes the allocated bandwidth for virtual network i in link l. For any set of virtual network links L_i we have $L_i \subseteq L$. Taking into account of the diversification and variability of virtual network topologies, we need to regulate all dimensions of L_i and make it equal to the dimension of L. Let L_i be the new link set of virtual network i, the two under sets are equivalent.

$$L_i' \equiv L_i, \forall l \in L, \begin{cases} L_{il}' = 1, \text{ if } l \in L_i; \\ L_{il}' = 1, \text{ if } l \notin L_i; \end{cases} \tag{1}$$

$L_{il}' = 0$ means that the bandwidth allocated for virtual network i on link l is 0. Then the overall optimal problem $P1$ can be rewritten as:

$$\text{MAX} \sum_{i=1}^{N} \sum_{l=1}^{L_i} U_i(b_{il}) = \sum_{l=1}^{L_i'} \sum_{i=1}^{N} U_i(b_{il})$$
$$= \sum_{i=1}^{N} U_i(b_{i1}) + \sum_{i=1}^{N} U_i(b_{i2}) + ... + \sum_{i=1}^{N} U_i(b_{iL})$$

Because the utility function $U(x)$ is nonnegative over the range $x \geq 0$, then the overall optimal problem $P1$ can be split into L sub optimization problems on each network port of all rack switches in data center. When the bandwidth allocation among virtual links for every virtual network on each physical links of the rack switch is optimal, the overall optimization problem $P1$ can be solved simultaneously. This distributed sub model is feasible in large-scale data center network, and can be easily achieved.

By now, we have demonstrated the overall problem $P1$ can be split into a number of distributed sub problems (in equation (2)) which can be solved on the rack switch in VDCs through a ware based solutions leveraging network virtualization technique over a programmable switches. Then we can discuss the details of the sub model for real system. Firstly, we can get the sub optimization problems:

$$\max \sum_{i=1}^{n} \omega_i \bullet U_i(b_i) \quad \text{s.t.} \quad \forall i, \ b_i \geq 0, \quad \sum_{i}^{n} b_i \geq C, \quad F(b) \leq D \tag{3}$$

Where ω_i is the weight of virtual network i, n is the number of virtual networks, b_i denotes the allocated bandwidth for virtual network i at link l. $U_i(b_i)$ is the utility function discussed in the previous section. Vector $\boldsymbol{b} = (b_1, b_2,..., b_n)$ is a solution for problem (3). Because the feasible region is compact and the objective function is continuous. We form the Lagrangian for the problem (3):

$$L(\boldsymbol{b}, \lambda, \mu) = \sum_{i=1}^{n} \omega_i \bullet U_i(b_i) + \lambda \bullet (C \sum_{i}^{n} b_i) + \mu \bullet (D - F(\boldsymbol{b})) \tag{4}$$

Here the second term is a penalty for the physical bandwidth capacity constraint; the third term is an additive penalty which is used for the fairness constraint of virtual bandwidth allocation in our model. Thus we need a method to describe the quantitative of the fairness for bandwidth allocation, which can be defined as follows:

$$F(\boldsymbol{b}) = \frac{n \bullet (b_1^2, b_2^2,..., b_n^2)}{(\sum_i^n b_i)^2} \tag{5}$$

Equation (5) can reflect the fairness of resource allocation in the system. The range of $F()$ is $[1, +\infty)$, and the more $F()$ is closer to 1 the more fair of the bandwidth allocation. But consider the weight ω_i of virtual networks, which the bandwidth allocation should according to. Obviously, if all virtual networks have the same weight then when $F() = 1$ the allocation is completely fair. But when the weight is unequal when $F() = 1$ the allocation seems not completely fair, thus we must consider the weight of each virtual network. Substitute ω_i to b_i:

$$f = \frac{n \bullet (\omega_1^2, \omega_2^2,..., \omega_n^2)}{(\sum_i^n \omega_i)^2} \tag{6}$$

In normal conditions $1 \le f < +\infty$, $f = 1$ means each virtual network has same weight. Thus, constant D in the optimal model (3) should be no less than f, because when the weights of virtual network are unequal, the completely fair allocation solution exits at $D = f$.

In the end of this section, we discuss the existence of the optimal solution of problem (3). Substitute (5) to (4):

$$L(\boldsymbol{b}, \lambda, \mu) = \sum_{i=1}^{n} \omega \bullet U_i(b_i) + \lambda \bullet (C - \sum_{i}^{n} b_i) + \mu \bullet \left(D - \frac{n \bullet (b_1^2 + b_2^2 + ... + b_n^2)}{(\sum_i^n b_i)^2}\right) \quad (7)$$

The Slater constraint qualification holds for the problem (3) at point $b = 0$, because then $0 = \sum_i^n b_i < C$, this guaran -tees the existence of Lagrange multipliers λ and μ. In other words, because the objective function is concave and the feasible region is convex, there exist at list one feasible vector $\boldsymbol{b}$ is optimal.

4. Performance Evaluation

We implemented the distributed virtual bandwidth allocation schema discussed in Section 3 using OpenFlow VM [11]. We made a simple but sufficiently persuasive experiment as shown in Figure 1. In virtual link 1, VM1 sends TCP traffic to VM3; in virtual link 2, VM2 sends UDP traffic to VM4. Both traffic are generated by Iperf, the physical link bandwidth is set to 1Gbps.

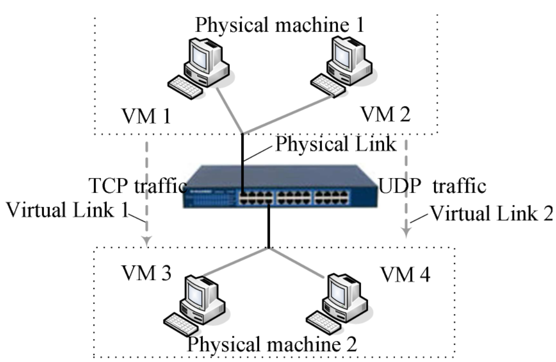

Figure 1. Experimental topology

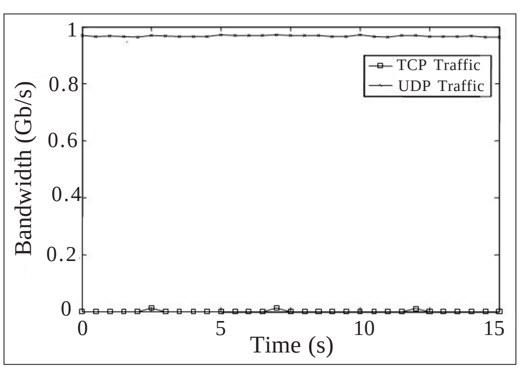

Figure 2. Without network virtualization

As well-known in network congestion control problem [9], without using any virtual bandwidth control mechanism, when the two virtual machines do their best to send the traffic, so as to occupy the full bandwidth of the physical link. UDP traffic which is non TCP friendly is dominant, and TCP traffic just rarely able to successfully send few packets (as the result shown in Figure 2 which is the first experiment on the topology in Figure 1). The main cause of this situation is because of the TCP congestion control mechanisms, when congestion occurs the sliding window will limit the sending rate to a small value.

Then, we built virtual networks for both traffic (*VN* 1 for TCP traffic and *VN* 2 for UDP traffic) and verified our mode for fair virtual bandwidth allocation discussed in this paper on the same topology (Figure 1). We set $n = 2$, $D = 1.1$, $C = 1$ and the weight of each virtual link $n1$ and $n2$ is 1. Figure 3 shows the bandwidth consumption of the two non-friendly data traffic after network slicing and the optimal model we proposed in this paper. From the experimental result we can see that the model presented in this paper can achieve fair bandwidth allocation between virtual links host on the same physical link.

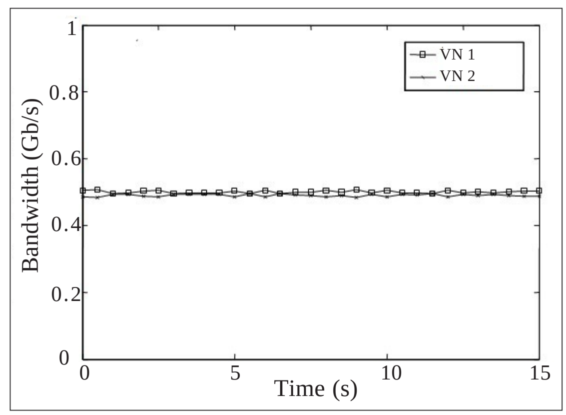

Figure 3. With network virtualization and fair bandwidth allocation model

5. Conclusion

This paper introduced a utility-maximization model for fair bandwidth sharing between virtual networks in VDCs to provide multi tenants mechanism for network resource under cloud computing. In the model, every physical link is associated with a fairness index constraint in maximize utility calculating. The goal is to limit the differences among the bandwidth allocation of the virtual links which share the same physical link. In our future work, we will test the trade-off between fairness and the bandwidth utilizing rate under some conditions such as not all virtual networks need bandwidth guarantee, i.e. the dynamic bandwidth allocation taking into account the fairness.

References

[1] Armbrust, M., Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., Zaharia, M. (2010). A View of Cloud Computing. *ACM Communications*, 53 (4) 50–58.

[2] Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., Brandic, I. (2009). Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility. *Future Generation Computer Systems*. 25 (6) 599–616.

[3] Xu, J., Zhao, M., Fortes, J., Carpenter, R., Yousif, M. (2007). On the Use of Fuzzy Modeling in Virtualized Data Center Management. ICAC'07: Proceedings of the 4th International Conference on Autonomic Computing, p. 395–400, Hong Kong, HK, China. *IEEE*.

[4] VMware. (2012). VMware vCloud. http://www.vmware.com/technology-/cloud-computing.html.

[5] Amazon, (2012). Amazon Elastic Compute Cloud (Amazon EC2). http://aws.amazon.com/ec2/.

[6] Yan, W., Zhou, L., Lin, C. (2011). VM Co-scheduling: Approximation of Optimal Co-Scheduling in Data Center. AINA'11: Proceedings of the 25th International Conference on Advanced Information Networking and Applications, p. 340–347, Biopolis, Singapore. *IEEE*.

[7] Khan, A., Zugenmaier, A., Jurca, D., Kellerer, W. (2012). Network virtualization: a hypervisor for the Internet?. *IEEE Communications Magazine*, 50 (1) 136–143.

[8] íñigo, G., Berral, J. L., Fitó, J. O., Julií, F., Nou, R., Guitart, J., Gavaldí, R., Torres, J. (2012). Energy-efficient and multifaceted resource management for profit-driven virtualized data centers. *Future Generation Computer Systems*, 28 (5) 718–731.

[9] Zhang, P., Wang, H., Cheng, S. (2011). Shrinking MTU to Mitigate TCP Incast Throughput Collapse in Data Center Networks. CMC'11: Proceedings of the 3th International Conference on Communications and Mobile Computing, p. 126–129.

[10] Alan, S., Srikanth, K., Albert, G., Changhoon, K. (2010). Seawall: Performance Isolation for Cloud Datacenter Networks. HotCloud'10: Proceedings of the 2nd USENIX Workshop on Hot Topics in Cloud Computing, p. 1–12, Boston, MA, USA. USENIX.

[11] Stanford Clean Slate Program. (2012). The OpenFlow Switch Consortium. http://www.openflowswitch.org/