Regional Load Balancing Circuitous Bandwidth Allocation Method Based on Dynamic Auction Mechanism

Wang Chao¹, Zhang Dalong², Ran Xiaomin¹

¹National Digital Switching System Engineering and Technological Research Center, Zheng Zhou, China; ²College of Information Engineering, Zhengzhou University, Zhengzhou, China

Abstract. Aiming at the problem of link congestion caused by the shortage of network bandwidth resources at the user end, this paper first proposes a regional load balancing idea. Then, for the problem of bandwidth resource allocation in regional load balancing, a bandwidth allocation model is established and a dynamic auction algorithm is proposed. The algorithm calculates the link quality and stability by constructing a link model, and introduces the auction bandwidth to the auctioneer's incentive degree to obtain the auction bidding function. The simulation results show that the algorithm can effectively improve the user's network status, reduce the service response delay, increase the throughput, and at the same time can effectively prevent the auction user's false bidding behavior, so that the auction quote quickly converges to the maximum quote, reduces the number of auctions, and reduces communication overhead.

1 INTRODUCTION

The rapid development of the Internet has put higher demands on the bandwidth. Although the link capacity is quite adequate, overconsumption of link congestion and bandwidth is still widespread in the Internet. Therefore, it is of great significance to efficiently use the available bandwidth resources in the network.

For the allocation of limited resources, the price control mechanism is an effective solution (Song Qizhu, 2010). The most typical control mechanism based on price is to set a static price for each unit resource, the most obvious advantage of this method is simple and easy to implement (Odlyzko A, 1997). However, the static pricing model can not adapt to the dynamic changes of network traffic in the network environment where the resource’s demand changes dynamically. In order to solve the problem of bandwidth resource allocation in dynamic network environment better, a dynamic pricing model is proposed. A classical dynamic pricing model uses the auction way to price the network bandwidth for maximum benefit (Semret N, 1999). Dramitinos M (2007) proposed a distributed network bandwidth auction mechanism for Dutch multi-unit auctions by setting a sufficiently high bidding price for each link's bandwidth and conducting price reduction auctions based on variable price reduction rules until the first bidder appears. The method uses different bidding prices to reflect the imbalance of link’s bandwidth requirements, meanwhile, the pricing method using VCG (Vickrey-Clarke-Groves) mechanism ensures the bidding behavior of the bidders. However, this mechanism requires bidders to monitor bidding progress at all time, which leads to large communication costs. Tipsuwan Y (2009) comparatively the advantages and disadvantages of three bandwidth allocation methods, and proposed a centralized auction mechanism based on game theory and dynamic bandwidth allocation method. As bidders initially submitted price may be different from the final payment price, bandwidth utilization and initial offer is used to determine its final offer and avoid fake quotations.

Link evaluation is a way to improve routing performance through link quality (Chan K P, 2003 and Lin-Yuan L, 2010). To solve the problem of network congestion caused by multiconcurrency services, this paper first proposes a solution of regional load balancing, then proposes a dynamic bandwidth auction method. Based on the centralized dynamic auction mechanism, the classification of bandwidth resources is firstly refined by evaluating the link’s quality to meet the needs of different bidders. The pricing rules based on link’s quality can effectively inhibit bidding's fake bidding behavior. At the same time, this paper also proposes auction incentive value mechanism, which determines the degree of stimulation or suppression of bandwidth resources to the bidders in each round of auctions, it avoids the communication overhead of multiple interactive interactions between auction parties, not only
ensures that the auction ends within limited rounds, but also makes auction quotes quickly tend to the maximum. Finally, the problem is simplified by separately solving the maximum benefit of each unit instead of calculating the global maximum benefit.

The rest of this paper is organized as follows. In Section 2, we introduce the concept of regional load balancing. In Section 3, we describe the model of bandwidth allocation and link quality. Then we describe the bandwidth auction algorithm in Section 4. Extensive simulations are discussed in Section 5. Section 6 concludes this paper.

## 2 REGIONAL LOAD BALANCING

In a multi-ISP (Internet Service Providers) network, the traditional multi-WAN router only considers the load balance within the device and does not consider the load balancing within the region, it easily leads to the user's data transmission between different ISPs. However, the bridge bandwidth between different ISP servers is often difficult to meet the demand of large amount of data transmission, which limits the implementation of load balancing. Meanwhile, due to massive data forwarding between ISP bridges, the transmission delay of user services under the load balancing policy increases, making it difficult to satisfy users' QoE.

In order to solve the above problem, this paper presents a regional load balancing technology - through multiple gateways (the gateway needs more than two WAN ports) to collaborate with each other, and to complete the round robin forwarding of the services of the congested gateways within a certain spatial domain. That is, using the corresponding WAN ports and links of idle gateways in the local area to implement detouring and forwarding of the congested links, and avoiding the preemption of load balancing services on the bridge bandwidth between ISPs. Enabling multiple users to reuse access bandwidth resources of a certain WAN port on demand to improve network access rate and throughput so as to complete distributed traffic balance at the bottom and enable efficient access to data, video, and other services on different ISP servers.

Regional load balancing is a load balancing technology that enables cooperative forwarding and traffic transfer among multiple devices within a certain spatial domain to round off congested links. Forwarding user's data on the congestion side by borrowing the links of other idle gateways, the load of the congested links and nodes can be release and the service response time of the congested gateway can be reduced, so as to achieve distributed traffic balance. The principle is shown in Figure 1. When multiple gateways in a network cooperate with each other to complete bandwidth renting and link bypassing, traffic load of the congested nodes is transferred to idle nodes. Adopting this circuitous strategy can avoid the transmission pressure of traffic between two different ISPs and implement the underlying multi-network convergence.

![Figure 1: Regional load balancing principle diagram](image)

## 3 MODEL BUILDING

### 3.1 Auction Model

Although the link capacity is sufficient, as the network user base and the user's level of network quality requirements increase dramatically, the user's demand for bandwidth is far greater than the bandwidth resources provided by the operator. While centralized allocation algorithm is easy to form a "heat island" effect, and the central server is easy to become a performance bottleneck, and there is a delay in the interaction and update of the information, which can not well adapt to the network with frequent dynamic changes. To rationally allocate resources on demand and ensure the fairness of resource allocation as well as the benefit of resource providers, this paper adopts the auction method for the bandwidth allocation of users with the same bandwidth requirements.

For the allocation of bandwidth resources to users requesting the same bandwidth, this paper describes the auction model as an optimization problem. According to bidders submit the offer information to determine the bandwidth allocation program. The model defines participants as three types: participants with free bandwidth resources are auctioneers, participants who request bandwidth resources are bidders, agents who decide on bandwidth allocation results and payment rules for both bidders.

During the auction, the total revenue obtained by defining $N$ auctioneers is $\sum_{i=1}^{N} W_i$, where $W_i$ is the revenue obtained by the auctioneer $i$, so the auction model can be abstracted as follow:

$$\text{max} \sum_{i=1}^{N} W_i$$

$$\text{s.t.} \sum_{j=1}^{M} B_j \leq C, B_j \geq 0, j = 1, \ldots, M$$

Where $B_j$ is the bandwidth requirement of bidder $j$. 

---

1. Draft writer: [Name]
2. Reviewer: [Name]
3. Editor: [Name]
In order to solve the maximum value, the above optimization problem is decomposed to solve the optimal value of each auctioneer so as to optimize the overall return. Therefore, Eq. (1) can be decomposed into the following formula:

$$\max W_i$$

s.t. $\sum_{j=1}^{M_i} B_j \leq B_i$, $B_j \geq 0$, $i = 1, \ldots, M_i$  \hspace{1cm} (2)

Where $B_i$ is the free bandwidth resource provided by the auctioneer $i$ and $M_i$ is the total number of bidders owned by the auctioneer $i$.

3.2 Link Quality

The purpose of the region load balancing strategy is to realize the traffic transfer from the congested node to the idle node, that is to select a better-quality path for data transmission. In the transmission of the service, the congestion status of the link is more important than the number of hops to affect users’ QoS.

To achieve regional load balancing, you need to measure the quality of the path between the client and the server. In the end-to-end multi-hop link topology shown in Figure 2, D is a server, S is a client gateway, and the others nodes are routers. For the client, the effect of the neighbor node on it’s QoS is higher than that of the other nodes.

Therefore, the link weight introduced is as the impact value of each link quality on the client. The weight of link $k$ is:

$$\omega(k) = \frac{1}{N + 1 - k}$$  \hspace{1cm} (3)

Where $N$ is the total number of end-to-end hops, $k$ is the link number, and the closer it is to the client $k$, the larger it is.

![Figure 2: End-to-end multi-hop forwarding link](image)

In this paper, the round-trip delay of packets is the main parameter of link quality. In normal conditions, we define the average roundtrip delay is \(RTT_0\), then the quality of link $k$ is:

$$Q(k) = \frac{RTT_k}{RTT_0} \omega(k)$$  \hspace{1cm} (4)

Where \(RTT_k\) is the average round-trip delay of link $k$ in the quality calculation period. Therefore, the client’s port $\tau$ to the operator's server path quality $Q'$ is:

$$Q' = \sum_{k = 1}^{N} \frac{RTT_k}{RTT_0} \omega(k)$$  \hspace{1cm} (5)

In the multi-hop path from the user terminal to the server, there may be multiple links with the same quality. In order to distinguish the quality of the links, a link’s stability parameter is introduced. Assume that the end-to-end path performs $m$ measurements within one RTT calculation period, so the link RTT matrix $R$ is obtained:

$$R = \begin{bmatrix}
RTT_{i1} & \cdots & RTT_{im} \\
\vdots & \ddots & \vdots \\
RTT_{N1} & \cdots & RTT_{Nm}
\end{bmatrix}$$  \hspace{1cm} (6)

Where \(RTT_j\) represents the $j$th RTT measurement of link $i$. Therefore, the link’s stability $S'$ is the variance of matrix $R$:

$$S' = D(R)$$  \hspace{1cm} (7)

4 BANDWIDTH AUCTION ALGORITHM

As the bidder often requests more bandwidth than the actual demand in bidding, in order to avoid the bidding behavior of the bidder and result in the waste of resources, the bidding party’s payment function $P_{bid}$ is defined, indicating that the requested link quality is $Q_{bid}$ and the bandwidth is $B_{bid}$, the bidder’s payment function is defined by the following formula:

$$P_{bid} = \alpha(Q_{bid} - Q_{max})B_{bid} - \beta S_{bid}$$  \hspace{1cm} (8)

Where $\alpha$ and $\beta$ are unit price coefficient constants of link quality and bandwidth. If a bidder requests a bandwidth greater than the actual demand, the cost to be paid by the bidder increases with a certain difference in the quality of the link, thereby suppressing the bidder’s fake bidding behavior. If the bidder needs a certain bandwidth, the higher the bidding link quality is, the higher the cost will be borne.

Bidding process, different bidders may be different, set a quotation in the highest offer is $F_{max}$, then the next quotation, the last highest bid is the reserve price of the offer. Suppose the $n$th bid quotation of bidder $i$ is $F_{i}(n)$, the highest quotation of the bid is $F_{max}(n)$, in order to reduce the bidding process information interaction, making bidding quickly converge to the highest quotation, we propose an incentive function $\theta(n)$, said the item is attractive or repulsive to the bidder in the quote. The incentive value is determined by the attraction generated by the ratio of the quality of the bidding object to the bidder’s own network quality and the price increase per bidding quote, as follows:

$$\theta(n) = \begin{cases} 
\frac{Q_{bid}}{Q_{max}}, & n = 0 \\
\theta(n-1) - \frac{F_{max}(n-1) - F_{i}(n-1)}{F_{i}(n-1)}, & n \geq 1 
\end{cases}$$  \hspace{1cm} (9)

When $\theta(n) > 0$, bidder $i$ will conduct the next bid, when $\theta(n) \leq 0$, bidder $i$ will give up bidding. From Eq.(9), we can see that when $n \geq 1$, $\theta(n)$ is a decreasing function, that is the bidding enthusiasm of bidders will gradually decrease as the number of bidding increases. After limited rounds of bidding, $\theta(n) \leq 0$, bidding participants give up auction and the auction ends.

From Eq.(8), we know that the maximum that an auctioneer can obtain in each round auction is
\[ \max P_{\text{bid}}(B_{\text{tot}}) \approx \alpha [\ln(n) + C] \cdot B_{\text{tot}}, \]

where \( C \) is the Euler's constant and \( C = 0.57722 \ldots \). In the case of the same bandwidth request and the number of hops between the client and the operator's server, \( B_{\text{tot}} \) and \( n \) is a constant. In order to make the bidding price quickly converge to the maximum bidding price, we set up bid \( L \) times per round. When \( \theta_i(n) > 0 \), the bidder bid increases to be \( \Delta P(n) \), so the bidding round of bidder \( i \) is:

\[
P^\prime_{\text{bid}}(n) = \begin{cases} P_{\text{bid}} & \text{if } n = 1 \\ F_{\text{max}}(n-1) + X_{\theta_i(n)} \cdot \Delta P(n), & \text{if } n \geq 2 \end{cases}
\]

In the Eq.(10), where \( X_{\theta_i(n)} = \begin{cases} 0, & \theta_i(n) \leq 0 \\ 1, & \theta_i(n) > 0 \end{cases} \) and \( 0 \leq \Delta P(n) \leq \max P_{\text{bid}}(B_{\text{tot}}) \). From Eq.(10), we can see that each round of bidding the quote is higher than the highest quote of last round of bidding, combined with Eq.(9), it can be obtained within limited rounds of auctions, the quotation will quickly converge to the highest final price.

5 SIMULATION ANALYSIS

In order to verify the advantages and dis-advantages of the proposed bandwidth allocation algorithm, NS2 is used to design the following experiment.

The experimental results are shown in Figure 3, There were six users who made fake quotes and compared the performance and cost before and after their quotes.

![Figure 3: Fake quote profit contrast](image)

We can see from the Figure 3 that when users make fake quotes, their performance improvement ratio is far less than the increase in cost, so the pricing rules provided in this paper can effectively avoid the problem of false quotes.

In order to evaluate the impact of the proposed algorithm on the performance of networks, the effect of link quality and bandwidth on users is simulated under two scenarios respectively: Case 1 is the effect on performance of users leasing different bandwidths under the same link quality; case 2 is the effect on the user's performance under the condition that the bandwidth of the user is the same and the link quality is different. In the case 1, the delay of three users and the throughput result are shown in Figure 4 and 5; In the case 2, the delay and throughput results shown in Figure 6 and 7.

![Figure 4: User's delay in case 1](image)

![Figure 5: User's throughput in case 1](image)

![Figure 6: User's delay in case 2](image)

![Figure 7: User's throughput in case 2](image)

Comparing Figure 4 and 6 shows that after the transmission is stable, when the link quality is different and the bandwidth is the same, the better the link quality is, the lower the delay is. As can be seen from the comparison between Figure 5 and 7, the throughput of the client is related to the bandwidth and the link quality. When the bandwidth is the same, the better the link quality is, the larger the throughput is. In the case of a certain link quality, the larger the bandwidth, the greater the throughput. Because the method proposed in this paper improves the link quality and increases the logical bandwidth at the same time, so can better improve user’s QoE.

To verify the advantages and disadvantages of the proposed algorithm, we simulate the bidding situation of
bidders under the same bandwidth and bidding under different bandwidth requests by multiple bidders. The simulation results are shown in Figure 8 and Figure 9.

![Figure 8: The same bandwidth's quote](image1)

![Figure 9: Different request bandwidth's quote](image2)

Where Figure 8 is a bidding result of four bidders bidding simultaneously with a bandwidth of 12M. From Figure 8, it can be seen that in the case of a bidding bandwidth of a certain size, bidders' initial quotations are not necessarily the same. In the bidding process, each Therefore, the bidders bid in the next round of bidding is also different, and the highest bid of the last bid is the lowest bid of the next round of bidding. Within the limited rounds of bidding, the quotation reaches the highest quotation of the bandwidth. When a bidder first reached the maximum quotation, the auction ended. As can be seen from Figure 9, for different auction bandwidths, the initial bidding price is different and the highest bidding price is also different. The price difference of bidding for two adjacent bidding rounds has nothing to do with the bidding bandwidth, but both reach the highest quotation within the limited bidding times. Therefore, when the highest bid price is obtained for each round of quotations, the total auction proceeds reach the maximum.

6 CONCLUSION

In this paper, we propose a regional load balancing scheme based on dual WAN routers from the perspective of users and propose an auction-based bandwidth allocation strategy for the link congestion in the network and the uneven allocation of bandwidth resources. Based on the idle state and link quality of the client routers in the logical area, the bandwidth resources are auctioned and allocated dynamically. The simulation results show that the bandwidth allocation strategy proposed in this paper can effectively ensure that the auctioneer can obtain greater benefits while improving bidder QoS.

References