Multi-objective Egress Router Selection Policies for Inter-domain Traffic with Bandwidth Guarantees

Kin-Hon Ho, Ning Wang, Panos Trimintzios, and George Pavlou

Centre for Communication Systems Research, University of Surrey,
Guildford, Surrey, England, United Kingdom GU2 7XH
{K.Ho, N.Wang, P.Trimintzios, G.Pavlou}@surrey.ac.uk

Abstract. The next generation Internet is designed to accommodate flows that span across multiple domains with quality of service guarantees, in particular bandwidth. In this context, destinations for inter-domain traffic may be reachable through multiple egress routers within a domain. In this paper, we formulate a bandwidth guaranteed egress router selection problem. The objective is to, for each aggregated inter-domain traffic flow, select an egress router that satisfies the end-to-end bandwidth requirement while optimizing the network resource utilization by which we consider three objective functions: minimizing the total bandwidth consumption, improving intra-domain and inter-domain load balancing in the network. We propose a heuristic algorithm with five egress router selection policies to solve this problem. The evaluation of these egress router selection policies through simulation benefits ISPs by choosing the one that fits their target objectives.

1 Introduction

As the Internet has grown in size and diversity of applications, the next generation Internet is intended to accommodate flows with end-to-end Quality of Service (QoS) guarantees across multiple domains. To provide efficient end-to-end QoS guarantees, QoS routing and Traffic Engineering (TE) have become indispensable: the former selects a path that meets the QoS requirements while the latter optimizes resource utilization in order to be able to carry more traffic flows in the network. In the past decade, there has been a considerable amount of work on QoS routing and traffic engineering at the intra-domain level. However, only little attention has been given to the inter-domain problem. We consider that inter-domain QoS routing and traffic engineering should be addressed for the following reasons.

Inter-domain QoS: End-to-end QoS over the Internet covers the intra-domain and inter-domain QoS. Even though research in intra-domain QoS is mature, the lack of inter-domain QoS support hinders the deployment of end-to-end QoS. Thus, together with the current QoS-aware intra-domain routing, inter-domain QoS routing will facilitate an end-to-end QoS-based Internet, which will benefit Internet Service Providers (ISPs) and their customers. The current inter-domain routing protocol, Border Gateway Protocol (BGP), however, does not cater for QoS support.

Inter-domain TE: Inter-domain traffic engineering [1] concerns forwarding traffic entering or exiting a network based on some optimization objectives. One of the inter-domain traffic engineering problems is to direct the inter-domain traffic flows to the 'best' egress router within a domain towards certain destination prefixes; this we call the "egress router selection" problem. The problem arises when a domain has multiple connections to neighboring domains, so that a situation can emerge that a destination prefix is reachable through multiple egress routers. In that context, selecting different egress routers for traffic flows can have diverse effects on network resource utilization. Addressing inter-domain TE is important because appropriate selection of egress routers for inter-domain traffic flows benefits ISPs by improving the network resource utilization. Inter-domain traffic engineering, however, is commonly applied today in a trial-and-error only fashion [1].

Based on this reasoning, we aim to develop a systematic approach to solve this inter-domain TE problem with end-to-end QoS guarantees, i.e. QoS guaranteed egress router selection. Assuming that most QoS requirements can be derived from bandwidth [2], our work only focuses on providing bandwidth guarantees. Thus, the problem we address becomes Bandwidth Guaranteed Egress Router Selection (BGERS).

Our goal is summarized as follows: *Given an aggregated customer traffic flow in an ISP network, select an egress router that satisfies the customer end-to-end bandwidth requirement while optimizing resource utilization in the network. Each customer traffic flow consists of a destination prefix that belongs to a remote domain and a bandwidth requirement. An egress router must be selected amongst egress routers that offer the guaranteed bandwidth to the destination prefix*.

With respect to optimizing network resource utilization, we consider three objectives: Minimizing the total bandwidth consumption, improving intra-domain and inter-domain load balancing in the network. We do not consider optimizing multiple objectives simultaneously since, for example, the objective of minimizing bandwidth consumption and load balancing are contradictory with each other.

Related work on inter-domain QoS routing and TE are as follows. Bonaventure [3] focuses on how to distribute flexible QoS information by BGP in different network scenarios. Cristallo and Jacqenet [4] propose a new attribute, the QoS_NLRI (Network Level Reachability Information), for the BGP UPDATE message to carry QoS information. Xiao [5] proposes a similar QoS extension to BGP to perform the bandwidth advertising and routing. On the other hand, research on egress router selection has only been done in the context of best-effort traffic. Bressoud [6] determines an optimal selection of outgoing links and associated border routers, where the selection optimizes the ISP's network resource utilization. The ISP, however, can only select an egress router based on prefix reachability and the egress link capacity information, without knowing whether the selected egress router can satisfy the traffic flow's end-to-end bandwidth requirement.

The key to solve the BGERS problem is the support for traffic engineering information (e.g. bandwidth) distributed within and between domains. In this paper, we propose a TE-enabled Internet architecture to achieve this, which includes traffic engineering extensions to the current intra-domain and inter-domain routing protocols. Our work extends the egress router selection problem presented in [6] by considering end-to-end bandwidth guarantees. We propose a heuristic algorithm to solve the

BGERS problem. In addition to this, we propose five egress router selection policies to address the optimization objectives that we consider. By evaluating the behavior of those egress router selection policies through simulation, we can provide an answer to the fundamental question of how an egress router is selected in order to give the best performance with respect to which optimization objective. This evaluation gives ISP insight into the relation of egress router selection policy and network resource utilization, based on which they can configure their networks in order to realize their target objectives. To the best of our knowledge, this work is the first attempt at inter-domain traffic engineering using BGP policies to control inter-domain traffic flows with end-to-end bandwidth guarantees.

The rest of the paper is organized as follows. In section 2 we present a TE-enabled Internet architecture. In section 3 we formulate the BGERS problem and propose a heuristic algorithm with five egress router selection policies to solve it. Section 4 presents the evaluation of those egress router selection policies through simulation. Finally, we conclude our work and discuss future research directions in section 5.

2 Traffic Engineering Enabled Internet Architecture

To address the BGERS problem, the TE-enabled Internet architecture requires that the current intra-domain and inter-domain routing protocols are able to convey TE information such as bandwidth. We assume Traffic Engineering extensions to OSPF (OSPF-TE) [7] as the intra-domain link-state protocol, which disseminates bandwidth information within a domain. Moreover, we assume using Constrained Shortest Path First (CSPF) [8] with unit link cost to calculate a bandwidth constrained path between an ingress router and an egress router in a domain. An explicit path is then configured for the selected path along with bandwidth reservation. This can be done by RSVP extensions for MPLS TE [8].

On the other hand, the lack of TE information support in the current BGP hinders the deployment of BGERS. To make up this deficiency, it is necessary to record bandwidth information in the BGP UPDATE message, which represents the ability of a neighboring domain to provide the route with such available bandwidth. In [4], a new attribute, the QoS_NLRI, is proposed for this purpose. We assume that bandwidth information, which takes a single value, is conveyed through a similar attribute and we call the extended BGP the Traffic Engineering extensions to BGP (BGP-TE). We simply extend BGP for the purpose of bandwidth advertisement and TE but do not use bandwidth information on the usual BGP path selection.

The bandwidth information advertised by BGP-TE is guaranteed by a Service Level Agreement (SLA) established between neighboring domains in a management time-scale. Each domain is configured to make sure that sufficient bandwidth is provisioned for other domains, conforming to the established SLAs. The BGERS is one of the issues in this bandwidth provisioning. The outcome of bandwidth provisioning is Bandwidth Capability (*BC*), which is the bandwidth that has been allocated to a path between an ingress router and an egress router within a domain towards certain destination prefixes. From the ISP point of view, establishing SLAs with neighboring domains can extend its services to customers by reaching more destination prefixes

that belong to the remote domains with bandwidth guarantees. Technically, this service extension is done by bandwidth capability binding: A domain binds its bandwidth capability to the bandwidth information advertised by the neighboring domains and forms an extended *BC* (*eBC*) as the basis for agreeing new SLAs with its customers. Bandwidth capability binding is unidirectional and is done by a simple algebraic method. The *eBC* or *BC* is advertised as bandwidth information to neighboring domains through BGP-TE.

We give a small example of bandwidth capability binding in Figure 1. The example also shows how ISP 1 provides bandwidth guarantees to the traffic flows of customer 1 destined to customer 2, conforming to the customer SLA established with customer 1. We denote by *BCX* the unidirectional bandwidth capability of ISP *X* towards destination prefixes that belong to customer 2. We assume that ISP 1 has established a provider SLA with ISP 2 for bandwidth guarantees to customer 2, thus ISP 2 will advertise *BC2* to ISP 1 through BGP-TE. When ISP 1 receives *BC2*, it performs bandwidth provisioning based on the customer SLA and then binds *BC1* to *BC2*. This binding forms a unidirectional *eBC*. The value of *eBC* is equal to the minimum of *BC1* and *BC2*. ISP 1 can then provide *eBC* bandwidth guarantees to the customer 1's traffic flows destined to customer 2, conforming to the customer SLA.

For a large scale Internet to provide bandwidth guarantees between edge domains, ISPs have to collaborate and provide bandwidth guarantees to other domains' traffic flows. The concatenation of SLAs between domains can ensure end-to-end bandwidth guarantees for end customers. In this context, the bandwidth information advertised by BGP-TE is the unidirectional cascaded effect of bandwidth capability binding between each two domains along a BGP path. It is also the concatenated bandwidth that is guaranteed starting from the downstream domain (i.e. the one which advertises the bandwidth information) until the destination domain. Each domain uses bandwidth information, provided by BGP-TE and OSPF-TE, to optimize its network resource utilization by selecting appropriate egress routers for inter-domain traffic flows with bandwidth guarantees.

3 Bandwidth Guaranteed Egress Router Selection

With the TE-enabled Internet architecture, we can provision bandwidth guarantees for inter-domain traffic flows. In this section, we formulate the BGERS problem and propose a heuristic algorithm with several egress router selection policies to solve it.

Figure 2 shows a general TE-enabled Internet architecture. The ISP domain has a set of border routers as well as a set of intra-domain and inter-domain links. An inter-domain link connects between border routers of the ISP and the downstream domain. Each border router may connect to multiple inter-domain links. We assume that the ISP has established SLAs with its downstream domains for bandwidth guarantees and border routers in the ISP and downstream domains support BGP-TE. Through BGP-TE, the border routers of the ISP receive route advertisements of destination prefixes associated with the bandwidth information from the downstream domains.

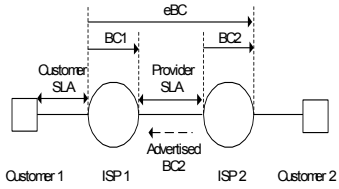

Fig. 1. BC binding

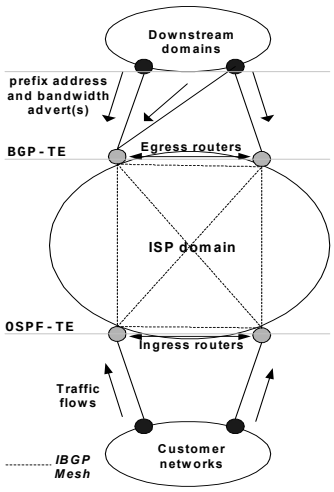

Fig. 2. A general TE-enabled Internet architecture

Symbol	Description
E	A set of intra-domain links
K	A set of destination prefixes
I	A set of ingress routers
J	A set of egress routers
$t(i,k)$	The aggregated bandwidth requirement of customer traffic flows destined to destination prefix $k \in K$ at ingress router $i \in I$
$Out(k)$	A set of egress routers that can reach destination prefix k
$NEXT_j$	A set of next hop addresses (addresses of border routers in downstream domains) that is connected to egress router $j \in J$
$f_k(j,n)$	True (1) / False (0); whether the prefix k can be reached through the inter-domain link between the egress router j and the next hop address $n \in Next_j$
C_{intra}^{l}	The capacity of intra-domain link $l \in E$
bw_{intra}^{l}	The current availability (unallocated bandwidth) on C_{intra}^{l}
$C_{inter}^{j,n}$	The capacity of the inter-domain link which is attached to egress router j and is connected to next-hop address n
$bw_{inter}^{j,n}$	The current availability (unallocated bandwidth) on $C_{inter}^{j,n}$
$p(k,j)$	The bandwidth advertised by BGP-TE on the egress router j to the destination prefix k after the BGP path selection
$bp(k,j)$	The current availability (unallocated bandwidth) on $p(k,j)$
$x_{(i,k)}^{j}$	True (1) / False (0); whether the customer traffic flow $t(i,k)$ has been assigned to the egress router j
$y_{(i,k)}^{l}$	True (1) / False (0); whether the customer traffic flow $t(i,k)$ has consumed bandwidth on the intra-domain link l
$d(i,j,k)$	Number of hops of the feasible shortest path (found by CSPF) between the ingress router i and the egress router j for $t(i,k)$
$bw_{intra}^{p:i,j}$	The bottleneck bandwidth of intra-domain path p between the ingress router i and the egress router j, i.e. $\min_{l \in p}\left(bw_{intra}^{l}\right)$

Fig. 3. Notations

Each border router that connects to the downstream domains selects the best route for each destination prefix based on the usual BGP decision process and then distributes the route to other border routers within the domain through Internal BGP (IBGP) mesh between border routers. It is possible that a border router receives multiple route advertisement with a common destination prefix through IBGP. Thus, an opportunity emerges to select the best among the appropriate egress routers for inter-domain traffic flows with bandwidth guarantees. The outcome of the BGERS can be realized by BGP policies such as using policy routing and manipulating BGP attributes. For the ISP's decisions on advertising bandwidth information to upstream domains such as how much bandwidth is advertised and where to advertise, we consider this as the subject of inbound inter-domain TE, which is out scope of this paper.

We assume that the traffic matrix of customer traffic flows is known through customer SLAs established in a management time-scale. Each customer traffic flow includes a bandwidth requirement to a destination prefix and the ingress router where it enters the ISP domain. Moreover, individual customer traffic flows are aggregated at each ingress router based on their destination prefixes. In the rest of this paper, we

refer customer traffic flow as the one which is aggregated (including bandwidth) from those individual flows destined to the same destination prefix at the ingress router.

3.1 Problem Formulation

We formulate the BGERS as an integer programming problem. Figure 3 shows the notations that are used in the rest of this paper. A solution of BGERS should compute a set of allocation between traffic flows and egress routers, which yields the best value for one or more objective functions. In this paper, we consider three:

1) *Minimizing the total bandwidth consumption:* The objective of minimizing the total bandwidth consumption in the network can be translated to the problem of minimizing the total number of hops that a traffic flow must traverse, i.e.

$$Minimize \sum_{k\in K}\sum_{i\in I}\sum_{j\in Out(k)} x^{j}_{(i,k)} \cdot d(i,j,k) \cdot t(i,k) \quad \textbf{(1)}$$

2) *Improving intra-domain load balancing:* The objective of improving intra-domain load balancing can be approximated with the problem of minimizing the maximum link utilization within the network, i.e.

$$Minimize \quad Max_{l\in E}\left(1-\frac{bw^{l}_{intra}}{c^{l}_{intra}}\right) \quad \textbf{(2)}$$

3) *Improving inter-domain load balancing:* The objective of improving inter-domain load balancing can be approximated with the problem of minimizing the maximum link utilization among all the inter-domain links, i.e.

$$Minimize \quad Max_{j\in J, n\in NEXT_j}\left(1-\frac{bw^{j,n}_{inter}}{c^{j,n}_{inter}}\right) \quad \textbf{(3)}$$

The BGERS is subject to the following constraints:

$$\sum_{k\in K}\sum_{i\in I} x^{j}_{(i,k)} \cdot t(i,k) \cdot f_k(j,n) \le c^{j,n}_{inter} \qquad \forall (j,n) \ \textit{where } j\in J \ \& \ n\in NEXT_j \quad \textbf{(4)}$$

$$\sum_{k\in K}\sum_{i\in I} y^{l}_{(i,k)} \cdot t(i,k) \le c^{l}_{intra} \qquad \forall l \in E \quad \textbf{(5)}$$

$$\sum_{i\in I} x^{j}_{(i,k)} \cdot t(i,k) \le p(k,j) \qquad \forall (k,j) \ \textit{where } k\in K, j\in J \quad \textbf{(6)}$$

$$x^{j}_{(i,k)}, y^{l}_{(i,k)} \in \{0,1\} \quad \textbf{(7)}$$

$$\sum_{j\in Out(k)} x^{j}_{(i,k)} = 1 \qquad \forall (i,k) \ \textit{where } i\in I, k\in K \quad \textbf{(8)}$$

Satisfy customer bandwidth requirement $\forall t(i,k)$ **(9)**

Constraints (4)-(6) are the capacity constraint respectively on each inter-domain link, intra-domain link and advertised bandwidth towards certain destination prefixes; constraint (7) ensures that the discrete variables to assume binary values; constraint (8) ensures that only one egress router is selected for each customer traffic flow. Con-

straint (9) enforces that there is sufficient end-to-end bandwidth to accommodate each assigned customer traffic flow. We define the following criteria to determine whether the bandwidth requirement of customer traffic flow, $t(i,k)$, is satisfied:

1. There exists a feasible path p from the ingress router $i \in I$ to the selected egress router $j \in J$ such that $bw_{intra}^{p,i,j} \geq t(i,k)$
2. $bp(k,j) \geq t(i,k)$
3. $bw_{inter}^{j,n} \geq t(i,k)$ *where n = FindInterdomainLink(j,k)*

The function *FindInterdomainLink(j,k)* returns the next hop address *n* (i.e. the address of border router in the downstream domain) to which a specific inter-domain link at the egress router *j* is connected to reach the destination prefix *k*. The first criterion is an intra-domain bandwidth constraint, which ensures that a feasible path exists between the ingress router and the selected egress router, and the bottleneck bandwidth of the path is no less than the bandwidth requirement of the customer traffic flow. The second criterion ensures that the advertised bandwidth at the selected egress router is sufficient to accommodate the customer traffic flow. This implies that in each domain along the corresponding BGP path towards the destination prefix there is sufficient bandwidth for the customer traffic flow. The third criterion ensures that the inter-domain link, connected between the selected egress router and the downstream domain reaching the destination prefix, has sufficient bandwidth for the customer traffic flow. If all the above criteria are met, the constraint on customer bandwidth requirement is satisfied.

The objectives of load balancing defined by (2) and (3) are orthogonal to each other. However, the objective of minimizing total bandwidth consumption defined by (1) and load balancing may lead to contradictory solutions. Based on this reasoning, we do not consider optimizing both objectives simultaneously. Instead, we study the implication of our egress router selection policies on each of these objectives.

The work in [6] has formulated the egress router selection problem as Generalized Assignment Problem (GAP) [9] and proved that the problem is NP-complete. Due to the fact our work extends that of [6] and has the additional constraints (5) and (6) as an intra-domain and a cascaded inter-domain capacity constraint respectively, we consider the BGERS problem to be a variant of GAP, which is also NP-complete. Hence, we propose a heuristic algorithm to solve it.

3.2 Heuristic Algorithm

We propose a greedy-based heuristic algorithm, namely Greedy-cost heuristic, to solve the BGERS problem. The Greedy-cost heuristic takes the following steps:

Step 1: Sort customer traffic flows in descending order based on their bandwidth requirements and selects one, $t(i,k)$, at a time in that order. This sorting can be assumed that large customer traffic flows have higher priority to be considered.

Step 2: Identify all the feasible egress routers for the customer traffic flow. The egress router *j* is feasible if it meets two criteria: (1) $j \in Out(k)$, and (2) Satisfy the bandwidth requirement of the customer traffic flow; i.e. the constraint (9). The first criterion ensures that the egress router *j* has a route to reach the destination prefix *k*

while the second criterion ensures that by selecting egress router *j* there is sufficient end-to-end bandwidth to provide bandwidth guarantees to the customer traffic flow.

Step 3: Compute the cost metric of selecting each feasible egress router. The cost metric encompasses two types of network information: (1) topology which includes the available bandwidth of each link, the number of hops and the bottleneck bandwidth on the path from the ingress to the egress router within a domain, and (2) resource at the egress router which includes the available bandwidth and the capacity of inter-domain links to which is connected. Among a set of feasible egress routers, the one that yields the best cost, determined by the adopted egress router selection policy which is discussed in section 3.4, is selected for the customer traffic flow.

Step 4: Once the egress router is selected, an explicit path is configured for the corresponding selected intra-domain path and the requested bandwidth is reserved on the explicit path, the corresponding selected inter-domain link and the advertised bandwidth capability for the assigned customer traffic flow.

Step 5: We consider the next customer traffic flow and repeat step 2 to step 5. The heuristic finishes when all the customer traffic flows have been considered.

3.3 Egress Router Selection Policies

With reference to step 3 of the Greedy-cost heuristic, we propose five egress router selection policies that use increasingly more network information to make a decision on egress router selection.

1) *Random-egress*: select an egress router randomly. We consider this policy as the behavior of the current BGP for the BGERS. The current non-TE BGP will select an egress router with respect to bandwidth information completely at random.

2) *Closest-egress-first*: selects the egress router which is the closest, in terms of number of hops, to the ingress router where the customer traffic flow enters the ISP domain. If there are several such egress routers, the selection tiebreak is in order of the maximum bottleneck bandwidth on the intra-domain path and the maximum available bandwidth on the inter-domain link.

3) *Widest-egress-first*: selects the egress router which has the maximum bottleneck bandwidth on the path leading from the ingress to the egress router. If there are several such egress routers, the selection tiebreak is in order of the path with the least number of hops and the maximum available bandwidth on the inter-domain link.

4) *Highest-availability-egress*: selects the egress router with which the selected inter-domain link, after being assigned the customer traffic flow, has the highest bandwidth availability. The bandwidth availability is the ratio of the available bandwidth to the capacity of a link.

5) *Shortest-dist-egress*: This policy uses the network distance between the ingress and the egress router as the selection parameter. The scope of network distance covers the intra-domain path between the ingress and the egress router, and the inter-domain link associated with the egress router. We define two distance functions to quantify the available bandwidth on each intra-domain and inter-domain link, and then compute the network distance. The distance of intra-domain link *l* is approximated by (10) while the distance of inter-domain link that is selected for the customer

traffic flow at the egress router *j* is approximated by (11). The parameter α in the distance functions represents the degree to which the greatest available bandwidth is favored over the least one. The network distance of selecting egress router *j* is defined by (12) where *l* is the link belongs to the selected intra-domain path *p* between the ingress router and the egress router *j*. The shortest-dist-egress selects the egress router with the minimum network distance. In the subsequent simulations, we use the notation *Shortest-dist-egress(*α*)* to denote this policy.

$$Dist_{intra}(l) = \frac{1}{\left[bw_{intra}^{l} - t(i,k)\right]^{\alpha}} \tag{10}$$

$$Dist_{inter}(j,n) = \frac{1}{\left[bw_{inter}^{j,n} - t(i,k)\right]^{\alpha}} \tag{11}$$

$$NetworkDist(j) = \sum_{l \in p} Dist_{intra}(l) + Dist_{inter}(j,n) \tag{12}$$

4 Performance Evaluation

We evaluate the Greedy-cost heuristic with the proposed egress router selection policies through simulation. Simulation results are based on 100-node transit domain topologies. The topologies are randomly generated by the Waxman's method [10]. The set of ingress and egress routers are disjoint. We set the number of ingress routers to 30, whereas the number of egress routers is a variable, as we will evaluate some effects by changing its value between 10 and 30. Each egress router is attached with a maximum of two inter-domain links. We assume that the inter-domain capacity is less than the intra-domain capacity, so the bottleneck resides at the former. The capacity of each intra-domain link is randomly generated between 400 and 500 while the capacity of each inter-domain link is randomly generated between 250 and 300.

Feamster [11] discovered that a typical default-free routing table may contain routes for more than 90,000 prefixes, but only a small fraction of prefixes are responsible for a large fraction of the traffic. Thus, we consider 1000 destination prefixes that belong to remote domains. As these prefixes are usually popular destinations, we assume that each egress router can reach all of them. This set of prefixes is randomly distributed on the inter-domain link(s) of each egress router. Each destination prefix is advertised with available bandwidth randomly generated between 200 and 250. For each customer traffic flow, the destination prefix and the ingress router are randomly generated and its bandwidth requirement is randomly generated between 10 and 40.

We evaluate the performance of the five egress router selection policies with respect to the three objective functions that we consider. Our evaluation consists of four scenarios. The first three scenarios evaluate the effects of selection policies on the total bandwidth consumption, intra-domain and inter-domain load balancing respectively. The last scenario studies the impact of α in the shortest-dist-egress on those objective functions. For the shortest-dist-egress, we use α=1.0 as the reference value.

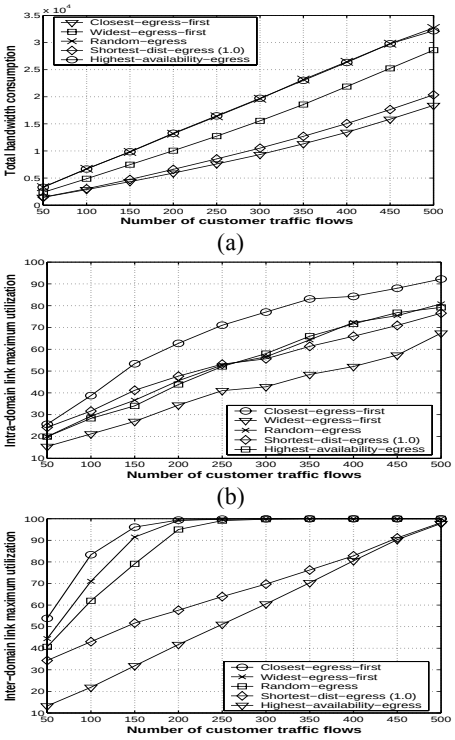

Fig. 4. Effects of egress router selection policies

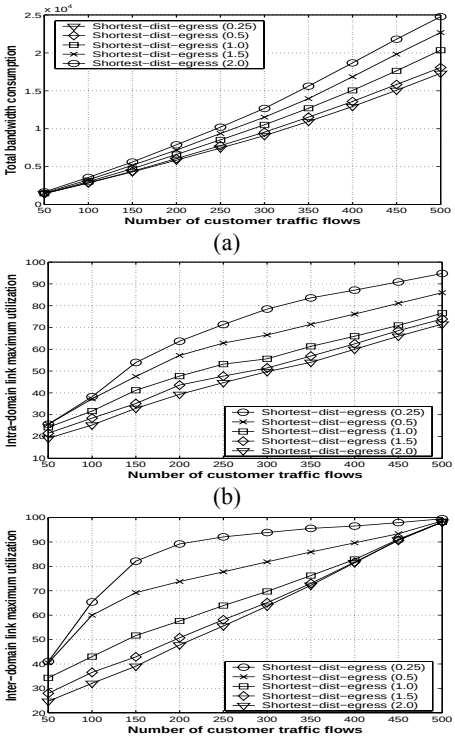

Fig. 5. Impact of α

Figure 4(a) shows the total bandwidth consumption as a function of the total number of customer traffic flows under the five egress router selection policies. The closest-egress-first consumes the least total bandwidth because customer traffic flows have always been directed to the nearest egress routers, which attempts to minimize bandwidth consumption by using least-hop paths. The other selection policies consume more bandwidth because they take load balancing into consideration, which may result in using longer paths. Specifically, the selection policies such as the highest-availability-egress and the random-egress, which do not entirely consider the network distance to make the selection decision, consume relatively much more bandwidth. While the shortest-dist-egress jointly considers number of hops and path bandwidth availability, it consumes slightly more bandwidth than the closest-egress-first but better load balancing is achieved as shown in the subsequent simulation.

Figure 4(b) shows the performance of intra-domain load balancing under the five egress router selection policies. The closest-egress-first exhibits the worst performance. This is because customer traffic flows are always statically directed to the nearest egress routers using a small subset of intra-domain shortest paths, thus causing the links on these paths to be heavily loaded. The widest-egress-first performs the best in intra-domain load balancing because customer traffic flows are always distributed to the paths with the maximum bottleneck bandwidth. This minimizes the maximum

utilization over all the links in the network. We find that an indirect way to achieve intra-domain load balancing is to evenly distribute traffic flows among all the egress routers. The random-egress and the highest-availability-egress are the examples. The rationale is that the more egress routers are evenly selected, the more different intra-domain paths are used to reach those egress routers. In general, this facilitates intra-domain load balancing. Since the shortest-dist-egress directly takes bandwidth availability as a factor to select egress routers, it performs better than all the other selection policies except the widest-egress-first especially under a heavily loaded network. Compared to the widest-egress-first, the shortest-dist-egress consumes much less bandwidth as this has been shown in the previous simulation.

Figure 4(c) shows the performance of inter-domain load balancing under the five egress router selection policies. Highest-availability-egress performs the best because customer traffic flows are distributed on the egress router with which the inter-domain link has the highest bandwidth availability. This minimizes the maximum utilization over all the inter-domain links. Selection policies such as the closest-egress-first, the widest-egress-first and the random-egress, which do not directly consider inter-domain link bandwidth availability to select egress routers, may direct traffic flows to a small subset of egress routers causing the inter-domain links at these egress routers to be heavily loaded while the other inter-domain links are barely used. This effect is more apparent for the closest-egress-first. As the shortest-dist-egress partly considers the inter-domain link bandwidth availability to select egress routers, it can achieve fairly good performance on inter-domain load balancing.

From the description of the shortest-dist-egress, α is an important selection parameter representing the degree to which the greatest available bandwidth is favored over the least one. The purpose of α is to balance the impact of path hops and path bandwidth availability on making the selection. This allows a selective adjustment between the objectives of minimizing the total bandwidth consumption and improving both the intra-domain and inter-domain load balancing. We investigate the impacts of α with respect to the three objective functions by adjusting its value. The simulation results are shown in Figure 5.

The graphs show that when α is small, the significance of path bandwidth availability to make egress router selection is reduced. As a result, the closest egress router is preferred and those heavily loaded intra-domain and inter-domain links may get higher chance to be selected. Thus, using small α would attempt to reduce bandwidth consumption at the expense of poor intra-domain and inter-domain load balancing. On the contrary, when α is big, the significance of path bandwidth availability to make the selection is amplified. This results in giving more emphasis on load balancing. In Figure 5, simulation results show that using large α would achieve better intra- and inter-domain load balancing at the expense of higher total bandwidth consumption. In summary, for the objective of minimizing the total bandwidth consumption and improving load balancing, small and big α are respectively used.

We summarize the implication of each egress router selection policy on the objectives that we consider as follows. The closest-egress-first, the widest-egress-first and the highest-availability-egress have the best performance with respect to the objective of minimizing the total bandwidth consumption, improving intra-domain and inter-domain load balancing respectively. The shortest-dist-egress has fairly good perform-

ance with respect to all the objectives. Specifically, it allows a selective adjustment between the objective of minimizing the total bandwidth consumption and improving load balancing by adjusting the parameter α. As minimizing the total bandwidth consumption and improving load balancing are contradictory, there is no policy that can achieve the best performance with respect to both objectives simultaneously.

5 Conclusions

In this paper, we presented the Bandwidth Guaranteed Egress Router Selection (BGERS) problem and solutions in the context of a TE-enabled Internet architecture. The architecture comprises traffic engineering extensions to the current intra-domain and inter-domain routing protocols. The objective of BGRES is that, for each customer traffic flow, select an egress router that satisfies the customer end-to-end bandwidth requirement while optimizing the network resource utilization by which we consider three objective functions: minimizing the total bandwidth consumption, improving intra-domain and inter-domain load balancing. We have developed a heuristic algorithm to solve the problem and studied the implication of egress router selection policies on the objective functions. This study benefits ISPs by choosing the egress router selection policy that fits their target objectives. As future work, we plan to extend the BGERS problem and solutions to accommodate the other QoS metrics, such as delay, with classes of services under the Differentiated Services network.

References

1. Awduche, D., Chiu, A., Elwalid, A., Widjaja, I., Xiao, X.: Overview and Prcinciples of Internet Traffic Engineering. RFC 3272 (2002)
2. Apostolopoulos, G., Williams, D., Kamat, S., Guerin, R., Orda A., Przygienda, T.: QoS Routing Mechanisms and OSPF Extensions. RFC 2676 (1999)
3. Bonaventure, O.: Using BGP to distribute flexible QoS information. Internet Draft <draft-bonaventure-bgp-qos-00.txt> (2001)
4. Cristallo, G., Jacquenet, C.: Providing Quality of Service Indication by the BGP-4 Protocol: the QoS_NLRI attribute. Internet Draft <draft-jacquenet-qos-nlri-04.txt> (2002)
5. Xiao, L., Lui, K.S., Wang, J., Nahrstedt, K.: QoS Extension to BGP. In the Proceedings of the 10th IEEE International Conference on Network Protocols. (2002) 100-109
6. Bressoud, T.C., Rastogi, R., Smith, M.A.: Optimal Configuration for BGP Route Selection. In the Proceedings of IEEE INFOCOM' 2003. (2003) 916-926
7. Katz, D., Kompella, K., Yeung, D.: Traffic Engineering Extensions to OSPF Version 2. RFC 3630 (2003)
8. Osborne E., Simha A.: Traffic Engineering with MPLS. Cisco Press (2002)
9. Martello, S.,Toth, P.: Knapsack Problems: Algorithms and Computer Implementations. John Wiley and Sons, New York (1990)
10. Waxman, B.M.: Routing of multipoint connections. IEEE Journal on Selected Areas in Communications. 6 (1988) 1617-1622
11. Feamster, N., Borkenhagen, J., Rexford, J.: Controlling the impact of BGP policy changes on IP traffic. AT&T Labs – Research, Technical Report HA173000-011106-02TM (2001)