A Self-Adaptive QoS Routing Framework for Multihomed Stub Autonomous Systems

M. Yannuzzi, A. Fonte, X. Masip-Bruin, E. Monteiro, S. Sánchez-López, and J. Domingo-Pascual

***Abstract*—Greedily managing interdomain traffic at the edge of the Internet is becoming a common practice in order to improve end-to-end QoS. However, the stability implications of such practices under massive utilization are completely unknown. Given that global stability is a must for the current and future Internet, self-adaptive tools will become unavoidable if masses of completely autonomous and uncoordinated stub Autonomous Systems (ASes) are allowed to simultaneously change their traffic patterns seeking only for the best of their own purposes in short, and even very short timescales. As a first step in this direction, we propose a novel and incremental self-adaptive interdomain QoS Routing (QoSR) algorithm, which helps BGP improving end-to-end QoS in a selfish but self-controlled manner. Our first results show that our algorithm not only improves end-to-end QoS, but also enhances overall throughput, while timely limiting the number of AS path shifts needed to accomplish these goals.**

***Index Terms*—Interdomain QoS Routing, BGP, Self-Adaptive, Stability**

I. INTRODUCTION

It is widely accepted that the main problem with end-to-end QoS provisioning is on the very foundations of the current interdomain network paradigm. This paradigm is based on a highly scalable and completely distributed network architecture, which relies on the Border Gateway protocol (BGP) as the glue that keeps the Internet together [1]. The central issue is that BGP has not inbuilt QoS capabilities given that it was designed with very different goals in mind by the early nineties.

Although some researchers have proposed to replace BGP, in practice, only incremental approaches are realistic and will have chance to become deployed. From this perspective, most of the interdomain heuristics proposed mostly tended to add QoS and Traffic Engineering (TE) extensions to BGP [2-3], but quite recently some research groups and manufacturers have started to avoid new enhancements to the protocol and proposed to decouple part of these tasks from BGP devices [4-7].

While heuristics extending BGP are only able to improve end-to-end performance for internets under low routing dynamics, the latter result much more effective, especially, when routing changes occur more frequently. The main difference between these two approaches is that the latter decouples part of the policy control portion of the routing process from BGP devices. Hence, the two approaches basically differ in how policies are controlled and signaled.

In-band QoS Routing (QoSR) and TE techniques, that is, those inherently supported by BGP can feasibly operate over long timescales which means they are appropriate for static or pseudo-static QoSR and TE provisioning. On the other hand, out-of-band techniques, that is, those decoupled from BGP are in fact able to operate in much shorter timescales so they result perfectly appropriate for dynamic or even highly dynamic QoSR and TE provisioning. However, the stability implications of rearranging interdomain traffic in very short time scales is not yet understood. Indeed, the effect of managing large amounts of interdomain traffic in this way is completely unpredictable. Thus, these kinds of solutions are definitively not applicable, for example, to large transit Autonomous Systems (ASes) such as Tier-1 or Tier-2 Internet Service Providers (ISPs). Additionally, the rearrangement of small fractions of interdomain traffic in short timescales, but magnified by the number of sources simultaneously injecting these perturbations to the network may also result unpredictable in terms of global stability.

Above all, multihomed stub ASes are those which could benefit the most from novel mechanisms providing them with dynamic QoSR and/or TE capabilities in medium or short timescales. This particular fraction of ASes crowds together mostly medium and large Enterprise Customers, Content Service Providers (CSPs), and small Network Service Providers (NSPs), which altogether actually represent more than 60% of the total number of ASes in the Internet.

Therefore, the blast of multihomed stub ASes in the last few years has gained huge interest in both research and commercial fields, and that's why several optimized edge routing proposals are starting to appear as commercial products. From our perspective, multihoming in combination with routing edge optimizers is a powerful tool in which stub ASes can rely on in order to improve their end-to-end QoS [8]. In this sense we foster this kind of interdomain TE approach for the most widely deployed AS in the Internet, but highlighting that the set of perturbations introduced by these ASes to the network should be timely controlled.

The rest of the paper is organized as follows. In Section 2 we illustrate the main motivations behind the design of self-adaptive tools in order to manage interdomain traffic at the edge of the network. We then present the design of our self-adaptive QoSR algorithm in Section 3, while Section 4

Manuscript received February 11, 2005. This work was partially funded by the MCyT (Spanish Ministry of Science and Technology) under contract FEDER-TIC2002-04531-C04-02, the CIRIT (Catalan Research Council) under contract 2001-SGR00226, the European Commission through Network of Excellence E-NEXT under contract FP6-506869, and the SATIN Grant "New Interdomain QoS Routing Algorithms based on a Distributed Overlay Entities Architecture".

M. Yannuzzi, X. Masip-Bruin, S. Sánchez-López, and J. Domingo-Pascual are with the Department of Computer Architecture and the Advanced Broadband Communications Laboratory (CCABA), Universitat Politècnica de Catalunya (UPC), Jordi Girona 1-3, 08034 Barcelona, Spain (e-mails: yannuzzi@ac.upc.edu; xmasip@ac.upc.edu; sergio@ac.upc.edu; jordid@ac.upc.edu).

A. Fonte and E. Monteiro are with the Laboratory of Communications and Telematics, CISUC-DEI, University of Coimbra, Pólo II, Pinhal de Marrocos, Postal Address 3030-290 Coimbra, Portugal. A. Fonte is also with the Polytechnic Institute of Castelo Branco, Av. Pedro Álvares Cabral, nº12, Postal Address 6000-084, Castelo Branco, Portugal (e-mails: afonte@dei.uc.pt; edmundo@dei.uc.pt)

shows our simulation results. Lastly, Section 5 concludes the paper and highlights our futures directions.

II. MOTIVATIONS

Two major motivations drive the design of self-adaptive interdomain tools at the edge of the network. First of all, it is widely known that network administrators of multihomed stub ASes are not willing to adopt complex mechanisms to control how their traffic is managed in medium and short timescales. Moreover, they do not want to get into the details of how and when their traffic should be rearranged. They simply want to take plain decisions, and they expect that such decisions last in time. This fact reveals that NSPs, CSPs, and medium/large Enterprise Customers are eagerly claiming for straightforward mechanisms that allow them to opportunistically manage their interdomain traffic in short timescales depending on the existing end-to-end performance. Thus, a major advantage behind self-adaptive mechanisms is that they are perfectly suitable for this kind of opportunistic and selfish demands.

The next figure depicts the significance of self-adaptive mechanisms from this perspective, given that they are able to hide the QoS dynamics from the traffic reallocation decision process. This approach supplies an appealing solution for multihomed stub ASes seeking for an opportunistic interdomain QoSR or TE mechanism, since they may decide how conservative or opportunistic they want to be by simply selecting a fixed threshold for example, and without worrying about the stability implications of their decision. It should become clear that how conservative or opportunistic such ASes will actually be strongly depends on the QoS dynamics, and so this may vary over time. In contrast, the use of self-adaptive mechanisms in combination with the selection of fixed thresholds allows these ASes to straightforwardly decide how opportunistic they are willing to be, and this decision will last in time.

In second place, the significance of a self-adaptive tool resides in its strengths in terms of guaranteeing local and global stability. Under highly unpredictable network conditions, such as link flaps, or routing misconfigurations it is imperative that each edge optimizer counts with a self-adaptive mechanism which allows it to learn from this dynamics and diminish or prevent the number of AS path shifts until the network conditions are stable once again. Indeed, multihomed stub ASes not using self-adaptive mechanisms may find that the number of traffic reallocations they are actually allowing may be much higher than the expected. In other words, under highly aggressive network dynamics even a conservative opportunistic approach may lead to network instability caused by an excessive number of path shifts. Thus, the assertion of being "conservative" strongly depends on the QoS dynamics. As an alternative, self-adaptive algorithms/metrics are able to adapt themselves to those changing conditions, so that they could be able to reflect the choices made an ISP independently of the QoS dynamics.

Furthermore, it should become clear that these self-adaptive tools will indeed become suitable in the near future not only to reallocate particular flows of interdomain traffic, but also to reallocate interdomain tunnels, or interdomain IP/MPLS Label Switch Paths (LSPs) in shorter timescales.

In the next section we present our first design of a self-adaptive interdomain QoSR algorithm, which was developed for an edge routing framework that we developed in a previous work [5].

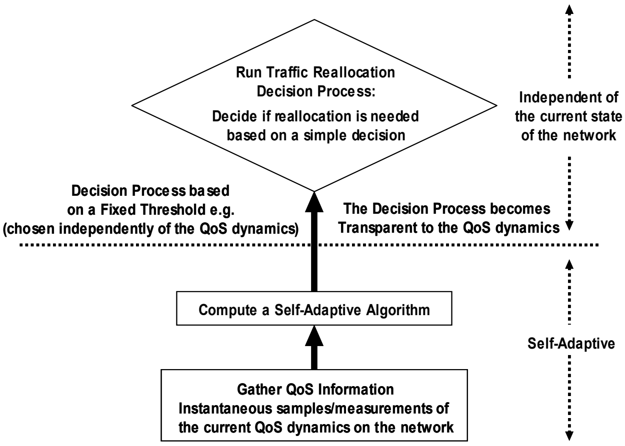

Fig. 1. A Self-Adaptive algorithm provides transparency to the traffic reallocation decision process

III. SELF-ADAPTIVE EDGE ROUTING

A. The Edge Routing Framework

In [5] we proposed a distributed architecture in which a pair of Overlay Entities (OEs) within two non-peering multihomed stub ASes were able to exchange soft Service Level Agreements (SLAs) regarding the traffic among them, examine the compliance with those SLAs, and accurately configure on-the-fly BGP to bypass network problems such as link failures, or service degradation for a given set of Classes of Service (CoSs). The foremost motivation for influencing traffic in this way is that with only a very small number of OEs, but located at strategically selected remote multihomed stub ASes is enough to control a significant part of the traffic of an AS [3]. A major advantage of this framework is that no OEs are needed in any transit AS connecting the remote ASes in our model. Thus, the complexity of dynamic QoS provisioning is pushed to the edge of the network by means of a completely distributed architecture. It is worth mentioning that large ISPs and a huge fraction of the Internet community eagerly support the idea of pushing the complexity of QoS provisioning to the edge of the network, releasing the core Internet switches/routers from these tasks.

Our approach is that an OE within a source AS dynamically manages the allocation of its outbound traffic towards a remote AS in our model, depending on the network conditions and QoS constraints for each CoS. This allows tweaking BGP even in very short timescales given that no BGP messages will be ever generated.

Furthermore, based on the recommendations given in [9, 10] the OEs are endowed with a mechanism to spawn probes targeting the reduced set of remote ASes in the architecture using a Pseudo-Random Poisson Process. In this process N_u random sampling times uniformly distributed are generated over consecutive intervals of duration T_u. The parameters N_u and T_u are selected so that $N_u / T_u = \lambda$ where λ^{-1} is the average sampling time of a classical Poisson process. We chose this approach to avoid the occasional lengthy spaces between sampling times that a Poisson process may create, since this could be unacceptable for many real-time applications.

In this paper we mainly propose to enhance two aspects of our routing framework. On the one hand, we propose to turn the QoS cost metric used by the OEs into a self-adaptive metric. This metric is based on One-Way Delay (OWD) [9]. On the other hand, the QoSR algorithm proposed in [5] was a reactive algorithm, which means that each OE triggers the reallocation of outbound traffic for a class j from its AS only when a violation to the SLA for that class is detected. Then, the algorithm prevents shifting traffic whenever the SLA is fulfilled, even though an alternative path with a better end-to-end cost may exist. Conversely, in this paper we propose a proactive or opportunistic routing algorithm so that the OEs will be able to take full advantage of the metric and reallocate traffic of class j each time an end-to-end path with a sufficiently lower cost exists. Any opportunistic algorithm imposes a trade-off in terms of the frequency of traffic reallocations, since an excessively conservative approach may under-utilize network resources, whereas frequently shifting traffic may lead to network instability. In order to cope with this problem our opportunistic routing algorithm feeds from the self-adaptive metric becoming also a self-adaptive routing algorithm.

B. *Self-Adaptive Cost Metric*

Equation (1) presents the self-adaptive metric to be used by each OE in our QoSR model. In terms of notation, M_{ij} represents the cost to reach a distant OE through the i_{th} egress link of the source AS for traffic of class j.

$$M_{ij} = \begin{cases} Floor\left(\alpha_j S_{ij}\right) & if \quad S_{ij} \le \overline{D_j} \\ \infty & if \quad S_{ij} > \overline{D_j} \end{cases} \tag{1}$$

With the aim of avoiding frequent changes of the metric, instead of using instantaneous values of OWD we introduce S_{ij}, which corresponds to a Smoothed OWD (SOWD). This smoothing process as well as the non-negative weight α_j will be essential to endow the metric with self-adaptive capabilities. This will become clear in the rest of this section. In addition, the bound $\overline{D_j}$ represents the maximum OWD tolerable to reach a remote AS for traffic of class j. This parameter is specified in the SLA exchanged between the OEs using the OE's protocol.

Equation (2) presents our former metric, that is, the one we used in [5], which corresponds to the average OWD through a sliding window of size W_j. The index n_{ij} in (2) simply represents the sequence number of the instantaneous samples of OWD.

$$\overline{OWD_{ij}}\left(n_{ij}\right) = \frac{1}{W_j}\left(\sum_{k=n_{ij}-W_j+1}^{k=n_{ij}} OWD_{ij}(k)\right) \quad \forall\, i, j \tag{2}$$

S_{ij} corresponds to a smoothing process applied to (2), and it is based on the introduction of a one-dimensional grid which provides some granularity to the samples of OWD in such a way that it can be exploited by an opportunistic QoSR algorithm. Our goal is to design this grid with self-adaptive capabilities, particularly depending on the current QoS conditions in the network.

To accomplish this goal, the interval [0, $\overline{D_j}$] is initially divided in N_{ij} subintervals, i.e.:

$$\left[m\frac{\overline{D_j}}{N_{ij}}, (m+1)\frac{\overline{D_j}}{N_{ij}} \right] \quad m \in Z \,/\, 0 \le m \le \left(N_{ij} - 1\right) \tag{3}$$

defining a grid $\forall\, i, j$. In order to design this grid we define the following parameters using the first W_j instantaneous samples of OWD_{ij}.

$$\begin{cases} \overline{K_{ij}} = \max\left\{OWD_{ij}(k)\right\} & \forall\, k = 1,\ldots,W_j \\ \underline{K_{ij}} = \min\left\{OWD_{ij}(k)\right\} & \forall\, k = 1,\ldots,W_j \end{cases} \tag{4}$$

Then, the interval $[\underline{K_{ij}}, \overline{K_{ij}}]$ defines our first estimation of the range of variation of the instantaneous samples of OWD_{ij}. Our aim is to prevent frequent variations in the metric, so the main idea behind the grid is that moderate variations of the samples given in (2) generate the same numerical value of S_{ij}, and thus the same cost M_{ij}. Our first criterion in the design of this grid is that the maximum variation $\left(\overline{K_{ij}} - \underline{K_{ij}}\right)$ fits into one subinterval of the grid. Moreover, we introduce an adjustable coefficient $\Delta_j \in R \,/\, \Delta_j \ge 1 \;\forall\, j$, which assures at least a percentage of separation between the grid lines and the parameters defined in (4) given by $(\Delta_j - 1)10^2$. In this sense Δ_j basically reflects the degree of conservativeness while defining the initial grid. In addition, Δ_j will also play a fundamental role when adding self-adaptive capabilities to the grid. Fig. 2a) shows the first W_j samples of OWD_{ij} and illustrates the grid design approach. Accordingly, N_{ij} is bounded by:

$$\left.\begin{aligned} \frac{m\overline{D_j}}{N_{ij}} &\le \Delta_j^{-1}\underline{K_{ij}} \\ \frac{(m+1)\overline{D_j}}{N_{ij}} &\ge \Delta_j \overline{K_{ij}} \end{aligned}\right\} \Rightarrow N_{ij} \le \frac{\overline{D_j}}{(\Delta_j \overline{K_{ij}} - \Delta_j^{-1}\underline{K_{ij}})} \tag{5}$$

$$N_{ij} \in Z \,/\, N_{ij} \ge 1$$

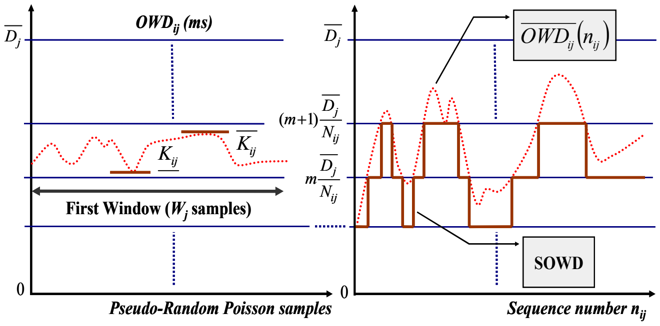

Fig. 2.a) The grid design

Fig. 2.b) SOWD S_{ij}

In order to provide a scalable design we impose the following restriction: for each CoS j, and $\forall\ i,\ k\ /\ i \neq k \ \Rightarrow N_{ij} = N_{kj}$. This is a reasonable decision since comparing costs M_{ij} and M_{kj} $\forall\ i \neq k$, only will make sense if the same grid is used for traffic of class j over every egress link from the source AS. Thus, we define: $N_j \equiv N_{ij}\ \forall\ i$.

Clearly, the advantages of this discrete arrangement may be lessened if the granularity is enough to cause that the opportunistic QoSR impels an OE to frequently re-configure the border BGP routers of its AS. Thus a trade-off exists in terms of the granularity of the grid, and how proactively the traffic will be reallocated.

Following a conservative approach, our second criterion is to use the $\max\limits_i [\Delta_j^{-1} \underline{K_{ij}}, \Delta_j \overline{K_{ij}}]$ when determining N_j. Thus:

$$N_j = Floor\left(\frac{\overline{D_j}}{\min\left\{ \overline{D_j}, \max\limits_i \left(\Delta_j \overline{K_{ij}} - \Delta_j^{-1} \underline{K_{ij}} \right) \right\}} \right) \qquad \textbf{(6)}$$

where (6) satisfies the restriction in (5) and provides a common grid over every egress link i for traffic of class j. Then, if we define $G_j = \frac{\overline{D_j}}{N_j}$ as the step of the grid for the j_{th} CoS, the S_{ij} to be used in (1) is defined as follows:

$$S_{ij} = \begin{cases} G_j Floor\left(\dfrac{\overline{OWD_{ij}}(n_{ij})}{G_j} \right) & N_j > 1 \\ \overline{OWD_{ij}}(n_{ij}) & N_j = 1 \end{cases} \qquad (7)$$

Equation (7) is the most general expression for S_{ij}, and anticipates an interesting feature of our approach. That is, in case the network dynamics are so aggressive in terms of OWD that no granularity is present on the grid for a given CoS j, i.e. $N_j = 1$, then the OE switches its behavior from proactive to reactive, computing the former metric and routing algorithm that we presented in [5], until some granularity is present once again on the grid. Fig. 2b) depicts the relation between $\overline{OWD_{ij}}$ and S_{ij}. As aforementioned, the development of a self-adaptive cost metric in terms of S_{ij}, demands that an OE should be able to dynamically adapt this grid depending on the current QoS conditions. Our approach is to avoid frequent recalculations of the grid during unstable network conditions, so we propose that each time a new grid is computed this is maintained for a superset of several windows S_{W_j} (several W_j). Then, we trigger the recalculation of the grid whenever:

$$\begin{cases} G_j^{(n)} < \max\limits_i \left(\overline{K_{ij}}^{(\mathrm{n})} - \underline{K_{ij}}^{(n)} \right) & \vee \\ \min\limits_i \left(\Delta_j \overline{K_{ij}}^{(\mathrm{n})} - \Delta_j^{-1} \underline{K_{ij}}^{(n)} \right) < \overline{K_j}^{(\mathrm{n})} - \underline{K_j}^{(n)} \end{cases} \qquad \textbf{(8)}$$

where the supra-index (n) represents the current grid, and:

$$\begin{cases} \overline{K_{ij}}^{(\mathrm{n})} = \max\left\{ OWD_{ij}(k) \right\} & \forall\ k \in S_{W_j}^{(n)} \\ \underline{K_{ij}}^{(n)} = \min\left\{ OWD_{ij}(k) \right\} & \forall\ k \in S_{W_j}^{(n)} \end{cases} \qquad \textbf{(9)}$$

with $\left[\underline{K_j}^{(n)}, \overline{K_j}^{(n)} \right] = \max\limits_i \left[\underline{K_{ij}}^{(n-1)}, \overline{K_{ij}}^{(\mathrm{n\text{-}1})} \right]$.

Then, a new grid (n+1) is obtained when the substitution of $\max\limits_i \left(\Delta_j \overline{K_{ij}}^{(n)} - \Delta_j^{-1} \underline{K_{ij}}^{(n)} \right)$ in (6) supplies a new $N_j^{(n+1)}$ / $N_j^{(n+1)} \neq N_j^{(n)}$. The first inequality in (8) reflects that the network conditions have become rather unsteady so the step of the grid $G_j^{(n)}$ needs to be increased, while the second inequality indicates that the conditions have become even steadier so the step of the grid could be diminished. It is possible that while an egress link i satisfies the first inequality in (8) for a given CoS j, another egress link k satisfies the second one. In such a case our decision is to follow a conservative approach so we choose to increase the step of the grid.

Finally, if the grid was not recomputed throughout $S_{W_j}^{(n)}$ instead of setting the current grid for a whole new superset, the OEs began to search for either of the conditions in (8) through a sliding window of size S_{W_j}. This mechanism will speed up the reaction of our self-adaptive cost metric when network conditions have changed.

The natural next step in the design of our metric is to link the weight α_j in (1) with the self-adaptive features of the previous grid. Moreover, this needs to be done in such a way that the metric and the routing algorithm using this metric could bring transparency to the traffic reallocation decision process as it was shown in Fig. 1. Given that the self-adaptive part of the grid is indeed its step, α_j should explicitly depend on the step G_j. The next proposition provides a reasonable criterion in order to choose an α_j that could supply transparency to the traffic reallocation decision process.

Proposition 1: If S_{ij} increases by one step of the grid, the cost M_{ij} increases at least by a fixed value Q and with maximum sensitivity if:

$$\alpha_j = \frac{Q+1}{G_j} \qquad \textbf{(10)}$$

Proof. Using (1):

$$\alpha_j \left(S_{ij} + G_j \right) - a_2 = \alpha_j S_{ij} - a_1 + Q \qquad \textbf{(11)}$$

where the parameters $a_h \in [0,1)\ h = 1,2$ come from the Floor function, hence:

$$\alpha_j = \frac{a_2 - a_1 + Q}{G_j} \ \Rightarrow\ \frac{Q-1}{G_j} < \alpha_j < \frac{Q+1}{G_j} \qquad \textbf{(12)}$$

Next, we choose the upper bound to get the maximum sensitivity. Therefore, α_j explicitly depends on the resolution of the grid G_j, and hence it evolves with it.

Corollary 1: Given that Q is fixed, the variations of M_{ij} in terms of S_{ij} are independent of the grid.

Proof. Using (10), the variation in M_{ij} when S_{ij} increases one step of the grid is:

$$\left(M_{ij}^{new} - M_{ij}^{old} \right) = (Q+1) + a_1 - a_2 \qquad \textbf{(13)}$$

Then, the variations of M_{ij} are independent of the QoS dynamics. In other words, while the resolution of the grid evolves in time according to those QoS dynamics, we adapt the calculus of M_{ij} depending on the state of the grid so that the opportunistic QoSR algorithm remains transparent to them. The advantage of a self-adaptive α_j like the one proposed in (10) is that it allows network managers seeking for an opportunistic approach, to transparently opt for a degree of conservativeness which is independent of the QoS conditions in the network, and this decision will last in time. This decision will be done by simply configuring a fixed threshold within the routing algorithm, so that a network manager does not need to get into the details of the metric nor worry about the stability implications of this decision.

C. An Opportunistic and Self-Adaptive QoSR Algorithm

The interdomain QoSR algorithm presented in this section uses the cost metric M_{ij} in an opportunistic way, so it triggers the reallocation of traffic even when no violations to the SLAs occur. Then, under this opportunistic routing algorithm, traffic of class j may be reallocated from the egress link i to link k if and only if their cost metrics satisfy:

$$\left(\frac{M_{ij} - M_{kj}}{Q+1} \right) \geq R_j^{th} \qquad \textbf{(14)}$$

Inequality (14) introduces the fixed threshold R_j^{th} that a network manager of a multihomed stub AS will need to choose according to its degree of conservativeness for each CoS. The motivation behind this selection is that with this criterion the threshold basically counts the number of steps that S_{ij} needs to increase in order to reallocate traffic of class j. The next piece of pseudo-code briefly summarizes the operation of our opportunistic QoSR algorithm:

```
        Opportunistic Interdomain
             QoSR Algorithm

if there exists a set of egress links $f$
which satisfy:
   $(M_{ij}-R_j^{th}(Q+1)) \geq M_{fj}$
   trigger traffic reallocation of class
   $j$ from $i$ to $k$ / $M_{kj} = \min\{M_{fj}\}, \forall f \neq i$
end.
```

Fig. 3. Pseudo-code summarizing the operation of the opportunistic QoSR algorithm

Finally, two additional advantages in our approach are that, first the parameters such as Q, Δ_j, N_u, T_u, or SW_j can be set to default values so they do not need to be explicitly configured by the network manager. Secondly, each OE could independently decide which QoSR algorithm to use depending on the network conditions. In other words, even within the same OE-to-OE connection, while one of the OEs shows a reactive behavior the other could behave proactive.

IV. EVALUATION RESULTS

Our simulations were performed using J-Sim [11] with the BGP Infonet suite [12] in which we have implemented all the functionalities of our edge routing architecture. As in [5] we have introduced a set of QoS extensions to BGP (QBGP) in the Infonet suite for pseudo-static QoS provisioning. This was done with the aim of contrasting the timescales in which in-band solutions (such as QBGP) and out-of-band solutions (such as our edge routing model) are able to operate and to manage interdomain traffic according to the SLAs established. For our simulations, and in order to compare with our previous results we used the same topology presented in [5] which is part of the GÉANT European Academic Backbone [13]. For complexity concerns, we modeled each AS as a single QBGP router with core DiffServ capabilities configured to support four different classes of traffic, namely EF, AF11, AF21 and Best-Effort or background traffic. To complete the scenario, on the domain where traffic was injected we used edge DiffServ capabilities to mark packets with a specific DSCP (DiffServ Code Point) depending on its corresponding CoS. These marks were applied both to regular IP packets, and to the probes generated by the OEs.

The maximum OWD tolerated per-CoS were heuristically chosen to allow the OEs to take advantage of alternative paths. During the evaluation we heuristically configured the following set of parameters: *i)* $\Delta_j = 1.05$; *ii)* the sliding window parameters $W_j = 10$, $SW_j = 5W_j \; \forall j$; *iii)* the degree of conservativeness $R_j^{th} = 1.0 \; \forall j$. In both routing algorithms the Pseudo-Random Sampling parameters were aggressively set to $N_u = 3$ and $T_u = 4$s, $\forall$ CoS j, and the traffic model used was Poisson for all CoSs.

The set of simulations presented here were carried for two different groups of values of the parameter Q, namely, Q=20 for Group 1, and Q=30 for Group 2. As performance indicators we used the number of AS path shifts needed to meet the SLA's constraints for each CoS, as well as a traffic transfer efficiency parameter $\rho_{nj} = C_{nj} / C_{sj}$. This last indicator assesses the traffic performance for each CoS, where C_{nj} is the throughput at a given destination n, and C_{sj} is the throughput at the source s for CoS j.

Fig. 4 shows the number of AS path shifts needed for each CoS, and for each of the groups described before. The average number of AS path shifts for each of the algorithms is also shown as Av to allow a global view of the algorithm dynamics. The comparison of the algorithms shows that while QBGP is unable to react to the SLA violations, the Proactive algorithm has, on average a much smaller number of AS path shifts. This fact is especially noticeable for the EF class, for which no path shifts were needed in the Proactive case confirming the contribution of this algorithm to the stability of the most important traffic class. The main reason for this lies in the opportunistic and self-adaptive capabilities of the cost metric used by the Proactive algorithm, given that the algorithm is able to react before congestion actually occurs. Fig. 4 also shows the efficiency ρ_{nj} of the algorithms under analysis.

The efficiency of EF traffic is not affected by the routing algorithm used and so it is independent of the number of AS path shifts. However, the impact of the reduction of the number of AS path shifts affects the end-to-end delay of EF traffic, which is smaller for the Proactive algorithm (not shown in this figure). The overall efficiency of QBGP is worse than the efficiency of the two proposed algorithms. Another important difference between QBGP and the proposed algorithms lies in the Best-Effort (BE) class treatment. The efficiency for this background traffic is highly improved by our routing algorithms due to the starvation avoidance mechanism we use, which is out of the scope of this paper. The Proactive algorithm has a better efficiency than the Reactive algorithm, yet much fewer path shifts were necessary to achieve this superior performance.

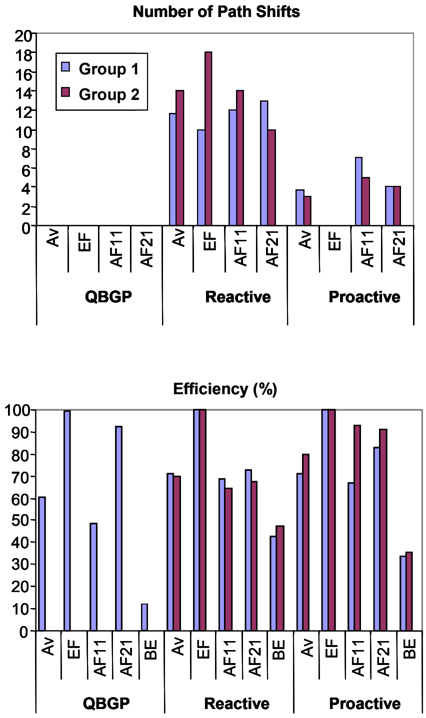

Fig. 4. Evaluation Results

V. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper we have surveyed the strengths of self-adaptive routing tools at the edge of the network, and we have also emphasized that this kind of tools will become unavoidable in the near future. As a first step in this direction, we introduced in this paper an opportunistic and self-adaptive interdomain QoSR algorithm supported by an innovative self-adaptive QoS cost metric based on OWD. This QoSR algorithm allows network managers of multihomed stub ASes to decide when and how to rearrange part of their interdomain traffic, in a straightforward and flexible way, by simply configuring fixed thresholds which are completely independent of the QoS dynamics. Indeed, the adaptability features of the metric leverages the opportunistic approach even in short timescales without worrying the network managers about the stability implications of their decisions. Simulation results show that the routing algorithm was able to limit the number of AS path shifts, and improve overall throughput efficiency while complying with the corresponding SLAs. We have also tested our algorithm under highly stressful network conditions, and even in such cases the number of path shifts was timely controlled by our OEs.

Our goals at this step are, firstly to extend the design of self-adaptive QoS cost metrics and self-adaptive QoSR and TE techniques for multihomed stub ASes, which will help BGP improving end-to-end QoS in a selfish, but globally stable manner. Secondly, we are working on the development of a stability model for the proposed architecture. We are applying physical similes (such as thermodynamic models for open systems) aiming at counting with an energy model which could be treated by means of the Lyapunov criterion. Finally, we have plans to test the response of the algorithms under different topologies, and different traffic loads. In fact, we plan to survey the scalability of our proposal and evaluate its effectiveness in an experimental test-bed.

REFERENCES

[1] Y. Rekhter, T. Li, “A Border Gateway Protocol 4 (BGP-4),” Internet Engineering Task Force, Request for Comments 1771, March 1995.
[2] Cristallo, G., C. Jacquenet, “An Approach to Interdomain Traffic Engineering,” in Proceedings of XVIII World Telecommunications Congress (WTC2002), France, September 2002.
[3] B. Quoitin, S. Uhlig, C. Pelsser, L. Swinnen and O. Bonaventure, “Interdomain Traffic Engineering with BGP,” IEEE Communications Magazine, May 2003.
[4] S. Agarwal, C. Chuah, R. Katz “OPCA: Robust interdomain policy routing and traffic control,” IEEE Openarch, April 2003.
[5] M. Yannuzzi, A. Fonte, X. Masip-Bruin, E. Monteiro, S. Sánchez-López, M. Curado, J. Domingo-Pascual, “A proposal for interdomain QoS routing based on distributed overlay entities and QBGP,” in Proceedings of the First International Workshop on QoSR (WoQoSR), co-located with QoFIS’04, LNCS 3266, Barcelona, Spain, October 2004.
[6] Internap Network Services Corp., Flow Control Platform, http://www.internap.com/
[7] Cisco Systems, Inc., Cisco Optimized Edge Routing, http://www.cisco.com/
[8] A. Akella, et al., “A Comparison of Overlay Routing and Multihoming Route Control,” in Proceedings of ACM SIGCOMM04, Portland, USA, August 2004.
[9] G. Almes, S. Kalidindi, M. Zekauskas, “A one-way delay metric for IPPM,” Internet Engineering Task Force, Request for Comments 2679, September 1999.
[10] S. Shalunov, B. Teitelbaum, “One-way active measurement protocol (OWAMP) requirements,” Internet Engineering Task Force, Request for Comments 3763, April 2004.
[11] J-Sim Homepage, http://www.j-sim.org.
[12] Infonet Suite Homepage, http://www.info.ucl.ac.be/~bqu/jsim/
[13] GÉANT Website, http://www.dante.net/server/show/nav.007