# **Towards Inter-Domain QoS Control**

Rui Prior<sup>1</sup> and Susana Sargento<sup>2</sup> <sup>1</sup>LIACC – University of Porto, <sup>2</sup>IT – University of Aveiro rprior@ncc.up.pt, ssargento@det.ua.pt

### Abstract

This paper describes a solution for inter-domain QoS to be part of the framework for end-to-end QoS control. The solution is based on the use of virtualtrunk type aggregates, defined by means of Service Level Agreements, for data transport between peering domains. These aggregates are defined for several "well-known" traffic classes, supporting different types of applications. We define an extension to the BGP routing protocol to transport three different QoS metrics (light load delay, assigned bandwidth and a congestion alarm), and a path selection algorithm using a combination of these metrics. Simulation results demonstrate the use of our solution, showing that it yields much better QoS results than those achieved with standard BGP or BGP with the QoS NLRI extension, since it is able to efficiently avoid congested paths, reducing delays and packet losses. Moreover, these results are achieved without significantly affecting route stability.

## 1. Introduction

Today's commercial Internet has far surpassed the initial concept of a data transport network mostly for academic and military purposes. Its multi-service nature becomes evident with the ever increasing number of online audio and video broadcasting services, podcasting, internet telephony and instant messaging with audio and video extensions, among others. Its success in replacing the traditional networks for multimedia services with real-time requirements and in providing new services is, however, conditioned by the ability to provide a high and guaranteed quality level. The introduction of Quality of Service (QoS) support in the Internet is, therefore, of major importance towards this goal, and the combined use of QoS routing, scalable QoS control mechanisms such as DiffServ [1] and admission control the means to achieve it.

Internet routing is performed in two layers: intradomain routing, within networks controlled by a single

operator or administrative entity, and inter-domain routing, for the interconnection and transport service between different administrative domains. Though much attention has been paid to QoS in packet switching networks in general and IP networks in particular, most of the effort has been centered on intra-domain OoS. Several different factors contribute to the difficulty of solving the inter-domain QoS problem. The Internet is a complex entity, comprised of Autonomous Systems (AS) managed by very diverse operators. If it is to be widely deployed, an inter-domain QoS routing mechanism must be capable of handling this heterogeneity while imposing minimum requirements on intradomain routing in order to be appealing to the different operators. The introduction of QoS routing should not disrupt the currently existing inter-domain routing: both should interoperate, allowing for incremental deployment, and the stability of intra-domain routing should not be significantly affected by the QoS mechanisms. Conciliating the required stability with the dynamic nature of QoS information is a major challenge in interdomain QoS routing. A related issue is scalability: a solution that does not scale to the dimension of the Internet cannot be deployed widely enough to be useful.

An appropriate choice of the path selection metrics is key to the behavior of QoS routing. Two fundamentally distinct types of metrics may be used for path selection: static or dynamic. Static metrics, such as the bottleneck link capacity, do not vary over time, meaning that the chosen routes are stable. However, they are unable to express the dynamic character of the networks, and therefore QoS routing based solely on static metrics is unable to prevent and/or react to network congestion. By using dynamic metrics, such as the available path bandwidth or the time window averaged instantaneous packet delay, it is possible to react to congestion. However, a high overhead is incurred in message exchange for updated metrics and in path recomputation, and the route stability property is lost.

To reduce the overhead associated with QoS routing with dynamic metrics, two main approaches [2] may be used: trigger QoS information updates only in case of significant changes in the metrics or convey longerterm information on path characteristics through statistical characterization of the route metrics. The first approach provides coarser-grained information in terms of scale, and the second in terms of time.

In a commercial Internet, data transport is a service subject to contracts between the interested parties. Network operators establish Service Level Agreements (SLA) between them specifying the conditions for data transport, and these agreements are translated to path selection policies in the routing protocol, frequently configured by hand. The ability of inter-domain routing to automatically adjust to these agreements should not be overlooked.

In this paper we propose an SLA-aware solution for inter-domain QoS routing based on both static and coarse-grained dynamic metrics. It uses light load delay and assigned bandwidth (both static) in order to improve the packet QoS and make better use of network resources, and a coarse-grained dynamic metric for path congestion to avoid overloaded paths. We define an extension to the Border Gateway Protocol (BGP) [3] to transport these OoS metrics and a path selection algorithm using a combination of them. Simulation results demonstrate the use of our solution, showing that it yields much better OoS results than those achieved with standard BGP or BGP with the QoS NLRI (Network Layer Reachability Information) extension [4], since congested paths are efficiently avoided, reducing delays and packet losses. These results are achieved without significantly affecting route stability.

The paper is organized as follows. The next section describes related work. Section 3 addresses the way to include this inter-domain QoS proposal into end-to-end QoS control approaches. Section 4 presents the proposed protocol and the associated path selection algorithm. In section 5 we discuss the results of simulations for protocol validation and analysis. Finally, section 6 concludes the paper.

## 2. Related work

The existing QoS inter-domain solutions can be basically divided into two categories: static and dynamic approaches. In the static approaches, cooperation between network providers is based on SLAs, which can specify the amount of resources available for each traffic class and the algorithms for QoS level mapping between domains. These solutions are relatively simple to achieve, but lack flexibility and usually result in inefficient resource utilization. The dynamic approaches are based on dynamic allocation of resources and dynamic QoS level mapping, depending on actual network traffic and other conditions. These solutions are more complex, but assure more efficient resource utilization.

The issue of inter-domain QoS was already addressed in several research projects, which specified and implemented frameworks for QoS reservation and signaling in inter-domain paths. As examples, the AQUILA approach [5] defined a set of "well-known services" for inter-domain QoS and used explicit signaling by the Border Gateway Reservation Protocol (BGRP) [6] protocol to reserve resources for flows; the TEQUILA approach [7] splitted the inter-domain QoS provisioning into QoS Service Level Specifications (SLS) management and QoS routing based on an extension of BGP to support the QoS NLRI attribute; the MESCAL [8] approach extended the TEQUILA one to provide different inter-domain QoS for different sets of services; finally, Internet2 OBone has developed the Simple Inter-domain Bandwidth Broker Signaling (SIBBS) [9] protocol for inter-domain signaling between bandwidth brokers. The common features of these solutions are the use of BGP extensions and SLSs between administrative domains. These approaches are either too complex to be used in the Internet [5] [8] [9], or ineffective in adapting to dynamic conditions and avoiding congested paths [7].

A framework for QoS-based Internet routing, adopting the traditional separation between intra- and interdomain routing, was defined by Crawley et al. [2]. They discussed the goals of an inter-domain QoS routing and the associated issues that must be addressed, and provided general guidelines that should be followed by any viable solution to QoS routing in the Internet. However, they did not specify the set of QoS metrics to be transported or the algorithms for using such metrics in the choice of inter-domain routes.

A series of statistical metrics for QoS information advertisement and routing, tailored for inter-domain QoS routing, though also applicable to intra-domain routing, were defined by Xiao et al. [10], along with algorithms to compute them along the path. These metrics, the Available Bandwidth Index (ABI), the Delay Index (DI), the Available Bandwidth Histogram (ABH) and the Delay Histogram (DH), convey information expressed in terms of one or more probabilistic intervals. Simulation results show that by using these metrics, selected routes are closer to optimality than when using static metrics, and the overhead is lower and the stability higher than when using the corresponding instantaneous metrics. However, these approaches consider only a single QoS parameter, making it difficult to simultaneously satisfy different requirements.

Cristallo and Jacquenet proposed an extension to the BGP with a new optional and transitive attribute,

QoS\_NLRI, for the transport of several types of QoS information [4]. This work is focused on the specification of the attribute, including the formats for transporting the different parameters and does not specify how the information is to be used in path selection. Some simulation results demonstrating its use with (static) information on one way packet delay are provided, though.

### **3. Inter-domain QoS control in E2E QoS**

In this section we describe the inter-domain scenario that we are aiming at for end-to-end QoS control in IP networks. This scenario is depicted in fig. 1 which illustrates a call between two users (which may be mobile, Mobile Terminal 1 and 2 - MT1 and MT2) connected to access networks belonging to two different administrative domains. We also consider that the MT domains are not directly connected but contain at least an inter-domain path (e.g., an international connection) between each other. Each administrative domain may contain several access networks, each of them capable of supporting different access technologies, and a core subdomain providing interconnection between the access networks and to other administrative domains.

To provide scalability the DiffServ framework [1] is used with different levels of resource control. In order to establish a reservation for a flow with fully end-toend QoS, QoS control at the different levels must be coordinated. Specifically, the admission control performed in the access needs to take into account the available resources in the end-to-end path, concerning the access, core and inter-domain links. In this paper we describe and evaluate the inter-domain QoS routing piece of the puzzle. Please refer to [11] for more information on the integration of resource control in the access and core networks.

## 4. Proposed protocol and algorithms

Our proposed approach to inter-domain QoS control is based on 3 main pieces: 1) a set of well-known traffic classes globally supported by all operators, 2) SLAs between adjacent domains, and 3) an inter-domain QoS routing protocol using a set of metrics for each of the well-known classes.

### 4.1. Class differentiation

The use of different traffic classes is important to simultaneously satisfy the requirements of applications with very different characteristics. The well-known classes are a small set of traffic class templates with



Fig. 1. Inter-domain call scenario

particular characteristics that should be globally supported by all network operators (e.g., a conversational class for small sized packets with very low delay and jitter). These classes have well-defined per-domain limits for QoS parameters other than delay, and limits for the complete path may be derived by combining the values of traversed domains. Since they are merely templates, operators have to map the well-known classes into specific DiffServ classes implemented in their own domains.

#### 4.2. SLAs and SLSs





In this proposal, the SLAs contain SLSs that specify a set of transport aggregates, each corresponding to a particular (*ingress point, egress point, service class*) triplet. These aggregates may be regarded as virtual trunks connecting two different domains across a third one directly connected to both. Fig. 2 illustrates the concept: an SLS between domain S1 and domain T1 specifies that X traffic may flow between S1 and domain T3 in a given traffic class; an SLS between domain T1 and domain T3 specifies that Y traffic may flow between T1 and domain D1 for the same class. Aggregates are managed internally within each (transit) domain, ensuring that enough resources are assigned.

#### 4.3. QoS routing

The proposed QoS routing protocol is an extension of BGP. Currently, version 4 of BGP has become the de facto standard for inter-domain routing in the Internet. BGP is a path vector protocol for the exchange of reachability information between connected ASs (External BGP) and to share information on learned routes among the different edge routers of a given domain (Internal BGP). Routes selected by BGP are propagated to the intra-domain routing protocol used in the AS, usually referred to as the Interior Gateway Protocol (IGP), by the edge routers. Reachability information is conveyed in UPDATE messages, each containing an advertisement of a new or changed route to a given network destination and/or a set of withdrawals of routes to destinations that may no longer be reached via the AS originating the UPDATE. The reception of an UP-DATE message triggers a three step decision process: (1) a degree of preference is assigned to the new route (if any) based on a set of policies; (2) one of the available routes to the destination is selected and propagated to the IGP; (3) if the route is different from the previously installed one, it is propagated to the peering ASs.

The most common policy for path selection is the minimum number of hops in the AS\_PATH. Though the AS\_PATH length bears only a very loose relation to QoS parameters, BGP can easily be extended to convey virtually any kind of relevant QoS information. The decision processes may also be changed to use the QoS information (if present) for path selection without breaking backward compatibility. We extended BGP to transport and use three QoS metrics for each of the well known traffic classes. These metrics are the assigned bandwidth, path delay under light load (both static) and a dynamic metric for path congestion, described below.

### 4.3.1. Metrics

SLA information is explicitly included using BGP to carry information on the amount of bandwidth contracted between two domains for data transport to a third one in each of the traffic classes. The assigned bandwidth, reflecting traffic contracts, is essentially static. It is updated along the path to be the minimum, that is, the bottleneck bandwidth (concave metric). Notice that our model does not require explicit and quantified agreements, only that transport operators assign a certain capacity for data transport between their connected peers; explicit SLAs are just a means to guarantee that reasonable assignments are performed.

Information on the expected delay in light load conditions is also carried. This information not only allows for the minimization of packet delays, but also for a more rational use of the network resources, since in high capacity links with significant length, such as today's inter-domain connections, it consists mostly on the sum of propagation delays, directly proportional to the traversed span of fiber, as long as there is no congestion. It also provides a lower bound for the expected packet delay. We argue that the minimization of this metric by the path selection mechanism will lead to a more rational use of the network resources, as well as better QoS to data packets, which would suffer smaller delay penalties<sup>1</sup>. The light load delay metric is static, and is summed along the path (additive metric). Contrary to the other metrics, only one delay value is transported by BGP, common to all classes.

The third QoS metric conveyed by our proposed extension is path congestion. The concept of congestion is deliberately vague and may, therefore, be translated into a coarse objective metric, minimizing the inherent overhead in message exchange and path recomputation typical of dynamic metrics. In our case, the congestion alarm is expressed by an integer with three possible values, with the following meaning: 0 - not congested; 1 - very lightly congested; 2 - congested. This metric is updated along the path to the maximum value (convex metric). In the most basic version, congestion may be inferred from the utilization of the aggregates; in a more advanced version, other parameters, such as packet loss, average length of traversed router queues or measured delay, may be used as inputs to the function that computes the alarm level of virtual trunk aggregates. The main requirement for the congestion alarms, the sole dynamic metric in our proposal, is that changes should be infrequent, for scalability and stability reasons; hysteresis and related techniques may be applied in the assignment of alarm levels to this end.

An effective value of the congestion alarm is used for path selection instead of the received one, aiming at reducing the fluctuations in the usage of the aggregates. The effective value is the same as the received value, unless the latter is 1 and the route is already in use, in which case the effective alarm is 0. This means that when level 1 (light congestion) is reached the route should not replace a previously established one, but domains that were already using the route should not switch to a different one unless a higher congestion level is reached. This behavior is meant to avoid the synchronized route flapping problem.

#### 4.3.2. Path selection algorithm

The three above mentioned QoS metrics are conveyed in UPDATE messages by a newly defined Path Attribute, QoS\_INFO, which is optional and transitive (meaning that ASs which do not yet support the extension simply forward the received value), and are updated by the BGP-speaking routers at each transit domain, taking into account the virtual trunks between the domain to which the route is advertised and the "next hop domain" for the route. The virtual trunks are shared among different source to destination routes: in fig. 2, for example, all traffic of a given class transported from T1 to D1 via T3 shares the T1:D1 virtual trunk for that class, independently of S1 or S2 origin.

<sup>&</sup>lt;sup>1</sup> If the links do not have these characteristics and the light load delay diverges significantly from the sum of propagation delays, the optimality regarding resource utilization will decrease, but the QoS optimization will remain intact.



Fig. 3 illustrates the propagation of the delay, bandwidth and congestion alarm metrics in the QoS INFO attribute of UPDATE messages. When the destination AS (AD2) first announces the route to an internal network, it may omit the QoS INFO attribute if this network is directly connected to the announced NEXT HOP. On receiving the UPDATE, the edge router at transit domain TD2 creates (or updates, if already present) the QoS INFO attribute with metrics of the outgoing link, for route selection purposes (this step is omitted in the figure). If the route is selected, it is propagated to all peering domains; the QoS INFO attribute sent to the different upstream domains is different, since the metrics are updated with respect to the virtual trunk aggregates. The same process is repeated at transit domain TD1. Notice that in the UPDATE sent from TD1 to AD1, the delay metric (19ms) is the sum of the delays of the concatenated virtual trunks (7ms and 12ms), the reserved bandwidth is the minimum along the path (300Mbps for the conversational class, for example) and the congestion alarm is the maximum (1 for the streaming class, for example). The virtual trunk values that contribute to the final values received by AD1 for this route are underlined in the figure.

Fig. 4 shows the algorithm for route comparison used in the decision processes in pseudo-code. Delay information is used to select the fastest/shortest route. The benefits of doing this, as previously mentioned, are twofold: packets will suffer lower delays and network resource usage will be more rational. The information on the reserved bandwidth is used to eliminate, from the set of possible choices, routes with insufficient bandwidth to support the current outgoing traffic aggregate from the local AS to the destination (including flows generated at the local AS and flows traversing it); it is also used as tie breaker when two routes for the same destination have the same announced delay. The alarm levels are used to eliminate congested routes from the set of possible choices. The elimination of routes with insufficient capacity from the set of possible choices prevents congestion of those routes to a

| set Traffic to dst = Local traffic to dst +                  |
|--|
| Transit traffic to dst                                       |
| for both routes  |
| if $Alarm_{rcv} = 1$ and route in use, set $Alarm_{eff} = 0$ |
| else set Alarm <sub>eff</sub> = Alarm <sub>rcv</sub>         |
| if both routes have Assigned BW < traffic to dst,            |
| choose the one with larger Assigned BW                       |
| else if one route has Assigned BW < traffic to dst,          |
| choose the other one   |
| else if Alarm <sub>eff</sub> is different,                   |
| choose the route with lower Alarm <sub>eff</sub>             |
| else if Delay is different,                                  |
| choose the route with least Delay                            |
| else if Assigned BW is different,                            |
| choose the route with larger Assigned BW                     |
| else use normal BGP rules (AS_PATH length, etc.)             |
| Fig. 4 Route comparison/selection function                   |
| ing. 4. Route companion/selection function                   |

certain degree, contributing to lower message and processing overhead and to route stability.

#### 5. Simulation results

In this section we present simulation results in ns-2 [12] of the QoS\_INFO proposal for inter-domain QoS routing, compared with standard BGP and to BGP with the QoS\_NLRI extension conveying static one-way delay information (the expected delay of the route in light load conditions). Note that the QoS\_NLRI extension can be used to convey QoS parameters other than delay: ref. [4] does not specify how that information is to be used for path selection; therefore, in this comparison we used the scenario therein illustrated.

Due to the extreme amount of traffic in inter-domain scenarios, we have chosen to simulate the signaling protocol at the packet level, but not the data traffic, which was mathematically simulated using the M/G/1 queuing model with three different packet sizes: 50% of packets with 40 bytes (4% of the traffic volume), simulating SYN, ACK, FIN and RST TCP segments; 20% of packets with 80 bytes, simulating packetized voice (3% of traffic volume); and 30% of packets with 1500 bytes, simulating full size TCP segments (93% of traffic volume). These packet sizes reflect the bimodality currently observed in Internet traffic [13], complemented with voice packets, whose frequency tends to increase. Queuing delays were obtained using the Po-

laczek-Khintchine formula, 
$$W_Q = \frac{\lambda E[S^2]}{2(1 - \lambda E[S])}$$
 where

 $W_Q$  is the queuing delay,  $\lambda$  is the traffic arrival rate and S is the service time [14], and computation of total packet delays was based on the Kleinrock independence approximation [14]. These simulations were performed using a single well-known class.

#### 5.1. Simulation scenario

To have meaningful results, a realistic topology and traffic matrix are required. We have used the topology

shown in fig. 5, with 26 nodes and 36 bidirectional links, where each city (node) simulates a different AS. The topology, link delay and traffic demand matrix are from the Portuguese backbone network of reference in [15]. The distribution of traffic demand values for the different routes is summarized in the figure, having a minimum of 1 Mbps, a maximum of 6.3 Gbps, an average of 88 Mbps and a standard deviation of 421 Mbps. Link bandwidth was assigned based on expected demands, and is shown, in bps, next to each link. The configuration of the virtual trunk type SLSs in our proposed model was performed automatically, based on the link bandwidth, the traffic matrix and a set of feasible routes (proportional distribution of link bandwidth). Not all triplets (a,b,c) such that a is connected to b and b to c have a corresponding SLS – for instance, there is no ('Lisboa', 'Ponta Delgada', 'Funchal') SLS, since all traffic going from 'Lisboa' to 'Funchal' is expected to use the direct link. For the sake of example, the following virtual trunks are established through 'São João da Madeira': ('Braga', 'São João da Madeira', 'Coimbra') with 9.7 Gbps; ('Coimbra', 'São João da Madeira', 'Braga') with 1.5 Gbps; and ('Coimbra', 'São João da Madeira', 'Porto') with 5.6 Gbps.



Fig. 5. Topology and CDF of traffic demand

Thresholds for setting alarm levels on virtual trunk usage were 50% of the SLS bandwidth for level 1 and 80% for level 2, except where otherwise stated. We ran simulations for 8200 simulated seconds, discarding data for the first 1000 seconds in order to filter out transient effects.

### 5.2. Link usage, route optimality and QoS

In the first experiment we compare the three interdomain routing mechanisms: standard BGP, BGP with



QoS\_NLRI and our proposed QoS\_INFO with respect to link usage, route optimality and QoS parameters.

Fig. 6 shows histograms with the distribution of the link offered traffic in the three approaches, averaged out of the 7200 useful simulation seconds. The overused class corresponds to links having an offered load above 100%, meaning that a significant portion of packets must be consistently discarded due to link capacity limitation. With the standard BGP, routes are normally chosen based on the lowest number of elements in the AS Path, disregarding path delay or congestion. For instance, the link 'Funchal' - 'Portimão' was used in many routes because the sub-path 'Lisboa' - 'Funchal' - 'Portimão' has a lower AS count than the much more sensible, in terms of both delay and bandwidth, 'Lisboa' - 'Setúbal' - 'Alcácer do Sal' - 'Sines' - 'Portimão' sub-path. As a result, 38% of the routes were suboptimal in terms of expected delay and 10% of the unidirectional links were overused, leading to an overall packet loss figure of 22%.

With the QoS\_NLRI BGP extension, results were much better, since routes are chosen in a more sensible way. However, since path congestion is not accounted for, 7% of the unidirectional links still had a utilization demand over 100%, leading to an overall packet loss of 10%. With our QoS\_INFO approach, traffic demand on all the links was always below their capacity: although overload may happen for short periods, the system reacts by changing the affected routes. As a result, there was no packet loss due to link overutilization.



Fig. 7. Percentage of routes with loss prob.  $\leq x$ 

Fig. 7 shows the packet loss probability CDF for the routes in the different scenarios. Once again, our proposed QoS\_INFO approach yields much better results, with 100% of the routes having a negligible packet loss probability, contrasting to only about 70% in QoS NLRI and 65% in the standard BGP.

Fig. 8 (left) shows CDFs of the summed propagation delays for the routes at the end of the simulation<sup>2</sup> in the three scenarios (in the cases of QoS NLRI and QoS INFO, they correspond to the announced delay values). As expected, QoS\_NLRI performs better in this respect, since the routes with the lower delay metric are always chosen, ignoring route congestion. The QoS INFO curve follows closely, but the standard BGP is clearly worse. It is worth noting that the light load expected delay holds little significance if routes are congested (heavily loaded); therefore, a much more meaningful parameter is the expected packet delay for the routes (sum of propagation and transmission delays with the expected queuing delays along the path), also plotted the figure (right). In this respect, our QoS INFO proposal did clearly better than the other two: about 30% of the routes in QoS NLRI and 35% in standard BGP had a delay larger than 0.5 seconds; in QoS INFO delays were always lower than 10 msec.



Fig. 8. Percentage of routes with delay  $\leq x$ 

#### 5.3. Signaling overhead and route stability

The drawback of the QoS\_INFO approach, as usual with dynamic QoS routing approaches, is increased signaling load and decreased route stability. The distribution of the frequency of sent and received updates is shown in fig. 9. With the other models, all routes are stable as long as there are no topology changes (due, e.g., to link failures). Regarding route stability, with the QoS\_INFO approach, 617 out of a total of 650 routes in the topology (ca. 95%) were stable, meaning that



Fig. 9. Distribution of UPDATE frequencies

they did not change during the useful simulation period; the other 5% did change, though with varying frequency. For example, 12 ASs sent less than 0.2 updates per second, whereas 2 ASs sent between 1.0 and 1.2 updates per second.

Since the choice of a new route is triggered by changes in the alarm levels, the SLS utilization thresholds used to assign a given alarm level have strong influence in the stability of the routes. In order to evaluate this influence, we performed a second experiment, with simulations using utilization values from 20% to 65% (x axis) of the bandwidth assigned to the SLSs as threshold for alarm level 1, and from 70% to 85% as threshold for alarm level 2 (different curves), whose results are shown in fig. 10. In the figure we may see that relatively low values of threshold for alarm level 1  $(th_1)$ tend to improve the route stability: indeed, 100% route stability was achieved in three cases with  $th_1$  at 25% and 30%. As  $th_1$  gets close to  $th_2$ , route stability decreases. Too high values for  $th_2$  also tend to reduce stability – the curve for  $th_2$ =85% is generally lower than the others. Since too low values for  $th_2$  tend to re-route traffic before it is really needed, a value of 80% for th2 seems to be a good choice.

Route stability is related to the frequency of update messages, since they are triggers for the BGP decision processes. Some reduction in the number of updates may be achieved by the introduction of hysteresis in the assignment of congestion alarm levels. We have introduced hysteresis by using two different values for each threshold – a change from an alarm level to a higher one occurs only when the high value is crossed, but in



Fig. 10. Route stability without hysteresis

<sup>&</sup>lt;sup>2</sup> Since routing with QoS\_INFO is based on dynamic information, routes do change in the course of the simulations; in standard BGP and BGP with QoS\_NLRI all routes are stable during the useful simulation period.



Fig. 11. Route stability with hysteresis

order to return to the lower alarm level, the utilization must drop below the lower value. Fig. 11 shows the stability of routes with different lower and higher levels for both thresholds. Compared to fig. 10, the results seem to be more predictable, and generally (though not always) better.

Fig. 12 shows the average number of updates per second per node without and with hysteresis. Similarly to route stability, results of the number of updates with hysteresis are not always better than without it, but they exhibit a more constant and predictable behavior.



Fig. 12. Frequency of BGP updates

## 5. Conclusions

This paper addressed the problem of inter-domain QoS routing. Our proposal is based on virtual trunk type aggregates, usually defined by SLAs, for the indirect transport of traffic between two different administrative domains across a third one. We proposed the QoS\_INFO extension to the BGP protocol using of a combination of three different metrics (assigned bandwidth, expected light load delay and congestion alarm) in order to simultaneously achieve different and conflicting goals: finding non-congested paths that satisfy the QoS requirements of the data flows, minimizing the network resources used to transport the flows, interdomain route stability and minimization of the message exchange and path computation overheads.

Simulations were performed to evaluate our proposal and to compare it to standard BGP and to the QoS\_NLRI BGP extension. The results show that, unlike the other cases, with our QoS\_INFO extension congested paths and their consequences on QoS are avoided. Although there is a penalty in overhead and route stability in doing this, most of the routes are stable, especially if the thresholds for alarm setting are appropriately selected. The introduction of hysteresis in the alarm level assignment slightly improves the route stability and makes it more predictable.

As future work we plan to improve the algorithm for assigning alarm levels to aggregates and to increase the scalability by introducing route aggregation, without compromising the accuracy of QoS information.

## 10. References

[1] S. Blake (ed.) et al., "An Architecture for Differentiated Services." IETF RFC 2475, Dec 1998.

[2] E. Crawley et al., "A Framework for QoS-based Routing in the Internet." IETF RFC 2386, Aug 1998.

[3] Y. Rekhter and T. Li (eds.), "A Border Gateway Protocol 4 (BGP-4)." IETF RFC 1771, Mar 1995.

[4] G. Cristallo and C. Jacquenet, "The BGP QOS\_NLRI Attribute." IETF Internet Draft, February 2004.

[5] M. Winter (ed.) et al., "Final System Specification." AQUILA Consortium Deliverable D1203 (v. b1), Feb 2003.

[6] P. Pan, E. Hahne and H. Schulzrinne: "BGRP: Sink-Tree-Based Aggregation for Inter-Domain Reservations." *Journal of Comm. and Networks, Vol. 2, No. 2*, Jun 2000,.

[7] T. Damilatis (ed.) et al., "D3.4: Final System Evaluation. Part B: D1.4: Final Architecture, Protocol and Algorithm Specification." TEQUILA Consort. Deliverable, Oct 2002.

[8] P. Flegkas (ed.) et al., "Specification of Business Models and a Functional Architecture for Inter-domain QoS Delivery." MESCAL Consort. Deliverable D1.1, Jun 2003.

[9] P. Chimento et al., "QBone Signalling Design Team – Final Report." Internet2 QoS Working Group, Jul 2002.

[10] L. Xiao et al., "Advertising Inter-Domain QoS Routing Information." *IEEE Journal on Selected Areas in Comm.*, *Vol. 22, No. 10*, Dec 2004.

[11] R. Prior, S. Sargento et al., "Providing End-to-End QoS in 4G Networks." *Proc. of the 3<sup>rd</sup> IASTED Intl. Conf. on Comm. and Computer Networks (CCN-2005)*, Nov 2005.

[12] The Network Simulator – ns-2, version 2.27. http://www.isi.edu/nsnam/ns/

[13] R. Sinha, C. Papadopoulos and J. Heidemann, "Internet Packet Size Distributions: Some Observations." Oct 2005.

http://netweb.usc.edu/~rsinha/pkt-sizes/

[14] D. Bertsekas and R. Gallager, "Data Networks," 2nd edition, Prentice-Hall, 1992.

[15] J. Pedro et al., "On a Portuguese Backbone Network of Reference." *Proceedings of the III Symposium on Enabling Optical Networks (SEON-2005)*, Jun 2005.