Abstract. Currently there is intensive research work attempting to obtain QoS provision on the Internet. Although existing proposals for the provision of QoS guarantees on Internet differ significantly, they share in most cases a common characteristic, that is, that all routers on the network are modified, thus allowing end-to-end QoS provision. Nevertheless, with this generalized requirement it is safe to say that QoS provision on the Internet can only become a reality in the long term. Therefore, taking this into account in this work we propose an alternative based on DiffServ in which it is only necessary to incorporate QoS provision mechanisms in a subset of routers within the network. In a more specific way, our model guarantees QoS provision whenever an operative path exists in which all intermediate routers incorporate certain mechanisms. In our proposal we locate an agent responsible for access control in each routing domain an agent responsible of access control and thus the management of the parameters related to QoS provision in each router within the network.

Keywords: QoS provision, Incremental QoS, Bandwidth Broker, QoS mapping, QoS requirements.

1 Introduction

From the IntServ specification it has been attempted to incorporate an architecture that allowed offering QoS on Internet. It is desirable that Internet could be used as a common infrastructure to support not only real-time but also non-real-time communication. A completely new infrastructure could be built for the real-time services, maintaining the Internet as it is at this moment. However, significant advantages like statistical sharing of resources among those real-time and non-real-time traffics would be lost, besides to become something more complex to build and to manage that a common infrastructure.

Nevertheless, current size of the Internet causes that any modification of its architecture to be able to offer QoS seems far from being a simple task. Each Autonomous System (AS) has its own hardware and software configuration and it could not always be possible to alter it to incorporate extensions to the protocols currently supported and/or configured. Diverse reasons may exist for which some organizations who, being responsible for one or several ASs, decide to migrate their architecture toward an architecture that can provide QoS, while others may postpone this decision. Without a complete implantation of the new architecture in the Internet, strict guarantees related to QoS obtaining could not be offered, although statistical guarantees under certain restrictions could be provided.

Our main premise is that the requirement of migrating completely the whole architecture of the Internet to be able to offer QoS guarantees is unacceptable. Proposals related to the Internet2 QBone architecture [12] are still far from arriving to end users, while obtaining QoS from the network is a current claim. In this sense, applications are ahead from the services that the Internet can offer.

In this work we propose an alternative based on DiffServ in which it only is necessary the incorporation of QoS provision mechanisms in a subset of routers of the network. More specifically, our model guarantees the provision of QoS requirements whenever an operative route exists in which all intermediate routers incorporate certain mechanisms for the QoS provision. In this proposal we locate at each routing domain an agent responsible of accomplishment of access control as well as of the management of those parameters related to QoS provision at routers of the network.

The architecture based on an agent to provide QoS over DiffServ is not new, since it was presented in [10]. In that work, Bandwidth Broker (BB) is denominated to the agent in charge of establishing in routers the necessary control policies for QoS provision. The BB establishes classification and policing functions at the edge routers, and policing functions at the core routers. It can also implement access control functions and can coordinate and automatically to negotiate QoS provision between BB of adjacent domains.

In our proposal we extend the functionality described for a BB to incorporate the following aspects: a) maintenance of an overlapped QoS topology to the best-effort one, b) it captures and processes the
IGP information, c) treatment of traffic statistics collected from the Edge Routers (ERs) for obtaining the QoS traffic matrix and d) negotiation of QoS capabilities between adjacent domains. This way we denominate Network Broker (NB) to the agent extended BB with the new functionality.

This paper is organized as follows. In Section 2, we describe our overlapped QoS topology proposal, at intra-domain and at inter-domain. Section 3 presents our Network Broker definition, its components and functions accomplished by it. Finally, Section 4 summarizes our conclusions.

2 Overlapped QoS Topology to Best-Effort

Best-effort and QoS traffics coexist in many networks in real world. If one main task of the routing function is to maximize efficiency of resources, for a network where both traffic types exists, this efficiency can be measured by the fulfilment of two objectives. On one hand, the number of QoS flows that can be admitted in the network should be maximized. This is equivalent to minimize the call-blocking ratio if admission control mechanisms are applied. On the other hand as much throughput as responsiveness of best-effort traffic should be optimized.

Both objectives could be contradictory, since whereas the first objective only considers QoS traffic, the second one only considers best-effort traffic. Finally, both traffic types can have very different distributions.

If resource reservations for QoS provision are made, QoS traffic will not be affected by best-effort traffic. However, if these reservations are overestimated, best-effort traffic will see reduced available throughput. For example, links with low QoS traffic load and high best-effort traffic load could exist. For many QoS algorithms, these links are considered good candidates to direct QoS traffic through them. When QoS traffic on these links is increased, best-effort traffic is inevitably affected by QoS traffic since routers congested along this path will discard best-effort packets in benefit of QoS traffic.

For Internet case, it seems reasonable that an implantation of QoS provision scheme should present minimum negative impact on the current operation of the best-effort traffic. Otherwise, perhaps a good subset of world ISPs and carriers would not be willing to assume the cost of the implantation on their functional best-effort ASs.

Current architecture of the Internet does not offer any guarantee regarding any QoS parameters. We assume then that any modification of the architecture that could allow the Internet to offer a greater guarantee will be considered an advance, but only if it does not negatively impact on the current operation of the network.

Therefore, this considers the following premises:

- To incorporate mechanisms that allow negotiating the QoS obtained from the network. If the network could not offer acceptable minimum QoS for clients, it should be indicated rejecting the negotiation [4].
- That the incorporation of these mechanisms has a minimum negative impact on the operation of current architecture. Best-effort and QoS traffics should coexist, not to be faced.
- That the incorporation of these mechanisms has a minimum impact on configuration changes of current equipment, providing automatic configuration mechanisms there where they can be available.
- That the best-effort protocol stacks do not have to be modified in any of the elements of the network.

Our main aim is to provide QoS to the clients who request and contract it. If that it could to be some way to provide this QoS, network must arrange it. If QoS parameters are unacceptable, the client will not meet a different case to the current operation of the Internet that does not make any guarantee regarding these parameters. If it is impossible to even guarantee the QoS service with these new mechanisms, we will not have arrived either at a worse situation to the present one.

This way,

- Users will be beneficiaries of the advantages that it supposes the possibility of contracting a QoS connection.
- ISPs have guaranteed a competitive advantage with regarding to the rest of ISPs that do not offer a QoS service.
Carriers that offer a QoS service will end up handling more traffic volume, since they will continue carrying best-effort traffic and additionally, all traffic of the QoS connections that non-QoS enabled carriers of the competition lose. This increment in the traffic volume directly or indirectly will take them to an increase in their profit [7].

Therefore, in our proposal we make a definition of two different planes at the network. This definition consists on a separation of a QoS plane from the best-effort plane, in a way that they are only overlapped at the control mechanisms of the data path. The overlapping in these control mechanisms is necessary because it is essential to limit resources destined to the best-effort traffic at those routers where QoS traffic and best-effort traffic may coexist.

QoS routing functions, resources reservation, access control and traffic classification are at the QoS plane. Best-effort routing functions are at the best-effort plane (table 1), using the same protocols that existed at the functional AS.

<table>
<thead>
<tr>
<th>Function</th>
<th>Best-effort plane</th>
<th>QoS plane</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routing</td>
<td>X (*)</td>
<td>X (**)</td>
</tr>
<tr>
<td>Control at data path</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Access control</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Resource reservation</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

(*) According to best-effort IGP in the AS
(**) According to specific QoS algorithm for QoS plane

Table 1. Functions at QoS and best-effort planes

The belonging of a router to an AS is equivalent to the belonging of the same router to the best-effort plane. The belonging of a router to the QoS plane comes given by the fulfillment of the following requirements:

1. If it is a core router:
   (a) Router is able to use different routing databases based on the DSCP field.
   (b) Router is able to apply policing mechanisms.
2. If it is an edge router:
   (a) Router is able to use different routing databases based on the DSCP field.
   (b) Router is able to apply classification and marking mechanisms.
   (c) Router is able to apply policing mechanisms.

The QoS plane is compound, besides the routers that are part of him, by an agent placed in the AS in charge of the accomplishment of the following functions:

- To establish in router the control policies at the data path for QoS provision. The agent establishes classification and policing functions at routers of QoS plane.
- It is responsible of access control. Users must to have negotiated the QoS parameters that they require from the network previously to the establishment of QoS connections. In case of a QoS service provision, if a user had not required the service or parameters would have not been accepted, connections only could be established at best-effort plane. It is responsibility of the best-effort plane to undertake access control at this point.
- Maintenance of the overlapped QoS topology. Computation of the classification and policing parameters at the data path, and computation of links metrics are required for the QoS plane to be maintained.
- Capture and information processing of the IGP. To be able of maintaining the overlapped QoS topology, the agent should consider the best-effort topology. Therefore, the agent will be destination of the routing protocol information at the AS.
- Treatment of traffic statistics collected from the ER for obtaining of traffic matrix. The treatment of these statistics, together with the SLA database and topology database allows the agent to accomplish the calculation of resource provisioning parameters in the network for QoS and best-effort planes.
Negotiation of QoS capabilities with adjacent domains. Since it is allowed the coexistence of QoS-enabled and non-QoS-enabled ASs, the QoS ASs will have to consider which they are not capable for QoS provision.

Automatically to coordinate and to negotiate the QoS provision between adjacent domains. Establishment of new QoS provision parameters could be requested by the agent, considering traffic observed in the AS and which of adjacent ASs are QoS-enabled.

2.1 Overlapped QoS topology at the Intra-domain

For end-to-end QoS provision, intra-domain QoS provision is one issue to solve. Although the overlapped QoS topology proposal presented in this work is applicable to other routing protocols, we will focus at the case of Open Shortest Path First (OSPF) [9] for intra-domain QoS provision.

OSPF protocol is an Interior Gateway Protocol (IGP) that belongs to the family of the link state routing protocols. Besides the group of mechanisms provided by OSPF to perform best-effort routing, it is also provided the possibility to use several link state databases that can be selected based on field TOS field of IPv4. Although the TOS-based routing has been eliminated recently of the OSPF specification due to its actual scarce implantation, routers are still able to inform of the specific costs associated to the TOS field in the LSA, to maintain backwards compatibility. This provides the possibility to implement the QoS routing like an extension of the standard characteristics of TOS in OSPF.

Regarding the maintenance of an overlapped QoS topology to best-effort, the agent at the QoS plane is responsible for the maintenance of a graph representing the overlapped QoS topology. This graph contains all routers of the AS except those ones that do not have necessary capabilities for QoS provision (for example, if they do not have the possibility of to differentiate QoS routing from best-effort one or they do not implement policing policies that allow to perform reservations for the no-best-effort classes).

For the construction and maintenance of the graph the collected routing information from the IGP is used. Once it is well-known the best-effort topology, the agent consults the QoS capabilities of routers directly via of a signalling protocol, eliminating from the graph all of them that do not have enough capabilities. In the figure 1 it is shown an example of an overlapped QoS topology on an AS, where one of the router of the core (CR3) does not support the requirements to remain in the QoS plane.

At the QoS plane, the implementation of the costs function of the SPF routing algorithm is proposed to be made in the agent. The reasons to undertake this location are the following two. On one hand the fact that the agent is going to have a precise knowledge of the traffic flows that cross the AS -as it will be exposed below-, making possible the optimization therefore of the QoS routing function. On the other hand, this location also allows to minimize the impact on the requirements of the routers of the AS, since they should not implement an additional routing protocol to the one used for best-effort routing, but rather it is enough with the selection of the proper link state database. Regarding the routing function out.
in the QoS plane -using OSPF in the best-effort plane-, the only requirement that routers must satisfy to belong to the QoS plane is the possibility of selecting link state databases. This capability is established in its specification [9] as a requirement.

2.2 Overlapped QoS Topology at the Inter-domain

Once mechanisms for QoS provision in the intra-domain are established, the following step it is to determine if this provision should also extend to the inter-domain. In [13] it is argued that most of important ISPs have direct peerings among them, be public or private. Therefore the application flows will cross no more than two different ISP for clients who use these ISPs, so it is not necessary to use mechanisms to choose the bests intermediate ISPs. Instead of it, the only requirement would consist on the correct selection of egress router so that the selected path does not become congested.

However, the studies presented in [8], based in empiric experiences, conclude that even not using QoSR but routing based on the shortest-path search, the number of ASs that crosses an application flow it is between 3 and 4 for more of 70% of the cases. Therefore it is necessary to search routes that necessarily cross ASs that are capable of to provide QoS.

In order to approach the aspects related to inter-domain constraint based routing it is necessary to know the logical organization of ASs within the Internet. ASs pairs are generally impaired in a client/provider or in a peer/peer way. ASs can be then classified in following way:

- Client AS. When an AS only has one connection to another AS (figure 2a).
- Transit AS. When an AS has two or more connections to other ASs and it transports the traffic so much local as transit traffic among the ASs to those that it is connected. It is the typical case of an ISP or of a carrier (figure 2a).
- Multi-homing AS. An AS is a multi-homing one (call it AS A) when it is connected with at least two AS of different providers (call them AS B and AS C). Usually, in this configuration, the AS A uses the AS of one provider as the primary connection to the Internet, happening to use a different provider AS if the primary connection is interrupted (figure 2b). This outline can also be used for load balancing through those AS B and AS C. Anyway, AS A does not allow transit of traffic between ASs B and C through him. As it is discussed in [6], the closer the ASs are to the core of the network, the percentage of multi-homing ASs increases. In fact, in this same work it is concluded that more than 73% of the clients connected to providers in the core are multi-homing ones. In addition, an upward evolution is appreciated related to the percentage of multi-homing AS in the time. Finally, more than 61% of the ASs present at the moment in the Internet they are multi-homing ones.

\[\text{Client AS} \quad \text{Provider AS} \quad \text{Multi-homing AS} \]

\(\text{Fig. 2. AS Classification based on traffic transit}\)

Indeed in the exposed conclusions concerning to current organization of the multi-homing ASs in the Internet we find one of most interesting scenarios of application of our work, as much if the AS belongs to a client or to a provider. In this scenario, the AS gets connectivity by several providers. If we suppose that some of these providers offer QoS services, the equipment of the multi-homing AS would be able to negotiate automatically the establishment of QoS routes through the providers that have an architecture
able to provide QoS connections. The non-QoS connections could then be directed well by the provider that offers QoS services or by the one that does not offer these QoS services (figure 3). The fact that a complete implementation of this architecture at all nodes of the Internet it is not necessary becomes a particularly advantageous characteristic under this scenario.

Anyway, the agent of the QoS plane in a domain must know which the adjacent domains that support the QoS provision are. Therefore, the use of capabilities advertisement mechanisms it is proposed like it is specified in [3] for the Border Gateway Protocol (BGP). It would also to be considered the use of signalling mechanisms among agents of adjacent domains to interchange this information.

3 Network Broker

One aspect that is object of an intense research currently is bandwidth assignment and control inside a DiffServ domain so that objectives of the organization are satisfied. A possible solution is making users individually decide what service to use, but is not probable that this approach is adequate to reach the wanted objectives. Another approach it is to have an agent for each domain, called Bandwidth Broker (BB), which registers the current assignment of traffic and process new service requests on watch in agreement with organization policies and current state of the traffic assignment [10].

A BB manages the bandwidth in a way that:

- It is responsible not only of the assignment of traffic to the different classes of service inside the domain, but also of establishing classifiers and droppers at edge routers of the network.
- The BB maintains an access control database that specifies what users are allowed to request what services in what instants of time.
- The BB authenticates each client first and then it decides if there is enough bandwidth for a particular service to be satisfied.

The BB also maintains bilateral agreements with other BBs at adjacent domains. It is expected, at least initially, that static bandwidth pre-allocation will be predominant in the bilateral agreements. However dynamic establishment of bilateral agreements would be possibly reached because the BB is able to treat a wide granularity range within bilateral agreements.

Opened decisions exist related to the BB architecture in DiffServ. Some implementations have been proposed and analyzed currently, although in most of the cases it is not arrived to a complete specification of the functionality BB to fulfil the objectives exposed above.

In our proposal we extend the functionality described for a BB, in the following aspects:

1. Maintenance of a QoS topology overlapped to best-effort.
2. Capture and compute the information of the IGP.
3. Process traffic statistics collected from ERs of the domain to calculate the QoS traffic matrix.
4. Negotiation of QoS capabilities between adjacent domains.

This way, we denominate Network Broker (NB) to the BB agent extended with the new functionality. For incorporating the described functionality, the NB is designed as it is shown in figure 4.
3.1 Components of the Network Broker

The components of the NB are the following ones:

- Users’ database. It is used for access control to allow the establishment or modification of the QoS provision agreements, reason why it is directly related to the SLA database.
- Network topology database. It contains the following information:
  - Identification, location and characteristics of the router of the AS.
  - Network interfaces of the routers of the AS and their characteristics.
  - Links states.
- SLAs database. It contains approved and established SLAs among users of the domain and network operator. It also contains established SLAs among the domain and its adjacent ones.
- Traffic matrix database. It contains data collected about the traffic matrix from the ERs of the domain, and together with the SLA and network topology databases, its purpose is the BB to be able to calculate traffic control parameters at the data path.
- Network operator interface. In those modifications of the SLA database in which it is required a human intervention, the NB provides an interface so that the network operator supervises and approves these modifications. Modifications will come caused by users’ requests to the NB, requiring human intervention for approval or supervision of those modifications that cannot be included in an automated QoS provision policy at the organization.
- User interface. Through this interface, the user can request the establishment of a new SLA or the alteration of an already existing one.
- Interfaces for configuration and control of routers of AS. This configuration and control can be made using different mechanisms:
  - RSVP [2], where all mechanisms of its specification are not used, but only the reservations signalling subset. Through these messages, the NB signals to corresponding router the reservations parameters for each one of the service levels.
  - COPS [5], a TCP-based protocol that uses a client/server model based on the following elements:
    - The PEP (Policy Enforcement Point) it is an entity where the established policies are applied. The PEP will generally be a router or a gateway located in a border of the AS.
    - The PDP (Policy Decision Point) it is entity that generates the rules to apply and agreement decisions for requests received from the PEP. The PDP acts therefore as a server for the PEP. It is supposed to be implemented at the NB of the AS.
SNMP (DiffServ MIB [1]), incorporated in most of routers, it allows the NB to control routers if other protocols like RSVP or COPS are not supported.

Telnet/SSH. If routers did not support COPS or RSVP and their SNMP agent do not implement the DiffServ MIB, the NB could still control routers using Telnet/SSH mechanisms to gain access to the CLI (Command Line Interface). This kind of control can only be made with a specific plug-in which depends on the command set of the CLI of each router.

### 3.2 Control of Data Path Mechanisms

The NB is the responsible of generating and indicating classification parameters to all ERs of the QoS plane. It must also generate and indicate the policing parameters to all the router of the QoS plane. For the NB to be able to determine these parameters, it would be necessary to know in a precise way the traffic matrix that traverses the domain. In the same way and due to the centralized approach of the cost assignment function to the links, the NB should also know the topology of the AS in which it resides. This information is obtained from the topology database. Finally, regarding the assignment of parameters for the control mechanisms at the data path, to make resources reservation in each router of the AS, it is necessary to know the QoS parameters that are guaranteed to each user. Therefore, parameters of control mechanisms at the data path are calculated based on the information obtained from the SLA DB, Topology DB and Traffic matrix DB.

### 3.3 Overlapped QoS Topology Management

Another of the functions of the NB is the maintenance of the overlapped QoS topology. On one hand, maintenance of the overlapped QoS topology implies the capacity for determining -by the NB- if a router is eligible to be part of the QoS plane. Using the configuration and control available interfaces, the NB can ask to the router the fulfilment of the requirements for the belonging to the QoS plane.

Starting off by the knowledge about the overlapped QoS topology, the NB assigns the links costs directly at routers in this topology. This way, it is possible to use different routes than the used in the best-effort plane for same origins and destinations. For OSPF domain case, it is proposed the use of at least one additional link state database than the best-effort one, making a correspondence between the DSCP field of the IP packets and the link state database selected like it is shown in the figure 5.

One important advantage of this method is that the existence of link state databases selected based on the TOS field is present in the OSPF protocol from their first version. Therefore, it is not necessary to modify the IGP at this point.

![Fig. 5. Correspondence between DSCP and link state databases](image)

The maintenance of the overlapped QoS topology also implies a supervision of the links states of the network to be made. In the event of a persisting failure of a group of them, it could be necessary to recalculate the links costs to balance the distribution of QoS traffic in the domain to adapt it to the new situation. Anyway, it can also be necessary to undertake this recalculation if they occur significant alterations regarding the QoS requirements of the users or regarding the SLAs negotiated between adjacent domains. In any one of these two cases, alterations are managed by the NB, and must be this element the one that determines if reconciles are needed. It must be considered that the proposed approach, where the NB establishes the links costs of the routers at the AS, does not indicate routers the paths of each flow. This strategy allows a certain level of autonomy to the routers at the QoS plane because if several link failures occur routers will continue sending the information on the operative route of smaller cost of the QoS plane.
3.4 Discovery Procedure

The NB could maintain the topology database if it is able to access the IGP protocol information of the best-effort plane. This discovery procedure avoids the necessity to alter an IGP protocol that already could be operative in a domain. Also, for the characteristics and links of those routers to be collected, some of the configuration and control mechanisms described above will be used.

3.5 Knowledge of Traffic Matrix

Combining the information stored in the SLA database with the topology information collected from the IGP -and stored in the topology database- it can be determined one end of the flows according to each SLA. In some cases the SLA could even allow to determine both ends (for example, in the VPN tunnel establishment, where the two ends could be specified for a tunnel). Therefore, although this information can help when establishing the policing parameters, to build these parameters only with this information can to take to suboptimal results, due to an imprecise knowledge of the way that flows will traverse the domain. Therefore it is proposed the incorporation of collection mechanisms of traffic statistics of the ERs. With this information, the NB can optimize the resources reservations made for the traffic classes in each one of the router in the QoS plane.

Actually, this collection of information at the ERs does not have that to be implemented as specific functionality incorporation. Instead of it, it can be used the RMON specification [11], widely implanted among the current commercial network equipment. Using RMON, the access to the traffic matrix is made through the OID mib-2.rmon.matrix.

4 Conclusions

In this work we have presented a proposal of architecture for an incremental provision of QoS based on Network Brokers. Its design enables the provision of QoS guarantees for application flows even if not all elements of the network are designed for QoS provision.

On one hand, in the intra-domain, it is enough whereupon at least one route exists from an ingress router towards an egress router where all intermediate network elements are enabled to provide QoS services, for end-to-end QoS to be provided to users. On the other hand, in the inter-domain, if at least one route exists in the one that all ASs in the path have the needed capabilities of QoS provision for QoS services be provided. Since all AS that allows the QoS provision must have a NB, it enables NBs of the adjacent ASs to exchange information for the establishment of QoS routes through them, ignoring the ASs that do not provide QoS services.

This scenario is particularly attractive for clients that obtain Internet connectivity by more than a provider (a multi-homing AS case) which corresponds to more than 61% of the ASs in the Internet. If any of the providers is enabled to provide QoS, the multi-homing AS can use it to redirect QoS traffic through while best-effort traffic can be redirected through any provider, be QoS-enabled or not.

References