On the Enforcement of Quality of Service Policies for IPTV Service Offerings

Christian Jacquenet and Julien Maisonneuve

Abstract—This paper discusses both Quality of Service (QoS) information signaling and IPTV traffic engineering options that may serve as key components of IPTV service-specific QoS policies, which are designed to accommodate with customers’ requirements and constraints (such as access capabilities) as well as ever-changing networking conditions that may possibly affect the level of quality associated to the delivery of such services.

Index Terms—IP multicast, QoS signaling, quality of service, traffic engineering.

I. INTRODUCTION

With the forthcoming yet massive deployment of IPTV service offerings over IP networking infrastructures, IPTV service providers are being challenged by the provisioning of the adequate level of quality, whatever customers’ access conditions to the network, whatever the technology used by the IPTV terminal, and whether customers are in motion or not.

From a design perspective, most of IPTV services are deployed over a broadband, multicast-enabled networking infrastructure. While the multicast transmission scheme represents a privileged way to optimize the use of switching and transmission resources because of the deterministic nature of its data replication procedures, the dynamic establishment and maintenance of multicast distribution trees remain receiver-initiated, without making any assumptions about their access capabilities, or the state of switching and transmission resources that are involved in the maintenance of such trees, at the risk of possibly degrading the quality of the IPTV service delivered to the customer (e.g. in the case where network congestion arises in DSL (Digital Subscriber Line) access environments).

The design and the (dynamic) enforcement of IPTV service-specific QoS policies that not only address customer’s requirements, but can also accommodate with evolving networking conditions has therefore become a key issue for IPTV service providers.

This paper discusses some of the options that can be seen as key components for the efficient and dynamic enforcement of IPTV service-specific QoS policies. Such options rely upon elementary capabilities (such as traffic marking and conditioning, or MPLS traffic engineering) that are commercially available but currently not combined and used for the purpose of delivering IPTV traffic with an adequate level of quality.

There is indeed no current deployment and operation of a combination of the techniques discussed in this paper at the scale of an IPTV service provider’s domain: Live broadcast IPTV services currently rely upon the use of the IP multicast transmission scheme, assuming its well-known limitations that have been briefly exposed in the previous paragraphs.

From this perspective, this paper proposes an IPTV QoS policy design and enforcement scheme that precisely relies upon the combination of a set of elementary capabilities that is meant to help IPTV service providers in better providing hard guarantees about the level of quality associated to the delivery of IPTV services that can benefit from the IP multicast transmission scheme.

The paper therefore investigates signaling options to propagate receiver-specific QoS information along the multicast-enabled routers involved in the dynamic computation and selection of the tree structure that will forward the IPTV traffic to the receivers, and then discusses multicast-inferred traffic engineering options that can accommodate receiver-specific QoS requirements and constraints.

II. SIGNALING QoS INFORMATION

A. The DiffServ Architecture

The DiffServ architecture relies upon a set of elementary capabilities—traffic classification and marking, conditioning and scheduling, traffic discarding—meant to differentiate the processing of datagrams forwarded through a set of DiffServ-enabled IP routers. Thus, the differentiated processing is local to a router and is based on the concept of PHB (per-hop behavior), which specifies, for each type of traffic characterized by a single criterion or a combination of these criteria (destination address, value of DS (differentiated services) field and DSCP (differentiated services code point), transport level protocol identifier, etc.), the behavior of the router from a traffic processing perspective.

DiffServ architectures do not provide an absolute guarantee of quality of service, but introduce differences in how each class of flows is forwarded, hence delivering relative QoS between different flows. DiffServ-based QoS policies are usually enforced at the scale of a domain,1 where the IPTV service

1In this paper, a DiffServ domain is defined as a set of DiffServ-enabled routers that are operated by a single, globally unique, administrative entity, such as an IPTV service provider.
provider controls the resources of the network that will be involved in forwarding IPTV traffic (among others). The use of DiffServ mechanisms at a larger scale (e.g. the Internet) raises consistency issues between domains, since the processing of a given PHB might very well differ from one domain to another.

The resources of a DiffServ domain are used according to the contracts established between the service provider and the clients, which describe the scope, the nature and the QoS expectations related to the delivery of a service. Such contracts can influence the way the planning policy of the DiffServ network will be designed and enforced. When this network is used also to forward IPTV multicast flows, the resources used by these flows may exceed what was negotiated between the client and the service provider.

Since the distribution trees are established at the initiative of the receivers, whose number and location are not known a priori, the amount of IPTV traffic forwarded along the trees can exceed the amount of traffic that will have to be processed by the routers of the domain, as negotiated between customers and the service provider. This problem is known as neglected reserved subtree (NRS, [1]) and it affects the level of quality associated to the forwarding of multicast flows within the DiffServ domain: whenever a new receiver joins a multicast group, the multicast distribution tree will be expanded by a new sub-tree. If the amount of resources that are used by such expansion are not taken into account by the network management, the level of quality that is provided to the already-connected receivers might be affected, if not violated. The negative effect on existing contracts because of such neglected resource reservation must therefore be avoided.

B. Toward DiffServ Multicast Distribution Trees

DiffServ-based IPTV traffic forwarding mechanisms can be used to influence the way multicast distribution trees are established and maintained, taking into account the access capabilities of the receivers, among other possible metrics. From this point of view, signaling such access capabilities throughout the multicast network ([2]) contributes to the enforcement of an IPTV service-specific QoS policy taking advantage of DiffServ capabilities.

With the use of DiffServ, it becomes possible to process differently IPTV traffic delivered to ADSL (asymmetric digital subscriber line) receivers (whose access rate generally do not exceed several Mbit/s) from IPTV traffic delivered to FTTH (Fiber-to-the-Home) receivers (whose access rate is typically measured in tens of Mbit/s) for example. Indeed, the marking of IGMP (Internet group management protocol) Report messages sent by the receivers, and possibly relayed by the access device (such as the BRAS, broadband remote access server, or the MSAN (multi-service access node)) toward the router in charge of processing these multicast group subscription messages, is meant to provide an indication about the access capabilities of the receiver.

As an example, such marking can be derived by the access device (BRAS, MSAN, etc.) from the IPTV customer-specific information that can be conveyed by means of an AAA (Authorization, Authentication and Accounting) protocol, such as RADIUS (remote authentication dial-in user service, [3]). RADIUS Access-Accept messages convey the information specific to the receivers’ profiles, by means of a specific AVP (attribute value pair). Contents of these RADIUS Access-Accept messages may include:

— Information pertaining to the capability of the receiver to subscribe to one or more multicast groups which correspond to several IPTV services;
— Information (e.g. bandwidth) pertaining to the marking of IPTV traffic. This information will be used by the access device (e.g. the BRAS, which embeds the AAA client) to enforce the proper Per Hop Behavior by means of selecting the proper marking ([4]).

The explicit activation of classification and marking functions for IPTV traffic in access devices can then be used by the group of routers involved in the forwarding of the multicast traffic toward the receivers:

— IGMP Querier routers will forward IPTV multicast traffic according to the access capabilities of the receivers they are connected to: hence ADSL customers’ IPTV traffic could be EF-marked (expedited forwarding, [5]), while FTTH IPTV customers’ traffic could be AF-marked (assured forwarding, [6]);
— Likewise, PIM (Protocol Independent Multicast, [7]) routers, involved in the establishment and maintenance of the IPTV-specific multicast distribution trees, will propagate QoS information to their upstream neighbors, by means of a QoS attribute that can be conveyed in PIM Join messages, as per the format described in [8], and which would carry DiffServ information that can be further derived by the PIM routers, yielding the establishment and maintenance of DiffServ-inferred multicast distribution trees.

Within such DiffServ-enabled multicast network environments, the manipulation of the QoS information propagated by PIM Join messages in a specific attribute is meant to influence the calculation of the routes provided by the IGP protocol (interior gateway protocol) and used by the RPF (reverse path forwarding) Check procedure enforced by the multicast routers.

Using this QoS information in the calculation of the shortest paths gives way to a modification of the process of route multicast computation and selection as illustrated in Fig. 1. The chronology of the different states mentioned in Fig. 1 is established as follows:

— Initialization (state S0): initialization of the routing processes (IGP, PIM, etc.) with constitution of the RIB (routing information base), for example the LSDB (link state database) database in the case of the Open Shortest Path First (OSPF, [9]) protocol and of the descriptive base of QoS policy-specific information as used by the routers to compute and select QoS-inferred routes. The routers exchange their information during the initialization procedure;
— Waiting (state S1): the routers are waiting for the processing of a new request (for example, the reception of an IGMP Report message and/or the modification of the value of the QoS information conveyed in PIM Join messages, motivated by adding or deleting an interface, a set of interfaces, a router, and/or a group of routers participating in
the calculation and maintenance of multicast distribution trees, and/or the overcoming of certain thresholds that will motivate a new route calculation;

— Processing (state S2): the router verifies if a route whose characteristics are compatible with the quality requirements identified by a DSCP marking is present in its table. The router performs the RPF Check procedure, on whose success depends the forwarding of IPTV traffic through the interfaces considered (entries of the outgoing interfaces tables of the router, typically).

This approach leads to the dynamic establishment of multicast distribution trees out of which at least a part of their branches (typically terminating branches) will be QoS-specific, as per the policy enforced by the routers of the multicast domain. In this case, the marking indication conditions the forwarding of the IPTV traffic along such or such branch, so that, for a given service, it will be possible to envisage as many distribution trees as DSCP markings, assuming the core of these trees is shared by the various types of IPTV traffic.

Fig. 2 is an example of these QoS-inferred multicast distribution trees established for a given IPTV service (identified by a single multicast group address or (S, G) channel), whose traffic will be marked differently according to the receivers’ access capabilities.

However, the marking of multicast flows according to the needs or constraints expressed by the recipients does not fully address the NRS problem, because the use of DiffServ mechanisms is not sufficient to actually guarantee the reservation of resources that will accommodate the dynamic grating of additional sub-trees while preserving the level of quality provided to already-connected receivers. The use of additional traffic conditioning capabilities can rely on a counting function [10] whose principle consists of estimating the volume of multicast traffic as a combination of unicast flows (based on the number of receivers connected to the distribution tree), then of enforcing the traffic conditioning policy, depending on whether the volume of estimated traffic at the level of the router that participates in the maintenance of the distribution tree is compatible, for example, with the number of available tokens if the algorithm used to condition the traffic is a token bucket algorithm.

The application of such a policy relies in this case on the maintenance of look-up tables which contain the group addresses, the values of the supported DSCP markings and the weight associated to each marking, as illustrated in Table I. Such weights typically correspond to the activation of scheduling algorithms such as Weighted Fair Queuing (WFQ), which will therefore prioritize traffic accordingly.

IPTV multicast flows are therefore processed by traffic classification capabilities embedded in multicast-enabled routers, which refer to the look-up table to determine the weight associated to each flow. This weight is a given value derived from the PHBs that correspond to the DSCP marking(s) assigned to these flows. The traffic conditioning capabilities activated by the corresponding routers are illustrated in Fig. 3.

Table 1

<table>
<thead>
<tr>
<th>Group Address</th>
<th>DSCP</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>233.12.144.2</td>
<td>EF</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>AF1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>AF2</td>
<td>2</td>
</tr>
<tr>
<td>233.12.144.4</td>
<td>EF</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>AF2</td>
<td>2</td>
</tr>
<tr>
<td>239.12.144.25</td>
<td>AF1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>AF2</td>
<td>2</td>
</tr>
</tbody>
</table>
Preliminary simulation works of these approaches made it possible to qualify their large scale deployment capabilities, while taking into account the multiplication of PIM (*, G) or (S, G) states, according to the number of DSCP markings considered. However, the dynamic establishment and maintenance of DiffServ-inferred distribution trees do not provide any strict guarantee regarding the level of quality associated to the IPTV service delivered to the customers, because this approach relies upon the activation of a set of capabilities that remain local to the multicast routers involved in the forwarding of IPTV traffic.

The provisioning of hard guarantees therefore encourages the use of additional capabilities, namely (IPTV) traffic engineering techniques that can be defined as a set of path computation techniques that help service providers in selecting specific routes whose characteristics will comply with IPTV-specific requirements (e.g. QoS requirements defined in a Service Level Specification contractually negotiated and invoked between a customer and a service provider) and constraints (e.g. the enforcement of the network planning policy that has been designed by the operator).

III. INTRODUCING TRAFFIC ENGINEERING CAPABILITIES FOR IPTV SERVICES

MPLS (Multi-Protocol Label Switching) is a switching technique that allows the enforcement of a consistent forwarding policy at the scale of a flow, where a flow can be defined as a set of IP datagrams that share at least one common characteristic, such as the destination address. Unlike the hop-by-hop IP routing paradigm case, MPLS makes sure that all the IP datagrams of a given flow (designated as a FEC (Forwarding Equivalence Class) in the MPLS terminology) will be conveyed over the very same path, which is called a Label Switched Path (LSP).

The use of MPLS as the switching technique is often seen as the cornerstone for the deployment of traffic engineering capabilities within a domain, based upon the use of a CSPF (Constraint-based Shortest Path First) algorithm for the dynamic computation and selection of traffic-engineered paths that will be entitled to convey different kinds of traffics, depending on the QoS requirements associated to such flows.

The capacities of traffic engineering associated to the MPLS switching technique are today the object of operational deployments as far as the establishment of point-to-point LSP (label switched path) meant to forward unicast traffic is concerned.

Naturally, such capacities represent an attractive field of research for the deployment of IPTV services with a guaranteed level of quality over MPLS infrastructures, including services that can naturally benefit from the IP multicast transmission mode [11].

A. Issues in Deploying IP Multicast in MPLS Networking Environments

The IP multicast transmission scheme relies upon the dynamic establishment and maintenance of distribution trees that are computed based upon the information derived from the activation of an IGP routing protocol. The adaptation of such multicast tree structures to MPLS Traffic-Engineered LSPs established thanks to the activation of a CSPF algorithm and of the RSVP protocol (resource reservation protocol) [12] raises the following difficulties:

— Calculation and establishment of LSPs: the multicast distribution trees suggest the establishment of point-to-multipoint LSPs. The functions of MPLS traffic engineering associated to the establishment of such routes are being specified within the IETF (Internet engineering task force), and the standardization effort is now close to stabilization. This implies that the establishment of such MPLS-based tree structures needs to be compatible with the dynamics of multicast subscription/cancellation procedures, characteristic of the maintenance of terminal branches of the multicast distribution trees;

— Multicast traffic routing within point-to-multipoint LSP tree structures: the multicast distribution trees are established and maintained according to the RPF Check procedure which is based on the routes computed by the IGP routing protocol in order to reach the source through the shortest path. From this perspective, the use of MPLS traffic engineering capabilities applied to the establishment of point-to-multipoint LSPs to forward the multicast traffic does not represent a sufficient condition to enable the routers to make the routing decision (and thus to ensure the deterministic replication of information). Indeed, the context of switching a shared tree in a PIM-SM environment (RPT, rendezvous point tree, a tree whose the root is a specific router in the network, called the Rendezvous Point (RP) router) toward a shortest path tree (SPT), a tree whose root is the source, can lead to a situation where a router participating in the establishment and maintenance of point-to-multipoint LSPs and of the multicast distribution trees for the corresponding group G will have to maintain two states—the state (*, G) for which it is assumed that the router forwards the traffic along the LSP, and the newly-created state (S, G) (following the RPT/SPT switching), for which the router will not have the capacity to associate an outgoing label in order to forward the multicast traffic on the proper LSP (assuming such LSP could have been established). Moreover, in the case when certain unicast flows share the same label as the multicast flows, the MPLS routers will have to analyze the content of the IP header in order to make the decision for a proper forwarding, which questions the advantages of the MPLS switching technique;

— Aggregation capabilities of the multicast traffic: the MPLS switching technique is based on the concept of flows. This technique makes it possible to aggregate the traffic in such a way that one IPTV flow will be forwarded over a given LSP within the MPLS domain, so that the forwarding is homogeneous for every datagram of the flow. Conversely, IP multicast datagrams cannot be easily aggregated into homogeneous flows, because of the hop-by-hop IP routing paradigm that conditions the results of the aforementioned RPF Check procedure.

Given these issues, the next section further discusses the computation and establishment of point-to-multipoint LSP tree structures.
B. A Three-Stage Approach

The capabilities of MPLS traffic engineering used for the establishment of point-to-point LSP paths can be extended to the establishment of point-to-multipoint trees [13]. The constraints associated to the deployment of such trees are the quality constraints defined for IPTV services, namely the QoS constraints specific to the delivery of multiple, real-time TV channels, for example (low if no inter-packet delay variation, low if no packet loss, etc.).

The establishment of point-to-multipoint LSP is done in three stages:

— Discovery of the topology of the network having traffic engineering capabilities: the dynamic routing protocols with traffic engineering extensions are used in order to announce the capabilities of the routers to calculate the paths with constraints, but also to disseminate the information on the available bandwidth on each of the links of the MPLS domain;

— Computation of the tree: this complex calculation can be done either by the router directly connected to the source (“head” router), or by an external component, designated as PCE (path computation element) [14];

— Signaling of the point-to-multipoint LSP: this signaling is done according to the specific extensions of the RSVP-TE protocol that will transport hop-by-hop the identification information of the point-to-multipoint LSP.

Fig. 4 describes the functional blocks used for the calculation and establishment of traffic-engineered point-to-multipoint LSPs in an MPLS network.

C. Computation of Point-to-Multipoint LSP-Based Trees

The point-to-multipoint LSPs are established according to cost considerations which may vary according to the usage constraints and situations.

Typically we distinguish:

— The lowest cost trees (Steiner’s trees), generally used in the context of greedy applications in terms of bandwidth (IPTV broadcasting services, for example);

— The shortest path trees, typical for multicast dynamic routing protocols and generally used in the context of applications having strong real-time requirements (video-conferences, for example).

Fig. 5 depicts such typology, where the cost is the typical hop count metric.

There are several options to compute the (point-to-multipoint LSP) tree, besides the capability to statically configure the LSP from the “root” router:

— Dynamic computation performed by the “root” router of the point-to-multipoint LSP, averaging the use of a CSPF algorithm adapted to this context. This option implicitly imposes constraints in terms of CPU, resources and memory which may be incompatible with the router’s capabilities;

— Dynamic computation performed by a tool external to the network (by means of a Path Computation Element,), whose results are then derived into configuration information used by the “root” router of the point-to-multipoint LSP. This option has the advantage of unloading routers of calculations which use a lot of CPU resources and makes it possible to use algorithms capable of taking into account complex heuristic methods, such as those used by Steiner’s trees. However, this option may be facing issues of reliability and reactivity if the PCE element is strongly and frequently solicited.

Since this approach is, however, still new, there are no industrial developments enabling the dynamic calculation of such trees capable of addressing scalability issues.

D. Establishment of Point-to-Multipoint LSP-Based Trees

The RSVP-TE protocol is used to signal information (identification of point-to-multipoint LSP, resource reservation, etc.) which will be used by the routers for the establishment of traffic-engineered point-to-multipoint LSPs. Typically, these tree structures are defined as a set of point-to-point LSPs established between the router directly connected to the source and the routers directly connected to the receivers.

Hence the notion of “sub-LSPs” ([13]), where a given Point-to-Multipoint LSP is composed of a set of sub-LSPs, which will be combined by the routers directly connected to
different levels of the tree in order to form the point-to-multipoint LSP. Sub-LSPs are used to minimize the amount of RSVP states that will have to be maintained by the participating routers, hence facilitating scalable deployments of this approach.

The information necessary to the routers to build the different sub-LSPs is signaled through a unique RSVP_PATH message if the number of sub-LSPs to build remains compatible with the size of the RSVP_PATH message, or even more RSVP-PATH messages, all of which containing a part of the sub-LSPs to be established.

The principle of establishing a traffic engineered point-to-multipoint LSP is illustrated in Fig. 6.

The resource reservation process used by the point-to-multipoint LSP is based on the transmission of RSVP_RESV messages from the tree leaves toward the “root” router (case of router S in Fig. 6). The router directly connected to the “root” router sends a single RSVP_RESV message which contains the set of information specific to the sub-LSP paths that have been established.

The establishment of point-to-multipoint LSPs (and the associated resource reservation process) with the help of the RSVP-TE protocol leads to the maintenance of a significant number of states which is unsurprisingly proportional to the number of subscription/withdrawal procedures to a multicast group.

The current works propose to address this issue by decomposing the RSVP states characteristic of the establishment and maintenance of a point-to-multipoint LSP. To each RSVP “sub-state” corresponds a sub-group of branches and leaves of the tree, so that each of these sub-states can be refreshed independently and so that the routers can graft themselves to the tree (or withdraw from the tree) through an incremental modification of the general RSVP state corresponding to the point-to-multipoint LSP to which the router had just been connected.

E. Several Design Approaches

The forwarding of IPTV flows along the point-to-multipoint LSP can then reflect several design approaches:

- Establishment of a point-to-multipoint LSP per broadcast service and per source S, a service being characterized by a channel (S, G):
- Establishment of a point-to-multipoint LSP unique per source, or the maintenance of a (*, *) FEC per source. By analogy with the multicast transmission mode, this would correspond to the maintenance of as many shared trees as the number of addresses of multicast groups G used by a source to send the multicast traffic;
- Aggregation of several channels (S, G) within the same point-to-multipoint LSP tree, or a FEC corresponding to a group {(S, G)}.

Each of these options has advantages and drawbacks. The use of a single tree for a given source has the advantage that it reduces the number of states to maintain, but it is exposed to the risk of wasting the resources according to the interest of the IPTV customers to receive such or such traffic and not another one. A possible use of such design may be to broadcast the IPTV programs which are the most popular, i.e. the ones for which there will always be receivers.

On the contrary, the establishment of a number of point-to-multipoint LSPs equal to the number of channels broadcast by a given source has the disadvantage that it has to maintain a large number of states, which may affect the performances of the network and the services supported. However, this option enables an optimization of resources in accordance with the requirements of quality of service expressed by each of these services. Broadcasting thematic IPTV programs may be an example of such a design.

IV. CONCLUSION

The combination of DiffServ mechanisms and MPLS traffic engineering capabilities addresses the need for hard guarantees as far as the delivery of IPTV services is concerned, hence solving the NRS problem that was introduced in Section II. The question of assigning the responsibility of calculating the distribution tree to the routers of the network (distributed approach) or to a centralized entity remains open. The complexity of such calculations is amplified by the need to take into account the qualities of service yielding several heuristics.

In addition, the design of DiffServ-inferred multicast distribution trees needs to make sure that only the termination branches and leaves actually deserve different colors: for the sake of scalability and depending on both the IPTV service and networking environments (importance of the mesh, number of accessible IPTV channels, etc.), care should be taken by network designers so that the core of such trees remains common to an IPTV traffic-specific service, whatever the access capabilities of the receivers hence restricting DiffServ-based differentiation to the access regions.

Finally, there is still work to be done on the standardization front (e.g. on the mechanisms for dynamic MPLS tree computation), but early deployments of MPLS-based design (with static establishment of point-to-multipoint LSP, [15]) confirms, if need be, the active involvement of the community of IPTV service providers in that area.
REFERENCES


Christian Jacquenet graduated from the Ecole Nationale Supérieure de Physique de Marseille. In 1989, he joined the national directorate of France Telecom where he was in charge of the specification and of the technical support related to the deployment of the first inter-networking service offerings of France Telecom. After several years of involvement in research and standardization activities within France Telecom, he’s now the Director of the Strategic Program Office for the evolution of the IP backbones operated by France Telecom’s Group. He authored and co-authored several Internet drafts and RFC in the field of dynamic routing protocols and provisioning techniques, several papers in the field of traffic engineering and automated production of IP services, as well as books that discuss QoS and dynamic provisioning techniques.

Julien Maisonneuve holds a PhD from University of Paris. He started researching distributed systems in INRIA before joining Alcatel-Alsthom laboratories where he worked on real-time and reliability issues, participating in the specification of Real-Time and Fault-Tolerant CORBA. He successfully led a project on router reliability that remains a valuable feature of current products. He is now Corporate Standardization Manager for Alcatel-Lucent, overseeing the applications and platforms domain, and is heavily involved in the field of IPTV and Mobile TV. He was co-chairman of the ITU-T IPTV Focus Group Architecture subgroup, architecture board member in the OMG and remains TISPAN liaison officer in the DVB forum.