This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the ICC 2008 proceedings.

Protocol Structure Overview of QoS Mapping over Satellite Networks

Mario Marchese, Maurizio Mongelli
DIST - Department of Communication, Computer and System Sciences
University of Genoa, Via Opera Pia 13, 16145 Genova, (Italy)

***Abstract*-The paper deals with protocol architectures to support mapping of *Quality of Service* (QoS) between protocol layers of telecommunication networks. The Technology Independent – Service Access Point studied to this aim within the ETSI committee is summarized and analyzed. Inherent QoS mapping operations introduce the generalization of the regular concept of *equivalent bandwidth* (EqB). Some performance evaluation is proposed to highlight EqB dimensioning.**

I. INTRODUCTION

Protocol stacks in telecommunications networks are composed of functional layers. *Quality of Service* (QoS) provision depends on the performance achieved at each layer and is based on functions performed at layer interfaces. Having in mind the OSI paradigm, QoS derives from reliable physical and link layers that can offer specific transport services to the upper network layers. The flows generated by the network layers (or bundles of them) are forwarded down to a physical interface that transports the information along a channel.

Even if network layer implements efficient QoS mechanisms (IP IntServ, IP DiffServ, MPLS), it is topical that layers below can assure connection to the channel with specific degrees of performance. Otherwise the implementation of complex QoS mechanism is useless. As a consequence, the QoS requirements must flow vertically and need to be received and satisfied by all the layers of the protocol stack. More specifically, the link layer must implement appropriate mechanisms to support the Service Level Agreement (SLA) defined at the network layer. In some cases, in particular in wireless environments, the link layer acts in cooperation with the physical layer through the application of specific cross-layer design solutions. The interaction between the layers in this context is called here "QoS Mapping". It leads to some technological problems that nowadays constitute open areas of standardization and research.

The paper proposes an insight into QoS Mapping issues from different viewpoints. A detailed analysis is reported for the protocol architectures necessary to support QoS Mapping. Recent results of the European Telecommunications Standardization Institute (ETSI) are considered in detail.

The remainder of the paper is organized as follows. The remainder of the paper is organized as follows. The next two sections introduce the technological elements of the ETSI architecture at both user and control plane levels. Section IV specifies the abstraction methodology used to hide the local implementation of the QoS within the satellite core. Section V outlines the technological problem inherent to this solution. Performance analysis is proposed in section VI to highlight how bandwidth dimensioning is a hard task in the presence of QoS mapping operations. Conclusions are summarized in section VII together with possible directions of future research.

II. THE CONCEPT OF TECHNOLOGY INDEPENDENT SERVICE ACCESS POINT

The protocol stack considered here is the ETSI BSM (*Broadband Satellite Multimedia*) architecture [1]. It is related to satellite communication, but it may have a wider application. This is the motivation to choose it as reference for the investigation of QoS mapping.

The considered protocol stack separates the layers between *Satellite Dependent* (SD) and *Satellite Independent* (SI). The interface between SI and SD is defined as *Satellite Independent – Service Access Points* (SI-SAPs). QoS requirements must flow through SI-SAPs and be implemented at SD layers. An appropriate set of primitives (called *Satellite Independent Adaptation Functions*, SIAF, and *Satellite Dependent Adaptation Functions*, SDAF) is defined to support resource reservation invocation at the different levels of the protocol stack. It allows decoupling responsabilities of resource control for each independent component of the system. A detail description of the approach is reported in the following.

Some issues are topical when traffic is forwarded from SI to SD: the change of encapsulation format, the possible need of aggregating traffic with heterogeneous performance requirements and satellite channel fading counteraction. On one hand, there is the need that SD layers provide a service to the SI layers, but, on the other hand, it should be done with the minimum information SD-SI bi-directional exchange. Ideally, the exchange should be limited to the performance requirements and its matching or not to simplify the structure of SIAF and SDAF primitives. The problem is therefore connected to automatic bandwidth adaptation. SD layer needs to compute the exact bandwidth to be assigned at SD buffer (i.e., the service rate) so that the performance requirements fixed by SI layers can be satisfied.

In the next sections, the key features of the architecture are analyzed, thus outlining the inherent problem of automatic bandwidth tuning.

III. ETSI BSM PROTOCOL ARCHITECTURE

A. User plane

The terrestrial portion of the network interconnects with the satellite portion through an earth station or Satellite Station (ST). The QoS paradigm on the terrestrial portion is supposed here to be DiffServ [2]. The presence of IntServ is implicitly considered since DiffServ flows may be the result of some IntServ over DiffServ aggregations performed before entering the SI-SAP. ETSI considers the application of IntServ directly at the SI-SAP [3], too. The mapping problems envisaged in the following are topical for both [2] and [3]. For the sake of simplicity, only the architecture of [2] is outlined in details.

DiffServ classes are classified using DSCP field. DiffServ management at BSM means that DSCP are mapped to SD queues. Since SD classes are system dependent, there is the need to abstract from the lower layers and to provide the SI layers with a common and agnostic interface. This concept leads to the definition of Queue Identifiers QIDs.

QID represents an abstract queue at the SAP level. Each QID is formally a relationship between IP queues and SD queues. Each QID defines a class of service for transfer of IP packets into the SD core. The application of QID principle allows hiding to SI layers the local implementation of the QoS within the SD systems, where, for instance, DVB may be applied. A QID is "virtual" in that it represents how a specific subset of IP queues are mapped over a specific subset of SD queues. For example, all Assured Forwarding (AF) queues may be aggregated together in a single SD queue or traffic shaping may be applied to IP flows before being forwarded to the SD queues. The application of shaping or other policing depends on the SLA between the satellite and terrestrial networks providers. On the other hand, the QID is "real" because it characterizes the real properties of the SD queues, for example, in terms of available SD resource reservations.

Fig. 1 depicts the relationship between traditional DiffServ IP queue management, which is performed above the SI-SAP, and the lower layers queues (SD queues and QIDs). The relationship is shown for user plane on the left side of Fig. 1 and for control plane on the other side. IP-to-QID mapping means the definition of which outgoing flow from IP queues is forwarded down to a specific subset of QID queues. The same definition holds true for QID-to-SD mapping.

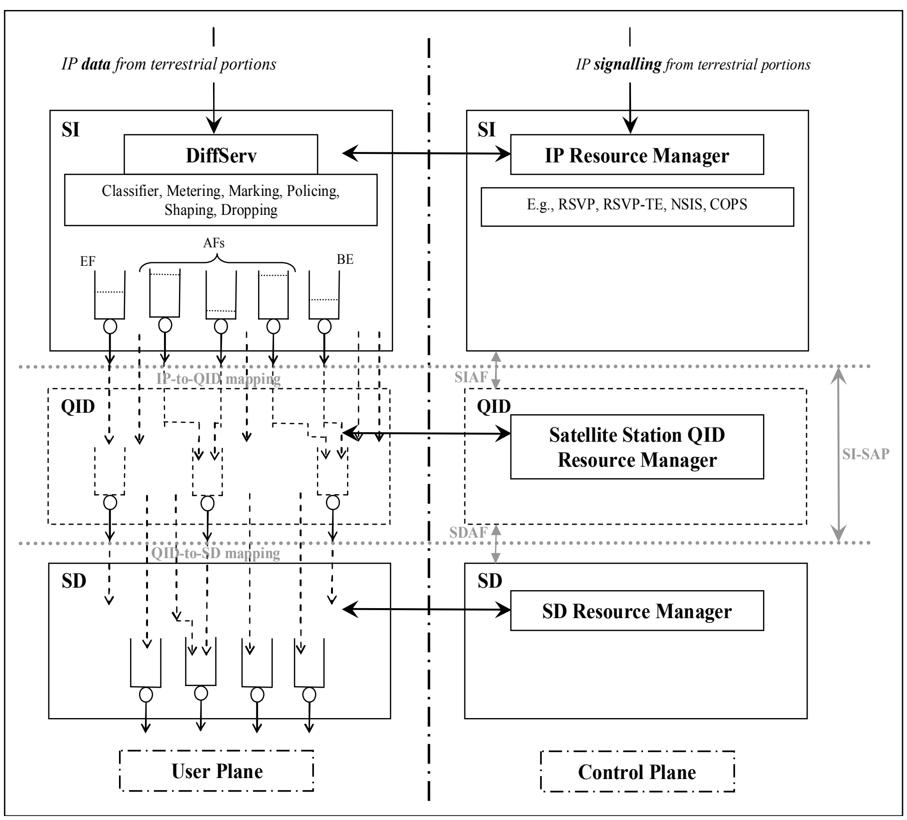

Fig. 3. ETSI SI-SAP protocol architecture.

B. *Control plane*

The control plane consists of one module above the SI-SAP (the IP resource manager), one module below (the SD resource manager), and one interfacing module (the Satellite Terminal QID Resource Manager, STQRM), which, through the set of SIAF and SDAF primitives described in [4]:

1) receives requests from the IP module to allocate, to release or to modify resources,;
2) translates these requests into QID allocation/release/modification actions;
3) checks with the SD module how the upper layer requests can be mapped to real resources.

The control plane may be statically or dynamically configured.

Static configuration. No specification to dynamically reserve resources or to receive indications of network resource availability may be implemented on the terrestrial portion of the network. In this case, user traffic entering the SI-SAP is distinguished through DSCP and served through static trunks. The term static denotes the manual management of network resources over large time scales. For instance, the pre-allocation of a bandwidth trunk for a specific traffic class. Neither traffic prediction nor real time reaction to congestion, including Call Admission Control (CAC), is provided. In this case, also the SD resources may be statically configured. The QIDs present to higher layers static resource reservations. Obviously, the SI static trunks need to be mapped on the private SD system in dependence of the specific IP-to-QID and QID-to-SD mappings implemented.

Dynamic configuration. Another approach consists of designing a dynamic control plane. It means triggering resource reservation on the basis of time varying SI traffic and satellite channel conditions. Thus, the control plane is based on some signalling scheme. RSVP may be used on the SI side. Usually, RSVP is applied in conjuction with IntServ. In case aggregation operations like IntServ over DiffServ are applied directly at the SI-SAP, RSVP is explicitly interfaced with DiffServ BSM [3]. So, SI-SAP can react to RSVP requests by allocating resources based on aggregated SI flows using regular RSVP agents as in common IntServ routers. The IP resource manager in the ST is therefore responsible for:

- translating RSVP requests into QID allocation requests;
- forwarding the QID allocation requests to the STQRM;
- receiving the reply from the STQRM;
- sending appropriate RSVP signalling.

In turn, the STQRM, which received a QID allocation request, performs the following operations:

-check for available SD resources to accommodate the request;
-if resources are available, make the required changes in the IP-to-QID or QID-to-SD mappings, and notify the IP resource manager;
-if SD resources are not available, request allocation of new resources to the SD resource manager, wait for response, make the required changes in the IP-to-QID or QID-to-SD mappings, and notify the IP manager.

It is worth noting that comparable operations may be performed if other signalling schemes are applied outside the SI-SAP. Possible examples may be RSVP-TE, if the SI-SAP interfaces MPLS networks, or extensions made on RSVP to support CAC information in DiffServ environments (called RSVP - Pre Congestion Notification) or priority information of the calls (Emergency RSVP). The Common Open Policy Service (COPS) may be also used as interfacing protocol between terrestrial and satellite portions to define policy-based SLAs. Future evolutions of the Next Steps In Signalling (NSIS) IETF working group may be of interest for the SI-SAP standardization, too.

IV. QID MANAGEMENT

Besides the possible tipologies of dynamic service invocations, the topical point is how the invocations are mapped over the SI-SAP control plane. The key mapping operation is QID management. QID management is the function that identifies the SD queues in play, the SD resource reservation, modifies the properties of the abstract queue that is associated with those queues, and makes the association with SI queues. In brief, QIDs interface the IP queues to a set of available SD queues. SD layers are responsible for assigning satellite capacity to the SD queues, and thus to the abstract queues. These operations occur below the SI-SAP, so they are hidden to the SI layers.

Also the SD capacity control may be static or dynamic.

A. *Static Resource Allocation*

According to the SLA and a pre-configured set of DiffServ queues, a consistent IP-to-QID mapping is created and statically maintained. No signalling is required across the SI-SAP. The QID-to-SD also has to be static. The way in which QIDs are (statically) associated to SD queues is left to the implementers and to the specifics of each SD technology.

B. *Dynamic Resource Allocation*

QID modification is supported by two possible cases:

1) *IP manager action*. Requests about allocating or releasing resources may come from SI layers through explicit signalling or by monitoring the IP queue occupancy. The requests may introduce changes of the IP-to-QID mapping or of the QID QoS characteristics (called QIDSPEC). The request is forwarded down to the STQRM, which, if accepted, adapts the QIDs to the new needs of the SI layers and, in turn, sends feedback to the SD manager if needed. The set of DiffServ queues remains unchanged.

2) *STQRM action*. Also without explicit messages coming from SI layer, QIDs can adapt to the time varying incoming IP traffic by means of feed-forward information received on SI queues state.

Also the QID-to-SD mapping can be either static or dynamic. It is static in the case the number of QIDs remains fixed, but other QID characteristics may change, namely, either the IP-to-QID mapping or the QIDSPEC. In any case, the SD resource manager must allocate/deallocate reservations accordingly. The QID-to-SD mapping or the related reservations are dynamic in the case the number of QIDs and/or of SD queues changes. For example, in the presence of time varying channel degradation conditions. Note that QID-to-SD mapping does not need to be signalled across the SI-SAP because QIDs hides SD state to the upper layers.

V. QOS MAPPING OVER BSM TECHNOLOGY

Dynamic QID management introduces a complex resource control problem for the SD manager. On the basis of the problem solution, the SD manager may update the SD queues reservation state. It may also trigger feedback up to the STQRM. For example, in case the SD manager is not able to support the desired level of performance. In turn, the STQRM may perform its specific reallocation decisions and/or forward feedback up to the IP manager.

Considering the bandwidth dimensioning viewpoint, the following operations lead to the generalization of the regular concept of equivalent bandwidth. The operations are:

1) Change of information unit. It is the consequence of IP traffic transport over a BSM SD portion that implements a specific technology, e.g., ATM as often done in industrial systems [5], or DVB.

2) Heterogeneous traffic aggregation. The due association of IP QoS classes to SD transfer capabilities is limited by hardware implementation constraints. As outlined in [6], "it is accepted in the BSM industry that at the IP level (above the SI-SAP interface) between 4 and 16 queues are manageable for different IP classes. Below the SI-SAP these classes can further be mapped into the satellite dependent priorities within the BSM which can be from 2 to 4 generally" (see also [7]). The problem is how much bandwidth must be assigned to each SD queue so that the SI IP-based SLA is guaranteed.

3) Fading counteration. Finally yet importantly, many transmission environments, as well as satellite an wireless link, needs to tackle time varying channel conditions due to fade.

VI. QOS MAPPING OVER BSM TECHNOLOGY

In the presence of problems mentioned at points **1)-3)** above, optimal SD bandwidth dimensioning faces the mentioned technological problems (change of information unit, heterogeneous traffic aggregation, fading counteration). A possible solution was firstly introduced in [8]. The algo-rithm of [8] is used here to give an example of bandwidth shift arising at the SI-SAP interface due to the mapping operations reported above. The presence of two traffic buffers at SI layer is considered without fading deg-radation at SD physical level. The first buffer offers a VoIP service. Each VoIP source is modeled as an exponentially modulated on-off process, with mean on and off times (as for the ITU P.59 recommendation) equal to 1.008 s and 1.587 s, respectively. All VoIP connections are modeled as 16.0 kbps flows voice over RTP/UDP/IP. The IP packet size is 80 bytes. The required end-to-end performance objective of a VoIP flow for ITU P.59 is end-to-end loss below 2%. The SLA here is . The second one is dedicated to video service. Different video traces, taken from the web site referenced in [9], are used. Video data are H.263 encoded and have an average bit rate of 64 kbps as well as a peak bit rate of 340 kbps. SI rate allocation (240 kbps) for video (fixed off line through simulations analysis) assures , which composes the video SLA (SI video buffer size is 75000 bytes). Both the outputs of the SI buffers are conveyed towards a single queue at the SD layer. DVB encapsulation (header 4 bytes, payload 184 bytes) of the IP packets through the LLC/SNAP (overhead 8 bytes) is implemented in this case. SD buffer is of 500 DVB cells.

Fig. 2 depicts the increase in percentage (with respect to the SI layer) of the bandwidth provision at the SD layer necessary to guarantee the SLAs outlined. Different video traces are applied (again taken from [9]). Video data are H.263 encoded and have an average bit rate of 260 kbps as well as a peak bit rate ranging from 1.3 to 1.5 Mbps, depending on the specific trace. The SI rate allocation for video is around 350 kbps (depending on the video trace in play) to assure the mentioned QoS thresholds.

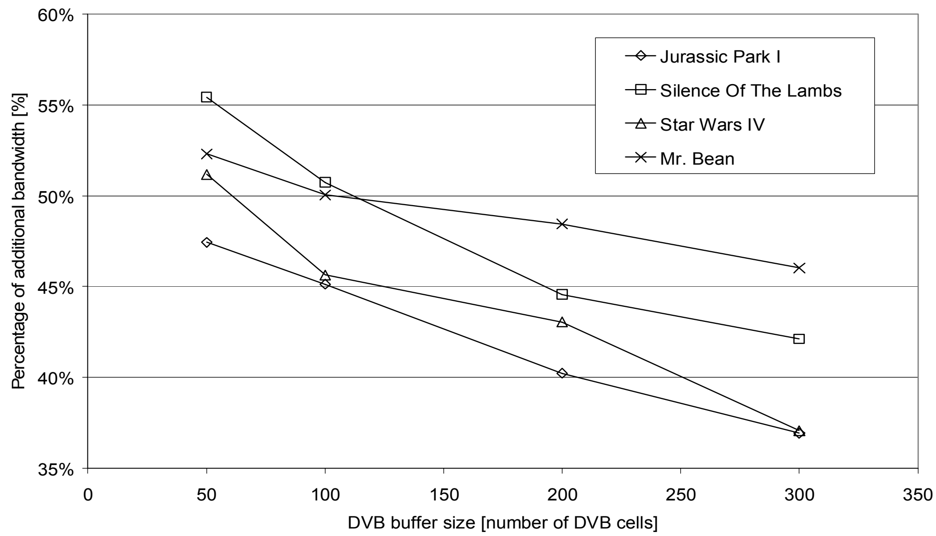

Fig. 2. Bandwidth increase at SD layer measured by algorithm in [8].

The exact SD rate allocation is obtained through the algorithm in [8] and guarantees the maintenance of the required SLA for VoIP and video ser-vices within the DVB portion, by tuning the bandwidth properly. As men-tioned in first part of the paper, the QoS guaranteed over network portions external to the SD core should be guaranteed at SD level. For example, the SLA between providers may consist of the guarantee of the assurance of the same levels of QoS achieved in each portions, despite technologies changes (in this case from IP to DVB). Actually, the QoS support must be implemented in each portion as if technology changes were transparent to end users. In this perspective, RCBC reveals to be a suitable instrument for network planning, too. From the operative viewpoint, if a SD static trunk must be provisioned for the aggregation case presented, the bandwidth provision of the traffic entering the SD queue must be increased up to 50% in order to assure the SI QoS. It is worth noting that the 50% of bandwidth increase holds true in the worst case, namely, with 50 DVB cells in the SD queue. For other buffer allocations, the increase is lower and depends on the specific trace in play. The procedure may help the network operator to measure resource reservation sensitivity as a function of video statistical behaviour. The results may be used to deploy a static trunk serving the ag-gregation of VoIP and video flows. Overprovisioning may be applied by considering the worst statistical behaviour, thus simplifying resource res-ervation over a large time scale.

REFERENCES

[1] ETSI. Satellite Earth Stations and Systems. Broadband Satellite Multimedia Services and Architectures: QoS Functional Architecture. Technical Specification, Draft ETSI TS 102 462 V0.4.2, Jan. 2006.

[2] ETSI. Satellite Earth Stations and Systems, Broadband Satellite Multimedia Services and Architectures: Interworking with DiffServ QoS, TS 102 464 V0.4.1, Sept. 2006.

[3] ETSI. Satellite Earth Stations and Systems, Broadband Satellite Multimedia Services and Architectures: Interworking with RSVP-based QoS (IntServ), TS 102 463 V0.4.2, Oct. 2006.

[4] ETSI. Satellite Earth Stations and Systems. Broadband Satellite Multimedia Services and Architectures: Interworking with IntServ QoS". Technical Specification, Draft ETSI TS 102 463.

[5] E. Lutz, H. Bischl, J. Bostic, C. Delucchi, H. Ernst, M. Holtzbock, A. Jahn, M. Werner, "ATM-based multimedia communication via satellite," Europ. Trans. on Telecomm., vol. 10, no. 6, Nov.-Dec. 1999, pp. 623-636.

[6] ETSI. Satellite Earth Stations and Systems (SES). Broadband Satellite Multimedia. Services and Architectures; BSM Traffic Classes. ETSI Technical Specification, TS 102 295 V1.1.1, Feb. 2004.

[7] ETSI. Technical Committee Satellite Earth Stations and Systems. ETSI Meeting n. 19, Sophia Antipolis, France. Material source: ESA project "Integrated QoS and resource management in DVB-RCS networks", June 2004.

[8] M. Marchese, M. Mongelli, "On-line Bandwidth Control for Quality of Service Mapping over Satellite Independent Service Access Points," Computer Networks, vol. 50, no. 12, Aug. 2006, pp. 1885-2126.

[9] http://www-tkn.ee.tu-berlin.de/research/trace/trace.html.