Solutions and Techniques for the Provision of End-to-End IP QoS in a GMBS System

C. Tocci1, P. Conforto1, A. De Carolis2, M. Femminella2, F. Pugini3

1Alenia Spazio, Italy
2D.I.E.I., University of Perugia, Italy
3University of Rome “La Sapienza”, Italy

Abstract: The present paper intends to provide an overview of the solutions and techniques defined in the framework of SUITED project in order to support Internet QoS sensitive mobile services over a satellite and terrestrial complemented system also referred to as Global Mobile Broadband System (GMBS).

The designed target system architecture consists of: the EuroSkyWay (ESW) Ka Band satellite system; the GPRS (General Packet Radio Service) land mobile system; the UMTS (Universal Mobile Telecommunication System) system representing the GMBS target solution aiming at complementing the GPRS system; a W-LAN system directly accessing the satellite termination nodes, i.e. the satellite Fixed Earth Stations (FESs), to prolong the satellite link in both indoor and short range outdoor environments; a terrestrial Internet network suitably upgraded with both mobility and QoS support capabilities.

The final objective of the proposed GMBS system is to allow a generic user, provided with a GMBS Multi Mode Terminal (GMMT), to access Internet QoS guaranteed services while moving across any possible user environment.

I. End-to-End IP QoS Architecture

Starting from the IntServ/RSVP [Wroclawsky97] and the DiffServ [Blake98] Internet QoS models, that essentially have complementary features, the approach followed in the framework of SUITED project to provide end-to-end IP QoS over the GMBS system, is based on a hybrid IntServ-DiffServ architecture. The latter, see Figure 1, foresees that in the GMBS system edge portion, consisting of the wireless segments and the edge subnetworks of the terrestrial Internet network, the RSVP protocol is implemented, whereas in the core portion of the terrestrial Internet network, where scalability is a stringent requirement, the scalable DiffServ model is adopted.

Nevertheless, as the DiffServ model, unlike the IntServ one, does not provide hard QoS guarantees, a new distributed admission control paradigm, called GRIP (Gauge&Gate Reservation with Independent Probing) has been devised to transparently operate over DiffServ networks so that strict QoS assurances can be provided while maintaining scalability. GRIP is a pure end to end distributed protocol operation which relies the decision to admit a new flow upon the successful and timely delivery, through the Internet DiffServ core network, of probe packets independently generated by the edge devices.

In order to deliver end-to-end IP QoS in such a complex scenario which foresees the use of the RSVP protocol over the GMBS wireless networks and over the enhanced DiffServ Internet core portion, two main issues have been investigated in the framework of SUITED project:

1) the IntServ operation over DiffServ network and the relevant service mapping;
2) the mapping between the IntServ/RSVP requests and the underlying capabilities of the GMBS wireless segments.

As far as the first issue is concerned, specific edge devices called Gateways (GWs) are introduced at the interface between the DiffServ region and the IntServ region of the terrestrial Internet network, (see Figure 1).

These Gateways are in charge of executing the mapping of an IntServ/RSVP request into a DiffServ GRIP class and then activating the GRIP protocol which aims at evaluating if adequate network resources are actually available to support the requested QoS.

For what concerns the second issue, each wireless segment, selected to support the data transfer between the Sender application and the Receiver application, in order to satisfy the QoS desired by the application, has to inter-work with the RSVP protocol and activate specific QoS support mechanisms. To this end, the approach adopted in the SUITED system envisages the introduction of a QoS Support Module implemented in the Terminal InterWorking Unit (T-IWU) of the GMBS Multi-Mode Terminal and, at the network side, in the InterWorking Unit of the satellite Fixed Earth Stations (FESs) and in the UMTS/GPRS GGSNs.

The basic function of the QoS Support Module is to perform interworking with the RSVP protocol and, consequently, to trigger segment specific admission control procedures to verify if the currently active wireless network can accept a new connection while assuring the respect of the desired QoS.

Moreover in order to harmonise the degree of QoS offered by the different wireless segments, the QoS Support Module provides additional QoS support functionality, (e.g. traffic shaping, policing, classification, scheduling and congestion control), to those wireless networks which are lacking them. In the following, the key aspects of the IntServ/RSVP operation over the GMBS network wireless portion and over the DiffServ core portion enhanced with the GRIP protocol will be described.

1.1. Provision of IP Integrated Services over the GMBS Multi-Segment Access Network

The considered model for the provision of IP Integrated Services over the GMBS multi-segment access network is the IntServ/RSVP model which includes RSVP signalling with IntServ parameters, IntServ admission control and per-flow traffic control at network elements.

RSVP is an end-to-end resource reservation protocol that requests resources in only one direction and logically distinguishes the role of data Sender from the one of data Receiver.

In the following, in order to give a definition of the RSVP Session Establishment procedures over the GMBS wireless networks it is assumed that the Sender of the application data is the mobile Terminal Equipment (TE) of the GMBS Multi-Mode Terminal (GMMT) and the Receiver of the application data is a Wired Host directly connected to the terrestrial Internet network. Moreover it is assumed that the currently active access network is the ESW satellite one, in any case the following considerations concerning interworking with RSVP and service mapping apply also to the GPRS/UMTS wireless segments.

At the GMMT side the Sender application formulates the Sender_Tspec parameters, describing the traffic the application expects to generate, and also constructs an initial Adspec object. This information is included in the RSVP Path message which is sent by the Sender to the Receiver.

When the T-IWU in the GMBS Multi-Mode Terminal (acting as an RSVP capable router) receives a Path message coming from the Terminal Equipment, it:
- updates Adspec objects by exporting the ESW error terms C and D and the local values for the general characterization parameters;
- sends to the ESW Segment Specific Mobile Terminal (ESW SS MT) a Session Establishment Request message to activate a ESW Connection Set-Up procedure.

It is important to note that the Connection Profile of the satellite connection established to support the Path message can not refer to the desired QoS as the latter is requested by the Receiver only after that it receives the Path message (the RSVP protocol is receiver-oriented). Therefore when a RSVP Path message has to be transmitted from the GMMT to the Wired Host, the satellite connection is created requesting a best effort service class; such a connection will be used to support all the RSVP control messages exchanged both during the RSVP set up phase and while the connection is going on.

As soon as the satellite Connection Set-Up procedure has been successfully completed, the Path Message generated by the Sender is carried over the satellite segment towards the Receiver.
As the Receiver (i.e., the RSVP module in the Operating System of the Wired Host) receives the Path message, the data in the Sender_Tspec object and Adspec object are passed across the RSVP Application Program Interface (API) to the application which interprets the arriving data and uses it to generate the Flowspec parameters. These parameters are included in the RSVP Resv message which is transmitted upstream towards the Sender.

When the T-IWU receives the RSVP Resv message, the T-IWU performs the IntServ-ESW Service Mapping and sends to the ESW SS MT a Session Establishment Request message to set up a satellite connection with the Connection Profile resultant from the Service Mapping.

If the T-IWU receives, from the ESW SS MT, a Session Establishment Confirmation message containing a positive response, the former forwards the RSVP Resv message towards the Sender which can start to transmit data, otherwise an RSVP ResvErr message is sent by the T-IWU to the Receiver Node which originated the RSVP request.

The message sequence chart diagram of the described RSVP Session Establishment procedure is shown in Figure 2.

![Fig. 2: RSVP Session Establishment procedure over ESW satellite network](image)

The critical point of this procedure is the execution of the IntServ-ESW Service Mapping function that occurs when an IP-level reservation (Resv message) triggers the ingress point of the ESW network to translate the RSVP service requirements into ESW Virtual Circuit semantics. This mapping function is executed by the QoS Support Module of the T-IWU or by the QoS Support Module of FES depending on if the Sender of application data is the mobile TE of the GMMT or a Wired Host respectively.

In both cases the QoS Support Module is in charge of (i) extracting, from the RSVP Resv messages, the RSVP Flowspec parameters containing the requested Service Class (Guaranteed Service or Controlled Load), Traffic Parameters (Tspec) and QoS Parameters (Rspec); (ii) mapping the requested Service Class onto the ESW Service Category and translating the RSVP Traffic Parameters and QoS Parameters onto the ESW ones.

It is worth highlighting that by means of a suitably defined Service Mapping for each wireless network (ESW, GPRS and UMTS) it is possible to provide effective end-to-end QoS for IP traffic that traverses the GMBS multi-segment access network.

I.II. The IntServ operation over the enhanced DiffServ core network

The core portion of the terrestrial Internet network is assumed to operate according to the DiffServ framework enhanced with the GRIP novel solution devised in the framework of SUITED project; it is a mechanism composed of the following three components: (i) GRIP Source Node Protocol (SNP), (ii) GRIP Destination Node Protocol (DNP), (iii) GRIP Routers.

The simplest SNP operation is the following. When a user terminal requests a connection with a destination terminal, the SNP starts a Probing Phase, by injecting in the network in principle just one packet, tagged as probe. Meanwhile, it activates a probing phase timer, lasting for a reasonably low time (from few tens of ms up to few hundreds ms). If no response is received from the destination node before the timer expiration, the SNP enforces rejection of the connection set-up attempt.
Otherwise, if a Feedback packet is received, the connection is accepted, the probing phase terminated, and a data phase is started, consisting in the transmission of data packets. The simplest GRIP Destination Node Protocol (DNP) operation trivially consists in monitoring the incoming packets, intercepting the ones labelled as probes and just relaying with the transmission of a feedback packet towards the source node. At each router output port, the GRIP router implements two distinct queues, one for data packets, and one for probes. Packets may be served according to an Expedited Forwarding priority discipline, i.e. probe packets are transmitted only when no data packets are waiting in the buffer. In addition, each GRIP router measures the aggregate data traffic that it is handling. On the basis of such running traffic measurements, the router implements a proprietary and arbitrarily sophisticated Decision Criterion (DC), which continuously drives the router output port to switch between two states: ACCEPT and REJECT. When a GRIP node is in the ACCEPT state, the Probing queue accommodates Probe packets, and serves them according to the described priority mechanism. Instead, when in the REJECT state, the router discards all the Probe packets contained in the Probing queue, and blocks all new Probe packets arriving. In other words, the router acts as a gate for the probing flow, where the gate is opened or closed on the basis of the DC estimates.

Fig 3 shows the simulation results of a GRIP node loaded with exponential On-Off sources and in which the maximum acceptable number of flows, K, is equal to 100. The number of admitted flows never exceeds the limit K and therefore the agreed service performance is guaranteed. The GMBS reference network, shown in Figure 1, includes a DiffServ region in the middle of a larger network supporting IntServ end-to-end. Therefore, requests for IntServ services must be mapped onto the underlying capabilities of the DiffServ network region. Aspects of such mapping include [Bernet00]: i) selecting an appropriate Per-Hop Behaviour (PHB), or a set of PHBs, for the requested service; ii) performing appropriate policing (including, perhaps, shaping or remarking) at the edges of the DiffServ region; iii) exporting IntServ parameters from the DiffServ region (e.g., for the updating of Adspec); iv) performing admission control on the IntServ requests that takes into account the resource availability in the DiffServ region. The first three functions are carried out by the Gateways (GWs) (see Figure 1). It has been recently discussed, in [Bianchi01], that an AF PHB class, as defined in RFC 2597, is capable of supporting explicit per flow admission control as defined by GRIP principles. The GRIP guidelines compatibility with EF PHB guarantees that, as far as the operation i) is concerned, the GRIP operation mode is robust to the particular service mapping strategy at IntServ/DiffServ borders that will be adopted. The operation iv) is executed by GRIP, which acts as the admission control agent to the DiffServ network region. This can be done according to two possible options, A and B. In option A, the GRIP control loop is executed between the end nodes (Sender and Receiver) and the functionality of the GRIP probe packets is executed by means of the RSVP Path messages. In other words the Path messages carry out both RSVP and GRIP related functions: the GRIP probe is piggybacked on the Path message. In particular, as far as GRIP is concerned, when a Path message (with the added significance of GRIP probe packet) is received by the upstream GW called GW1, the latter device executes the mapping of the RSVP QoS request into a DiffServ GRIP class (and marks the relevant DSCP). It also starts the GRIP Probing Phase by injecting in the DiffServ subnetwork the Probe (=Path) Packet relevant to the selected DiffServ class. If such packet succeeds in reaching the
egress GW, called GW2, then it means that all the involved DiffServ/GRIP routers are found in the ACCEPT state and that the DiffServ region can support the requested connection. The GW2 forwards the Probe (=Path) message to the next RSVP router belonging to the adjacent IntServ region, which continues the RSVP operation, till reaching the Receiver node. If the Receiver node is willing to accept the connection, it answers with a Resv message, which has also the added significance of a GRIP Feedback packet. In the meanwhile, the Sender node has activated a Probing Phase timer. If the Sender receives the Feedback packet, (i.e. the Resv message) before this timer expires, the Probing Phase is successfully completed; in this case, control is given back to the user application which starts a Data Phase, simply consisting in the transmission of data packets, which will then be marked with a suitable DSCP by the upstream GW. This operation is reported in Figure 4 as “Case A”.

In option B, the GRIP control loop is executed between the GWs. In other words, these routers assume the role of GRIP source and destination. The mapping between RSVP requests and DiffServ GRIP classes is executed as above. The probing procedure can be carried out:

1. By keeping on holding the Path message in GW1, starting a GRIP operation between GW1 and GW2 and letting the Path message go ahead through the DiffServ network only if the GRIP operation is successful (that is if GW1 receives a Feedback packet from GW2, within the GRIP timeout). Otherwise, an RSVP error message is sent by GW1 to the Sender node that originated the Path message; this operation is reported in Figure 4 as “Case B1”.

2. By simply forwarding the Path message through the DiffServ network, without GRIP operation; then, when the Resv message is eventually received by GW1, the latter starts a GRIP operation and keeps on holding the Resv message; if the GRIP operation is successful, GW1 forwards the Resv message upstream towards the Sender node. Otherwise, an RSVP error message is sent by GW1 to the Receiver node that originated the RSVP request. In this alternative, GW1 can also exploit the information contained in the Resv message (i.e., requested bandwidth and slack term, in addition to Tspec) to fine-tuning the GRIP request. This operation is reported in Figure 4 as “Case B2”.

Note that the last two alternatives have the con of increasing the set-up delay.

Fig. 4: GRIP Distributed Admission Control

II. QoS Aware Mobility Management Scheme

The proposed GMBS system, by means of its satellite and terrestrial complemented service area, gives the opportunity to provide the users with global mobility capability. To this end a suitable Mobility Management scheme has been devised to allow the GMBS users to access Internet services while moving across different environments maintaining active application sessions without impacts on the perceived QoS.

This implies that the user should not perceive significant QoS degradation due to the change of access segment. Actually the possibility to successfully terminate an Inter-Segment Handover (ISHO) without QoS degradation is strictly related to several conditions: i) the current link does not have to experience an abrupt drop, ii) there should be enough available resources on the new path (consisting of a wireless and a fixed portion), iii) the overall procedure should be fast etc.. Due to the fact that these conditions are not always respected, two different kinds of ISHO can be executed:
Forced handover: taking place when a sudden drop of the link layer QoS (e.g., the radio link is lost due to fading or interference) occurs. In this case, the IP QoS temporarily degrades until a new end-to-end QoS procedure is executed.

QoS Aware Handover: allowing to seamlessly switch the point of attachment to the Internet without any drop in the IP QoS and executed when "old" and "new" access segments are simultaneously available for a while.

The QoS Aware Handover procedure consists of two phases: a "test and reservation" phase and then the actual handover phase where the change of the access segment supporting the data traffic is executed and the IP flow is switched on the new path.

The test and reservation phase is executed to determine if along the new path there are enough resources to guarantee the same QoS provided by the old access segment and, in the affirmative case, to reserve these resources.

This is realised by activating a new RSVP session on the new path and by introducing a new mobility agent, called the Secondary Home Agent (SHA), supporting, at network side, the service which allows QoS IP connectivity to be transparent to mobility issues. The SHA is just a software module running in suitably selected routers placed at the interface between the IntServ region and the DiffServ region of the terrestrial Internet network.

Figure 5 depicts how the RSVP session takes place during a QoS Aware Handover procedure in the case of mobile originated RSVP session start up. Flow diagram is self-explanatory. What is worth pointing out is the role of SHA and of T-IWU which are in charge of executing, depending on the direction of RSVP session, the duplication of the RSVP Path messages on the new path or the selection of the "worst case" RSVP Path message that will require the greater resource reservation.

III. Conclusions

The proposed hybrid IntServ-DiffServ approach, envisaging the implementation of GRIP novel solution, allows a generic GMBS user to access Internet services with effective end-to-end QoS guarantees. The presence of a satellite/terrestrial multi-segment access network imposed specific QoS and Mobility requirements that have been satisfied by defining a suitable InterWorking between the IP-level and wireless segments QoS support mechanisms and by designing a QoS Aware Mobility Management Scheme that allows to preserve the QoS perceived by the GMBS user while moving across different segments.

References


