A Dynamic Service Level Negotiation Mechanism for QoS Provisioning in NGEO Satellite Networks

Tarik Taleb*, Kazuo Hashimoto, Nei Kato, and Yoshiaki Nemoto
Graduate School of Information Sciences Tohoku University, Japan *taleb.tarik@ieee.org

Abstract—Satellite communication systems exhibit important and unique features that qualify them to be an integral part of a global ubiquitous information system. Given the universality of the Internet Protocol (IP), traffic over satellite network is expected to be all IP. Success of these all-IP satellite systems depends on their abilities to guarantee QoS. QoS provisioning has been a hot topic in terrestrial wired networks. It has been, however, highly overlooked in wireless networks. An efficient QoS provisioning in wireless networks in general, and in satellite networks in particular, can be possible only with the development of new schemes that are able to dynamically (re)negotiate service levels in an adaptive manner to changes in network conditions upon handoff occurrences. This paper surveys major dynamic service level negotiation schemes proposed for terrestrial wireless networks and discusses their limitations when applied to satellite networks. As a solution, a dynamic service level negotiation scheme specifically tailored to satellite networks is portrayed. Comparison of the proposed scheme to other dynamic negotiation approaches, via a qualitative and quantitative analysis, is also presented.

I. INTRODUCTION

Because of their extensive coverage area and inherent multicast capabilities, satellite communication systems are seen as an attractive solution for the realization of global ubiquitous information systems. Indeed, an efficient integration of satellite systems with existing terrestrial networks where the former functions as a high speed backbone network to support and/or back up the latter can result in a cost-effective global ubiquitous system. Communications via satellites have first commenced with the use of satellites in geostationary orbits [1], [2]. The need for lower propagation delays and lower terminal power requirements, along with the need for the coverage of high latitude regions, have given birth to Non-Geostationary (NGEO) satellites, called as Low Earth Orbit (LEO) or Medium Earth Orbit (MEO) satellites.

First NGEO systems, such as the Iridium system, were initially designed for the provision of only voice or low data rate services. The recent financial failure of the Iridium systems has, however, made researchers realize that such systems may turn out unfavorable. Along with the worldwide acceptance of the Internet and the exponential growth of bandwidth-intensive services, next-generation NGEO satellite systems should be designed with abilities to provide broadband data rate applications similar in spirit to today’s Internet. In light of the universality of the Internet Protocol (IP) in the Internet, this will derive satellite systems into all IP. A key challenge in realizing the vision of such all-IP satellite systems consists in how to guarantee Quality of Service (QoS).

Without loss of generality, communication via satellites can be either direct or indirect. In the former, terminals communicate directly to satellites, whereas in the latter the communication goes through a gateway. To grant users with full mobility and to play a major role in the design of an efficient global ubiquitous system, next generation satellite systems are expected to directly serve a large population of terminals. This is highly feasible given the ongoing advances in wireless technologies and the resultant massive deployment of cost-effective very small aperture terminals (VSATs) and ultra small aperture terminals (USATs).

For an efficient provision of Internet applications, there is need to ensure QoS. For this purpose, different architectures, such as Differentiated Services (DiffServ) and Integrated Services (IntServ) have been proposed. Most of these architectures are specifically designed for wired networks and are inapplicable, in their current versions, to satellite communication systems for a number of reasons related to the features of satellite systems. Indeed, current QoS architectures are based on centralized and highly static Service Level Agreement (SLA) mechanisms where SLAs are usually agreed by both the client and the service provider when a client signs up for a service. The contractual duration of such SLAs is in a large time scale, typically in order of months or years, i.e., one to three years in case of IP/Digital Video Broadcasting for Satellites (DVB-S).

Given the mobility of users, heterogeneity in wireless technologies, and diversity of user terminals, applying static service level agreement approaches to wireless users may raise unfavorable performance. Indeed, due to users’ mobility, mobile users freely and sometimes frequently change their points of attachment to the network, an operation referred to as handoff henceforth. Upon a handoff occurrence, the amount of resources available at the new point-of-attachment may be different from that of the old one. This disparity of resources can be due to difference in traffic load or due to the use of different wireless access techniques.

Assigning a constant level of service to a mobile user, all the time during its contract period, may lead to an unfair service towards the user. Effectively, upon a handoff occurrence, it is likely that a user is offered a service level higher than what it can be actually provided by the link layer or is bearable by the user’s device. In such an overbooking scenario, the customer will be unfairly charged for a service level he/she can not fully utilize. In case of multiple users from different traffic classes, this unfairness issue becomes more aggravated as the service provider can not fulfill its QoS commitments to all customers.

As a remedy to this issue, a dynamic negotiation of SLA in a small time scale is highly desirable. This dynamic negotiation of SLA should offer users with only what they are seeking for or what is allowable by the current network conditions. This
should be beneficial for both users and service providers. From the customer’s perspective, a dynamic negotiation of service level is beneficial as users will be charged for only what they have actually requested or indeed used. At the service provider side, the system scalability can be improved as savings in the network resources become possible and more users can be then served.

In short, for an efficient QoS provisioning in mobile environments, dynamic negotiation of service level agreement is mandatory. In this context, different dynamic service level negotiation schemes have been proposed for terrestrial wireless networks. Applying such schemes to NGEO satellite networks results in different issues related to scalability and handoff management as will be explained later. This is due to several reasons related to the unique features of satellite systems, namely their long propagation delays and significantly frequent handoff occurrences as both end terminals and the network are on move. There is thus need for a dynamic service negotiation mechanism that is specifically tailored to the needs and features of NGEO satellite networks. This underpins the research work outlined in this paper. It should be stressed out that to the best knowledge of the authors, there has been no previous research work investigating the feasibility of dynamic service level negotiation in NGEO satellite networks.

The paper is organized in the following fashion. Section II highlights the relevance of this work to the state-of-art of dynamic service level negotiation in terrestrial wireless networks. Section III discusses the applicability of current service level negotiation mechanisms to NGEO satellite systems, highlights their pros and cons, and proposes an approach that best suits the features of NGEO satellite networks. A qualitative evaluation of the discussed protocols is presented in Section IV. The paper concludes in Section V.

II. RELATED WORK

For QoS provisioning in IP networks, the Internet Engineering Task Force (IETF) has proposed various frameworks. Differentiated Services (DiffServ), Integrated Services (IntServ) with Resource reservation Protocol (RSVP), and Multi-Protocol Label Switching (MPLS) [3] with Constraint-based Label Distribution Protocol (CR-LDP) [4] are notable examples. Among these architectures, DiffServ is the most scalable and has thus been considered for implementation in different projects, such as the 3rd Generation Partnership Project (3GPP).

Generally speaking, QoS provisioning consists of two major operations: resource allocation and service level negotiation or management. The former attempts to find optimal allocations of network resources to meet the service contract between a client and a service provider. For resource allocation, a large body of research work has been done [5]. In mobile environments, the usefulness of resource allocation hinges on an efficient negotiation of the service level as network conditions change and users’ terminals are diverse. Indeed, in current wired networks a user and a service provider agree verbally or in writing for a given service level. The service level remains static during all the contract period and is changeable only manually after a request from the user. In next-generation wireless communications systems, this scenario is not beneficial neither for clients nor for service providers as stated above. To cope with this issue, several researchers have examined new ways to enable dynamic negotiation of the service level between clients and service providers. A detailed survey on these mechanisms is available at [6].

In the context of dynamic QoS negotiation in satellite networks, the PROQOS group has developed a dynamic SLA management scheme [7] for the provision of DVB in satellite networks. The SLA negotiation occurs between the Space Link Providers (SLP) and their customers, namely Internet Service Providers (ISPs). In this research work, the authors do not consider the case of NGEO satellite networks, nor do they describe the mechanics of QoS negotiation in their considered network topology. Given the limited work in the field of dynamic SLA negotiation in satellite networks, we survey, in the following, major techniques that have been devised for dynamic QoS negotiation in terrestrial wireless networks.

The basic idea behind the Resource Negotiation And Pricing (RNAP) protocol [8], an extension of RSVP, consists in an integration of pricing with resource reservation. It enables users to dynamically negotiate their contracted services with service providers according to their needs and the resources availability of the network. It accordingly adjusts the service fees as well. For this purpose, RNAP maintains a state table at routers along the path of a data flow. To notify routers of their current grades, users are also requested to periodically send signaling messages to routers. These two operations incur significant storage overhead and results in waste of bandwidth and energy at user terminals. The Service Negotiation Protocol (SrNP) [10] is initially developed for wired networks. It functions independently of any SLA format and is applicable to the negotiation of any document in the format of attribute value pairs. SrNP involves exchange of a number of messages for negotiation. These messages are encoded using different techniques. This encoding operation incurs significant overhead and limits the applicability of the protocol to mobile nodes where battery power is scarce.

Common Open Policy Service (COPS) - Service Level Specification (COPS-SLS) [9] enables dynamic negotiation of service level between two ISPs, and an end-user and an ISP. The latter interaction occurs in three phases: client-type establishment phase, configuration initiation phase, and negotiation phase. The system first starts with the client-type establishment phase followed by the configuration phase. During the configuration phase, the end terminal is instructed, via a number of messages, on how to perform the negotiation by being provided information on the negotiation mode, the set of negotiation parameters, and the renegotiation interval. This phase is introduced so as to make the negotiation adaptive to a variety of customers, service providers, and QoS architectures. Upon a successful initiation of the configuration, the client
enters the negotiation phase where the service levels can be negotiated. Given the installation procedure during the configuration phase and the exchanged messages, COPS-SLS results in additional complexity at end-terminals. This fact puts in question its applicability to mobile devices with scarce resources such as Personal Digital Assistant (PDA) devices.

Simple Inter-domain Bandwidth Broker Signaling (SIBBS) [11] is proposed for experimental deployment in the QBone test bed. It manages network resources of different QBone domains by enabling communication between bandwidth broker agents that manage the network resources on behalf of each domain. SIBBS uses long lived TCP connections to aggregate signal messages from all bandwidth brokers. For broker agents of non adjacent networks, a virtual peering between the agents is considered. Similar to COPS-SLS and SrNP, SIBBS requires periodic exchange of signal messages, above all over long lived TCP connections. This consumes both the wireless bandwidth and the battery power of mobile devices, and is thus not suitable for mobile environments.

QoS NSIS (Next Steps in Signaling) Signaling Layer Protocol (QoS-NSLP) [12] is conceptually similar to RSVP. Similar to RNAP, it employs a soft state approach. However it does not carry out the negotiation procedure in a centralized manner. Indeed, negotiation is performed between pairs of adjacent NSLP nodes and not in an end-to-end fashion along the complete signaling path. QoS GANS (Generic Ambient Network Signaling) Signaling Layer Protocol (QoS-GSLP) [13] is also developed on top of the NSIS protocol suite. It reduces the time required for setting a SLS by defining SLSs in advance. This early setting of SLS is done by referring to mobility patterns of mobile users and traffic patterns of the entire network.

Whilst most of the schemes mentioned above demand periodic signaling between network entities for SLA negotiation, Dynamic Service Negotiation Protocol (DSNP) [14] takes into account the wireless environments constraints on bandwidth and power and reduces the frequency of signaling messages. Indeed, in DSNP, mobile users are not required to maintain any continuous TCP connection, nor are they expected to send any additional signaling upon handoff. They are not required to send periodic refresh messages either. For these reasons, DSNP is a light weight protocol and is seen as the most suitable scheme for dynamic mobile environments. The QoS architecture considered in DSNP is based on DiffServ and consists of several domains. Each domain is managed by a QoS Global Server (QGS), an Authentication, Authorization, and Accounting (AAA) server, and a Dynamic Host Configuration Protocol (DHCP) server. Each domain is divided into a number of sub-domains, each managed by a QoS Local Node (QLN). The basic concept behind the working of DSNP consists in an immediate dissemination of the QoS profile of each user, negotiating for a given service, to not only the QLN of the sub-domain where the user is currently located but also to all the QLN s of adjacent sub-domains. This intuitively requires storing state information of users at QLN s. By so doing, when the user enters the area of an adjacent sub-domain, it will be immediately served at the same service level with no need for additional signaling from the user.

Although the above mentioned protocols have the potential of carrying out dynamic service negotiation, they differ in terms of their scalability, their signaling overhead, and the complexity of their operations. Aside from DSNP, all protocols are extensions of other existing protocols. While having protocols extended from existing protocols does not introduce entirely new protocols to the network and accordingly does not cause any further burden to network administrators, it may add complexity to the negotiation operation of the original protocol. In this sense, having a new protocol exclusively designed for only the objective of dynamic service negotiation may be beneficial. Indeed, from the above discussion, it can be deduced that most protocols that are extended from existing protocols require significant signaling and ultimately increase the overall complexity. This signaling consumes the scarce battery power of mobile devices and wastes an important portion of the network bandwidth. With this respect, the two protocols that are seen light weight and more suitable for mobile devices, mainly those with limited resources, are DSNP and QoS-GSLP. In the remainder of this paper, we show that the applicability of these two schemes to NGE0 satellite networks raises significant issues. As a solution we propose a new service negotiation protocol specifically designed for NGE0 satellite networks.

III. DYNAMIC SERVICE NEGOTIATION FOR NGE0 SYSTEMS

This section gives a detailed description of the proposed dynamic service negotiation mechanism for NGE0 systems. Before delving into details of the proposed scheme, we first enlist the key components of the whole architecture on top of which the proposed scheme is designed.

A. Architecture Description

The components of the architecture are schematically depicted in Fig. 1. The figure portrays the coverage area of a NGE0 satellite system. While the NGE0 satellite constellation can serve users via terrestrial gateways, we consider users, mobile as well as fixed, that are directly communicating with satellites in an interactive fashion.

The coverage area is divided into a number of wide domains.
countries). Each domain is administrated by a Global Service Negotiation Manager (GSNM) and an AAA server. The latter is used to verify whether mobile users are authorized to enjoy their requested services whereas the former carries out the service negotiation procedure. Indeed, upon receiving a service initiation/renegotiation request from a mobile user, GSNM uses information about outstanding requests and the availability of resources at satellites to accept or reject requests. Practically, when a request for a service level arrives, the negotiation manager entity decides whether to deny or accept this request. If the request can not be taken without degrading QoS of already existing users or causing network congestion, the manager may deny the request and send an immediate negative acknowledgment to inform the customer that the request has been turned down. Simultaneously, a list of available services can be sent along with the negative acknowledgment to induce the subscriber for a renegotiation of the requested service level. The mechanisms by which GSNMs admit or turn down requests, or allocate resources for a user is outside the scope of this research work. For this purpose, different approaches have been proposed [15], [16].

All GSNMs are assumed to be managed by the same service provider, i.e., operator of the satellite constellation. It should be noted here that while it is possible to manage the service negotiation over the entire Globe at a single GSNM, such a centralized management strategy may incur significant burden on the GSNM and may add significant delay to the service initiation or negotiation. An efficiently distributed architecture of GSNMs is thus more beneficial. Indeed, since GSNMs maintain the QoS profile of all subscribers in their coverage areas, they should be formed according to the geographical proximity and the density of subscribers. Their determination should be also performed in a way that the mechanisms for initiating and renegotiating a service level are reliable and easy to adapt to users’ needs. Above all, the service negotiation mechanisms should be sufficiently fast as they will be applied to environments well characterized by their long propagation delays.

B. Service initiation or renegotiation

At the initial setup of a connection, a subscriber first searches for the satellite it can communicate through and then sends a service initiation request to the detected satellite. The satellite then forwards the message to the GSNM of the domain it is currently covering. Upon consulting its corresponding AAA server, GSNM makes the decision on whether to admit or reject the request. If the requested service level can be guaranteed, the GSNM acknowledges the subscriber of a successful registration of its required SLS. The acknowledgment is sent in a negotiation response message via the same satellite.

In case the requested SLS can not be provided, a negative acknowledgment is sent to the subscriber along with a list of available service levels. The subscriber then notifies the GSNM of its desired SLS via a SLS negotiation request message. A successful negotiation of SLS is acknowledged

![Diagram](image)

(a) Service initiation or renegotiation

(b) Service negotiation upon downlink-to-satellite handoff

Fig. 2. Major procedures of the proposed dynamic service negotiation mechanism.

to the subscriber via a negotiation response message. Once the service initiation procedure is agreed upon, the subscriber starts enjoying the service by exchanging actual data traffic. While enjoying the service, a subscriber may desire to renegotiate a different service level at a later instant. This operation is possible following the same procedure.

Service renegotiation can be intrigued also by the GSNM in a proactive way. Indeed, in case of network congestion, GSNM can offer some privileges to subscribers that accept to downgrade their current service levels. Similarly, if sufficient network resources become available, GSNM can encourage subscribers to join high service levels for better QoS. The major operations of the service initiation and renegotiation mechanism are conceptually depicted in Fig. 2-(a).

C. Service negotiation upon downlink-to-satellite handoff

When a mobile node becomes outside the visibility area of a satellite and performs handoff to a new satellite, an operation referred to as downlink-to-satellite handoff henceforth, the mobile node issues a service negotiation request to the new
satellite. This satellite verifies whether the requested service should be guaranteed by consulting the previous satellite via a SLS confirmation request message. Once the requested SLS is confirmed, the satellite informs the mobile user of the negotiation results via a service negotiation response. Assuming that the mobile user is currently (or soon to be) outside its coverage area and is thus (or will be) served by a different satellite, the previous satellite erases the profile of the mobile user from its database. This mechanism informs satellites when to erase profiles of departing users. Satellites are thus not obliged to eternally store profiles of users. Fig. 2-(b) portrays the major operations of the service renegotiation upon a downlink-to-satellite handoff. While the fact of consulting previously used satellites about the SLSs of newly joining users add two hops delay to the service renegotiation latency, this delay can be minimized by adding an encrypted key to SLSs. Indeed, upon a downlink-to-satellite handoff, a user can send its SLS to the new satellite along with a key. The new satellite uses this key to authenticate the user and the requested SLS. This is similar in spirit to the idea of the distributed system proposed in [17]. In this manner, if resources are available, the mobile user can immediately start enjoying its requested service with no need for confirmation from the previous satellite. Whilst this operation minimizes the negotiation delay, it causes some flaws in terms of security as malicious users (i.e., man-in-the-middle) can crack keys and use them to claim a service level they are not eligible for. Additionally, in order to minimize the effect of the service negotiation delay on the overall handoff delay, mobile users can initiate the service negotiation procedure as soon as they enter the overlapping area between the coverage areas of the two satellites; the currently used satellite and the one to be next used.

IV. ANALYSIS AND DISCUSSION

Having described the details of the proposed dynamic service negotiation scheme for NGEO systems, we now direct our focus on its evaluation. This section first discusses the limitations of the current dynamic service negotiation schemes when they are applied to NGEO systems. It next presents simulation results of the proposed scheme and compares them against results of DSNP.

A. Qualitative analysis

As stated above, different protocols have been proposed for dynamic service negotiation in terrestrial networks. Apart from DSNP and QoS-GSLP, most of these schemes require significant overhead and are thus not suitable for wireless environments, particularly satellite networks with mobile subscribers. Concerning QoS GSLP, it bases its service negotiation on early setting of SLSs which is, in turn, based on the prediction of users’ mobility and network traffic patterns. Whilst this operation is somehow doable in wireless terrestrial networks, it becomes highly complex in case of satellite networks as both end terminals and the network are on the move.

On the other hand, the application of DSNP to satellite networks gives rise to a number of issues. Indeed, in terrestrial wireless networks, whenever a user subscribes for a service level, global negotiation managers implementing DSNP should acknowledge potential QoS local nodes, i.e., the currently serving edge router and the neighboring edge routers, of the new service requirements of the subscriber. This implies that global managers should have knowledge on the topology of the entire network. In satellite networks, this operation is not feasible as both the network and users are on the move. Effectively, as a NGEO satellite can turn over the Earth in a period of time of few hours, a user in a given location can be served by different satellites during its communication time. Therefore, applying DSNP to satellite networks would necessitate dissemination of users’ profiles to all satellites. This implies that satellites should maintain a significantly huge state table on users’ profiles. Obviously, some of these profiles may never be referred to as their corresponding users may never be served by some satellites. In addition, given the fact that in DSNP, there is no mechanism that informs local nodes when to erase profiles of departing users, local nodes have to eternally maintain information on profiles of all users. This puts the scalability of DSNP in question when applied to NGEO satellite networks.

In contrast, our proposed scheme presents solutions to all the aforementioned issues. Indeed, satellites are required to maintain information on SLSs of only users that are currently residing in their coverage areas. This information is removed once users move to the coverage area of a different satellite. Additionally, as newly visited satellites confirm SLSs of users with the satellites they were previously using, GSNMs are not required to maintain a prior knowledge of the satellite network topology, nor do they have to predict its changes during the service time.

B. Simulation results

To illustrate the idea of the qualitative analysis at hand, we conducted a simple simulation using Network Simulator (NS2) [18]. We consider a satellite constellation made of a number of satellites. For the sake of simplicity, we consider a part of the constellation made out of only five satellites. In general, a satellite constellation is considered dynamic through the concept of dynamic virtual topology. It can be thus modeled as a set of time-discrete snapshots of satellite positions over one system cycle, which can be divided, in turn, into a number of time intervals with variable lengths. Over each interval, the topology is considered to be constant; the link state changes take place at only discrete times. To evaluate the performance of the proposed scheme with more clarity, we consider one of these time intervals where the satellite network topology is constant and the satellite motion can be neglected, in other words when inter-satellite link handoff does not occur among the considered satellites. The five satellites are assumed to cover a geographical region that is managed by the same GSNM and the same AAA server. The number of mobile nodes roaming in the coverage area of the five
satellites is varied from five to 400. All mobile nodes are assumed to have already initiated their services. Focus is thus on service negotiation upon downlink-to-satellite handoff. In the performance evaluation, DSNP is used as a comparison term and the proposed scheme is termed as SAT-DSNP.

The simulation results have intuitively shown that the proposed scheme exhibits higher negotiation delays than the DSNP scheme. This result is obvious as the proposed scheme needs to confirm with the previous satellite before granting a mobile node its requested service level. As previously stated, a solution to this issue is possible by adding a key to the service level that can be used by the current satellite to authenticate the issuing user. However, this method can be used if the system is sure that there is no risk of attacks such as man in the middle. While the drawback of the proposed scheme consists in its relatively long negotiation delay, its major advantages consist in its reduced signaling overhead and its shorter SLS storage table as it can be deduced from Fig. 3. Indeed, the figure shows that DSNP necessitates high number of signaling messages and causes the replication of most of the users’ profiles at all satellites. Taking into account the motion of satellites, application of DSNP to satellites systems may end up by replicating the profiles of the majority of users at all satellites. This is ineffective as the database in satellites will be huge. Furthermore, many satellites may never have traffic coming from or going to a given user. On the other hand, the proposed scheme tends to be more scalable as it reduces the number of signaling messages and does not require huge storage tables. Effectively, the reduced signaling of the proposed scheme is intuitively attributable to the fact that the proposed scheme locally negotiates service requirements of users, without involving GSNMs, upon a downlink to satellite handoff, whereas its short SLS storage table is due to the protocols and features of satellite networks. As a remedy to the inefficiencies of these dynamic negotiation approaches in NGEO systems, a dynamic service level negotiation scheme specifically designed for satellite networks is portrayed. Evaluation of the proposed scheme is performed via a qualitative and quantitative comparison of the proposed scheme against former service negotiation protocols. While the presented performance evaluation demonstrated the strengths of the proposed scheme via a small satellite constellation, extension of the simulations to large satellite constellations, along with some mathematical analysis, forms the focus of our future research work.

V. CONCLUSION

Due to users’ mobility and diversity in transmission technologies and user devices, dynamic service negotiation is vital in any wireless communications system, including satellite networks. In this paper, we have surveyed major protocols proposed in the recent literature for dynamic service negotiation in terrestrial wireless networks. We have shown that applications of these protocols to NGEO satellite networks may turn ineffective due to semantics related to the protocols and features of satellite networks. As a remedy to the inefficiencies of these dynamic negotiation approaches in NGEO systems, a dynamic service level negotiation scheme specifically designed for satellite networks is portrayed. Evaluation of the proposed scheme is performed via a qualitative and quantitative comparison of the proposed scheme against former service negotiation protocols. While the presented performance evaluation demonstrated the strengths of the proposed scheme via a small satellite constellation, extension of the simulations to large satellite constellations, along with some mathematical analysis, forms the focus of our future research work.

REFERENCES