

Selection of QoS Monitoring Points in a New QoS Control Framework for SIP-Based NGN

F.Weber^{1,2}, W.Fuhrmann³, U.Trick¹, U.Bleimann³ and B.V.Ghita²

¹ Research Group for Telecommunication Networks, University of Applied Sciences Frankfurt/M., Frankfurt/M., Germany

² Centre for Information Security and Network Research, University of Plymouth, Plymouth, United Kingdom

³ University of Applied Sciences Darmstadt, Darmstadt, Germany
e-mail: weber@e-technik.org

Abstract

This paper introduces a concept for the selection of QoS (Quality of Service) monitoring points in SIP-based NGN (Session Initiation Protocol; Next Generation Networks). It is assumed that the majority of effects influencing the QoS of a single media session (such as a VoIP session (Voice over IP)) also has an effect on the QoS experienced by coexistent media sessions held by other subscribers under similar conditions (such as if subscribers share the same access network, and comparable access and session parameters). An Artificial Neural Network is intended to be used for determining the comparability of sessions, and hence, of user terminals, regarding their QoS conditions. User terminals affected by comparable QoS conditions are dynamically assigned to virtual groups, each representing the QoS conditions effective for its respective members. Hence, from monitoring network performance parameters at one terminal associated with a particular group, conclusions can be drawn on the QoS experienced by other group members. In order to minimise the network traffic resulting from both the querying for and the communication of QoS conditions experienced by user terminals the concept does not only consider the time already elapsed of a session to be potentially monitored but also the mean duration of sessions held by the subscriber associated with a potential monitoring point.

Keywords

Session Initiation Protocol, Next Generation Networks, Quality of Service, Monitoring

1. Introduction

Recently, the concept of NGN has become widely accepted within the field of fixed and mobile telecommunications. SIP as an IP-based signalling protocol has been emerged as the major signalling protocol for NGN-based telecommunication networks.

Generally, the NGN concept can be outlined by several main key features (Trick and Weber, 2007; ITU-T Y.2001, 2004), one of which is the provision of Quality of Service (QoS). Unfortunately, as shown in (Weber *et al.* 2007) the NGN QoS architecture defined by ETSI TISPAN (European Telecommunications Standards

Institute Telecoms & Internet converged Services & Protocols for Advanced Network) in (ETSI TS 185 001, 2005) is not scalable and does not work efficiently regarding the network resources required for both QoS provision and control.

Within our previous research work new requirements to QoS provision in SIP-based NGN have been defined (Weber *et al.* 2007). Based on these requirements tasks have been derived (Weber *et al.* 2008) to be performed by an optimised NGN QoS framework. Subsequently the fundamentals of this framework have already been defined and introduced. Section 2 of this paper provides a brief framework overview. One of the main tasks to be fulfilled by this framework is the continuous evaluation of QoS conditions effective for both ongoing and upcoming communication sessions. In order to perform this evaluation in a scalable and efficient way a concept has been developed that minimises the traffic resulting from both the transmission of information on effective QoS conditions and the related signalling. Section 3 of this paper describes several aspects of this concept, resulting in the selection of adequate QoS monitoring points. Section 4 provides related conclusions and an outlook on the next steps regarding the implementation of the selection concept.

2. Framework for comprehensive QoS control in SIP-based NGN

2.1. Framework overview and components

In order to satisfy the general requirements on the provision of QoS in SIP-based NGN denoted in (Weber *et al.* 2007) the “Integrated framework for comprehensive QoS in SIP-based NGN” has been defined. This framework aims to fulfil the following tasks.

- Task 1: Evaluation and near future prediction of QoS conditions effective for sessions held by subscribers of the respective NGN
- Task 2: Advanced Admission Control - Integration with admission control for SIP-based services
- Task 3: Exertion of influence on QoS conditions within the NGN’s transport network (QoS control)

Figure 1 presents a typical NGN transport infrastructure based on IP with the framework for comprehensive QoS control being implemented. The user end systems (Users A, B, ...) are connected to the access networks, each of which is linked to the core network. A centralised SIP session control function is directly connected to the core network. If required other networks (such as NGN operated by other providers, or other types of telecommunication networks such as the ISDN (Integrated Services Digital Network)) can be connected to the NGN.

If an NGN subscriber wants to initiate a media session with another user SIP messages are exchanged between their related user end systems via the SIP Call Server. Once the session has been established media data are exchanged in a peer-to-peer manner over the IP transport infrastructure between the respective end systems.

The centralised session control infrastructure is typically not involved in the media data exchange.

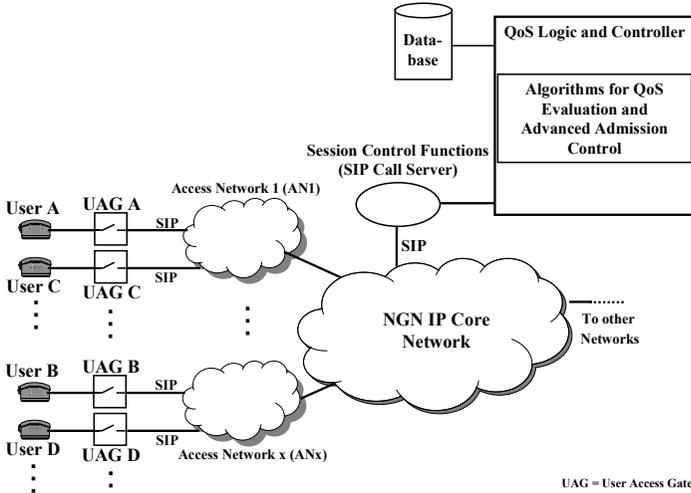


Figure 1: Exemplary NGN after framework implementation (Weber and Trick, 2008)

To provide an NGN with the integrated framework for comprehensive QoS control, the following logical entities are added.

- **User Access Gates (UAGs):** This entity can be located either at the subscribers’ residence or at the border of the access network. It works as a mediation entity for all data exchanged between a user end system and the network and covers the functionalities shown in Figure 2. UAGs are SIP-aware, and are able to monitor QoS conditions (such as jitter and packet loss progressions) related to sessions held by the respective subscriber. By the use of the SIP message SUBSCRIBE a UAG can be queried to communicate QoS-related information using SIP NOTIFY messages to the QoS Logic and Controller via the SIP Call Server. Note that the UAG must be trusted by the SIP service provider.

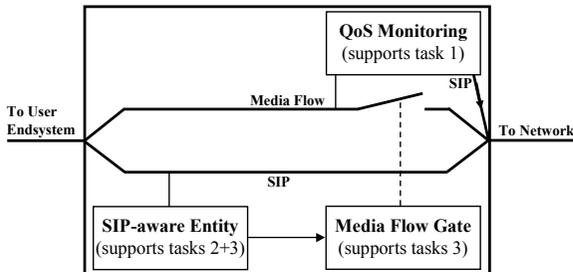


Figure 2: User Access Gate (UAG) block diagram (Weber and Trick, 2008)

QoS Logic and Controller (QoS L&C): This centralised entity is provided with bidirectional interfaces to the SIP session control function and to a database. It covers the functionalities shown in Figure 3. Note that, in order to analyse and control QoS, the QoS L&C depends on information provided by both the Call Server and the database. On the other hand the QoS L&C provides the SIP Call Server with admission decisions and uses the Call Server as an interface for querying and receiving QoS information, provided by subscriber’s UAGs upon request.

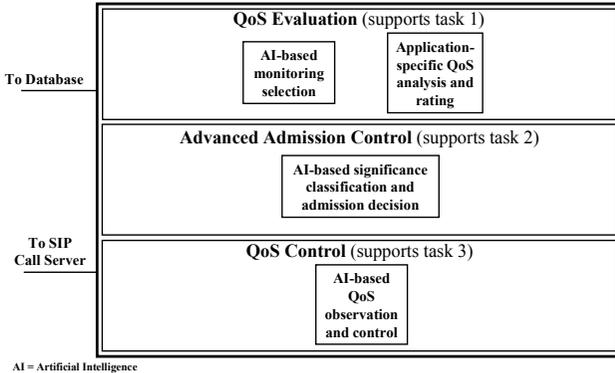


Figure 3: QoS Logic and Controller (QoS L&C) block diagram (Weber and Trick, 2008)

2.2. General framework assumption

The following assumption is made regarding the definition of the framework described within this chapter. Two subscribers A and C, both provided with IP connectivity with comparable conditions (such as identical uplink and downlink bandwidth) to the same Access Network 1 (see Figure 1) experience similar QoS conditions for media sessions established with two other subscribers, B and D, connected to another Access Network x, providing that both sessions have in common the same type of medium and codec (such as two VoIP sessions, both encoded with the G.711 audio codec). In this case, in order to collect information on the QoS experienced by subscribers A and C for incoming media data, it is sufficient to obtain QoS-relevant information (such as jitter and packet loss rates) of only one of these two sessions (note that QoS information can be obtained by querying the UAG of the respective subscriber). This information is likewise meaningful regarding the statement on the QoS experienced by both subscribers.

2.3. Framework functionality

As stated in section 2.2, information on the QoS experienced by a number of subscribers can be obtained from querying the QoS Monitoring entity of only one UAG for QoS-related information such as jitter and packet loss rates, providing that the QoS experienced by the respective subscribers under similar conditions is subject to the same influences (such as the same routing path within the transport infrastructure). In order to minimise the bandwidth required for the collection of QoS

information, UAGs to be monitored have to be selected well-considered (AI-based monitoring selection performed by the QoS L&C, as pictured in Figure 3). Note that the procedure of selecting the UAGs to be monitored is further described in section 3 of this paper. Furthermore, the QoS information determined for all active media sessions is analysed and rated by the QoS L&C subject to the respective media and codecs (the application-specific QoS analysis and rating block from Figure 3).

In order to provide every NGN subscriber with the best possible service in respect of the effective QoS conditions (while considering both user policies and the service availability formerly experienced by a respective subscriber) the Advanced Admission Control entity of the QoS L&C (see Figure 3) classifies all ongoing and requested sessions by their respective objective significance. Both the significance ranking and the QoS conditions effective for the respective subscriber are considered within the admission decision for new sessions requested. Possible outcomes of the admission process are rejection, granted as-is, or granted under downgraded conditions (using a lower-bit rate codec). The Advanced Admission Control can also decide to reject or downgrade competing media sessions that are objectively less relevant. In any case the SIP signalling originating from the SIP Call Server is recognised by the concerned UAG's SIP-aware Entity (see Figure 2) which ensures that the subscriber end system follows the signalled directives. If required the UAG's Media Flow Gate (see Figure 2) is controlled by the SIP-aware Entity regarding bandwidth limitation or media flow cut-off.

The QoS L&C's QoS Control functionality (see Figure 3) is in charge of reacting on debasing QoS conditions affecting already existing media sessions. It reacts in order to maintain or recover satisfying QoS conditions for as much high priority media sessions as possible. Like the Advanced Admission Control functionality, the QoS Control functionality can decide that objectively less relevant competing media sessions can be rejected or downgraded, respectively, based on the ranking list of session significance.

3. Selection of QoS monitoring points

3.1. General prerequisites for the selection of monitoring points

The following prerequisites exist for the selection of a UAG as a QoS monitoring point.

- The monitoring of QoS conditions should result in an exhaustive overview of QoS conditions effective for any existing and upcoming session between subscribers connected to any connection point of the respective NGN.
- In order to minimise the network traffic required for both QoS information querying and exchange as less UAGs as possible should be monitored simultaneously.
- Because interchanging monitoring points causes signalling traffic, a UAG once selected as a monitoring point should perform this task as long as possible.

3.2. Introducing virtual UAG groups

As mentioned in section 2.2 it is assumed that a number of NGN subscribers having the same network access conditions in common will also experience similar QoS conditions. In order to identify subscribers experiencing the same QoS conditions each subscriber's UAG has to be associated with a virtual UAG group which represents the respective QoS conditions. An UAG group is defined by the development of certain measurable network performance parameters (such as packet jitter and packet loss rates) effective for media data received from defined sources. Hence, once a UAG joins the NGN for the first time (or after relocation of the UAG, respectively) it is allocated to a certain UAG group by comparing its jitter and packet loss characteristics with the reference characteristics effective for each respective UAG group within a given time frame. In order to consider the imperative of real-time processing, because value dispersions and uncertainties are to be expected (amongst others, resulting from the need to compare time-line based progressions of values obtained from discrete measurements), the use of Artificial Intelligence (AI) is preferred over mathematical analysis for the comparison of jitter and packet loss progressions. According to (Görz *et al.*, 2003) Artificial Neural Networks (ANN) in general can be used to organise input vectors into classes ("classifying"), or to associate input vectors with output vectors, respectively. (Fortuna *et al.* 2001) describe that ANN can tackle the problem of classification by learning to distinguish characteristics (such as individual jitter and packet loss progressions) of elements (such as UAGs) to be assigned to specific classes (such as UAG groups). Hence, an ANN learns to identify the correct membership class of any given element by its characteristics. That is why the application of ANN has been considered useful for the allocation of UAGs to UAG groups. If a monitored UAG shows jitter and packet loss progressions which can not be satisfyingly allocated to any existing UAG group a new UAG group is created. The respective UAG becomes the first member of the new-defined group. On the other hand, if the jitter and packet loss characteristics of two or more UAG groups converge the groups can be merged. In order to validate the once performed allocation of each UAG to a respective UAG group affiliation checks should be performed periodically.

Once a UAG is associated with a UAG group statements can be made regarding the QoS experienced by the related subscriber for communications with a dedicated media data source.

3.3. Evaluation of QoS conditions

Generally an overview on QoS conditions must be obtained for sessions among all connected subscriber end systems. Hence, QoS conditions must be monitored for connections among UAGs of all UAG groups in either communication direction.

As shown in Table 1, assuming that four different UAG groups (group 1 ... group 4) have been instantiated, 16 different communication situations among all UAG groups potentially coexist (media senders associated with UAG group 1: receivers associated with UAG group 1, 2, 3, and 4, respectively. Media senders associated with UAG group 2: ...).

In order to obtain a sufficient overview on the QoS conditions experienced by any subscriber at least one session for each potential communication situation must be monitored. Thus, referring to Table 1, if four UAG groups exist, the QoS conditions of at least 16 media sessions have to be continuously monitored at the same time (providing that all communication situations are given at any time), regardless of the number of UAGs (and hence, subscribers) associated with each respective UAG group.

Group number of UAG group the media senders are associated with	1	2	3	4
Group numbers of UAG groups the media receivers are associated with	1	1	1	1
	2	2	2	2
	3	3	3	3
	4	4	4	4

Table 1: Potential communication situations with four UAG groups identified

Note that, because the number of sessions that have to be monitored rises quadratic with the number of UAG groups, the number of UAG groups should be minimised.

3.4. Relevance of monitoring points

In order to activate the monitoring function of a UAG during an ongoing session, additional SIP signalling is required. To increase the efficiency of the monitoring process, the relevance of each active UAG as a monitoring point should be rated at any point in time.

Generally a UAG typically involved in long-lived sessions should be considered more relevant than UAGs typically involved in short-lived sessions in order to minimise the amount of signalling per time required for monitoring point selection. Hence, the mean session duration of sessions per UAG should be explicitly taken into account regarding the relevance calculation. Furthermore the ratio of (overall session duration to be expected (represented by the already experienced mean session duration)) over (already elapsed session time) has to be considered as a factor of relevance regarding the selection of monitoring points. That is why the relevance of a UAG as a potential monitoring point can be defined as follows.

- $R_{UAG_x} = MSD_{UAG_x} * (MSD_{UAG_x} / EOS_{UAG_x}) = (MSD_{UAG_x})^2 / EOS_{UAG_x}$
- R_{UAG_x} = (Relevance of UAG_x as a QoS monitoring point)
- MSD_{UAG_x} = (Mean duration of sessions in which UAG_x has been involved)
- EOS_{UAG_x} = (Elapsed time of ongoing session in which UAG_x is involved)_{UAG_x} > 0

Table 2 shows an exemplary relevance calculation for a defined point in time (“snapshot”) for UAGs associated with a certain UAG group (UAG group 1) of an NGN with three UAG groups existing at this time.

UAG Group 1					
UAG Identifier	In session with subscriber associated with UAG group no.	MSD [s]	EOS [s]	R	Relevance position (for monitoring QoS of incoming media data associated with UAG group no.)
UAG 1-A	1	126	32	496,1	1 (group 1)
UAG 1-B	2	45	164	12,3	
UAG 1-C	3	264	562	124,0	1 (group 3)
UAG 1-D	-	437	-	-	
UAG 1-E	1	146	278	76,7	
UAG 1-F	2	81	26	252,3	1 (group 2)
UAG 1-G	3	196	526	73,0	

Table 2: Exemplary calculation of monitoring point relevance

Within Table 2 every row represents an independent UAG (UAG 1-A to UAG 1-G) associated with UAG group 1. For UAGs that are currently involved in a media session the second column provides the number of the UAG group that the related communication partner is associated with. Both the mean session durations (MSD) and, if applicable, the elapsed time of ongoing sessions (EOS) are provided and the relevance as a potential monitoring point is calculated for each UAG. The last column indicates the position in relevance ranking of the respective UAG subject to the UAG group number the respective communication partner is associated with. As an example, UAG 1-C is the most relevant potential QoS monitoring point associated with UAG group 1 for media sessions with subscribers associated with UAG group 3 at the defined point in time that Table represents.

Note that the selection of UAGs providing a high relevance as monitoring points results in longer monitoring periods per UAG, and hence, prevents from a high frequency of monitoring point interchanging.

The relevance calculation for a respective UAG is assumed to be performed generally independent of the affiliation of related communication partners to a specific UAG group. If required this relationship can be taken into account additionally. As a result the relevance of a UAG as a potential monitoring point could be determined by the UAG group affiliation of the respective communication partner.

If required incoming and outgoing session initiations can be treated separately within the relevance calculation regarding the mean session duration.

3.5. General selection rules and variants

The selection of UAGs as monitoring points is based on the following general rules.

- If only one media session between subscribers associated with two UAG groups is active the involved UAGs represent the only available QoS monitoring points and, hence, are selected.
- A UAG, once selected as a QoS monitoring point, communicates to the QoS L&C information on the experienced QoS while the related media session is active.
- If a media session of a UAG acting as monitoring point is terminated and further media sessions exist between subscribers associated with the related UAG groups, new QoS monitoring points for the related communication situation have to be assigned.
- UAGs are typically chosen as QoS monitoring points based on their respective relevance (see section 3.4).

As a variant of the selection of monitoring points a set of exceedingly relevant UAGs within a certain UAG group could be preselected as QoS monitoring points. These UAGs are queried in advance to share QoS information of all media sessions in which they will be involved in future with the QoS L&C. If the QoS information obtained from the collection of preselected monitoring points covers communication situations among all active UAG groups no additional UAGs have to be queried.

Preselecting QoS monitoring points can help reducing the signalling traffic from the explicit querying of UAGs for QoS information. If required, preselection can be used in combination with the formerly-mentioned method of per-session selection and in-session querying of monitoring points by their relevance. Depending on the characteristics given in a respective NGN regarding the distribution of UAGs among different UAG groups and the subscribers' communication behaviour this combination can possibly result in a comprehensive overview on the QoS conditions effective among all identified UAG groups and, at the same time, minimise the amount of signalling traffic required for QoS information querying.

4. Conclusions and outlook

The introduced QoS monitoring point selection concept comes as a part of a framework for comprehensive QoS control in SIP-based NGN introduced in (Weber *et al.* 2008). The framework has been briefly described within this paper. In comparison with the NGN QoS architecture standardised by ETSI TISPAN in (ETSI TS 185 001, 2005) the implementation of this framework will result in a more efficient and scalable way to provide trustworthy QoS to NGN subscribers.

The selection concept introduced within this paper allows for the comprehensive evaluation of QoS conditions effective for any ongoing and upcoming media session within a SIP-based NGN. The concept is optimised to be scalable with the number of NGN subscribers by virtually grouping subscriber terminals experiencing comparable QoS conditions. Hence, from only a small subset of terminals information on effective QoS conditions has to be queried, received, and analysed in order to draw conclusions on the QoS experienced by any other ongoing media session.

The next step in order to deploy this concept as a prototype is the definition of both architecture and configuration of a Neural Network required for the comparison of network performance progressions and, subsequently, the allocation of user terminals to virtual groups.

The concept of QoS monitoring point selection introduced within this paper is particularly suitable for integration with the framework for comprehensive QoS control and, as such, is intended to enrich and facilitate the framework's functionalities.

5. Annotation

The research project providing the basis for this publication was partially funded by the Federal Ministry of Education and Research (BMBF) of the Federal Republic of Germany under grant number 1715B05. The authors of this publication are in charge of its content.

6. References

ETSI TS 185 001 V1.1.1 (2005), Technical Specification, “*NGN Quality of Service (QoS) Framework and Requirements*”, ETSI TISPAN

Fortuna, L.; Rizzotto, G.; Lavorgna, M.; Nunnari, G.; Xibilia, M. G.; and Caponetto, R. (2001), *Soft Computing*, Springer, ISBN: 1-85233-308-1

Görz, G., Rollinger, C. and Schneeberger, J. eds. (2003), *Handbuch der Künstlichen Intelligenz*, (Handbook of Artificial Intelligence), Oldenbourg, Munich, Germany, ISBN: 3-486-27212-8

ITU-T Y.2001 (2004), Recommendation, “*General overview of NGN*”, ITU-T

Trick, U. and Weber, F. (2007), *SIP, TCP/IP und Telekommunikationsnetze*, (SIP, TCP/IP and telecommunication networks), Oldenbourg, Munich, Germany, ISBN: 978-3-486-58228-4

Weber, F.; Fuhrmann, W.; Trick, U.; Bleimann, U. and Ghita, B.V. (2007), “QoS in SIP-based NGN – state of the art and new requirements”, *Proceedings of the third collaborative research symposium on Security, E-learning, Internet and Networking SEIN 2007*, pp. 201-214, Information Security & Network Research Group – University of Plymouth, Plymouth, ISBN: 978-1-8410-2173-7

Weber, F.; Fuhrmann, W.; Trick, U.; Bleimann, U. and Ghita, B.V. (2008), “A framework for improved QoS evaluation and control in SIP-based NGN”, *Proceedings of the seventh International Network Conference, INC 2008*, pp. 27-37, Centre for Information Security and Network Research – University of Plymouth, Plymouth, ISBN: 978-1-84102-188-1

Weber, F. and Trick, U. (2008), “Optimizing and simplifying SIP-based NGNs’ QoS architecture”, *International SIP 2008 Conference*, Paris, France, upperside