ARTÍCULOS DE REVISIÓN

Review of Quality of Service (QoS) mechanisms over IP Multimedia Subsystem (IMS)

Revisión de los mecanismos de Calidad de Servicio (QoS) sobre IP Multimedia Subsystem (IMS)

Mauricio Alexander Vargas Rodríguez* , Ernesto Cadena Muñoz**

1* Researcher and Telecommunications Instructor, Instituto Europeo de Posgrado - Universidad CEU San Pablo Business School. MBA with specialty in project management, Universidad Nacional de Colombia Sede Bogotá, Colombia. MSc. Candidate in Telecommunication engineering, Escuela Colombiana de Ingeniería Julio Garavito, Colombia. Electronic Engineer. vrmauricio4@misena.edu.co

2** Researcher and Telecommunications Instructor, Universidad Nacional de Colombia Sede Bogotá. MSc. in Telecommunication engineering, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. Specialist in engineering Project Management and Telecommunication Engineer. eca denam@misena.edu.co

ABSTRACT

The present state of the art establishes the basis for studying different mechanisms to provide end-to-end Quality of Service on a network over the IP Multimedia Subsystem (IMS). Unlike traditional networks which ensure QoS individually depending on the service provided, IMS seeks the convergence of fixed and mobile services with Quality of Service through a set of policies defined by operators and Service Level Agreements established with their users.

Key words: Heterogenous network; IMS; NGN; QoS; SLA

RESUMEN

El presente estado del arte establece las bases para el estudio de diferentes mecanismos para proporcionar la calidad de servicio (QoS) de extremo a extremo en una red sobre el subsistema multimedia IP (IMS). A diferencia de las redes tradicionales que aseguran la Calidad de Servicio de forma individual dependiendo del servicio que proveen, en IMS se busca la convergencia de servicios móviles y fijos con la Calidad de Servicio a través de una serie de políticas que son definidas por los operadores y los Acuerdos de Nivel de Servicio establecidos con sus usuarios.

Palabras-clave: IMS; NGN; QoS; Red heterogénea; SLA

INTRODUCTION

The concept of next-generation networks (ngns) was introduced to face the reality of the telecommunications industry, characterized by open competition among operators due to the deregulation of markets, the convergence between networks and services, and the increasing demand for multimedia applications. This reality creates challenges for operators both in their infrastructure and in their portfolio of services; operators must quickly adopt new technologies and the ability to develop services in a short time, at low cost, to meet market needs.

Growing mobility needs and customizing services provided to consumers have given rise to more efficient ways to provide services using any technology available at any given time; therefore, network infrastructures must provide sufficient resources to offer value-
added services. ngns provide a model that allows the operator to improve the provision of resources to integrate all types of telecommunications services under a single network infrastructure adopting internet protocol (ip); this model is referred to as All-ip 2. Therefore, operators can move from a model of vertical architecture, in which each service they provide has a separate infrastructure (mobile phones, switched telephony, data networks, television networks, etc.) with corresponding types of access, transport, control and application infrastructure, to an independent landscape architecture in which these services are integrated (figure 1).

Figure 1 Horizontal and vertical telecommunication architectures

The telecommunications industry is adopting the IP Multimedia Subsystem (ims) as the reference architecture for the development of all telecommunications services for ngns 4 that are able to support current telecommunications services and that will be useful in the future. One advantage of ims is that it integrates the concept of the convergence of services supported by heterogeneous networks, i.e., networks of different types such as fixed, mobile or the internet. However, ims does not standardize applications but facilitates access to multimedia applications and voice access to different types of terminals and access technologies.

One of the most important aspects concerning the convergence of services is Quality of Service (QoS) because IP networks originally lacked QoS control mechanisms. These networks were designed to provide service delivery without guaranteeing the reliability of the corresponding information, depending on the amount of bandwidth required, especially for services requiring real-time connections (5). Based on recommendation Y.2001 (6) of the International Telecommunication Union (itu), the necessary features were established to ensure end-to-end (e2e) QoS in ngns.

Globally, QoS is currently one of the most researched areas because of its interest to users, operators and regulators 7. ims is an IP-based architecture that is constantly evolving and is heterogeneous 8 in operator networks; it is characterized by a variety of protocols used in operator network infrastructures. Therefore, QoS policy control has become an important research topic 9, 10.

IP MULTIMEDIA SUBSYSTEM

The IP Multimedia Subsystem is a reference architecture for next-generation service provision standardized by the 3rd Generation Partnership Project (3gpp) and introduced into umts Releases 5 and 6 (March 2003). ims allows telecom operators to offer multimedia services, such as voice, data, video and combinations thereof, under the same infrastructure through packet switching networks based on IP 11. ims is considered a subsystem because it is part of a complete network in which other components are required, such as an access network, to fully function as a service deployment system 12.

This subsystem is important because it allows different types of access networks to be integrated regardless of technology or internet services combining fixed and mobile networks 12; however, for the 3gpp ims described from the point of view of mobile operators (those that support new applications), a body of standardization telecommunication members of the European Standards Institute (etsi) called the Telecommunications and Internet Converged Services and Protocol for Advanced Networking (tispan) adds the necessary ims compatible with networks of fixed-operator (convergence) specifications. The flexibility of this architecture allows for modifications and extensions in the subsystem 14, and it also simplifies the application design by harmonizing the ability to obtain session control through the Session Initiation Protocol (sip) 15.

Figure 2 illustrates the ims architecture and its three main levels, 16, which are as follows: the multimedia services layer; the control session layer; and the ip transport layer, in which an ims-based network architecture enables the convergence of different technologies for access networks, such as a fixed network pstn, a broadband xdsl, a wireless lan, or 2G, 3G and 4G cellular mobile networks.
A distinction must be made between "ims core" and "ims" because the ims architecture refers to a "core" or "ims core" (defined by 3gpp) in addition to a non-ims subsystem number (defined by tispan), such as the Network Attachment Subsystem (nass), a Resource Admission Control Subsystem (racs) or the pstn Emulation Subsystem (pes). The central components of the ims architecture are the Call Session (or State) Control Function (cscf) entities, which are actually sip servers. These entities have specific functions for signaling and routing traffic and are located at the core level of the Proxy-cscf (p-cscf), the Interrogating-cscf (i-cscf), the Serving-cscf (s-cscf), the Home Subscriber Server (hss) and the Application Server (as).

Figure 3 shows the different types of components used in the ims architecture; the main elements found in a basic solution are the CSCF entities. Similarly, various functional entities for managing fixed and mobile networks are defined, and some entities are responsible for contacting the transport network to ensure QoS and prevent misuse of the services provided.

Different protocols are used by these components according to their signaling function, such as session control through the Session Initiation Protocol (sip), authentication control via Diameter or policy control by the model Open Policy Service (cops); the components have been standardized by the Internet Engineering Task Force (ietf) for the purpose of adopting open protocols used in the tcp/ip architecture to create a flexible platform that can be scalable and compatible with legacy networks.
of the functional architecture of an ims subsystem can be found in 3gpp 23.228, where interfaces, protocols and applications that can be offered to users under the ims network are defined.

QUALITY MECHANISMS

One of the essential criteria for evaluating a system lies in the measurement of network performance from the points of view of deployment, operation and customer satisfaction. For quality assessment, there are two approaches, Quality of Experience (QoE) and Quality of Service (QoS).

1. Quality of Experience is related to the satisfaction levels centered on the user according to their expectations, perceptions and underlying effect on QoS. The overall effect of these factors has a subjective evaluation that can vary between low, medium or high percentages. These parameters are focused on the user and are of great research interest.

2. Quality of Service provides a set of metrics focused on network performance, which can be technically evaluated and negotiated based on the possible factors that define QoE.

QUALITY OF SERVICE

According to ITU Recommendation E.800, Quality of Service (QoS) is defined as the totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service. ims architecture focuses on ensuring that the applied QoS policies between specific requests for applications (Session Initiation Protocol - sip / Session Description Protocol - sdp) and multimedia flows (Real-time Transport Protocol - rtp) are defined by the network operators according to the type of business and service requirements at the application level that are used to manage the network resources and enhance QoS.

In IMS QoS, control plays an important role through a series of indicators that allow for evaluations according to service type, such as bandwidth, e2e delay (delay), delay variation (jitter), data rate and bit error rate. Within the context of an IP network, a service refers to the overall traffic handling of a client through a particular domain; a service is useful only if it meets the requirements of the end user. The organizations 3GPP and tispan (Telecoms & Internet Converged Services & Protocols for Advanced Networks) define a set of four classes of QoS for transport networks. Corresponding to the ims services, these requirements also apply to the interconnection of networks that are IP-based (ip-can) access networks and ims. These four classes are listed in Table 1.

Table 1 QoS classes for transport networks

<table>
<thead>
<tr>
<th>Class</th>
<th>Key Attributes</th>
<th>Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversational</td>
<td>Responsive to delay variation, limited tolerance to packet loss</td>
<td>Audio/Video conversation</td>
</tr>
<tr>
<td>Streaming</td>
<td>Responsive to but tolerant of delay variation, limited tolerance to packet loss</td>
<td>Video streaming</td>
</tr>
<tr>
<td>Interactive</td>
<td>Responsive to round-trip delay, packets transferred transparently with low bit error rate.</td>
<td>Collaboration, conference</td>
</tr>
<tr>
<td>Background</td>
<td>Insensitive to delay transparently with low bit error rate.</td>
<td>Email, IM, chat</td>
</tr>
</tbody>
</table>

An e2e multimedia session can traverse a series of heterogeneous administrative domains in NGNs, in which the control system of policies must be able to guarantee the QoS resources in all domains involved. Each domain defines its own mechanisms and policies for the provision of QoS depending on the technologies that are accessible to the operator; however, in an e2e session, a mutual negotiation of Service Level Agreements (slas) between the domains involved is required. An sla is a formal contract negotiated between two parties that defines the commitment levels associated with network performance and service responsiveness. The two parties can be a user and an operator or two operators, in which case one takes the role of customer to purchase services from another provider.
Protocols for handling QoS

To provide end-to-end IP QoS at the network level, ietf has defined the reserve model-based Integrated Services (IntServ) model for resources and traffic prioritization based on Differentiated Services (DiffServ) [33], [34]. DiffServ is based on the division of traffic into different classes and the prioritization of these aggregates with a code in the datagram when congestion occurs [35]. The nodes in the network must know and identify the code to prioritize packets [36].

IntServ implements a channel reservation and admission control packets through the nodes in the network using the rsvp (Resource Reservation Protocol); the implementation consists of three types of services, controlled, guaranteed and best effort load [37]. However, there are issues concerning scalability and complexity, which are overcome by implementing DiffServ [38]; all cases of interoperability between operators are based on the use of SLAs, which is an integral part of DiffServ [39].

Policy-based QoS Architectures

To provide QoS in NGNs, the IETF organizations ETSI TISPAN and 3GPP have defined different architectures for the control of data flows. The ietf describes a policy framework through RFC 2753 (39) (Figure 5), which sets policy rules defined as models that become network configurations or devices in an administrative domain. These rules are stored in repositories called Policy Decision Points (pdps) and Policy Decision Functions (pdfs); these repositories recover the rules of appropriate policies in response to requests for policies that are generated by the QoS requirements of services [40]. The Policy Enforcement Point (pep) refers to the point where a server enforces policy admission control and policy decisions in response to a transaction request from a user who wants to access a resource on a server network.

The ETSI TISPAN architecture developed its own QoS assurance through the standard ETSI ES 282 003 [41]; the architecture is called the Resource and Admission Control Subsystem (racs, Figure 6), which is a subsystem responsible for NGN control element policy, resource reservation and admission control. It is the main component that interacts with the access network and the core network that carries a service that can affect the priorities of the packets through the DiffServ protocol and book resources with RSVP [42].

racs is the logical element of the largest network that allows for interaction between the service layer, the transfer functions of the control resources and QoS support within the respective NGN [43].
3gpp is defined in ts 23.203 and specifies one policy-based QoS called PPC (Policy and Charging Control, figure 7) from the standpoint of the mobility architecture that was introduced in 3GPP R7. The entity PCRF (Policy and Charging Rules Function) is responsible for policy and charging control; the Execution Policy and Charging Function (pcdf) associates reference points to them. The concept and the QoS architecture used by the 3gpp specification are described in 3gpp ts 23.107. The concept of end-to-end QoS architecture used by the 3gpp specifications is described in 3gpp ts 23.207. The pcc architecture is flexible and applicable to a variety of services, access networks and load models; however, it is not well adapted to multimedia service requirements that are dynamically negotiable with QoS according to changing conditions in the network.

**QUALITY OF SERVICE MANAGEMENT**

To implement the various policies of quality of service management architectures described above, it is necessary to make a study of the parameters involved in ensuring the selected service according to 3gpp ts 23.107, 3gpp ts 23.203 and 3gpp ts 23.207.

**Traffic Policing**

This mechanism refers to the packets that are discarded by limits or operating policies, dropping packages that have the lowest priority while being sent. This mechanism is used in routers and is applied to ip packets on input and output interfaces, depending upon the following variables according to the rfc 2698:

1. PIR (Peak Information Rate): maximum transmission rate of a client in bits/s and previously agreed between customer and operator with any type of contract or service level agreement (sla). The PIR can never be greater than the capacity of the line.
2. CIR (Committed Information Rate): average rate of long-term traffic that the operator undertakes to provide a customer with a contract or service level agreement (sla). This parameter is measured in bits/s and is generally less than pir. In any case, the cir can never be greater than the PIR.

3. CBS (Committed Burst Size): Maximum burst size allowed on the network. Specifies the maximum number of bytes that can be transmitted to the PIR, while complying with the agreement of the CIR.

4. PBS (Peak Burst Size): similar to CBS but defined with respect to PIR instead of regarding the CIR parameter.

Traffic Shaping

Traffic shaping uses a policy based on queueing and subsequent extraction of packets to maintain the rate of the traffic. Unlike traffic policing, the original characteristics and traffic delays disappear due to the queue. It is widely used by service providers and users to always guarantee the contracted bandwidth.

Quality of Service Attributes related to Policy Decision Point

The Release 99 of 3gpp recommendations defines QoS attributes such as Traffic class, Delivery order, SDU format information, SDU error ratio, Maximum SDU size, Maximum bit rate for uplink, Maximum bit rate for downlink, Residual bit error ratio, Transfer delay, Traffic-handling priority, Allocation/retention priority, and Guaranteed bit rate for uplink and Guaranteed bit rate for downlink.

QoS parameters related to policy control

The following mechanisms are managed by the pcrf to ensure the traffic over the lte (Long Term Evolution) and lte-Advanced networks:

- QoS Class Identifier (qci) is a scalar value that controls packet forwarding treatment (for example, scheduling weights, admission thresholds, queue management thresholds, link layer protocol configuration and so on).

Table 2 QCI standardized values

<table>
<thead>
<tr>
<th>QCI</th>
<th>Resource Type</th>
<th>Priority</th>
<th>Packet Delay Budget</th>
<th>Packet Error Loss</th>
<th>Example Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GBR</td>
<td>2</td>
<td>100 ms</td>
<td>10-2</td>
<td>Conversational Voice</td>
</tr>
<tr>
<td>2</td>
<td>GBR</td>
<td>4</td>
<td>150 ms</td>
<td>10-3</td>
<td>Conversational Video</td>
</tr>
<tr>
<td>3</td>
<td>GBR</td>
<td>3</td>
<td>50 ms</td>
<td>10-3</td>
<td>Real Time Gaming</td>
</tr>
<tr>
<td>4</td>
<td>Non-GBR</td>
<td>5</td>
<td>300 ms</td>
<td>10-6</td>
<td>Non-Conversational Video (Buffered Streaming)</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>1</td>
<td>100 ms</td>
<td>10-6</td>
<td>IMS Signalling</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>6</td>
<td>300 ms</td>
<td>10-6</td>
<td>Video (Buffered Streaming) TCP-Based (for example, www, email, chat, ftp, p2p and the like)</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>7</td>
<td>100 ms</td>
<td>10-3</td>
<td>Voice, Video (Live Streaming), Interactive Gaming</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>8</td>
<td>300 ms</td>
<td>10-6</td>
<td>Video (Buffered Streaming) TCP-Based (for example, www, email, chat, ftp, p2p and the like) Typically used as default bearer</td>
</tr>
</tbody>
</table>

1. Allocation and Retention Priority (arp) It is used as an indicator of priority setting processes or modification of an EPS bearer service enabling the network to decide whether to accept these processes depending on available resources.

2. MBR (Maximum Bit Rate) indicates the maximum number of bits delivered to the network or by the network within a period of
DISCUSSION

The central problem is to provide consistent end-to-end QoS to an IP service in such a way that the requested QoS requirements are met when the deployment of a service infrastructure involving two or more operators, being autonomous networks whose administrative domains are managed according to their own policies.

Although operators must agree on QoS requirements for a particular service among a set of IP services, operators do not configure their networks in the same way. On the other hand, operators and service providers have their own internal topologies and QoS mechanisms that depend on their devices and other management requirements that are non-technical, and it is observed that the IMS has become the de facto standard in the deployment of LTE networks, due to its open interfaces for deploying converged services. Further studies should consider this reality taking as baseline the Policy and Charging Control framework defined by 3GPP. That is why it is necessary to develop a guideline that allows operators to maintain a consistent level of QoS interconnections without depending on the complexity of the network.

Similarly, the regulatory aspects regarding the quality of service must be taken into consideration by the regulator. In the case of Colombia, the CRC has been working on issues relating to the adoption of next-generation networks that are managed by an industry group, which aims to promote "...cooperation among all industry players involved in the development and deployment of NGN as well as to track the evolution of the same..." in order to guarantee SLA compliance for the user and address the implications of the adoption of these technologies in terms of the convergence of services and markets.

CONCLUSIONS

The evolution of operator infrastructure has led to the implementation of softswitches that allow migration to an All-IP network. However, the IMS architecture has been dominating industry due to its open interfaces for the deployment of converged services. The main challenge is to provide a consistent end-to-end Quality of Service through an IP service, such that the requested QoS requirements are satisfied when the deployment of a service involves the infrastructure of two or more operators, which are autonomous networks whose administrative domains are managed according to their own policies. Although operators must agree on the QoS requirements for a particular service among a set of IP services, operators do not configure their network devices in the same way because they have their own internal topology and QoS mechanisms that depend on their network devices and other non-technical management requirements.

REFERENCES


2 () telefonica, Las Telecomunicaciones y la Movilidad en la Sociedad de la Información. Albadalejo, 2005. [Links]

3 () P. Podhradsky, "New Multimedia Applications based on IMS NGN Architecure", in Systems, Signals and Image Processing (IWSSIP), 2011 18th International Conference on, pp. 1 - 4. [Links]


9 () M. Jian-jun and X. Zi-ping, "Research of the QoS guarantee system in IMS based on MPLS", in 2011 International Conference on Electronics, Communications and Control (ICECC), 2011, pp. 403-406. [Links]


16 () M. Navarro and V. Rodríguez, "An IMS Architecture with QoS Parameters for Flexible Convergent Services", in Computers...
2.1 IP Multimedia Subsystem architecture 3GPP has considered a layered approach for IMS architectural design. As shown in Figure 1, IMS network comprises transport, session control, and application layers. The signaling plane of the IMS is comprised of several functions: CSCF (Call Ses- Fig.1 IP Multimedia Subsystem logical architecture. The IMS QoS mechanism is largely based on the interaction between IMS and the underlying access network. When an UE hands off among IMSs over hetero-geneous access networks, the QoS parameters have to be re-negotiated between the newly visited IMS and its underlying access network. End-to-end Quality of Service (QoS) signalling flows (Release 6). Jun. 2007. Electrical and Communication Engineering. The concept of IP Multimedia Subsystem (IMS) is open network architecture, introduced for Next Generation Networks (NGN). It supports specter of services for channel and packet switching net works. Main signaling protocol in IMS is SIP Session Initiation Protocol [1], which allows user service managing and mixing media streams within a session. Also, it is universal for any type of user access. Quality of Service (QoS) is one of the main characteristics of any next generation network. In the process of services providing, signaling messages delays occur in nodes of IMS-network. The article ana QoS - Quality of Service. indigoo.com. 3. QoS at layer 2.5: MPLS MultiProtocol Label Switching. § MPLS switches IP traffic flows on layer 2 thus improving network performance (more. throughput). § MPLS combines IP routing (addressing) and fast forwarding of traffic (layer 2 switching). § MPLS LSPs (Label Switched Paths) can be assigned certain QoS (like ATM PVCs). IP forwarding. Active Queue Management is the intelligent drop of packets in Internet routers (congestion control mechanism). When queues become full, packets must be dropped (discarded). AQM algorithms try to drop packets with the goal of alleviating the buffer congestion as quickly as possible. Quality Of Service QOS General Principles. • The QOS is associated with transfer classes : ATM ATCs, MAC MTCs, DiffServ » service classes », • The QOS may be found within several layers of the protocols stack (between the MAC and the application layers) up to the user level (subjective QOS). • The QOS is defined by a. • Multiplexing of AAL2 connections over ATM VC. • Needs of scheduling mechanisms able to satisfy real time constrains. • CAC for radio and AAL2 network resources at the call set up and the HO levels. • Dynamic bandwidth allocation and statistical multiplexing of flows at the ATM level. IP Multimedia Subsystem : this domain is the innovative part of UMTS Rel5. It is expected to support IP multimedia applications between two end users. You can use policy-based Quality of Service (QoS) within Group Policy to set the source port range for the predefined DSCP value in the Teams client. The port ranges specified in the following table are a starting point to create a policy for each workload. Table 1. Recommended initial port ranges. Media traffic type. Client source port range. Protocol. DSCP value.