End-to-End Quality of Service in Next Generation Networks

Lav Gupta
TRAI, New Delhi, India
lavgupta@gmail.com

***Abstract*— Unlike traditional disparate legacy networks, Next Generation Networks (NGN) offer functional separation of services from the underlying transport enabling convergence in many forms. This paper discusses the requirement of QoS and QoE in Next Generation Networks (NGN), as defined by various international standards organizations, in order to set up the background for discussing the requirements in the Indian context.**

Keywords—Convergence, NGN, QoS, end-to-end, QoE, parameters

I. INTRODUCTION

Convergence is manifesting at technology, network, device and service levels where one technology or one network today makes it possible to deliver many services which were possible only with diverse networks [1]. Devices are becoming capable of doing more than what they were initially designed to do. For example, a mobile phone is also an image and video capture and transmission device, Internet access device, a jukebox, an office assistant and a veritable application platform. Next Generation Networks (NGNs) provide the technology to implement convergence. An NGN represents a single transport platform on which the carriage of distinct service types, viz., video, voice, and data "converge", together with new and emerging services and applications[2]. NGN is a paradigm shift in the architectural sense when contrasted with traditional networks. The traditional vertically integrated "stove pipe" networks give way to a network with horizontal functional separation [3]. NGNs provide the service providers ways to reduce capital and operational expenses, enable faster rollout of new and old services and equip operators to compete in a market where voice is not the sole source of revenue. NGN is partly driven by the increased demand for ubiquitous packet based multimedia services in the broadband domain.

II. DEFINITION OF NGN

An NGN liberates access to services from the underlying transport and allows faster development and rollout of new integrated applications for the end user. As they were developed, the core and access have high bandwidth that supports multiple services.

International Telecommunications Union (ITU) and European Telecommunications Standards Institute(ETSI) have integrated work from other standards bodies like Internet Engineering Task Force (IETF), Cablelabs, WiMAX forum and the broadband forum into the NGN framework. The ITU terminology is rather general and applicable to different specific technological implementation.

A. *ITU-T definition of NGN*

As per ITU ITU-T Recommendation Y.2001 (12/2004), an NGN) is a packet-based network able to provide Telecommunication Services to users and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent of the underlying transport-related technologies. It enables unfettered access for users to networks and to competing service providers and services of their choice. It supports generalised mobility which will allow consistent and ubiquitous provision of services to users [4].

B. *ETSI definition of NGN*

NGN is a concept for defining and deploying networks, which due to their formal separation into different layers and planes and use of open interfaces, offers service providers and operators a platform which can evolve in a step-by-step manner to create, deploy and manage innovative services [5].

III. ARCHITECTURAL ASPECTS OF NGN

Traditionally, there were different networks for specific services; for example, PSTN for voice; X.25for data and cable networks for video. The transport, service and control were integrated and designed for the defined service set. A key aspect of NGN is the separation of service aspects from the transfer aspects. It implies that the mechanisms for service provision will no longer need to be embedded with the transfer mechanisms[6][7].

A. *Layers and planes in NGN*

The functional separation in NGN architecture is represented by two distinct strata of functionality as shown in Fig 1.

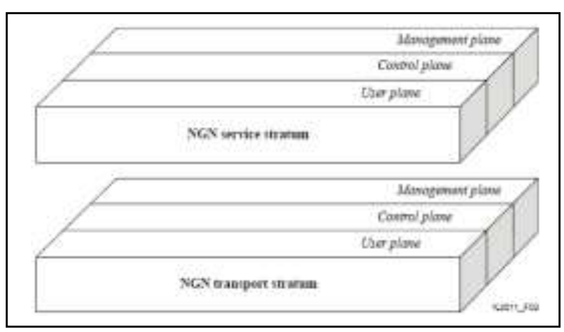

Figure 1 Layer and Planes in NGN Architecture

The service stratum provides the user functions to transfer service-related data and control and manage service resources and network services to enable user services and applications. Services may be developed by recursive use of multiple service layers e.g. voice, data or video application modules could be called in a recursive manner to provide multimedia applications.

The transport stratum provides the user functions that transfer data and control and manage transport resources to carry such data between terminating entities. Dynamic or static associations may be established to control and/or manage the information transfer between such entities. In general, any and all types of network technologies may be deployed in the transport stratum, including connection-oriented circuit-switched (CO-CS), connection-oriented packet-switched (CO-PS) and connectionless packet-switched (CLPS) layer technologies.

Both the service and transport strata have control and management planes. The control and management planes of the two strata could be separate or unified. In NGNs employing unified control or management plane technology, equivalent control or management plane functions at multiple layers can be instantiated in a single protocol.

B. Architectural Framework

The concept of layers and planes is incorporated in the architecture defined by ITU -T Y.2012 (2010-04) and shown in Fig 2[8].

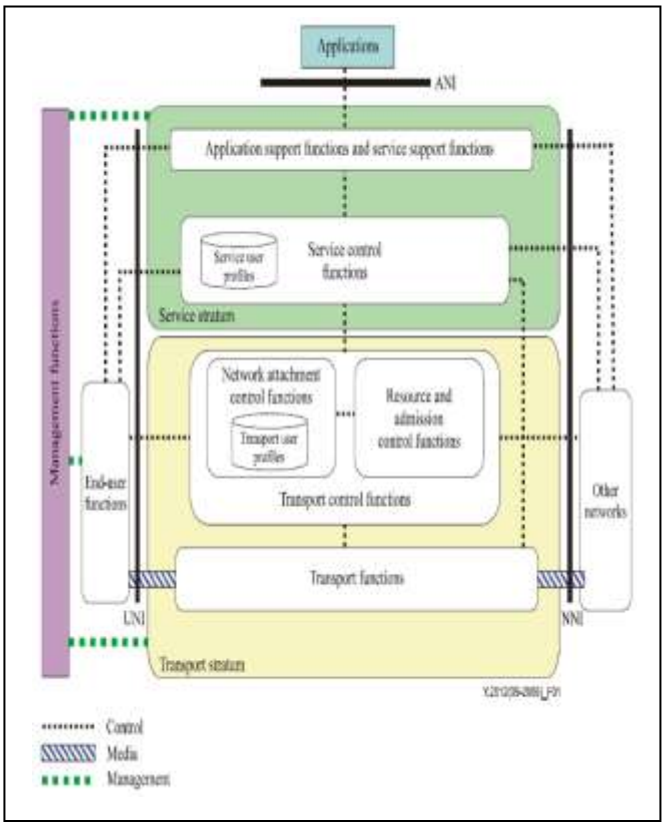

Figure 2 Architecture of NGN

The service control functions (SCF) include resource control, registration, authentication and authorization functions at the service level. They can also include functions for controlling media resources, i.e., specialized resources and gateways at the service-signalling level. The transport stratum provides the IP connectivity services to the NGN users under the control of transport control functions, including the network attachment control functions (NACF), the resource and admission control functions (RACF) and mobility management and control functions (MMCF). The service stratum functions includes – the service control and content delivery functions including service user profile functions and, the application support functions and service support functions. An NGN connects to other networks – IP and legacy, through the network-network interface (NNI). The NNI supports both a control level type of interaction and a media level type of interaction. While it is important to maintain QoS within a single operator domain, for end-to-end (E2E) QoS it is necessary to guarantee secure, QoS enabled inter-operator interconnections [9].

IV. QUALITY OF SERVICE IN NGN

ITU has defined QoS in Recommendations E.800 as "*the collective effect of service performance, which determines the degree of satisfaction of a user of a service".* Expanding on the E.800 QoS concept, ITU-T Rec. G.1000 breaks down service quality into functional components and links it to network performance. According to ITU-T Rec. No. I. 350 Network Performance is measured in terms of parameters which are meaningful to the network provider and are used for the purpose of system design, configuration, operation and maintenance. It is useful to consider the concept of Quality of Experience (QoE) to capture the experience of the user**.** *QoE is defined as the overall acceptability of an application or service, as perceived subjectively by the end-user. Quality of Experience includes the complete end-to-end system effect* [10]

Different services and applications have different requirements of quality that are described by different sets of parameters. Real time voice and video calls are delay sensitive and require steady flow of packets. File transfer may tolerate packet delay but may be sensitive to errors requiring rejection of the entire file in case of errors or retransmission or erroneous packets at the cost of increasing network load and delaying response. Video streaming may be able to tolerate some errors and in any case there may not be enough time for retransmissions. Bandwidth requirements can also vary from real time video to file transfers to voice. While NGN can take advantage of unused bearer periods, they need to be able to correctly determine quality and amount of resources various applications need. The QoS concept has been created to do that and it is defined in such a way that various types of users and service applications with different bearer requirements can be supported in Next Generation Networks.

A network QoS class creates a specific combination of bounds on the performance values. Various existing defined network QoS classes in ITU –T in Y.1541 are give below [11].

TABLE I. IP QoS CLASS DEFINITIONS AND NETWROK PERFORMANCE OBJECTIVES

Parameter	IPTD	IPDV	IPLR	IPER
Performance Objective	Upper bound on mean IPTD	Upper bound on mean IPTD-minimum IPTD	Upper bound on packet loss probability	Upper bound
Class 0	100ms	50ms	$1*10^{-3}$	$1*10^{-4}$
Class 1	400ms	50ms	$1*10^{-3}$	$1*10^{-4}$
Class 2	100ms	U	$1*10^{-3}$	$1*10^{-4}$
Class 3	400ms	U	$1*10^{-3}$	$1*10^{-4}$
Class 4	1s	U	U	$1*10^{-4}$
Class 5	U:Unspecified			

The table below gives guidance for the applicability of the network QoS classes.

TABLE II. GUIDANCE ON CLASS USAGE

QoS class	Applications	Network Techniques
0	Real-time, Jitter, Sensitive, High Interaction (VoIP, VTC)	Constrained routing and distance
1	Real-time, Jitter, Sensitive, High Interaction (VoIP, VTC)	Less constrained routing and distance
2	Transaction data, highly interactive (Signalling)	Constrained routing and distance
3	Transaction data, Interactive	Less constrained routing and distance
4	Low Loss Only (Short transactions, bulk data, Video streaming)	Any route/path
5	Traditional Applications of default IP Networks	Any route/path

Latency [IP Packet Transfer Delay (IPTD)]: time between the occurrence of two corresponding IP packet reference events.
Jitter [IP Packet Delay Variation (IPDV)]: variation in IP packet transfer delay.
Packet Error [IP Packet Error Ratio (IPER)]: ratio of total errored IP packets outcomes to the total of successful IP packet transfer outcomes plus errored IP packet outcomes in a population of interest.
Packet Loss [IP Packet Loss Ratio (IPLR)]:ratio of total lost IP Packets outcomes to total transmitted IP packets in a population of interest.
Toll Quality: The voice quality resulting from the use of a nominal 4-kHz telephone channel.
Call Completion Rate: It is the ratio of established calls to call attempts during time consistent busy hour (TCBH).
Availability of Network: Measure of the degree to which network is operable and not in a state of failure or outage at any point of time for all users.
Latency: Latency or IP Packet Transfer Delay (IPTD) is the time between the occurrence of two corresponding IP packet reference events.
Jitter: Jitter or IP Packet Delay Variation (IPDV) is the variation in IP packet transfer delay.
Packet Error: Packet Error or IP Packet Error Ratio (IPER) is the ratio of total errored IP packets outcomes to the total of successful IP packet transfer outcomes plus errored IP packet outcomes in a population of interest.
Packet Loss: Packet Loss or IP Packet Loss Ratio (IPLR) is the ratio of total lost IP Packets outcomes to total transmitted IP packets in a population of interest.

Toll Quality: The voice quality resulting from the use of a nominal 4-kHz telephone channel.
Call Completion Rate: It is the ratio of established calls to call attempts during time consistent busy hour (TCBH).
Availability of Network: Measure of the degree to which network is operable and not in a state of failure or outage at any point of time for all users.

The basic functions used in most QoS mechanisms are *resource reservation*, *admission control* and *traffic regulation*. To implement resource reservation, a program or service has to specify the type of traffic generated and QoS requirements. The exact method may differ but most QoS implementations define classes of traffic. Admission control decides whether or not a resource request can be satisfied given the current (and possible future) conditions. When a service has been granted access to the network it can commence its file transfer. To implement functionality of regulating traffic during transfer, generally the classes used in the resource reservation are each given a distinct manner of handling traffic, such that high requirement classes get some form of priority over low requirement classes.

While research in many aspects of end-to-end QoS in NGN is in the initial stages, general methods for implementation of QoS have been widely researched [12]. IntServ uses resource reservation and admission control. IntServ, however, requires every node to maintain the flow state which overloads routers and results in scalability issues. Diffserv implements the concept of aggregation of flows. It uses only the edge nodes of a network does all the processing and intermediate nodes need only know some different classes of forwarding behaviour known as per-hop forwarding behaviour (PHB). DiffServ is more scalable than IntServ. There are methods that combines IntServ and DiffServ. EuQoS (European Quality of Service) is one of the newer projects. EuQoS is a path-based QoS solution, chaining several domains together to form a path from receiver to sender.newer projects in the area of QoS. ITU-T QoS work revolves around Resource and Admission Control Subsystem (RACS) while ETSI's around Resource and Admission control Function (RACF). As both the organisations have co-operated in this effort both the systems have a lot of similarities.

V. INDIAN NETWORK MIGRATION AND QOS REQUIREMENTS

A. Status of Migration of Indian Network to NGN

In the Indian telecommunications network service providers have incorporated elements of NGN to varying degree. In some cases there does not seem to be any conscious plan to migrate to NGN. Major telecommunications operators have already implemented Internet Protocol (IP) based core transport network for carrying voice and data traffic. As IP alone does not offer traffic engineering and QoS required by many services some of the telecom service providers like BSNL, Airtel and Reliance and Internet service providers like Sify have deployed MPLS/IP based core transport networks. Following the softswitch approach to migration, the existing fixed line providers sometimes go for replacement of Class 4(transit) and Class 5(local) switches by IP switches that would implement the required functionalities. A major public sector service provider has started a project of introduction of 6 million IP TAX ports with installation of about 32 softswitches and trunk media gateways at 119 locations. The service provider's tentative plan for Class 5 NGN is to deploy 22 million PSTN ports in next 2-3 years.

The Inter Multimedia Subsystem (IMS) route to migration has been slow to pickup but some of the access service providers have indicated that IMS based implementation is on their roadmap. The deployment considerations has so far mainly been economic and replacement is being done to phase out life expired legacy switching equipment. In some cases the changes are being made because of technological evolution and introduction of International Mobile Communications (IMT) Advanced network.

B. Requirement of QoS in the Indian Network

A cross industry NGN expert group consisting of members drawn from the service providers, original equipment manufacturers, academicians, researchers and the regulator was constituted by the regulatory authority. The expert group worked during the period 2006-08 to address various issues related to NGN. The author was part of the expert group first as the representative of the largest PSU telecommunications service provider and then as part of the regulatory authority. The expert group noted that since NGN is a packet based network, quality of service (QoS) in this context is a complex issue and international standards are still evolving. There was complete agreement in the expert group that fixing the framework for QoS at early stages would help the service providers to deploy and engineer their networks and equipment manufacturers to tune up their assembly lines to meet the overall QoS objectives so decided. In general, ensuring QoS would require control of a range of features such as latency, packet loss, jitter and bandwidth but beyond this basic minimum, QoS is meaningful to a user only if it exists on an end-to-end basis i.e. across the networks involved in a session. It was also decided that QoS must cover not just the transmission quality but also parameters like reliability, fault tolerance, service availability, security, call set-up, scalability, service provisioning, service restoration and the like. In a heterogeneous environment, that is likely to exist for a ling time, achieving the desired QoS requires co-operation among elements of disparate networks. If networks do not work on a set of uniform metrics or if the interconnection among networks does not offer adequate QoS then end-to-end delivery of QoS would be impaired.

QoS Parameters are service specific. For real-time voice service, call set-up delay, call completion rate and speech quality are some of the parameters whereas for IPTV service, Jitter and the channel change time could be important parameters. Concerning parameters and values the common opinion was that it will be necessary to identify and to agree on parameters like latency, delay and jitter but there were different points of view on how to achieve it. In its pursuit to concretize the requirements, the expert group felt that QoS requirements should split into two parts viz. network centric and customer centric parameters .

(1) Network Centric Parameters

During the deliberations, the service providers were a divided lot. Those who were in favour IP interconnection and regulatory intervention to standardize quality of service parameters said this should be done by establishing a joint working group. The other who felt that the regulation of QoS should be left to the operators said this was necessary to respond to the dynamics of the environment.

The expert group members felt that QoS was too important a consideration, in provision of various services to the customers, to be left entirely to competitive behaviour. They decided that mandatory quality parameters need to be defined and their expected values specified as has been done for fixed, mobile and broadband services. The expert group members were in favour of IP Network QoS Classes definitions and Network Performance objectives defined in ITU T Y.1541 recommendation to be taken as a standard for deciding the QoS parameters of applicable services in NGN network centric issues in Indian context[*See IV above*].

Following extensive deliberations, the expert group identified the following network centric parameters for E2E QoS:

1) Network QoS Classes: this needs to be done in conformance with ITU-T Y.1541. One example of such a class is QoS Class 0 recommended for real time, jitter sensitive, high interaction services for voice calls.

2) Latency/IPTD for real time/non real time voice, data, video and streaming multimedia services. This should be defined for various classes of service separately.

3) Jitter/IPDV for real time/ non real time voice, data, video and streaming multimedia services. This should be defined for various classes of service separately.

4) Packet Error/IPER for real time/ non real time voice, data, video and streaming multimedia services. This should be defined for various classes of service separately.

5) Packet Loss/IPLR for real time/ non real time voice, data, video and streaming multimedia services. This should be defined for various classes of service separately**.**

6) In case of VoIP, toll quality and non toll quality parameters shall be defined. Customers should be made aware of the difference in Quality and tariff between the two services, by service providers.

7) Interconnection congestion limit should be specified. Some percentage level should be defined for bandwidth utilization.

8) Call Completion rate within network and across networks.

9) E2E QoS across networks in a multi-operator scenario. Apportionment of impairment objectives among operators and number of operators that could be allowed in a particular scenario also needs to be considered.

(2) Customer Centric Parameters

The expert group has identified the following customer centric parameters

a) Service related: Service Activation Time, Service De-activation Time, Service, Restoration Time,

b) Tariffs and billing: Clarity of Tariff Plans, Ease of switching between plans, Ease of getting Billing information, Billing Accuracy, Ease of Bill payments, Ease of getting refund

c) Network Availability,

d) Security of customer information,

e) Grievance Redressal: Access to senior executives/ officers, Round the clock availability of customer care, Fault Repair Service, Redressal of Excess Metering Cases, Service availability etc.

For many of the services like voice on fixed and mobile networks the regulator has laid down the QoS Standards through the "Standards of Quality of Service of Basic Telephone Service (wireline) and Cellular Mobile Telephone Service Regulations, 2009"[1]. With the introduction of 3G services by the service providers, the regulatory authority has amended these regulations to the extent of introducing new nomenclature and modifying some measurement methodologies. some of the changed parameters are Node Bs Accumulated downtime, Worst affected Node Bs due to downtime, Traffic Channel (TCH) and Circuit Switched Radio Access Bearer Congestion and Circuit Switched Voice Drop Rat.

VI. NGN QOS: CURRENT RESEARCH AND FUTURE WORKS

There is abundant research in the area of NGN and QoS[13]. ITU-T Study Group 13, the group working on future networks, is working on a number of research questions including Q4/13 which deals with the "Requirements and framework for QoS enablement in the NGN". One of the important ongoing study item includes, "What new Recommendations are needed to provide the resource control and management for achieving end-to-end QoS in a heterogeneous environment involving different QoS mechanisms and multiple provider domains?" Research on NGN QoS at ITU-T itself is at the moment mainly focused on the Resource and Admission Control Function (RACF), which is the infrastructure that controls resource allocation and admission control, found in ITU-T Recommendation Y.2111[14].

In ETSI TISPAN has begun work, since early 2008, on the third release of NGN specifications and are now presently looking at IPTV enhancements, IP network interconnection, NGN security enhancements, and QoS with overload control amongst other permanently evolving NGN requirements[15]. Besides these standards, which are currently the most accepted QoS standards, an initiative by the European Union called EuQoS is under development.

The current RACF specifies the functional architecture and control procedure at the IP level. There are still many open issues, and continuing effort is under way to solve the issues. QoS control in the transport technology dependent aspect is an issue that needs further work. The general framework for flow aggregation and signal aggregation is one of the major issues for the overall QoS architecture. The high complexity and scalability of the control mechanism is an issue that needs to be resolved to make large scale migration to NGN successful. A lot of work needs to done in the area of reliability and security in the core network is another issue.

REFERENCES

[1] Thomas B Fowler, "Convergence in Telecommunications", www.noblis.org/News Publications/Publications, 2002

[2] OECD Ministerial Background Report, "Convergence and Next Generation Networks," DSTI/ICCP/CISP, 2007, pp. 7-10

[3] IBM, "Services over IP- delivering new value through Next-Generation Networks," IBM Institute for Business Value, 2005

[4] ITU-T "General Overview of NGN" in ITU-T Recommendations Y.2001, 2004, p 5

[5] ETSI TISPAN ES 282 001 "TISPAN NGN Functional Architecture"(Version 3.4.1), September 2009

[6] Knightson, K. Morita, N. Towle, T, "NGN architecture: generic principles, functional architecture, and implementation," IEEE Communications, Oct 2005

[7] Hassan Yeganeh, Maryam Shakiba, and Amir hassan Darvishan, "Functional Architecture for Resource Allocation and Admission Control," International Journal of Hybrid Information Technology Vol.2, No.3, 2009

[8] ITU-T "Functional Requirements and Architecture of the NGN" in ITU-T
Recommendations Y.2012, 2006

[9] Jasmina Barakovic and Himzo Bajric, "QoS Aspects in NGN Interconnection," BH Telecom, Sarajevo, Bosnia and Harzegovnia, 2009

[10] Jasmina Baraković, Himzo Bajrić, Nedžad Rešidbegović; BH Telecom; Sabina Baraković, Ministry of Security of Bosnia and Herzegovina; "QoE Aspects of NGN Services," Sarajevo, Bosnia and Herzegovina, Sep 2010

[11] ITU-T "Network performance objectives for IP-based services," in ITU-T Recommendations Y.1541, 2006

[12] Peter Wagenaar, "Next Generation Networking Quality of Service On real-time Quality of Service requirements and their fulfilment in Next Generation Networking'" Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, 2009

[13] D. Mustill and P.J. Willis, "Delivering QoS in the next generation network – A standards perspective," BT Technology Journal, vol 23, pp. 48-60, 2005

[14] ITU-T, "Resource and Admission Control Functions in NGN," in ITU-T Recommendations Y.2111, 2008

[15] ETSI "TISPAN- Defining the Next Generation Network," 2008

[1] www.trai.gov.in