Next Generation Network – A Study On QOS Mechanisms

Dr. B. B. Jayasingh[1] and Sumitra Mallick[2]
[1]CVR College of Engineering, Department of IT, Ibrahimpatan, R.R.District, A.P., India
Email: bbjayasingh9@rediffmail.com
[2]Guru Nanak Institute of Technology, Department of IT, Ibrahimpatan, R.R.District, A.P., India
Email: sumitra_mallick2008@rediffmail.com

Abstract-**Next Generation Network (NGN) being an IP based network enables customers to receive voice, data and video over the same network. NGN offers reduced network and operational complexity resulting in better and reliable service. It offers unrestricted access by users to different service providers also supporting generalized mobility. Next Generation Network is capable of converging heterogeneous networks and provides converged services. Resource and Admission Control Function(RACF) is needed to support QoS of the SIP based converged services, which are per session based real time services, such as IP telephony and video telephony. We study the architecture of NGN though the transition from the legacy PSTN to an IP based NGN is an important issue and the QoS control scheme has a strong tendency of focusing on the edge and the access networks. We also present a hierarchical QoS control architecture for improvement of simplicity and scalability in the whole view of the NGN using a divide and conquer strategy which separates interesting objects that are the core and the access networks. We present the Markova modeling of the per session based centralized control scheme and the distributed traffic engineering scheme(e.g. RSVPTE, CR-LDP) for the verification of control costs.**

Index Terms—**NGN, RACF, SCF, CCM, System Sojourn Time.**

I. INTRODUCTION

Nowadays the IP telecommunication market is characterized by an increasing demand for multimedia and real-time communication services, such as Video on Demand, IPTV and Grid Computing, with strict connectivity requirements about bandwidth, packet delay and jitter [1]. Unfortunately, current Internet architecture does not fully support the provisioning of end-to-end emerging broadband services since the Internet service model does not envisage a generalized end-to-end Quality of Service (QoS) support.

Next Generation Network (NGN) is a packet oriented network architecture, standardized by the International Telecommunication Union – Telecommunication Standardization Sector (ITU-T) that supports end-to-end service provisioning based on different QOS-enabled transport technologies. The NGN services include multimedia services (e.g., IPTV), content delivery services (e.g., audio and video streaming), and existing fixed and mobile telephone.

NGN being an IP based network enables customers to receive voice, data and video over the same network. NGN offers reduced network and operational complexity resulting in better and reliable service. It offers unrestricted access by users to different service providers also supporting generalized mobility. In the course of transition from the legacy PSTN to an IP based NGN there are many issues [2] which need to be addressed. We would be addressing the issues relating to regulation and interconnection which arise in the course of migration to NGN.

Next Generation Network (NGN) has been on the implementation phase, primarily focused on a replacement of PSTN. In Japan, the largest national telecom carrier NTT group has announced to start NGN services in March 2008. NTT has also released a set of documents on the preliminary interface condition and service specification for connecting to their NGN networks.

The primary target for NGN is to replace the existing PSTN and ISDN, by introducing highly-reliable networks based on Internet Protocol (IP) and the related technologies. For example, telephone signaling network will be replaced by Session Initiation Protocol (SIP), and the voice transmission will use connectionless protocols such as Realtime Transfer Protocol (RTP) [3].

The *current Internet* is a set of multiple networks which are arbitrarily connected together in various Internet exchanges (IXes), under multitude of bilateral and multilateral agreements between individual Internet Service Providers (ISPs). Some ISPs own the physical links while some use the links provided by the others. Forwarding data between different ISPs are controlled by the policy-based routing protocol, such as Border Gateway Protocol (BGP) [4]. While the current Internet allows open and diverse connectivities as an inter-network of multiple ISPs, the routing has become too complex and a high-cost task for each router. Routing between ISPs are only controlled by the forwarding path between the Autonomous Systems (ASes), a set of

DOI:10.32377/cvrjst0215

multiple IP networks representing an ISP, since BGP is a path vector routing protocol based on policies and rule-sets.

As the number of networks connected to the current Internet increases rapidly, the minimal service conditions between two arbitrary networks get worse with higher latency of packets, instability of multi-ISP routes, and the financial and social conditions of transit ISPs and IXes [5]. Another NGN's perspective is to provide a *reliable* set of services which have already been commercialized on the current Internet, under the control of single management entity, such as a telecom carrier company, which is a completely different management model from the current Internet. NGN networks will introduce prioritized packet forwarding based of Differentiated Services (DiffServ) by using the priority field in each IP packet and with strict priority queueing strategy, so that it can provide real-time services such as telephony and video multicast with no interruption by other services with less real-time demands, such as email and Web.

NGN has layered architecture which consists of a service stratum as a session control layer and a transport stratum as a packet transmission layer [6]. MPLS (Multi Packet Label Switched) is adopted as a packet core network technology of transport stratum [7] and the QoS control on MPLS-TE which is considered to be the standardization. NGN needs QoS control in the view of traffic engineering for the purpose of providing SIP based multimedia services such as IP telephony, video IP telephony, and video conference etc. Herein, NGN architecture defines Resource and Admission Control Function (RACF) in the transport stratum [7].

II. NGN Architecture

This section provides an outline of the NGN Release architecture and the current state-of-the-art of the standardized features. An extensive outline of the ITU-T NGN is already provided in [8], while this section is strictly focused on the NGN mechanisms for the provisioning of QOS-guaranteed connectivity services.

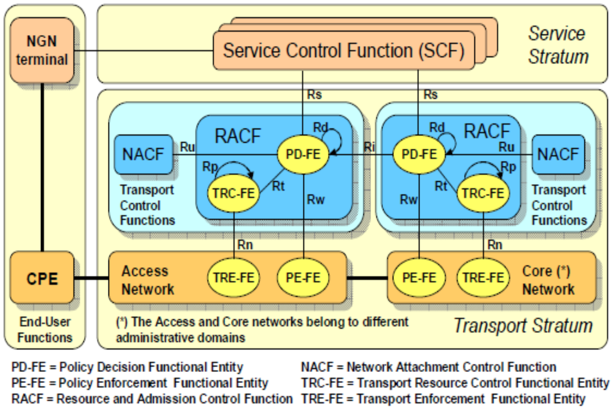

Figure. 1. The NGN architecture

Referring to Fig. 1, the NGN architecture is composed of two sets of functionalities, named Service Stratum and Transport Stratum. To the purpose of this work, the main Service Stratum functional entity is the Service Control Function (SCF). SCF functionalities are specific of a given class of services. SCF performs end-users Authentication, Authorization, and Accounting (AAA) and processes service requests issued by authorized end-users. In particular, by interacting with the Transport Stratum, SCF checks the availability of the network resources and authorizes the network resources reservation needed for the provisioning of the requested service. An example of SCF is the Core IP Multimedia Subsystem (IMS) that inherits and possibly extends a subset of the IMS functionalities standardized by 3rd Generation Partnership Project (3GPP) for the provisioning of Session Initiation Protocol (SIP)-based multimedia services [9].

The Transport Stratum provides IP connectivity services for the benefit of SCFs under the authority of the Network Attachment Control Function (NACF) and the Resource and Admission Control Function (RACF). The NACF is a functional entity that, on receiving a service request issued by an end-user, authenticates the user identity and authorizes the use of network resources based on user profiles. In addition, NACF may supply RACF with the network configuration parameters needed to support the service provisioning. The RACF is the functional entity that enables the SCFs to control network functionalities such as the bandwidth reservation and allocation, the packet filtering, the Network Address and Port Translation (NAPT) while hiding the network technology and topology details. A RACF instance is able to control the network resources belonging to the same administrative domain as shown in Fig.1.

The RACF consists of two specialized functional entities, namely the Policy Decision Functional Entity (PD-FE) and Transport Resource Control Functional Entity (TRC-FE). The PD-FE is the single contact point between any SCF and the Transport Stratum. It makes the final decision about admission, reservation, and control of the network resources supporting the provisioning of the SCF services. The PD-FE decisions are based on (i) the preloaded policy rules decided by the network operator, (ii) the service information provided by the SCF via Rs interface, (iii) the result of resource authorization provided by the NACF via Ru interface, (iv) the outcome of resource admission provided by the TRC-FEs via Rt interface. In addition, the PD-FE may enforce decisions by interacting with a set of Policy Enforcement Functional Entities (PE-FEs) via Rw interface. The PE-FEs are the functional blocks that control the technology-independent network service functionalities implemented at the boundaries of the network such as the NAPT. A set of PD-FEs may interoperate for the seamless provisioning of connectivity services across a multidomain network. PD-FEs within the same administrative domain communicate via Rd interface while PD-FEs belonging to different administrative domains communicate via Ri interface.

The TRC-FE performs technology-dependent admission decisions over the network resources on behalf of PD-FE. Those decisions are based on the requirements received from the PD-FE and on the information previously collected by the TRC-FE about the network topology and the status of the network resources. TRC-FE may interact with a set of Transport Resource Enforcement Functional Entities (TREFEs) via the Rn interface to enforce its decisions over the network. A network domain may contain multiple instances of TRC-FEs controlling different areas. The TRC-FEs belonging to the same administrative domain directly communicate via the Rp interface. TRC-FEs belonging to different administrative domains indirectly communicate through the corresponding PD-FEs via the Ri interface.

The NGN Transport Stratum comprises both access and core networks. In NGN an access network is meant as a network that “takes care of end-users' access to the network as well as collecting and aggregating the traffic coming from these accesses towards the core network” [10]. An NGN terminal interacts with the NGN access network through the corresponding Customer Premise Equipment (CPE) to exchange signalling messages with an instance of SCF. Also, the NGN terminal sends data traffic through the CPE to the Transport Functions, defined in [10] as the set of Transport Stratum entities that interact with the RACF.

III. QOS CONTROL ARCHITECTURE

Although NGN architecture which provides a carrier grade constraint route scheme is different from Internet architecture to support a liberal route scheme, the researches of QOS control in Internet are still of use. The research of Broad Band (BB) architectures gives a suggestion of hybrid architecture as well as an explanation of a centralized and a distributed Architecture. The hybrid architecture suggested in [11] approaches to improve the resource utilization of the admission mechanism while balancing it with the BBs’ processing loads through the adaptation of the centralized and the distributed architectures. However the hybrid architecture pays attention to information synchronization and work load distribution. This architecture provides a coordinate function between two areas. As the coordinate function does not take part in database accesses, there is no issue of consistency. This architecture is expected to handle real time resource and admission control operations for supporting SIP based converged services such as IP telephony and video IP telephony services in NGN. We anticipate that this architecture solves a scalability issue through process load balancing and has advantages of the centralized control scheme which provides high resource utilization, strong consistency, and simplicity. Figure 2 shows the suggested hierarchical architecture in this paper. The Control Coordination Layer (CCL) is deployed in the upper layer and the Resource and Admission Control Layer (RACL) is in the lower. The Control Coordination Manager (CCM) is located in the CCL. RACL approaches a divide and conquer strategy so that the complexity of traffic engineering is decreased. Access Resource and Admission Control Manager (ARACM) takes charge of the access network and Core Resource and Admission Manager (CRACM) takes charge of the core network.

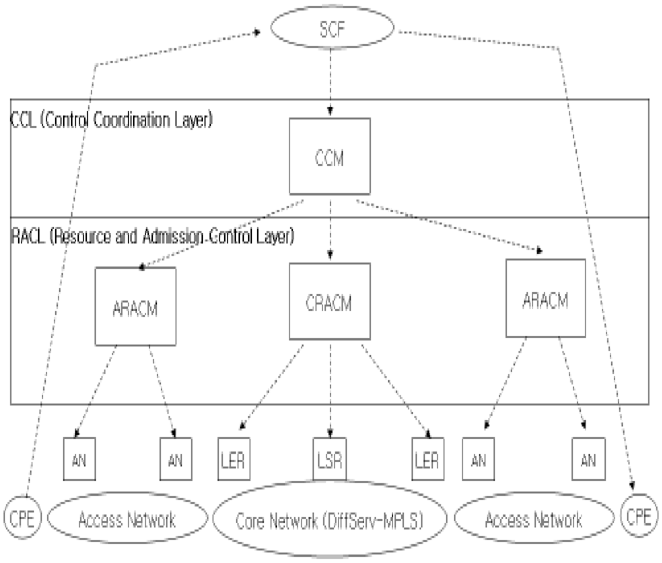

Figure 2. Hierarchical resource and admission control architecture

The operation mechanism of this architecture is as follow:

1. Whenever a calling CPE initiates SIP signaling to a called CPE, the Session Control Function (SCF) transfers information(e.g. bandwidth constraint) for acquiring an appropriate route to the CCM.
2. CCM request to CRACM and ARACM which handles the access networks of the calling and the called CPE belonging to.
3. CRACM searches a matched ER-LSP route in the core network and ARACM finds an appropriate route in the access network concurrently.
4. CRACM and ARACM return admission results to CCM.

5. CCM decides final admission result based on the results of CRACM and ARACM.
6. If the final admission result is positive, CCM responds to SCF a call admission and requests to CRACM and ARACM resource reservations concurrently. And then CRACM and ARACM reserve resources.
7. If the result is negative, CCM responds to SCF a call rejection.

A. QOS Mechanisms

The ways to assure end-to-end QoS is through priority scheduling, resource reservation and admission control mechanisms.

Priority scheduling:NGN is based on IPv6 and RACS calculate the number of hops and remarks hop limit.

- If the number of hop limit is large it provides higher priority.
- If the number of hop limit is small it provides lower priority.
- Let’s assume average minimum bound of packet delay is r ms
- Average propagation delay of one hop is p ms
- There are two QOS class: high and low
- Average queuing delay of high priority per one hop is h ms
- Average queuing delay of low priority per one hop is l ms
- Let there are two end to end connections: longest and shortest routes
- Longest takes d hops to reach destination
- Shortest takes s hops to reach destination

Longest route-worst case:

(p+h)*d ms must be smaller than r ms

shortest route-worst case:

(p+h)*s ms and it is much lower than r ms if (p+l)*s ms is much smaller than r ms, it can use low priority for the shorted route means there are the route which is (p+h)*(l-d)/s - p hops to get low priority QOS request to high priority.

B. Centralized QOS Control Scheme

In this section, we introduce a Centralized MPLSTE (CMPLS-TE (Traffic Engineering)) which uses a centralized scheme for the majority of NGN carriers who use DiffServ aware MPLS in the core. MPLS-TE is mainly dealt in the management plane so far, however the suggested scheme approaches from the control plane of the transport stratum as fig 1 shows. CMPLS-TE has advantages compared with the distributed MPLS-TE(e.g. RSVP-TE). It improves reliability because control messages such as LSP (Label Switch Path) setup and release are transferred through a control channel separated from a data channel. It improves efficiency as the scheme supports the setup, release, and modification of a bidirectional LSP and a multicast LSP just a one-shot control. It is expected to be a fast reroute because it is possible to control a reroute in an affected problem area only instead of a crank-back of whole paths along the LSP. It has flexibility of LSP route selection algorithm so that it adopts several algorithms simultaneously and is easier to update a new algorithm. On the other hand, it needs consideration with regards to a router failure or a failure of whole link which consists of a control channel. A manual management approach is needed in these cases. We also expect that the scheme overcomes a scalability issue in the one NGN carrier domain scale because of nowadays’ blade server capacities and management of LSP paths limitations (e.g. LSP merge, LSP modification).

System sojourn time (E(T)) - the total waiting time from an arrival to a departure from a system - is a representative performance metric [12]. Therefore we define E(T) for LSP setup as a control cost and present models of RSVP-TE and CMPLS-TE for acquiring the control costs. In the RSVP-TE scheme, RSVP-TE transfers a PATH message from ingress LER (Label Edge Router) to egress LER through intermediate LSRs (Label Switch Router) along the route. Then it allocates bandwidths of the LERs and LSRs of the LSP through transferring a RESV message backward for a completion of ER-LSP setup as shown in figure 3. We presume a bidirectional symmetric LSP setup and assume that processing time for the messages is much bigger than transmission and propagation delay for the messages. EDist(T) of RSVP-TE represents equation which means sum of processing time of the PATH and the RESV message for the downstream uni-direction and for the upstream uni-direction LSP.

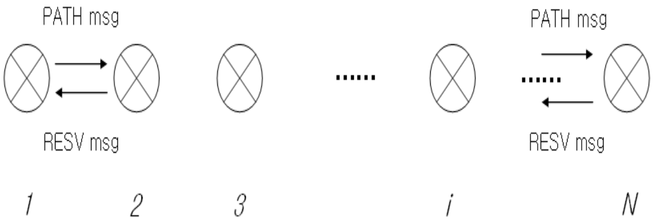

Figure 3. RSVP-TE ER-LSP setup scheme

$$E_{Dist}(T) = \sum_{i=1}^{N-1}\left(T_{\overset{\rightarrow}{i,PATH}} + T_{\overset{\leftarrow}{i,RESV}}\right) + \sum_{i=1}^{N-1} T_{\overset{\leftarrow}{i,PATH}} + T_{\overset{\rightarrow}{i,RESV}})$$

The central manager sends an LSP_Setup_Request message to ingress LER via intermediate LSRs to egress LER and reports to the CCM the result of bandwidth reservation after each LER and LSR allocates bandwidth resources for setting up the bidirectional ER-LSP concurrently in the CMPLS-TE as shown in figure 4. ECent (T) of CMPLS-TE as a control cost represents equation which means the longest latency of processing LSP_Setup_Response message.

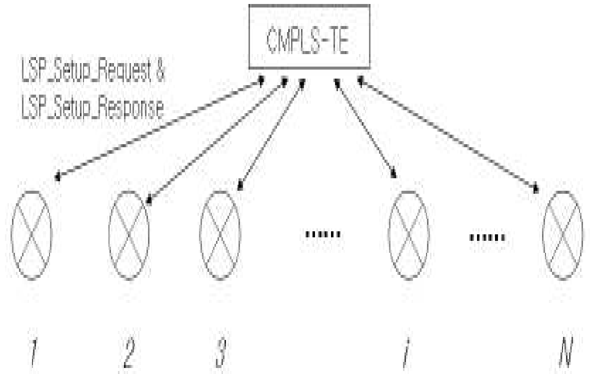

Fig. 4 CMPLS-TE ER-LSP setup scheme

$$E_{Cent}(T)=Max\{\prod_{i}^{N}(T_{i,Request}+T_{i,Response})\}$$

We make the model of M/M/1 open Markova network as shown in figure 5 from a RSVP-TE ER-LSP setup scheme as shown in figure 3 for solving the EDist(T). And then EDist(T) of equation gets a solution from inter-arrival rate of PATH messages(λPATH), inter-arrival rate of RESV messages(λRESV), and service rates of message processing(μPATH, μRESV).

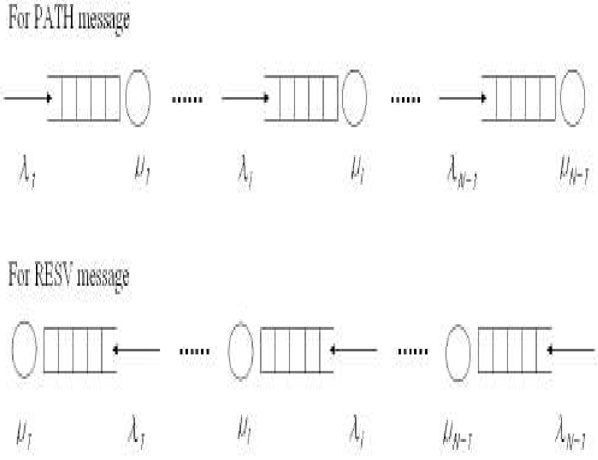

Figure 5. M/M/1 Open markovian network model for RSVP-TE LSP setup

EDist(T) for a bidirectional ER-LSP setup scheme of RSVP-TE results to equation which is two times of summarization of processing time of PATH messages and RESV messages.

$$E_{Dist}(T)=2\times\left[\sum_{i}^{N-1}\left(\frac{\rho_{i,PATH}}{1-\rho_{i,PATH}}\right)\cdot\frac{1}{\lambda_{i,PATH}}+\sum_{i}^{N-1}\left(\frac{\rho_{i,RESV}}{1-\rho_{i,RESV}}\right)\cdot\frac{1}{\lambda_{i,RESV}}\right]$$

$$\rho_{i,PATH}=\frac{\lambda_{i,PATH}}{\mu_{i,PATH}},\qquad \rho_{i,RESV}=\frac{\lambda_{i,RESV}}{\mu_{i,RESV}}$$

We make the model M/M/1 queue for the process of a request (LSP_Setup_Request) message and MX/M/1 queue for the process of response(LSP-Setup_Response) messages as shown in figure 6 for solving the ECent (T) from equation.

The arrival of a request message is a poisson process and that of response messages is a compound poisson process. ECent (T) for a bidirectional ER-LSP setup of CMPLS-TE gets from inter-arrival rate of request message(λReq), inter arrival rate of response message(λResp, the size of group = N), service rate for request message(μi), and the service rate for response message(μResp). ECent(T) results to equation which is summarization of the processing time of the request message and the response messages. Therefore equation presents the solution.

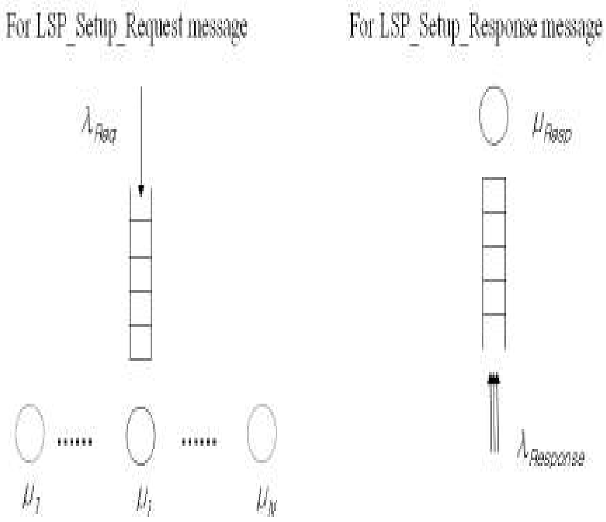

Figure 6. M/M/1 and Mx/M/1 model for CMPLS-TE LSP setup

$$E_{Cent}(T)=\left[E(S_{Req})\cdot\frac{1}{1-\rho_{Req}}\right]+\left[\frac{N+1}{2(\mu_{Resp}-\lambda_{Resp}\cdot N)}\right]$$

$$\rho_{Req}=\lambda_{Req}\cdot E(S_{Req}),\qquad E(S_{Req})=\frac{1}{N}\cdot\left(\sum_{i=1}^{N}\frac{1}{\mu_i}\right)$$

We assume the followings for comparing of the ECent(T) and the EDist(T) according to the N size(the number of the LERs along the arbitrary ER-LSP).

1. Assuming an arbitrary absolute value of μ, and setting up $\mu_{PATH}=\mu_{RESV}=2\mu$, $\mu_i=\mu$, $\mu_{Resp}=\alpha\cdot\mu$ (α is a constant coefficient).
2. Assuming an arbitrary absolute value of λ, and setting up $\lambda=1$, $\lambda_{PATH}=\lambda_{RESV}=\lambda_{Req}=\lambda_{Resp}=\lambda$.

Figure 6 shows that the value of the ECent(T) and the EDist(T) according to the size of N, when the value of μ is varied. Dotted lines show when the value of μ is 30, and solid lines show when μ is 50. In the condition of general assumptions, the ECent(T) is lower than the EDist(T) therefore we conclude that ECent(T) is superior for the control cost in general.

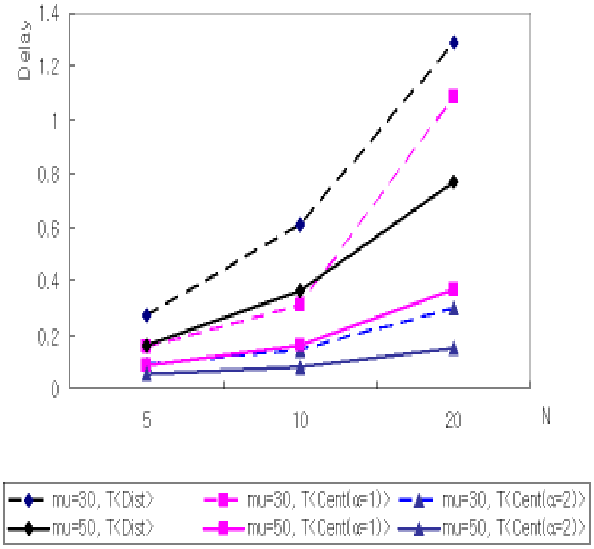

Figure 7. Comparison of control cost on condition of mu=30 and 50

CONCLUSION

We reviewed the novel NGN architecture with RACF entity, able to control dynamically and to coordinate the QOS-guaranteed connectivity. NGN has brought a revolution in mobile networks and some questions are yet to be addressed in implementing it but the situation is changing fast and NGN is capturing the mobile market. The demand for various services to be provided on a single platform has increased. We understand the primary motivation of NGN is the replacement of the legacy telecom carrier networks, with the inexpensive equipments available for the Internet protocols. Many Internet users expect NGN to be better than the current Internet and ensuring access to the existing Internet services. The advantage of NGN is having admission control mechanism that reduces network resource reservation complex in unit of log scale. The reserved resources in a route has less delay in centralized approach compared to distributed approach that shows the NGN QOS.

REFERENCES

[1] [1] F. Baroncelli, B. Martini, V. Martini and P. Castoldi, “Supporting Control Plane-enabled Transport Networks within ITU-T Next Generation Network (NGN) architecture”, 978-1-4244-2066-7/08/$25.00 ©2008 IEEE, pp. 271-278.

[2] Dr. Mustafa Shakir, “Challenging Issues in NGN Implementation and Regulation”, 978-1-4244-3709-2/10©2010 IEEE

[3] H. Schulzrinne and S. Casner and R. Frederick and V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications,” July 2003, RFC3550.

[4] Y. Rekhter, T. Li, and S. Hares, “A Border Gateway Protocol 4 (BGP-4),” January 2006, RFC4271.

[5] Kenji Rikitake and Koji Nakao, “NGN and internet: from coexistence to integration**”,** First ITU-T Kaleidoscope Academic Conference, 92-61-12441-0/CFP0838E © 2008 ITU, japan.

[6] Keith Knightson, Naotaka Morita, and Thomas Towle, “NGN Architecture: Generic Principles, Functional Architecture, and Implementation”, IEEE Communications Magazine, vol. 43, no. 10, pp 49-56, Oct. 2005.

[7] Keith Knightson, Naotaka Morita, and Thomas Towle, “NGN Architecture: Generic Principles, Functional Architecture, and Implementation”, IEEE Communications Magazine, vol. 43, no. 10, pp 49-56, Oct. 2005.

[8] Keith Knightson et al “NGN Architecture: Generic Principles Functional Architecture, and Implementation” Communications Magazine, IEEE, Oct. 2005, Volume 43, Issue 10.

[9] ITU-T Y.2021, “IMS for Next Generation Networks”, September 2006.

[10] ITU-T Y.2012, “Functional requirements and architecture of the NGN release 1”, September 2006.

[11] Ch. Bouras and K. Stamos, “Examining the Benefits of a Hybrid Distributed Architecture for Bandwidth Brokers”, in Proc. IPCCC’05, 2005, Paper 10.1109/PCCC, pp. 491-498.

[12] Kwon Cho, Koji OKAMURA, “*A Centralized Resource and Admission Control Scheme for NGN Core Networks*”, Graduate School of Information Science and Electrical Engineering, Kyushu University 6-10-1 Hakozaki, Higashi-ku, Fukuoka, 812-8581, Japan.