Mobile QoS provisioning by Flow Control Management in Proxy Mobile IPv6

Taihyong Yim, Tri M. Nguyen, Youngjun Kim and Jinwoo Park
School of Electrical Engineering
Korea University
Seoul, Rep. of Korea
{autonome, trinm, youngjun2, jwpark}@korea.ac.kr

Abstract- **In the next-generation mobile wireless networks, mobility support and QoS provisioning are two critical issues. When it becomes much easier to access the internet from mobile devices, the real-time service over mobile network will be on high demand. To satisfy these requirements we must consider finest level of QoS guarantee in the mobile network. In this paper, we propose a QoS Provisioning Method based on flow-level traffic management for guaranteed service in Proxy Mobile IP.**

Keywords-QoS; Mobility; PMIP; Flow-based traffic management; Admission Control

I. INTRODUCTION

Mobility and QoS mechanism is the key issue in future wireless mobile networks. Future wireless mobile networks are expected to provide efficient mobility support with quality-of-service (QoS) guarantees. Services are required to maintain their network connectivity with the same QoS during handoff.

From Mobile IP (MIP)[1] to Proxy Mobile IP (PMIP)[2], many mobility management protocols are proposed to maintain the session continuity for higher layer in the IP-based networks. Especially PMIP [2], the Internet Engineering Task Force (IETF) standard, can serve as the basic network-based mobility management in the IP-based mobile networks. PMIP aims to solve the host-based mobility support scheme such as MIP. PMIP relies on the proxy mobility agents in the network to detect the MN's attachments and detachments and then signal this information, in the form of binding updates without the active participation of the MN itself. However PMIP is not enough to support the service continuity for guaranteed service. If the network suffers from congestion on the specific link, connections to networks may be broken and QoS also may be degraded. It is because PMIP is a legacy of IP, which is based on "best effort service".

To cope with this problem, QoS mechanisms have been largely studied in both wired and wireless environments. For example, Integrated Service (IntServ) [3] and Differentiated Service (DiffServ) [4] can be used in IP based wired and wireless networks. IntServ can provide QoS through admission control, classifier, packet scheduler and resource reservation. In IntServ, a QoS signaling protocol, Resource Reservation Protocol (RSVP) is used for this purpose. RSVP enables end applications requiring certain guaranteed services to signal their end-to-end QoS requirements to obtain service guarantees from the network. In IntServ network resources are reserved for a session according to a specific QoS requirement can support QoS per flow level though the reservation by exchanging explicit signaling messages. However, it has a scalability problem since it requires signaling messages to be exchanged between terminals periodically. Moreover, it results additional delay during handoff. Therefore RSVP is not suitable for mobile networks.

On the other hand, DiffServ is a direct extension to the work done by IntServ. While IntServ provides per-flow guarantees, DiffServ follows the Class of Service (CoS) of mapping multiple flows into a few service levels. DiffServ controls only traffic classes rather than each session within a traffic class. For CoS the SLA (Service Level Agreement) a central component of DiffServ, which is a service contract between a customer and a service provider. The SLA specifies the details of the traffic classifying and the corresponding forwarding service a customer should receive. DiffServ uses code point (DSCP) values in the IP header to deliver the CoS according to the SLA. However, DiffServ lacks controllability such as admission control and it cannot satisfy of per-flow QoS required in the various services.

For these reasons, simple and efficient QoS architecture is needed with a traffic management schemes in flow-level admission control, packet scheduling, policing, which do not use expensive signaling messages. In this manner, Flow-Aware Networking (FAN) is introduced by France Telecom in [5][6] as a new way of providing QoS in the IP networks. The main goal of this proposal is to ensure the proper QoS in packet networks in an implicit way. That is, no signaling is required to control the network. Each node makes locally optimal decision based on local observation. In the congestion state, new flows are blocked to protect existing flows by flow-level admission control of IP packets. On the other hand, Flow-State-Aware (FSA) technologies were

 ISBN: 978-1-61208-133-5

developed for NGN transport technologies [7]. FSA defines the service types based on typical examples of Internet services: maximum rate (MR), guaranteed rate (GR), variable rate (VR), and available rate (AR), and divides the network resource into two portions: fixed rate (FR) and network rate (NR). In FSA, signaling procedure requires every node to exchange requests and responses according to service types. Through this signaling capability in controlling transit nodes FSA can support QoS in flow-level.

In this paper, we proposed a Mobile Flow-Aware access network which can provide a mobile QoS provisioning of flow-level for PMIP. The rest of this paper is organized as follows. In Section II, we discuss the related work to this research. Section III describes our proposed network architecture and scheme. Finally, the conclusion and further work are presented in Section IV.

II. Related Work

A. *Proxy Mobile IP*

Proxy Mobile IPv6 protocol is intended for providing network-based IP mobility management support to a mobile node, without requiring the participation of the mobile node in any IP mobility related signaling [2]. The mobility entities in the network will track the Mobile Node (MN)'s movements and will initiate the mobility signaling and set up the required routing state. Therefore, an MN is exempt from participation in any mobility-related signaling, and the proxy mobility agent in the serving network performs mobility-related signaling on behalf of the MN. Once an MN enters its PMIPv6 domain and performs access authentication, the serving network ensures that the MN is always on its home network and can obtain its HoA on any access network. That is, the serving network assigns a unique home network prefix to each MN, and conceptually this prefix always follows the MN wherever it moves within a PMIPv6 domain. From the perspective of the MN, the entire PMIPv6 domain appears as its home network. Accordingly, it is needless (or impossible) to configure the CoA at the MN. The new principal functional entities of PMIPv6 are the mobile access gateway (MAG) and local mobility anchor (LMA). The MAG typically runs on the AR. The main role of the MAG is to detect the MN's movements1 and initiate mobility-related signaling with the MN's LMA on behalf of the MN. In addition, the MAG establishes a tunnel with the LMA for enabling the MN to use an address from its home network prefix and emulates the MN's home network on the access network for each MN. On the other hand, the LMA is similar to the HA in MIPv6. However, it has additional capabilities required to support PMIPv6. The main role of the LMA is to maintain reachability to the MN's address while it moves around within a PMIPv6 domain, and the LMA includes a binding cache entry for each currently registered MN. The binding cache entry maintained at the LMA is more extended than that of the HA in MIPv6 with some additional fields such as the MN-Identifier, the MN's home network prefix, a flag indicating a proxy registration, and the interface identifier of the bidirectional tunnel between the LMA and MAG. Such information associates an MN with its serving MAG, and enables the relationship between the MAG and LMA to be maintained.

B. *QoS mechanisms based on flow-based traffic management*

To cope with limitation of IP based on best effort service, QoS mechanisms have been largely studied in both wired and wireless environments such as IntServ and DiffServ. IntServ enables end applications requiring certain guaranteed services to signal their end-to-end QoS requirements. On the other hand DiffServ controls only traffic classes rather than each session within a traffic class, which enables network to be scalable. However, both InServ and DiffServ have the limitation of scalability and controllability respectively.

For these reasons, simple and efficient QoS architecture is needed with a traffic management schemes in flow-level admission control, packet scheduling, policing. As the network processor and memory technologies developed, routers can recognize packets as a flow which is sequence of packets with the same 3-tuples or 5-tuples information. This enables the network to associate packets dynamically. That is, traffic control can be done at flow-level. The definition of flow is not fixed, but it could be defined in various ways according to the requirements of the user or service provider. A flow could be defined as a traffic flow which shares the 5-tuple IP header fields. Several schemes have proposed in this manner.

FAN is a new way of providing QoS in the IP network [4][5]. It is designed for providing state information to conventional IP router with stateless information for specific classification of the IP packet. In the FAN, packets are treated by the flow level. Through CAC per flow ongoing service can be maintained even in the situation of overload. It can guarantee more specific level of QoS compared with class-based traffic control architecture such as DiffServ. The main goal of FAN is to ensure the proper QoS in packet networks in an implicit way [6]. That is, no signaling is required to control the network. Each node makes locally optimal decision based on local observation. In the congestion state, new flows are blocked to protect existing flows by flow-level admission control of IP packets. If a packet comes into the system, the selected hashing function will generate a hash value. The hash value is used to find the flow state entry for the flow of the packet. If the packet is the first packet of the flow, no flow state entry for the flow exists. Therefore, a new flow state entry must be created for the flow with the appropriate forwarding and QoS information. On the other hand, if there is already a flow state entry for the flow, the packet is just processed according to the information in the flow state table. Since a flow is uniquely identified by its 5-tuple fields, the lookup for the flow state table should be an exact match instead of longest-prefix match as in the IP forwarding table lookup. A flow state entry is created and maintained when the first packet enters the system [8]. Once flows are identified and maintained in the system, traffic management can be done for each flow.

III. PROPOSED FLOW CONTROL MANAGEMENT

In this section, we describe our proposed QoS provisioning scheme for guaranteed service in Proxy MIP. As mentioned earlier, our ultimate aim is to overcome the limitation of Proxy MIP and benefit from the QoS support capability of flow-based traffic management. Proposed access networks is based on the integration of PMIP [2] and Flow-aware technologies [5][6][7]. That is, mobility management is performed according to PMIP and QoS provision is obtained by Flow-aware technologies. In the following, we present the operation of our proposed network architecture, namely, Mobile Flow-Aware Access Network and mobility management schemes and QoS provisioning method in flow-level.

A typical architecture for Mobile Flow-Aware access network is shown in Fig.1. We assume that a Mobile Flow-Aware access network exist between the Mobile Flow-Aware Local Mobility Anchor (MFA-LMA) and the Mobile Flow-Aware Mobile Access Gateway (MFA-MAG). The architecture is based on a two-level hierarchy. At the higher level is the MFA-LMA that performs the role of the LMA as it of PMIP [2] with flow-based traffic management function. At the second level is the MFA-MAG that is responsible for tracking the MN's movements to and from the access link as conventional MAG in PMIP. MFA-MAG also has a function of flow-level traffic management The MFA is an intermediate node that route packets with function of flow-level traffic management.

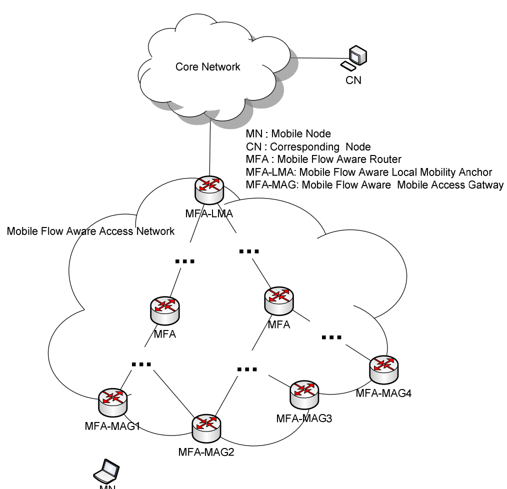


Figure 1. Architecture of a proposed flow- aware access network

A. QoS Provisioning for Guaranteed Service

For QoS provisioning two types of flows are defined: elastic and inelastic. Elastic flow is usually used for data transmission, served with the best effort regime such as web service. On the other hand, inelastic flow is used for delay-sensitive services, served with the specific fixed data rate like VoIP services. The packets of the latter have a higher priority than that of the former. The goal of the proposed QoS provisioning is to guarantee the inelastic flows even though the congestion is occurred at the link. For this purpose, each MFA node should store the list of the ongoing inelastic flow, namely, Flow Cache Entry (FCE) at each interface. Fig. 2 shows the structure of FCE. FCE include the 5-tuple of packets (Source/Destination IP address, Source/Destination port number, higher layer protocol) and interface of the MFA link. Flow entry is maintained by soft state, that is, no explicit signaling is needed.

The main elements of the proposed flow-based traffic management scheme are shown in Fig..3. On a packet arriving at node incoming interface, the packet is classified into elastic or inelastic flow by the classifier. While elastic flows are forwarded to admission control block directly, inelastic flows are checked whether new flows or not. If there are matching entries in the FCE, flows are forwarded to sub queue at the outgoing interface directly. If a packet of flow is determined to new flow, then FCE may be updated according to result of the admission control. The admission control uses traffic measurement of waiting time in sub-queues for inelastic flows. Congestion state can be defined as the state that satisfy the following inequation,

$$\sum_{n=1}^{k} \frac{l_i(n) \times 8}{r_i} \geq \varepsilon_i \qquad (1)$$

where li(n) is the length (byte) of n-th packet of inelastic flow i, ri is the service rate (bits/sec) of flow i and εi is the delay constraint for flow i. If the total sum of the expected service time for each flow which are waiting in the sub-queue is longer than the delay time for QoS of flow i, new elastic flows are blocked to protect existing inelastic flows. According to this admission control Each MFA node makes locally optimal decision based on local observation. The

Flow Cache Entry

interface	Source IP address	Destination IP address	Source port number	Destination Port number	Protocol	time
...	...	...	...	...	...	...

Figure 2. Flow Cache Entry

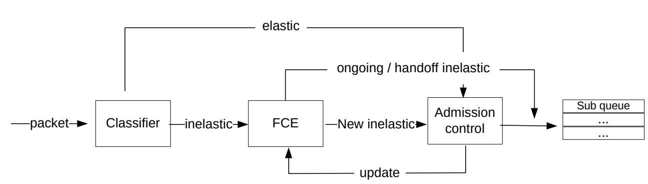


Figure 3. Flow-based traffic management

main advantage of proposed QoS provisioning is simplicity. It requires no signaling for QoS. Only implicit admission control is required upon congestion.

B. QoS Provisioning for handoff flow

When MN moves between MAGs, the flow path may be changed. Mobility can cause some problem in flow-based QoS control such as the failure of flow identification. To protect ongoing inelastic flows in the congestion, MFA node should keep the FCE. However, FCE is maintained locally, therefore some MFA nodes on the newly changed path according to handoff do not have flow list in FCE for the handoff flow. That is, handoff flow can be treated as a new inelastic flow and blocked in the congestion state. To avoid handoff flows treating as new flows, two types of FCE are proposed: Local FCE and Global FCE. Both Local FCE and Global FCE have the same structure as shown in Fig. 2. The only difference between two FCEs is the coverage of the contained flows in the list. That is, the Local FCE is the FCE that is managed by each node respectively while the Global FCE is the FCE that is managed by MFA-LMA. In other words, Local FCE contains the list of inelastic flows that are treated independently by a MFA and Global FCE contains the list of all inelastic flows in the domain. Fig. 4 shows the admission control for handoff flow identification. MFA checks Local FCE first and then checks Global FCE additionally. Through this simple mechanism handoff flow can be detected at the node. Therefore the QoS for handoff flows can be support like ongoing flows.

Basically local MFA do not need to maintain the Global FCE. Although the FCE is maintained in soft state, to maintain the FCE is a burden to the MFA. Therefore the small size of the FCE is good for MFA. For this reason, MFA refer the Global FCE only when MN moves to its local area through the PBU. The Global FCE request/reply procedures are shown in the Fig. 5. Handoff procedures follow the procedures of conventional PMIP [2] basically. The Global FCE is requested and replied with PBU and PBA. Instead of noticing list of all the flows, MFA-LMA just sends the list of flows of MN for reducing the burden of unnecessary work.

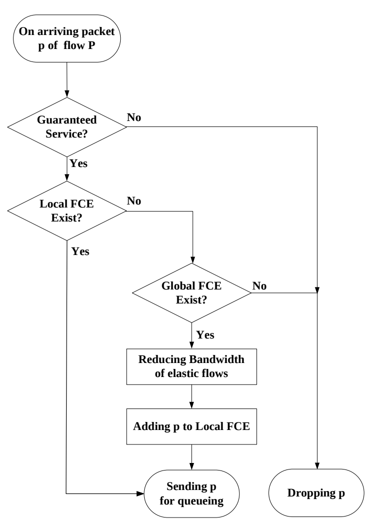


Figure 4. Admission control

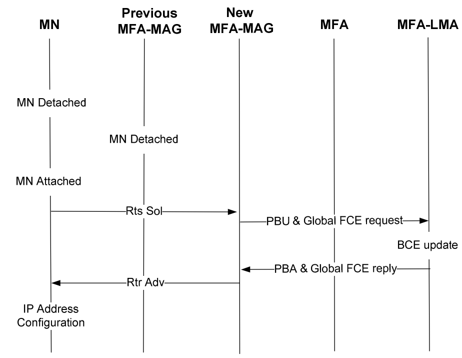


Figure 5. Global FCE request/reply during handoff procedures

IV. PERFORMANCE ANALYSIS

In this section, we analyze the QoS provisioning of flow-based traffic management mechanism that has been proposed in the previous sections. We did not build any concrete numerical methodologies to analyze the queue management schemes; therefore, we provide computer simulation results. The network topology for the simulation is shown in Fig. 6. The links between nodes are set to have a link speed of 1Gbps. Background traffic of 100 elastic flows with CBR 10 Mbps is generated by MN1 and sent to CN1, and MN2 and MN3 each generate 50 inelastic flows with CBR 20 Mbps to send to CN2. Therefore, a total of 2Gbps traffic is trying to be sent between MFA and MFA-LMA, which causes congestion at the link. The packet size was set to 1,000 bytes.

Fig. 7.(a) shows MFA-MAG1's throughput when proposed QoS provisioning is not applied. As we can expect, the rates of the flows are fairly distributed; 100 flows share 1Gbps fairly; therefore, each flow receives about 10 Mbps. This means QoS of inelastic flows is not provided. If we need to guarantee the bandwidth of a certain flow to 20 Mbps, it is not possible in that architecture.

In Fig. 7.(b), our mechanism can guarantee 20 Mbps for a inelastic flow, and the rest of the bandwidth can be fairly shared among the other flows.

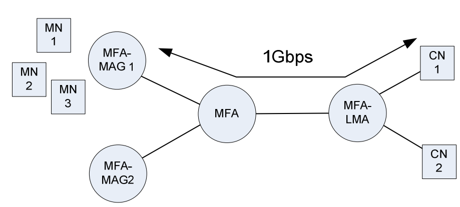
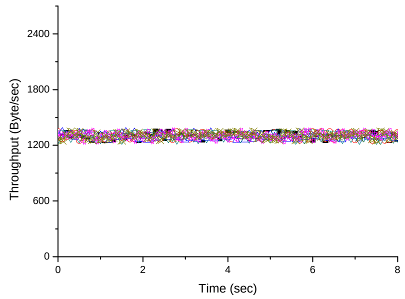
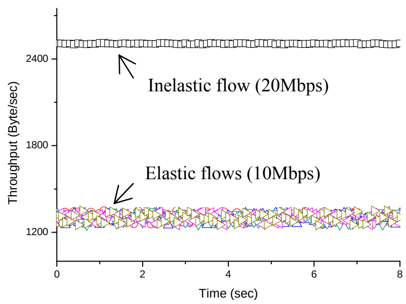


Figure 6. Simulation network topology



(a)



(b)

Figure 7. Throughput : (a) conventional scheme, (b) proposed scheme

V. CONCLUSION

This paper described a new QoS provisioning based-on flow-level traffic management in the PMIP for guarantee the QoS of inelastic flows. The proposed scheme shows how flow-level QoS provisioning can be evaluated in the PMIP domain. Through proposed classification, admission control, buffer managements, QoS for inelastic flows are guaranteed even when the network link is congested. Also through managing the two kinds of FCE, mobile inelastic flows also can be treated with the same priority of ongoing flows and protected on congestion state.

As a further works, we'll analyze the performance of proposed scheme with numerical methodologies to prove the advantage of our proposed scheme.

ACKNOWLEDGMENT

This work was supported by the IT R&D program of MKE/KEIT[KI001822, Research on Ubiquitous Mobility Management Methods for Higher Service Availability] and partly by the BLS project from the Seoul Metropolitan City[WR080951, Seoul R&BD Program]

REFERENCES

[1] C. Perkins, "IP Mobility Support for IPv4," RFC 3220, Jan. 2002.

[2] S.Gundavelli,Ed et al., "Proxy Mobile IPv6," RFC5213, Aug. 2008.

[3] R. Braden et al., Integrated Services in the Internet Architecture: An Overview, IETF RFC 1633, June 1994.

[4] S. Blake et al., An Architecture for Differentiated Services, IETF RFC 2475, Dec. 1998.

[5] Lawrence G. Roberts, "The Next Generation of IP - Flow Routing", Proceedings of SSGRR 2003, Italy, 2003.

[6] S. Oueslati, J. Roberts, A new direction for quality of service: Flow-aware networking, Proceedings of NGI 2005, Rome, June 2005.

[7] ITU-T Recommendation Y.2121, Requirements for the Support of Stateful Flow-Aware Transport Technology in an NGN, Sept. 2007.

[8] Nam-Seok Ko, Sung-Back Hong, Kyung-Ho Lee, Hong-Shik Park, and Nam Kim, "Quality-of-Service Mechanisms for Flow-Based Routers," ETRI Journal, Volume 30, Number 2, April 20

[9] N. Yamagaki, H. Tode, and K. Murakami, "DMFQ Hardware Design of Flow-Based Queue Management Scheme for Improving the Fairness," *IEICE Trans. Comm*., vol. E88-B, no. 4, Apr. 2005, pp. 1413-142