Seamless QoS Handling Mechanism for Macro and Micro Mobility[1]

Sangheon Pack and Yanghee Choi

School of Computer Science and Engineering, Seoul National University, Seoul, Korea
Email: shpack@mmlab.snu.ac.kr and yhchoi@snu.ac.kr

Abstract. One of the most challenging issues facing designers of wireless and mobile networks is the proper provisioning of Quality of Service (QoS), and is probably the most difficult problem in terms of meeting users' QoS requirement at various mobility levels. In this paper, we propose a seamless QoS handling mechanism for diverse mobility situations. This architecture consists of a Differentiated Service (DiffServ) model and an Integrated Service (IntServ) model. We propose the class upgrade scheme in packet forwarding for smooth QoS handling of macro mobility and propose a multiple path reservation scheme based on the Resource Reservation Protocol (RSVP) for seamless QoS handling of micro mobility. The proposed mechanism can meet constant QoS requirements after mobility. Furthermore, it is so flexible and scalable that it can be utilized as the QoS model in the next wireless and mobile networks.

1 Introduction

To guarantee Quality of Service (QoS) is the most critical problem in the current Internet. To address the QoS problem , several models have been proposed by the IETF, such as the Integrated Service (IntServ) model based on RSVP (Resource Reservation Protocol) and the Differentiated Service (DiffServ) model. However, these models are designed for wired networks and they are unsuitable in some respects for wireless and mobile networks. Unlike wired networks, wireless networks have several different characteristics, namely, a high loss rate, bandwidth fluctuations, and mobility. Therefore, the QoS mechanism in the wireless and mobile networks should consider these characteristics.

Recently, there has been much research on the provision of QoS in wireless and mobile networks. These studies have focused on several features of wireless networks such as low bandwidth, high loss rate, and the constraints of terminals [1]. However, due to the increased number of portable devices, the QoS guarantee mechanism must consider the mobility of terminals as well as wireless network characteristics.

[1] This work was supported in part by the Brain Korea 21 project of the Ministry of Education, in part by the National Research Laboratory project of the Ministry of Science and Technology, 2001, Korea.

In this paper, we propose a seamless QoS handling mechanism under various mobility situations. We assume a QoS architecture combined DiffServ model with IntServ model. The former handles the QoS guarantee in a wired core network and the latter provides the QoS guarantee in a wireless access network. In case of macro mobility, it is possible to support smooth handoff by the packet-forwarding scheme using a class upgrade mechanism. In terms of micro mobility, we use the multiple path reservation scheme for seamless QoS handling.

The rest of the paper is structured as follows. Section 2 describes IntServ/RSVP and DiffServ models in brief and introduces related works. In Section 3 we propose the overall architecture based on these models. Section 4 describes the QoS handling mechanism in terms of macro and micro mobilities. Section 5 evaluates the performance of the proposed mechanism. Finally, Section 6 concludes our approaches.

2 Background and Related Work

In this section we introduce the IntServ model based on the RSVP and DiffServ models, and summarize some related work upon the QoS mechanism in wireless/mobile networks utilizing these models.

2.1 Intserv and DiffServ model

In the IntServ model, a path between sender and receiver is reserved before establishing a session. In the path setup phase, a method of informing the application's requirements of the network elements along the path and a method of conveying QoS-related information between network elements and the application is needed. This process is achieved by using a Resource Reservation Protocol (RSVP) [2]. RSVP is a signaling protocol to carry the QoS parameters from the sender to the receiver to make resource reservations.

The protocol works as follows: (i) The sender of an application sends PATH message containing the traffic specifications to the receiver(s) of the application. (ii) The receiver on receiving this PATH message sends on RESV message to the sender specifying the flow it wants to receive. (iii) As the RSVP message flows back to the sender, reservations are made at every intermediate node along the required path. If any node along the path cannot support the request, that request is blocked. (iv) Path and reservation state are maintained at every router along the way for every session. To refresh the path and reservation states, PATH and RESV are sent periodically.

The IntServ model based on RSVP can provide three types of services to users: (i) Best effort service is characterized by the absence of a QoS specification and the network delivers at the best possible quality, (ii) Guaranteed service provides users with an assured amount of bandwidth, firm end-to-end delay bounds, and no queuing loss for flows, and (iii) Controlled load service assures that the reserved flow will reach its destination with a minimum of interference from the best-effort traffic [3].

One drawback of IntServ using RSVP is that the amount of state information increases with the number of flows. Therefore it is considered as a non-scalable solution for the Internet core network. On the other hand, DiffServ maps multiple flows into a few service classes. The 8-bit TOS (Type of Service) field in the IP header supports packet classification. The TOS byte is divided into a 6 bit Differentiated Services Code Point (DSCP) field and a 2-bit unused field [4]. DiffServ is realized by mapping the DSCP contained in the IP packet header to a particular treatment, also described as per-hop behavior (PHB). DiffServ defines various PHBs. For example, Assured Forwarding (AF) Service gives the customer the assurance of a minimum throughput, even during periods of congestion. DiffServ does not have any end-to-end signaling mechanism and works based on a service level agreement between the provider and the user. All packets from a user are marked in a border router to specify the service level agreement and are treated accordingly.

2.2 Related Works

[5] proposed a QoS supporting mechanism in mobile/wireless IP networks using a DiffServ model. [5] assumed the hierarchical FA (Foreign Agent) structures and the fast handoff [6] for mobility management. Based on these assumptions, [5] describes the Service Level Agreement (SLA) procedures in inter-FA handoff and inter-domain handoff. However, DiffServ only provide PHBs for aggregated flows so that it is impossible to meet per-flow QoS requirements, and reactive mobility managements like fast handoff are unsuitable to guarantee QoS for mobile hosts that move between small-size cells frequently.

To support the per-flow QoS requirement in a micro mobility environment, [7] proposed the extended RSVP supporting mobility. [7] proposed a path reservation scheme using a multicast tree. This scheme makes resource reservation in advance at the locations where it may visit during the lifetime of the connections. But, since [7] does not assume a specific mobility management scheme, the extended RSVP handles not only resource reservation but also mobility management. Therefore the RSVP proposed in [7] requires excessive overhead for the implementation and adaptation of the protocol.

3 Architecture

In this section we describe the overall architecture for the QoS guarantee in a wireless/mobile network, and explain the initial QoS negotiation procedure in this architecture.

3.1 Overall Architecture

Figure 1 shows the overall QoS provisioning architecture. The architecture is composed of two parts: a wired core network and wireless access network.

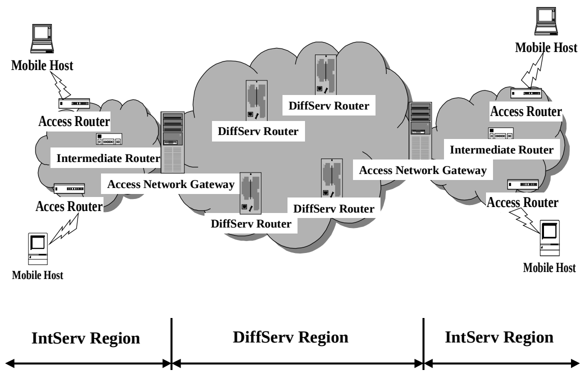


Fig. 1. Overall QoS Architecture

The core network uses the DiffServ model for QoS provisioning, which is composed of Access Network Gateways (ANGs) and DiffServ Routers (DRs). ANG plays an interface role between the IntServ region and the DiffServ region, and should classify incoming packets and mark DSCP field in the IP header. In addition, the ANG is operated as either a Foreign Agent(FA) or a Home Agent(HA) for mobility management. DR is a general DiffServ router that forwards the received packet according to corresponding PHBs.

On the other hand, the access network uses the IntServ/RSVP model. Routers in an access network are organized in a hierarchical structure, similar to Cellular IP [8] and HAWAII [9]. ANG is a root node in a wireless access network and administers resource management and admission control for all Mobile Hosts (MH). The Intermediate Router (IR) is a general router that reserves according to the QoS specification, as described in the RSVP message. Access Router (AR) is a router that acts as a Base Station (BS) for MHs. For seamless handoff, we assume the host specific routing scheme [10] is used within an access network. In host specific routing, an MH keeps its IP address while it moves between cells included in the same access network, and IP packets are routed by an entire IP address not by a network prefix. Since the host specific routing has the problem of non-scalability, it can be used only in limited regions, such as in the single access network domain.

By combining two models it might be possible to build a scalable network that would provide predictable end-to-end QoS services.

3.2 Initial QoS Negotiation Procedure

The initial signaling process to obtain end-to-end quantitative QoS starts when an MH generates an RSVP PATH message. The generated PATH message is forwarded to ANG along the AR and IRs. Then ANG sends the received PATH message towards the DiffServ region. To forward the PATH message through the DiffServ region, ANG must map the message to a DiffServ service class. The service mapping is possible by either defining a new service class [14] or utilizing existing classes. After the service mapping procedure, the PATH message is tunneled based on PHB. When the PATH message gets to the ANG in the destination access network, the message is processed according to the standard RSVP processing rules. When the PATH message reaches the destination MH, the MH generates an RSVP RESV message, and the RESV message is routed to the source MH along the reverse path. Figure 2 shows the initial QoS negotiation procedure.

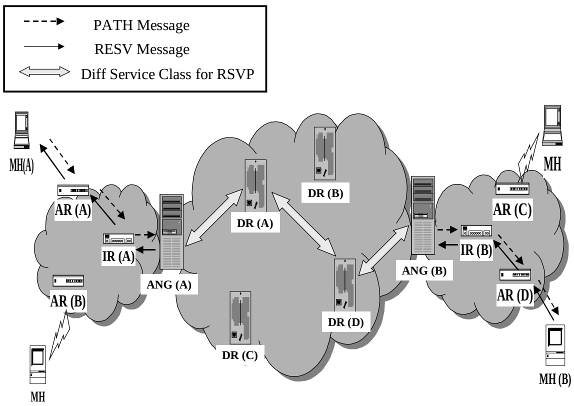


Fig. 2. Initial QoS Negotiation Procedure

Of course, the request may be rejected at any node in the IntServ region according to IntServ admission control. Also, ANG triggers the DiffServ admission control when it receives the RESV message. The ANG checks whether the resources requested in the RESV message exceed the resources defined in SLA or not. If the request fits the negotiated SLA, it is granted, if not, the RESV message is not forwarded and an appropriate RSVP error message is sent back to the receiver.

4 QoS Handling Mechanism for Mobility Management

For more efficient mobility management, [15] proposed the hierarchical mobility management. According to this proposal, the mobility management is divided into macro and micro categories. In this section, we propose the seamless and smooth QoS handling mechanism in each case.

4.1 QoS Handling Mechanism for Macro Mobility Management

Macro mobility means that an MH moves from an access network to another access network. In such a case, it is necessary to inform the home domain of the movement because of security, billing, and other considerations. Macro mobility is generally handled by the Mobile IP [12]. When an MH moves to a new access domain, it sends a registration message to a new foreign ANG in that domain.

In terms of the QoS guarantee, since an MH's IP address is changed after macro mobility, a new end-to-end QoS reservation phase is needed. The moved MH reserves a path within an access network by sending a new PATH message to the current ANG and performs an SLA procedure with the ANG for packet tunneling in the DiffServ region. In addition, the SLA between the previous ANG and the current ANG is required for packet tunneling in Mobile IP. Except that two end-point MHs move together, this QoS handling mechanism is done within the new access network. Figure 3 shows the QoS re-negotiation procedure for macro mobility management.

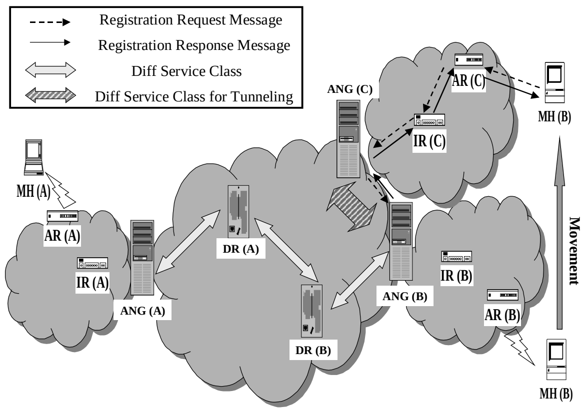


Fig. 3. QoS handling mechanism in macro mobility without route optimization

In addition to QoS re-negotiation, a packet forwarding scheme is needed for smooth handoff [11]. To minimize packet losses, the previous ANG must buffer packets during handoff and forward the buffered packets to the current ANG. Since the packet forwarding is based on the tunneling scheme in the Mobile IP, an additional IP header with same DSCP field as in the original IP header is added ahead of the existing IP packet. However, in many cases, especially in real-time multimedia applications, on time delivery of the forwarded packets is important. Therefore, forwarded packets should have a higher priority than the incoming packets with the same DSCP field value. To address this problem, we propose the class upgrade scheme. In this scheme, data packets are tunneled as packets with higher priority DSCP field values not the same DSCP field values.

Figure 4 shows the concept of the class upgrade scheme. Not to disturb packets in upper class level, the class upgrade is possible only within the same class level. And, the class update scheme is used until all buffered packets have been forwarded to the current ANG.

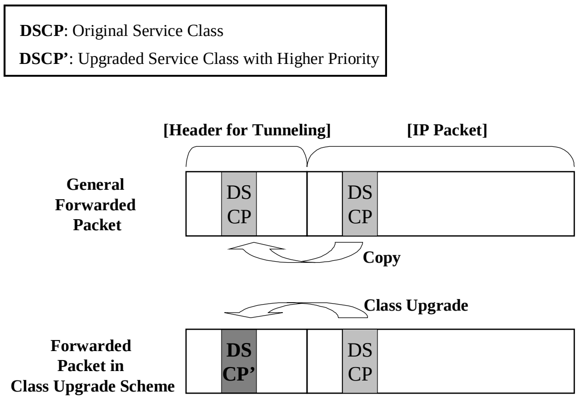


Fig. 4. Class Upgrade Scheme

If route optimization [16] by a binding procedure is supported in Mobile IP, the tunneling between the previous ANG and the current ANG is not required. In this case, the additional SLA between the DiffServ region, belonging the current ANG, and the DiffServ region belonging to the corresponding ANG is performed with a binding update procedure.

4.2 QoS Handling Mechanism for Micro Mobility Management

Unlike macro mobility, micro mobility means that a terminal moves from one cell to another adjacent cell while retaining its IP connectivity. Since a cell size will be smaller and smaller in the next mobile network, micro mobility occurs more frequently. Therefore, QoS handling for micro mobility management is a very important issue [17]. QoS negotiation in micro mobility is performed within an access network by using RSVP.

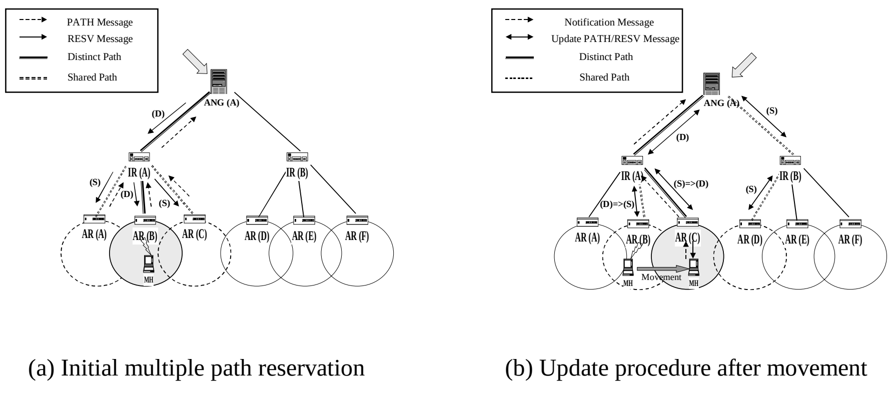
However, the standard RSVP does not consider mobility. So, we present a modified RSVP mechanism that is more suited to the mobile environment. In the modified RSVP mechanism, one PATH message reserves multiple paths. All paths are reserved only to access routers due to scarce resources in the wireless network. To minimize wasted wired network resources, we utilize two reservation styles in RSVP. Reservation in RSVP can be categorized as distinct and shared types [3]. Distinct reservation is appropriate for those applications in which multiple data sources are likely to transmit simultaneously. It requires separate admission control and queue management on the routers along its path to the receiver. On the other hand, shared reservation is appropriate for those applications in which multiple data sources are unlikely to transmit simultaneously.

In the proposed scheme, when a PATH message arrives at ANG, the ANG reserves one distinct path to the AR in the current location of the destination MH and multiple shared paths to the ARs of adjacent cells. Multiple shared paths are reserved by multicasting at the ANG.

Figure 5 presents the multiple path reservation mechanism. ANG(A) receiving a PATH message sends the PATH messages to several ARs. First, it unicasts a PATH message to the AR(B), which is an access router that attaches the destination MH. Then the MH sends an RESV message to the ANG(A) and the RESV message reserves a distinct path from the destination MH to ANG(A). In addition to unicasting, ANG(A) undertakes multicasting to adjacent ARs such as AR(A) and AR(C). AR(A) and AR(C) send the RESV message for shared path reservations.

When a MH enters into a neighbor cell, the reservation update procedure is performed as described in Figure 5. The MH sends the notification message to a new access router, AR(C), which relays the notification message to ANG(A), and ANG(A) then updates the routing path and multicasts the changed reservation styles. Using this procedure, the path styles to the AR(C) and AR(B) are changed into a distinct style and a shared style respectively. Since the AR(A) is not an adjacent access router with respect to the current access router, AR(B), the previously reserved resources are released. Also, a path to the new adjacent access router, AR(D), is reserved in shared style.

In our mechanism, multiple paths are reserved in not only the specified cell but also adjacent cells in advance so that it may provide the seamless QoS handling for micro mobility management.



(a) Initial multiple path reservation (b) Update procedure after movement

Fig. 5. Multiple Path Reservation Scheme

5 Performance Evaluation

In this section, we evaluate the proposed mechanisms. First, we simulate the class upgrade scheme for macro mobility using relatively long handoff latency. For an accurate simulation, we use a DiffServ patch in NS-2 (Network Simulator). Second, we analyze the performance of the multiple path reservation scheme with respect to, latency time and resource usages.

5.1 Macro Mobility Management

To demonstrate the influence of the class upgrade scheme in smooth handoff, we measure packet loss rates for packet forwarding between the previous ANG and the current ANG.

DiffServ model defines two types of PHB; Assured Forwarding (AF) and Expedited Forwarding (EF). AF PHB guarantees low delay and has allocated bandwidth according to SLA. AF is based on the RIO (RED with In and Out) scheme [18]. In the RIO scheme, packets are classified as in-profile or out-profile according to whether the traffic exceeds the bit rate specified by the SLA. During congestion the packets tagged as out-profile will be dropped first. On the other hand, EF PHB guarantees minimum delay, low latency, and peak-limited bandwidth. The EF traffic should have rates of independency of the intensity of other traffic attempting to transit the node. Therefore, it is handled with higher priority than AF. EF uses a priority queuing scheme with a token bucket. The token bucket is used to limit the total amount of EF traffic so that other traffic will not be starved by bursts of EF traffic.

In simulation, one DiffServ node has three independent queues for each PHBs; a RIO queue for AF PHB, a simple Drop-tail queue with a token bucket for EF, and a RED queue for best effort service. We controlled the conditioner located in the incoming node to adjust the number of forwarded packets due to mobility. The packet loss was then measured by changing a drop priority in one AF service class. Since EF packets are handled independently of the other packets, we only use AF packets in simulation. Figure 6 shows the simulation results.

In Figure 6 AF11 and AF12 belong to the same service class but have different drop priorities. AF12 has a higher drop priority than AF11. In simulation results, we know that the overall packet drop ratio is decreased by giving forwarded packets a lower drop priority, which means that the class upgrade scheme can meet the requirements of time-critical multimedia applications more efficiently.

5.2 Micro Mobility Management

In the case of micro mobility, we reserve multiple paths for seamless QoS handling. In addition, we utilize two reservation types to minimize resource usages. To evaluate

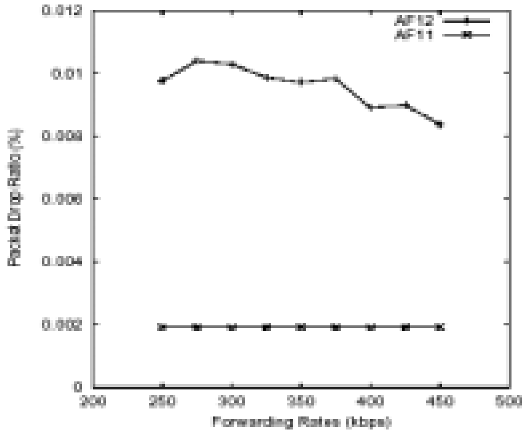


Fig. 6. Multiple Path Reservation Scheme

this scheme, we analyze the latency time of mobility quantitatively. For this analysis, we assume the N-dimension cell model and the K paths reservation scheme.

The total latency time of mobility is composed of the location setup time and the path setup time. Since all packets during the latency time are delivered to the previous AR or to the current AR without QoS limitations, the minimization of latency time becomes the most important problem in QoS handling in the wireless/mobile network.

Equation (1) and (2) show the latency time in the general RSVP scheme and in the proposed RSVP scheme respectively.

$$T_{LATENCY} = T_{UPDATE} + T_{PATH\_SETUP} \quad (1)$$

$$T_{LATENCY'} = T_{UPDATE} + T_{PATH\_UPDATE} \quad (2)$$

In Equation (1) and (2), T_{UPDATE} denotes the time of the location update using micro mobility management. For the general RSVP, a new path should be reserved after mobility, and therefore, a new path setup time is needed. Since the new path setup is done between two end hosts, T_{PATH_UPDATE} is proportional to the end-to-end round trip time in the overall network. On the other hand, in the proposed scheme, no new path is reserved and only path types are updated after mobility. All procedures for path updates are done within an access network so that T_{PATH_UPDATE} is proportional to the round trip time in an access network. In other words, the latency time in the proposed scheme is much less than that of the general RSVP. Therefore, packet losses and miss-deliveries due to extended latency time can be efficiently eliminated.

It might be expected that the reservation of multiple paths requires more network resources. However, in our scheme, since the shared reservation is used for adjacent paths, additional resource usages can be minimized. Specifically, resources in only non-overlapped cells and not in all adjacent cells are reserved. Equation (3) shows the resource usages when all adjacent paths are reserved as distinct types, and equation (4) shows the resource usages in the proposed scheme, which utilizes distinct and shared reservation types.

$$L \cdot (Rd + K \cdot Rs) \quad (3)$$

$$L \cdot Rd + K' \cdot Rs \quad (4)$$

L: Number of flows
K': Number of non-overlapped cells to be reserved
Rd: Resources for distinct reservation
Rs: Resources for shared reservation

Although the proposed scheme requires more resource than the general RSVP, it can provide a seamless QoS handling mobility mechanism. Furthermore, since the number of reserved adjacent cells can be adjusted according to mobility patterns and packet priorities, unnecessary resource usages can be diminished.

6 Conclusion

In this paper, we propose seamless QoS handling mechanisms in diverse mobility situations. We assume a combined architecture consisting of a DiffServ model and an IntServ/RSVP model. This design can provide a scalable per-flow QoS provisioning service in a wireless/mobile network.

For macro mobility, a new service level agreement and packet tunneling between ANGs are required, and accordingly, we propose a class update scheme to meet the QoS specification of time-critical multimedia applications. For micro mobility, a multiple path reservation scheme is used for seamless mobility management, which reserves multiple paths to keep QoS specifications and connections of a moving MH. Because we use distinct and shared reservation types, resource wastage is minimized.

The most important entities in our architecture are the ANGs. For macro mobility, they play the role of the foreign agent or home agent and negotiate SLA with incoming MH for packet tunneling in the DiffServ core network. In addition, they modify the service class of forwarded packets to meet delay constraints. Besides, for micro mobility, ANGs are a root node in a hierarchical access network so that they perform packet routing for seamless handoff and multicast for multiple path reservations. Since all nodes, except ANG, keep the standard DiffServ and IntServ/RSVP mechanisms, the proposed mechanism is so flexible and scalable. The performance evaluation results show that the proposed mechanisms are better than the existing mechanisms in the wireless/mobile QoS environment, and that they could be applied as a QoS model in the next generation of wireless and mobile networks.

References

1. I. Mahadevan and K. Sivalingam, "Quality of Service Architecture for Wireless Networks: IntServ and DiffServ Models," I-SPAN, 1999.
2. R. Braden et al., "Resource ReServation Protocol (RSVP) - Version 1 Functional Specification," IETF RFC 2205, Sep. 1997.
3. R. Braden et al., "Integrated Services in the Internet Architecture: an Overview," IETF RFC 1663, June 1994.
4. S. Blake et al., "An Architecture for Differentiated Services," IETF RFC 2475, Dec. 1998.
5. S.U. Yoon et al., "QoS Support in Mobile/Wireless IP Networks using Differentiated Services and Fast Handoff Method," WCNC 2000.
6. K.E. Malki et al., "Hierarchical mobile IPv4/v6 and fast handoffs," Internet draft, draft-elmalki-soliman-hmipv4v600.txt, March. 2000
7. W.T. Chen et al., "RSVP Mobility Support: A Signaling Protocol for Integrated Services Internet with Mobile Hosts," INFOCOM 2000.
8. A. Valk, "Cellular IP: A New Approach to Internet Host Mobility," ACM Computer Communication Review, January 1999.
9. R. Ramjee et al., "IP-Based Access Network Infrastructure for Next-Generation Wireless Data Networks," IEEE Personal Communications, Aug. 2000.
10. A. Oneill et al., "Host Specific Routing," Internet draft, Nov. 2000.
11. C.E. Perkins et al.," Optimized Smooth Handoffs in Mobile IP," Computer and Communications, 1999.
12. J. Solomon, "Mobile IP: The Internet Unplugged," Prentice Hall, Sep. 1997.
13. X. Xiao et al., "Internet QoS: A Big Picture," IEEE Network, Mar. 1999.
14. G. Mamais et al., "Efficient Buffer Management and Scheduling in a Combined IntServ and DiffServ Architecture: A Performance Study," ICATM, 1999.
15. R. Caceres et al., "Fast and Scalable Handoffs for Wireless Internetworks," MOBICOM, 1996.
16. D. Johnson et al., "Route Optimization in Mobile IP," Internet Draft, Nov. 1998.
17. A. Cambell et al., "IP Micro-Mobility Protocols," Mobile Computing and Communication Review, Oct. 2000.
18. D. Clark et al., "An Approach to Service Allocation in the Internet," Internet draft, draft-clark-different-svc-alloc-00.txt, July 1999.