All-IP 무선망을 위한 에이전트 기반의 멀티캐스트 핸드오프 메커니즘
(Agent based Multicast Handoff Mechanism for All-IP Wireless Network)

김 병 순 † 한 기 준 ††
(Byung-Soon Kim) (Ki-Jun Han)

요 약 이 논문은 IP 멀티캐스트 상에서 핸드오프 지연을 줄이기 위해 Multicast Handoff Agent (MHA)라는 에이전트에 기반을 둔 새로운 멀티캐스트 핸드오프 메커니즘을 제안한다. MHA는 기지국에서 이동 호스트의 Internet Group Management Protocol (IGMP)에 대한 프록시로서 동작되고 셀 내의 멀티캐스트 그룹의 구성원에 대한 정보를 유지하는 역할을 한다. 이동 호스트가 다른 셀로 이동할 때, MHA는 IGMP query 메시지를 기다리지 않고 즉시 리포트 메시지를 전송한다.

제안하는 메커니즘은 시뮬레이션과 분석을 통해 성능평가 되고 마이크로 이동성과 IGMP 트래픽에 대하여 IGMPv2와 비교분석 한다. 시뮬레이션 결과로서 마이크로 이동성에 대한 핸드오프 지연이 크게 줄일 수 있고 또한 그룹의 구성원으로 있는 동안 무선 링크상의 IGMP 제어 트래픽을 제거될 수 있음을 보인다. 따라서 제안하는 메커니즘은 마이크로 이동성에 대해 기존의 메커니즘보다 뛰어나고, IGMP 질의에 대한 응답이 불필요하여 배터리 지속시간이 오래갈 수 있다.

키워드 : MHA, 멀티캐스트 핸드오프, 마이크로 이동성, IGMP, IP 멀티캐스트

Abstract This paper proposes a new agent based mechanism called a Multicast Handoff Agent (MHA) to reduce handoff latency for IP multicast. The MHA acts as a proxy for an Internet Group Management Protocol (IGMP) of Mobile Nodes (MNs) at each Base Station (BS) and keeps information for members of multicast groups in a cell. When an MN moves to a next cell, the MHA immediately sends unsolicited reports without waiting for the IGMP query.

The mechanism was evaluated through simulation and analysis and compared with the IGMPv2 for micro-mobility and the IGMP traffic. Simulation results show that handoff latency for micro-mobility can be largely reduced and the IGMP control traffic on the wireless links during the duration of membership can be eliminated. Thus, this mechanism is superior to the existing mechanism in both micro-mobility and battery duration, as the need not to reply to a query conserves battery power.

Key Words : MHA, Multicast Handoff, micro-mobility, IGMP, IP multicast

1. Introduction

Multicasting is an efficient paradigm for transmitting data from a sender to a group of receivers called "group members". Multicasting is much more advantageous than multiple unicasts as it reduces the communication costs. Multicast communications are widely utilized in various applications including information dissemination, multimedia conferencing, shared whiteboards, multicast file transfer, multi-party games and distributed computing.

A wide spectrum of portable, personalized computing devices ranging from laptop computers to handheld personal digital assistants have recently

† 비 회 원 : 대구공업대학 컴퓨터정보계열 교수
bskim@ttc.ac.kr
†† 종신회원 : 경북대학교 컴퓨터공학과 교수
kjhan@bh.kyungpook.ac.kr
논문접수 : 2001년 6월 20일
심사완료 : 2002년 1월 17일

been introduced. Their explosive growth has led to considerable interest in providing continuous network coverage to such MNs regardless of their location. Wireless communications have attracted interest in the integration of wireless networks with the Internet. More and more users would like to maintain Internet access without disruptions while in transit. This can be achieved by Mobile IP [1] that hides mobility from the transport services.

The IP multicast provides a mechanism for location independent addressing and packet delivery to a group of hosts that belong to a multicast group. It also provides efficient mechanisms for hosts to join and leave multicast groups. In Mobile IP, an MN must join the multicast group via either a local multicast router (MR) on the visited subnet or a bi-directional tunnel to its Home Agent, assuming that its HA is an MR [1]. When the Foreign Agent (FA) is an MR, it gathers the IGMP information from the MNs and multicasts the datagrams.

When an MN is a multicast group receiver, it experiences disruption in the reception of the multicast stream immediately after handoff due to the multicast handoff latency. The handoff latency for the IP multicast varies from zero up to minutes. This paper examines the problem of the inefficient IP multicast handoff with the IGMPv2 [2] and proposes an agent based mechanism to reduce the latency. Several studies in [3]-[5] approached a pre-join mechanism that establishes the new multicast tree before an MN moves to other cell in order to reduce latency. Unfortunately, these cannot eliminate an overhead on the wireless links due to the periodic IGMP messages. And these studies were assumed that the MNs join the group via the tunnel.

This paper is organized as follows: Section 2 presents IP multicast mechanism, Section 3 provides detail into how long handoff latency takes in IP multicast. Section 4 presents a new mechanism that reduces the latency and eliminates the IGMP control traffic on the wireless links. The evaluation of the mechanism is presented in Section 5 and finally, Section 6 concludes the paper.

2. IP Multicast

IP multicast is based on the concept of the host group [6]: a dynamic set of hosts identified by a single class D IP address. A host can join or leave a group at any time in order to start or stop receiving datagrams sent to the group. IP multicast consists of two parts: group membership management and multicast routing. The group management protocols collect the local membership information in the leaf networks. This information is later used by the routing protocols such as DVMRP [7], CBT [8], MOSPF [9], PIM-DM [10] and PIM-SM [11] and reflects the dynamic changing nature of membership in the local networks.

The IGMP is responsible for locating hosts belonging to a multicast group. The IGMP is used by the routers to periodically check whether the known group members are still active for Local Area Networks (LANs), where packets are broadcasted on the physical medium. If more than one MR exists in its local area, one of the routers is elected as the querier and assumes the responsibility of keeping track of the membership state of the multicast groups that have active members [2]. The querier periodically sends the IGMP queries for group membership in the local area and the attached hosts reply with the IGMP reports identifying the groups in which they participate. The querier consequently constructs a membership table to effectively manage a number of multicast groups.

These IGMP queries are periodically repeated and if no reports are received for a previously reported group after a number of queries, the group is removed from the membership list. When joining a group, each host sends a number of unsolicited reports to reduce join latency, i.e. the time between a host joining the group and the router starting multicast propagation towards the host. In the IGMPv1 [6], after the last member of a group departs, datagrams for the group are still forwarded to the LAN until membership times out, thus wasting bandwidth. The time between these two events is called the leave latency. To reduce this, a host in the IGMPv2 sends a leave message when abandoning a group when it was the last host to send a membership report for that group. While join latency is avoided by unsolicited reports, the IGMPv2 leave messages reduce but do not eliminate leave latency, since they are only hints rather than authoritative information.

3. Micro-Mobility for IP Multicast

We assume that all of the FA is an MR and an MN must join the multicast group via a local MR. Also, assume that an MN may join a number of multicast groups. When an application in the MN wishes to join a multicast group, an unsolicited IGMP membership report is sent. Thus, the joining delay is reduced by using the event triggered protocol messages. When an MN moves to other cell in a subnet, the MN can no longer receive multicast traffic in the next cell if there is no active member of the group in that cell. In order to establish a new path from the FA to the entering MN, the MN must wait for the query message sent by the FA, as neither the multicast applications nor the IGMP has the mechanism to detect the handoff. The multicast handoff latency in such a handoff is sometimes unacceptably long [5, 12] and as a result, the handoff for the IP multicast may cause potential heavy traffic loss.

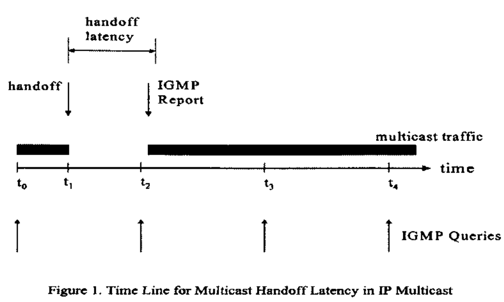

Figure 1 Time Line for Multicast Handoff Latency in IP Multicast

Figure 1 shows a time line for a handoff in IP multicast. After a handoff at a time t_1, the MN must wait for a new IGMP query for group membership until a time t_2. It is this message that triggers the MN to send an IGMP report to the FA. After receiving the IGMP query, a delay timer is set up. An IGMP membership report will be sent when the timer expires. It is assumed that the value of the timer is α. Thus, the multicast handoff latency for micro-mobility will be $(t_2-t_1)+\alpha$. The handoff in IP multicast will probably result in a large traffic gap in receiving, since the average waiting time for an IGMP membership query by an MN after a handoff is approximately half of the IGMP query membership interval. The waiting time for a query message could range between 0 and 125 seconds since the IGMP query interval is by default 125 seconds. Further, after the MN receives the query, it will have to back off for another random time which ranges between 0 and 10 seconds [2]. It is found that the MN will wait for approximately 67 seconds on average to resume receiving the multicast traffic after the handoff. Thus, the moved MN will have a big multicast traffic gap in the next cell.

As mentioned before, the MR sends and receives the IGMP query/report messages for the group membership periodically. This is of no concern for a fixed network that has sufficient bandwidth. However, this results in an overhead on the wireless links that have narrow bandwidth. In addition, the periodic IGMP queries prohibit MNs from using the sleep mode to conserve battery power when no multicast traffic is presented. Therefore, a mechanism is required to reduce the multicast handoff latency in the IP multicast and the overhead on the wireless links.

4. Multicast Handoff Agent

To support multicast micro-mobility in an all-IP mobile network, periodic IGMP query/report messages can be sent only in the fixed network and an MR is able to know information about handoff of MNs before the next IGMP membership cycle. We can eliminate the traffic on the wireless links due to the IGMP and also reduce the multicast handoff latency.

We propose a new mechanism called Multicast Handoff Agent (MHA) to reduce the multicast handoff latency and the control overhead. This agent will be located at each BS and acts as a proxy for the MNs to the IGMP query message sent by the MR. This agent has a Multicast Group Membership Table (MGMT) which keeps group membership information for multicast groups in its cell using a Multicast Group Identifier (MGID) and a Mobile Node Identifier (MNID).

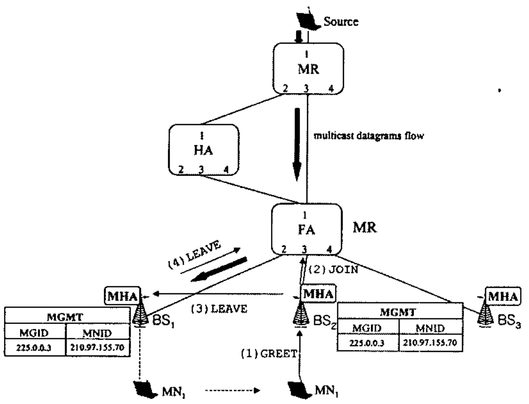

Figure 2 Proposed Mechanism for Multicast Micro-Mobility of Receiver

Figure 2 shows the states of MGMT and exchanged messages for micro-mobility. Before MN_1 moves, the MHA at BS_1 has membership information for it. When MN_1 moves from BS_1 to BS_2, a GREET message (1) is sent to BS_2 including MGID, MNID and IP address of BS_1. The MHA at BS_2 then stores the values of MGID and MNID in its MGMT. Following this, a JOIN message (2) is sent to the FA. If the GREET includes a number of group addresses, it must send a JOIN for each group since the IGMPv2 includes a single group address. At this time, the FA adds MN_1 as a new entry to the membership table of MR. So, the FA can now send the multicast datagrams to both interface 2 and 3 for the group member. After a new path from the FA to MN_1 has been established, BS_2 sends a LEAVE message (3) to BS_1 to remove MN_1 from the MGMT at BS_1, and then forwards the LEAVE (4) to the FA. The old entry will be removed from the MGMT at BS_1.

Group ID	Membership List	Interface
225.0.0.3	210.97.155.70	2

(a) Before Handoff

Group ID	Membership List	Interface
225.0.0.3	210.97.155.70	2
225.0.0.3	210.97.155.70	3

(b) While Handoff

Group ID	Membership List	Interface
225.0.0.3	210.97.155.70	3

(c) After Handoff

Figure 3 Membership Table of MR

Figure 3(a) shows the membership table when the datagrams are sent through interface 2 of the MR before MN_1 moves. When the FA receives the IGMP report message from the MHA at BS_2, it adds a new entry as shown in Figure 3(b). Now, MN_1 can receive the datagrams through interface 3. After that, the old entry will be removed as illustrated in Figure 3(c).

If an MN wants to join a group, it sends a JOIN message to the MHA. Then, the MHA stores the MN's information and forwards it to the FA. Alternatively, when an MN wishes to leave a group, the MHA deletes its information from the MGMT and forwards the message to the FA.

Using the MHA mechanism, the new MNs do not need to wait for the IGMP query to be sent by the MR and to back off for a random time. In addition, since the MHA receives periodic IGMP queries from the MR for group membership, it can reply to the queries by simply looking up the MGT without any interaction with the MNs on the wireless links.

5. Performance Analysis

For a performance analysis of the MHA, some assumptions were made. First, it is assumed that no protocol message is lost and each IGMP datagram is 256 bits long (20 bytes for the IP header, 8 bytes for the IGMP payload and 4 bytes for link layer overhead). Multicast handoff call arrivals form a Poisson process with an average of 0.5 calls/minute while a dwell time of the MN is assumed to be exponentially distributed with an average value of μ at 10 minutes. For simplicity, the processing time of the IGMP messages and the executing time at the MHA are ignored. Some important parameters used for performance evaluation are as follows:

B : the transmission rate
T_I : the IGMP query interval
T_T : the random backoff time
n : the number of multicast groups joined by the MN
T_l : the radio link establishment delay
T_d : the transmission delay of the wired link
mT_d : the transmission delay of the wireless link
P_m : the probability that the MN belongs to the group established in the new cell

The multicast handoff latency with the IGMPv2 for micro-mobility, denoted by T_L, is given by:

$$T_L = T_l + 2mT_d + ((mT_d + T_d) + n * (mT_d + T_d + E[T_I]) + E[T_T])(1 - P_m) \quad (1)$$

where $E[.]$ means the mean value.

In contrast, the latency with the MHA mechanism, denoted by T_P, is given by:

$$T_P = T_l + 2mT_d + (mT_d + nT_d)(1 - P_m) \quad (2)$$

From these two equations, it can be seen that our mechanism provides a much shorter latency than the IGMPv2.

Let's consider the IGMP control traffic overhead while MNs dwell in cells. The number of IGMP queries during the duration of membership in a cell is $\frac{\mu}{T_I}$ and the maximum amount of received data in a cell is μB bits. So, the IGMP control overhead normalized by the transmission rate, denoted by O_I, is given by:

$$O_I = \frac{512}{BT_I}. \quad (3)$$

A simulation study was performed to verify the analytical models. The same assumptions are used for simulation as for the mathematical analysis. Also, it was assumed that the number of cells is 100, T_l=3 ms, T_d=1 ms, and m=5. We consider a single multicast group, but the multicast establishing time when the MN moves to other subnet was not considered. A lot of simulation runs are performed with various sizes in the multicast group ranging from 5 to 60, that is, the group density from 0.05 to 0.6.

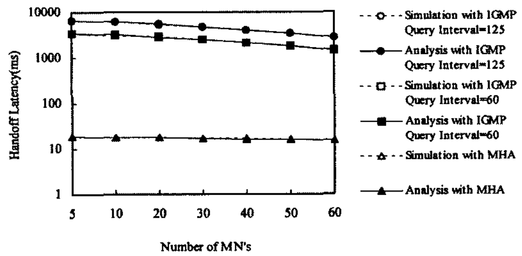

Figure 4 Multicast Handoff Latency for the IGMP and the MHA

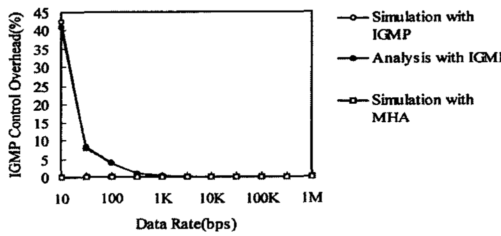

Figure 5 IGMP Control Overhead for the IGMP and the MHA as a Function of Data Rate

In Figure 4, we can see that the simulation results are in a sound agreement with the analytical model given by (1) and (2). As shown in this figure, when the size of the group is 5, the IGMPv2 offers the latency of about 64 and 34 seconds with the IGMP query interval values of 125 and 60 seconds, respectively. Our scheme provides a much shorter latency at only 19 ms. We can see that as the size of the multicast group increases, the latency decreases, sharply.

Figure 5 shows the IGMP control overhead on the wireless links as the data rate varies. It can be seen that the IGMP control overhead is considerably large under 100 bps and the result is also in a sound agreement with the analytical model given by (3).

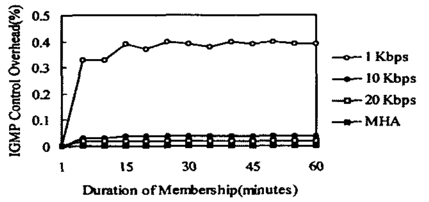

Figure 6 IGMP Control Overhead for the IGMP and the MHA as a Function of the Duration of Membership

Figure 6 shows the IGMP control overhead on the wireless links as the duration of the membership varies. When the duration of membership is under 5 minutes, the IGMP control overhead sharply increases.

From both Figure 5 and Figure 6, it can be seen that there is some overhead on the wireless link due to the periodic queries/reports in the IGMP. As the transmission rate or the query interval decreases, the overhead increases. However, there is no overhead in the MHA since it does not send any queries to the MNs.

6. Conclusion

Some of the problems with existing implementations of the IP multicast have been presented. When an MN belong to a multicast group moves to other cell in an all-IP wireless network, an MR finds its arrival after some delay in the next IGMP membership cycle. Thus, due to this, the multicast handoff latency becomes unacceptably long.

We propose the MHA mechanism that does not wait for a IGMP query for the micro-mobility support in order to reduce the handoff latency. The MHA approach was compared with the standard query/report approach with regard to protocol performance, as measured by transmission overhead and handoff latency. Also, it can be seen that the simulation results are in agreement with the results obtained from the presented previously analytical model. From these results, it is clear that the MHA mechanism is superior to existing models for micro-mobility.

Since MHA is used rather than the host replied report, the battery powered MNs can employ sleep mode when there is no traffic of interest. The hosts only have to wake up in order to process multicast data, not to reply to queries by the group management protocol.

References

[1] C. Perkins, "IP mobility support," RFC 2002, Oct. 1996.

[2] W. Fenner, "Internet Group Management Protocol, Version 2," RFC 2236, Nov. 1997.

[3] Yutaka Ezaki and Yuji Imai, "Mobile Ipv6 handoff by Explicit Multicast," Internet Draft (work in progress), Nov. 2000.

[4] Jiang Wu and Gerald Q. Maguire Jr., "Agent Based Seamless IP Multicast Receiver Handover," IFIP Conference on Personal Wireless Communications PWC'2000, Sep. 2000.

[5] Jiang Wu, "Seamless IP multicast Receiver Mobility Support," Internet Draft (work in progress), Oct. 2000.

[6] S. Deering, "Host Extensions for IP Multicasting," RFC1112, Aug. 1989.

[7] S. Deering and C. Partridge, "Distance vector multicast routing procotol," RFC 1075, Nov. 1988.

[8] A. Ballardie, J. Crowcroft and P. Francis, "Core based trees (CBT)-An architecture for scalable inter-domain multicast routing," Proc. of the ACM SIGCOM Vol.23, pp. 85-95, Oct. 1993.

[9] J. Moy, "Multicast extensions to OSPF," RFC 1584, Mar. 1994.

[10] S. Deering, D. Estrin, D. Farinacci, V. Jacobson, A. Helmy and L. Wei, "Protocol independent multicast version 2, dense mode specification," Internet Draft, May 1997.

[11] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma and L. Wei, "Protocol independent

multicast sparse mode (PIM-SM): protocol specification," RFC 2362, Jun. 1998.
[12] George Xylomenos and George C. Polyzos, "IP Multicast Group Management for Point-to-Point Local Distribution," Computer Communications, Vol. 21, No. 18, pp.1645-1654, 1998.

김 병 순
1991년 서강대학교 전자계산과(공학사). 1993년 서강대학교 전자계산과(공학석사). 1993년 ~ 1996년 현대전자 정보시스템사업본부. 2000년 경북대학교 컴퓨터공학과 박사과정 수료. 1996년 ~ 현재 대구공업대학 컴퓨터정보계열 교수

한 기 준
정보과학회논문지 : 정보통신
제 29 권 제 1 호 참조