MobiCast: A Multicast Scheme for Wireless Networks

Cheng Lin Tan [a]

[a] *Centre for Wireless Communications, National University of Singapore*
Computer Science & Electrical Engineering, Luleå University of Technology
E-mail: chenglin@acm.org

Stephen Pink [b]

[b] *Computer Science & Electrical Engineering, Luleå University of Technology*
E-mail: steve@cdt.luth.se

1. Introduction

As wireless technology matures, small wireless cells at the fringes of the wired network are fast becoming part of the Internet infrastructure to connect users to the wired internetwork. These wireless cells enable mobile users to enjoy the benefits of mobility within the coverage area while having access to the Internet. Each wireless cell has a base station which acts as the gateway to the wired network for the mobile nodes. These wireless cells tend to be small so as to provide higher data throughput, better frequency reuse, more accurate location information, and the power requirement of the transceiver at the mobile node is lower. The other emerging trend is the increasing importance of multicast applications for the purpose of multiparty conferencing [16] [14], and mobile users will expect similar kinds of multicast applications to be supported on their portable devices with wireless capability. In such an environment, handoffs between cells during a multicast session are common. Hence an efficient scheme is needed to support multicast for mobile nodes roaming among small wireless cells. This scheme should minimize re-computation of the multicast delivery tree and reduce packet loss when a mobile group member crosses cell boundaries during a multicast session.

Providing multicast support for mobile nodes in an IP internetwork is a challenging problem for several reasons. First, the addition of mobility to the host group model [9] implies that the multicast routing algorithm must deal not only with dynamic group membership, but also with dynamic group member location. Second, most multicast routing protocols, such as DVMRP [8], CBT [2] and MOSPF [17], implicitly assume that hosts are static when setting up a multicast delivery tree. Reconstructing the delivery tree every time a mobile node moves is expensive due to the overhead incurred, but leaving the multicast tree unchanged may result in inefficiency or even failure in multicast data delivery. Third, when the mobile node is a multicast source, the multicast stream to other group members is often disrupted after the mobile node moves to a new cell. Besides handoff latency, the new base station also has to find a route to the multicast delivery tree. On the other hand, when the mobile node is a multicast group receiver, it experiences disruption in the reception of the multicast stream immediately after the handoff due to handoff latency as well as the join and graft latencies involved in the new base station's subscription to the multicast group.

In this paper, we propose a multicast scheme known as MobiCast that is suitable for mobile hosts in an internetwork environment with small wireless cells. Our scheme adopts a hierarchical mobility management approach to isolate the mobility of the mobile hosts from the main multicast delivery tree. Each foreign domain has a domain foreign agent. To send a multicast packet, the mobile host encapsulates the multicast packet and sends it to the domain foreign agent. The domain foreign agent decapsulates the multicast packet and sends it out on behalf of the mobile host. A subscription to a multicast group by a mobile host is relayed by the serving base station to the domain foreign agent. The domain foreign agent subscribes to the requested multicast group and forwards the requested multicast packets to the mobile host using another multicast address known as the translated multicast address. This translated multicast address should be unique within the domain and corresponds to this requested multicast group. The base station receives the multicast packets by subscribing to the translated multicast group (the multicast group with the translated multicast address), and forwards the received multicast packets to the mobile host. As long as the mobile host remains within the domain of the domain foreign agent, the mobility of the mobile host is hidden from the rest of the multicast group, and no re-computation of the main multicast delivery tree is needed.

Our scheme minimizes disruptions to the multicast session due to handoffs of mobile group member by organizing physically adjacent cells into Dynamic Virtual Macro-cells (DVM). When a mobile host subscribes to a multicast group through the domain foreign agent, the serving base station informs the other member base stations in its DVM to subscribe to the same translated multicast group. While only the serving base station actively forwards multicast data to the mobile host, the other base stations in the same DVM buffer recent packets and quickly forward them to the mobile host should a handoff occur. This provides a short handoff latency, and the use of buffers at the BSs reduces packet loss due to handoff. It also eliminates multicast group join and graft latencies since the new base station has already subscribed to the multicast group prior to the handoff. Hence the disruptions to the multicast session due to handoffs of mobile host group members are minimized.

Our scheme is developed to work with IP and is compatible with existing multicast routing algorithms such as DVMRP [8], CBT [2], MOSPF [17], PIM-DM [7] and PIM-SM [6]. Our base stations are network-layer routers with buffers, and are capable of subscribing to multicast groups. Compared to a link-layer solution adopted by most wireless LAN, a network-layer base station is capable of forwarding only those multicast packets with interested mobile receivers in its cell, thus achieving efficient utilization of wireless bandwidth. Furthermore, a network-layer base station is able to differentiate packets with different service types for IPv6 so as to support QoS for mobile hosts in its cell. Our scheme aims to support best effort IP multicast efficiently for mobile hosts in an environment with small wireless cells, while maintaining the quality of the multicast session during handoffs.

We have simulated our scheme using the Network Simulator [20] and the measurements show that our multicast scheme is effective in minimizing disruptions to a multicast session due to the handoffs of the mobile group member, as well as reducing packet loss when a mobile host crosses cell boundaries during a multicast session.

The rest of our paper is organized as follows. Section 2 presents related work. Section 3 describes our hierarchical mobility management architecture to support mobile multicast. Section 4 presents the simulation of our mobile multicast scheme using the ns tool. Section 5 reports the performance measurement results of our simulation. Section 6 highlights some areas for future work, and Section 7 concludes the paper.

2. Related Work

The problems of providing multicast support for mobile hosts have been studied by various research groups. Xylomenos and Polyzos highlight some of these problems in supporting multicast for mobile hosts in [22] and [21]. Arup et al. [1] propose solution to solve the conflict of combining IP multicast with the Columbia University Mobile-IP scheme [12] [13]. The Columbia Mobile IP scheme allocates care-of address from a virtual subnet to mobile hosts (MHs), and different MHs may attach to the wired network via different Mobility Support Routers (MSRs) located at different subnets. This often results in multicast packets from MHs being discarded by routers as the packets arrived from the wrong interface. Their proposed solution is to use a multicast tunnel to connect all the MSRs located at the different subnets. Multicast packets from a MH is encapsulated by the local MSR and sent to all MSRs through this multicast tunnel to ensure that this multicast packet reaches all MSRs and receivers. Some of the problems raised are peculiar only to Columbia University Mobile IP scheme and do not apply to the IETF Mobile IP model. Furthermore, the scheme requires modifications to IGMP and is limited in terms of scalability.

In the IETF Mobile IP standard [18], two approaches, namely, remote subscription and bi-directional tunneling, are proposed to support multicast for MHs. In the remote subscription approach, the MH has to re-subscribe to the desired multicast groups when it visits a foreign network. The main advantage of this scheme is that it is simple and offers the shortest path for the multicast data delivery to the MH. This scheme works well when the MH stays at each foreign network for a relatively long time, as compared to the join and graft latency. The main disadvantage of this is that if the MH is to send a multicast packet to the group, the source address of this packet has to be the MH's care-of address. If the home address of the MH is used instead (as required by some multicast applications), the incoming interface check of most multicast routers will discard this packet.

For the bi-directional tunneling approach, MH receives and sends multicast packets through its home agent (HA) via unicast IP tunnel. The advantage of this approach is that it hides the mobility of the mobile group members from the rest of the multicast group. However, this scheme has some serious drawbacks. First, the routing of the multicast packets may be sub-optimal, especially when the multicast source and the MH are near to each other, but both are far away from the MH's HA. Second, the scheme can be bandwidth inefficient. If the HA has two mobile hosts which reside at the same or nearby foreign networks and subscribe to the same multicast group, the HA has to forward the same multicast packet twice to the foreign networks located at the same region via two different IP tunnels to the respective MHs, resulting in a waste of network bandwidth and extra processing at the HA. Third, the bi-directional tunneling approach also suffers from the tunnel convergence problem [4] due to the numerous multicast IP tunnels from the different HAs terminating at a particular foreign agent (FA). These home agents all have mobile hosts residing at this FA's network and subscribed to the same multicast group. This again results in inefficient use of network bandwidth as well as computing resources. We believe that the bi-directional tunneling approach is not a good idea as multicast forwarding for MHs then degenerates into multiple unicasts, and is likely to face serious scalability problems.

As noted in [15], that while Mobile IP (and IPv6) is designed to handle the "macro" mobility management problem, such as supporting host mobility over wide-area networks, it does not address micro-level mobility issues such as packet loss and delay due to handoffs. Hence using either the remote subscription or the bi-directional tunneling approach recommended by Mobile IP alone is not a good enough solution to support multicast in an environment of MHs roaming among small wireless cells.

Harrison et al. [11] propose the MoM protocol to solve some of the problems associated with the bi-directional tunneling approach of the IETF Mobile IP (Mobile IP). In the MoM protocol, when a HA has more than one MHs residing at the same foreign network subscribed to the same multicast group, only one copy of the multicast data is forwarded from the HA to the FA. The protocol solves the tunnel convergence problem [4] by having the FA appoint a HA as the designated multicast service provider (DMSP) for a given multicast group to forward multicast packets to that FA. This avoids duplicating multicast packets from the different IP tunnels originating from the different HAs. This scheme eliminates some of the inefficiencies of the bi-directional tunneling approach of Mobile IP. However, this scheme has various drawbacks. First, the scheme still does not address the problem of a HA having to forward duplicate multicast packets to neighboring foreign networks when the HA is chosen as the DMSP for both FAs. Second, if the number of HAs eligible to be chosen as the DMSP is small, the route for the multicast packets from the DMSP to the FA may still be sub-optimal. Third, when the last MH of a DMSP of a given multicast group roams out of the FA's network, the FA has to reselect a new DMSP. Meanwhile, the multicast data delivery to all other mobile group members served by that FA is disrupted until the DMSP handoff is completed. Fourth, by adopting the bi-directional tunneling approach, the scheme avoids the delays associated with joining and grafting, but not the handoff latency. The problem of packet loss when mobile host crosses cell boundaries is not addressed in this scheme. Finally, the scheme requires modifications of both HA and FA codes.

In contrast to the MoM protocol, our scheme adopts an approach similar to the remote subscription method recommended by the Mobile IP. However, we use a hierarchical mobility management architecture to shield the mobility of the MHs from the multicast delivery tree. MH subscribes to a multicast group via the domain foreign agent. This eliminates sub-optimal routing and avoids tunnel convergence problem. Unlike the original Mobile IP remote subscription method, no re-computation of the main multicast delivery tree is needed due to mobility of mobile group member within the domain. We minimize disruption to multicast session due to handoffs by forwarding the multicast data from the domain foreign agent to multiple base stations in the vicinity of the MH to achieve fast handoff performance. This is again something which is not addressed by the original Mobile IP remote subscription approach. We aim to show that our scheme is efficient and suitable for handling multicast for mobile hosts in an environment with small wireless cells.

3. Mobility Architecture

A typical campus network with wireless extensions is shown in Figure 1. The various network segments are interconnected by routers (R), and mobile hosts (MH) access the Internet via the base stations (BSs) over the wireless links. As each MH moves, it leaves the wireless coverage of one cell and enters into another, resulting in a handoff between the BSs. For our scenario of small wireless cells at the fringes of the Internet, such handoffs during a multicast session will be frequent as wireless cells may be of the size of a few meters.

In Mobile IP, after a MH arrives at a foreign network and obtains a care-of address, the MH sends a location update message to inform its home agent (HA) of its care-of address. The care-of address identifies a foreign agent (FA) in the foreign subnet where the MH is. The FA can be a separate node or reside in the MH.

In the Mobile IP remote subscription method of supporting multicast, the MH has to subscribe to the desired multicast group via FA1 when it is at subnet A, and via FA2 when it moves to subnet B. Immediately after a handoff from one subnet to the other, the multicast session to the MH is disrupted until the handoff is completed. The handoff latency may be long as the location update message has to traverse the wide-area network from the new FA to the HA.

In the Mobile IP bi-directional tunneling method of supporting multicast, the HA intercepts packets from multicast groups that the MH has subscribed to and tunneled them to the care-of address of the MH. If the foreign agent is a separate node, it decapsulates the tunneled packets and forwards them to the MH. This method has various inefficiencies and scalability problems as explained in the earlier section. It also suffers from disruption to the multicast session due to long handoff latency when the MH roams from one subnet to another.

Our approach avoids these drawbacks while providing the desired performance of a mobile multicast scheme suitable for mobile nodes roaming among small wireless cells.

3.1. Hierarchical Structure

We propose to adopt the domain foreign agent (DFA) concept [3] [19] to shield all mobility within the foreign domain from the main multicast delivery tree. For the purpose of exposition, we will use the scenario of a campus environment. As shown in Figure 2, a campus domain may have a DFA which is responsible for all foreign mobile hosts within the campus. At the subnet level, agent advertisement messages containing the IP address of the DFA are broadcast periodically. Note that this functionality of broadcasting an agent advertisement may be subsumed by the BSs for our case. When a mobile node hears a beacon and decides to attach to the wired network, it registers with the DFA and sends the IP address of the DFA to its HA as its care-of address. Subsequent multicast subscriptions or sending of packets to a multicast group by the MH are done through the DFA. When the MH moves from one cell to another, resulting in a change of serving BS (but still within the campus), this change is shielded from the rest of the multicast group. No re-computation of the main multicast delivery tree is needed due to mobility of the MH within the domain. The details of how our hierarchical mobility management architecture handles unicast communications for the MH is not within the scope of this paper, and can be found in [19].

3.2. Mobile Host as a Multicast Sender

If the MH is a multicast source of a multicast group, the MH cannot use its care-of address (DFA's address) as the source address for its multicast packet, as this packet may be discarded during the incoming interface check by a multicast router (e.g. if DVMRP is used). The MH cannot use a co-located care-of address to avoid this problem since our mobile multicast scheme requires the use of DFA's address as the MH's care-of address to handle mobility. Our scheme overcomes this problem by requiring the MH to encapsulate its multicast packet and unicast it to the DFA. The DFA decapsulates the packet and sends out the multicast packet on behalf of the MH, with the DFA's address in the source address field of the multicast packet. Here we are assuming that the higher level protocols in the application will provide the identification of the original source of the packet. This is the preferred method for our scheme. If this is not possible because the IP address of the originating node is required by the application to be at the IP source address field of the multicast packet, the other alternative is for the MH to tunnel back the multicast packet to its HA and send out the multicast packet via the HA. Since this is a unidirectional tunnel from the MH to the HA, bandwidth wastage and the tunnel convergence problem associated with bi-directional tunneling are avoided. The disadvantage is the possibility of inefficient routing of multicast packets from the MH to the multicast delivery tree. This happens when the MH is near the multicast delivery tree, but has to tunnel its multicast packet to its HA.

3.3. Mobile Host as a Multicast Receiver

To support a MH which is a multicast receiver, our scheme works in the following manner. The MH sends a IGMP report [5] to the BS to subscribe to multicast group X. This subscription is relayed by the serving BS to the registered DFA at the vicinity of the MH. Upon receiving the subscription request, the DFA supplies a different multicast address known as the translated multicast address to the BS. The DFA subscribes to the requested multicast group X and forwards the requested multicast packets to the MHs in its domain using the translated multicast address. BSs with MHs which are members of the multicast group X receive the multicast packets by subscribing to the translated multicast group (the multicast group with the translated multicast address), and forward the received multicast packets to the interested MHs in their cells. As long as the MH remains within the domain of the DFA, the mobility of the MH is hidden from the rest of the multicast group, and no re-computation of the main multicast delivery tree is needed.

3.4. Handoff Protocol

To achieve fast handoff performance so as to reduce disruption to multicast reception when a MH crosses cell boundaries, we organized physically adjacent cells into Dynamic Virtual Macrocells (DVM). When a serving BS relays a multicast group subscription request of a MH to the DFA, a translated multicast address is given to the serving BS by the DFA. Besides subscribing to the translated multicast group itself, the serving BS also informs the other member BSs in its DVM to subscribe to the same translated multicast group. While only the serving BS actively forwards multicast data to the MH, the other BSs in the same DVM buffer recent multicast packets and can quickly forward them to the MH should a handoff occur. The thick arrows from the DFA to the various BSs in Figure 3 show the paths of the translated multicast group data from the DFA to the physically adjacent BSs. If the MH is to move to either the left or right cell, both BS 1 & 2 already have the recent few multicast packets for the MH in their buffers. No forwarding of buffers from the old to the new BSs is needed. No join or graft latency is involved as the new BS is already on the multicast tree prior to the handoff. No long handoff latency due to the location update message traversing the wide-area network to the HA is needed, as in the case of Harrison et al. [11]. Handoff is fast and multicast packet loss due to handoff is minimized. As a result, disruption to the multicast session when the MH is a multicast receiver is minimized when the MH crosses cell boundaries.

The problem of reducing packet loss due to handoff when the MH is a multicast source is more tricky. During the period when the MH is outside the coverage area of the old cell to the time when MH initiates a handoff, all multicast packets transmitted by the MH are lost. As the manner of coping with data loss is highly dependent on the needs of the multicast applications [10], we believe that it will be enough if the underlying network can just provide relevant information[1] to the MH immediately after a handoff, and leave it to the multicast application running on the MH to decide on its course of action. Our scheme provides the IP ID of the last packet from the MH received by the old BS prior to the handoff to the MH immediately after a handoff.

Our MobiCast scheme has the following advantages. First, the use of hierarchical mobility management architecture separates the mobility of the MH from the main multicast delivery tree. As long as the MH remains within the domain of the DFA, the mobility of the MH is shielded from the rest of the multicast group. No re-computation of the main multicast delivery tree is needed. The re-computation of the multicast delivery tree from the DFA to the vicinity of the MH when the MH crosses cell boundaries is trivial, as this multicast delivery tree spans across networks within the domain only.

Second, our scheme requires MHs which are interested in receiving multicast data to re-subscribe again via the DFA when they are in a foreign domain. This approach is somewhat similar to the remote subscription method proposed by Mobile IP, and network routes taken by the multicast pack-

---

[1] e.g. IP ID of the last packet received by the old BS prior to the handoff

ets to the MHs are efficient. The inefficiencies and scalability problems associated with bi-directional tunneling in Mobile IP are avoided totally, unlike the case of Harrison et al. [11], whereby complex solution such as DMSP are needed to solve the tunnel convergence problem.

Third, our scheme uses multicast to forward the multicast packets from the DFA to the interested MHs within its domain. Each multicast group is associated with a translated multicast group address, and serving BSs of interested MHs only need to subscribe to this translated multicast group to receive the desired multicast data. Besides delivering multicast data in an efficient manner (as compared to multiple unicasts to interested MHs within its domain), the use of multicast as the forwarding mechanism from the DFA to interested MHs in its domain also alleviates the DFA from the task of keeping track of the exact location of the MH to ensure correct multicast data delivery. It also allows the multicast forwarding algorithm used within the foreign domain to be different from that of the main multicast delivery tree. Furthermore, the use of multicast also enables fast handoff performance to be achieved at a relatively low cost.

Fourth, our scheme requires base stations in the same DVM as the serving BS to subscribe to the same translated multicast group so as to provide fast handoff and minimize disruption to the multicast session when the MH crosses cell boundaries. Since physically adjacent cells are most likely to reside on the same network segment, the extra network load generated due to the other member BSs in the same DVM subscribing to the same multicast group is negligible. This is especially so for the case of shared-medium networks such as Ethernet.

We believe that our hierarchical mobility management architecture coupled with the use of DVM, is a scalable and efficient solution to support multicast for mobile hosts in an environment with small wireless cells.

4. Simulation

We simulated our mobile multicast scheme on the Network Simulator tool (ns2) version 2.1b3 [20]. All code is written and simulated on a Sun Sparc 5 workstation running on Solaris 2.5.1 software platform. We highlight the important components of our simulation in the next few sections.

4.1. Dynamic Virtual Macro-cells

In our simulation, we organized the base stations logically into Dynamic Virtual Macro-cells (DVMs). DVMs are formed by clusters of base stations adjacent to each other, and these DVMs overlap each other as shown in Figure 4. A BS may belong to more than one DVM, but each BS is a core of only one DVM. Only the core can transmit information while the other member BS in the same DVM should only listen. Two BSs in the same DVM need not necessarily be in the same subnet. As an illustration, in Figure 4, BS2 is the core of DVM A and also a member of DVM B. Similarly BS 3 is the core of DVM B as well as a member of DVM A. A handoff can only happen between the core and any of its member BSs in the same DVM. For example, a MH can only handoff from BS2 to BS1 or 3.

In Figure 4, assuming that a MH is in BS2's cell and subscribes to multicast group X. The translated multicast group address Y that corresponds to group X is given to BS2 by the DFA. BS2 will inform all other member BSs in DVM A to subscribe to multicast group Y as well. Since the member BSs of a DVM are likely to reside on the same network segment, we have chosen to use multicast as the communication mechanism within the DVM as well. Each DVM core sends control information to the other member BSs using multicast and each DVM has its own multicast address. This reduces significantly the network load due to the control information flow between the members of the DVM when a MH crosses cell boundaries. For example, the control message from BS2 to other member BSs (BS1 and 3) in its DVM is sent out as a multicast packet using multicast address M, and a control message from BS3 to its member BSs (BS2 and 4) is sent using multicast address N.

In the simplest form of DVM, all physically adjacent cells reside in the same DVM. However, with the knowledge of the actual building layout, we can plan our DVM membership to make our scheme more efficient. If we know that two physically adjacent cells are separated by a wall and there is no way a mobile node can move between the two cells, then these two cells should not be in the same DVM even though they are physically adjacent to each other. A smarter version of DVM is one where BSs do not need manual configuration and can learn about their neighboring BSs over time. Trade off is the non-optimal performance during start up or the transient period when the BSs are not aware of newly added neighboring BSs.

A more complex form of a DVM is one where the shape of the DVM can change based on the movement of the mobile node. The membership in the DVM is dynamic depending on the movement of the mobile node. When the mobile node is static, the DVM has the shape of a circle. When the mobile node is moving, the shape of the DVM may change to an oval with the longer end in the direction of movement so as to include BSs which are two cells ahead.

4.2. Multicast Group Subscription

When a MH is in a foreign network and registered with the DFA as its foreign agent, any subscription request for a multicast group will go to the DFA via the serving BS. Assuming that the MH wants to subscribe to multicast group X, the following message exchanges take place as shown in Figure 5.

- The MH sends an *IGMP report* message containing the group X multicast address.
- The BS receives the *IGMP report* from the MH. The BS sends a *subscription request* message to the DFA. This message contains the group X multicast address that the MH is interested in joining.
- The DFA processes the *subscription request* message and replies to the BS with a *subscription reply* message containing the translated multicast address Y corresponding to multicast group X. If the DFA has already subscribed to group X (i.e. other MHs in the domain have already subscribed to multicast group X), the translated address Y is taken from the translation table. If group X is not on the DFA's existing subscription list, then address Y is a new multicast address created by the DFA corresponding to multicast group X, and the translation table is updated. Multicast address Y has to be unique within the DFA's domain.
- If group X is not on the DFA's existing subscription list, the DFA subscribes to group X.
- Upon receiving the *subscription reply* message, BS subscribes to multicast group Y. It informs the other BSs in its DVM to subscribe to group Y as well, so as to achieve fast handoff performance.

4.3. Multicast Data Delivery

After the MH has subscribed to group X via the DFA, multicast data delivery is achieved in the following manner. When the DFA receives a multicast group X packet from its upstream multicast router, it replaces the IP destination address of the multicast packet with the translated multicast address Y, and forwards it to its domain as a multicast packet. The serving BS receives the packet since it has already subscribed to group Y. It replaces the destination address of the packet back to X, and forwards it to its wireless cell. The other member BSs in the same DVM as the serving BS receives the packet as well since they also subscribed to group Y. They buffer the multicast packets in a First-in-First-out (FIFO) manner.

Note that the MH uses the normal IGMP membership report procedure to subscribe to the multicast group, and the multicast address translation is shielded from the MH. This is to maintain compatibility with the existing IGMP standard and to avoid code changes at the MH, at the expense of extra processing (changing destination address of multicast packets) at the BS.

4.4. Handoff when MH is a Receiver

Our handoff protocol is designed to be fast and efficient. Referring to the scenario shown in Figure 6, the MH is in BS2's cell and BS 1 and 3 are buffering the recent packets of the multicast group that the MH has subscribed to. Assuming that the MH moves to the right, hears the beacon from BS3, and decides to do a handoff to BS3. The following are the message exchanges, as shown in Figure 7.

- After hearing the *beacon* from BS3, the MH decides to switch from BS2 to BS3 based on factors such as radio signal strength and quality of connection.
- MH sends a *greet* message to BS3, indicating its intention to switch over to BS3's cell. The *greet* message contains the IP address of the old serving BS (BS2), the IP multicast group address that the MH has subscribed to, and the IP ID of the last forwarded multicast packet received by the MH from the old BS.
- BS3 sends back a *greet ack* message to the MH to confirm the handoff. BS3 also sends a *notify* message to BS2 to inform the latter about the handoff. At the same time, BS3 starts forwarding multicast packets to the MH, beginning with packets in its buffer received immediately after the packet whose IP ID is indicated in the *greet* message received from the MH. The use of IP ID is to minimize missing or duplicate packets being forwarded to the MH after the handoff. The protocol critical to a correct handoff can be considered as finished after the exchange of the above messages.
- BS3 multicasts a control message to inform the other member BSs (BS2 & 4) in its DVM to subscribe to the translated multicast group associated with the newly arrived MH. After receiving this message, BS4 subscribes to the translated multicast group corresponding to the multicast group that MH has subscribed to and starts buffering the recent packets meant for this multicast group. BS2 does nothing since it is already subscribed to the multicast group.
- After receiving the *notify* message from BS3, BS2 sends back a *notify ack* message to BS3 and stops forwarding multicast packets to the MH.
- BS 2 multicasts a control message to inform the other member BSs (BS 1 & 3) in its DVM that MH has left its cell, and these member BSs can prune themselves off the translated multicast tree corresponding to the multicast group that MH subscribes to if necessary. In this case, since the serving BS is no longer adjacent to BS1 after the handoff, BS1 can unsubscribe from the translated multicast group corresponding to the multicast group that MH has subscribed to.

Note that our scheme does not require the forwarding of multicast packets in the buffers from the old to the new BS after the handoff, as the new BS's buffer has the same content as that of the old BS. The control information packet sent out by BS2 and 3 (as multicast) due to the handoff contains minimum information and has a short packet size. Hence the extra network load generated by these control packets in our scheme is negligible as compared to other handoff schemes which forward data packets from the old BS to the new BS.

The flow of the multicast streams from the DFA to the various BSs after the handoff of MH from BS2 to 3 is shown in Figure 8. Note that after the handoff, BS1 is no longer on the translated multicast tree corresponds to the multicast group that MH subscribes to, and instead BS4 is subscribed to that multicast group.

4.5. Handoff when MH is a Sender

Referring to the same scenario shown in Figure 6, the MH is in BS2's cell and now is a multicast sender to a multicast group X. Assuming that the MH moves to the right, hears the beacon from BS3, and decides to do a handoff to BS3. The following message exchanges take place, as shown in Figure 9.

- After hearing the *beacon* from BS3, the MH decides to switch from BS2 to BS3 based on factors such as radio signal strength and quality of connection.
- MH sends a *greet* message to BS3, indicating its intention to switch over to BS3's cell. The *greet* message contains the IP address of the old serving BS (BS2), the address of the IP multicast group that this MH presently sends packets to (group X).
- BS3 sends back a *greet ack* message to the MH to confirm the handoff. BS3 also sends a *notify* message to BS2 to inform the latter about the handoff. The *notify* message contains a request to BS2 to send back the IP ID of the last received packet from MH for group X. The protocol critical to a correct handoff is considered finished after the exchange of the above messages.
- BS2 sends back a *notify ack* message to BS3 containing the IP ID of the last received packet from MH for group X.
- BS3 relays this information to the MH. The multicast application at the MH can make use of this information to minimize the effect of handoff to the ongoing multicast transmission.

For a MH which is both a multicast receiver and sender, the *greet* message sent by the MH during handoff will contain information on the IP multicast groups that the MH has subscribed to, as well as the IP multicast groups that the MH presently sends packets to. Upon receiving this *greet* message, the new BS reacts accordingly as described in the earlier sections.

4.6. Beacon Period and Buffer Size

The beacon period and buffer size of the base stations are important parameters that determine the handoff performance of our scheme. Every BS transmits a beacon periodically. Besides serving as an agent advertisement message, this signal is important to a mobile host as an aid to detect its own movement. A beacon signal from a base station that gets weaker over time indicates that the mobile host is moving slowly away from that BS's coverage, and a beacon from a BS that gets stronger over time shows that the mobile host is moving nearer to that BS. A mobile host can listen to transmissions from BSs to other mobile hosts in the same region to identify its own location. If such transmissions are absent, the mobile host can conclude that it has roamed out of the coverage area of its serving BS either by detecting a missing beacon from its serving BS, or hearing a beacon from another BS. In our simulation, a mobile host sends out a BS solicitation message once it detects a missing beacon. BSs who heard this solicitation message must send out a beacon, and if the mobile host hears multiple beacons, it can decide which BS to handoff to by looking at criteria such as received signal strength of the beacon messages.

Consider the general case of a small wireless cell environment where cells are overlapping with no coverage gap, and the BSs involved in the handoff are of the same network hierarchy (meaning that a multicast packet from the DFA will arrive at both BSs at about the same time). For the ideal case of eliminating packet loss due to handoff, the amount of buffers needed at the BSs should be equivalent to the maximum possible amount of packet loss due to the handoff. We define the rendezvous time as the time taken for a mobile host to hear a beacon from a new BS after roaming out of the old BS's cell. Hence the rendezvous time determines how soon a mobile host can detect its movement out of a wireless cell and initiate a handoff. In a wireless environment with approximately-synchronous beacon system (all BSs send out beacons approximately at the same time, with a small time offset just enough to prevent collisions of beacons between adjacent base stations), the worst case rendezvous time is equal to the beacon period. Equation (4.1) shows the relationship between the rendezvous time (RT), the packet inter-arrival time, and the maximum possible number of packet loss per connection during a handoff (without any buffering scheme). The amount of buffer needed at the base station to support handoff for multiple multicast sessions can be extrapolated easily.

$$\text{max. num. of pkt loss during a handoff} = [\text{RT} \ / \ (\text{pkt inter-arrival time})] + 1 \quad (4.1)$$

The worst case scenario can be a situation where a multicast packet is transmitted immediately after the beacon from BS X, and the mobile node leaves the wireless cell of BS X before receiving the complete packet. While the mobile host is outside the coverage area of BS X, all multicast packets for the mobile host which arrive at BS X before any handoff is initiated are lost (assuming no buffering is done), and the mobile host can only initiate a handoff after hearing the beacon from a new BS. The amount of packet loss can be reduced if the mobile host can detect its movement out of a wireless cell sooner and initiate a handoff earlier to a nearer BS. Hence a shorter rendezvous time will help to reduce the amount of buffers needed at the BSs to eliminate packet loss during handoff.

For the case of a small wireless cell environment with coverage gap, the maximum possible amount of packet loss is dependent on the mobile host's mobility pattern. Infinite amount of buffers are needed if the mobile host decides to stay put at the coverage gap indefinitely.

In our scheme, each BS buffers the recent multicast packets from multicast groups that the mobile hosts at neighboring cells have subscribed to. The fast handoff is achieved by a BS forwarding these packets from its buffers should the mobile host move into its cell.

A short beacon period consumes more wireless bandwidth, increases processing overhead, but reduces the amount of buffers required at the BSs because the number of expected lost packets is lower. On the other hand, a longer beacon period increases the number of buffers needed at the BSs but consumes less wireless bandwidth and reduces processing overheads. In our simulations, we aim to find combinations of beacon periods and buffer sizes that can achieve smooth handoff for a mobile host engaging in interactive multicast communications like multi-party conferencing during handoff, and which does not require too many resources.

5. Performance Measurement

We have written simulation test scripts to simulate the different mobility scenarios to evaluate the performance of our mobile multicast scheme. Our aim is to determine whether our scheme can reduce packet loss as well as meeting the requirements to support real-time multicast applications like Internet multi-party conferencing during handoff. Our simulations were restricted only to mobility scenarios involving BSs of the same network hierarchy.

5.1. Simulation Scenario

Our simulation scenario is shown in Figure 10. In our ns simulations, node A, node B, the DFA and the BSs are connected by wired links, and the MH is attached to the BSs through wireless links. The wired and wireless networks are simulated using 10 Mbps duplex links and 2 Mbps duplex links respectively. For our multicast packet audio source, we chose pulse code modulation (PCM) as the multicast conferencing audio coding format to simulate the most resource demanding case. Using Mbone applications like vat as the yardstick, the shortest packet inter-arrival time for PCM format is 20 ms, and the average packet size is 200 bytes. We use these parameters in our simulations of a multicast conferencing audio source.

5.2. Handoff Latency

The objective of this simulation is to find out the time needed for our scheme to complete a handoff. The time to complete a handoff has two components: the rendezvous time and the handoff latency. The rendezvous time refers to the time taken for a mobile node to hear a beacon from a new BS after roaming out of the old BS's cell. When the MH is a multicast receiver, the handoff latency is defined as the difference in time between the arrival of the first new packet from the new BS's buffer and the time at which the MH sends a handoff request to the new BS. In our case, this includes the exchange of *greet*, *greet ack* messages and the arrival of the first packet from the new BS at the mobile node.

Our experiment involved node A sending out UDP packets of 200 bytes each at an interval of 20 ms, to simulate the multicast audio source. These multicast packets are received by the DFA and forwarded as another multicast stream with a different multicast address to the MH. We measured the handoff latency when the MH does a handoff from BS1 to 2 while receiving this multicast stream. To gauge the handoff latency performance of our scheme for the different possible wireless networks, we performed our simulations using wireless networks of different bandwidths and link delays. The results are plotted in Figure 11. For a typical wireless network of 2 Mbps bandwidth and 4 ms link delay, the handoff latency for our scheme is 9.2 ms. Hence the total handoff time for our scheme in such a wireless network is equal to the rendezvous time plus 9.2 ms.

Note that in our case, as long as both BSs are approximately at the same network hierarchy, the handoff latency is the same for handoff between BSs located at the same subnet and BSs located at different subnet, since no forwarding of buffers is needed. However, extra network load generated due to the handoff is slightly higher for handoffs between BSs of different subnets.

For all the subsequent simulations, the wireless networks simulated have a bandwidth of 2 Mbps and 4 ms link delay.

5.3. Beacon Period

The goal of this simulation is to find the minimum beacon period to give a short rendezvous time so as to minimize multicast packet loss when the MH crosses cell boundaries, and yet one that will not involve too much processing and affect the overall performance of the system. This simulation involved the sending of 4 Mbytes of data from the DFA to the MH (via BS1), with a receiver advertised TCP window size of 15 Kbytes and 1024 bytes segment. Each beacon is 50 bytes long. We measured the throughput of this transfer for the different beacon periods. The results are plotted in Figure 12.

As shown in Figure 12, the throughput is above 99% of the maximum for beacon periods of 30 ms or longer, and it drops to 98% of the maximum when the beacon period is 13 ms. Hence we conclude that the minimum beacon period we can choose to give a low rendezvous time and involved minimum overhead is 30 ms for a wireless network of 2 Mbps and 4 ms link delay.

5.4. Multicast Packet Audio Performance

The objective of this simulation is to find out whether our scheme can minimize the effects of a handoff while the MH is in an interactive multicast session.

In our experiment to simulate the handoff scenario when MH is a multicast receiver, node A multicasts a stream of packets of 200 bytes each every 20 ms to simulate a real-time Internet multicast audio source. The DFA subscribes to this multicast group on behalf of the MH and re-multicasts the stream to the vicinity of the MH. The MH moves from BS1 to 2 while receiving this multicast stream. Various human factors studies have shown that the maximum tolerable delay for an interactive conversation is approximately 200 ms. This helps to set the maximum tolerable rendezvous time and beacon period the system can allow.

Figure 13 shows the maximum number of lost packets at the MH during a handoff while the MH is a multicast receiver of an interactive multicast session, against the buffer size at the BS for the different rendezvous time. Note that the worst case rendezvous time is equal to the beacon period in a wireless environment with a synchronous beacon system. To satisfy the maximum tolerable delay for an interactive conversation, the rendezvous time has to be below 200 ms. A possible combination that will eliminate multicast packet loss during handoff is to choose a buffer size of 6 packets and a rendezvous time of 100 ms. This rendezvous time can be satisfied by having a system with synchronous beacons of 100 ms beacon period.

Figure 14 shows the packet inter-arrival time at the MH before, during and after a handoff for a wireless system with synchronous beacon period of 100 ms, with and without buffer. The handoff scenario simulated involved the worst rendezvous time of 100 ms. There are several points to highlight regarding these results. First, the jitter introduced by our handoff scheme to support mobile multicast receiver is below the maximum tolerable delay for an interactive multicast conversation, for both with and without buffer. For the case with buffer, the longest packet inter-arrival time is 122 ms, of which 100 ms is contributed by the rendezvous time. Packets with sequence number from 16 to 20 are transmitted back-to-back from the new BS and have short packet inter-arrival times.

Second, for the case with buffer, we can see that no packets are lost, duplicated or have arrived at the MH in the wrong sequence. The provision of enough buffers at the BSs ensures that packets not received by the MH during handoff are stored, and the use of IP ID in our scheme ensures that these stored packets are forwarded to the MH from the new BS in the correct sequence. For the case with no buffer at the BSs, we can see that packet 15 to 20 are missing. This results in a quality degradation to the reception of the multicast stream at the MH when it crosses cell boundaries.

In the experiment to simulate the handoff scenario when MH is a multicast source, MH sends a stream of packets of 200 bytes each every 20 ms to the DFA to simulate a real-time Internet multicast audio source and the DFA sends the packets out on behalf of MH. Node A and B subscribe to this multicast group and receive multicast packets from the DFA. The MH moves from BS1 to 2 while sending this multicast stream to the DFA. The beacon period and the rendezvous time are both 100 ms. For our simulation, we assume that the application's strategy of dealing with packet loss can be either to ignore this 'last received packet' information, or to make use of this information and retransmit those missing packets immediately.

Figure 15 and Figure 16 show the multicast packet inter-arrival time and packet sequence at node B respectively, before, during and after a handoff. The 'last received packet ' information arrives at the MH before the next transmission from MH to the multicast group is due. First, referring to Figure 15, we can see that the jitter introduced by our handoff off scheme is below the maximum tolerable delay for an interactive multicast conversation, when the MH is the multicast source, with or without the use of the 'last received packet' information. Second, if the 'last received packet' information arrives early enough at the MH, the multicast application at MH can react accordingly to minimize the effect of handoff, as shown in Figure 16. In this case, this information is used to eliminate packet loss from the MH due to handoff. An example of such applications that can benefit from our scheme is an interactive multicast conferencing session.

Figure 17 shows the packet arrival at node B before, during and after the handoff of MH, for the case when the 'last received packet' information from the old BS (BS1) reaches the MH after the next transmission to the multicast group is due. If this information arrives at the MH too late, it may be better off for some applications (e.g. interactive conferencing) to simply ignore this 'last received packet' information. However, this late 'last received packet' information may still be useful for applications which are not very time critical and sequence of packet arrival is not important. In this case, the application can backtrack from the latest transmission and retransmit the missing packets, as shown in Figure 17.

From these simulations, we have shown that our MobiCast scheme is effective in minimizing packet loss and disruption to the multicast session when the MH is either a multicast receiver or sender and crosses cell boundaries during an interactive multicast session.

6. Future Work

The present scheme supports best effort multicast for mobile hosts in an internetwork environment with small wireless cells. It will be interesting to extend the architecture to support reliable multicast for mobile hosts in a similar environment. One possible method for the mobile hosts to recover from packet loss is for them to ask for missing multicast packets within the domain they reside in first, before sending the repair request to the other group member outside the domain.

7. Conclusions

In this paper, we have presented MobiCast, our multicast scheme for mobile hosts in an internetwork environment with small wireless cells. Our scheme adopts the hierarchical mobility management architecture coupled to the use of the Dynamic Virtual Macro-cells concept, to handle multicast to and from mobile hosts roaming among small wireless cells.

Our scheme requires the mobile hosts to subscribe or send packet to a multicast group via the domain foreign agent. This avoids the recomputation of the main multicast delivery tree due to mobility of the mobile host within the foreign domain. Fast handoff performance to reduce packet loss due to mobility is achieved by forwarding multicast packets from the domain foreign agent to the base stations at the vicinity of the mobile group members. We have designed and simulated MobiCast in ns2. From simulation results, we have shown that our scheme can maintain the quality of a multicast conferencing session when a mobile group member crosses cell boundaries during the multicast session. When the mobile host is a multicast receiver, the first packet from the new base station arrives within 9.2 ms at the mobile host after a handoff is initiated. When the mobile host is a multicast sender, our scheme provides "last received packet" information to the mobile host immediately after a handoff. We have demonstrated how this information can be used by a multicast application at the mobile host to minimize degradation due to handoffs.

Multicast applications are popular due to their efficient network model of data delivery to multiple members. The multicast scheme for mobile hosts presented in this paper enables mobile users to enjoy the benefits of such multicast applications on portable devices with similar quality of service as their wired counterparts, without the constraint of a tether.

Acknowledgments

The authors would like to thank the anonymous reviewers for their comments on the earlier draft of the paper.

References

[1] A. Acharya, A. Bakre, and B. R. Badrinath. IP multicast extensions for mobile internetworking. In *Proc. of INFOCOM*, San Fransisco, California, Mar. 1996.

[2] A. Ballardie, J. Crowcroft, and P. Francis. Core based tree (CBT) - an architecture for scalable inter-domain multicast routing. In *Proc. of ACM SIGCOMM*, volume 23, pages 85–95, Oct. 1993.

[3] R. Cáceres and V. N. Padmanabhan. Fast and scalable handoffs for wireless internetworks. In *Proc. of ACM/IEEE MobiCom*, pages 56–66, New York, USA, Nov. 1996.

[4] V. Chikarmane, R. B. Bunt, and C. L. Williamsom. Mobile IP-based multicast as a service for mobile hosts.

In *Proc. of the Second International Workshop on Services in Distributed and Networked Environments*, pages 11–18, June 1995.

[5] S. Deering. Host extensions for ip multicasting. RFC, RFC 1112, Aug. 1989.

[6] S. Deering, D. Estrin, D. Farinacci, A. Helmy, D. Thaler, M. Handley, V. Jacobson, C. Liu, P. Sharma, and L. Wei. Protocol Independent Multicast-Sparse Mode (PIM-SM): protocol specification. Request for Comments (Experimental) 2362, Internet Engineering Task Force, June 1998. (Obsoletes RFC2117).

[7] S. Deering, D. Estrin, D. Farinacci, V. Jacobson, A. Helmy, and L. Wei. Protocol independent multicast version 2, dense mode specification. Internet Draft, Internet Engineering Task Force, May 1998. Work in progress.

[8] S. Deering and C. Partridge. Distance vector multicast routing protocol. RFC, RFC 1075, Nov. 1988.

[9] S. E. Deering and D. R. Cheriton. Multicast routing in datagram internetworks and extended LANs. *ACM Trans. Comput. Syst.*, 8(2):85–110, May 1990.

[10] S. Floyd, V. Jacobson, C.-G. Liu, S. McCanne, and L. Zhang. Reliable multicast framework for light-weight sessions and application level framing. In *Proc. of ACM SIGCOMM*, pages 342–356, Cambridge, Massachusetts, Sept. 1995.

[11] T. G. Harrison, C. L. Williamsom, W. L. Mackrell, and R. B. Bunt. Mobile multicast (MoM) protocol: Multicast support for mobile hosts. In *Proc. of ACM/IEEE MobiCom*, Sept. 1997.

[12] J. Ioannidis, D. Duchamp, and G. Q. Maguire, Jr. IP-based protocols for mobile internetworking. In *Proc. of ACM SIGCOMM*, pages 235–245, Zürich, Switzerland, Sept. 1991. ACM. also in *Computer Communication Review* 21 (4), Sep. 1991.

[13] J. Ioannidis and G. Q. Maguire, Jr. The design and implementation of a mobile internetworking architecture. In *Proc. of USENIXW*, pages 491–502, San Diego, California, Jan. 1993. Usenix.

[14] V. Jacobson and S. McCanne. The LBL audio tool vat. Manual page, July 1992.

[15] D. Johnson and C. Perkins. Mobility support in IPv6. Internet Draft, Internet Engineering Task Force, Mar. 1998. Work in progress.

[16] S. McCanne and V. Jacobson. vic: A flexible framework for packet video. In *Proc. of ACM Multimedia '95*, Nov. 1995.

[17] J. Moy. Multicast extensions to OSPF. Request for Comments (Proposed Standard) 1584, Internet Engineering Task Force, Mar. 1994.

[18] C. Perkins. IP mobility support version 2. Internet Draft, Internet Engineering Task Force, Nov. 1997. Work in progress.

[19] C. L. Tan, S. Pink, and K. M. Lye. A fast handoff scheme for wireless networks. In *Proc. of the 2nd ACM International Workshop on Wireless Mobile Multimedia (WoWMoM'99)*, pages 83–90, Aug. 1999.

[20] UCB/LBNL/VINT Network Simulator NS. http://www-mash.cs.berkeley.edu/ns/.

[21] G. Xylomenos and G. C. Polyzos. IP multicasting for wireless mobile hosts. In *Proc. of the IEEE MILCOM Conference on Military Communications*, volume 3, pages 933–937, Oct. 1996.

[22] G. Xylomenos and G. C. Polyzos. IP multicast for mobile hosts. *IEEE Communications Magazine*, 35(1):54–58, Jan. 1997.

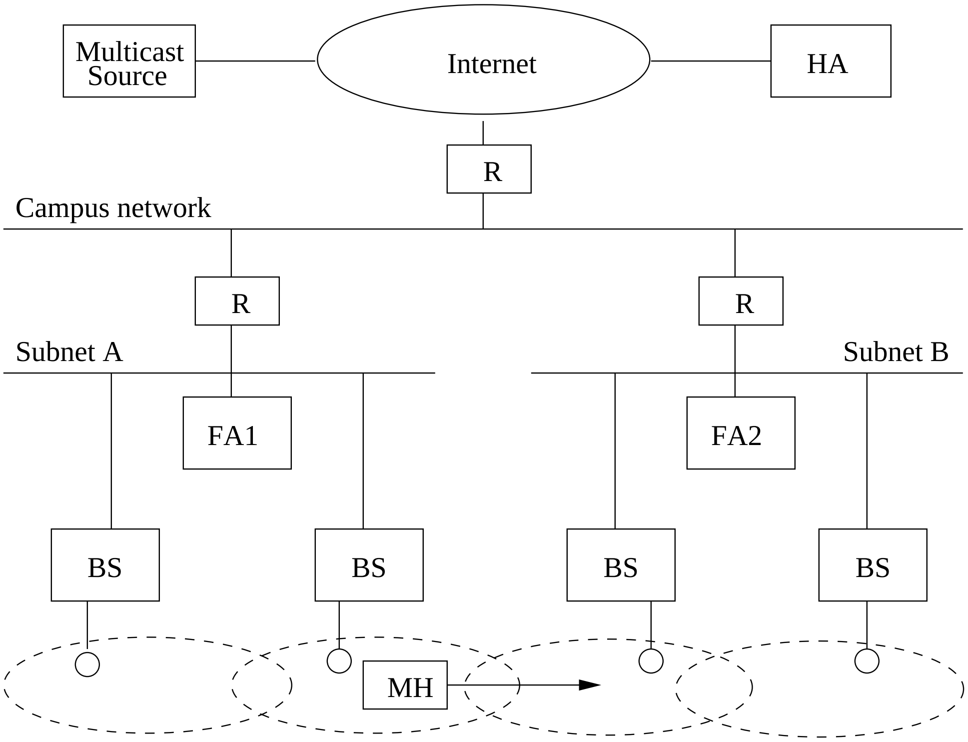

Figure 1. *Mobile IP entities in a wireless internetwork.*

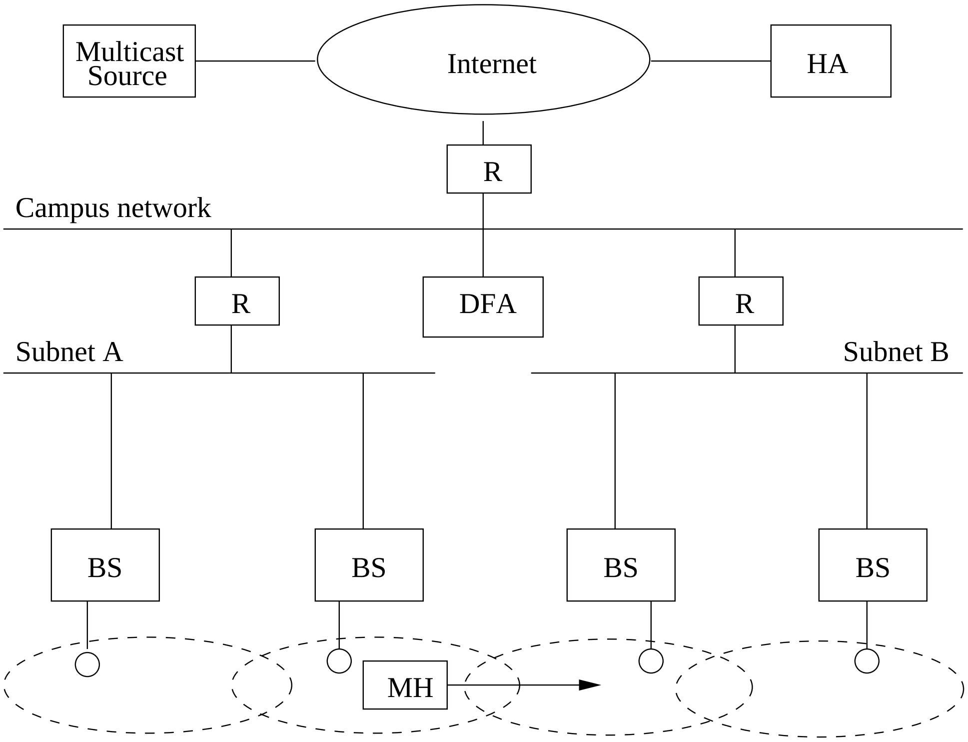

Figure 2. *Hierarchical mobility management approach using Domain Foreign Agent.*

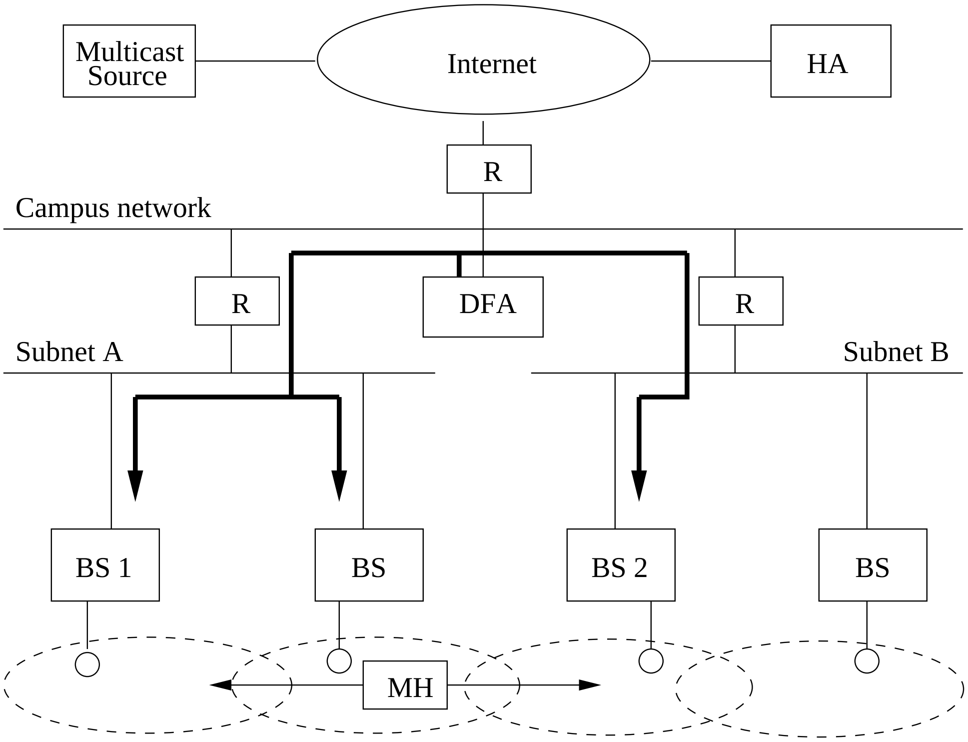

Figure 3. *Use of multicast as the multicast packet forwarding mechanism from the DFA to base stations in the same DVM*

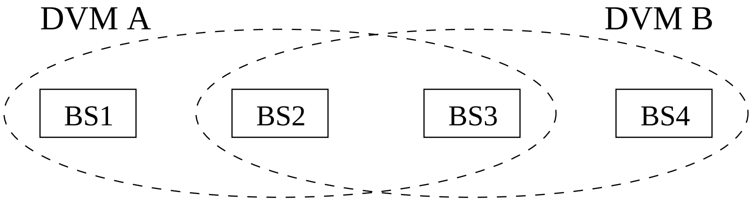

Figure 4. *Dynamic Virtual Macro-cell.*

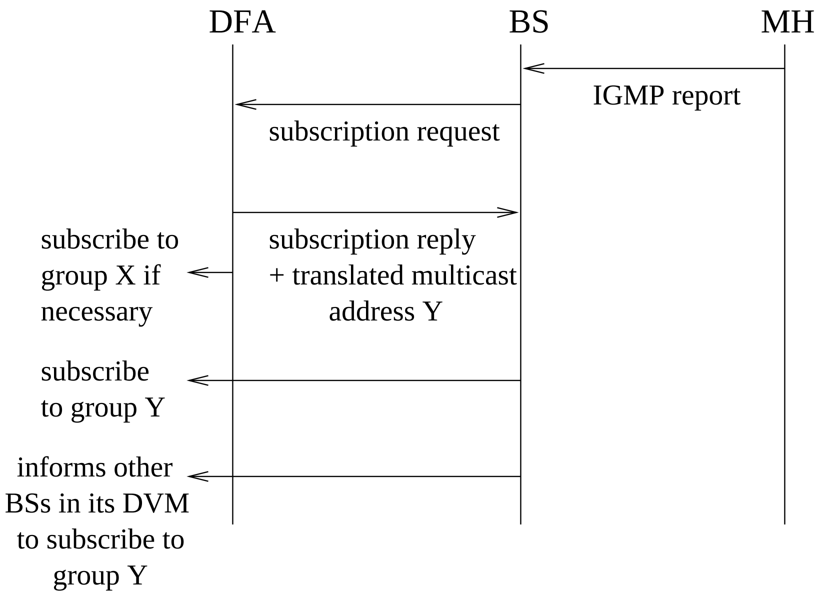

Figure 5. *Message exchange when a MH subscribes to a multicast group.*

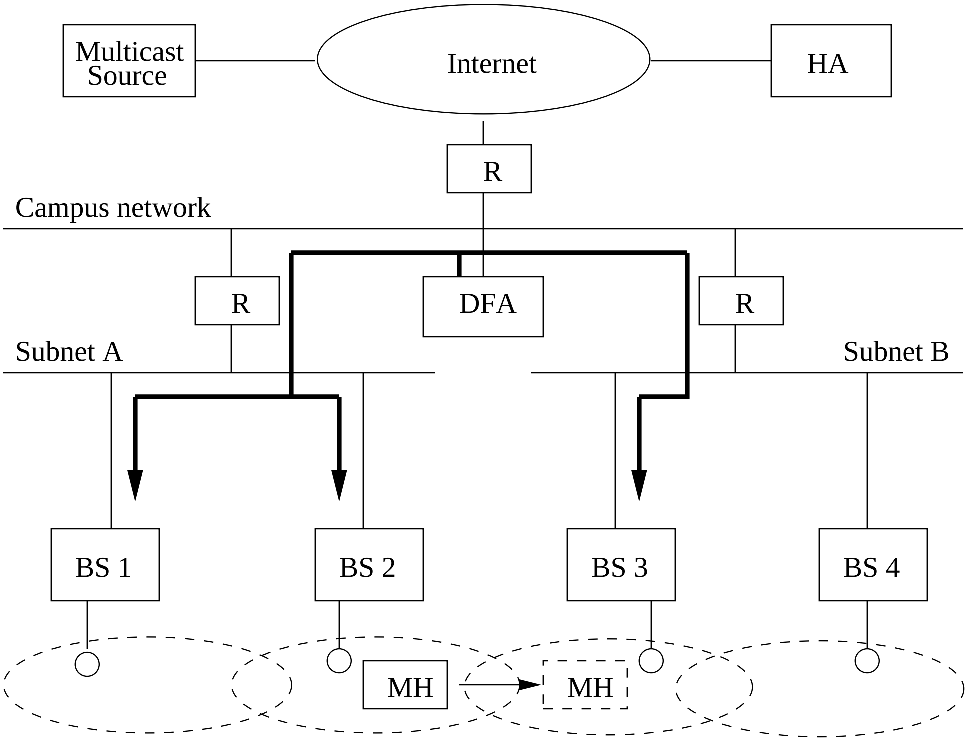

Figure 6. *Before the handoff of MH from BS 2 to 3*

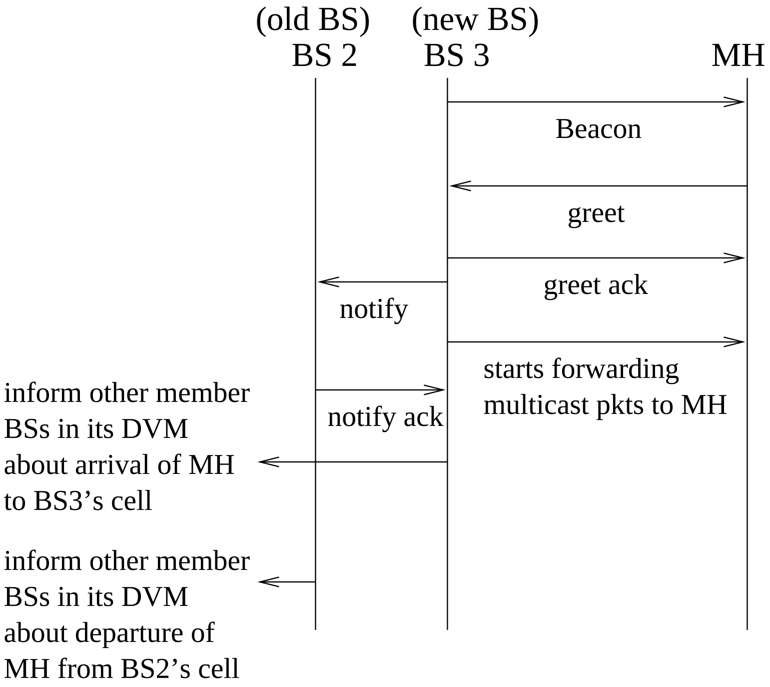

Figure 7. *Message exchange during a handoff when MH is a multicast receiver.*

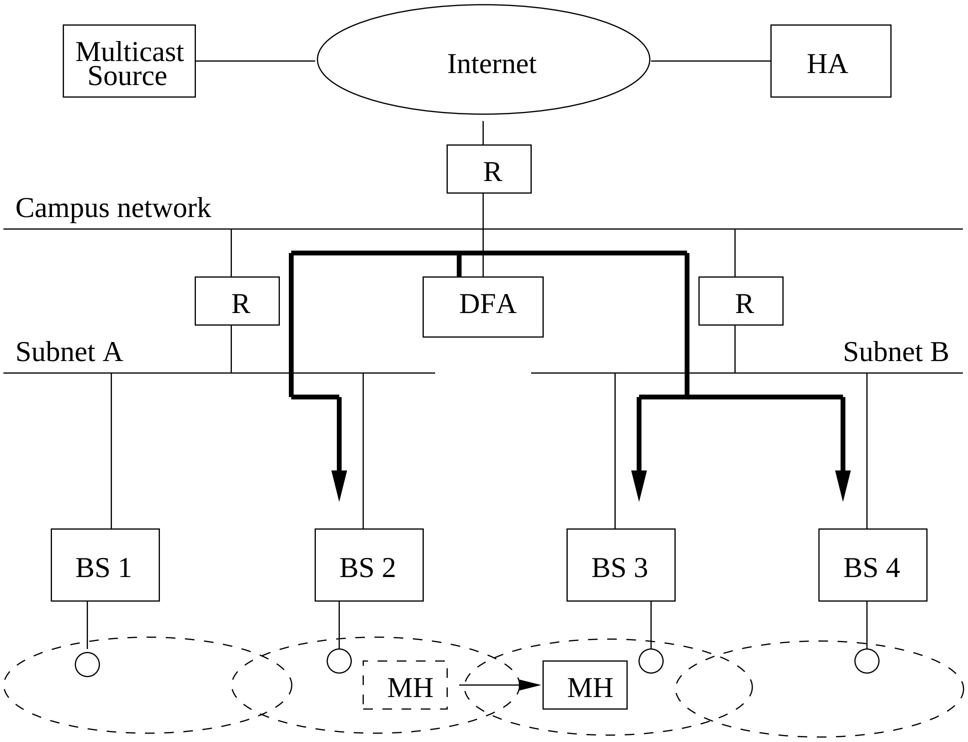

Figure 8. *After the handoff of MH from BS 2 to 3 when MH is a multicast receiver.*

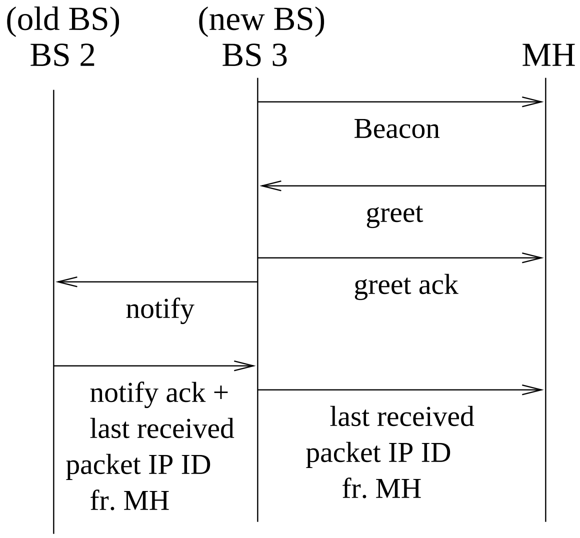

Figure 9. *Message exchange during a handoff when MH is a multicast sender.*

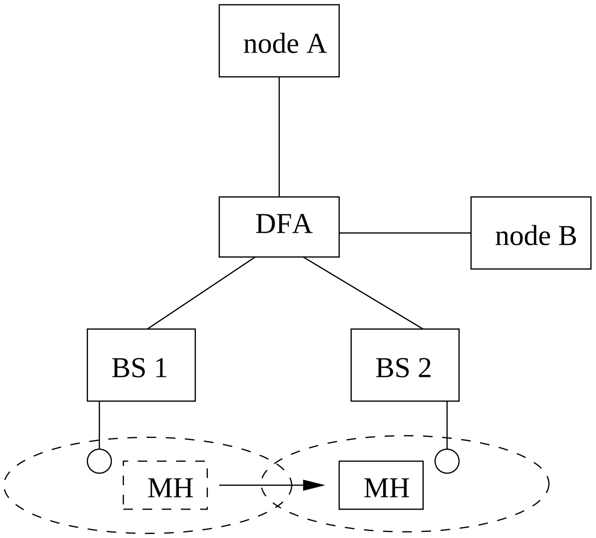

Figure 10. *Simulation Scenario.*

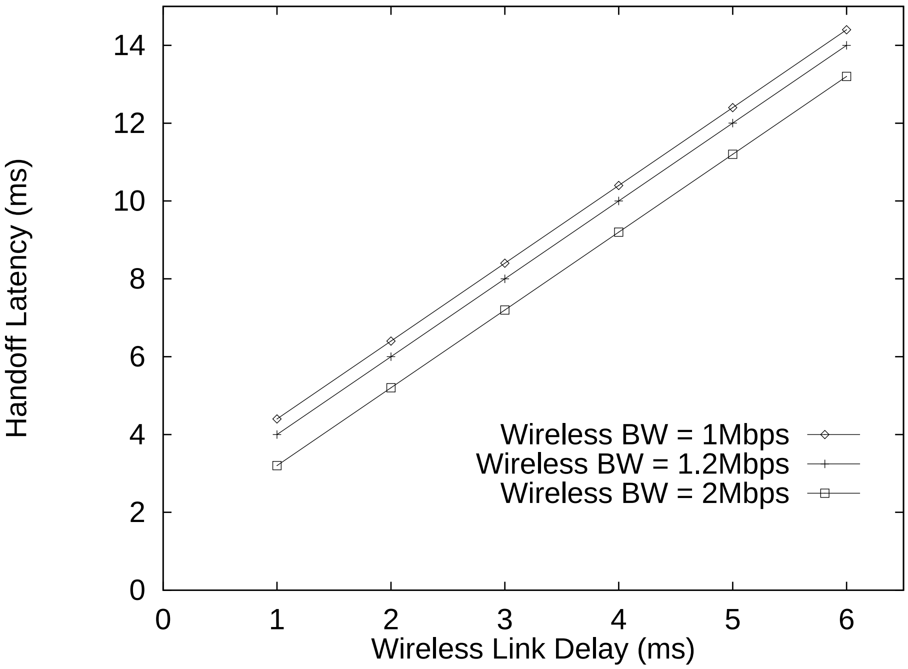

Figure 11. *This graph shows the handoff latency of our scheme for wireless networks of different bandwidth and link delay parameters. (BW = bandwidth)*

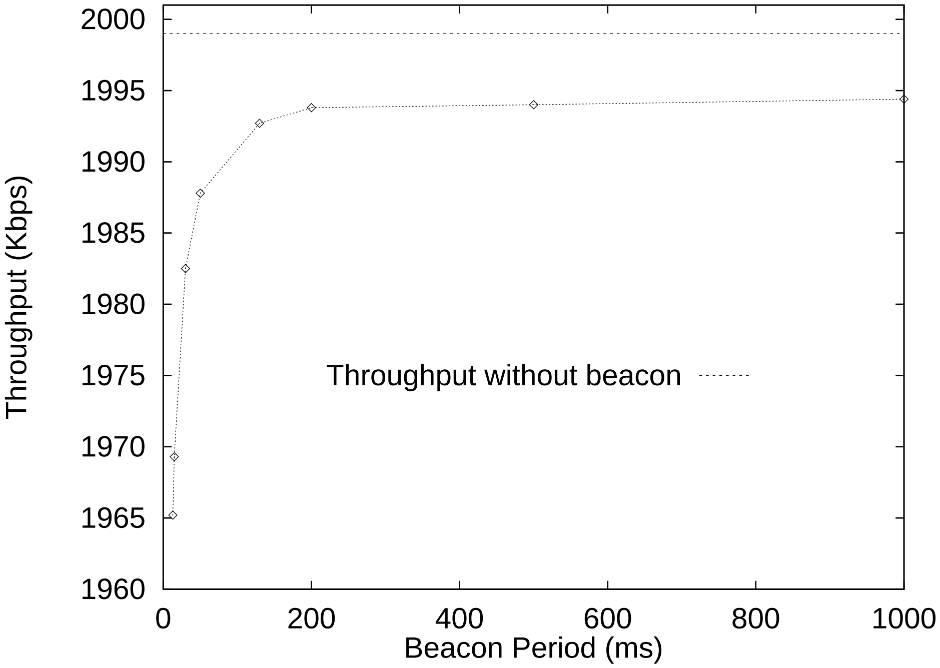

Figure 12. *This graph shows the effect of the beacon period on the throughput of a FTP transfer over a wireless link of 2 Mbps bandwidth and link delay of 4 ms.*

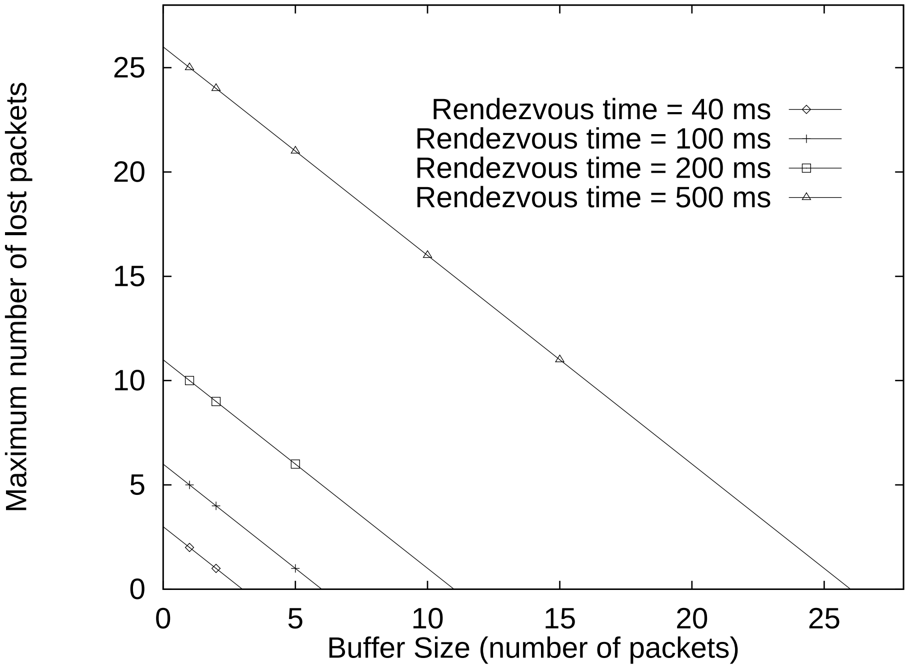

Figure 13. *This graph shows the maximum number of lost packets due to a handoff if the BS does not have the required number of buffers for the different rendezvous time, when MH is a multicast receiver.*

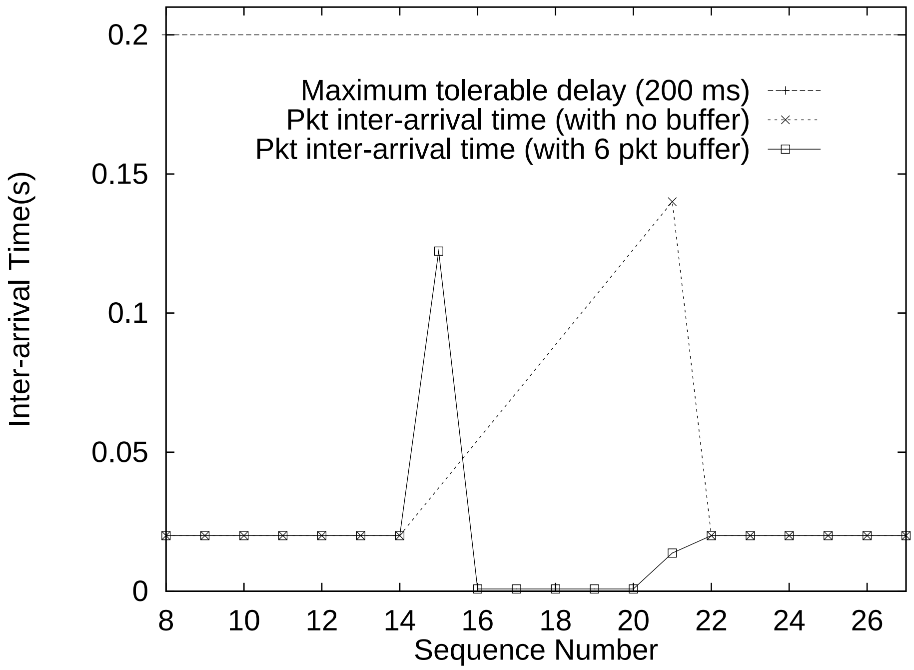

Figure 14. *This graph shows the packet inter-arrival time at MH before, during and after a handoff when MH is a multicast receiver, with and without buffer.*

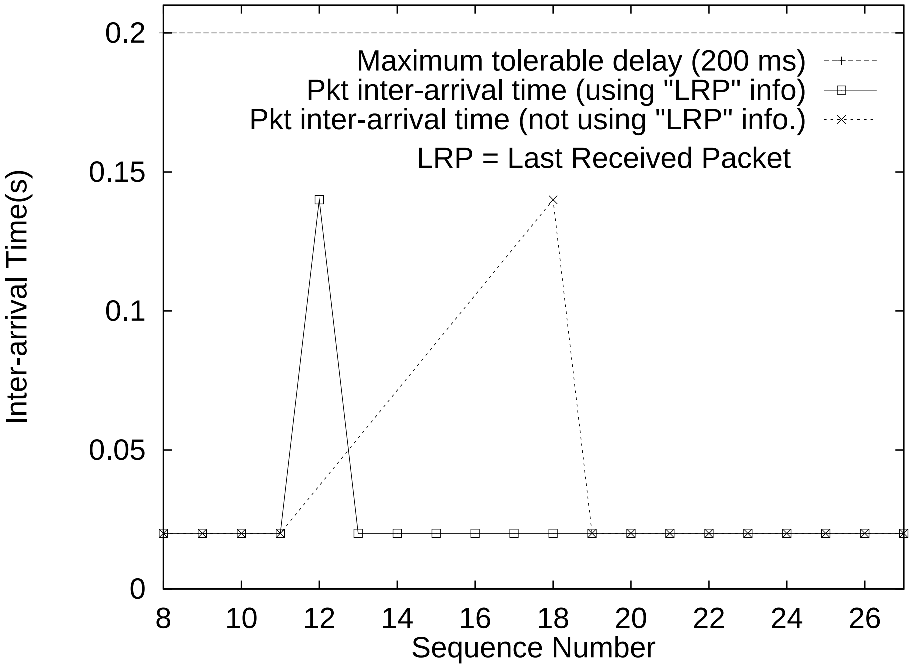

Figure 15. *This graph shows the packet inter-arrival time at a multicast receiver (node B) before, during and after a handoff when MH is a multicast sender, and the 'last received packet' information from the old BS arrives at the MH before the next transmission from MH to the multicast group is due.*

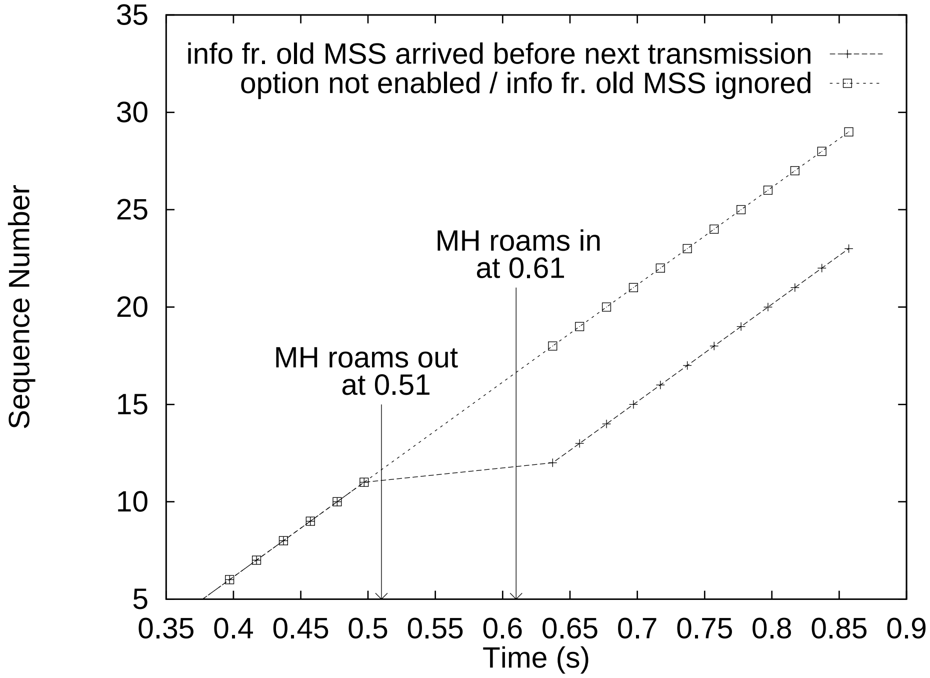

Figure 16. *This graphs shows the packet arrival at a multicast group receiver (node B) when MH is a multicast source. The last received packet information from the old BS arrives at the MH before the next transmission from MH to the multicast group is due.*

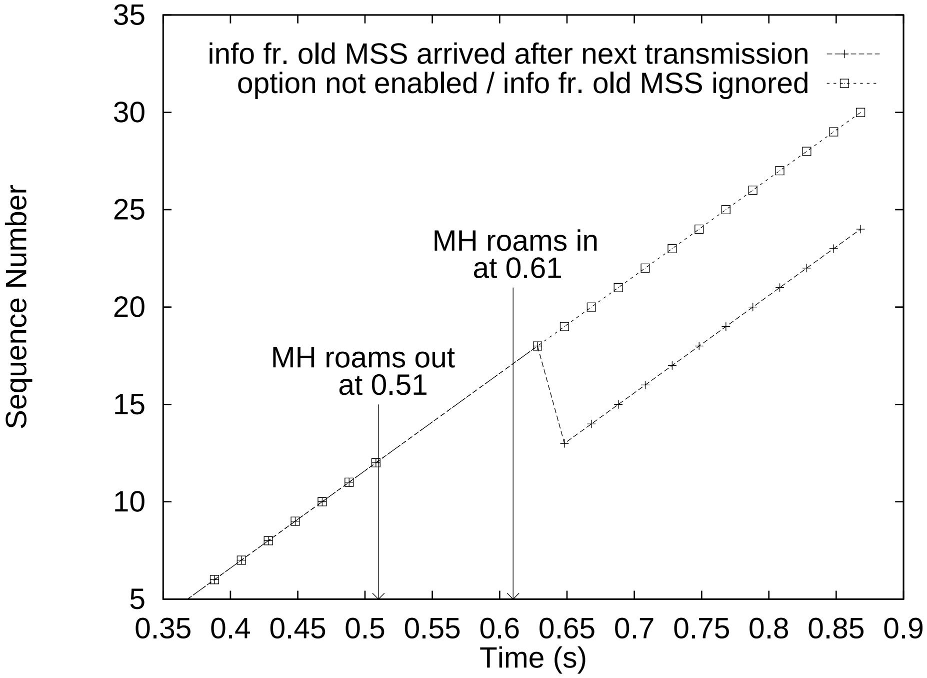

Figure 17. *This graphs shows the packet arrival at a multicast group receiver (node B) when MH is a multicast source. The last received packet information from the old BS arrives at the MH after the next transmission from MH to the multicast group is due.*