Group Management for the Multimedia Broadcast/Multicast Service

George Xylomenos
Mobile Multimedia Laboratory
Department of Informatics
Athens University of Economics and Business
Patision 76, Athens 104 34, Greece
Email: xgeorge@aueb.gr

***Abstract*— This paper compares the group management mechanisms used in the IP and the MBMS multicasting models. After outlining the design of each model, we describe the group management protocols that they employ. We then examine how the IP group management protocols can be adapted for MBMS and finally evaluate the group management approach adopted by MBMS. Our main findings are that IGMP v.2 is preferable for use with MBMS, that the join/leave group management approach of MBMS outperforms the query/report approach of IP and that the reliability of the MBMS approach can be enhanced by upcalls.**

I. Introduction

As the bandwidth of cellular systems is enhanced, they are becoming feasible platforms for the delivery of multimedia services, but the amount of bandwidth that such services require makes them too expensive for most users. A dramatic cost reduction is however possible when many users receive the same service simultaneously. By transmitting data only once in each cell, all interested users can share the cost of the radio resources consumed. This can be achieved either by *broadcast*, where all users receive the service, or by *multicast*, where only a selected set of users receives the service. Multicast is especially appropriate for targeting users that have explicitly subscribed to (and paid for) a service.

Support for multicast has long been added to the Internet, in the form of IP multicasting [1]. In the *Universal Mobile Telecommunications Systems* (UMTS) specified by the *3rd Generation Partnership Project* (3GPP), support for IP multicasting was first introduced in the Release 99 specifications. This service used unicast tunnels to transmit IP multicast packets to each receiver, therefore it did not offer any resource savings. In contrast, the *Multimedia Broadcast/Multicast Service* (MBMS), first introduced in the Release 06 specifications, allows resource sharing throughout the UMTS network, an especially critical issue over the air interface [2]. Applications envisaged for MBMS include multimedia streaming and file downloads in multicast or broadcast mode [3].

We will only consider the multicast mode of MBMS, which we will simply call *MBMS multicasting*, as it is expected to be far more important commercially than the broadcast one. Even though the MBMS specifications are still evolving, the intention is for MBMS multicasting to be compatible to some extent with IP multicasting, upon which it is loosely based. Since MBMS multicasting exclusively targets UMTS networks however, it diverges from IP multicasting in some areas. In this paper we focus on the group management aspects of IP and MBMS multicasting. Our first goal is to examine how the IP multicasting group management mechanisms can be adapted for MBMS multicasting. Our second goal is to assess whether the design choices made by the MBMS multicasting model are appropriate for UMTS networks. The results of this evaluation will be used in the MBMS simulator that is currently under development by the IST B-Bone project.

In Section II we briefly describe the IP multicasting model and contrast it with the MBMS multicasting model. In Section III we describe the group management protocols used for IP and MBMS multicasting. In Section IV we discuss how the group management protocols of IP multicasting can be adapted for MBMS multicasting. In Section V we evaluate the MBMS group management approach in terms of performance and reliability. We summarize our findings in Section VI.

II. Multicast models

A. The IP multicasting model

In the IP multicasting model each multicast group is identified by a class D IP address. Any host can *join* the multicast group in order to receive packets sent to it and later *leave* the group to stop receiving such packets. Any host can send packets to the group by using its IP address as the packet destination. In the original IP multicasting model the groups are open in both directions, that is, anybody can receive data transmitted to a group by becoming a member, and anybody can send data to a group, even without being a member [1].

The use of open groups means that, on the receiver side, a commercial content provider cannot ensure that its content is only received by paying subscribers without additional mechanisms, while on the sender side, groups are vulnerable to denial of service attacks or, at least, annoying senders. For this reason the IP multicasting model has been extended with *source filtering*, that is, the ability for group members to specify which sources they want to receive data from.

The mechanisms implementing the IP multicasting model are split into local and global ones. The *local* mechanisms track group membership and deliver multicasts within a local network, while the *global* mechanisms route multicast packets between local networks. Regarding local mechanisms, the multicast router of each network is responsible for discovering the groups with local members, as well as for forwarding multicasts originating from the local network to external networks and vice versa. Regarding global mechanisms, multicast routers execute a distributed routing protocol to deliver multicast packets originating from any network to those multicast routers serving local members for the group addressed.

The only local mechanism that has been standardized for use over the Internet was designed for local area networks, supporting native multicasting and broadcasting, such as Ethernets. In these networks, multicast packet transmission and reception are trivial, but group management still requires an appropriate protocol. This is the *Internet Group Management Protocol* (IGMP) that we will examine in Section III. In contrast, many global mechanisms have been deployed on the Internet, but they are beyond the scope of this paper.

B. The MBMS multicasting model

The functional entities of a UMTS network relevant to the MBMS multicasting model are shown in Figure 1. The *Broadcast/Multicast Service Centre* (BM-SC) is an entity used only by MBMS, while the *Gateway GPRS Support Node* (GGSN), the *Serving GPRS Support Node* (SGSN), the *Radio Access Network* (RAN) and the *User Equipment* (UE) are entities that must be modified in order to support MBMS. We also show MBMS content sources, both internal and external to the UMTS network, which are not specified by the 3GPP.

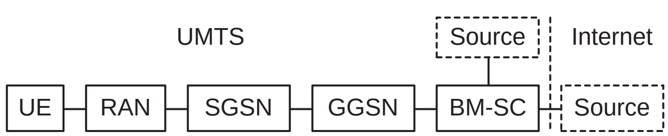

Fig. 1. Components of MBMS.

In the MBMS multicasting model each group is identified by a class D IP address which is used as the destination address for packets sent to the group. Each group is also identified by an *Access Point Name* (APN), which effectively identifies the GGSN serving a specific UMTS network [4]. As a result, MBMS multicasting groups are defined with respect to a particular UMTS network. It is also possible for the group members to receive different content depending on the cell they are currently residing in [2], thus allowing location based services to be offered.

The major deviation from the IP multicasting model is that MBMS multicasting groups are closed in both directions. On the sender side, only the GGSN identified by the APN may send data to the group. These data, whether originating inside or outside the UMTS network, are first processed by the BM-SC and then delivered by the GGSN to each UE in the group. On the receiver side, a UE must first subscribe to the group in order to be allowed to join it. The subscription mechanism is beyond the scope of the 3GPP specifications, but it must enable the BM-SC to verify whether a UE attempting to join a group is subscribed to it. Combined with the ability to charge the group members [2], the MBMS multicasting model offers an attractive business model to commercial content providers.

The mechanisms implementing the MBMS multicasting model roughly correspond to those of IP multicasting. Regarding global mechanisms, the GGSN may act as a multicast router in order to receive packets from an IP multicasting group and forward them to an MBMS multicasting group. However, the content forwarded to an MBMS multicasting group may also reach the GGSN by other means. Regarding local mechanisms, a variant of IGMP is used between the UE and GGSN for group management purposes, complemented by considerable additional MBMS specific signalling that we will discuss in Section III. This additional complexity has three main causes. First, before a UE may join a group its subscription to the group must be verified. Second, the UE and the GGSN are not directly connected for either signalling or data transport. Third, while the GGSN is able to send multicast packets to all UEs in a cell, a UE can only send unicast packets to the GGSN, due to the nature of the UMTS air interface.

III. Group management protocols

A. IGMP version 2

The goal of IGMP is to determine which IP multicasting groups have local members, so that the multicast router can use the global multicasting mechanisms to receive data addressed to these groups. IGMP assumes that the local network supports native multicasting and broadcasting. In this type of network it does not matter how many members exist for a group or who they are; as long as at least one member exists, data destined to the group should be locally delivered by native multicast.

IGMP versions 1 and 2 are used with the original IP multicasting model, where any host is allowed to send to a group. IGMP v.1 has long become obsolete [1], but IGMP v.2 is in active use for IPv4 [5]. A nearly identical protocol called *Multicast Listener Discovery* (MLD) version 1 [6] has been designed for IPv6. Apart from the different IP address sizes, IGMP v.2 and MLD v.1 are practically the same, so we will only discuss IGMP v.2 below.

<table>
<tr><td>Type</td><td>Response</td><td>Checksum</td></tr>
<tr><td colspan="3">Multicast Address</td></tr>
</table>

Fig. 2. IGMP version 2 messages.

All IGMP v.2 messages have the format shown in Figure 2. They are encapsulated in IP packets that are never forwarded outside a local network. The multicast router periodically sends *general query* messages to the all multicast enabled hosts group, leaving the multicast address field blank. The response field contains a *query interval* expressed in 10 msec units. On receiving a general query, each multicast receiver schedules a separate *report* message to be sent for each group that it is a member of, after a random interval that is less than the query interval. Each report includes the multicast address reported and is sent to the all multicast enabled hosts group.

The first report sent for a group suppresses other reports for the same group. Thus only a single report is sent per group after each general query. A receiver is also allowed to send unsolicited reports when joining a group. Since general queries are periodically repeated, the multicast router continuously refreshes its group membership list. If no reports are received for a group after a few consecutive queries, the group is dropped from this list. Therefore, data for a group may continue to be forwarded to the local network for a considerable time after the last local member has left it.

To avoid this waste of resources, when a receiver leaves a group for which it was the last one to report membership, it sends a *leave* message to the all multicast routers group, specifying the group that it left. As the last report sent may have suppressed others, the multicast router sends a *group specific query* to this group with a shorter query interval. If no receivers report membership to this group after a few consecutive group specific queries, the group is dropped.

B. IGMP version 3

IGMP version 3 diverges from IGMP v.2 in its support for source filtering, that is, the ability of receivers to specify which sources they desire to receive data from [7]. As each receiver may have different source filtering preferences, it is quite complex for the multicast router to decide which sources to forward for each group. In this paper we are only concerned with group management issues relevant to MBMS, where only a single source exists for each group, therefore we will not deal with the source filtering rules of IGMP v.3. As for IGMP v.2, IGMP v.3 has a nearly identical IPv6 counterpart, MLD version 2 [8], that we will also not discuss any further.

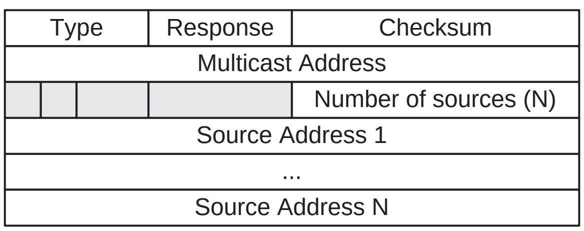

Fig. 3. IGMP version 3 query message.

The format of the IGMP v.3 query messages is shown in Figure 3, with fields that are irrelevant for our discussion grayed out. Both general and group specific queries exist, differentiated by whether the multicast address field is left blank or not. The interpretation of the response field depends on its value: values less than 127 are interpreted as 10 msec units, similar to IGMP v.2, while higher values are interpreted as numbers in an exponential notation which allows very large query intervals to be set. A new type of query message also exists, the *group and source specific* query. This is a group specific query with $N > 0$, followed by a N IP source addresses. It is used to query the receivers whether they are interested in receiving the specified sources for the group.

The format of the IGMP v.3 report messages is shown in the top part of Figure 4. Each report includes $M > 1$ group records, the format of which is shown in the bottom part of the figure. Each record specifies the source filtering rules requested by the receiver for the multicast group indicated. The record

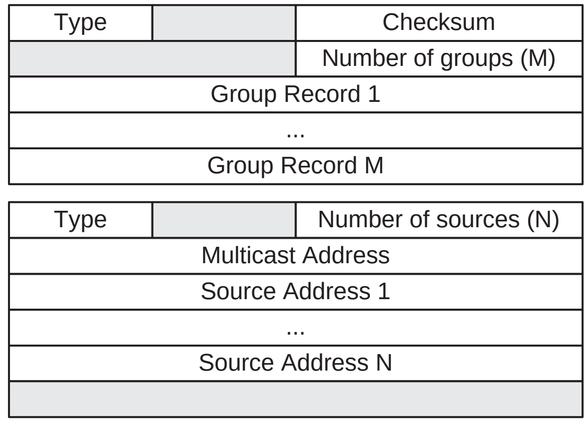

Fig. 4. IGMP version 3 report message.

includes $N \geq 0$ IP source addresses, the interpretation of which depends on the value of its *type* field. The basic types are *include* and *exclude*, meaning that the receiver wants to receive or does not want to receive the source IP addresses indicated, respectively. Other types allow the receiver to switch its list from include to exclude mode or vice versa, and to extend its list with additional IP source addresses.

Rather than providing separate leave messages, IGMP v.3 uses report messages of type *include* with an empty IP source address list. Conversely, the IGMP v.2 report messages that allow all sources to be received are equivalent to reports of type *exclude* with an empty IP source address list. That is, the IGMP v.2 report is equivalent to an *exclude none* report and the IGMP v.2 leave is equivalent to an *include none* report.

With source filtering it is unlikely that many receivers will send the same report, therefore IGMP v.3 reports are sent only to the all multicast routers group and do not suppress other reports. This means that each multicast receiver answers *all* queries related to its groups. The large query intervals of IGMP v.3 are used to spread these reports in time. The policy of not suppressing reports has two other implications: the multicast router is aware of *all* group members and each report may include information about many multicast groups.

C. Group management in MBMS

According to the MBMS specifications, IGMP (for IPv4) or MLD (for IPv6) messages are used by a UE to notify its GGSN that it wants to join or leave an MBMS multicasting group [4]. In particular, when a UE desires to join a group, is sends an IGMP join message to its GGSN stating the corresponding IP multicast address. When the UE desires to leave the group, it sends an IGMP leave message to its GGSN stating again the corresponding IP multicast address. An IGMP join message is followed by a sequence of MBMS signalling messages which form the *MBMS multicast activation* procedure, shown in Figure 5 [4]. An IGMP leave message is followed by a corresponding *MBMS multicast deactivation* procedure. Due to space limitations, we will only outline below the steps of the activation procedure, as the deactivation procedure is nearly its exact opposite.

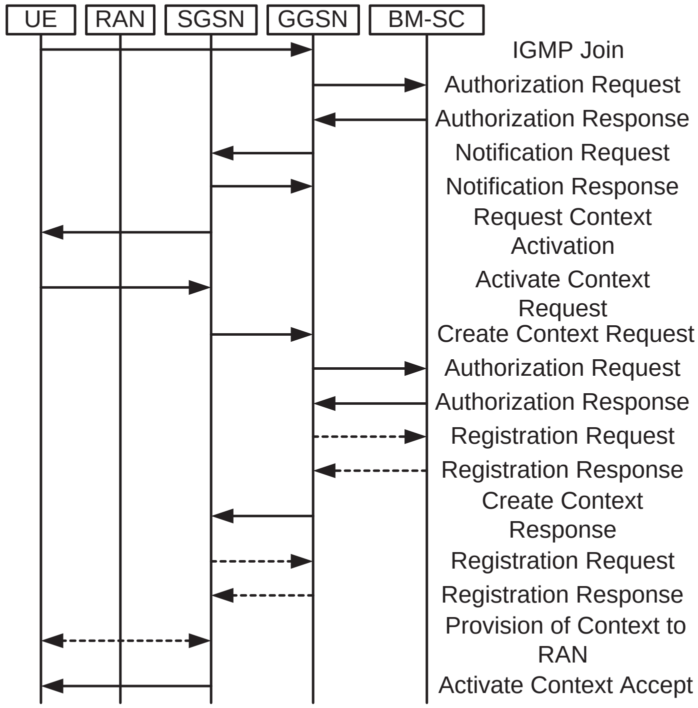

Fig. 5. MBMS multicast activation.

In the first phase of the activation procedure, the GGSN asks the BM-SC whether the UE is a subscriber to the group and the BM-SC returns the APN corresponding to the GGSN that acts as the source for the group. The GGSN then asks the SGSN whether it can handle the MBMS multicasting group. The SGSN responds to the GGSN and then notifies the UE that it can proceed with the second phase. At this point the UE knows the APN of the source, which may map to a different GGSN than the one initially contacted, and the UMTS network knows that the SGSN serving the UE can handle MBMS multicasting.

In the second phase of the activation procedure, the UE requests the SGSN to start sending it multicast data. The SGSN notifies the GGSN corresponding to the APN and the GGSN verifies with the BM-SC whether the UE is a subscriber to the group. These messages are eventually acknowledged, completing the second phase. If this is the first request received by the GGSN (SGSN) for a particular group, it registers with the BM-SC (GGSN), indicating that the BM-SC (GGSN) should start forwarding to it data addressed to the group. If the group is active, that is, transmitting data, the SGSN notifies the RAN that is should create radio bearers to transport the multicast data to the UE, if the RAN is not already transmitting these data in the cell where the UE is residing.

One obvious difference between IP and MBMS multicast group management is that there is no direct correspondence between the join and leave messages of MBMS and the IGMP v.2 and v.3 messages. We will address this issue in Section IV by mapping the MBMS messages to IGMP ones. Another obvious difference is that in MBMS the join and leave requests of applications lead to the transmission of unicast join and leave messages from the UE to the GGSN, rather than to multicast report messages triggered by queries from the GGSN. We will evaluate the MBMS group management approach in Section V. A non obvious similarity is that, even though MBMS groups are identified by an IP multicast address and an APN, the APN is not included in join and leave messages; it is instead supplied by the BM-SC, allowing the BM-SC to select an appropriate source for each group.

IV. Compatibility between IGMP and MBMS

The MBMS specifications [4] state that IGMP join and leave messages should be used in order for the UE to indicate the multicast addresses of the groups that it wants to start or stop receiving. Since there are no join messages in IGMP v.2 or v.3 and no leave messages in IGMP v.3, in this section we will map these messages to appropriate IGMP v.2 and v.3 messages and assess which IGMP version is preferable for MBMS.

The mapping to IGMP v.2 is quite simple: join messages are mapped to IGMP v.2 report messages and leave messages to IGMP v.2 leave messages. Sending unsolicited report and leave messages is allowed by the IGMP v.2 specification [5]. Unsolicited report messages are used by receivers to immediately notify their multicast router that they have joined a group, without waiting for the next query message. Leave messages are always unsolicited, and a receiver is allowed to send them in all cases, so as to avoid keeping track of whether it was the last one to report membership for a group.

The mapping to IGMP v.3 can exploit the mapping of IGMP v.2 to IGMP v.3 messages described in Section III. In this manner, the join messages of MBMS are mapped to IGMP v.3 *exclude none* reports, while the leave messages of MBMS are mapped to *include none* reports. Referring to Figure 4, this means that group records should only indicate the multicast address of a group and a type of *include* or *exclude*, without any IP source addresses. Even though multicast groups in MBMS are also identified by their source, the identity of the source should not be included in IGMP messages as it is determined by the BM-SC. Sending unsolicited reports of all types is also allowed by the IGMP v.3 specification [7].

In terms of overhead, the IGMP v.2 report and leave messages are 8 bytes long, while the IGMP v.3 report messages of type *exclude none* and *include none* are 16 bytes long, including 8 bytes for a group record without any IP source addresses. Since IGMP messages are encapsulated in IP messages with a header at least 20 bytes long, the real difference in favour of IGMP v.2 at the IP level is 28.6%. If however a UE wants to join or leave many groups at the same time, IGMP v.3 is more economical. For example, when two group records are included in an IGMP v.3 message, its IP level length is 52 bytes, while IGMP v.2 requires two 28 byte messages, a 7.7% advantage in favour of IGMP v.3. While some applications may prefer using multiple groups, for example, separate groups for audio and video, this increases the MBMS signalling overhead shown in Figure 5. We thus expect that most applications will use a single MBMS group, making IGMP v.2 more economical in most cases.

V. Evaluation of MBMS group management

While the join/leave group management model adopted by MBMS multicasting is very different from the query/report model adopted by IP multicasting, it is a natural extension of the UMTS service model. In a UMTS network users must explicitly activate a service in order to be charged for it. Therefore, for the network to start or stop forwarding multicast data to a UE, the UE must first send an explicit join or leave request to the GGSN responsible for authorizing the use of network resources by the UE. As a result, join and leave messages must always be sent from the UE to the GGSN in order for the UE to start and stop receiving data addressed to an MBMS multicast group.

In order to compare the performance of the join/leave and the query/report group management model for UMTS networks, we must take into account the overhead generated by the IGMP messages exchanged between the UE and the GGSN and the possible waste of resources due to the delay between a UE leaving a group and the GGSN stopping forwarding data to it. Fortunately, a similar join/leave approach has been previously proposed for networks where the multicast router is connected by unicast links, such as telephone lines, to the receivers [9], and its performance has been compared with the query/report approach in that environment [10]. These studies indicate that the join/leave model outperforms the query/report model over unicast links in nearly all circumstances. We can apply the results of these studies to MBMS if we can show that the differences between a unicast link network and a UMTS network do not invalidate the assumptions of the studies.

One difference is that these studies only evaluated IGMP v.1 and v.2, not IGMP v.3 which did not exist at the time. Since source filtering is irrelevant for MBMS multicasting where only a single source exists for each group, when IGMP v.3 is used for MBMS it operates in the same manner as IGMP v.2. Another difference is that in UMTS networks the GGSN is able to send multicast messages to all UEs in the same cell, even though the UE can only send unicast messages to the GGSN. Indeed, the main attraction of MBMS multicasting is that it can economize on resources by multicasting data over the air interface. While this capability allows the GGSN to send queries in multicast mode, the previous studies already ignored the cost of queries, as this cost is shared among the multicast groups that each host has joined. We conclude then that the emergence of IGMP v.3 and the ability of UMTS to send multicast query messages do not invalidate the results of the previous studies, therefore the join/leave model will outperform the query/report model in UMTS networks.

An important difference of the join/leave model proposed in [9] and the one proposed for MBMS is that the former uses explicit acknowledgments for each join and leave message transmitted in order to guard against IGMP message losses. The MBMS join/leave model does not provide acknowledgments, thus cutting the number of messages in half, but it also does not allow the UE to detect the loss of join and leave messages, so that it may retransmit them. This reliability issue can be addressed in MBMS without adding explicit acknowledgments, by exploiting the MBMS specific signalling messages. As shown in Figure 5, a join message is implicitly acknowledged by the MBMS message terminating the first phase of the activation procedure; the same is true for a leave message. While these messages are not visible to the IGMP layer as they arrive at the MBMS layer, it is straightforward for the MBMS layer to perform a cross-layer upcall to the IGMP layer when an acknowledgment arrives. Thus, after sending a join or leave message the IGMP layer should start a retransmission timer for the message, cancelling it on reception of an upcall from the MBMS layer.

VI. Conclusion

In this paper we have compared the group management mechanisms used by the IP and the MBMS multicasting models. We can summarize our findings as follows. First, while both IGMP v.2 and v.3 can be easily adapted for use with MBMS, IGMP v.2 should be preferred as it is more economical in this environment. Second, previous studies showing that the join/leave approach to group management outperforms the query/report approach over unicast links, remain valid even for UMTS networks, justifying the use of the join/leave model for MBMS. Third, even though the group management approach of MBMS potentially suffers from reliability problems at the IGMP layer, these can be easily addressed by exploiting a cross-layer upcall from the MBMS layer. These results will be used in the MBMS simulator that is currently under development by the IST B-Bone project.

Acknowledgements

The work reported in this paper has been supported by the IST B-Bone project under contract IST-2003-507607.

References

[1] *Host Extensions for IP Multicasting*, S. Deering, *RFC 1112*, August 1989.
[2] *Multimedia Broadcast / Multicast Service (MBMS); Stage 1*, 3GPP TS 22.146, V6.6.0, September 2004.
[3] *Multimedia Broadcast / Multicast Service (MBMS) user services; Stage 1*, 3GPP TS 22.246, V6.2.0, September 2004.
[4] *Multimedia Broadcast / Multicast Service (MBMS) user services; Architecture and functional description*, 3GPP TS 23.246, V6.5.0, December 2004.
[5] *Internet Group Management Protocol, Version 2*, W. Fenner, *RFC 2236*, November 1997.
[6] *Multicast Listener Discovery (MLD) for IPv6*, S. Deering, B. Fenner and B. Haberman, *RFC 2710*, October 1999.
[7] *Internet Group Management Protocol, Version 3*, B. Cain, S. Deering, I. Couvelas, B. Fenner and A. Thyagarajan, *RFC 3376*, October 2002.
[8] *Multicast Listener Discovery Version 2 (MLDv2) for IPv6*, R. Vida and L. Costa (Eds.), *RFC 3810*, June 2004.
[9] *IP Multicasting for Point-to-Point Local Distribution*, G. Xylomenos and G.C. Polyzos, *Proceedings of the IEEE INFOCOM 1997*, pp. 1382–1389.
[10] *IP Multicast Group Management for Point-to-Point Local Distribution*, G. Xylomenos and G.C. Polyzos, *Computer Communications*, vol. 21, no. 18, 1998, pp. 1645–1654.