the MH has moved to a new location in the meantime). This could lead to another round of multicast (over the xed network) by MSS_h, and so on. Finally note that this message overhead is incurred just to deliver a message to a single destination; the overhead multiplies when the message delivery protocol is repeatedly executed, once for each destination.

3. The M CAST protocol

3.1 Overview

The M CAST protocol is initiated from a MSS to deliver a message M to Dests(M), a set of destination MHs, with the guarantee that the message be delivered to each destination exactly once. The MSS initiating an execution of the protocol to deliver M will be referred to as the initiator, or $MSS_{M\ init}$. Each MSS maintains its own local counter to assign sequence numbers to multicast messages initiated by it, which is incremented whenever it initiates a new execution of M CAST. M id refers to the message-id of M; it consists of the initiating MSS's id, M init, concatenated with the sequence number, M seq, assigned to M by its initiator. A MH may also invoke an execution of M CAST: in this case, the local MSS initiates the protocol on behalf of the MH.

An execution of M CAST consists of three distinct interactions, each handled by a different module. The WIRED module handles the initial exchange of control messages over the xed network between the initiator and the participant MSSs. The WIRELESS module schedules transmission of the multicast messages over the wireless cell to its local MHs and informs the initiator (over the xed network) of the list of local MHs to which a multicast message has been delivered. When a MH moves from one cell to another, the HAND-OFF module transfers state information associated with the MH, from the MSS of the previous cell to that of the new cell.

All messages exchanged between MSSs over the xed network are assumed to be delivered in sequence with a nite, but arbitrary delay. Each MSS is equipped with a wireless interface to communicate with mobile hosts within its local cell; this interface is responsible for delivering messages in sequence to a local MH using protocols appropriate for the speci c wireless communication medium in use. All three M CAST modules use the wired network interface of a MSS to exchange messages over the static network; additionally, the WIRELESS module relies on the wireless interface to send messages to local MHs, and receive noti cation from the wireless interface on delivery of such a message to a local MH.

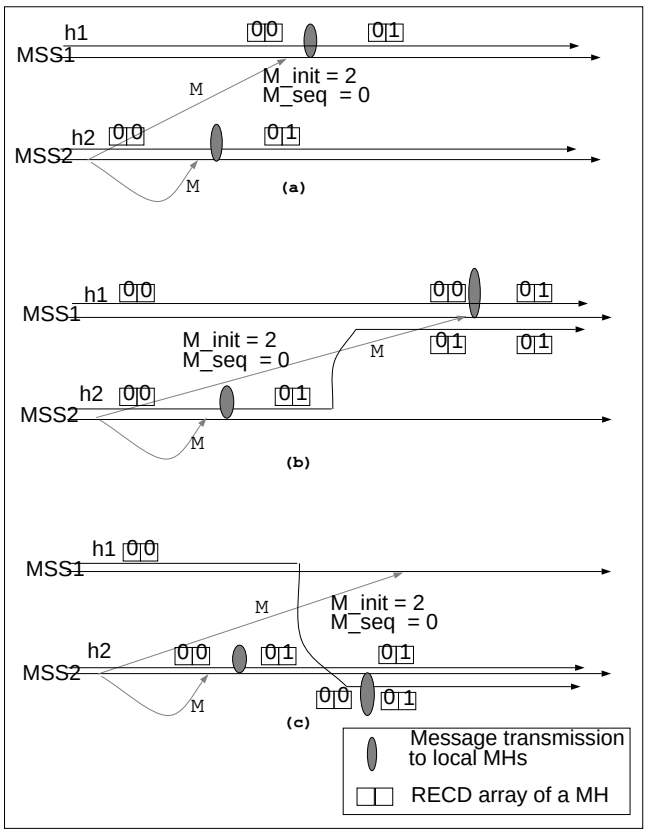

Fig. 4

The basic idea behind M CAST is simple. The protocol associates an array h RECD[1 . . .N] with every MH h, which is kept only with the local MSS. This array records the sequence number of the next multicast message from each of the N MSSs in the system, that might include h as a recipient. The HAND-OFF module hands over this array to the MSS of the new cell when a MH switches its cell. When a message M is received at the local MSS from MSS_i and its sequence number, M seq, is equal to h RECD[i], then it is delivered to h over the local wireless cell and h RECD[i] incremented if h is listed in Dests(M). If h is not included in Dests(M), then h RECD[i] is incremented without delivering M to h. In g.4, M CAST is initiated from MSS2 to deliver M to the MHs h1 and h2. Let the sequence number assigned to M by MSS2 be 0. In (a), M is forwarded to h1 by MSS1 on receipt of M, since M seq equals h1 RECD[M init]. If h1 was not included in Dests(M), then MSS1 would simply increment h1 RECD[2] on receipt of M. In (b), h2 moves to the local cell of MSS1 having already received M from MSS2. Since h2 RECD[2] is now greater than M seq, MSS2 does not deliver M to h2 for a second time. In (c), h1 moves into the local cell of MSS2 with h RECD[2] equal to 0; thus, MSS2 delivers M to h1.

M CAST is a two-phase protocol. In the rst phase,

the initiating MSS sends a copy of a message M and Dests(M), a list of destination MHs, to all MSSs over the xed network. Each recipient MSS is responsible for forwarding M over its wireless network to the local MHs that belong to Dests(M). A recipient MSS then sends a reply message to the initiating MSS containing an ack for every local MH to which it could deliver a copy of M. Since a destination MH may already have received a copy of M in an earlier cell, a recipient MSS must deliver M to only those local MHs listed in Dests(M) that have not received it in a previous location. On the other hand, a MH listed in Dests(M) might move into a new cell, without having received M in any previous cell. However, in the new cell, the MSS may already have delivered M to its local MHs. Thus, a MSS needs to keeps M "live" (even though it my have delivered it to its local MHs) till it is noti ed by the initiator to do so otherwise. When the initiator receives an ack for every destination MH, it informs all MSSs to delete M from their respective message buffers, i.e. M is no longer "live" since it is no longer possible that a MH will move into a new cell without having received a copy of M.

To multicast a message from a xed host to a group of destinations, which include both xed and mobile hosts, a copy of the message is rst sent via the xed network to each static destination. A copy of the message along with a list of mobile destinations is then sent to some MSS; this MSS invokes M CAST to deliver the message exactly once to each mobile destination. To multicast a message from a MH, the message is rst forwarded to the local MSS. This MSS sends a copy of the message to each static destinations via the xed network, and then invokes M CAST to deliver the message to mobile destinations.

3.2 Data structures

At MSS_i, the list of MHs that are local to its cell is maintained as $Local_i$. An array h RECD[1. . .N] with each element initially set to zero, is maintained for every local MH h. The signi cance of the value h RECD[j] is that any message M multicast from MSS_j with a sequence number less than h RECD[j] has been delivered to h if Dests(M) included h. The text of a multicast message M together with a list of local MH ids, ack $list_i$(M), that have acknowledged receipt of M, is kept in $buffer_i$; this is later deleted on receipt of delete(M) message from $MSS_{M\ init}$. A queue of multicast message-ids and their target local MHs are kept in $transmit_i$; the wireless interface is responsible for transmitting these messages over the local cell to the corresponding MHs, and notify the WIRELESS module on successful delivery of a message to a local MH.

MSS_i keeps track of the sequence number of the most recently received multicast message from each MSS in an array, mss $recd_i$[1. . .N], with each element initially set to zero. The signi cance of the value mss $recd_i$[j] is that the sequence number of the next multicast message expected from MSS_j is mss $recd_i$[j]; all multicast messages with a lower sequence number have been received (over the xed network) at MSS_i.

The data structures are shared by all the three modules of M CAST. To ensure consistency of the data structures, it will be assumed that a module will lock all shared data structures, perform relevant updates and then release the lock.

3.3 WIRED module

To initiate an execution of M CAST for a message M, the WIRED module at $MSS_{M\ init}$ executes the following actions:

1. Send a message over the wired network multicast(text of M, Dests(M), M id) to all MSSs (including itself).
2. Create an empty list, ack recd(M). For each ack $list_i$(M) received from MSS_i, the ids of the MHs listed in ack $list_i$(M) are added to ack recd(M). When ack recd(M) equals Dests(M), a delete(M) message is sent to all MSSs.

On receipt of multicast(M), the WIRED module of a recipient MSS, say MSS_i, executes as follows :

1. Increment mss $recd_i$[M init].
2. M is inserted in $buffer_i$ along with Dests(M). A list, ack $list_i$(M), initially empty, is also associated with M to keep track of the local MHs that have acknowledged receipt of M.
3. Create a list local dests(M), initially empty.
   a. $\forall h \in Local_i$, **if** (h_RECD[M_init] = M_seq) **then**
      if ($h \in Dests(M)$)
      then insert h in *local_dests(M)*
      else increment h_RECD[M_init]
      else
      if (h_RECD[M_init] > M_seq) **and** ($h \in Dests(M)$)
      then delete h from *Dests(M)*
   b. If *local_dests(M)* is not empty, then insert <*M_id, local_dests(M)*> in $transmit_i$, i.e. *M* is enqueued for transmission[2] over the wireless network to all local MHs listed in *local_dests(M)*. The $transmit_i$ queue is kept sorted in increasing order of message sequence numbers, with ties broken in favour of the one with a higher initiator-MSS id. Thus, *M_id*

[2] The wireless medium is physically a broadcast medium and therefore, a single message transmission should be targeted at as many local destinations as possible.

is used to correctly insert <*M_id, local_dests(M)*> in $transmit_i$.

4. Upon receipt of the *delete(M)* message at MSS_i, the entry *(text of M, Dests(M), M_id, ack_list$_i$(M))* is deleted from $buffer_i$.

Wireless interface : The wireless interface is not strictly a part of M_CAST. However, its role needs to be defined in the overall context of the protocol. First, it looks up the entry, say <*M_id, local_dests(M)*>, at the head of $transmit_i$ and transmits *M* over the wireless network to *local_dests(M)*. The wireless medium is capable of supporting broadcast transmissions; however, M_CAST does not specify whether *M* is sent in the local wireless network via point-to-point messages to each destination MH or as a broadcast transmission to all local MHs out of which only those listed in *local_dests(M)* pick up the message and respond with an ack. On receipt of such an ack, the wireless interface notifies the WIRELESS module.

The wireless interface is thus responsible for delivering messages to MHs in sequence using protocols suitable to the specific wireless medium in use.

3.4 WIRELESS module

The wireless interface notifies the module of a successful delivery of a message *M* to a MH *h* within the local cell of MSS_i, resulting in the following updates :

1. Delete *h* from *local_dests(M)* in $transmit_i$; if *local_dests(M)* now becomes empty, then the entry <*M_id, local_dests(M)*> is removed from $transmit_i$.
   Insert *h* in *ack_list$_i$(M)*;
   Increment *h*_RECD[*M_init*].
2. If the condition

   $Dests(M) \bigcap Local_i \subseteq ack_list_i(M)$ **and** $ack_list_i(M) \neq \emptyset$

   is satisfied, then at the present time, there are no local MHs that belong to *Dests(M)* and have not yet received a copy of *M*. Further, there is atleast one local MH whose ack has not been sent back to the initiator. A message containing *ack_list$_i$(M)* is now sent to MSS_{M_init} and the lists, *Dests(M)* and *ack_list$_i$(M)*, are re-initialized as :

   a. $Dests(M) := Dests(M) - ack_list_i(M)$
      i.e. the copy of *Dests(M)* stored in $buffer_i$ no longer contains those local MHs to which *M* has been delivered.
   b. $ack_list_i(M) := \emptyset$

   At a participant, MSS_i, the updates to *Dests(M)* by the WIRED and WIRELESS modules are performed on a copy of *Dests(M)* kept in $buffer_i$. Since the initiator MSS is also a participant (it receives a copy of *multicast()*), the WIRED module at MSS_{M_init} must perform both as the initiator and a participant; the WIRED module in the initiator mode uses a private (read-only) copy of *Dests(M)*.
3. **while** ($mss_recd_i[M_init] >$ *h*_RECD[*M_init*])
   { Let *M'* denote the message from MSS_{M_init} with sequence number *h*_RECD[*M_init*].
   if $M' \in buffer_i$ **and** $h \in Dests(M')$
   then if $transmit_i$ contains an entry <*M'_id, local_dests(M')*>, add *h* to *local_dests(M')*; else such an entry is inserted in $transmit_i$ with *local_dests(M')* initialised to *h*.
   Exit *while* loop.
   else *h*_RECD[*M_init*] is incremented. }

3.5 HAND-OFF module

This module relies on the *beacon protocol* of [5] to detect that a new MH has moved into a cell. Each MSS periodically broadcasts its identity (*beacon*) in the local cell. When a MH receives a beacon packet different from the previous packet it received, it responds with a *greeting* message to the new MSS containing its own identity and the identity of the MSS of the previous cell. The MSS then acknowledges receipt of the *greeting* message to the new MH.

Using the beacon protocol, the HAND-OFF module first detects the entry of a new MH *h* in the local cell of MSS_i. Then, MSS_i sends a *deregister(h)* message to MSS_j, the MSS previously handling *h*. On delivery of this message, the HAND-OFF module at MSS_j executes as follows:

1. The WIRELESS module may have delivered messages to *h* before it left the cell, but for which the corresponding updates to the *h*_RECD array (at MSS_j) hadn't been completed. The HAND-OFF module waits till the updates to the *h*_RECD array (corresponding to the messages already delivered to *h*) are completed. Note that since no new messages can be delivered to *h* in MSS_j's cell, the number of pending updates to *h*_RECD array are bounded.
2. If any entry in $transmit_j$, say for message *M'*, contains *h* in its *local_dests(M')* list, then *h* is deleted from the list; if such an entry is at the head of $transmit_j$, then the wireless interface is also informed of *h*'s departure from the local cell given that it is no longer possible to deliver any message to *h* in this cell. The entire entry is removed if *local_dests(M')* becomes empty due to *h*'s deletion.

3. Send a message *register(h, h_*RECD*)* to MSS_i; then delete the array *h*_RECD[] and remove *h* from $Local_i$.

The HAND-OFF module at MSS_i takes the following actions on delivery of *register(h, h_*RECD*)* :

1. $\forall M' \in buffer_i$
   if $h \in Dests(M')$ **and** $M'_seq <$ *h*_RECD[*M'_init*]
   then delete *h* from *Dests(M')*
2. **for** k = 1 to N **do**
   while (*h*_RECD[k] < mss_recd_i[k]) **do**
   if $buffer_i$ contains a message *M* such that
   M_init = k **and** *M_seq* = *h*_RECD[k] **and**
   $h \in Dests(M)$
   then
   if $transmit_i$ contains an entry <*M_id, local_dests(M)*>
   then *h* is inserted in *local_dests(M)*
   else <*M_id, local_dests(M)*> is inserted in $transmit_i$, with *local_dests(M)* initialised to contain *h*.
   else increment *h*_RECD[k]
3. Insert *h* in $Local_i$, and install the *h*_RECD array received with the *register()* message, in the system structures.

4. Correctness

Assumptions : It is possible that a MH that changes its cell so frequently that it does not reside within a cell long enough to receive any (multicast) message transmissions over the local wireless network. We introduce the following realistic assumption to eliminate such an occurrence :

If for a given mh *h*, every MSS has at least one "deliverable" message *M* in its *buffer* (not necessarily the same message), i.e.

- *h*_RECD[*M_init*] = *M_seq* and
- $h \in Dests(M)$

then, *h* may make only a finite number of moves without receiving any "deliverable" message.

Multicast messages are transmitted over the wireless cell in the order in which they are inserted in the *transmit* list. It is assumed that these messages are received in the same sequence at a target MH. Before a MH moves away from its current cell, it is required that it acknowledge receipt (to the local MSS) of all those multicast messages that it received within this cell.

Correctness sketch : First, note that a message *M* will eventually be received by every MSS and will be kept in their respective buffers till an explicit *delete(M)* message is received. This *delete()* is sent by the initiator MSS only after it has received an ack for the delivery of *M* to every MH listed in *Dests(M)*.

Second, *M* is delivered to a MH *h* only if *h* belongs to *Dests(M)* and *h*_RECD[*M_init*] = *M_seq*, i.e. it is "deliverable". On delivery of *M* to *h*, the local MSS increments *h*_RECD[*M_init*]. Consequently, even if *h* were to move away to a new cell after acknowledging receipt of *M*, the *h*_RECD[] array will be handed over to the MSS of the new cell only after *h*_RECD[*M_init*] has been updated, i.e. the updated *h*_RECD[] array received at the new cell as part of handoff, makes it known to the new MSS that *h* has already received *M*. Thus, the role of *h*_RECD[] is to ensure *atmost once* delivery of a multicast message to *h*.

Finally, consider how could it be that a message *M* is never be delivered to a target mh *h*. Assume that a copy of *M* is present at every MSS and is "deliverable". Earlier, it was stated that a mh may change its cell only a finite number of times without picking up *any* "deliverable" message. Then, without violating this assumption, the only way through which *M* never gets to be delivered to *h* is if it were to always pick up deliverable messages other than *M*, i.e. *M* is "starved out"[3]. To prevent such an occurrence, the HAND-OFF module always schedules the transmission of messages "deliverable" to *h* in increasing order of messages sequence numbers (either by adding *h* to *local_dests(M)* in an existing entry for *M* in the *transmit* queue, or by inserting a new entry for *M*). To facilitate this process, the WIRED module also inserts entries to the *transmit* queue in increasing order of message-ids; this ensures that, at a later time, if MH *h* were to enter a cell and messages *M* and *M'* were both "deliverable" to *h* with *M_seq* < *M'_seq*, then the HAND-OFF module will not encounter a situation where *M'* is ahead of *M* in the *transmit* queue. Thus, M_CAST ensures that, once a message *M* has been received by every MSS, it will be delivered to a target MH before the delivery of any message (to the same MH) with a sequence number greater than *M_seq*. This ensures *atleast once* delivery of *M* while the use of *h*_RECD array guarantees *atmost once* delivery: together, M_CAST ensures *exactly once* delivery of multicast messages.

5. Performance issues

In a single execution of M_CAST, the initiator must send a *multicast()* message and later a *delete()* message to all MSSs; this requires *2N* messages over the fixed network. Additionally, a participant MSS replies to the initiator with *ack_list()* messages. For MSS_i to send a

[3] *Mobility* of recipients plays a very crucial role here, e.g. consider that *h* shuttles between the local cells of MSS1 and MSS2 such that it resides in any cell long enough to receive only one message transmission. Let MSS3 issue multicast messages at such a frequency that whenever *h* switches cells, a message from MSS3 is available for delivery to *h* at the new cell. Now, if "deliverable" messages are scheduled for transmission in increasing order of initiator-MSS ids, then delivery of messages with initiator-MSS ids greater than MSS3 could be forever postponed.

ack_list$_i$(M) message, the following condition needs to be first satisfied :
Dests(M) $\bigcap$ *Local$_i$* $\subseteq$ *ack_list$_i$(M)* **and** *ack_list$_i$(M)* $\neq \emptyset$
i.e., $\mathsf{MSS_i}$ will send a non-empty *ack_list$_i$(M)* message only if there is no local MH that is an intended recipient of *M* but has not yet received *M*. The role of this condition is therefore, to collect a maximal set of acknowledgments at $\mathsf{MSS_i}$ before forwarding them to the initiator. However, in the worst case, depending on when *M* reaches every participant MSSs and the mobility of the destination MHs, the number of *ack_list()* messages sent may equal |*Dests(M)*|.

In our system model, besides the messages sent over the fixed network, the number of wireless messages is also an important parameter of performance. The M_CAST protocol associates a list of local MHs, viz. *local_dests(M)*, with an entry for *M* in the *transmit* queue. This allows for a message *M* to be sent to multiple local MHs with a single transmission using the broadcast capability of the wireless medium. However, as with sending *ack_list()* messages, a combination of message latency over the fixed network and mobility pattern of the individual target MHs could force a separate wireless message to be transmitted for each target MH.

The M_CAST protocol requires the initiator to contact every MSS. Reliable communication between the initiator and other MSSs can be implemented on top of "best-effort" multicast routing protocols for static-host groups, such as [3]. From a scalability standpoint, M_CAST is intended for delivering multicast messages to mobile hosts within a single *campus* and needs to be suitably modified for *inter-campus* operation.

6. Conclusions

This paper first presented a system model for incorporating mobile hosts within a network of fixed hosts, derived from the architecture of [5]. It was shown that mobility of hosts coupled with network latency, introduce a new dimension to the problem of reliably delivering multicast messages without duplication, even in the absence of failures. Multicast routing algorithms such as [3] and different message delivery abstractions[1] have hitherto been limited to delivering messages to static hosts. It is our thesis that even with a reliable environment, mobility of hosts offers a new set of problems that need to be solved in order to incorporate mobile hosts within static networks.

In this paper, we focussed on how delivery of multicast messages is affected by a change in location of a mobile recipient. The ability of a mobile host to connect to the fixed network from different locations at different times combined with varying message latencies between static hosts (MSSs) in the fixed network, allows for the possibility that a mobile host may receive a multicast message at more than one location or may altogether miss receiving the message. This paper presented the M_CAST protocol for efficiently multicasting a message to multiple mobile destinations. It ensures that a multicast message will be delivered to a mobile destination *exactly once* even though it may move between different MSSs during the execution of the protocol. The protocol maintained an array, *h*_RECD[*N*], for each mobile host *h* that is always stored within the static portion of the network; when *h* swicthes its cell, *h*_RECD is transferred to the local MSS of the new cell as part of the *handoff* procedure. The size of this array is proportional to the number of MSSs in the system and independent of the number of MHs; MHs can be expected to vastly outnumber MSSs.

Acknowledgements

The authors would like to thank Ajay Bakre, Navin Budhiraja, Brian Davison, Vipul Gupta, Tomasz Imielinski, James Kistler, Partha Pal, Krithi Ramamritham, Tony DeSimone, Girish Welling, Wei Zhao and the anonymous referees for their constructive comments.

Bibliography

[1] K. Birman, A. Schiper, and P. Stephenson. Lightweight causal and atomic group multicast. *ACM Trans. Comput. Systems*, 9(3):272–314, 1991.

[2] D. Cohen, J. B. Postel, and R. Rom. Ip addressing and routing in a local wireless network. Manuscript, July 16, 1991.

[3] S. E. Deering and D. R. Cheriton. Multicast routing in datagram internetworks and extended lans. *ACM Trans. Computers*, May, 1990.

[4] T. Imielinski and B. R. Badrinath. Querying in highly mobile distributed environments. In *18th Intl. Conference on Very Large Databases*, pages 41–52, 1992.

[5] J. Ioannidis, D. Duchamp, and G. Q. Maguire. Ip-based protocols for mobile internetworking. In *Proc. of ACM SIGCOMM Symposium on Communication, Architectures and Protocols*, pages 235–245, September 1991.

[6] James Kistler and M. Satyanarayanan. Disconnected operation in the coda file system. *ACM Trans. on Computer Systems*, 10(1), Feb. 1992.

[7] Yakhov Rekhter and Charles Perkins. Optimal routing for mobile hosts using ip's loose source route option. In *Internet Draft*, October 1992.

[8] Carl D. Tait and Dan Duchamp. Service interface and replica management algorithm for mobile file system clients. In *Proc. First Intl. Conf. on Parallel and Distributed Information Systems*, 1991.

[9] Hiromi Wada, Takashi Yozawa, Tatsuya Ohnishi, and Yasunori Tanaka. Mobile computing environment based on internet packet forwarding. In *1992 Winter Usenix*, Jan. 1993.