ICT-MobileSummit 2009 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2009
ISBN: 978-1-905824-12-0

Overlay Multicast Assisted Mobility for Future Publish/Subscribe Networks

Konstantinos KATSAROS[1], Nikolaos FOTIOU[1],
George C. POLYZOS[1], George XYLOMENOS[1]
[1]Athens University of Economics and Business, Patision 76, Athens, 104 34, Greece
Tel: + 30 210 8203693, Fax: + 30 210 8203686
Email: ntinos@aueb.gr, fotiou@aueb.gr, polyzos@aueb.gr, xgeorge@aueb.gr

Abstract: This paper shows how the multicast nature of proposed future publish/subscribe network architectures can assist mobility. Publish/subscribe is an information-centric networking paradigm. Unlike the current send/receive based Internet architecture which which favours the sender of information, the publish/subscribe paradigm leads to a more balanced relationship between entities. Publish/subscribe is considered a promising architecture for the future Internet as it can potentially address various current Internet problems, such as spam and (Distributed) Denial of Service attacks. In a pub/sub architecture all data are transmitted via multicast and end host identification, which has a less critical role to begin with, is decoupled from location identification. In such an environment, fast mobility can be supported effectively, particularly for continuous media distribution (such as mobile TV) and other types of real-time multimedia applications. In this paper, a prototype overlay multicast scheme based on Scribe is extended to support mobility. We compare this scheme against an alternative solution based on Mobile IPv6 and demonstrate the effectiveness of the proposed approach and the basic trade-offs.

Keywords: Future Internet, content networking, handoff, mobility management

1. Introduction

In the early days of the Internet mobility was not an issue, as the Internet was designed to resemble a telephone network where endpoints are fixed and uniquely addressed. As the Internet evolved, the need for mobility appeared, but instead of redesigning the Internet architecture, mobility was offered as an "add-on" with protocol enhancements such as Mobile IP. Nowadays however, most researchers agree that in order to harvest the full potential of the Internet a clean slate approach is needed [1]; increasing user demands, content evolution, portable networking devices, bandwidth hungry applications, real time constraints and many more, create an environment with which the current Internet architecture can barely cope. Such a clean slate approach should be organized around information, not endpoints [2], taking into consideration mobility from the very beginning rather than adding it as an afterthought.

Publish/subscribe is a communication paradigm that has been suggested as a candidate for a future Internet architecture due to the way it decouples end hosts from each other, as well as to the way it focuses on information rather on endpoints. Publish/subscribe has received great attention and many variants of the basic idea have been proposed, each adapted to different application or network models [3]. A publish/subscribe architecture divides the roles of the endpoints into two categories; publishers and subscribers. Publishers are data owners and subscribers are data consumers. Subscribers express interest on particular pieces of information, or on information patterns, in the form of subscription messages which are stored in an event notification service. Publishers make information available in the form of publications, which are also handled by the event notification service. The network is responsible for matching subscriptions with publications and notifying subscribers when a publication matches their interest. Subscribers are not necessarily aware of publishers, and vice versa, as the event notification service can transparently handle all messages exchanged. Moreover there is no need for synchronization or coordination between publishers and subscribers, as they do not have to interact with the event notification service at the same time.

The anonymity and the asynchrony of publish/subscribe systems allows them to adapt quickly to frequent connections and disconnections, making them advantageous in a mobile network [4]. At the same time, this shift towards an information centric paradigm means that multicast will become the norm rather than the exception, as it is the most appropriate mechanism for the efficient delivery of content to groups of subscribers. Hence, multicast assisted mobility [5] re-emerges as a promising research direction, albeit in a new context. In this paper we explore this new context by focusing on an overlay variant of that paradigm based on Scribe [6]: an overlay publish/subscribe system providing multicast routing. Our ultimate target is to investigate both the intrinsic characteristics of publish/subscribe and multicast network architecture, as well as the specific complications introduced by an overlay realization, with respect to mobility support. The remainder of this paper is organized as follows. Section 2 summarizes related work while Section 3 gives an overview of the publish/subscribe architecture envisioned for the future Internet. An overlay realization of the proposed architecture is presented in Section 4 which we analytically evaluate in Section 5 against Mobile IPv6. Finally, Section 6 presents conclusions and future work.

2. Related Work

Mobility and its inherent problems have been widely studied over the past years resulting in a wide range of approaches. The ***Mobile IP*** (MIP) protocol, whose most recent version was designed for IPv6 [7], was a first attempt to add mobility on top of IP. In MIP whenever a ***Mobile Node*** (MN) connects to a foreign network it is assigned a ***Care of Address*** (CoA) and it informs its ***Home Agent*** (HA) - a router in its home network which represents the mobile node when it is away - about that address. Every message towards the mobile node has as a destination address the mobile node's initial home address, and it is tunnelled by the HA to the CoA of the mobile node. This leads to inefficient triangular routing and has triggered the emergence of various optimizations such as Route Optimization. In this scheme, the ***Correspondent Node*** (CN), that is, the node currently communicating with the mobile node, is informed of the mobile node's CoA so that traffic can be directly destined there. Even with Route Optimization, MIP suffers from various problems. Encapsulation adds significant overhead to each packet and handover is not always fast enough. Furthermore, MIP handles local host mobility the same way it handles global host mobility, so the same signalling load is imposed to the Internet regardless of a MN's mobility pattern, making MIP not scalable.

In order to tackle these problems various extensions to MIP have been proposed such as Hierarchical MIPv6a [8], an extension to MIP that aims to create a clear division between local and global mobility, and Cellular IP [9] which aims to provide fast local mobility. In order to perform faster handovers various approaches, such as the fast Handover for MIPv6 scheme [10], try to predict user movement and pre-configure the next access point. However the accuracy of these prediction schemes does not reach a satisfactory level [11] and sometimes the burden they impose at the link layer is higher that the handoff time improvement they offer.

An alternative approach, multicast assisted mobility, has been widely studied in the context of IP multicast [5]. This approach is based on the creation of multicast trees that enable the wide distribution of data around the area a mobile user resides in. This wider distribution can be either statically determined i.e. data are forwarded to all neighbouring areas, or based on the expectation of possible movements of the mobile node. Multicast based mobility solutions offer the advantage of fast local handoff as well, as they provide an effective mechanism of simultaneous data delivery to various destinations. However, these approaches suffer from IP multicast's drawbacks, such as its limited scalability, which have hindered its deployment [12].

3. A Publish/Subscribe Future Internet Architecture

In the envisioned future Internet architecture all networking functions, from the lowest to the highest level, will be implemented using the publish/subscribe communication paradigm and the main data delivery method will be multicast. Endpoints will be identified by flat, location independent identifiers and routing and forwarding will take place as in the routing on flat labels [13] and data-oriented network [14] architectures, where only data identifiers are used for routing and forwarding decisions. Multicasting and the loose relationship between endpoint identifiers and their location will create an environment where mobility will be handled in a fast and effective way.

In this architecture, subscription and publication matching takes place in ***Rendezvous Points*** (RPs). RPs are organized in a hierarchical manner or via a Distributed Hash Table (DHT) in order to achieve publication resolution. Whenever a subscriber wishes to subscribe to a publication, it sends a subscribe message to its local RP (equivalent to an access router) which forwards it to the next level RPs until it reaches the RP responsible for the desired publication. Each information item is accompanied by a statistically unique identifier, called a ***Rendezvous Identifier*** (RId). Each RP is responsible for a certain range of RIds. RPs are also responsible for defining the path that a publication's packets should follow in order to reach subscribers. This path is denoted by the ***Forwarding Identifier*** (FId) label added to the header of each packet. Routers along the path can thus forward each packet along its path based on its FId.

This type of network organization favours the use of multicast and caching. Whenever a RP receives multiple subscriptions for the same RId, it can forward only one of them to the next level RP, so that there will only be a single data flow from the publisher towards this RP. Moreover a RP may cache published data for a period of time in order to serve future subscriptions. In this network architecture the notion of endpoint addressing does not exist, as every routing and forwarding decision is based on data packet identifiers. Whenever a subscriber wants to express its interest for a certain publication, it only needs to know its RId; it neither needs to know who is the publisher nor where this publication is located. In a similar way, publishers do not have to know who is accessing their publications, they only have to set an appropriate RId for their publication and advertise it to the correct RP. Finally, routers along the forwarding path do not know what the final destination of a packet is; they simple forward it towards its next hop.

4. An Overlay Based Implementation

As a first step towards the investigation of the above notions we have created a prototype publish/subscribe implementation on top of Scribe. The use of a network overlay makes the transition to the new architectural model smoother, as not all network nodes have to support it and it can coexist with current technologies. Moreover, regarding multicast support, the limited deployment of IP multicast [12] urges for an alternative solution.

4.1 Pastry and Scribe

Scribe is based on Pastry [15], an efficient and scalable DHT substrate. Unlike other DHT schemes Pastry attempts to employ proximity metrics, such as the number of IP hops or the round trip time towards other nodes, when choosing among the potentially large number of DHT nodes that may relay the data in question. Due to the use of proximity metrics the average distance a unicast message travels does not exceed 2.2 times the distance between the source and the destination in the underlying network [6]. Furthermore, according to the ***Route Convergence*** property of Pastry, the distance travelled by two messages sent to the same identifier before their routes converge is approximately equal to the distance between their respective source nodes in the proximity space. This property is of particular importance in mobile environments as will be argued and analytically demonstrated in the following sections.

Scribe exploits Pastry by mapping the name of each multicast group to an identifier and making the node responsible for that identifier the rendezvous point for that group. A subscriber to a group issues a Scribe JOIN message towards the RV of the multicast group. As the JOIN message propagates over the Pastry overlay, reverse path routing state is established until a node already in the tree is found, thus forming a multicast tree rooted at the RV point. When the RV receives a publication it forwards it to all nodes that have expressed their interest for this publication.

One of the main reasons for choosing this particular overlay multicast mechanism is that multicast routing state is maintained in a completely decentralized fashion: each node in a tree is only aware of its immediate ancestors and descendants, thus eliminating the signalling traffic required in order to maintain global state information. Hence, in a highly dynamic environment with many MNs where group membership (i.e. subscriptions) not only changes due to content-related reasons but also due to mobility, state management is simplified, thus favouring scalability.

4.2 Mobility Support

In our architecture every router participates in the Pastry and Scribe overlays. Every router is assigned a unique identifier and hence, a unique position in the identifier space. On the other hand, (mobile) end nodes neither participate in Pastry nor carry an IP address. This clearly reflects our target of breaking the end-to-end semantics of today's communication. Every node is directly connected to an *Overlay Access Router* (OAR), that is, the router providing access to the overlay network and the multicast communication substrate. Mobile nodes are connected to OARs through the currently associated *Access Point* (AP). APs may act as simple bridges to the wired part of the network, they may form groups connected to a single OAR so that link layer mobility (roaming) is provided in certain parts of the network or they can act as OARs themselves. For simplicity, we will consider the first option in the remainder of this paper.

Whenever a (mobile) node wishes to act as a publisher, it simply delivers its publication to its OAR which is then responsible to deliver it to the proper RV point. Whenever a (mobile) node wishes to subscribe to a publication it sends a SUBSCRIBE RId message to its OAR. When an OAR receives a SUBSCRIBE RId message for the first time, it issues a Scribe JOIN message in order to join the multicast group of the publication. What actually changes with mobility is that as MNs move from one AP to another it is possible that they will change their OAR as well. In this case they must inform their new OAR about the publication they are interested in, so that the OAR may join the corresponding multicast tree, unless if it is already a member of that tree. This may happen because another end node that resides on that OAR has already subscribed to the same publication or because, due to the overlay nature of Scribe, this OAR has joined the tree in order to serve another OAR as a forwarder. In that case the OAR only needs to forward the publication towards the MN.

An OAR leaves a multicast group by sending a LEAVE Scribe message *after* a period of time from the moment it anticipates the need to leave the multicast tree. This anticipation is based on the state of the association of each currently attached MN to the wireless AP. In practice, an OAR schedules the LEAVE Scribe message for a specific group to be sent some time after the last MN that is a member of that group has disassociated from the AP. The reason why an OAR does not immediately leave the group is that whenever an OAR leaves a group the multicast tree for this publication is (partially) destroyed so it will not be possible for a moving MN to take advantage of the multicast nature of Scribe and the Route Convergence property of Pastry, unless if another MN in the same area has subscribed to the same tree.

The support for mobility in the described architecture is based on the following two characteristics. First, by employing multicast as the main routing mechanism, routing information updates can be localized. This means that when a MN moves from one OAR to another, traffic can be diverged to the new point of attachment at the lowest *Common Ancestor* (CA) of the two visited OARs in the tree. Second, due to the Route Convergence property of Pastry, this CA is expected to be close enough to the new point of attachment so that the routing update can be performed without large delays. The intuition here is that when a MN moves from one AP to another, these APs are expected to be close enough in the proximity space, so the re-subscribe message of the MN will meet the CA after travelling a short distance in the proximity space.

5. Performance Analysis

In this section we investigate the performance of the proposed architecture and compare it analytically against Mobile IPv6. We chose Mobile IPv6 as it is in itself a standardized solution, and it can also serve as a means for an indirect comparison of our approach with other mobility assisting schemes. We chose to use the handoff delay as the primary performance metric since it reflects the service disruption experienced by the user. Specifically, we consider the amount of time required for a MN to be able to resume communication after a change of network location, that is, once the MN has associated with the new wireless AP. We will refer to this metric as the ***Resume Time*** (RT). Obviously, this time is heavily affected by the signalling required for the involved network entities to be informed about the MN's change of position i.e. the time required for the routing substrate to adapt to the movement of the MN.

Figure 1 shows a generic network topology in which a MN, initially residing at its home network, moves around the network from OAR to OAR while communicating with a CN via an RV. Our target is to investigate the impact of the signalling procedure required in the cases of Mobile IPv6 and the proposed architecture so that the MN can again become reachable after changing its point of attachment to the network. In the case of Mobile IPv6, we do not consider the case of Route Optimization as it incurs additional signalling overhead, thus increasing the RT.

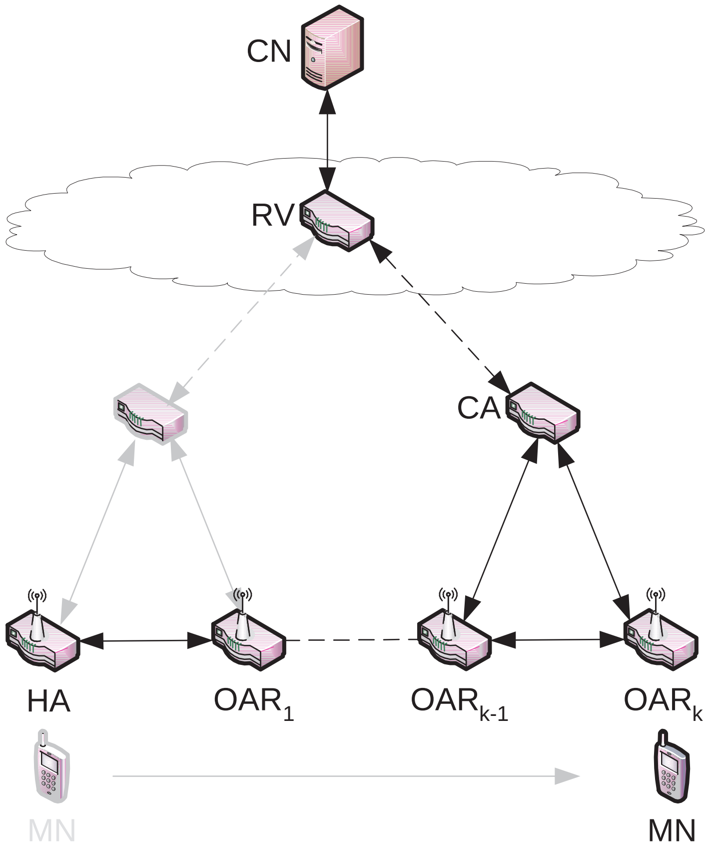

Figure 1: Example topology.

We denote by $d_{x \to y}$ the delay of a message sent from network entity x to y. In the case of Mobile IPv6, a Binding Update message is sent towards the HA of the MN and a Binding Acknowledgment message is returned to the MN. We choose to omit the latter message from the analysis in order to provide a fair comparison, as the proposed scheme's signaling does not involve acknowledgements. Hence:

$$RT_{MIPv6} = d_{MN \to OAR_k} + d_{OAR_k \to HA} \quad (1)$$

In the case of our ***Overlay Multicast Assisted Mobility*** (OMAM) scheme, a re-subscribe message is sent over the wireless medium by the MN towards the newly visited OAR which in turn generates a Scribe JOIN message towards the CA. Hence:

$$RT_{OMAM} = d_{MN \to OAR_k} + d_{OAR_k \to CA} \quad (2)$$

The following equation expresses Pastry's route convergence property.

$$d_{OAR_k \to CA} = a \times d_{OAR_{k-1} \to OAR_k}, a \to 1 \quad (3)$$

Since the delay of wireless medium transmissions is the same in both scenarios, our scheme results in a smaller RT value when:

$$RT_{OMAM} < RT_{MIPv6} \Leftrightarrow a < \frac{d_{OAR_k \rightarrow HA}}{d_{OAR_{k-1} \rightarrow OAR_k}}$$

Since according to the route convergence property $\alpha \rightarrow 1$, it is clear that our architecture results in a reduced RT compared to MIPv6, since in most cases the distance between neighbouring OARs is expected to be smaller that the distance between the current OAR and the home network of the MN. Moreover $RT \rightarrow 0$ when OAR_k is already a member of the publication's multicast tree. This may happen because another MN attached to OAR_k has already expressed interest for the same publication, or because OAR_k acting as a forwarding node for another OAR.

6. Conclusions and Future Work

In this paper we have shown how an overlay multicast architecture can assist mobility in a clean slate Internet architecture based on the information centric publish/subscribe paradigm. As this work concerns an architecture that is still under development, it leaves a lot of questions open for future investigation. First, the overhead induced by the overlay architecture needs to be measured. Moreover we have not yet considered caching strategies which are expected to accelerate data dissemination and further support reliable communication. Prediction patterns in conjunction with the delayed leave messages and proactive multicast group joins also need to be researched as they may possibly lead to even faster handoffs. Finally we need to investigate how mobility is affected when multicast has to cope with reliable transfers and two-way communications.

Acknowledgements

The work reported in this paper was supported by the ICT PSIRP project under contract ICT-2007-216173.

References

[1] A. Feldmann, "Internet clean-slate design: what and why?," *ACM Computer Communication Review*, vol. 37, no. 3, pp. 59–64, 2007.

[2] V. Jacobson, *If a Clean Slate is the solution, what is the problem?* Available at http://cleanslate.stanford.edu/seminars/jacobson.pdf.

[3] P. T. Eugster, P. A. Felber, R. Guerraoui, and A.-M. Kermarrec, "The many faces of publish/subscribe," *ACM Computing Surveys*, vol. 35, no. 2, pp. 114–131, 2003.

[4] Y. Huang and H. Garcia-Molina, "Publish/subscribe in a mobile environment," *Elsevier Wireless Networks*, vol. 10, no. 6, pp. 643–652, 2004.

[5] A. Festag, H. Karl, and A. Wolisz, "Investigation of multicast-based mobility support in all-IP cellular networks," *Wiley Wireless Communications and Mobile Computing*, vol. 7, no. 3, pp. 319–339, 2007.

[6] M. Castro, P. Druschel, A.-M. Kermarrec, and A. Rowstron, "SCRIBE: A large-scale and decentralized application-level multicast infrastructure," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 100–110, 2002.

[7] D. Johnson, C. Perkins, and J. Arkko, "Mobility support in IPv6." RFC 3775, June 2004.

[8] C. Castelluccia, "HMIPv6: A hierarchical mobile IPv6 proposal," *ACM Mobile Computer Communication Review*, vol. 4, no. 1, pp. 48–59, 2000.

[9] A. T. Campbell, J. Gomez, S. Kim, A. G. Valko, C. yih Wan, and Z. R. Turanyi, "Design, implementation, and evaluation of cellular IP," *IEEE Personal Communications*, vol. 7, pp. 42–49, 2000.

[10] R. Koodli, "Fast Handovers for Mobile IPv6." RFC 4068, July 2005.

[11] L. Song, D. Kotz, R. Jain, and X. He, "Evaluating location predictors with extensive Wi-Fi mobility data," in *Proc. of the IEEE INFOCOM*, vol. 2, pp. 1414–1424 vol.2, 2004.

[12] C. Diot, B. Levine, B. Lyles, H. Kassem, and D. Balensiefen, "Deployment issues for the IP multicast service and architecture," *IEEE Network*, vol. 14, pp. 78–88, Jan/Feb 2000.

[13] M. Caesar, T. Condie, J. Kannan, K. Lakshminarayanan, and I. Stoica, "ROFL: routing on flat labels," in *Proc. of the ACM SIGCOMM*, pp. 363–374, 2006.

[14] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, "A data-oriented (and beyond) network architecture," in *Proc. of the ACM SIGCOMM*, pp. 181–192, 2007.

[15] A. Rowstron and P. Druschel, "Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems," in *Proc. of the 18th IFIP/ACM International Conference on Distributed Systems Platforms (Middleware)*, vol. 2218 of *LNCS*, pp. 329–350, Springer-Verlag, 2001.