Available online at www.sciencedirect.com

ScienceDirect

Computer Communications 31 (2008) 980–998

computer communications

www.elsevier.com/locate/comcom

Towards transport-layer mobility: Evolution of SCTP multihoming

Ł. Budzisz [a,*], R. Ferrús [a], A. Brunstrom [b], K.-J. Grinnemo [b,1], R. Fracchia [c], G. Galante [d,2], F. Casadevall [a]

[a] Department of Signal Theory and Communications (TSC), Universitat Politecnica de Catalunya (UPC), Cl. Jordi Girona 1-3, 08034 Barcelona, Spain

[b] Department of Computer Science, Karlstad University (KaU), Universitetsgatan 2, 651 88 Karlstad, Sweden

[c] Motorola Labs – Paris, Parc Les Algorithmes – Saint Aubin, 91193 Gif sur Yvette, France

[d] Networking Lab, Istituto Superiore Mario Boella (ISMB), via Pier Carlo Boggio 61, 10138 Torino, Italy

Available online 25 December 2007

Abstract

Recently, growing availability of emerging wireless technologies has pushed the demand to integrate different wireless-network technologies such as: wireless local-area networks, cellular networks, and personal and short-range networks. The inter-working of heterogeneous radio access networks poses many technical challenges, with mobility management being one of the most important. In this paper we survey the existing proposals and show that transport-layer mobility is a viable candidate for implementing seamless handover in heterogeneous wireless access networks. Since the mobile Stream Control Transmission Protocol (mSCTP) is at the core of most relevant transport-layer mobility schemes being currently studied, we identify the key scenarios where the protocol can effectively leverage the multihoming feature to enhance handover support. Moreover, to provide the reader with a complete overview of the mSCTP's application area, we also survey the situations where the use of mSCTP-based schemes is not possible or has some limitations. Then, in one of the identified key scenarios, we investigate several challenging open issues related to path management and path-transition optimization by considering bandwidth-estimation schemes and link-layer support. Finally, we consider introducing concurrent multipath transfer (CMT) into mSCTP-based mobility schemes, as a future research direction.

Keywords: SCTP; Transport-layer mobility; Handover management; Concurrent multipath transfer; Bandwidth estimation

1. Introduction

The latest evolution and successful deployment of different wireless-network technologies (such as wireless local-rea networks (WLAN), cellular, personal and short-range) has spurred a strong demand to develop the framework for co-existence of heterogeneous wireless networks within, so-called fourth-generation (4G) mobile data networks. According to [1], one of the most important technical challenges that the development of 4G networks poses is to provide seamless mobility that can guarantee service continuity for multi-mode mobile terminals like cellular phones, personal digital assistants (PDAs), and notebook computers. Seamless mobility requires the deployment of inter-system mobility-management solutions, so that users and service providers are hidden as much as possible from the complexity of inter-networking wireless access networks. In this sense, the development of mobility-management solutions over the Internet Protocol (IP) is a key enabler to provide seamless mobility between heterogeneous wireless access networks.

Earlier work [2] on mobility management in heterogeneous networks discussed solutions affecting different layers of the IP stack. This paper, however, will mainly focus on

* Corresponding author. Tel.: +34 934015948; fax: +34 934017424.

E-mail addresses: lukasz@tsc.upc.edu (Ł. Budzisz), ferrus@tsc.upc.edu (R. Ferrús), Anna.Brunstrom@kau.se (A. Brunstrom), karl-johan.grinnemo@kau.se (K.-J. Grinnemo), roberta.fracchia@motorola.com (R. Fracchia), giulio.galante.ext@nsn.com (G. Galante), ferranc@tsc.upc.edu (F. Casadevall).

[1] K.-J. Grinnemo is now a consultant at TietoEnator R&D Services AB, Kanikenasbanken 12, S-651 15 Karlstad, Sweden.

[2] G. Galante is now a consultant at Nokia Siemens Networks, via Monfalcone 1, 20092 Cinisello Balsamo, Milan, Italy.

0140-3664/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2007.12.014

transport-layer handover schemes, as this topic has still not been given enough attention by the research community. As the main contribution of the paper, we identify the key scenarios and challenging issues in handling seamless mobility at the transport layer in heterogeneous wireless access networks. Namely, we survey the scenarios where it is possible to apply transport-layer solutions like the mobile Stream Control Transmission Protocol (mSCTP) [3], having some benefits over other existing mobility solutions, and the scenarios where it is not recommended. Specifically, multihoming support is analyzed as the new protocol feature that lays the foundations of transport-layer mobility. In this context, the key topics are supported by the development of specific experiments on path management and path-transition optimization. In particular, after an initial analysis to assess the suitability of relying on the legacy SCTP failover mechanism to handle mobility, the use of link-layer information and end-to-end bandwidth estimation are considered in the protocol-optimization process.

The rest of the paper is organized as follows: Section 2 introduces the details of mobility management and provides the reader with an overview of the existing mobility-management solutions, focusing on transport-layer schemes. Section 3 is devoted to the description of mSCTP, a transport-layer protocol that is at the core of the most relevant transport-layer mobility schemes, and to the identification of mSCTP's use for mobility management in heterogeneous wireless networks. Then, the main issues identified are illustrated by a set of experiments in Section 4. Finally, conclusions are drawn in Section 5.

2. Mobility management

2.1. Related work

In the near future, most Internet hosts will be mobile, so mobility should be supported throughout the Internet. In this context, an open challenge is the design of *mobility-management* solutions that take full advantage of IP-based technologies to achieve the desired mobility between the various access technologies, and, at the same, time provide the necessary Quality of Service (QoS) guarantees.

Mobility can be classified into terminal, personal, session and service mobility. *Terminal mobility* is the ability of a mobile host (MH) to move between IP subnets, while continuing to be reachable for incoming requests and maintaining sessions across subnet changes. *Personal mobility* refers to the ability of addressing a user that can be located at several terminals. *Session mobility* refers to maintaining a session when moving between terminals. Finally, *service mobility* can be defined as the ability of users to maintain access to their services even when moving and changing terminals or service providers. Hereafter we concentrate on terminal mobility, since it is the foundation of the analysis addressed in this paper.

Management of terminal mobility includes two fundamental operations: location and handover management. According to [3], *handover-management* deals with all the necessary operations to change a MH's point of attachment (PoA) to the IP network, while maintaining the communication with the correspondent node (CN). An IP-address change gives rise to challenges for maintaining an uninterrupted data flow, minimizing packet loss, and maintaining security. On the other hand, *location-management* focuses on keeping track of a MH's current IP address, and providing this address to any entity needing to communicate with the MH, while being transparent to its peers.

Many proposals aimed at solving the problem of terminal mobility management in heterogeneous wireless networks providing IP connectivity can be found in the literature. A good survey on the current state of the art for mobility management in next-generation all-IP-based wireless systems can be found in [2]. Currently, the most representative mobility-management solution are Mobile IP (MIP) [4] and Mobile IPv6 (MIPv6) [5]. Both MIP schemes are usually classified as macro-mobility schemes, and are tailored to follow a MH's movement across different subnets within an Administrative Domain (AD), or across different subnets belonging to different ADs. Yet, if the MH's PoA changes frequently, the MIP tunneling mechanism may lead to unacceptable network overhead in terms of increased delay, packet loss, and, especially, signaling. In this context, so-called micro-mobility schemes, such as Hierarchical Mobile IP (HMIP) [6], Cellular IP [7], Hawaii [8], and Fast Handoff [9], have been proposed to handle the movement of a MN within or across different PoAs in a subnet within an AD. A discussion of different micro-mobility protocols can be found in [10].

Another interesting approach to mobility management is introduced by Yabusaki et al. in [1]. The proposed solution advocates for the network itself to transparently handle mobility for mobile terminals. Thus, Yabusaki et al. suggest a network-centric solution to handle IP mobility in analogy with conventional 2G/3G networks, where mobility management has mainly been implemented as *network intelligence*, a concept just opposite to the *end-to-end intelligence* architectural principle of the Internet [11]. In this approach, IP addresses are used separately as host addresses and routing addresses. Thus, a *host address* is semi-permanently assigned to a MH and a *routing address* is temporarily assigned to the MH when datagrams are delivered to it. Datagrams are sent from a CN to a MH with the host address of the MH but then, within the IP mobile network, datagrams are transported using the routing address generated from the host address. All in all, user terminals are unaware of this rerouting management that is handled entirely in the network.

Handling mobility at the application layer has also received a lot of attention since a solution that is almost independent of the underlying wireless or wired access technologies and network-layer elements can be envisaged. In this context, the Session Initiation Protocol (SIP) [12] can be used for mobility management. Thus, when a MH moves during an active session into a different network, it