Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT
Computer Communications 27 (2004) 697–707
computer communications
www.elsevier.com/locate/comcom

Analysis of SIP-based mobility management in 4G wireless networks

Nilanjan Banerjee*, Wei Wu, Kalyan Basu, Sajal K. Das

Center for Research in Wireless Mobility and Networking (CReWMaN), Department of Computer Science and Engineering, The University of Texas at Arlington, Arlington, TX 76019-0015, USA

Received 23 October 2003; accepted 24 October 2003

Abstract

Providing seamless mobility support is one of the most challenging problems towards the system integration of fourth generation (4G) wireless networks. Because of the transparency to the lower layer characteristics, application-layer mobility management protocol like the Session Initiation Protocol (SIP) has been considered as the right candidate for handling mobility in the heterogeneous 4G wireless networks. SIP is capable of providing support for not only terminal mobility but also for session mobility, personal mobility and service mobility. However, the performance of SIP, operating at the highest layer of the protocol stack, is only as good as the performance of the underlying transport layers in such a heterogeneous environment. In this paper we analyze the handoff performance of SIP in a IP-based 4G network with Universal Mobile Telecommunication System (UMTS) and Wireless LAN (WLAN) access networks. Analytical results show that the handoff to a UMTS access network introduces a minimum delay of 1.4048 s for 128 kbps channel, while for handoff to a WLAN access network the minimum delay is 0.2 ms. In the former case the minimum delay is unacceptable for streaming multimedia traffic and requires the deployment of soft-handoff techniques in order to reduce the handoff delay to a desirable maximum limit of 100 ms.

Keywords: Application layer mobility management; Streaming multimedia; Handoff

1. Introduction

Fuelled by the advancement of wireless technologies and the emergence of multimedia data services, cellular wireless networks have evolved to their third generation (3G) in just two decades. However, comprehensive 3G wireless networks are yet to be available due to the costly deployment and upgrade of already deployed system equipment. It may also be possible that 3G technology will never be fully deployed. Other predictions foresee a 'generation jump' directly to 4G wireless networks [18,24]. The major task towards 4G architecture is system integration [19,22], where a unified wireless access system is to be established through the integration of the services offered by current access technologies such as General Packet Radio Service (GPRS), CDMA2000 or Wireless LAN (WLAN) as well as future wireless access technologies such as Universal Mobile Telecommunication System (UMTS). The trend towards packet switched technologies and increasingly general use and acceptance of the Internet Protocol (IP) indicate that different wireless access networks are to be connected to an IP-based core network, namely the Internet. Conceptually, a 4G wireless network architecture can be viewed as many overlapping wireless Internet access domains as shown in Fig. 1. In this heterogeneous environment, a mobile host (MH) is equipped with multiple (often called multi-mode) wireless interfaces to connect to any or all wireless access networks anytime anywhere. Therefore, providing *seamless mobility support* is one of the most challenging problems for the system integration in 4G wireless networks.

Several mobility protocols have been proposed for wireless Internet [8,10,11,15,17,21,25]. Although these protocols have the common goal of location transparency, they differ a lot from each other due to choices made during design and implementation phases. These protocols can be broadly classified based on the layer of their operation, such as those operating in the network layer [15], transport layer [21] and application layer [11]. The dependency of these mobility protocols on the access networks reduces progressively as we move up on the protocol stack [5].

* Corresponding author.
E-mail addresses: banerjee@cse.uta.edu (N. Banerjee), wuwei@cse.uta.edu (W. Wu), basu@cse.uta.edu (K. Basu), das@cse.uta.edu (S.K. Das).

0140-3664/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.comcom.2003.10.009

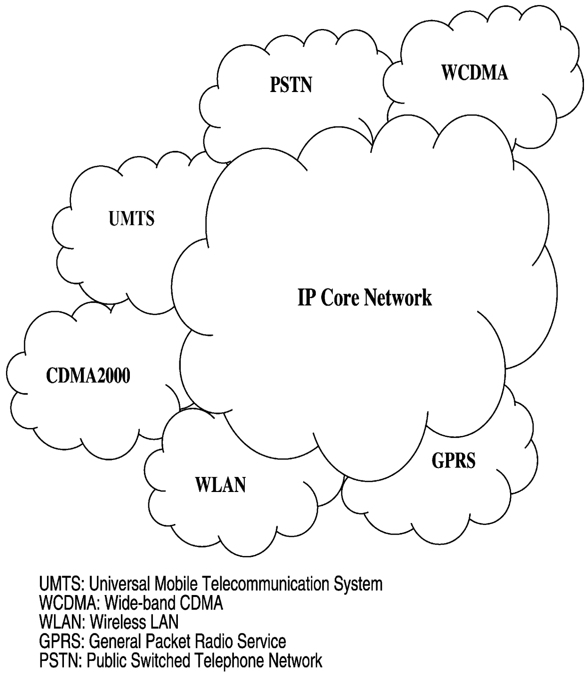

Fig. 1. Conceptual view of a 4G wireless network architecture.

Among them, Mobile IP [15] and Session Initiation Protocol (SIP) [11] have been standardized by Internet Engineering Task Force (IETF) [23] as the mobility solutions for the network layer and application layer, respectively. Although Mobile IP seems to be the architecturally right protocol for providing IP Mobility in the wireless Internet, it requires significant changes in the underlying networking infrastructure. Application layer protocols, however, are transparent to the lower layer characteristics. They maintain the true end-to-end semantics of a connection and are expected to be the right candidate for handling mobility in a heterogeneous environment. Indeed, SIP has been accepted by the third Generation Partnership Project (3GPP) as a signaling protocol for setting up real-time multimedia sessions. SIP is capable of supporting not only terminal mobility but also session mobility, personal mobility and service mobility. Therefore, SIP seems to be an attractive candidate as an application layer mobility management protocol for heterogeneous 4G wireless networks. However, SIP uses TCP or UDP to carry its signaling messages and hence is limited by the performance of TCP or UDP over wireless links. In addition, SIP entails application layer processing of the messages, which may introduce considerable delay. These are the prime factors behind the handoff delay while using SIP as the mobility management protocol.

European Telecommunications Standards Institute (ETSI) [2] has defined in a quantitative way four different classes of performance—*best, high, medium, and best effort*—for voice traffic and streaming media over IP networks [4]. The first two classes specify the type of IP telephony services that have the potential to provide a user experience better than the Public Switched Telephone Network (PSTN). Medium class has the potential to provide a user experience similar to common wireless mobile telephony services. Best effort class includes the type of services that will provide a usable communications service but may not provide performance guarantees. The specification for the end-to-end media packet delay for the best and high classes of services is less than 100 ms, while for medium and best effort classes the delay is less than 150 and 400 ms, respectively. In fact, a handoff delay of more than 200–250 ms makes voice conversations annoying. Clearly, the handoff delay, being a component of the total end-to-end delay, should also abide by these delay limits. Thus, it is evident that for quality of service (QoS) sensitive streaming multimedia traffic belonging to either *best* or *medium* class, the handoff delay should be less than 100 ms.

In this paper we investigate the performance of SIP as a mobility management protocol in a heterogeneous access networking environment predicted for 4G wireless networks. In particular, we perform a case study of SIP-based handoff delay analysis using SIP to handle terminal mobility in a IP-based network. Two different types of access technologies, viz. UMTS and IEEE 802.11b based WLAN, have been considered for the IP-based network. Analytical results show that for WLAN networks the handoff delay is suitable for streaming media but for UMTS network the minimum handoff delay does not meet the specifications. More precisely, handoff to a UMTS network from either another UMTS network or a WLAN, introduces a minimum delay of 1.4048 s for 128 kbps channel, while a handoff to a WLAN access network from another WLAN or a UMTS network, the minimum delay is 0.2 ms. Clearly, in the former case the minimum delay is unacceptable for streaming multimedia traffic and requires the deployment of soft-handoff techniques to reduce the handoff delay and keep it within a desirable maximum limit of 100 ms.

The rest of the paper is organized as follows. In Section 2, we describe the system architecture for the case study. In Section 3, we model and analyze the handoff performance of SIP for the sample architecture and present numeric results to evaluate the performance of SIP-based terminal mobility management. Section 4 concludes the paper.

2. System architecture

Telecommunication networks are gradually shifting from circuit switched to packet switched networks. At the same time the applications are converging to multimedia based applications. For our case study, we have considered an architecture conceptually similar to IP-based 4G networks in terms of heterogeneity in access network technologies. A logical view of the architecture considered is presented in Fig. 2. The architecture is primarily focused on wireless