Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE DIRECT®

Computer Communications 26 (2003) 1975–1989

computer communications

www.elsevier.com/locate/comcom

Proactive seamless mobility management for future IP radio access networks

T. Pagtzis[a,*], P. Kirstein[a], S. Hailes[a], H. Afifi[b]

[a]*Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK*
[b]*2004 route des Lucioles 06802, INRIA, Sophia Antipolis, Cedex, France*

Received 25 June 2001; accepted 25 March 2002

Abstract

Cellular communications are aligning rapidly to adopt suitable network models for supporting packet switched services over IP networks. Pure third generation (3G) wireless architectures have already adopted IP as the network layer bearer within their core network component specification. This article presents an architecture and protocol in support of *proactive mobility management* for future IP radio access networks. It encompasses a novel approach for seamless handoff and proactive allocation of PDP context with respect to *IP roaming state*. The latter establishes a generic substrate for proactive *state relocation* of different context classes relating to the state of IP connectivity for a mobile node (MN).

To address such form of IP mobility, the proposed model identifies a tentative mobility–routing matrix (TMM), which represents an accurate mapping between a mobility neighbourhood vector (MNV), surrounding the current GPRS-attachment point of an MN and the correct underlying routing neighbourhood vector (RNV), over arbitrary routing topologies. Sustained IP connectivity is achieved by introducing a 1-neighbour-lookahead (1-NL) view of PDP roaming state derived from the established TMM component; seamlessness is pursued through mapping of the 1-NL component to some handoff care-of address onto the IP Multicast domain; this allows abstracting a plurality of candidate care-of address instantiations of the MN onto a single handoff routing identifier.

Keywords: Proactive mobility; Seamlessness; Cell bounce effects; Ping-pong effects; Context transfer; State relocation; Handoff care-of address; Multicast; Mobility neighbourhood; Routing neighbourhood; PDP roaming state; Tentative mobility matrix

1. Introduction

Advances in packet-switched wireless network technologies [1,2] and portable computing terminals [3] are reaching a stage of maturity; users expect convergence in the wireless/wired IP network infrastructure, enabling true diverse access capabilities: access 'on the move', global span, constant connectivity, uniform performance characteristics, seamlessness, IP transparency.

In this engendered paradigm shift in traditional access practices emerging as *mobile networking*, users in command of portable wireless computing devices *envisage* fast and simple access through an abstract all-IP radio access network (RAN) [4].

The vision remains however, as Internet and the developments in the mobile domain have fuelled only the first stage towards an all-IP radio access paradigm shift. Radio access technologies and standards are taking a very specific view of user needs disregarding the heterogeneous composition of current networks providing mobile access, be it 802.11 [5], Bluetooth [6], GSM [7], or UMTS [8]. As a result, discontinuous connectivity is considered currently the norm rather than the exception. It becomes evident that mobility goes far beyond the ability to effect communications while wireless or on the move.

Instances of such discontinuity are manifested as mobile users associated with different wireless access networks, serve by another (or no) operator, or located under a different RAN environment. This presents a need to acquire many different terminals operating under different signalling and data protocols, whereas the sample applications are not guaranteed to be available on all of them. From the perspective of both the mobile network or application

* Corresponding author. Tel.: +44-20-7679-3666; fax: +44-20-7387-1397.

E-mail addresses: t.pagtzis@cs.ucl.ac.uk (T. Pagtzis), p.kirstein@cs.ucl.ac.uk (P. Kirstein), s.hailes@cs.ucl.ac.uk (S. Hailes), afifi@sophia.inria.fr (H. Afifi).

0140-3664/03/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0140-3664(03)00162-2

services the user remains *wired* rather than *wireless*, without the unifying properties of the IP layer abstraction, as a result of cross-standardisation disparity emergent between existing, new or forthcoming wireless access technologies.

Abstracting the individual wireless access technology (cellular, WLAN, bluetooth, etc.) behind the notion of a *network protocol* such as the Internet protocol (IP) provides an open, unifying protocol substrate for transparent operation of current as well as future application services. In this fashion, vertical integration of new wireless access technologies can be IP-embedded successfully beyond the boundaries of a typical 3G core network (CN) [9], under the horizontal integration of the application domain. This can provide the safeguards for a practical open mobile system *now*, without dismissing existing wireless access technologies in favour of a truly open, universal wireless access standard [10], which is bound to lead to yet another lengthy as well as costly infrastructural network evolution exemplar.

The benefits from all-IP based RAN architecture become instantly apparent: realistic potentials for a service-rich mobile domain; through IP the service creation cycle across the wireless technology spectrum, is shortened significantly, thus, cultivating economies of scale in introduction of new services. Significant savings on ownership and management across RAN infrastructures as signalling and the data are abstracted as IP flows rather than control and data channels. Capacity enhancement under an IP transport is easier and cheaper.

In addition, an all-IP RAN, can prove an extremely efficient transition moderator towards a truly open wireless access standard towards advanced third (3G +) or subsequent fourth (4G) generation wireless networks. This is so because the radio access domain can be incrementally augmented with new wireless access technologies which over IP can still claim a fair share of the application services domain, while a universal wireless access technology [11] is being developed.

1.1. IP-RANs and real-time services

Coupled with the notion of ubiquitous computing [12] and *nomadic* communications [13], all-IP RAN infrastructures and protocols enable a wealth of opportunities for novel kinds of user-level IP services in the application domain of interactive multimedia: navigation, personal locator services, interactive audio/video, telemetry and guidance.

Real-time dissemination of multimedia information becomes, thus, of major importance, as mobile devices and users integrate information retrieval as a peripheral task of their ‘primary’ activity (driving, operating, pursuing, walking, or generally ‘acting’). These activities require bounded latencies if IP communications are to sustain real-time guarantees in terms of both acted task performance and supporting communicated information. It has been shown extensively in Ref. [14], that for one-way delays in excess of 150 ms[1] the quality of interactive audio/video traffic degrades significantly, while beyond 200 ms it is rendered unacceptable [15].

Towards all-IP mobile networking practices, Perkins [16] proposed extensions to protocol considerations for network-layer host mobility, originally devised by Ioannidis and Maguire [17], known as Mobile IP. Proposed as the dominant standard for mobile networking, Mobile IP effects a transparent mapping between the home IP address of a mobile node (MN) and a care-of IP address (CoA) acquired at the visited point of attachment; it is characterized as a *reactive* IP mobility protocol since IP connectivity provisions at a visited link are *initiated* upon detection of an incoming MN.

Despite its wide acceptance, Mobile IP has been found [18], to be insufficient for support of real-time IP traffic. The specification for Mobile IP version 4 [19] restricts the MN in changing points of attachment not faster than once every 1000 ms. Over IPv6 networks, Mobile IP [20], continues to lack of support for delay-sensitive IP traffic, due to network layer switching latencies incurred either by core IPv6 protocol functions or due to *latency externalities* impacting directly its reactive character.

With respect to core IPv6 protocol functions, Finney and Scott [21] verify such deficiency by showing that, irrespective of the IPv6 router advertisement interval, the allocation of an IP address requires a minimum of 160 ms;[2] such delay period is accounted from the moment that the IPv6 stack of the MN[3] is notified for stateless address autoconfiguration until the moment that a Binding Update is transmitted. The above imply that allocation/activation of IPv6 addressing state generates by itself enough latency to place any active IPv6 flow on the boundaries of acceptable guarantees for real-time traffic delivery.

IP addressing state is only one of the types of context required to admit an MN and its active IP flows into a visited network; for instance, security context may have to be re-established prior to any packet flow transmissions: in the case of resource reservation for the purposes of Integrate Services QoS provisioning, the entire end-to-end path needs to be re-established at the new network link. For AAA-based admission control, credential verification must be effected with the home network prior to admission of the MN at the new network link. Each of these context states requires one or more round trip times in terms of protocol interactions at the new GPRS service node (GSN), or between visited and home network, before they are established; this is clearly beyond the control of core IPv6

[1] More accurately around 200 ms.

[2] Assumes no duplicate address detection hits, which can worsen latency although very rarely.

[3] This is orthogonal to any hardware optimizations at the access router, but implementation dependent on the IPv6 stack of the MN.