THE ROLE OF INTERNET TECHNOLOGY IN FUTURE MOBILE DATA SYSTEMS

I. Guardini[†]
Tel. +39 011 228 5424
e-mail: ivano.guardini@cselt.it

P. D'Urso[†]
Tel. +39 011 228 7745
e-mail: paolo.durso@cselt.it

P. Fasano[†]
Tel. +39 011 228 5071
e-mail: paolo.fasano@cselt.it

[†]CSELT (Centro Studi E Laboratori Telecomunicazioni)
via G. Reiss Romoli, 274
10148 Torino (Italy)
Fax. +39 011 228 5069

Key Theme: T4

THE ROLE OF INTERNET TECHNOLOGY IN FUTURE MOBILE DATA SYSTEMS

Abstract

Mobile telephony and the Internet are the fastest growing businesses in the telecommunications market. This is why most operators and service providers are looking after the development of new services in both sectors and newcomers are expected to enter the arena. The mobile operators foresee an increasing share of their revenues coming from new data services, whilst Internet Service Providers (ISPs) are attracted from wireless technology and mobility services both to reduce costs within the last-mile segment and to enrich their market share providing ubiquitous access to the Internet and corporate intranets.

In this scenario several wireless overlay networks will coexist and their interworking will be a challenging objective. The employment of the Internet technology, with its novel mobility and security extensions, seems to be the most attractive option for achieving that goal. In addition, the migration to a full IP network architecture even within each specific wireless domain will be another promising opportunity, already under consideration within several technical and standardization bodies.

The envisioned role of Internet technology makes it worthwhile undertaking significant research efforts on the development of innovative IP based mobile data systems and opens promising opportunities for both telcos and Internet Service Providers.

1 Introduction

Mobile telephony and the Internet are the fastest growing businesses in the telecommunications market. This is why most operators and service providers are looking after the development of new services in both sectors and newcomers are expected to enter the arena. The mobile operators foresee an increasing share of their revenues coming from new data services, whilst Internet Service Providers (ISPs) are attracted from wireless technology and mobility services both to reduce costs within the last-mile segment and to enrich their market potential providing ubiquitous access to the Internet and corporate intranets.

The present deregulated market will foster the emergence of many actors offering a variety of mobile data services over several wireless overlay networks to match different specific user needs. In this scenario, the deployment of innovative wireless data networks, the integration with the Internet and the interworking between different wireless technologies will be challenging objectives for competitive service providers.

This paper, starting from the analysis of the current development of service offerings and the available technologies, focuses on the main evolutionary trends, envisioning a key role for the Internet technology in the future mobile data systems.

2 IP services within mobile networks

Approaching the market of data communications, during the last few years the PLMN (Public Land Mobile Network) operators have begun to deliver data services to their customers. Currently available solutions for data transfer over the GSM network include the Short Message Service (SMS), that allows a basic e-mail exchange, and the traditional low speed Circuit Switched Data (CSD), that may be used to access Internet services. The main drawbacks of CSD are the very limited bandwidth capacity (9.6 kbps or 14.4 kbps depending on the employed coding scheme) and the sub-optimal use of the radio interface. A straightforward extension of CSD is the High Speed Circuit Switched Data (HSCSD) [1], which increases the transmission capacity by allocating to a single user up to 8 CSD channels. Anyway this solution still makes use of the radio resources in circuit switched mode and is therefore poorly optimized for the bursty traffic profile generated by most of Internet applications.

Aiming to provide more bandwidth as well as a more efficient spectrum usage, most of GSM operators are about to deploy the General Packet Radio Service (GPRS) [2], which is a fully packet oriented technology designed to support both IP and X.25. A further medium term evolution will be the migration to the 3rd generation mobile system (UMTS) [3], which will include a new W-CDMA radio access and will be able to accommodate a wide range of data rates (from 144 kbps to 2 Mbps). This evolution will also encompass relevant changes in the user terminals technology. In the meanwhile, as a short-term approach to enable advanced data capabilities even on current technology mobile handsets, GSM operators are beginning to deliver WWW like services based on the WAP (Wireless Access Protocol) technology [4].

3 Mobility within the Internet

The increasing diffusion of portable devices, such as laptops, PDAs and smart phones, have recently led to a growing demand for access to the Internet and corporate intranets independent of the technology and the point of attachment. Today's ISPs cope with these new user needs by offering a set of dial-up services including remote Internet access as well as secure access to corporate intranets established by means of tunneling protocols like PPP [5], L2TP [6] and IPSec [7]. Besides, most of current ISPs have joined together in confederations (e.g. iPass [8] and GRIC [9]) to provide network access on a wider area and in a cost effective way. New technologies for user roaming among ISPs have been deployed so that the professional on the move can access the services he has subscribed with his home ISP without having to afford a long distance call to his hometown, but just by setting up a dial-up connection to the nearest Network Access Server (NAS) managed by any of the confederation members.

A further ongoing evolution is the provision of wireless access to mobile Internet users, so that they can stay on-line even while moving, and take advantage of seamless user mobility. Available options include the use of cheap WLAN solutions (e.g. IEEE 802.11 [10], Bluetooth [11], HomeRF [12], etc.) in indoor environments and the exploitation of the wireless coverage provided by existing satellite or cellular operators in urban and rural outdoor areas. In addition, the first mobile ISPs offering wireless access to the users by their own are appearing as well. For example, Metricom [13] is currently delivering metropolitan wireless IP services (up to 28.8 kbps) in the US by means of a spread-spectrum wireless system operating within the US license-free (902-928 MHz) portion of the radio spectrum. More of such metropolitan mobile ISPs are expected to appear in the near future even outside the US, where the license-free ISM (Industrial, Scientific and Medical) bands available in the 2.4 GHz and 5-60 GHz spectrum range could be exploited.

The main drawback of the above solutions is that mobility management is not performed at the IP layer but is handled almost completely by the underlying wireless infrastructure. This may lead to sub-optimal traffic routing and does not allow seamless user mobility across different wireless media. To overcome this problem, innovative protocols for handling mobility at the IP layer are being developed by the IETF (Internet Engineering Task Force) [14].

4 Advanced IP mobility protocols

Mobile IP (MIP) [15] is the IETF proposed standard solution for handling terminal mobility among IP subnets and was designed to allow a host to transparently change its point of attachment to the Internet. MIP works at the network layer influencing the routing of datagrams and for this reason it can easily handle mobility among different media (LANs, dial-up links, wireless channels, etc.). As a main feature, MIP is completely transparent to applications in that no IP address changes are needed to allow mobility, but the mobile node (MN) can communicate using its home address regardless of its actual position in the Internet. This in turn also means that

every active TCP session survives to movements, thus allowing uninterrupted communications to wireless terminals.
The basic MIP protocol requires to assign two IP addresses to the MN: the first one is its home address (HAddr), which never changes and is used for identification purposes, and the other one is a care-of address (CoAddr), which is an address lent by the visited subnet used to determine the actual location of the mobile node. The CoAddr changes at every movement and is normally the address of a Foreign Agent (FA), e.g. a base station, acting as a local relay point for the MN. Packet routing with Mobile IP (Fig. 1) is based on the presence of a Home Agent (HA) within the home network that keeps trace of the MN CoAddr by means of a registration procedure and takes care of forwarding data traffic addressed to the MN while it is away from home.
A relevant drawback of basic MIP is the so-called triangular routing problem: all packets sent to the mobile node must transit through its home agent causing increased load on the home network and high latency. Although routing optimization solutions have been proposed to overcome this problem [16], they would require the update of every host in the Internet, which means that they might be used in practice only after a breakdown event such as the advent of IPv6 [17].

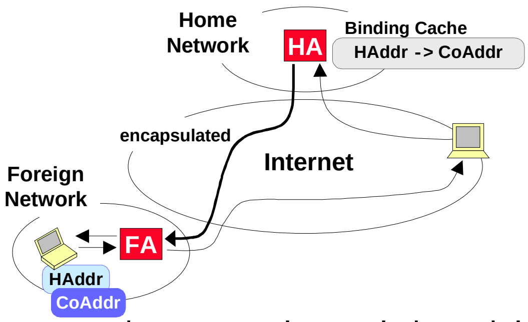

Fig. 1 - Packet routing with Mobile IP

Another significant problem with MIP is that packets in flight during a handoff are often lost because they are tunneled based on out-of-date location information. This is because the registration with the HA may involve very high delays especially when the MN is far away from home. MIP is therefore bad suited for managing the frequent local handoffs typical of micro/pico cellular environments. A promising solution is to move to a hierarchical mobility management scheme. This can be achieved by deploying new protocols (e.g. Cellular IP [18]) for managing micro-mobility in relative small areas without involving the HA, letting MIP manage only macro-mobility amongst those areas. A further protocol enhancement is to perform smooth handoffs by having the new FA notify the previous one of the new CoAddr assigned to the MN, so that packets in flight can be diverted to the correct location even before the registration with the HA has been completed. By coupling this feature with buffering in the previous FA [19] it is possible to reduce the loss of traffic that may occur whenever crossing domain boundaries, with potential benefits for UDP/RTP based real-time applications (e.g. video/audio conferencing).
Finally, it must be noted that MIP was originally designed for a flat Internet where end-to-end communications were always

possible. When dealing with ISPs domains, Intranets, Security Gateways and Proxies the MIP architecture has to be enhanced to supply roaming capabilities between different administrative domains. In fact, whenever a mobile user roams to a hosting administrative domain, an authentication procedure is required between the home and the visited domain, so that proper user authorization and accounting can be carried out. More than enhancing MIP, the IETF is facing here the great challenge of the definition and the deployment of a common AAA (Authentication, Authorization and Accounting) infrastructure for the Internet [20].

5 Evolutionary trends

The current development of innovative mobile data services is leading to the coexistence of several wireless overlay networks. On the other hand, the availability of advanced IP mobility protocols makes the Internet technology attractive for both the interworking between different networks and the definition of new wireless network architectures.

5.1 Overlay networks

IP mobility services for wireless users are currently delivered over a variety of networks and technologies, including GSM, GPRS, satellite, wireless MAN (e.g. Metricom), WLAN and many others. However, none of these solutions can be considered really universal because each one is targeted to a specific set of services and applications and has therefore different characteristics in terms of geographical coverage, bandwidth, delay, etc. As a consequence, allowing transparent user roaming amongst different wireless networks will be the best way to deliver the widest range of services anywhere and in a cost effective way for both the user and the service provider. In particular it is expected that the same geographical region will be covered by several wireless overlay networks, so that it will be up to the user to decide when to switch from one wireless access to the other based on availability or cost/performance considerations [21]. For example, when in the office, the mobile user will typically perform data retrieval from the Internet through the local 10 Mbps WLAN, even if satellite, metropolitan or GSM coverage is available. Nevertheless the same user will need to perform a seamless handoff to a lower speed wireless WAN in order to continue working almost undisturbed even when moving outside.

5.2 Interworking between overlay networks

Nowadays there is almost no integration between available mobile data networks, but each one has its own user authentication and mobility management procedures. Therefore seamless mobility among overlay networks cannot take place, but any user roaming to a new wireless domain is typically assigned a new identity (i.e. a new IP address) and any previously active communication session gets lost. To efficiently solve these problems, a common protocol for handling inter-domain user mobility has to be deployed and it is expected that in future mobile data systems the novel IP technologies described in section 4 will be used for that purpose. In particular it will be possible to deploy a common IP backbone to interconnect heterogeneous wireless IP networks and to rely on Mobile IP to manage user mobility among them

(Fig. 2). In this way any mobile user equipped with a multi-mode handset or laptop PC will be allowed to transparently roam amongst wireless domains being constantly reachable at his home address.

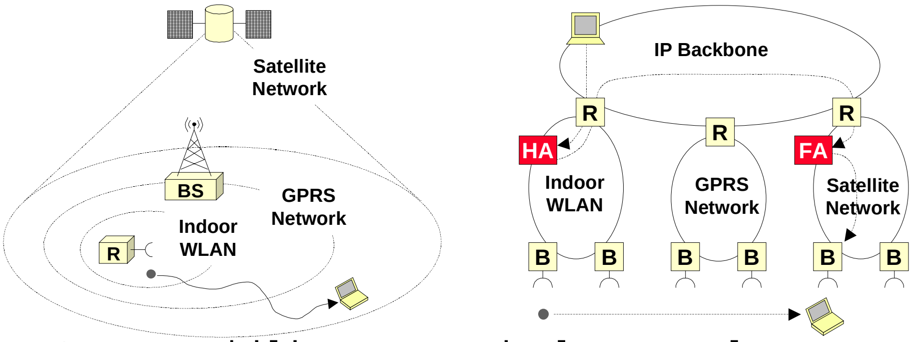

Fig. 2 - IP mobility among wireless overlay networks

The basic Mobile IP protocol enriched with the smooth handoff feature should be well suited for accommodating this kind of user mobility, in that vertical handoff between wireless overlay networks will not be a very frequent event. However, it is worth noting that this suitability is bound to the development of a common AAA infrastructure over the Internet and further enhancements in MIP to gain roaming capabilities between different administrative domains.

5.3 A full IP perspective

A key role of IP in future mobile data systems will probably be the provision of efficient and cost effective interworking between overlay networks. Moreover, both cellular operators and ISPs, together with the relevant standardization bodies (mainly 3GPP [22] and IETF), are eagerly looking at the possibility of employing IP and its mobility and security extensions even within the specific wireless networks [23]. Starting from traditional circuit switched data services on today's PLMNs (e.g. GSM CSD), where there is no IP at all within the network, several architectural enhancements can be identified to end up with more IP oriented solutions (Fig. 3).

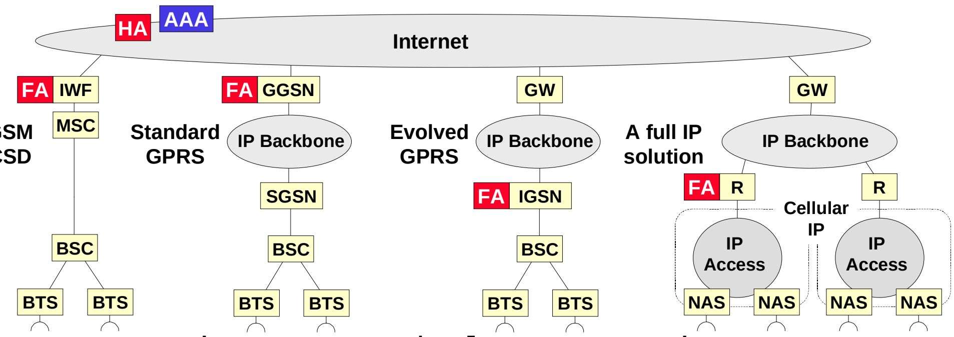

Fig. 3 - IP deployment options

The first possibility is to keep on using legacy mechanisms (e.g. HLR/VLR and SS7 signalling) to handle mobility and to deploy a new packet oriented radio interface and a common IP core network to achieve a more optimized data transport among Radio Access Networks (RANs). This is for example the solution that has been chosen for the standard GPRS system, which will be deployed by most of GSM operators by the end of year 2000. A further evolution, which is being considered by 3GPP for 3rd generation mobile systems, is to use Mobile IP for handling macro-mobility among radio access networks, thus relying on cellular telephony facilities (e.g. GPRS) just for managing micro-mobility between base stations belonging to the same RAN. Finally the most revolutionary alternative is to push IP up to the base stations by deploying a full IP infrastructure both in the core and in the access networks. With this solution each base station can be considered as being similar to an IP Network Access Server (NAS) equipped with IP mobility features and one or more wireless interfaces. The idea is to rely on Mobile IP and its hierarchical extensions (e.g. Cellular IP) for handling both macro and micro mobility and to exploit IP based authentication mechanisms (e.g. RADIUS [24]) to perform user authorization and accounting.

This full IP wireless solution has some valuable advantages, including optimal IP routing, easier integration with the Internet and other fixed IP networks and no need for SS7 signalling. In addition the provider has the opportunity to support both data and voice traffic with a single IP network by exploiting emerging technologies such as VoIP [25] and IP QoS [26]. Anyway, it is expected that during the next years the need to optimize the radio access for voice traffic and to reuse as much as possible the existing network infrastructure will continue to be the primary constraint for outdoor cellular telephony systems. For this reason, cellular operators are likely to undertake the migration to a full IP solution on PLMNs only beyond 3rd generation mobile systems.

6 Conclusion

The evolution of mobile data services outlines a trend towards the coexistence of a variety of wireless overlay networks managed by several actors and covering both indoor and outdoor environments. The Internet technology, with its novel mobility and security extensions, appears to be the most attractive for the interworking between wireless data networks, as it has proved for the wired data networks since the past two decades. The migration to a full IP network architecture even within each specific wireless domain will be another promising opportunity, already under consideration within several technical and standardization bodies. Anyway, it is foreseen that the deployment of a full IP wireless network will be feasible in the medium/short term only in indoor environments or in the case of newcomer operators. Established PLMN operators, which are delivering voice services over densely populated areas, will probably face the migration to full IP network architectures only in the longer term.

The envisioned role of Internet technology makes it worthwhile to undertake significant research efforts on the development of innovative IP based mobile data systems and opens promising opportunities for both telcos and ISPs. The mobile telephone operators will have the chance to enter the Internet market providing wireless access to third party IP networks or becoming themselves full Internet ISPs. The ISPs will have the opportunity to offer seamless mobility services either by themselves or relying on existing wireless infrastructures provided by traditional operators. Finally, a significant level of integration between the mobile data services deployed by the mobile telephone operators and the Internet Service Providers is likely to take place in the near future.

References

[1] "High Speed Circuit Switched Data (HSCSD); Stage 1", ETSI GTS GSM 02.34 v5.2.0.

[2] "General Packet Radio Service (GPRS); Service description; Stage 1", ETSI GTS GSM 02.60 v6.1.0.

[3] Special Issue, "Third Generation Mobile Systems in Europe", IEEE Personal Communications Magazine, vol. 5, n.2, April 1998.

[4] http://www.wapforum.org

[5] W. Simpson, "The Point-to-Point Protocol (PPP)", Internet Engineering Task Force, RFC1661 (Standard), July 1994.

[6] W. Townsley, A. Valencia, A. Rubens, G. Pall, G. Zorn, B. Palter, "Layer Two Tunneling Protocol - L2TP", Internet Engineering Task Force, RFC2661 (Proposed Standard), August 1999.

[7] S. Kent, R. Atkinson, "Security Architecture for the Internet Protocol", Internet Engineering Task Force, RFC2401 (Proposed Standard), November 1998.

[8] http://www.ipass.net

[9] http://www.gric.net

[10] http://grouper.ieee.org/groups/802/11/index.html

[11] http://www.bluetooth.com

[12] http://www.homerf.org

[13] http://www.metricom.com

[14] http://www.ietf.org

[15] C. Perkins, "IP Mobility Support", Internet Engineering Task Force, RFC 2002 (Proposed Standard), October 1996.

[16] C. Perkins, D. B. Johnson, "Route Optimization in Mobile IP", Internet Engineering Task Force, draft-ietf-mobileip-optim-08.txt (work in progress), February 1999.

[17] C. Perkins, D. B. Johnson, "Mobility Support in IPv6", Internet Engineering Task Force, draft-ietf-mobileip-ipv6-08.txt (work in progress), June 1999.

[18] A. Valko, A. Campbell, J. Gomez, "Cellular IP", Internet Engineering Task Force, draft-valko-cellularip-00.txt (work in progress), November 1998.

[19] C. Perkins, K.-Y. Wang, "Optimized Smooth Handoffs in Mobile IP", Proceedings of ISCC'99, Sharm El Sheikh, July 1999.

[20] http://www.ietf.org/html.charters/aaa-charter.html

[21] M. Stemm, R. H. Katz, "Vertical Handoffs in Wireless Overlay Networks", ACM Mobile Networking, Special Issue on Mobile Networking in the Internet, Winter 1998.

[22] http://www.3gpp.org

[23] T. Hiller et al.,"3G Wireless Data Provider Architecture Using Mobile IP and AAA", Internet Engineering Task Force, draft-hiller-3gwireless-00.txt (work in progress), March 1999.

[24] C. Rigney, A. Rubens, W. Simpson, S. Willens, "Remote Authentication Dial In User Service (RADIUS)", Internet Engineering Task Force, RFC2138 (Proposed Standard), April 1997.

[25] D. Minoli, E. Minoli, "Delivering Voice over IP Networks", Wiley Computer Publishing, 1998.

[26] P. Ferguson and G. Huston, "Quality of Service: Delivering QoS on the Internet and in Corporate Networks," Wiley Computer Publishing, 1998.