An architecture to access data services through cellular phone networks

J.M. Rivadeneyra Sicilia and J. Miguel-Alonso*

***Abstract*-- In this position paper, we present our research work focused on the design of a communication architecture to access TCP/IP networks (mainly Internet) through GSM. We exploit the existing and widely used TCP/IP communication architecture, but also take into account the specific features of wireless links and GSM networks. Our proposal is based in the indirect model of client/server interaction, where a intermediate element acts as a bridge between GSM and the rest of the data network. This interceptor provides enhanced functionality, improving the performance of standard TCP/IP applications in a mobile environment, maintaining the protocols in server machines, and requiring minimum changes in those running in the mobile clients. GSM operators are not required to change their network infrastructure in order to implement the architecture.**

***Index terms*-- GSM, indirect model, mobile computing, Internet access**

I. INTERNET ACCESS FROM CELLULAR TELEPHONE NETWORKS

There is an increasing interest in using the cellular phone networks for data access, as it presents some advantages with respect to other options for data access:

1. Those networks already exist, and are increasingly familiar to the user; data transmission becomes just another facility of the mobile phone.
2. They are able to provide ubiquitous access to information.
3. Mobility issues are managed by the network.

On the other hand, the characteristics of cellular telephone networks are quite different from those of wired links. New problems arise for existing applications designed to run in fast and reliable connections. Currently we are working in the design and implementation of a framework to provide the functionality required to run standard Internet applications using a cellular system as the access network, with an acceptable performance levels. We will focus on the GSM [1] (Global System for Mobile Communications) technology for the wireless access network, as it currently forms one of the most comprehensive mobile phone systems, thus being able to provide access to data services to a wide range of nomadic population.

A. Internet access using GSM

Nowadays, GSM supports circuit-switched connections, with a single error-correcting link per handset. The nominal transfer rate is 9.6 Kb/s, far from the 33.6 Kb/s of the wired telephone network (in fair conditions). Besides, actual sustainable transfer rates handled are even lower: slightly over 7 Kb/s for bulk data transfer over a good quality GSM link [2]. The use of all-digital connection from GSM to ISDN and data compression makes that rate to go up, but for some considerable time most traffic will continue to be handled by the PSTN.

Figure 1 gives an overview of the use of a GSM network to access Internet. Three different networks are traversed in the way from mobile client to fixed server (and vice-versa): GSM (digital, circuit switched), PSTN (analog, circuit switched) and Internet (digital, packet switched).

In this arrangement, the only "new" step is the first one. GSM specifies two modes of asynchronous bearer service: transparent and non transparent. In the first one the bit error rate on the radio link, within normal coverage, is assumed not to exceed the level of 10^{-3} at the line speed of 9.600 b/s and after mere forward-error correction. The non-transparent is implemented using the Radio Link Protocol (RLP), a member of the HDLC family. RLP includes error correction mechanisms, using selective retransmissions, which reduce the average bit error rate down to 10^{-8} [3]. The price to pay is an increase in latency while throughput and transmission delays become variable, remarkably so under poor conditions of the air link.

Figure 2 shows the protocol architecture used of this access. It is assumed, as it is common practice, that the non-transparent GSM data service is chosen.

* Departamento de Arquitectura y Tecnología de Computadores. Universidad del País Vasco / Euskal Herriko Unibertsitatea.
P.O. Box 649, 20080 San Sebastián, SPAIN. Tel.: +34 43 448000
E-mail: {pepe, miguel}@si.ehu.es

Location updating, and subsequent call routing, use the MSC and two location registers: the Home Location Register (HLR) and the Visitor Location Register (VLR). The HLR contains the actualized position of every mobile user in the network. The VLR contains information about every mobile station present in a particular location area. Typically there is a VLR per MSC: the group of cells under control of a MSC defines a VLR location area. When a mobile station moves to a new location area, it must announce itself to the network by means of a location update message sent to the new MSC. The mobile is registered in the VLR and a message is sent to the subscriber's HLR to update the location information.

B. *TCP/IP applications over GSM*

Several researchers have identified transport protocol problems that arise in wireless networks [4][5]. The problem relies on some design premises of TCP: it was intended for relatively fast and reliable fixed networks, and not for unstable, wireless networks. For this reason, some of the algorithms used in TCP to deal with congestion and error situations are inappropriate in a wireless environment. In fact, these mechanisms result in unnecessary waste of capacity and, at the same time, inefficient use of the (scarce) available bandwidth. Another problem of classic TCP over cellular networks is the large amount of redundant control information in the segment headers, which wastes a good deal of the available bandwidth. Some techniques are already available to reduce the amount of control information in TCP headers sent through SLIP (CSLIP) or PPP.

C. *The indirect model*

This model [6] allows to improve the performance of standard TCP/IP implementations in a mobile environment, maintaining the protocols in server machines, and with minimum changes in those running in the mobile clients. The main idea of the indirect model is simple: whenever two different environments, such as wired and unwired, are involved in a client/server connection, split the connection in two parts, one per environment. An intermediate system is located between the two connections, participates in both of them, and manages each one in a different way, suitable for its characteristics. The intermediate element looks as a server to the mobile client, and as a client to the fixed server, so the classical client/server model becomes a client/intermediary/server model. A datagram traveling from client to server (or vice versa) will traverse some layers in the protocol stack of the intermediary before being forwarded to its destination. The number of layers above the network layer traversed at the intermediary depends on the nature of the application. The applicability of the model can be exploited at any layer, but its benefits are specially appreciated at the transport [7] and application layers.

II. Proposed communications architecture

The proposed design has the following targets:

- Compatibility with TCP/IP protocols, making client mobility transparent to the common TCP/IP user.
- To establish a framework for the development of new applications, or the adaptation of existing ones, in mobile environments.
- To improve the quality of the data service of the network as it is perceived by the user, improving the throughput of network protocols. The quality of service should be close to that achievable from wired telephone networks.
- Scalability to afford new services and the increment of demand that accompany the spreading of GSM networks.
- GSM operators are not required to change their network infrastructure in order to implement the architecture, even when the new GSM standards for high speed data service (GPRS and HSCSD [8][9]) are available.

Our proposal consists of an indirect architecture, with client-server communications divided in two parts: the first one for the cellular network, controlled by special network protocols adapted for the wireless environment, and the other one for the rest of the connection, managed by conventional TCP/IP protocols. In the frontier of both sides is the intermediary, known as GSN (Gateway Support Node) using GPRS conventions. There will be one or more

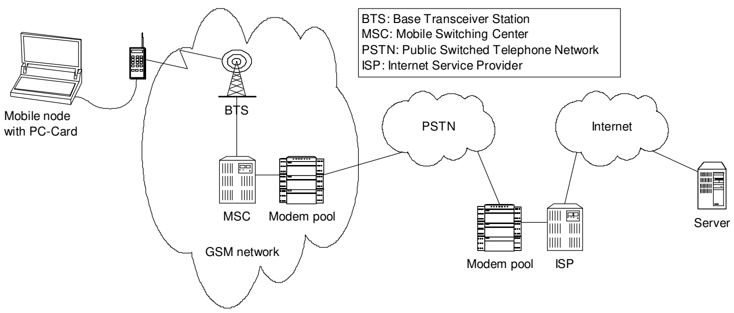

Fig. 1. Conventional access to Internet through GSM networks

GSNs in the cellular network, and the user will be connected by a circuit switched connection to one of them (it will be this way until GPRS is available). The rest of the communication, from the GSN to the server, uses packet switching. Therefore, in this proposal the intermediary is not located in the border between the wireless and wired environments (as it is usual in indirect proposals for mobile computing) but in the border between circuit and packet switching environments (when GSM data services are circuit switched) or, more generally, between the GSM network and the Internet. Next we will describe, using a bottom-up approach, the different layers in the architecture for all the constituting entities, shown in Figure 3.

A. *Network access*

Taking Figure 2 as the reference point, no changes are proposed at the lowest layers: PPP and RLP protocols are used to access the GSN. The main difference is that in this arrangement, the GSN also acts as an Internet Service Provider, thus simplifying the user access to the Internet and increasing the value added by the GSM operator. However, this is just an option. Other ones are to locate the GSN and ISP in the MSC (avoiding the use of the PSTN), or to distribute all the functions (MSC, GSN and ISP) along separate elements, with the last one out of the GSM network. Which configuration to use is a question to decide by the operator; one of our tasks is to evaluate their relative advantages and problems.

B. *Mobility management*

The mobility management procedures described in Section 2 are enough to cope with the common scenario of a mobile client who establishes a connection with a fixed server. First, the client will connect with its ISP and will get a dynamic IP address. Then, IP datagrams will flow through the circuit switched connection (or will be packet switched using GPRS) to and from the mobile client. The maintenance of that connection is guaranteed by the GSM handover mechanism.

In the case of a user who wants to access the Internet with a fixed IP address, the above solution is not appropriate. In such situations incoming datagrams cannot be correctly routed through the appropriate GSN, as the destination address is not belonging to the set of addresses managed by the GSN. Those datagrams will be routed to the home network of the mobile station. For those situations is necessary to use Mobile IP [10]. The main difference with HLR and VLR functionality is that these GSM registers are used for location updating and call routing into the GSM network, and mobile IP is used for dynamic routing in the other side of the GSN, this is, the Internet.

C. *Transport layer*

In the GSM part of the connection, a new, simplified version of TCP replaces the standard TCP (to be implemented; the proposed name is STP, Simplified Transport Protocol). The basic challenge for STP is to maximize the throughput in the connection, efficiently using all the available bandwidth. The main limitation is the modest performance and irregular behavior of the wireless link, so the protocol overhead must be minimized, and the flow control mechanism must be tailored to the characteristics of the link. STP will provide communication services to client applications in the mobile node and to proxy agents which control the execution of filters in the GSN.

The design objectives of the STP protocol are high performance, efficient recovering from expected and unexpected disconnections, and tolerance to delay fluctuations. The indirect model allows us to implement the required functionality as a simple, optimized point to point protocol. The combination of STP and TCP in both sides of the GSN constitutes an implementation of indirect TCP.

D. *Proxy layer*

The use of the indirect model at the application layer means to interpose application specific programs, called filters, in the middle of client-server connections, running in the GSN. These filters, and its management, is what we

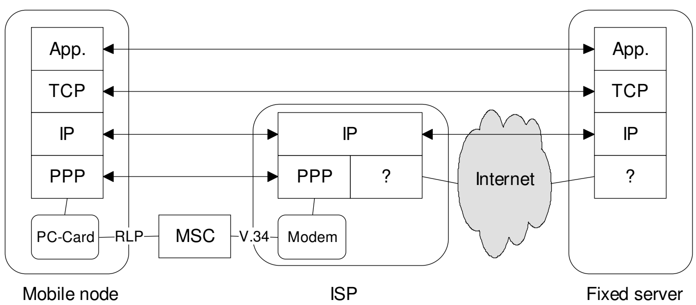

Fig. 2. Protocol architecture of a conventional connection to Internet through GSM.

call the proxy layer.

There can be general purpose filters (those that are supposed to perform well for any application) and specific filters (specifically designed for a particular application). In fact, the transport entity running at the GSN can be seen as a general purpose filter; it is just a question of naming. Actions that filters might take include selective dropping of superfluous header information (e.g., in e-mail), delay data for later pickup (e.g., POP mail), re-segmentation of data, reducing “frivolous” traffic addressed to the mobile host (e.g., replying ICMP echo requests). All of these advantages can be achieved without the use of intermediaries, but in most cases client, server or both would have to be changed. In an indirect architecture like this, changes are few and just in the client, not in the server.

E. *Adapted protocols*

The action of filters is not the only manifestation of the indirect model at the application layer of the architecture. It can support the option of mobility awareness to applications (i.e. non-transparency), allowing the use of modified application protocols in the client-proxy connection, always maintaining standard protocols in the proxy-server side. The WWW is an application where this approach suits particularly well: several efforts are devoted to the development of simplified versions of HTTP protocol [11] which would be used by specific clients (for example, a personal digital assistant) when connecting to specific proxy servers running in the GSN.

III. CURRENT WORK

Our current efforts are mainly devoted to the design and implementation of the Simplified Transport Protocol for the mobile-intermediate part of the connection. A main objective of that work is to guarantee an easy integration in the most commonly used environments for Internet access, not requiring the installation of a new version of the operating system in the client machine.

Other areas to be investigated are:

- The recent evolution of GPRS and HSCDS standards, and its future integration with our architecture, with special attention to possible changes in the bottom layers of it.
- The number of GSN to install in a GSM network, and the optimal ubication of them.
- The integration of mobile-IP home and foreign agent functionality in the GSN.
- The strategy for filter managment, considering the interaction between client and proxy. It supposes the design of an application layer protocol for it.

Acknowledgements

This work has been supported by Gipuzkoako Foru Aldundia under grant number OF 134/1997.

References

[1] M. Rahnema. *Overview of the GSM System and Protocol Architecture*. IEEE Communications Magazine, 31(4) pp. 92-100, April 1993.
[2] T. Alanko, M. Kojo, H. Laamanen, M. Liljeberg, M. Moilanen, K. Raatikainen. *Measured Performance of data transmission over cellular telephone networks.* Computer Communications Review, 24(1994)5.
[3] ETSI. *Quality of service*, GSM specification 02.08, Version 3.0.0., March 1990.
[4] H. Balakrishnan, S.Seshan, R.H. Katz. *A comparison of mechanism for improving TCP performance overwireless links* . SIGCOMM 96, August1996
[5] R. Caceres, L. Iftode. *Improving the performance or reliable transport protocols in mobile computing environments.* IEEE JSAC Special Issue on Mobile Computing Networks, 1994.
[6] B.R. Badrinath, A. Bakre, T. Imielinski, R. Marantz. *Handling mobile clients: A case for indirect interaction*. In 4th Workshop on Workstations Operating Systems, 1993.
[7] A. Bakre, B.R. Badrinath. *I-TCP: Indirect TCP for mobile hosts*. Rutgers University DCS-TR-314, 1994.
[8] Italian Cellular Home Page-GPRS. http://www.venus.it/cellular/dati/gprs/index.html
[9] Ericsson Radio Systems AB. *The mobile multi-media kit*. 1996. http://www.ericsson.se/systems/gsm/future.html
[10] M. Kojo, T. Alanko, M. Liljeberg, K. Raatkainen. *Enhanced communications services for mobile TCP/IP networking.* University of Helsinki, Department of Computer Science, Series of publications C, No C-1995-15, Apri1995.
[11] M. Liljeberg, H. Helin, M. Kojo, K. Raatkainen. *Enhanced services for WWW in mobile WAN environment*. Proc.. of the IEEE Global Internet 1996 Conference, London, England, November 1996.

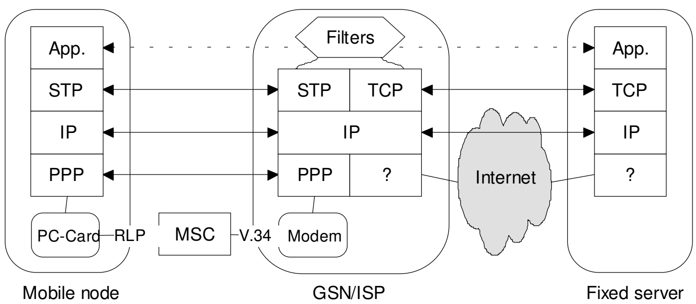

Fig. 3. Protocol architecture with an intermediary located at the GSN.