Enabling technologies for the 'always best connected' concept

Article *in* Wireless Communications and Mobile Computing · June 2006

DOI: 10.1002/wcm.392 · Source: DBLP

10 authors, including:

Nikos Passas
National and Kapodistrian University of …
140 PUBLICATIONS **1,127** CITATIONS

Alexandros Kaloxylos
University of Peloponnese
97 PUBLICATIONS **706** CITATIONS

Mairtin S. O'Droma
University of Limerick
241 PUBLICATIONS **886** CITATIONS

Ivan Ganchev
University of Limerick
236 PUBLICATIONS **796** CITATIONS

WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2005; **5**:175–191
Published online 23 August 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/wcm.207

Enabling technologies for the 'always best connected' concept

Nikos Passas[1,*,†], Sarantis Paskalis[1], Alexandros Kaloxylos[1], Faouzi Bader[2], Renato Narcisi[3], Evangelos Tsontsis[4], Adil S. Jahan[4] and Hamid Aghvami[4]

[1]*Communication Networks Laboratory, University of Athens, Panepistimiopolis, Ilisia, 15784 Athens, Greece*
[2]*Centre Tecnològic de Telecomunicacions de Catalunya, c/ Gran Capità 2-4, 08034 Barcelona, Spain*
[3]*Department of Informatics and Telecommunications, University of Catania, Catania, Italy*
[4]*Centre of Telecommunications Research, King's College London, Strand, WC2R 2LC London, UK*

Summary

'Always Best Connected' (ABC) is considered one of the main requirements for next generation networks. The ABC concept allows a person to have access to applications using the devices and network technologies that best suits his or her needs or profile at any time. Clearly, this requires the combination of a set of existing and new technologies, at all levels of the protocol stack, into one integrated system. In this paper, a considerable set of the technologies, that are expected to play a key role towards the ABC vision, are presented. Starting from a reference architecture, the paper describes the required enhancements at certain levels of a traditional protocol stack, as well as technologies for mobility and end-to-end Quality of Service (QoS) support. The paper concludes with a case study that reveals the advantages of the ABC concept. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: Always Best Connected; next generation networks; enabling technologies

1. Introduction

Second generation networks, referred to as 2G (e.g. GSM), and their successors, usually referred to as 2.5G (e.g. GPRS), provided the concept of 'Always Connected' to mobile users, offering voice and limited data services in wide areas. 3G (e.g. UMTS) and, better yet, 4G systems are expected to provide the concept of 'Always Best Connected' (ABC). ABC means that the network offers a set of access technologies and mechanisms that allow the users to be connected with the most appropriate available technology at all times, in order to enjoy the best possible service. 'Best' is usually defined separately for each user, as part of his/her profile, and it can be a function of service quality, cost, terminal capabilities, personal preferences etc. In any case, the network should have the flexibility to adjust the access technology and activate the appropriate mechanisms, in order to be consistent with the user's profile. This should be performed with no or minimum intervention of the user, leading to what is referred to as 'invisible network'. Consequently, a set of available access technologies and supporting mechanisms should be integrated in a single architecture, supporting multiple services, adjustments at all layers and

*Correspondence to: Nikos Passas, Communication Networks Laboratory, University of Athens, Panepistimiopolis, Ilisia, 15784 Athens, Greece.
†E-mail: passas@di.uoa.gr

vertical handover capabilities between different technologies [1].

The first step towards the ABC vision is the availability of a wide range of access technologies, able to support all kinds of environments. Towards this direction, the latest evolution of wireless local and personal area networks (WLANs and WPANs) will be a key enabler for in-door coverage. On the other hand, the internet protocol (IP) is considered today the basic transport technique for next generation networks, but faces serious limitations in its use for ABC provision, especially in terms of Quality of Service (QoS) and mobility support. QoS supporting mechanisms are currently developed, aiming at extending IP from a 'best-effort' technology to a QoS provision system. Additionally, IP mobility extensions will satisfy the need for roaming, as well as horizontal and vertical handovers.

This paper focuses on the main enabling technologies that aim to make ABC a reality. Section 2 describes the basic reference architecture, considered throughout the rest of the paper, that integrates different technologies in one system. Section 3 includes available enhancements for different layers of the protocol stack (communication layers, convergence, TCP, middleware). Section 4 discusses solutions for mobility support. In Section 5, end-to-end QoS mechanisms are briefly described. In Section 6, a case study aims at revealing the advantages of the integrated system for the end user. Finally, Section 7 contains our conclusions.

2. ABC Reference Architecture

The development of technologies based on the ABC concept will imply a gradual migration from today's vertically closed networks to future horizontally 'all-IP' layered networks, sharing the same backbone (Fig. 1). Integration will impact the perception of the end-user towards the provided services. Today, users are able to access their services, either by dialling in through a wired line from home to browse the web, or using the LAN at their office to read company email, or listening to voice messages using the mobile phone while waiting for the bus. In most of these cases, the set of available services depends on the access technology used by the user. With an increasing number of internet-based services, users will require having transparent and permanent access to these services regardless of the access technology they use. However, some service degradation caused by possible limitation of some access systems should be acceptable.

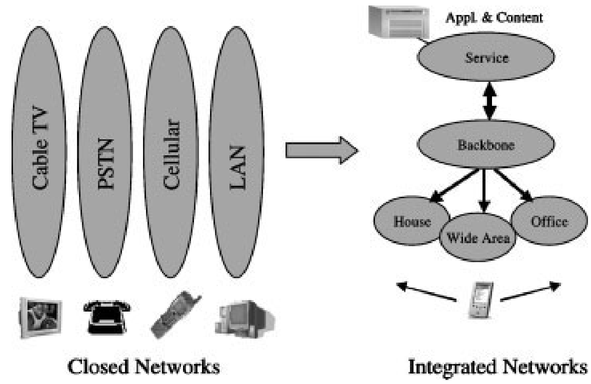

Fig. 1. Evolution to integrated networks.

From an architectural point of view, this objective drives a great effort towards three main directions:

- enhancements of the existing architectures to provide the necessary features (seamless handovers, advanced QoS, adaptable services, flexible charging policies etc.),
- integration of existing architectures (e.g. more advanced network management systems, vertical handovers, roaming etc.) and
- development of architectures for mobile terminals (MTs) able to support multi-standard access.

In Figure 2, an indicative architecture is depicted that illustrates the required enhancements of the network functionality, as well as the integration of separate networks, together with a possible terminal configuration. Although not the only alternative, this architecture gives a view of some of the required enhancements in today's networks.

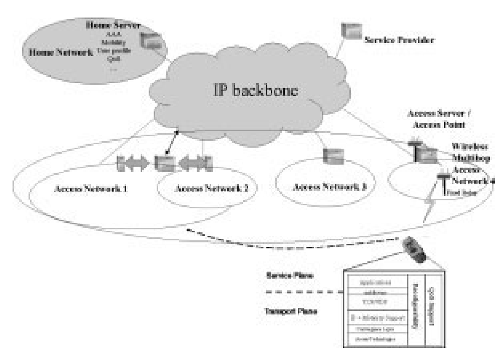

Fig. 2. Always Best Connected (ABC) reference system architecture.

In addition to the well-known layers, the future terminal architecture must include the following:

– A set of 'Communication layers', to support different access technologies, ranging from WPANs (e.g. Bluetooth) and WLANs (e.g. 802.11), to 2.5G (e.g. GPRS) and 3G systems (e.g. UMTS). This set should be able to efficiently cover environments ranging from a few meters to many kilometers. On the other hand, the minimum required set should be implemented, to avoid unnecessary increase of the required cost. Today, at least WLANs and UMTS are seen as very promising technologies for that purpose.
– A 'convergence layer', aiming at providing to the upper layers a unique link-layer interface, basically in terms of the offered QoS. As different access technologies offer different QoS capabilities, this layer will have the required degree of functionality and flexibility, in order to enhance the QoS, as seen by the higher layers to a unique and acceptable level.
– A 'middleware layer', acting as an interface between the application layer and the access selection process. Its purpose is to pass application requirements to the lower layers and inform the applications about the network conditions of the lower layers.

3. Protocol Stack Enhancements

3.1. Communication Layers

Next generation networks will contemplate the integration of a number of communication engines, placing them in a strategic position towards the ABC vision. Recent advances in various access technologies show the benefits of such an integration. Especially in the area of WLANs, the activities of the 802.11 task groups reveal the will for improving the performance in local area environments by extending the functionality to cover traditional weaknesses of these networks, such as security, advanced QoS, handover support etc. More specifically, 802.11g uses the same PHY scheme as 802.11a in the 2.4 GHz band, aiming to offer transmission speeds beyond 20 Mbps. The 802.11f task group is currently working on specifying the inter access point protocol (IAPP) that provides the necessary mechanism for information exchange between access points (APs) needed to support the 802.11 distribution system functions (e.g. handover). The 802.11e task group is currently adding extra functionality to the 802.11 MAC layer to improve QoS for better support of a larger set of applications. Finally, 802.11i incorporates stronger encryption techniques to enhance the security of 802.11, in order to be suitable for confidential information exchange. More details about these enhancements can be found in Reference [3].

Reconfigurability and adaptability on the other hand, are considered as essential parts to achieve the interoperability between the different technologies. The main targets are:

(i) to achieve full interoperability between the different communication technologies (GSM, UMTS, WLAN, ad-hoc networks etc.),
(ii) to use adaptable and reconfigurable physical layer resources, able to absorb environmental changes, and
(iii) to use the optimum power consulting mode.

To reconfigure any part of the communication layer, it is necessary for the network to have some intelligence and reconfiguration control (Fig. 3). The intelligence decides what part(s) of the network should be reconfigured, based on the relevant information supplied to it, and then instructs the reconfiguration controller to implement these decisions in the most appropriate way. Intelligent reconfigurability for the ABC concept should take into account the following essential components: reconfigurable network, software reconfigurable languages, radio environment, user status (i.e. the applications profiles) and network status (i.e. the current states of the different hardware and software components of the physical (PHY), and the medium access control (MAC) layers).

A key architectural component supporting reconfigurability control and application adaptability, as well

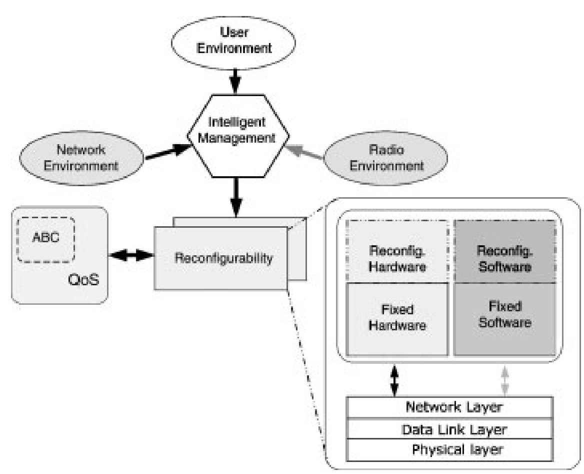

Fig. 3. Essential elements of reconfigurability involved in the communication layers for ABC.

as adaptability of link, physical and other layers can be through the development of a cross-layer protocol entity [4–6]. This emerging idea has been motivated by the need to introduce major degrees of adaptability and efficiency to variations of the actual communication systems, and thus to take a step forward the challenges of adaptability/reconfigurability and the ABC concepts required for the 4G systems. The cross-layer approach aims at introducing a degree of knowledge, offering optimization between the physical and link layers and taking into account both PHY and MAC characteristics. The exchanged information in the cross-layer can be classified as:

- **channel state information** (CSI) (i.e., estimation of the channel impulse response, location information, signal strength, interference level etc.),
- **physical layer resources** (i.e. number of antennas, spatial processing etc.),
- **QoS** (i.e. throughput, delay, bit error rate (BER), etc.).

On the other hand, it becomes more and more evident that elements such as smart antennas or multiple input multiple output (MIMO) elements and scalable detections will play an important role in modern wireless systems and they will be the main physical layer support resources for achieving the ABC strategy in the future communication systems. Smart antenna reconfigurability will be enforced by algorithms that implement adaptive channel and bandwidth allocation, as well as power control. However, multiple antennas can offer substantial spectrum efficiency and link capacity. Transmit and receive algorithms as single detection, multi-user detection, or scalable detection are also very important for the ABC, since the performance (of the Tx/Rx schemes) can vary due to the use of a specific algorithm [4,5].

Reconfigurable software is also one of the essential elements in the reconfigurability process at the communication layer level, and it must be carried out by the introduction of new program code in the user terminal, with the aim of modifying its configuration and/or contents (Fig 3). The downloading process encompasses not only the protocol or the software entities to be downloaded, but also the method and performance of the download [7]. Software reconfigurability for the ABC strategy could be divided in two categories:

- lower-level software components (e.g. physical protocol entities for more structural modification of the air interface),
- software components and parameters for modification of the PHY layer, including DSP algorithms and FPGA reconfiguration (addressing framing and channelising issues, modulation schemes, power amplifier efficiency and linearization algorithms and settings etc.).

The evolution of software downloading for ABC software radio reconfigurability may move through the following stages [7]:

- *out-call (static download)*: Software components are downloaded into a secure sandbox for installation at an appropriate time;
- *in-call (dynamic download)*: Software reconfigurability components are downloaded and installed during a call to support dynamic service reconfiguration (for ABC) or distributed processing, requiring over-the-air download.

3.2. Convergence Layer

The different characteristics of wireless links compared to fixed links, pose special requirements on the interworking between the network layer and the wireless link layer. Radio resources are typically scarce and packet loss may be extensive. An important aspect for QoS is radio resource management (RRM). In traditional cellular networks, RRM refers to such functions as layer-2 admission control, congestion control, handover management and power control. The concept of a well-defined RRM functional split allows inter-technology handover (i.e. vertical handover). Additionally, the point of attachment to an access network may change suddenly, which inflicts fluctuations in QoS and may cause a need to change the routing path. Although many operations, such as handover, can be entirely handled at layer-3, layer-2 information (such as signal strength) can improve performance significantly. It has been widely recognized that assistance from link layer mechanisms is a prerequisite for devising efficient fast handover solutions for wireless IP access networks. To address the above stated needs for cooperation between layers 2 and 3, research efforts are focusing on the standardization of the various convergence layers (CL) and interfaces. Below, we describe the wireless adaptation layer (WAL) as an indicative example.

The WAL architecture was developed in the context of the European project IST-WINE, and is considered as an intermediate layer located between the IP and the lower layers (Fig 4) [8]. The WAL incorporates a set of functional modules, viewed as generalized

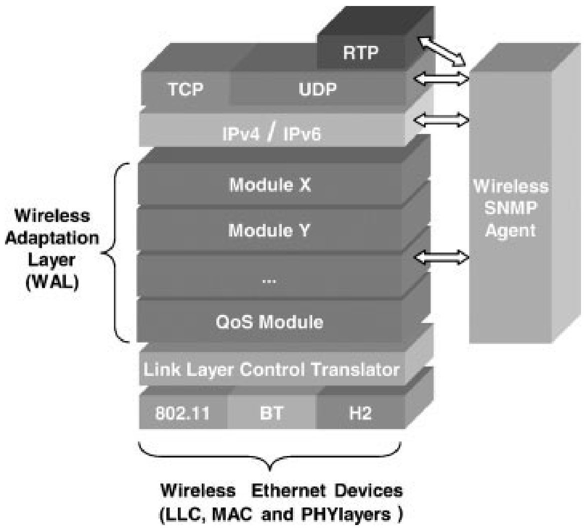

Fig. 4. The wireless adaptation layer.

performance enhancing proxies (PEPs) [9], which can be dynamically combined and adapted to the special characteristics of the wireless link and the transport protocol. Both classical link layer protocols and transport protocol specific techniques can be applied in this fashion, along with service differentiation techniques.

A main feature of the WAL is an abstraction used for service provisioning at the link layer. Each IP datagram is classified into classes and associations. Service provision in the WAL is based on these two concepts. A WAL class defines the service offered to a particular type of application traffic (e.g. audio/video streaming, bulk transfer, interactive transfer, web), and the sequence of link layer modules (protocol components) that provide such a service. The module list for every class is completely defined so that every WAL MT uses the same module order within the same class. This approach allows the WAL packet classification to be mapped onto existing internet QoS classes. An association identifies a stream of IP datagrams belonging to the same class and destined to a specific MT, i.e. WAL_Association = < WAL_Class, MT_Id >. A fair allocation of bandwidth can be easily achieved if based on a per-user operation. In addition, services for particular users can be customized to meet their QoS requirements and to implement a differentiated-charging policy. Another advantage of distinguishing IP streams with respect to their destination is that channel state conditions can be taken into account. In fact, as the condition of each wireless channel varies independently, the parameters of the modules defined for a class will be adjusted dynamically to adapt them to changes occurred in a channel.

The WAL coordinator shown in Figure 5 may be viewed as the central 'intelligence' of the WAL. Both downstream and upstream traffic passes through the WAL coordinator before being processed by other modules. In the downstream flow, the WAL coordinator intercepts IP datagrams and decides on the sequence of modules that these datagrams should pass through, as well as the parameters of these modules. The sequence of modules for each IP flow is chosen on the basis of specific fields in IP datagrams' headers, identifying the 'class' to which the datagrams belong. In the upstream flow, the WAL coordinator accepts WAL frames (encapsulated IP datagrams) and passes them through the sequence of modules associated with the class in the reverse order. Information about the modules' sequence as well as the required module parameters is contained in the WAL header described later in this section. To determine the optimum module parameters, the WAL coordinator monitors the channel conditions through continuous measurements. The WAL configuration parameters can be setup remotely via simple network management protocol (SNMP) in the local 'wireless' management information base (MIB).

The WAL coordinator maps the internet QoS classes onto WAL classes in order to provide flow isolation and fairness guarantees through traffic shaping and

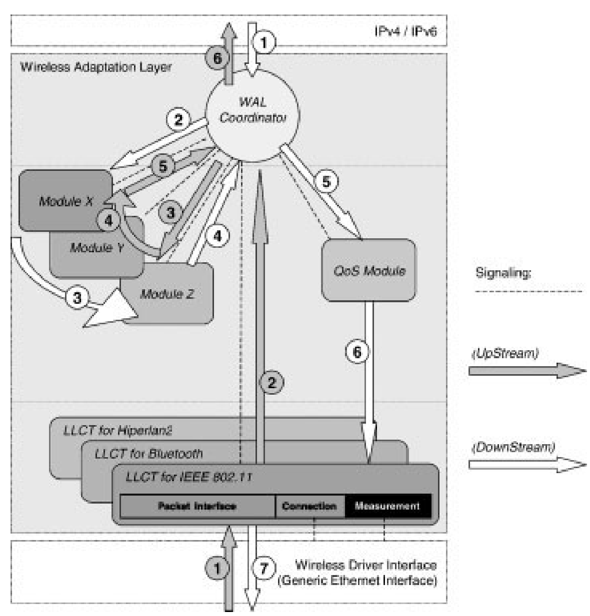

Fig. 5. The wireless adaption layer (WAL) architecture.

scheduling. The module performing flow regulation and scheduling is referred to as the QoS module. A packet scheduler is the core of the QoS module as it allows wireless resource sharing among traffic classes, according to their association. It is divided in two levels. The first level of the scheduler is implemented both in the AP and MT and consists of a class-based queuing mechanism preceded by a traffic shaper for each traffic class. The second or main level scheduler is only implemented in the AP and is responsible for allocating the wireless network (or MAC level) bandwidth to each MT. The objective of the main scheduler is to attain both throughput maximization and fairness in bandwidth allocation. Modules X/Y/Z comprise a pool of modules, aiming to improve performance in a number of ways. This pool includes error control modules such as forward error correction (FEC), a Snoop module for TCP performance improvement [10], header compression module, an Automatic Repeat ReQuest (ARQ) module and a fragmentation module to reduce packet loss probability. Other modules may be included in later versions of the WAL, to further improve the overall performance. Finally, in order to interface with a number of wireless drivers of different platforms, a wireless technology specific logical link control translator (LLCT) module for each different platform has been introduced. The main functions of this module are: (1) to manage the connection status with the wireless driver; (2) to ensure the stream conversions toward the wireless driver; (3) to perform channel measurements, via the driver; (4) to control MT registration and termination processes.

To measure the performance improvement of WAL, several simulations were performed with the use of the OPNET simulation tool. For example, the case of two MTs was investigated, each one communicating through the AP with an FTP server and a fixed host. Each MT requested to download a 2 MB file from the FTP server every 10 s, and had an active bi-directional VoIP connection with the fixed host. The UDP protocol was used for the voice transfer and each voice source was generating traffic at the rate of 64 kbps, simulating a PCM quality speech. HIPERLAN/2 was used as the access technology, operating at 6 Mbps. The overall VoIP delay observed in the system with and without WAL is presented in logarithmic scale in Figure 6(a). As shown in the figure, the VoIP delay is below the threshold of 50 ms with the use of the WAL (assuming a 100 ms delay as the maximum acceptable value for round trip delay in voice communications), while the absence of any adaptation mechanism (NO_WAL) results in an undesirable delay up to

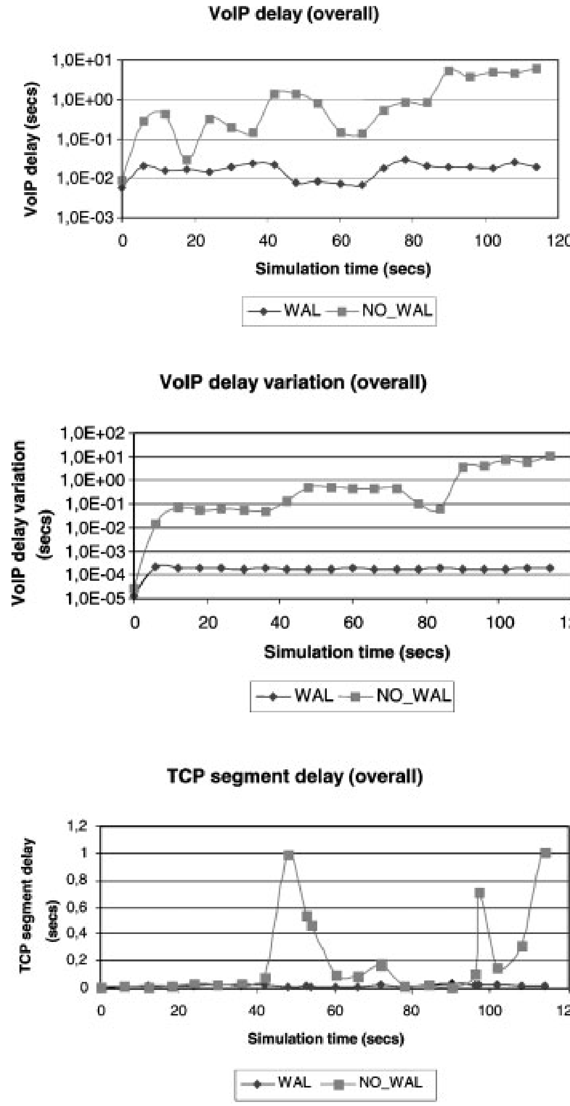

Fig. 6. WAL performance results.

10 s. The main reason for this is that the WAL (QoS module) always handles UDP traffic (VoIP) with higher priority than TCP traffic. The delay variation in the system using WAL follows the delay statistics behavior and reaches an almost fixed value of 0.2 ms, as shown in logarithmic scale in Figure 6(b). Without the WAL, voice experiences variation of more than 10 s, which makes the communication impossible. The most interesting result in a high loaded system is the way the WAL manages to schedule the transmission of both the TCP and UDP traffic without exceeding the QoS limits. The WAL seems to be able to control not only the quality of UDP flows, but also of the TCP flows. As shown in Figure 6(c), in an

overloaded system the WAL keeps TCP segment delay stable and low, while without the WAL the TCP segment delay has great variances and this causes more timeouts and retransmissions. This can be explained by the fairness and the stability that the WAL achieves by reserving a fixed bandwidth for FTP applications. In contrast, without the WAL FTP traffic is sent through the link in an abnormal way.

3.3. TCP Enhancements

Transmission control protocol (TCP) is a connection-oriented transport layer protocol that provides reliable delivery of data streams. TCP connections experience very low throughput in wireless networks, primarily due to bandwidth limitations, long round trip times (RTT), high BER and user mobility. In order to enhance the performance of TCP in wireless environments, three different approaches have been proposed in the literature: (i) link layer (LL) solutions, (ii) TCP modifications and (iii) new transport protocols.

3.3.1. LL solutions

These solutions operate at the LL in such a way that the TCP connection takes place in a dependable communication environment, with characteristics comparable to wired communications. We can distinguish between:

- *TCP-aware LL protocols*: The most important one is SNOOP [10], which is applicable to wireless cellular networks. Its major goal is to improve the performance of communication over wireless links without triggering retransmission and window reduction policies at the transport layer. A SNOOP agent, residing at the base station, buffers unacknowledged data segments destined for the mobile hosts and deals with eventual duplicate acknowledgments, instead of forwarding them to the data source.
- *TCP-unaware LL protocols*: The most important one is the TULIP [11], which was designed for half-duplex wireless channels with limited bandwidth. TULIP is service-aware in that it provides reliability only for those packets that require such service. It buffers packets locally in order to recover from losses on the wireless link, before the TCP sender times out. Performance results show that TCP-unaware LL solutions have better performance than TCP-aware LL protocols over half-duplex radio links.

3.4. TCP Modifications

In this kind of solutions, the algorithms of TCP are modified to overcome specific problems. Three main representatives of this category are the following:

- TCP selective acknowledgments options (TCP SACK) [12] were proposed to overcome TCPs ineffective handling with bursts of packet drops in a single window of data. The TCP layer at the receiving side sends back SACK packets to the sender notifying the data that have been received. The sender implements a mechanism to retransmit only the missing data segments. The standard congestion control algorithms are not affected by this modification.
- Indirect TCP (I-TCP) [13] splits the TCP connection at the base station. The base station runs a TCP connection with the fixed host and a connection with the mobile host using a protocol optimized for wireless links. Although straightforward in its implementation, if faces a number of disadvantages that can reduce its performance significantly.
  - It violates the TCP end-to-end semantics.
  - I-TCP does not handle handovers efficiently.
  - The wireless link should be the last part of the connection path.
  - It cannot be used if end-to-end IP encryption is utilized.
- **M-TCP** [14] also splits TCP connections at the base stations, but preserves TCP semantics and is more robust than TCP in handling high BERs, disconnections due to user roaming, blackouts, etc. However, M-TCP requires a LL protocol to recover from losses in the wireless link.

3.4.1. New transport layer protocols

The most significant representatives of this category are the following:

- *Wireless transmission control protocol* (WTCP):- WTCP [15] is designed to provide a reliable transport in low bandwidth and high latency WWANs when a mobile host needs to connect through a proxy ('split connection' fashion). Its main characteristics are:
  1. WTCP is rate-based with the rate control performed at the receiver. The receiver communicates through cumulative ACKs the appropriate transmission rate to the sender.

2. WTCP attempts to predict when a segment loss is due to transmission errors or to congestion and signals the sender to continue transmitting with the same rate if the loss is estimated to be due to transmission errors.
3. In order to assure reliability WTCP employs a scheme with SACKs and probes instead of using an ARQ scheme [15].

The main disadvantage of WTCP is that the receiver is considerably more complex than in traditional TCP. This could lead to increased power consumption, since usually the mobile host plays the role of the receiver.

- *TCP westwood* [16] introduces sender-side only modification. The key innovative idea is to continuously estimate, at the TCP sender side, the packet rate of the connection by monitoring the ACK reception rate. The estimated connection rate is then used to compute congestion window and slow start threshold to be set after a congestion episode. This makes the protocol more robust to sporadic losses. Experimental studies show significant improvements in throughput performance over NewReno and SACK, particularly in mixed wired/wireless networks over high-speed links.
- TCP peach [17] was developed for communication scenarios where long RTTs and/or lossy links were involved. The sender transmits low priority packets called dummy segments to probe the availability of network resources in the end-to-end path and uses their ACKs (if any) to set the congestion window. TCP peach is an end-to-end solution but priority mechanisms are required in the intermediate routers.

Studying the literature on TCP in wireless networks, it is clear that each of the proposed solutions has characteristics which best suit a given environment. Next generation wireless networks, which will support the ABC concept, will integrate heterogeneous wireless communication environments with dramatically different characteristics. Therefore, in ABC systems the transport protocol will be reconfigurable to best adapt itself to the current environment. Moreover, note that LL solutions can coexist with TCP modifications as well as new transport layer protocols. Performance results show that LL solutions can completely solve the problems due to the high BER. However, they require modifications to be introduced in the wireless access provider infrastructure. Moreover, such solutions do not cope with long delays. Accordingly, LL solutions, if available, can be combined with others dealing with the problem of long delays, for example TCP-peach [17] if priority mechanisms are supported in the end-to-end path, or WTCP [15] if a proxy is available. Otherwise, if LL solutions are not available, then more aggressive protocols are required. As an example, M-TCP [14] can be used if a proxy is available, or TCP-westwood [16] if the end-to-end semantic must be guaranteed and modifications can be introduced only in the end terminals.

3.5. Middleware

Middleware can be defined as a reusable, expandable set of services and functions that are commonly needed by many applications to function well in a heterogeneous network environment. The above phrasing could further be refined to include persistent services, such as those found within an operating system, distributed operating environments (e.g. JAVA/JINI), the network infrastructure (e.g. DNS) and transient capabilities (e.g. run time support and libraries) required to support client software on systems and hosts. In any case, it can have different meaning to different network professionals.

Middleware is particularly useful in heterogeneous environments. Mobile, pervasive applications, delivered over highly diverse contexts, present challenging problems to designers. Devices face temporary and unannounced loss of network connectivity when they move, while they are likely to have scarce resources, such as low battery power, slow CPUs and little memory. They are required to react to frequent changes in the environment, such as change of location or context conditions, variability of network bandwidth, which will remain by orders of magnitude lower than in fixed networks.

When developing distributed applications, designers do not have to deal explicitly with problems related to distribution, such as heterogeneity, scalability, resource sharing and the like. Middleware developed upon network operating systems provides application designers with a higher level of abstraction, hiding the complexity introduced by distribution. Existing middleware technologies, such as transaction-oriented, message-oriented or object-oriented middleware have been created, trying to hide distribution as much as possible, so that the system appears as a single integrated computing facility. The interaction primitives, such as distributed transactions, object requests or remote procedure calls, assume a stable and constant connection between components. In mobile systems, unreachability is the norm rather than the exception. On the other hand, synchronous point-to-point communication supported by object-oriented middleware systems, such as CORBA, requires the client asking for a service, and the server delivering that service, to be up and running simultaneously. In a mobile environment, it is often the case that client and server hosts are not connected at the same time, because of intended disconnections (e.g. to save battery power) or forced disconnections (e.g. no network coverage). Finally, traditional distributed systems assume a stationary execution environment, characterized by stable and high bandwidth, and fixed location for every hosts. Recent developments in object-oriented middleware have introduced asynchronous primitives in order to allow a more flexible use, which could be a better choice in mobile scenarios.

In mobility-enabled systems, look-up service components are used to hide service location in order to allow reconfiguration with minimal disruption. In mobile environments, where the location of a device changes continuously, and connectivity fluctuates, service and host discovery becomes even more essential, and information on where the services are might have to reach the application layer. While in stationary systems it is reasonable to completely hide context information (e.g. location) and implementation details from the application, in mobile settings it becomes both more difficult and less beneficial. By providing transparency, the middleware must take decisions on behalf of the application. In constrained and dynamic settings, however, such as mobile ones, applications can make more efficient and better quality decisions based on application-specific information.

In order to cope with these limitations, many research efforts have focused on designing new middleware systems capable of supporting the requirements imposed by mobility. As a result of these efforts, a pool of mobile middleware systems has been produced [18]. It is notable that most of these approaches do not conceptualise middleware as hierarchical, or strictly layered, since this approach has been sometimes proven problematic and unproductive. Middleware can be better considered as a collection of components (such as resources and services) that is to some extent unstructured, often orthogonal that could be utilized either individually or in various subsets. This assumption enables work and study on various middleware issues to proceed independently and yield clearer results.

4. Mobility Support

One of the key attributes of the ABC concept is the capability to support users with an appropriate end-to-end QoS. To fulfil such a requirement is not an easy task. The problem is even more difficult in the ABC architecture where users are able to change their location as well as the network technology used, while they are in communication. The ABC concept contains the idea of ubiquitous connectivity at any time and any place. To achieve this goal, the underlying assumption is that the 'always connected' user is not hindered by geographic or movement restrictions. The users, and their connecting devices, are allowed to move freely either on foot or by other means (car, train, ship etc.) and still maintain the best level of connectivity possible. Mobility support is inherent for any ABC architecture.

The first level of mobility support focuses on the infrastructure design. The mobile access networks usually consist of geographically dispersed base stations, connected in a hierarchical fashion that allows the mobile device to connect successively to neighboring base stations as it moves. This is the model that all current cellular networks employ both telco-oriented (GSM, GPRS, UMTS) or internet-focused (mobile IP [19]).

As the internet technology penetrates more and more into every connectivity aspect in the research community, there has been much work done in optimising mobility support for internet-enabled devices. One of the first observations was that the mobile IP standard was not suitable for high-mobility, small geographic areas circumstances. To handle the needs for such applications, the so-called micro-mobility protocols evolved (Cellular IP, HAWAII etc.). These protocols operate within an administrative domain to achieve optimum mobility support for fast moving users within the domain's boundaries. Most micro-mobility protocols establish and maintain soft-state host-specific routes in the micro-mobility enhanced routers. However, the inter-domain mobility support is left to standard mobile IP.

Other approaches follow a different path, more closely coupled to the internet philosophy. They are not altering the routing tables for each moving user, but rely on the mobile to take care of the burden of different routing infrastructure. They use (as mobile IP does) tunneling between mobility endpoints. The Regional Registrations approach [20] follows this paradigm, and is actively developed in the IETF mobile IP working group. Several compromises had to be made in the first mobile IP design, because of the legacy IPv4 node support issues. Security was specified as an add-on, and interaction with other IPv4 nodes had to travel through the mobility agent at the home network of the mobile. The IPv6 protocol design took input from these drawbacks and provides the necessary mobility interaction functionality in every IPv6 node [21]. Moreover, mobile IPv6 [22] provides tighter integrated security and authentication options, since it reuses the mandatory functionality imposed by the core IPv6 protocol. On top of that, a hierarchical solution exploiting local mobility characteristics has already been defined, the hierarchical mobile IP protocol [23]. The diverse approaches are consolidating now to the localized mobility management architecture [36].

Mobility management has also been researched from a different angle. Specifically, the mobility support functionality is proposed to be included in higher layers, such as transport or even application layer. The argument to that kind of schemes is that the management of mobile hosts in an end-to-end fashion would simplify the infrastructure necessary for dealing with mobile hosts. Therefore, various higher layer solutions are available in the research literature, trying to tackle the mobility issue from a different angle. The TCP migrate extension [24] adds mobility support to TCP sessions. Similarly, mobile SCTP [25] builds upon the features offered by the SCTP transport protocol to offer transport layer mobility. In higher layers, the best-known scheme utilizes SIP [26] to achieve mobility management. In this approach, the SIP infrastructure is reused for mobility purposes (Registrar, Redirect Server).

The common factor in these approaches, though, is that they apply to specific protocols and applications and do not cover the full spectrum of internet applications. Some upper layer mechanisms support reliable transport, while others support real-time traffic, but usually not both. If the mobility mechanisms are not built for the least common denominator (the internet protocol layer) they are bound to exclude some types of applications and users.

Nowadays, it is relatively standardized for a mobile device to roam seamlessly through the infrastructure of a single provider, i.e., using its base stations, accounting services and other facilities. The ABC concept however supports the use of the fittest existing infrastructure at any time. According to the ABC idea, the user is aware of the multiple surrounding mobility support infrastructures and can choose to be serviced from the optimum at any time (possibly judging from multiple parameters, such as cost, bandwidth, technology capabilities etc.) Multiple connections may be initiated and deployed for the final result. The connections can stem from the same mobile device to multiple access networks or even from cooperating mobile devices to different access network technologies. Such multi-homing capabilities must be built into the operating system of the terminal handling the devices, and are still in research stages. Examples of available infrastructure alternatives include the base stations of another network provider (of the same technology), the base station in an access network of a different technology (e.g. WLAN) or even connectivity shared in an ad-hoc manner from a peer mobile terminal.

To achieve that ubiquitous connectivity, however, several problems must be overcome, most of which are not technical in nature. Whereas the multi-homing capabilities of modern terminals are expected to mature in the near future, roaming between different providers of the same service is only possible after a sound business model for all the players emerges. The technological capabilities are nevertheless huge and are currently being researched into prototypes and research demonstrations.

5. End-to-End QoS Support

The need for efficient support of real-time services is the major drive behind research efforts for enhancing internet with appropriate end-to-end QoS support. However, the QoS concept is still ambiguous, including a large variety of network quality aspects. There are, though, some common elements identified that are thought to be common among diverse QoS interpretations. These are per hop packet processing characteristics (router functionality to differentiate packet treatment and to utilize underlying links), the necessary signaling and the respective accounting of the service offered.

The internet community soon realized the vision of end-to-end QoS services and introduced the Integrated Services (IntServ) architecture [27] to implement this vision into specifications. IntServ supports end-to-end signaling, QoS state establishment and management for per-flow differentiated treatment in intermediate routers along the data path. The signaling protocol designed to meet the integrated services requirements is the RSVP (Resource reSerVation Protocol) [28]. However, the IntServ architecture and the RSVP received a lot of criticism, mainly due to the state maintenance for every data flow in intermediate routers across the end-to-end path. To minimize the state space needed for RSVP, the RSVP aggregation signaling was proposed [29]. Note here that RSVP was not designed to support handovers and thus its interworking with mobility schemes is quite poor.

As an alternative to RSVP, engineers shifted their target to a lightweight QoS architecture putting as little burden in the routers as possible and providing coarse-grained traffic prioritization. The outcome was Differentiated Services (DiffServ) [30]. DiffServ networks support only a small set of QoS levels (PHBs, per hop behaviors), perform packet classification according to a 6-bit field in the IP header (DSCP, DiffServ code point) and do not use QoS signaling for QoS state establishment and maintenance in routers. This coarse-grained traffic prioritization had also some disadvantages thus, several techniques have been proposed for the interworking of RSVP, deployed in the access network and DiffServ, deployed in the core network [31]. In terms of mobility support, DiffServ does not provide any means. This is especially true in the case of statically configured trunk reservations.

In light of heterogeneous QoS techniques flourishing and being deployed in different situations and needs, the end-to-end QoS framework needed to be re-evaluated. The paradigm of routing protocols classification into inter-domain (e.g. OSPF, open shortest path first) and intra-domain (e.g. BGP, border gateway protocol) protocols, lends itself naturally to a similar classification for QoS frameworks and signaling protocols. Thus, a two-tier resource management model was proposed [32], with the lower-tier QoS signaling performing resource management inside a domain, and the upper-tier one managing resource allocation between domains. The two tiers must be closely coordinated in order for the network to provide the necessary end-to-end QoS support. The two-tier model increases the degrees of freedom regarding end-to-end QoS support, since each domain is free to choose any QoS support mechanism for allocating resources internally, as long as proper co-operation takes place with the respective inter-domain signaling protocol. The two-tier signaling architecture implies that each domain is allowed to use its own QoS mechanism or protocol internally, allowing for concatenation of various heterogeneous domains. However, at the domain boundaries appropriate mapping should take place between the intra- and the inter-domain signaling QoS parameters, which introduces complexity.

Intra- and inter-domain signaling can either follow the same path with the subsequent data flow (path-coupled signaling), or follow a different route (path-decoupled signaling). In case of path-coupled signaling, QoS parameter mapping, admission control and resource management for each domain take place in a distributed fashion by enhanced edge (border) routers situated at the domain boundaries. Two such examples of inter-domain path coupled signaling are presented in [33,34].

The first proposal, referred to as border gateway reservation protocol (BGRP), operates end-to-end only between domain border routers. BGRP mainly aims at aggregating reservations between domains and improving scalability. BGRP performs reservation aggregation by building a sink tree for each destination domain. Reservations from different initiating domains belonging to the same sink tree are aggregated along the path to the destination domain.

The second protocol is designed for supporting end-to-end QoS through several DiffServ domains. The protocol is called DiffServ PHB reservation protocol (DPRP) and is a modified version of RSVP that enables transport and negotiation of the QoS requirements (in terms of DSCPs) between source and destination, as well as reservation of resources inside the respective QoS level for each data flow. DPRP is implemented only in DiffServ domain edge routers, where it stores per-QoS level soft states. Note that both proposals however do not cater for terminal mobility, since they were designed to support solely end-to-end QoS support.

In comparison to path coupled signaling, the path-decoupled signaling is strongly related with domain architectures where the resources of the domain are managed by one or more entities that are not necessarily situated on the data path. Instead, they can be located in central points inside the domain and perform QoS parameter mapping functions, admission control functions and resource management functions for the domain. Among these architectures, the most representative is the bandwidth broker (BB) architecture [35], where each domain avails a BB being responsible for intra- and inter-domain dynamic resource provisioning and admission control management.

Various end-to-end protocols have been designed that allocate resources between neighboring domains. In addition, edge-to-edge protocols for allocating resources inside a single domain have also been proposed. No general consensus exists up till now in the research community for the prevalent QoS protocol amongst the proposed ones. The suitability of a specific QoS protocol seems to be dependent on network specific parameters. However, a critical factor in a QoS protocol's efficiency seems to be the flexible balancing between reservation granularity and aggregation. Moreover, the ability of the QoS protocols to cope well with the user's mobility and security issues is an important protocol evaluation factor that only lately is being seriously considered in the protocols under design.

6. Case Study

In this section, we present and analyze two user scenarios in terms of network actions, to better describe the functionality and effectiveness of a system integrating the aforementioned enabling technologies. The scenarios are presented in the form of user action and respective network reaction.

6.1. Professor's Case

Brian is a University Professor, who works both at home and in the University. Most days, he goes to the University for lectures, meetings, library work etc. The scenario below focuses on possible communication requirements and solutions meeting Brian's needs in his attendance to the University. The following equipment is utilized/deployed:

- *In Brian's possession*:
  - UMTS cell phone with Bluetooth card;
  - laptop with IEEE 802.11a, IEEE 802.11b, UWB and Bluetooth cards;
  - car with intelligent ergonomic vehicle communication terminal system (VCTS).
- *In the University Campus*:
  - IEEE 802.11a/b islands;
  - WLAN/Bluetooth gateways;
  - Information Stations (InfoStations) at the University gateways (input/output points), used for fast network application updates, such as web caching and email downloads, providing fast wireless network access for short periods of time. InfoStations utilize specific transmission radio technologies, such as Ultra WideBand (UWB), to attain high transmission speeds for short ranges and periods of time;
  - SMS gateway;
  - satellite gateway.
- *In Brian's office*:
  - WLAN (IEEE 802.11a/b) access point providing access to the wired (fixed) campus network (intranet) and internet;
  - Bluetooth (or UWB) connection between all office devices;
  - WLAN/bluetooth internetworking gateway.

User action	Network reaction
Arriving by car, Brian approaches a University gateway, deployed with an InfoStation. While Brian is passing through, his laptop (always-on in a semi-sleeping mode) and the InfoStation detect each other's presence and initiate communication. The laptop immediately starts downloading user specific information, such as urgent messages (e.g. about meetings rescheduled for this day), and general information, such as local announcements etc. Messages may be played by voice one after another over the vehicle's VCTS, which will contain a text to voice system. For example, the first information Brian may desire is the location of free car park spaces (sorted in order of suitability to his office—information which is taken from Brian's profile). By voice, he may also indicate (overriding the profile information) the car park he would prefer and receive oral instructions, and screen feedback, on the location within the chosen car park of the free spaces. In this way, he can drive directly to the suitable car park space.	Strict delay constraints (<10 s) must be enforced. The best connection is obtained by establishing an UWB connection between the laptop and the InfoStation for the time of passing through (Fig 2). Laptop core profiles will have the capability of reconfiguring overall laptop profile as a function of user location, and a variety of user defined and network defined variables. The car parking assistance service will require the support of fine-grained location service if it is to meet Brian's needs to be directed to a car park space. The laptop will establish an IP connection with an intelligent VCTS via Bluetooth or other standard. This will be especially ergonomically designed for safe visual and oral communication with the driver (as well as other passengers). It could also be the source of the location-based information for the vehicle and other devices and have (reconfigurable) network access technology for communications with vehicle support services (traffic patterns, maps etc.).

Continues

(Continued)

User action	Network reaction
One of the messages Brian receives is an urgent notification about a meeting (with the University rector) rescheduled for this day. After parking his car, Brian checks his on-line diary and sees that at the same time with the meeting he has a lecture (alternatively, an intelligent diary automatically announces this to him). While in the car Brian urgently needs to broadcast a message to the entire student class about canceling/postponing the lecture. He types and sends an SMS on his mobile phone. What Brian may not realize is that communications between his mobile phone and laptop have established that the quickest and cheapest way to handle this SMS message to the University's SMS gateway is through the laptop to an InfoStation or a car park WLAN access point. All registered users (professors and students) have profiles at the SMS gateway containing (among other things) information about the best way of forwarding urgent messages to them at any particular moment, for example, by SMS, email, fax, voice mail or otherwise.	No strict delay constraints must be enforced. A Bluetooth connection between the UMTS cell phone and the laptop is established. Both will know other's presence and both will be 'up to date' in their awareness of the other's network(s) accesses (characteristics, QoS, cost etc.). Laptop's IEEE 802.11a/b card 'senses' the presence of a wireless island in the car park, with its access to the SMS gateway via the campus network, and this information is automatically conveyed to the mobile phone as an alternative, reliable, cheap SMS service access. To support the mobile phone's request for SMS service, IP session is established between the laptop and the SMS gateway to deliver the SMS received from the mobile phone via a Bluetooth. SMS is received by the SMS gateway. To the students already on the campus, SMS is forwarded to their laptops/mobile phones over the campus network. The SMS is broadcasted to all WLAN/Bluetooth gateways on the campus network. Each WLAN/Bluetooth gateway forwards the SMS only to the laptops/mobile phones of the intended recipients in its own picocell. Others (not currently present at the campus but having mobile phones) will receive the SMS via the UMTS network. The rest of the students (without mobile phones) will receive email/fax/voice mail message, as currently specified in their user profile. To support this, the SMS gateway converts the SMS into an email/fax/voice mail and sends it over the campus network/internet/PSTN.
While Brian is walking to his office through campus WLAN islands, all new emails (or new ones meeting his campus-profile filter conditions) are automatically being downloaded to his laptop.	Connection is established between the laptop and the nearest wireless LAN access point with seamlessly handover/roaming from one access point to another. Brian's movement is supported by micromobility protocols implemented both in the Campus WLAN and Brian's laptop (e.g. Cellular IP), to offer continuous connectivity.
Brian walks into his office where all devices are connected via Bluetooth (or UWB) to each other. Automatic update and synchronization starts between his laptop and office PC. While in the office, Brian initiates a videoconference (over the campus network) with some of his colleagues, in order to discuss and prepare for the new issues included in the agenda of that day's meeting.	Bluetooth (or UWB) connection is established between the laptop and office PC (or other office device). A videoconference is organized between PCs/laptops of the users. For any colleague not possessing a multi-media PC/laptop, a phone connection is initiated. The connection decision is dictated by the network (or the videoconference server), based on the user-profile of that colleague. Network awareness of the location of a colleague (through up-to-date location-based information about him/her) will dictate the connection decision to be made to that colleague for video conference call from among the possibilities and types of connection available.
During the meeting, Brian has to call his colleague to clarify some issues. Brian uses his UMTS cell phone. Depending on his colleague's location, the network can provide the most suitable connection (in terms of cost, availability or user's profile): ● Using VoIP (cheapest option with worst QoS). If the colleague is in the campus (with a laptop and mobile phone, or currently working on multi-media PC in the lab without a phone) the call is made without a charge. ● Using UMTS. If the colleague is outside the campus, the connection can be established through the UMTS networks, charged by the network operator.	Based on Brian's profile, an intelligent agent in the network decides on the best available connection. ● In the first case, the following connection is established: UMTS phone—laptop—campus network—laptop—mobile phone—multimedia PC. ● In the second case, the following connection is established: UMTS phone—UMTS network—UMTS phone.

6.2. Student's Case

Alice is a Ph.D. student, who is taking the opportunity of the reading week to go back to her home city for 4 days. She is carrying her laptop computer with her, so that she can pass the journey time more entertainingly. Alice's mode of travel is by train. The following equipment is utilized/deployed:

- *Alice*:
  - UMTS cell phone with Bluetooth card,
  - a powerful laptop with big screen and IEEE 802.11b WiFi card and Bluetooth connector on-board.
- *Train station*:
  - IEEE 802.11b WiFi islands are deployed in the train stations offering connectivity to the internet.
- *Train*:
  - A Bluetooth access point responsible for managing a vehicular area network (VAN) inside every wagon.

- Through the VAN, it is possible to obtain interconnection with the external internet infrastructure and to access the following services provided by the train operator:
  - tourist information about the localities through the travel path,
  - travel planned and actual schedules,
  - train timetables,
  - order meals and drinks from the train or upcoming station restaurants.
- A satellite station acting as gateway between the VAN and the overall network infrastructure for internet access.

pected to contribute in converting this vision to reality. From the above discussion, it is clear that a number of extensions to today's networks are required, affecting most of the layers of a traditional protocol stack, in order to introduce the required functionality. This functionality, focuses mostly on adding a considerable degree of flexibility to the network for adjusting to different 'conditions', in terms of traffic, transmission quality, user preferences, available tariffs etc. The next big challenge will be to integrate these technologies in a single network architecture, which has the intelligence to perform the required adjustments.

User action	Network reaction
Alice is in the train station cafeteria with her laptop. She wants to check the timetable again for possible delays.	The IEEE 802.11b card 'senses' the WLAN of the train station and the Bluetooth connection with the mobile phone. In her profile, Alice has given priority to WLAN for internet access, which is cheaper than UMTS. Her HTTP browser is tuned to the train station home page, where she can retrieve the info she requests.
The mobile network operator pushes a message to Alice's mobile phone, announcing the availability of new patch software for her mobile's device OS. Alice decides to update her mobile.	Using the station WLAN gateway to the internet Alice downloads the patch to her laptop computer. Then using the graphical interface of her mobile and the Bluetooth connection, finds the file stored in the laptop's hard drive and upgrades her software.
Alice is sitting in the train wagon when she is notified about the satellite gateway the train is equipped with and about the cost of using it.	Both laptop and the mobile sense the Bluetooth access point and are connected to the VAN network. The van network provides Alice laptop with information about the available satellite gateway but she decides not to use the service due to cost constrains.
While waiting for departure, Alice chats with her instant messenger application with three friends of hers, who are online and decide to play a real-time strategy multiplayer game.	Alice is still using the WLAN gateway of the station. The most important QoS requirements are delay, which must not exceed 300 msec and an IP address that does not change during the game. Alice however is able to play using mobile IP. The game application auto-configures itself to advertise the Home Agent the core-of address of Alice's laptop.
The train departs and Alice gets out of the WLAN range. She then moves through UMTS cells.	The laptop automatically uses the UMTS of the mobile phone through Bluetooth as its internet gateway. A vertical handover occurs during the game but the application is tolerant of small packet loss. A small pause in the game happens but is acceptable. During cell changes, horizontal handovers occur with seamless impact to the game.
The train is not express so it makes a stop at a number of stations equipped with WLAN access points.	Whenever the WLAN card senses an available connection, the laptop switches to the WLAN network which is faster and cheaper. This is done in accordance with the QoS profile of Alice.
The train is moving into an area with no UMTS coverage.	The mobile phone of Alice flashes and makes a sound that it is moving out of coverage. Alice informs the other players of the problem but thanks to the warning they have time to save the game and continue whenever Alice is back online.
The game ends and Alice decides to print some post-game statistics. She uses the printer in the cafeteria computer of the train for only a small fee.	The laptop connected to the Bluetooth VAN network has been informed through proper discovery protocols of the availability of the printer.

7. Conclusions

Moving from 'Always Connected' to 'Always Best Connected' is considered critical for next generation networks. In this paper, we briefly presented a considerable set of enabling technologies that are ex-

Acknowledgements

This work has been produced in the framework of the project 'ANWIRE' (IST-2002-38835), which is funded by the European Community. The authors would like to acknowledge the contributions of their colleagues from University of Athens, King's College London, University of Helsinki, Rheinisch-Westfaelische Technische Hochschule Aachen, Universita Degli Studi di Catania, Universidad Politécnnica de Madrid, Instituto Superior Técnico, Universite Pierre et Marie Curie, University of Cyprus, University of Limerick, Ecole Nationale Supérieure des Télécommunications, NEC Europe Ltd., Thales Communications, Thales Research Limited, University of Surrey, Centre Tecnològic de Telecomunicacions de Catalunya.

References

1. Gustafsson E, Jonsson A. Always Best Connected. *IEEE Wireless Communications Magazine* 2003; **10**(1): 49–55.
2. Campbell AT, Gomez J, Kim S, Valkó AG, Wan C-Y, Turányi ZR. Design, implementation, and evaluation of cellular IP. *IEEE Personal Communications Magazine* 2000; **7**(4): 42–49.
3. Passas N. Internet access over wireless LANs. In *Mobile Internet: Advanced Technologies and Services*. CRC Press: ISBN: 0849316316, scheduled for March 2004.
4. Anton-Haro C, Lagunas MA. Array processing techniques for wireless: a cross-layer perspective. *International Forum on Future Mobile Telecommunications and China-EU Post Conference on Beyond 3G*, Beijing, China, November 2002.
5. Bartolomé D. Reconfigurable receivers on the basis of channel classification. (GIRAFA—Spanish project, under the contract TIC2002-04594-C02-02), http://www.cttc.es/training/index.html#ws.
6. Realp M, Pérez-Neira AI. PHY-MAC dialogue with multi-packet reception. *ETSI Workshop on Broadband Wireless Ad-Hoc Networks and Services*, Sophia Antipolis, France, September 2002.
7. Farnham T, Clemo G, Haines R, Seidel E, Benamar A, Billington S, Greco N, Drew N, Le TH, Arram B, Mangold P. A perspective on the reconfiguration of future mobile terminals using software download. IST-TRUST project, available at http://www.ist-trust.org/Publications/PIMRC_WP4_Paper-FINAL.pdf
8. Mähönen P, Passas N, Orphanos G, Munoz L, Marshall A, Melpignano D, Inzerelli T, Lucas F, Vitiello M, Garcia M, Saarinen T. Platform-independent IP transmission over wireless networks: the WINE approach. *IEEE Personal Communications Magazine* 2001; **8**(6).
9. Border J, Kojo M, Griner J, Montenegro G, Shelby Z. Performance enhancing proxies. Internet draft (work in progress), ⟨draft-ietf-pilc-pep-04.txt⟩, October 2000.
10. Balakrishnan H, Seshan S, Katz RH. Improving reliable transport and handoff performance in cellular wireless networks. *Wireless Networks* 1995; **1**(4): 469–481.
11. Parsa C, Garcia-Luna-Aceves JJ. Improving TCP performance over wireless networks at the link layer. *Mobile Networks and Applications* 2000; **5**(1): 57–71.
12. Mathis M, Mahdavi J, Floyd S, Romanow A. TCP selective acknowledgment options, RFC 2018, April 1996.
13. Bakre A, Badrinath BR. I-TCP: Indirect TCP for mobile hosts. *Proceedings of 15th Internal Conference Distributed Computing Systems (IDCS)*, May 1995.
14. Brown K, Singh S. M-TCP: TCP for mobile cellular networks. *Computer Communication Review* 1997; **27**(5): 19–43.
15. Sinha P, Venkitaraman N, Sivakumar R, Bharghauan V. WTCP: a reliable transport protocol for wireless wide-area networks. *Proceedings of ACM MOBICOM '99*, Seattle, WA, August 1999.
16. Mascolo S, Casetti C, Gerla M, Sanadidi M, Wang R. TCP westwood: bandwidth estimation for enhanced transport over wireless links. *Proceedings of ACM Mobicom 2001*, July 16–21 2001.
17. Akyildiz IF, Morabito G, Palazzo S. TCP-peach: a new congestion control for satellite IP networks. *IEEE/ACM Transactions on Networking*, 2001; **9**(3): 307–321.
18. Aiken B, Strassner J, Carpenter B, Foster I, Lynch C, Mambretti J, Moore R, Teitelbaum B. *Network Policy and Services: A Report of a Workshop on Middleware, RFC 2768*, February 2000.
19. Perkins C. (ed.). RFC 3344, "IP Mobility Support for IPv4", August 2002.
20. Gustaffson E, Jonsson A, Perkins C. Mobile IPv4 regional registration. Internet draft, work in progress, October 2002, draft-ietf-mobileip-reg-tunnel-07.txt.
21. Deering S, Hinden R. RFC 2460, Internet protocol version 6 (IPv6) specification. December 1998.
22. Johnson D, Perkins C, Arkko J. Mobility support in IPv6. Internet draft, work in progress, February 2003, draft-ietf-mobileip-ipv6-21.txt
23. Soliman H, El-Malki K, Bellier L. Hierarchical mobile IPv6 mobility management (HMIPv6). Internet draft, work in progress, October 2002, draft-ietf-mobileip-hmipv6-07.txt
24. Snoeren A, Balakrishnan H. An end-to-end approach to host mobility. In Proceedings of ACM Mobicom, August 2000.
25. Riegel M, Tuexen M. Mobile SCTP. Internet draft, work in progress, February 2003, draft-riegel-tuexen-mobile-sctp-02.txt.
26. Schulzrinne H, Wedlund E. Application-layer mobility using SIP. *Mobile Computing and Communications Review (MC2R)* 2000; **4**(3): 47–57.
27. Braden R, Clark D, Shenker S. Integrated services in the internet architecture: an overview. IETF RFC 1633, June 1994.
28. Braden R, Zhang L, Berson S, Herzog S, Jamin S. Resource reservation protocol (RSVP). IETF RFC 2205, September 1997.
29. Baker F, Iturralde C, Le Faucher F, Davie B. Aggregation of RSVP for IPv4 and IPv6 Reservations. IETF RFC 3175, 2001.
30. Blake S, Black D, Carlson M, Davies E, Wang Z, Weiss W. An architecture for differentiated services. IETF RFC 2475, December 1998.
31. Bernet Y, Ford P, Yavatkar R, Baker F, Zhang L, Speer M, Braden R, Davie B, Wroclawski J, Felstaine E. A Framework for integrated services operation over DiffServ networks. RFC 2998, November 2000.
32. Terzis A, Wang L, Ogawa J, Zhang L. A two-tier resource management model for the internet. Global Internet, December 1999.
33. Pan P, Hahne E, Schulzrinne H. BGRP: a tree-based aggregation protocol for inter-domain reservations. *Journal of Communications and Networks* 2002; **2**(2): 157–167.
34. Qian W, Jing W, Shiduan C, Jian M. A PHB reservation mechanism in DiffServ network. *International Conference on Communication Technology*, WCC-ICCT 2000.
35. Nichols K, Jacobson V, Zhang L. A two-bit differentiated services architecture for the internet. IETF RFC2638, July 1999.
36. Williams C. (ed.). Localized mobility management requirements. Internet draft, work in progress, March 2003, draft-ietf-mobileip-lmm-requirements-03.txt.

Authors' Biographies

Nikos Passas received his Diploma (honors) from the Department of Computer Engineering, University of Patras, Greece, and his Ph.D. from the Department of Informatics and Telecommunications, University of Athens, Greece, in 1992 and 1997, respectively. From 1992 to 1995, he was a research engineer at the Greek National Research Center 'Democritus'. Since 1995, he has been with the Communication Networks Laboratory of the University of Athens, working as a sessional lecturer and research associate in a number of national and European research projects in the ACTS and IST frameworks. He has also served as a guest editor and technical program committee member in prestigious magazines and conferences, such as IEEE Wireless Communications Magazine, Wireless Communications and Mobile Computing Journal, IEEE Vehicular Technology Conference, IEEE Globecom etc. Dr Passas has more than 30 papers published in peer-reviewed journals and international conferences and has also two book chapters published. His research interests are in the protocol design and performance analysis for mobile multimedia communications. He is particularly interested in QoS for wireless networks, mobility management, mobile network architectures and protocols etc. Dr. Passas is a member of the IEEE and a member of the Technical Chamber of Greece.

Sarantis Paskalis received his B.Sc. degree (with honors) in Computer Science from the Department of Informatics at the University of Athens in 1995 and his M.Sc. (with honors) in Computer Science (Communication Systems and Networks) from the same department in 1998. He is currently a Ph.D. candidate, working in the area of services for wireless IP networks). He is a staff member of the Communication Networks Laboratory of the University of Athens. His research interests include mobility management, QoS issues, protocol performance analysis, network security as well as design and evaluation for wireless networks.

Dr Alexandros Kaloxylos received his B.Sc., from the Department of Computer Science at the University of Crete, Greece in 1993, his M.Phil. from the Department of Computing and Electrical Engineering at the Heriot-Watt University in Edinburgh, Scotland in 1994 and his Ph.D. from the Department of Informatics and Telecommunications at the University of Athens in 1999. During 1990–1993, he was a staff member of the Computer Centre of the University of Crete, and a researcher in the network team of the Foundation of Research and Technology Hellas (FORTH). During 1994–1995, he worked as a research associate at the University of Wales. From 1995 till today, he has been a senior researcher in the Communications Network Laboratory of the University of Athens. Currently, he is an assistant professor in the University of Peloponnesus. He has participated in numerous projects realised in the context of EU programs as well as National Initiatives. His research interests include mobile and wireless networks, broadband networks and formal description techniques.

Faouzi C. Bader graduated in Electronic Engineering in Communication Area from the Faculty of Engineer Sciences of the University Mentouri of Constantine in 1996 (Constantine, Algeria), and has completed his Ph. D. in November 2002 at ETSI of Telecommunication of the Polytechnic University of Madrid (UPM) in Spain. Since December 2002, he is an associate researcher at the CTTC's centre in Access Technologies Area. Prior to joining the CTTC center, he was a member of the Applied Signal Processing Group (GAPS) (1998–2000) at the ETSI of Telecommunication where he developed the hexagonal pilot pattern and the MUD detection for the MC-CDMA system in the uplink transmission mode. He has also worked as a development engineer (2000–2001) in Massana Technologies Company (Spain), where his work involved the development of the *Gigabit* Network prototype. His primary research interests areas include the Multi-Carrier Spread Spectrum (MC-SS), OFDM, MC-CDMA UTRA-TDD/FDD, ADSL, VHSL and UMTS systems, channel estimation, multiple access strategies, synchronization and multiple user detection (MUD).

Renato Narcisi was born in Catania, Sicily (Italy) on 20 May 1975. He received the laureat degree in Electronic Engineering (spec. telecommunication) from the Istituto di Informatica e Telecomunicazioni, University of Catania, Catania, Italy, in 2001. From September 2001 to August 2002, he was with RSI Sistemi SPA (Altran Group) in Rome as a junior consultant. Since September 2002, he is with the Dipartimento di Ingegneria Informatica e delle Telecomunicazioni of the University of Catania. His research interests focus on wireless networks.

Evangelos Tsontsis received a combined B.Sc. and M.Sc. from the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, in September 2000. In 2002 he received his M.Sc. in Mobile and Personal Communications from King's College, London. In September 2002, he commenced his Ph.D. in the Centre for Telecommunications Research, King's College, London. He is currently a member of the European IST project ANWIRE and is doing research in Signaling Protocols (COPS), Mobility Issues, IntServ-DiffServ architecture and Policy Based Management Networks.

Adil S. Jahan received his B.Sc. (Hons) degree in Electronic Engineering from King's College, University of London in 1998. He continued his post-graduate studies in Centre for Telecommunications Research, King's College, London and is currently in the process of writing his Ph.D. thesis. As of September 2002, he has been employed at King's as a research associate working in the European IST-ANWIRE project. He is currently one of the workpackage leaders in the project, with research work in the area of adaptability, reconfigurability and Always Best Connected concepts.

Hamid Aghvami joined the academic staff at King's in 1984. In 1989, he was promoted as reader and in 1992 as professor in Telecommunications Engineering. He is presently the Director of the Centre for Telecommunications Research at King's. Professor Aghvami carries out consulting work on Digital Radio Communications Systems for both British and International companies. He has published over 300 technical papers and given invited talks all over the world on various aspects of Personal and Mobile Radio Communications as well as giving courses on the subject worldwide. He was visiting professor at NTT Radio Communication Systems Laboratories in 1990 and senior research fellow at BT Laboratories in 1998–1999. He is currently executive advisor to Wireless Facilities, Inc., U.S.A. and managing director of Wireless Multimedia Communications LTD. (his own consultancy company). He leads an active research team working on numerous mobile and personal communications projects for third and fourth generation systems, these projects are supported both by the government and industry. He is a distinguished lecturer and a member of the Board of Governors of the IEEE Communications Society. He has been member, chairman, vice-chairman of the technical program and organizing committees of a large number of international conferences. He is also founder of the International Conference on Personal Indoor and Mobile Radio Communications (PIMRC). He is a fellow of the Royal Academy of Engineering, fellow member of the IEE and senior member of the IEEE.