IP QOS AND MOBILITY EXPERIMENTATIONS WITHIN THE MIND TRIAL WORKPACKAGE

Jean-Christophe Rault [1], Louise Burness [2], Emilio Garcia [3], Tomas Robles [4], Jukka Manner [5], Nikos Georganopoulos [6], Pedro Ruiz [7]

[1] France Telecom R&D, 4, rue Clos Courtel, 35510 Cesson, France, jeanchristophe.rault@francetelecom.com
[2] BT Exact, Rigel House, B28/pp2b, Adastral Park, Martlesham, Ipswish IP35 3RE, UK, louise.burness@bt.com
[3] Agora Systems, S.A., c/Aravaca 12 3°B, 28040 Madrid, Spain, emilio.garcia@agoratechnologies.com
[4] UPM, DIT - ETSI Telecomunicacion, E-28040 MADRID Spain, robles@dit.upm.es
[5] University of Helsinki, CS dept, Teollisuuskatu 23, Helsinki, FIN00014, Finland, jukka.manner@cs.helsinki.fi
[6] Kings College London, Strand, London WC2R 2LS, UK, nikolaos.georganopoulos@kcl.ac.uk
[7] Agora Systems, S.A., c/Aravaca 12 3°B, 28040 Madrid, Spain, pedro.ruiz@agoratechnologies.com

Abstract - In addition to the conceptual work, the IST MIND project aims to practically verify a set of concepts related to the use of broadband, multimedia services by mobile users in an IP network. This paper presents an overview of the trials, and focuses on a few particular topics that are dedicated to the experimentation of adaptive applications within a wireless IP access network including advanced QoS and mobility protocols. The novelty of the experimentation is underlined, and first results are presented.

Keywords – QoS, Mobility, Trials, Multimedia application.

I. INTRODUCTION

The goal of the IST MIND project [1] trials work group is to practically implement, investigate and verify technical concepts that are required to support multimedia services in a mobile IP network. The multimedia applications should be fully supported and customized when accessed by users from a wide range of access technologies. In this type of situation, the most critical applications are real time ones, which call for a high degree of Quality of Service (QoS), as perceived by the user. The applications should behave the same way, even when accessing different types of networks.

The concepts being implemented were designed in the IST BRAIN project [2]. This project provided a wide set of concepts that were driven by the requirements of end-users. Because these concepts were developed together, it is expected that they will work together in an understandable, controlled and beneficial fashion. In keeping with IP networking principles, these concepts were developed with a strict understanding of the importance of correct layering and modularity. Thus it should be possible to implement and test each concept individually. However, few of the individual concepts have actually been implemented earlier, and no practical verification exists that any combination of the different elements will succeed in meeting user expectations.

A. Trials overview

As expressed in [3], it is easiest to consider the trial concepts as belonging to one of three categories: those that directly relate to the application layer, those that relate to the networking (IP) layer, and those that relate to the lower (link) layer.

At the application layer, we study the development of mechanisms to support QoS-enabled applications and application adaptation, utilising an API between the application layer and the IP layer.

At the networking layer we investigate both QoS mechanisms suitable for use in a wireless IP radio network with mobile terminals and the development of IP mobility management solutions. A key result is the implementation of a new state-of-the-art local mobility management protocol.

At the link layer, we have been examining the relationship between a Brain Access Network and public cellular networks (GPRS and/or UMTS). Cooperation between different networks is needed in order to be able to provide a continuous service to a mobile user. Further there is also implementation of a Hiperlan/2 physical layer. This is a very powerful radio interface, especially regarding broadband multimedia IP wireless communication, as it includes advanced QoS feature on the radio bearers.

The trials are being performed by the MIND partners mentioned in the acknowledgements section. In the first stage of the project, the experimentations have been set up at different premises across Europe. However, to demonstrate more significant concepts than individuals could have made, some cooperation has been initiated with the aim of suppressing the overlaps between the various stand-alone platforms. Cooperative development work has been undertaken in the early stage to favour the integration work, which is planned in the last phase of the project for demonstrating a complete user scenario.

B. Supporting User Requirements

Amongst the whole set of trials, a particular attention has been paid to the need for a top-down approach. A user who attends a videoconference session on his mobile terminal illustrates one of the most challenging issues, as explained hereunder.

To support a real-time videoconference we need a tight control of the delay and jitter of the audio and video packets between the terminals. This in turn produces a requirement for low delay and low jitter on the network, which must be communicated to the network layer from the application. The network layer in turn has to set up the right services on the wireless and wired links.

Furthermore when the user moves, the access network must provide support for mobility and keep delivering the requested QoS. The network and the terminal have to co-operate to re-establish points of attachment, and then update the packet path through the access network so that packets can still be delivered to the right place; all this must minimize any perceptible disruption as seen by the end user.

Wireless links have irregular performance, with variable levels of packet loss and delay. Further, the handover process may result a terminal attaching to an access point with a significantly different set of transmission characteristics. Thus, the application and the network protocols have to co-operate to adapt the behaviour of the application to the changing characteristics of the wireless link in use.

The following sections of the paper will tackle the topic of how integration of QoS aware applications with advanced QoS and mobility protocols is expected to achieve a seamless and continuous service, even with frequent handovers and alterations of the networking conditions.

II. SERVICE CONCEPTS: QOS-ENABLED APPLICATION DEVELOPMENT

Adaptability of applications and services to context conditions is a key issue for the successful application of the ‘always on’ concept in systems beyond 3G. In such systems, the applications and services have to deal with highly changing context conditions, regarding both users preferences and terminal and network characteristics and capabilities. In this environment, applications have to show a very flexible behaviour in order to react to the changing conditions in the best way, however this concept may be defined by the end-user. In this section we focus on supporting service concepts identified during BRAIN, specifically QoS and mobility issues.

The creation of this kind of flexible applications requires actualisation of several key elements:

- Applications must be network aware- that is, capable of reacting to the information coming from lower layers.
- Further, this information needs to be made available by the lower layers. Mobility protocols can be modified to provide upper layers with information regarding handovers.
- As identified in BRAIN, a middleware in form of a Control Entity is needed for the coordination of both elements, and also for enabling the management of various QoS properties.

A. Application

The application has to implement a flexible adaptation mechanism in order to react in real time to the changes in the context conditions. Several adaptation mechanisms may be implemented by the applications. For example, in a usual videoconference application, adaptation may be based on selection of the suitable media codecs, resizing buffers, or just dropping components.

Adaptation should be as instantaneous and transparent to the end user as possible, thus maintaining perceived quality with the minimum impact on the service. This adaptation will be strongly based on the management of precise and on time information about the context conditions, by the means of the suitable interfaces and feedback mechanisms in the network and in the applications.

B. Control entity

The control entity serves a clear purpose: decoupling the lower layers from the application itself. This way, and if the interfaces are properly designed, it is expected that the same control can provide support for distinct applications, while hiding the underlying the details of the network and its protocols. The control entity has two main functions:

- **Mobility Support.** Mobility issues, although decoupled from the services layer, present a hard influence on user perceived QoS, because of the disruption induced when a handover takes place. A suitable interface has been identified (indicating for example, handover imminent, handover complete) which, if implemented by the lower layer protocols, would enable the application to prepare itself, so that the effect of the drastic disruption is minimized (e.g. buffering some data, changing to lower bandwidth data rates or, in the worst case, simply preparing to renegotiate the QoS parameters for the new network).
- **QoS Support.** Quality of Service is understood as the collective effect of service performance, which determines the degree of satisfaction of a user of a service [4]. To this extent, at least two ways can be envisioned to achieve a better experience. Firstly there is the need to receive feedback from the network layer regarding the actual network conditions (such as available bandwidth, jitter or error rate), and secondly there is the need to be able to enforce or request some preferred settings.

The control entity provides a bi-directional communication between the applications and the lower layers. Primitives coming from lower layers are to be translated into a precise command indicating the application the adequate adaptation procedure to follow. The application, on its side, can enforce some parameters, which the control entity must inject in whatever QoS mechanism is implemented for supporting the adaptation.

In order to illustrate the benefits of such an adaptive approach, a videoconference application has been selected as the paradigm of the new bandwidth consuming real-time services. The conceptual model is shown in Fig. 1:

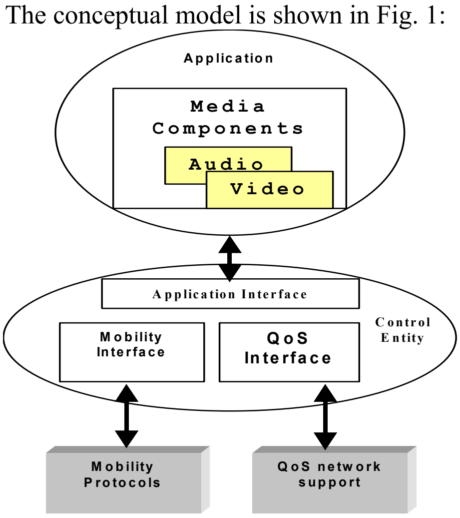

Fig. 1: Conceptual Model for service adaptation

The integration between mobility and QoS supporting layers with the application is done through interfaces in the control entity. This decouples the solution from the very specific service selected for these trials. Tests will be performed in order to evaluate the feasibility of the approach, evaluate and refine the proposed interfaces, evaluate the performance of several applications adaptation strategies well as the improvement over non-coupled version. Finally the trial will allow producing recommendations for applications, mobility protocols and control entities.

III. NETWORK CONCEPTS: QOS AND MICRO MOBILITY MECHANISMS

The access network must be able to support both seamless mobility (i.e. no significant packet delays or losses during handovers) and provide QoS for a wide variety of applications. The QoS mechanism must be able to provide explicit resource reservations (e.g. for real time services) as well as the more flexible priority-based services (e.g. to differentiate an ftp download from interactive web browsing traffic). From a user perspective, the QoS provision should not be disrupted by mobility.

Mobility management is divided into global and local management. The global mobility management is supported with Mobile IP (MIP) [3]. This is particularly suitable for managing "inter-domain" handovers – where the routable IP address changes and additionally authentication is required. MIP is not suitable for managing mobility within a single domain of a wireless, cellular style network however: it has large latencies for mobility updates, and could lead to a large signaling overhead in the core network. MIP is therefore not suitable for providing either fast (minimal packet delays) or smooth (minimal loss) handovers. Thus, local mobility mechanisms, which confine the impact of the move, are required. Local mobility should be managed as transparently as possible to the nodes outside of the access network. This is where the BRAIN Candidate Mobility Management Protocol (BCMP) [6] is used (Fig. 2). BCMP allows mobile nodes and the access network to keep the effect of mobility management local so that end-to-end mobility management updates are needed only seldom. In a simple IP network, this protocol could support seamless (i.e. smooth and fast) handovers. Only a limited number of nodes need be modified to support the BCMP. The network consists of legacy IP routers with added mobility aware functionality in just two types of nodes. Anchor Points (ANPs) own and allocate IP addresses, authenticate users, maintain user records, and tunnel packets towards Mobile Hosts (MHs). Access Routers (ARs) terminate tunnels from ANPs and forward packets to/from mobile hosts. By supporting the ability for the access network gateway (ANG) to change, the BCMP provides a resilient and scalable solution.

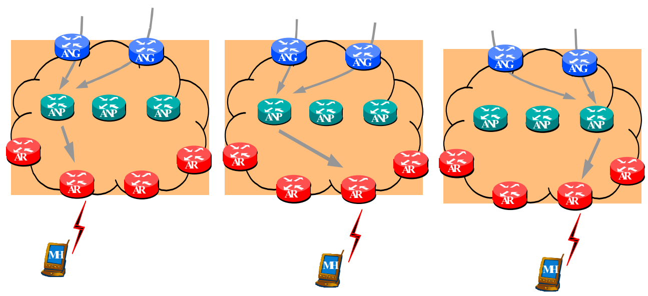

Fig. 2: BCMP operation as mobile node moves between different access routers

The solution to provide QoS is based upon the IETF Integrated Services (IntServ) over Differentiated Services [7] architecture. This supports IntServ parameters and classes within the network, using the Resource Reservation Protocol (RSVP) [8] to request QoS. However, the network is actually provisioned using the flexible and scalable Differentiated Services (DS) [9] network. Only edge routers need to provide per-flow admission control functionality.

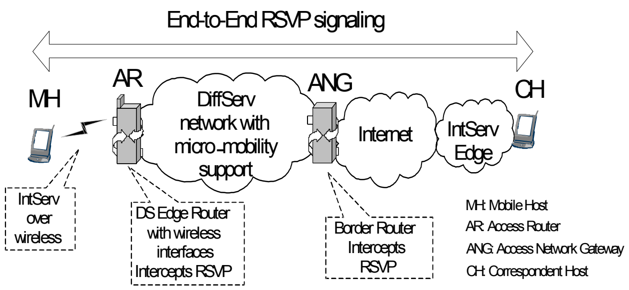

Fig. 3: BRAIN Architecture for Scalable, Reservation-based QoS

However, this solution would result in QoS being disrupted by a mobility event. Further, because of the poor quality of wireless networks, this QoS solution would be essential for any mobile wireless user, yet in the foreseeable future, RSVP will not be supported end to end across the Internet.

Proposals have been made to solve these problems. Of these, within our testbed we will focus on:

- Localized RSVP protocol (LRSVP) [11], to provide QoS in the absence of end to end QoS signaling
- Coupling of mobility and QoS signaling, to ensure that a disrupted reservation is quickly repaired
- Network specific DiffServ Per Hop Behaviour (PHB) to provide access to statically configured guard bands for use of traffic during handover to ensure that any QoS disruption is minimized

IV. TESTBED IMPLEMENTATION

As stated in section I.A, the trials implemented within the MIND project include diverse communication and application aspects, from link-specific issues up to application adaptation. In a first stage, the individual concepts have been developed independently. However, a number of these individual elements are being brought together in a single testbed [Fig. 4] that will allow us to demonstrate a more complete framework for a mobile communications beyond 3G. The hardware infrastructure of this testbed consists of:

- Desktop PCs , used as routers (ANGs, ANPs and ARs).
- Portable computers and PDAs used as mobile hosts. Actually, whilst the PDA provides a more realistic portable terminal, the portable computers are more suitable for software development.
- Wired links provided by 10/100M Ethernet.
- Wireless links provided by IEEE 802.11b wireless LAN PCMCIA cards

The following software set-up is being implemented on the platform:

- The mobile and correspondent hosts both have an implementation of the ISABEL application. ISABEL is a Computed Supported Collaborative Work (CSCW) application provided by ASSA. Within MIND, this application is being developed to enable it to use the control entity and to respond to the information from the network regarding handover and QoS changes.
- The PC routers are all running the latest Linux 2.4 kernel, which has built in support for DiffServ. The routers are configured to support both EF and AF per-hop behaviours. ISI RSVP code [10] is used to implement the RSVP functionality required by the AR and ANG nodes.
- RTP/RTCP is used to enable the application to identify network QoS violations
- BCMP implementation for Linux is installed on ANPs, ARs and Mobile hosts.

However, as described above, this testbed cannot be considered as complete. The following components are essential to demonstrate innovative mobile user scenarios; they are still under development

- The control entity can be further extended, for example to use information obtained from its local resources, particularly the quality of the link layer, in order to make decisions regarding the adaptation of the application.
- L-RSVP is still under development, and all network tests to date assume that the Correspondent Node (CN) is RSVP aware
- Mobility management and RSVP coupling has still to be achieved. This is required in order to provide fast re-establishment of QoS after a handover has occurred.
- All tests to date have been using IPv4, however we believe that IPv6 is more suitable for use in a new mobile network.

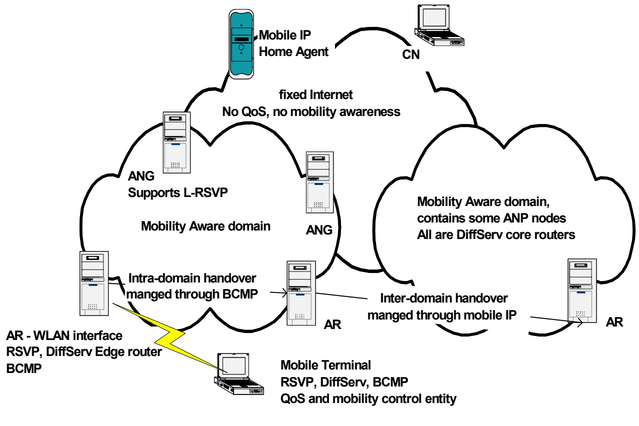

Fig. 4: Testbed Infrastructure

This testbed allows us to demonstrate various network and application scenarios:

- QoS support when there is more traffic trying to cross the access network than resources are available. In this situation, sensitive applications must either try to request specific resources or indicate a relative priority of their communication and rely on application adaptation to smoother the effect.
- Handover scenarios, where the mobile node moves to the coverage area of access points that are under heavier or lighter load than the old access point. This allows us to study how adaptive applications can try to adapt to congestion situations but also to provide a higher quality service to the user by using more resources after than during the handover.

V. CONCLUSIONS

We are developing a testbed that can show how a real end user can have broadband and multimedia services in a wireless mobile environment. This requires a broadband wireless network – such as provided by a wireless LAN, a QoS and mobility enable network in which mobility events do not lead to large disruptions in QoS, and applications that can identify and adapt to the changing environment. Whilst many of these concepts have been previously discussed, implementations are rare, and little effort has been expended in attempting to demonstrate how these different elements could co-operate to support the end user. In many cases, it is clear that individual concepts (for example RSVP and MIP) do not co-operate at all, and so further development is required which provides the correct level of support, without producing "wireless specific" protocols and applications.

As of writing, tests have only been completed on specific elements of the whole testbed. These include tests that demonstrate that a broadband wireless network can, under the test conditions, deliver packets with controlled delay and jitter for the benefit of real-time applications, and tests that demonstrate how adaptive behaviour can aid application – for example by reducing packet loss after handover to a lower capacity network by reducing buffer overflow problems. These tests have been performed using HMIP [12] and MIP as the mobility management protocols. Whilst the two protocols are similar, it is intended to replace HMIP when the BCMP is fully implemented, as the later allows for arbitrary topology within the access network. Further, this protocol supports multiple access network gateways and anchor points, thus making it a more scalable and resilient solution.

ACKNOWLEDGEMENTS

This work has been performed in the framework of the IST project IST-2000-28584 MIND, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues from Siemens AG, British Telecommunications PLC, Agora Systems S.A., Ericsson AB, France Télécom S.A., King's College London, Nokia Corporation, NTT DoCoMo Inc, Sony International (Europe) GmbH, T-Systems Nova GmbH, University of Madrid, and Infineon Technologies AG.

REFERENCES

[1] IST-2000-28584 MIND, Mobile IP-based Network Developments, www.ist-mind.org.

[2] IST-1999-10050 BRAIN, Broadband Radio Access for IP-based Networks, www.ist-brain.org.

[3] MIND deliverable D6.1, "the MIND trial concepts", November 2001

[4] ITU-T Recommendation E.800 (08/94) "Terms and definitions related to quality of service and network performance including dependability"

[5] C. Perkins, "Mobile IP", IETF RFC 2002, October 1996.

[6] K. Ceszei, N. Georganopoulos, Z. Turanyi, A. Valko, "Evaluation of the BRAIN Candidate Mobility Management Protocol", Proceeding IST Summit '01, September 2001.

[7] Y. Bernet, et al, "A Framework For Integrated Services Operation Over DiffServ Networks", IETF, RFC 2998, November 2000.

[8] R. Braden, L. Zhang, S. Berson, S.Herzog, S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1, Functional Specification", IETF RFC 2205, September 1997.

[9] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Services", IETF RFC 2475, December 1998.

[10] http://www.isi.edu/rsvp/release.html

[11] MIND Deliverable D2.1, March 2002.

[12] K. El Malki and H. Soliman, "Hierarchical Mobile IPv4/v6 and Fast Handoffs", Internet Draft (work in progress), draft-elmalki-soliman-hmipv4v6-00.txt, March '00