A Radio Access Network for Next Generation Wireless Networks Based on Multi-protocol Label Switching and Hierarchical Mobile IP

Vasos Vassiliou
Department of Computer Science
INTERCOLLEGE
Nicosia, Cyprus

Henry L. Owen, David A. Barlow
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia, USA

Jochen Grimminger, Hans-Peter Huth, Joachim Sokol
Siemens AG, CT IC2 Corporate Technology
Munich, Germany

Abstract - **This paper presents an option for a radio access network (RAN) suitable for next generation wireless networks. The proposed RAN is based on multi-protocol label switching (MPLS) and hierarchical mobile IP (HMIP). The methods for establishing and modifying LSPs in the proposed architecture are detailed and the operation of the integrated architecture is outlined. The paper explains what benefits can be expected by the use of IP and MPLS in a RAN. The combined benefits are low cost, fast deployment, scalability, reduced latency handoffs, fast table lookup, and fast switching. This type of RAN can be used in current 3G networks with additional gateway elements and can easily be used without modifications in evolved 3G and 4G networks as well.**

I. INTRODUCTION

Next generation wireless networks are expected to be the culmination of all the efforts aimed to create a single type of network architecture and a single suite of services for future communications. Up to now, second generation (2G) networks have been successfully deployed around the world using different architecture standards, but with similar capabilities. These capabilities have been improving in their latest incarnations, but the whole technology seems to have reached its limits. Second generation systems have started to be inadequate to support the requirements for higher data rates and advanced services expected by users nowadays. Most of the additional requirements are created by the demand for seamless roaming among different systems and worldwide access to the Internet. Additional requirements are also placed on networks because user expectations are increasing to more sophisticated services like multimedia, (image and video) content, and professional services typically related to broadband wired systems.

Third generation cellular wireless networks (3G) were created with the aim to support these requirements. A serious effort has been made to create only one standard under the International Mobile Telecommunication 2000 (IMT-2000) umbrella. This would have solved the problem of global mobility. However, in the process, two major standards have been accepted for the wireless interface. The standardized 3G solutions are WCDMA (wideband CDMA) based on UMTS (Universal Mobile Telecommunication System)[1] and CDMA2000 an extension of IS-95[2]. The two standards may have different air-interface methods, but have a lot of similarities in the access and core networks.

3GPP[3] and 3GPP2[4], the bodies developing and promoting the WCDMA (UMTS) and CDMA2000 (IS-95) solutions respectively, are currently driving toward an IP core network harmonization. Such harmonization will enable the creators and the operators of the two systems to capitalize on common network entities, common protocols, and common APIs. This commonality will ensure service transparency, application portability (context transfer), and seamless mobility. Convergence is mostly sought in the multimedia domain of the IP core; however, it assures a coordinated evolution plan towards All-IP wireless networks (4G) that share more common entities besides a common transport type.

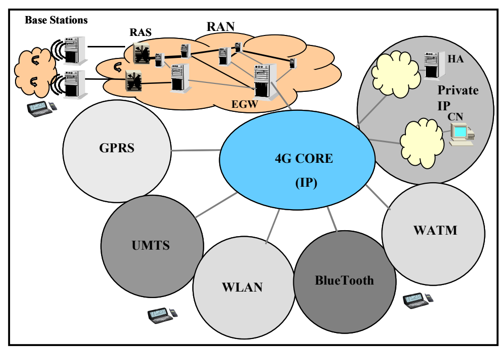

Figure 1. 4G Example Network

All-IP is the concept of moving the cellular wireless network architecture from the current circuit-based concept to a packet-based architecture utilizing IP protocols and technology where possible. According to [5], there are four main reasons for this design choice. First, an IP-aware RAN can give better support to IP applications. IP is the basis for packetized data, voice and signaling, making the use of IP based protocols and technology desirable to be utilized wherever possible. Second, an IP infrastructure is widely available, reducing the cost of deployment. Third, IP-style engineering is faster and cheaper. Fourth, by constructing networks based on IP technology, seamless connections can exist between 2G, 3G, 4G, WLANs, Bluetooth, and other access networks as illustrated in Figure 1.

This paper proposes the integration of hierarchical mobile IP (HMIP) and multi-protocol label switching (MPLS) and its use in the radio access network (RAN) of a future wireless system. The proposed scheme merges the intelligence, scalability, and reduced latency handoffs of HMIP with the performance and traffic management capability of MPLS. More specifically, MPLS is a good enhancement of an IP-based RAN because, among others, it has faster table lookup, less control overhead, and it can be applied over networks using any Layer 2 switching which proves to be very desirable in the multi-radio access architectures considered for future wireless networks.

Section II of this paper provides a brief background on the basic protocols and architectures used in our proposal. Section III presents the proposed radio access network based on MPLS and HMIPv6. Section IV relates our proposal with current and future cellular standards and finally Section V provides a conclusion of all the ideas presented.

II. BACKGROUND

Multiprotocol Label Switching: MPLS [6] is a packet forwarding technology that assigns packet flows to Label Switched Paths (LSPs). Packets are classified at the network edge based on Forwarding Equivalence Classes (FECs). FECs summarize essential information about routing the packet such as destination, precedence, VPN membership, QoS information, and the route of the packet chosen by traffic engineering (TE). Based on the FEC, packets are labeled, and then transported over a label switched path (LSP) based on that label. Packets belonging to the same FEC get similar treatment by all intermediate nodes in the path.

MPLS operates between layer two (data link) and layer three (network) of the protocol stack, thus it is referred to as a 2.5 layer architecture. MPLS provides faster delivery of data packets through switching than IP routing. This capability is made possible because the forwarding paradigm of MPLS is exactly the same as that provided by ATM. If an ATM switch is used, MPLS has the capability to translate its label information to VCI information and take advantage of the switches full capabilities. If IP routers are used, MPLS inserts the label information in a shim header between the Layer 2 and Layer 3 headers. Whether MPLS runs over an ATM switch router or an IP router it can provide a faster service than pure IP because it does not have to send the packets to the IP layer for processing. The savings in this case are realized first by the fact that the forwarding entity deals with a smaller amount of data (20-bit labels in MPLS instead of 128-bit addresses in IPv6) and that MPLS has less processing delay at a router because a label table search is faster than longest-bit- matching.

Mobile IP: Mobile IP (MIP) allows a mobile node (MN) to move from one link to another without changing the mobile node's home IP address [7]. A home address is an IP address assigned to the mobile node within its home subnet prefix on its home link. Packets may be routed to the mobile node using this address regardless of the mobile node's current point of attachment to the Internet, and the mobile node may continue to communicate with other nodes (stationary or mobile) after moving to a new link. While a mobile node is attached to some foreign network, it is also addressable by one or more care-of addresses (CoA). When away from home, a mobile node registers one of its care-of addresses with a router on its home link; requesting this router to function as the home agent (HA) for the mobile node. The HA intercepts, encapsulates, and forwards packets to the mobile node through its registered CoA.

Hierarchical Mobile IP: Hierarchical Mobile IP (HMIP) is a micro-mobility management model. Its purpose is to reduce the amount of signaling to correspondent nodes and the home agent and improve the handoff speed performance of mobile IP. With the increase in operating frequencies for 4G, cell radii decrease to half that of 3G. This means that to cover the same area as in 3G, fourth generation systems will require four times the number of base stations. More base stations are also required to support the increase in capacity. [8]. The implication of having smaller and more base stations is that mobile nodes will cross cell boundaries more often and the amount of signaling they exchange will increase proportionally. HMIPv6 [9] introduces a Mobility Anchor Point (MAP), and minor extensions to the mobile node and home agent operations. The major idea is that the mobile node registers the MAP's CoA with its home agent instead of its own or the FA/Access Point it is currently attached to. Therefore, when the mobile node moves locally (i.e. its MAP does not change), it only needs to register its new location with its MAP, but nothing needs to be communicated with the home agent or any other correspondent nodes outside the RAN. Using this method, signaling is contained in a smaller area and does not overwhelm the core network.

III. PROPOSED RADIO ACCESS NETWORK

In Sections I and II we have discussed how future cellular networks have increasing demands, not only from the physical architecture standpoint, but also from the types of services needed to be supported. We believe that the increased requirements of an IP-based RAN can be solved when the scalability and reduced latency of HMIP is combined with the switching performance and traffic engineering capabilities of MPLS.

A. Radio Access Network

The basic topology for this research work is a radio access network (RAN) as shown in Figure 1. The RAN consists of at least three layers of label switched routers (LSRs). The edge components of the architecture (outmost circle) are the radio access routers (RAS), which are the first IP aware devices of the network seen from the mobile terminal. One, or more, base stations (BS) are attached to a RAS. Several RASs are interconnected to one or more Edge Gateways (EGW - innermost circle), which in turn provide access to outer (backbone) networks including other RANs. The RASs and the EGWs are linked through a network of label switched routers (LSRs). We assume that all routers in the RAN can act as mobility agents (MA) to support mobility management based on hierarchical mobile IPv6.

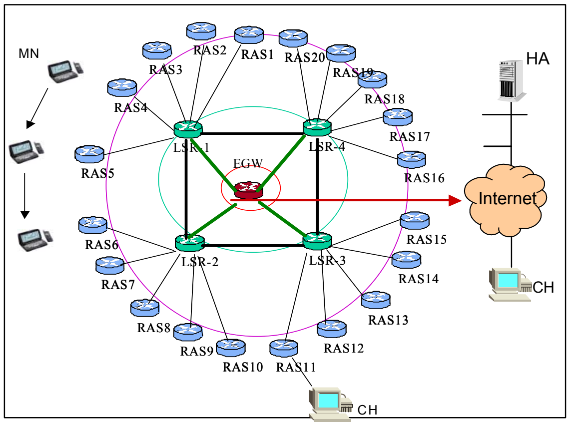

Figure 2. MPLS-HMIP Integrated Radio Access Network

Enhancements to the network may include additional layers of hierarchy and more complex interconnection like full mesh or double homing. We believe that the reference network in Figure 2 has the required characteristics to support an integrated HMIP-MPLS network with traffic engineering capabilities.

B. MPLS-HMIP Integration

The main idea of this integration proposal is that the tunneling procedures of MIP can be replaced by label switched paths (LSPs) in the MPLS network similar to [10]. Figure 3 illustrates the steps required to setup an LSP for communication from the correspondent node to the mobile node for the first time, when the mobile node is at a foreign network.

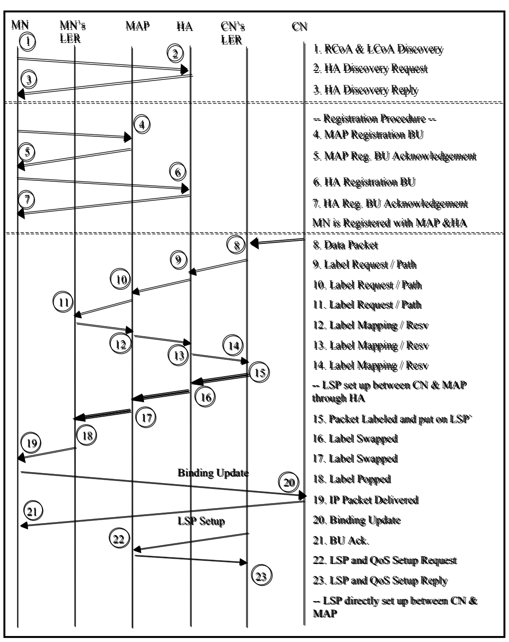

Figure 3. Data Driven LSP Setup in HMIPv6-MPLS

When a correspondent node initiates communication toward a mobile node it first examines its binding cache for an entry of mobile node's new CoA. If the correspondent node does not have an entry it sends the packet to the mobile node's home address. If there is an LSP to the mobile node's home address the CN's LER pushes an appropriate label on the packets and sends them to the mobile node. If no label is bound for that FEC, the CN's LER will have to set up an LSP to the mobile node's network with that FEC. The edge router on the mobile node's domain (which also serves as a home agent) will receive the request, but since it knows the mobile node's current address it forwards the request toward the RCoA registered for the mobile node. The MAP also has a binding for the mobile node, but no label bound for it, therefore, before it can send a label mapping to the home agent, the MAP must get a label from the LER serving the mobile node at its new location. At the end of this operation the correspondent node's LER will have an outgoing label for sending packets to the mobile node (through the home agent) and the mobile node's LER an incoming label for receiving packets for the mobile node.

After a mobile node receives messages to its new location with its home address as the destination address it understands that the correspondent node does not have an updated binding of its location. It then sends a binding update to the correspondent node. The correspondent node's LER will decide if a request/path message needs to be sent to establish an LSP with the mobile node's MAP. This decision depends on the amount of data traffic in the downstream direction. The LER will recognize if a persistent data stream exists and will initiate an LSP setup. The LSP setup request will create a shorter LSP directly from CN's LER and the MAP.

When a mobile node sends packets to a correspondent node, it sends the packets directly – without using the home agent. At the beginning of the communication the mobile node's LER will have to create an LSP to the correspondent node's LER before it can forward any packets. Figure 4 shows the details for this scenario.

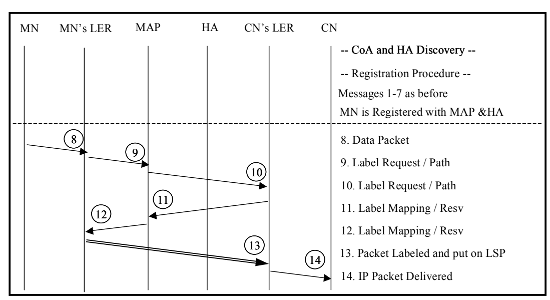

Figure 4. Data driven LSP Setup in HMIPv6-MPLS – Mobile Initiated

C. Quality of Service

MPLS has a unique capability of being able to set up routes not only based on shortest path, but also on other metrics like load and availability. By utilizing explicit routes, MPLS provides benefits to service providers in the field of traffic engineering. This term refers to the ability to control where traffic flows in a network, with the goal of reducing congestion and getting the most use out of the available resources. In addition, MPLS provides in-sequence packet delivery, which is particularly useful in real time (voice or video) communication.

A QoS method that is very compatible and popular with MPLS is differentiated services. For the support of DiffServ in an MPLS environment, the different DiffServ classes of service need to be communicated using the MPLS header. The recently approved IETF RFC on the matter [11] suggests that a Forward Equivalence Class (FEC) can be used to denote the route's DiffServ class in addition to the route's destination. There are two methods to achieve this functionality: (a) Establish a different LSP for a single FEC-aggregate pair, called the Label-only-inferred-PSC LSPs or L-LSP, or (b) use a single LSP to support up to eight different aggregates for the same FEC, called the EXP-inferred-PSC LSPs or E-LSPs because the 3 EXP bits in the MPLS header are used.

The proposed HMIP-MPLS integrated framework could support any traffic engineering algorithm based on DiffServ, provided that the mobile node remains or roams within an area supported by the same service level agreement. If not, we assume that the agreement is communicated between the old and the new domain through some higher layer entity (bandwidth broker) that is usually co-located either with the EGW node or an authorization, authentication, and accounting server (AAA) called during the mobile node's first entrance into the new domain. Investigating specific traffic engineering options like the interaction between bandwidth brokers and AAA servers is out of the scope of this work.

IV. MPLS-HMIP RAN IN 3G – 4G NETWORKS

A radio access network encompasses all functions that enable a user to access core network services. It can be used to hide all access specifics from the core network and all core management issues from the terminals.

In the proposed RAN the base stations are not IP based entities. They just provide last hop wireless connectivity to the mobile node. The first MPLS and IP capable node is the radio access router (RAS). We have seen that the RAN can have more than one layer between the edge routers and the gateway to the core network. This gateway can be a simple edge router if the core of the network is also IP-based, or can be an entity with interworking features able to connect the MPLS/HMIP based RAN to whatever technology the core uses. Let us first briefly examine the architectures of 3GPP and 3GPP2.

The European version of IMT-2000, the Universal Mobile Telecommunications System (UMTS) has evolved from the already-existing GPRS network and shares a lot of its components. It is comprised of a core network (CN) and a radio access network called the UMTS Terrestrial RAN (UTRAN). The UTRAN is the section connecting to the user equipment. The core network has two parts; the circuit switched part (CS) and the packet switched part (PS). In this work we are only interested in the packet switched part. The main entities in the core PS network are the gateway GPRS support node (GGSN), the serving GPRS support node (SGSN), the home location register (HLR), and the border gateway (BG). The core network performs location control, call control, and service control. The SGSN is the first IP-aware point of contact for the MS. It provides mobility management, authentication, ciphering, routing, and admission control. The SGSN connects the Base Station System (BSS) to the GGSN and maintains a PDP context for every mobile node attached to it. The GGSN is a gateway node that connects to the packet data network (outside network – Internet). It supports tunneling, accounting, mobility management, routing, encapsulation, and authentication. The GGSN maintains an entry on the relationship of a mobile and the corresponding SGSN. The UMTS terrestrial radio access network (UTRAN) consists of a set of radio network systems (RNS). Each RNS consists of a radio network controller (RNC) and one or more node B's (base stations). The RNC provides radio resource management and connects to the SGSN and the MSC. The SGSN and RNC may be interconnected with one or more IP routers.

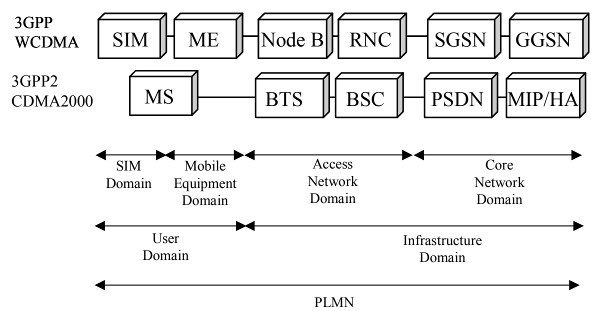

Figure 5. Simplified 3GPP and 3GPP2 Network Architectures and Domains

The architecture based on 3GPP2 is very similar in structure to 3GPP. However, this architecture was created without any existing packet data service and incorporated well researched and developed IETF protocols (like MIP) in its architecture. In the 3GPP2 architecture there is also a core and an access network (cdma2000 access network). The core network is also separated to circuit-switched and packet-switched parts. The packet switched part includes the packet data serving node (PSDN) and the home agent (HA). The PSDN is the first IP-based router seen from the mobile node and performs DiffServ shaping, ingress filtering, accounting and loacl handoffs. The home agent performs similar functions as in mobile IP. The access network (cdma2000 access network) is comprised of the base station controller (BSC) and the base transceiver station (BTS).

Figure 5 shows the simplified network architectures of 3GPP and 3GPP2. The figure also shows the different domains of a public land mobile network (PLMN). The radio access network in these architectures is below the RNC or the BSC.

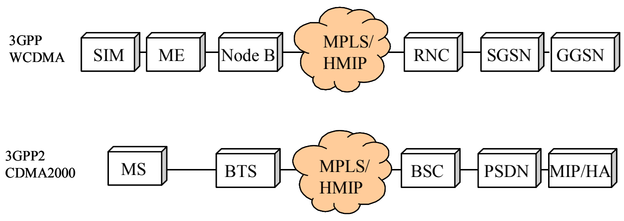

Figure 6. 3GPP and 3GPP2Network Architectures with MPLS/HMIP RAN

Therefore, a network using the proposed framework as the transport mechanism in the RAN will conceptually look like in Figure 6. The integrated MPLS/HMIP framework will operate in the UTRAN or cdma2000 access network environment as described in Section III. Instead of setting up LSPs based on requests from correspondent nodes, LSPs in these networks are set up between base stations (BTS or Node Bs) and the controlling nodes (BSC and RNC respectively) only. There is no home agent in this case and the controlling nodes play the role of the MAP for the outside network. Depending on the quality of service requirements of the traffic, the LSPs can be set up to carry only one or many DiffServ classes using the L-LSP and E-LSP methods discussed in Section III. The LSPs could also be bi-directional if the controlling node recognizes that the upstream and downstream traffic are identical.

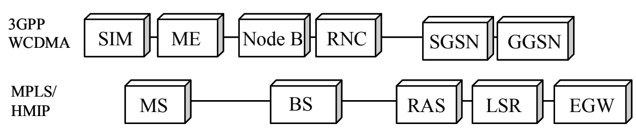

Figure 7. MPLS/HMIP 3GPP PLMN

Given that IP transport has been considered also for the core network [15][16], besides the RAN, we have compared our proposed framework with the whole PLMN. The relationship is shown in Figure 7. In order to support IP-based mobility using the HMIPv6 protocol, the 3GPP architecture must give mobility agent capability to the SGSN and GGSN. Therefore the GGSN automatically becomes the EGW/MAP and the SGSNs and any other intermediate routers form the lower hierarchy just as LSRs and RASs do in the integrated HMIPv6/MPLS network. A similar relationship exists between the 3GPP2 architecture and the MPLS/HMIP framework as shown in Figure 8.

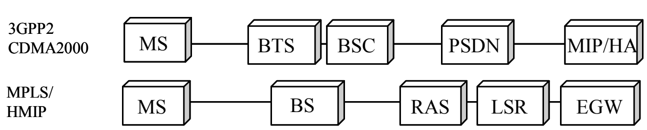

Figure 8. MPLS/HMIP 3GPP2 PLMN

In this case, the architecture already includes a home agent node, which can serve as the MAP. The only change is to make PSDN a mobility agent. Even if the SGSNs and PSDNs assume mobility agent capability there is no further hierarchy in the core network limiting the usefulness of thierarchical MIP and reducing the options for traffic engineering.

V. CONCLUSIONS

In this paper we proposed a scheme that introduces MPLS and micro-mobility in a packet based wireless cellular network. We have shown how MPLS and HMIP can be integrated to support mobility and QoS in a RAN and what options exist to making them efficient in that role. This type of RAN can be used in current 3G RANs without changes. With additional mobility elements in the core cellular architectures this type of framework can replace all or parts of current 3G networks and can be easily used without modifications in future 4G networks as well.

REFERENCES

[1] 3GPP Network Architecture (Release 5), 3GPP TS 23.002 v5.6.0, Mar. 2002.

[2] 3GPP2 IP Network Architecture Model for cdma2000 Spread Spectrum Systems, May, 2002.

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

[4] http://www.3gpp2.org/

[5] T. Robles et al.,"QoS Support for an All-IP System Beyond 3G," *IEEE Communications Magazine*, pp64-71, Aug. 2001

[6] A. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol Label Switching Architecture," Request for Comment 3031, IETF, Jan. 2001.

[7] C. Perkins, " IP Mobility Support", Request for Comment 3220, Internet Eng ineering Task Force, Jan. 2002.

[8] T. Otsu, I. Okajima, and Y. Yamao, "Network Architecture for Mobile Communications Systems Beyond IMT-2000," *IEEE Personal Communications Magazine*, pp..31-37, Oct. 2001.

[9] H. Soliman, et al., "Hierarchical MIPv6 Mobility Management," *IETF Internet Draft*, Nov. 2001.

[10] Z. Ren, C-K Tham, C-C Foo, and C-C Ko, "Integration of Mobile IP and Multi-Protocol Label Switching," *In Proceedings of IEEE International Conference on Communications (ICC) 2001*, pp. 2123-2127, Jun. 2001.

[11] F. Le Faucheur et al, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services," IETF RFC 3270, May. 2002.

[12] MWIF Technical Report MTR-006, June 2001

[13] G. Swallow, "Simple Header Compression," IETF Internet Draft, Sep. 2000.

[14] L. Berger and J. Jeffords, "MPLS/IP Header Compression," IETF Internet Draft, Jul. 2000.

[15] ETSI, Combined GSMa dn Mobile IP Mobility handling in UMTS IP CN, ETSI TR 123 923 v3.0.0, May 2001.

[16] 3GPP2 Wireless IP Network Standard, 3GPP2 P.S0001-A v3.0.0, July 2001.