FAST HANDOFF SUPPORT IN AN IP-EVOLVED UMTS ARCHITECTURE

Lila Dimopoulou, Georgios Leoleis, Iakovos S. Venieris
School of Electrical and Computer Engineering, National Technical University of Athens, 9 Heroon Polytechniou str, 157 73 Athens, Greece

Abstract: IP technology will play a key role in beyond 3G systems, which face the great challenge of integration in order to provide seamless service to users anywhere and anytime. Apart from its natural role as a unifier, IP also comprises the main drive for network evolution towards all-IP network infrastructures. In this regard, we exploit IP as an enabler for the evolution of the UMTS packet-switched core network, eliminating its duality at user and transport level. We focus on mobility management in the core network, which is handled by pure IP mechanisms (Mobile IPv6, MIPv6), and on the support of fast handoff across UMTS access networks by means of the IETF's Fast MIPv6 proposal. Emphasis is put on identifying the proper interaction points between the Fast MIPv6 operation and the UMTS-specific Serving Radio Network Subsystem (SRNS) relocation procedure in order to provide a seamless handoff service to the user while not compromising the network's performance and scalability.

Key words: Fast Handoff; Fast Mobile IPv6; Beyond 3G System; SRNS Relocation.

1. INTRODUCTION

IP technology has evidently played a key role in cellular mobile systems, such as the UMTS, where it has been adopted as the transport means of the packet switched (PS) core network, as an effort to support packet-based services in an efficient and cost effective manner. Although its impact has been primarily economic, considering the homogenization it achieves on the network resulting in the latter's easier maintenance and operation, it has become an imperative need from technological standpoint in future mobile networks – beyond 3rd Generation (3G) networks – facing the great chal-

lenge of integration[1] in order to provide seamless service to users anywhere and anytime.

Such integration involves all networks that simply co-exist nowadays, namely 3G cellular systems, Wireless Local Area Networks (WLANs), mobile ad-hoc networks, Personal Area Networks etc. World-wide roaming between various network technologies, continuously growing in heterogeneity, necessitates their seamless integration. IP is naturally the best choice to serve the internetworking and unification of these diverse technologies. However, apart from its role as a unifier, IP technology will also be the main drive for network evolution towards all-IP core networks. By that we mean that cellular infrastructures are directed to adopting IP as the basis for networking in the core network, for the latter to become access technology agnostic, and to allow a more efficient and straightforward integration with diverse access technologies.

In this paper, we address the IP technology as an enabler for network evolution, on the basis of the UMTS infrastructure. We propose a target architecture, where the UMTS Terrestrial Radio Access Network (UTRAN) forms the user's access means towards an all-IP core network. The duality of IP in application (user) and transport level[2], as it is the case in the UMTS PS core network, - jeopardizing network efficiency and performance - is eliminated, and a single IP layer is employed for all functionalities. Our focus has been brought on mobility management in the core network where pure IP-based solutions, and in particular Mobile IPv6[3] –MIPv6–, are adopted. Our contribution however concerns the application of the IETF's Fast MIPv6 mechanism[4] at the borders of the core network for achieving seamless handoffs across UMTS access networks. This necessitates an in depth study of the UMTS procedures[5] executed during handover for rendering their seamless interworking with IP handoff procedures feasible.

The rest of the paper is structured as follows. Section 2 presents the background in Mobile IPv6, IPv6 address autoconfiguration and Fast Handoff issues with the aim to identify the main contributions of handoff delay. In section 3, the target architecture is introduced and the Fast Handoff procedure in the context of the proposed environment is described in great detail. Certain issues regarding the handoff procedure and enhancements to it are examined in section 4 while section 5 concludes our paper.

2. BACKGROUND

2.1 Mobile IPv6

Mobile IPv6[3] comprises the IETF solution for handling the mobility of hosts in IPv6 networks. It extends the basic IPv6 functionality by means of

header extensions rather than being built on top of it, as it is the case with MIPv4. Its fundamental principle is that a mobile host should use two IP addresses, a permanent address – the *home address*, assigned to the host and acting as its global identifier – and a temporary address – the *care-of address* (CoA), providing the host's actual location –. A mobile host (MH), while attached to its home network, is able to receive packets destined to its home address, and being forwarded by means of conventional IP routing. When the host crosses the boundaries of its serving network, movement detection is performed in order to identify its new point of attachment and further acquire a new CoA (nCoA). Once configured with a CoA, the MH needs to send a Binding Update (BU) message to its Home Agent (HA) to register this 'temporary' address. This CoA is obtained through IPv6 address autoconfiguration mechanisms; however, the time needed for autoconfiguration and for the Binding Management to complete sets the MIP operation inefficient for fast intra-domain movements.

According to the typical Mobile IP operation, the correspondent host (CH) addresses the MH at the latter's home address, and consequently does not need to implement the specific IPv6 extensions, which actually form MIPv6. In the opposite case – when the CHs are augmented with the MIPv6 functionality – then route optimization can be used for the direct delivery of packets to the MH without the intervention of the HA. The CHs are able to associate the MH's home address with a CoA – via BUs transmitted by the MH to CHs. However data packets will not be encapsulated for delivery to the MH, as is the case in MIPv4, but instead an IPv6 Routing header will be used for this purpose. These packets have as destination address the MH's CoA. The 'home address' information, required to preserve transparency to upper layers and ensure session continuity, is included in the routing header. In the reverse direction, packets have as source address the host's CoA while the home address is included in the newly defined *home address* destination option.

2.2 Handoff View of IPv6 Address Autoconfiguration

There are two mechanisms defined for the allocation of an IPv6 address to a node: the stateless and the stateful autoconfiguration. The stateful mechanism requires a Dynamic Host Configuration Protocol server to perform the address assignment to the IPv6 node. On the other hand, the stateless autoconfiguration procedure does not need any external entity involved in the address autoconfiguration (apart from the entity functioning as the first hop IP router, referred to as the access router –AR–). The stateless mechanism[6] allows a host to generate its own addresses using a combination of locally available information and information advertised by ARs. The latter advertise prefixes that identify the subnet(s) associated with a link, while

hosts generate an 'interface identifier' that uniquely identifies an interface on each subnet. A global address is formed by combining the two. The formation of an address must be followed by the Duplicate Address Detection (DAD) procedure in order to avoid address duplication on links. Since it is the interface identifier that guarantees the uniqueness of the address –all hosts on the link use the same advertised prefix–, it suffices to perform DAD on the link local address. The latter is formed by appending the interface identifier to the well-known link-local prefix (FE80::) and only allows for IP-connectivity with nodes located at the same link. Global addresses formed from the same interface identifier need not be tested for uniqueness. In brief, the address autoconfiguration is composed of the following steps:

1. The host generates a link-local address for its interface on a link
2. It then performs DAD to verify the uniqueness of this address, i.e. verify the uniqueness of the interface identifier on the link
3. It uses the prefix(es) advertised by routers for forming a global address. DAD is not needed if the same interface identifier as in the link-local address is used.

When a host handoffs to a new subnet, it needs to be configured with a new global address, which is topologically correct, for being able to receive data on the new link. As it is expected, DAD needs to be performed for this address so as to verify its uniqueness on the link. As before, DAD can be once executed on the host's link-local address, given that the newly formed global address uses the same interface identifier. The basic deficiency coming along with DAD execution is that it adds delay to the handover. For the IP connectivity to be regained, after the establishment of link-level connectivity, some additional time is needed. In particular, during DAD, the host transmits a *Neighbor Solicitation* for the tentative link-local address and waits for *RetransTimer* milliseconds[7] till it considers the address unique. More precisely, the exact number of times the *Neighbor Solicitation* is (re)transmitted and the period between consecutive solicitations is link-specific and may be set by system management. DAD only fails if in the mean time, the host receives a *Neighbor Advertisement* for the same address, meaning that another host is using the being questioned address or if another host is in the progress of performing DAD for the same address and has also transmitted a *Neighbor Solicitation*. From the above it is deduced that at least a link-wide round-trip is needed for performing DAD while 1.5 - 2 round-trips are required in total for the whole autoconfiguration procedure if the router discovery (step 3) is performed in the sequence.

2.2.1 UMTS links

Let us examine how the IPv6 stateless address autoconfiguration mechanism is supported in UMTS[8], as it is recommended by the IETF IPv6 WG[9], and which of its inherent delays are eliminated. We should first present the

UMTS architecture in an IP centric view and understand how UMTS links are viewed from the IP layer. According to UMTS terminology, the Packet Data Protocol (PDP) context defines a link between the cellular host and the GGSN (Gateway GPRS Support Node), over which packets are transferred. An IP address is initially assigned to a primary PDP context while zero or more secondary PDP contexts use the same IP address. Thus, all PDP contexts using the same IP address define a link (a point-to-point link, referred to as UMTS link). A host may have activated more than one primary PDP contexts, i.e. may have more than one links to the GGSN(s).

Over the UMTS links, the host should configure a link-local address for on-link communication with the GGSN, and global address(es) for communication with other hosts. What is most important here is that the GGSN assigns the interface identifier to the host for forming its link-local address (corresponding to a UMTS link). The GGSN only has to ensure that there will be no collision between its own link-local address and the one of the cellular host, i.e. between its own interface identifier and the one assigned to the host. As a consequence, the host does not need to perform DAD for this address. Moreover, the cellular host must form a global address, based on the prefix(es) advertised by the GGSN. However here, the GGSN assigns a *prefix* that is *unique within its scope* to each primary PDP context. The uniqueness of the prefix suffices to guarantee the uniqueness of the MH's global address. This approach has been chosen taking into account that hosts may use multiple identifiers (apart from the one assigned by the GGSN) for forming global addresses, including randomly generated identifiers (e.g. for privacy purposes). This avoids the necessity to perform DAD for every address built by the MH. To sum up, the way address autoconfiguration is performed in UMTS eliminates the need for DAD messages over the air interface and therefore removes this factor of delay. The afore-described concepts are summarized in Table 1.

Table 1. IPv6 autoconfiguration concepts

	Shared Links	UMTS Links (Point-to-Point)
Interface Identifier	Generated by the host	Assigned by GGSN
Link-local address uniqueness	Guaranteed, DAD is performed	Guaranteed, GGSN is the only neighbor on the link with a different Interface ID
Prefix	Assigned by the Router to all hosts on the link	Unique prefix assigned to host by GGSN
Global address uniqueness	Guaranteed if using the Interface ID of link-local address	Guaranteed due to Unique Prefix
DAD	Needed for the link-local address and for global addresses using other Interface IDs	Not needed

2.3 MIPv6 Enhancements towards Fast Handoffs

As obviated earlier, MIPv6 presents some deficiencies due to the inherent delays introduced by address autoconfiguration and binding management. Fast MIPv6[4] (FMIPv6) comes to address the following problem: how to allow a mobile host to send packets as soon as it detects a new subnet link, and how to deliver packets to a mobile host as soon as its attachment is detected by the new access router. In other words, FMIP's primary aim is to eliminate the two factors of delay in address autoconfiguration. It achieves this by informing the mobile host, prior to its movement, of the new AR's advertised prefix, IP address and link layer address. The mobile host is already configured with the new address at the time it attaches to the new link. Supposing that the uniqueness of the address is guaranteed, the host can start sending packets in the uplink direction, setting the new address as the source address of these packets. In the downlink direction, a factor of delay is yet introduced before the new AR (nAR) can start delivering packets to the host. The nAR typically starts the Neighbor Discovery operation as soon as it receives packets for a host, in order to detect its presence and resolve its link layer address. This operation results in considerable delay that may last multiple seconds. In order to circumvent this delay, the FMIPv6 procedure requires from a MH to announce its attachment through a *Fast Neighbor Advertisement* (FNA) message that allows nAR to consider it reachable.

FMIPv6 is also essential for de-correlating the packet reception and transmission capability of the host from the time needed for the Binding Updates to HA and CHs to complete. This is required for two reasons:

1. The MH cannot start sending packets to CHs it communicates with, setting as source address the new CoA, prior to sending a BU to them, since the CHs will drop these packets.
2. The MH will not be able to receive packets from CHs at its new address, till the CHs update their caches for the host.

These two problems are basically addressed by setting up a bidirectional tunnel between the old AR and the MH at its nCoA. The tunnel remains active until the MH completes the Binding Update with its communicating hosts. To CHs, the mobile host is located at the old subnet; the old path is temporarily extended with the branch *old AR – nCoA of host* for allowing communication to continue during the IP handoff transition period. The full path is reestablished when the Binding Update procedure completes.

In brief, the operation of the protocol is as follows: the host sends a *Router Solicitation for Proxy* (RtSolPr) message to its AR so as to obtain information – e.g. prefix – related to available access points. The AR serving the user responds with a *Proxy Router Advertisement* (PrRtAdv) containing the requested information and thus allowing the mobile host to perform address autoconfiguration as if it had already migrated to a new link. The host,

after formulating a prospective new CoA, sends a *Fast Binding Update* (FBU) to its AR for requesting the tunneling of packets addressed to its old CoA (oCoA) towards its nCoA. The AR serving the host (referred to as old AR, oAR) exchanges *Handover Initiate* (HI) and *Handover Acknowledge* (HAck) messages with the nAR for initiating the process of the MH's handover, while possibly validating the nCoA formed by the host. The oAR responds to the MH with a *Fast Binding Acknowledge* (FBack) message on both links (old and new) and starts the tunneling of arriving data. The MH, as soon as it attaches on the new link, transmits an FNA for informing the nAR of its presence. Packets from this point on can be delivered to the MH.

3. TARGET ARCHITECTURE

Before presenting our reference architecture, we shall elaborate on the requirements that the architecture should meet. We aim at a fully homogenized, access-agnostic core network (CN) solution that will cover the majority of access technologies under the IP suite umbrella. As a first step, we shall use the advanced radio access technologies offered by UTRAN. Moreover, the proposed solution should leave the UTRAN intact, that is, the core network should only view the standard Iu interface[10] and utilize its standardized capabilities. Further, the architecture, being IP-centric, should support all functionalities implemented by legacy UMTS protocols. Our focus however is placed on mobility management. Last, Fast Handoff should not be an access-dependent capability. Instead, it should be supported by the access-agnostic core network for handling the movement of the users across any access network.

Based on these requirements, we propose a target architecture, being evolved from the legacy UMTS network, with a view to bring the IP layer closer to the access network while maintaining performance, efficiency and scalability. To this aim, we integrate both GPRS Support Nodes (GSNs) of the UMTS PS core network, to form a single node, named as UMTS Access Router (UAR), and situated at the border of the RAN and the CN of the cellular infrastructure. In this way, the UMTS PS core network is replaced with a fully IP-compatible backbone where the UAR acts as the first-hop IP router. Traditional UMTS mechanisms in the CN, such as mobility management, being based until now on the heavy GPRS Tunneling Protocol (GTP) operation[5], are handled by IP-oriented ones. Mobile IPv6 is adopted as the ultimate mobile IP solution and hence, the user's movement across UARs will trigger the Binding Update procedure with the HA and/or CHs and not the re-establishment of GTP tunnels. However, the basic strength of the proposed architecture lies in that UARs run Fast MIPv6 for enabling the fast handoffs of Quality-of-Service (QoS) stringent sessions across RANs that do

not allow for soft handoff[1]. This IP-based fast handoff capability also renders the UMTS Gn interface[2] among UARs redundant, since location management functionalities are now covered by FMIPv6.

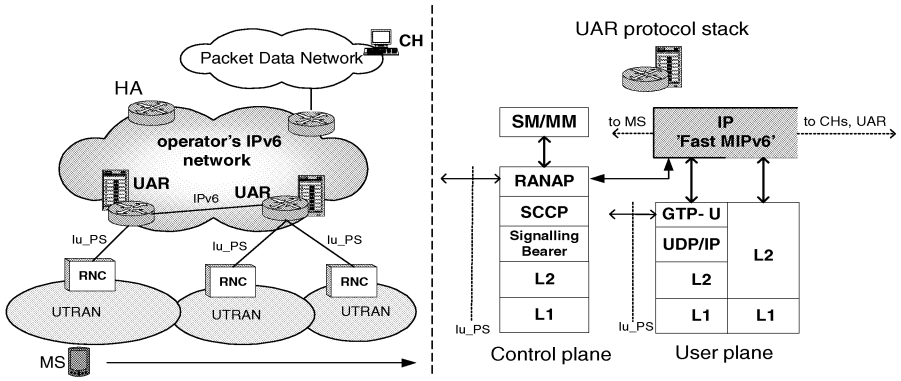

Figure 1. Target Architecture and UAR protocol stack

Fig. 1 depicts the target architecture; as it is shown, the UAR is mediated among UTRAN and the IP-based core network while trying to hide the peculiarities of the former towards the latter (by terminating legacy UMTS Non-Access-Stratum protocols for Session and Mobility Management – SM, MM). Moreover, no change in the UMTS protocol stack over the communication links with UTRAN (Iu-PS interface) has been effected, in both control and user planes. Fast MIP runs over the UTRAN user plane towards the Mobile Stations (MSs) while it also interoperates with RANAP (Radio Access Network Application Protocol), as it will be described below, for efficiently handling the MS's handoff across UARs.

3.1 Fast Handoff Procedure

In this section, we provide the full details of interoperation between UMTS and IP mechanisms during the handoff procedure. Our main objective has been to find the appropriate points of interaction (triggers) between the two layers – UMTS and IP – for enhancing overall handoff performance. The SRNS (Serving Radio Network Subsystem) relocation[5], as performed in UMTS when the user is being served by a new RNC (Radio Network Con-

[1] *Soft handoff* – macro-diversity – enables a seamless type of handover at layer 2. This is not applicable to inter-technology handoffs or even intra-technology ones, such as UTRAN-UTRAN handoffs, where the RANs are not interconnected (Iur interface not available).

[2] The Gn interface lies between two SGSNs and is implemented by GTP. It mainly carries out location management functionality.

troller), is the basis of the Fast Handoff procedure introduced, while interoperation with the FMIP protocol additionally takes place.

The handoff scheme, as depicted in Fig. 2, is triggered by UTRAN functions – sbased on measurements between the MS and the RNC serving the user. The SRNS relocation procedure starts and is composed of two main phases; a) the reservation of resources on the new link –between the target RNC and target UAR–, b) the target RNC takes over the serving RNC (SRNC) role and starts delivering data to the MS. According to the UMTS standardized procedure, a temporary GTP tunnel is used between the concerned RNCs for the forwarding of data following the old path until the new one – GGSN to target RNC – is updated. Although this forwarding capability is a built-in functionality of SRNS relocation, we choose instead to use the IP tunnel –used in FMIP– between the involved UARs.

Looking at the procedure in Fig. 2, the *Relocation Required* message comprises the L2 trigger at oUAR for transmitting a *PrRtAdv* to the MS. For the sake of simplicity, we will assume that the MS has activated one primary PDP context with possibly multiple secondary ones and therefore has been assigned one prefix for forming its global address(es). Before transmitting the Advertisement, the oUAR performs a target RNC – nUAR resolution for being able to fill the advertisement with the relevant information for the nUAR. At this point, we will assume that the oUAR knows the new prefix to be assigned to the MS for the corresponding PDP context, and we explain later how this is performed. The advertisement reveals to the MS information, such as the prefix used for address autoconfiguration and the nUAR's IP address. The message is sent unsolicited, in which case it acts as a network-initiated handover trigger, and not in response to a *Router Solicitation for Proxy* message from the MS where it would be required that the MS be aware of an identification of its attachment point (e.g. RNC).

The MS, in response to the PrRtAdv, formulates a new CoA based on the advertised prefix and sends a Fast Binding Update message to the oUAR. The source address is set to the oCoA while the alternate care-of address option is set to the nCoA. This message declares to the oUAR that it should forward data addressed to the oCoA towards the nCoA. The oUAR needs however to ensure that the new address can be used; note that address –prefix– assignment is performed by nUAR. Moreover, the oUAR will only start the forwarding if it is assured that resources have been reserved in the new path. Optionally, it may start buffering data, for example, in the case of lossless PDPs, in parallel to forwarding the data to the source RNC. The procedure continues with a *Handover Initiate* to the nUAR. This message requests from nUAR to verify the validity of the host's nCoA, and in the negative case the assignment of a new one. The nUAR may also create a host route entry for oCoA in case the nCoA cannot be accepted or assigned.

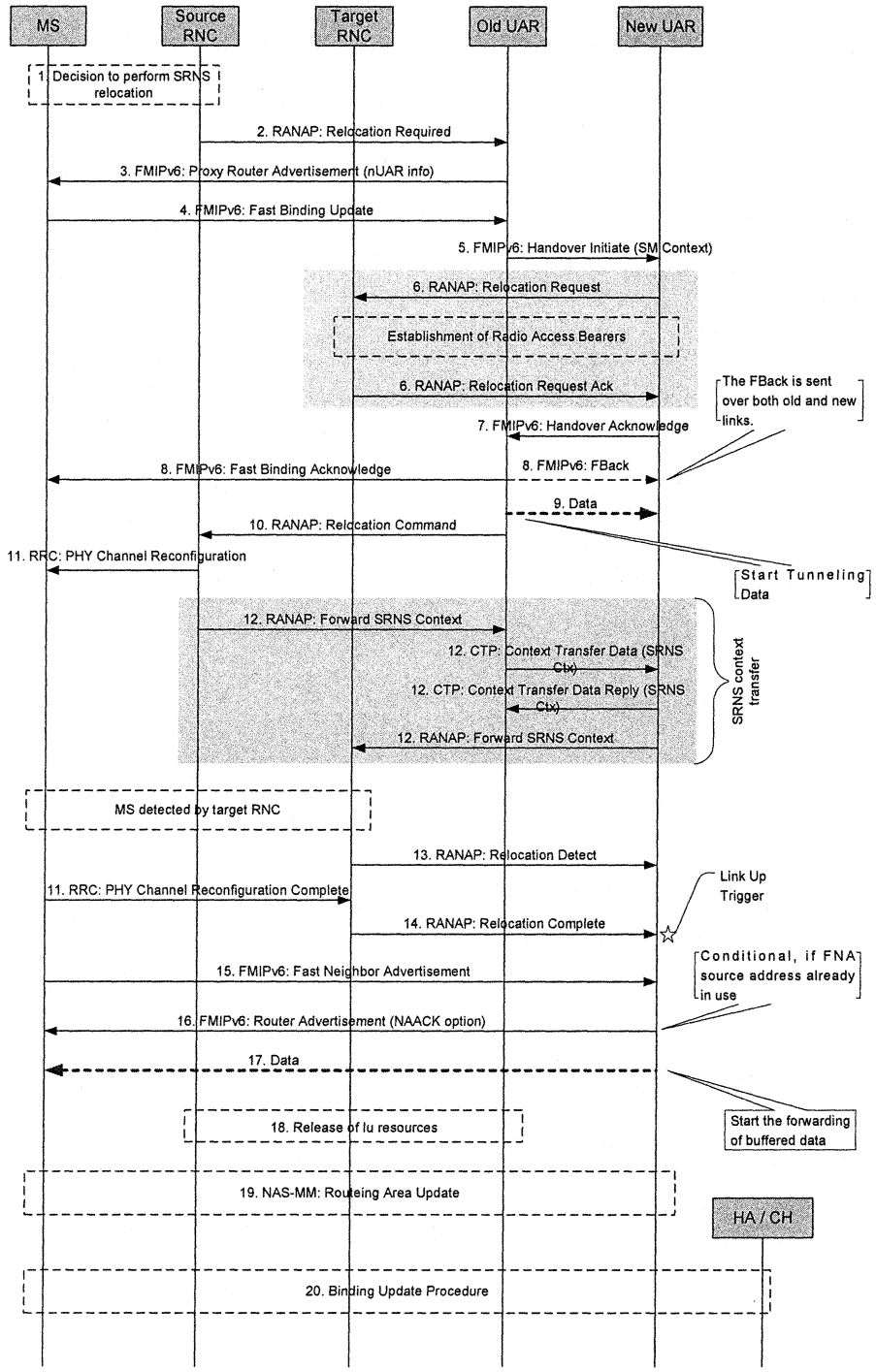

Figure 2. Fast Handoff scheme

The nUAR will only respond to HI after it relocates the Iu bearers. Note that the context information for these bearers is carried in the HI message, by means of a new option defined for this purpose. This mainly involves Session Management parameters for the MS's PDP contexts, such as QoS related information etc. The nUAR instructs the reservation of resources towards the new RNC by means of the *Relocation Request* message. In parallel, it checks whether the nCoA is valid and if not, it assigns a new address to the host – i.e., it only modifies the prefix –. After receiving the confirmation for the relocation of Iu bearers –*Relocation Request Acknowledge*–, it can respond to the oUAR with a *Handover Acknowledge* indicating that it accepts handover and that the forwarding of data can start. The new CoA is also included in the message if the nUAR does not accept the one assigned by the oUAR. Note here that the nUAR has to update the PDP contexts with the new address assigned to the host and also creates a proxy neighbor entry for this address and starts defending it.

The oUAR, after receiving the HAck, responds to the MS with a *FBack* where it indicates the nCoA – carried in the *alternate care-of address* option – to be used, if needed. This message is sent both over the old and new link. In the former case, it has as destination the oCoA while in the latter case the nCoA is used. Typically, this will be the first packet buffered at the nUAR for the MS's nCoA. The oUAR at this point instructs the source RNC to start the execution of the relocation – *Relocation Command* –, meaning that the source RNC from this moment will no longer be able to deliver packets to the host. In the meantime, the oUAR after receiving the *HAck* may also start the forwarding of data to the nCoA. After receiving the *Relocation Command*, the source RNC will instruct the MS to reconfigure itself and set the target RNC as its serving one. This is performed by means of a Radio Resource Control (RRC) message sent to the MS; the latter informs accordingly the target RNC, with an RRC message, that it has been reconfigured, which means that theoretically the target RNC is capable of delivering data to the MS. In parallel to the physical channel reconfiguration, the involved RNCs also exchange SRNS contexts for the supported Radio Access Bearers (RABs). This context information will allow the target RNC to support delivery order or lossless service for RABs, when this has been requested.

When the target RNC detects the MS, it notifies accordingly the nUAR – *Relocation Detect* – and starts the SRNC operation. This means that it may start processing upstream packets coming from the MS and forward them towards the nUAR. Moreover, it can start forwarding downstream packets, if available, towards the MS, given that the latter has reconfigured itself. The *Relocation Complete* comprises the link-up L2 trigger to the nUAR, which is informed of the MS's arrival on the new link. However, the forwarding of buffered packet towards the new RNC can not immediately start upon receipt of this message. Recall that the nUAR buffers packets addressed to nCoA after the request to do so by oUAR via the HI message.

At this point, we should examine when exactly packet forwarding to the MS is performed. If the MS has received the FBack message over the old link, then it has already updated its PDP context with the nCoA. In this way, packets buffered at the nUAR and addressed at the nCoA can be delivered over this PDP context. This is communicated to the nUAR by means of an FNA message, with the expected nCoA as source address, transmitted by the MS as soon as it is reconfigured. If a different address is identified by the nUAR (e.g. when FBack is not received over the old link and a non-valid address was assigned by oUAR), then the latter responds with a Router Advertisement with the *Neighbor Advertisement Acknowledge* option for indicating an alternate prefix to the MS. The MS subsequently forms a new CoA and modifies the respective PDP context. The nUAR may now start the forwarding of packets and also updates its neighbor cache. As expected, the MS will receive the FBack message on the new link. At this point, the MS receives packets addressed to its oCoA and being encapsulated by the oUAR towards its nCoA. In the reverse direction, packets have as source address the oCoA for not interrupting communication with the CHs, and are reverse tunneled towards the oUAR. Next, the MH may proceed with the MIPv6 Binding Update procedure with its HA and CHs.

3.2 Further Issues

3.2.1 Interface Identifier and Prefix Assignment

We shall analyze how the assignment of interface identifiers and prefixes may be performed by UARs in the context of an operator's IP domain. Each UAR is configured with an interface identifier which is used for its link-local address. Moreover, the UAR has to assign a unique interface identifier to each MS, for the latter to form its link-local address. The only requirement is that the UAR's identifier must not collide with the one assigned to the MS. Taking into consideration that the MS might roam to a new UAR within the IP domain, we propose that each UAR knows the Interface Ids belonging to all UARs and assigns Interface Ids to MSs that do not collide with the ones assigned to UARs. When the MS camps away from its serving UAR, it can continue using the same link-local address with the new UAR.

As for the Prefixes advertised by PrRtAdvs, we present an example of how they can be calculated by UARs, trying to avoid collisions with already assigned ones. Assuming that the operator's network, comprised of four UARs, has been assigned an n-bit prefix, then the operator assigns 2^{64-n-2} prefixes –links– to each UAR. Prefixes are identified as belonging to each UAR by means of the two bits *64-n-1* and *64-n-2*. When an UAR proxies an advertisement from a neighboring router, it simply sets these 2 bits for the advertised prefix to belong to the router being proxied. Certainly, collisions

may arise, if this prefix is already assigned to a MS being served by the neighboring router. However, we can avoid collisions if neighboring ARs are configured to start assigning different prefixes to MSs – e.g. one starts from the low order bits and the other from the high order bits –. In areas where roaming between UARs is a frequent event, the operator may even resort to configuring UARs to set non-overlapping bits in their prefixes – i.e. one sets only the low bits while the other sets only the high bits.

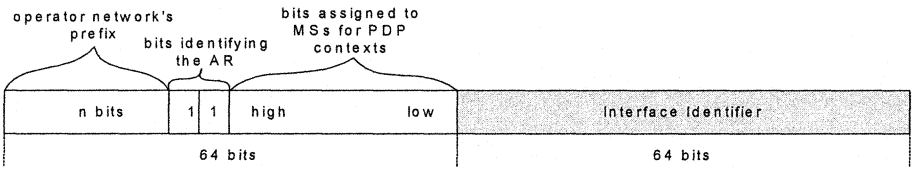

Figure 3. Prefix Assignment to MSs

3.2.2 Context Transfer

During the SRNS relocation procedure in UMTS, context transfer is performed at two levels of the architecture; between SGSNs –MM, PDP contexts– and between RNCs –SRNS contexts–. We attempt to identify which of these contexts are still meaningful within the new architecture. MM and PDP contexts are kept unchanged and are transferred between UARs by means of an extension to the HI message, as indicated earlier. As for the SRNS contexts, their main parameters are the GTP-U sequence numbers (SNs) next to be transmitted in both directions and the PDCP[11] (Packet Data Convergence Protocol) SNs to be used for data transmission from and to the MS. This information refers to each Radio Access Bearer –corresponding to a PDP– activated for the MS. Note that GTP-U packet sequencing is used when 'delivery order' is requested. In this case, the GGSN and RNC GTP-U protocol entities have to maintain the sequence of GTP-U packets transmitted in both directions. In our architecture, this requirement is relaxed due to the fact that *both* GTP-U entities are initialized during handoff – in UMTS, the GGSN GTP-U entity continues running – and therefore there is no problem in starting sequencing from scratch[3]. The PDCP SNs, on the other hand, are used when lossless PDCP service is requested. The target RNC needs this information in order to synchronize with the MS for the next PDCP packet expected in the uplink and downlink direction for each lossless radio bearer. Since the concerned RANs are not interconnected, this context transfer follows a path via the involved UARs. As an option, we can apply the IETF Context Transfer Protocol at the inter-UAR interface for transferring SRNS context.

[3] Delivery order is not required for IPv6 PDPs and it is not of our concern in the context of the proposed fast handoff procedure.

4. CONCLUSIONS

We have presented an evolved UMTS architecture, where the GSNs have been integrated to form one node, the UMTS Access Router, and the PS core network has been replaced by a fully IP-compatible backbone. Mobility in the core network is handled by Mobile IPv6, which has replaced the heavy GTP operation, while fast handoff across UARs is also supported by means of the IETF Fast MIPv6 protocol. This feature is particularly important in the cases of not interconnected RANs, e.g. UTRANs where the Iur interface is not available, where soft handoffs cannot take place. We have detailed the fast handoff procedure, while trying to identify the proper points of interaction between the SRNS relocation procedure and the Fast MIPv6 operation. Future work includes the execution of simulations with varying topologies, number of hosts, user moving patterns and active sessions regarding their QoS needs in order to evaluate the protocol's performance from both user and network perspectives.

REFERENCES

1. Jun-Zhao Sun et al., "Features in Future: 4G Visions From a Technical Perspective", in proceedings of IEEE Globecom'01, vol. 6, 2001.
2. Fabio M. Chiussi et al., "Mobility Management in Third Generation All-IP Networks", IEEE Communications Magazine, vol. 40, no. 9, September 2002, pp. 124-135.
3. D. Johnson et al., "Mobility Support in IPv6", Internet Draft, draft-ietf-mobileip-ipv6-21.txt, August 2003.
4. Rajeev Koodli et al., "Fast Handovers for Mobile IPv6", Internet Draft, draft-ietf-mipshop-fast-mipv6-01.txt, January 2004.
5. 3GPP TS 23.060 v6.3.0, "General Packet Radio Service (GPRS); Service description; Stage 2", December 2003.
6. S. Thomson et al., "IPv6 Stateless Address Autoconfiguration", Internet Draft, draft-ietf-ipv6-rfc2462bis-00.txt, February 2004.
7. T. Narten et al., "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.
8. 3GPP TS 29.061 v5.8.0, "Interworking between the Public Land Mobile Network (PLMN) supporting packet based services and Packet Data Networks (PDN)", December 2003.
9. M. Wasserman et al., "Recommendations for IPv6 in Third Generation Partnership Project (3GPP) Standards", RFC 3314, September 2002.
10. 3GPP TS 25.413 v6.1.0, "UTRAN Iu interface RANAP signalling", March 2004.
11. 3GPP TS 25.323 v6.0.0, "Packet Data Convergence Protocol (PDCP) Specification", December 2003.