A SIP based concept for NAS signaling in combination with an all-IP core network architecture

Christoph Pollak[1], Seoug Hoon Oh[2], Arnoldo Giralda[3], Vaggelis Nikas[4], Georgios Lioudakis[4], Jose Manuel Sanchez[5]

[1]Siemens AG Österreich, Autokaderstraße 24, 1210 Wien, Austria
[2] ComNets, RWTH Aachen, Kopernikusstr. 16, 52074 Aachen, Germany
[3] Telespazio Via Tiburtina 965 00156 Rome, Italy
[4] Ericsson Hellas, Zeppou 33 str., 16675 Athens, Greece
5 Integrasys, Esquilo 1, 28230 Las Rozas-Madrid, Spain

ABSTRACT

In this paper a novel concept for NAS signaling, which replaces the currently standardised GPRS mobility and session management by SIP signaling, is presented. In order to fulfill the functionalities covered by current NAS protocols (e.g. to determine the current location of the terminal within the access network or to manage the assignment of bearers), extended or additional fields will be used in the SIP messages, being denoted as SIP_{RAN} (indicating that they contain information being only relevant in within the radio access network) The concept aims at the avoidance of the repetitive transmission of information serving the same purpose (like User Identity, QoS parameters etc.) in order to save capacity on the used channels. In this paper, we show how to introduce the SIP_{RAN} protocol to be run between the mobile terminal and the network node between access network and core network (being denoted as RASN – Radio Access Support Node) in order to on the one hand covering the currently standardised features of session and mobility management while reducing the expected signaling traffic on the other.

I. INTRODUCTION

Due to the general interest to increase the usage of the IP protocols suite in wireless networks and to provide both real time as well as non-real-time services via the packet switched domain, intense research activities are performed in order to determine the optimal network architecture, a convenient distribution of functions and to identify the protocols or protocol extensions for this purpose.

Current 3GPP standardisation provides the following two instances for the control of PS services (besides the functions provided by the lower layer protocols, which are not in the focus of this paper):

- Bearer signalling between TE, SGSN, GGSN and HSS/HLR (being controlled mainly by the so-called NAS [Non Access Stratum] protocols), this assigns or modifies bearers with respect to the desired QoS of the user and manages AAA issues, as well as the provision of mobility related functions allowing a subscriber to change the terminal (personal mobility) or to move free throughout the coverage area of the PLMN. [1][2]
- SIP protocol for PS services like VoIP, Presence, etc. : The controlling instance for these service is the IMS (IP Multimedia Subsystem), which also performs AAA functions, negotiates QoS parameters, and provides personal mobility, i.e. to use different devices while being reachable by always the same identifier [8][9]

Furthermore, current 3GPP Architecture uses the GGSN as mobility anchor point by requiring the use of special equipment and protocols to be used within the core network of the PLNM [3][4][5]. As current interest has focused on the establishment of more flexible architectures and on multi-homing scenarios in order to fully exploit the advantages of the expected co-existence of different access technologies, new concepts have to be developed in next generation mobile networks to achieve these aims.

Considering the concepts of application layer mobility, which aim at the provision of terminal mobility by applying SIP, we consider it possible to enhance the current SIP standard in order to allow the creation of a purely SIP based solution for covering the functions of both, current NAS as well as SIP protocols. [6][7]

As a consequence, the concept being pursued in this paper is based on the following principles:

- Concentrate the non-native-IP part of the PLMN core network within one node. (Which in the following will be called “RASN” – Radio Access support Node)
- Avoid any non-singular allocation of functions as far as possible and do not send (signalling) information serving the same purpose/containing the same IEs more than once.

- Achieve the same set of features for the user as it is now provided in systems being standardised in 3GPP

See figure 1 depicting the network architecture being used as the baseline for our concept:

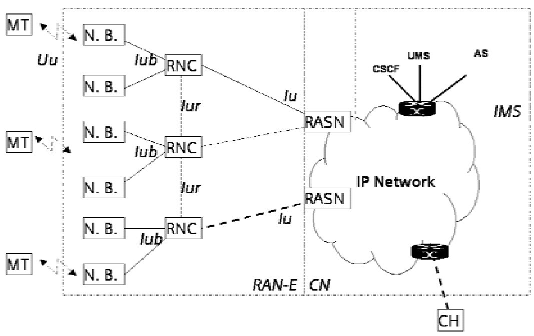

Figure 1; Underlying Network Architecture

In particular, we propose a smooth evolution to the standard UMTS Core Network, by integrating the GPRS Supporting Nodes (GSNs) into the RASN. In this new architecture, the SIP protocol is adopted and further extended, so as to replace Mobility and Session Management procedures, currently undertaken by the legacy UMTS NAS protocols. The extended SIP messages are used by both, the bearer and the application layer control functions.

II. MAPPING OF MESSAGES

If SIP shall cover the functions of the NAS protocols, It must be clarified, which SIP methods should be mapped to corresponding NAS messages. The following table gives an overview on how it is envisaged to achieve this.

3GPP SM/MM messages	**Possible SIP message**
GPRS Attach Procedure	**SIP Registration procedure**
ATTACH REQUEST	REGISTER
ATTACH ACCEPT	200 OK
ATTACH COMPLETE	INFO (related to TMSI procedure) or ACK[1]
ATTACH REJECT	e.g. 401 unauthorised
P-TMSI reallocation	**T-URI reallocation**
P-TMSI REALLOCATION COMMAND	SIP-NOTIFY
P-TMSI REALLOCATION COMPLETE (up)	200 OK
Authentication and ciphering	
AUTHENTICATION AND CIPHERING REQUEST	SIP-NOTIFY/401 unauthorised
AUTHENTICATION AND CIPHERING RESPONSE	SIP-REGISTER
AUTHENTICATION AND CIPHERING REJECT	Response codes to REGISTER
AUTHENTICATION AND CIPHERING FAILURE)	Response codes to REGISTER
GPRS Detach	**SIP Deregistration Procedure**
DETACH REQUEST (uplink)	REGISTER, exp.time =0
DETACH ACCEPT	200 OK
DETACH REQUEST (downlink)	REGISTER, exp.time =0 (Network initiated deregistration)
DETACH ACCEPT	200 OK
RA Update	**Register Procedure**
ROUTING AREA UPDATE REQUEST	REGISTER, e.g. contact field incl. RAI
ROUTING AREA UPDATE ACCEPT	200 OK
ROUTING AREA UPDATE COMPLETE	INFO or ACK
ROUTING AREA UPDATE REJECT	Response codes to REGISTER

Table 1 Mapping of SIP to NAS messages

[1] The current SIP standardisation foresees to use the ACK request only for session setup (INVITE). As a consequence this would introduce a further extension of SIP.

III. FUNCTIONAL DESIGN OF SIP_{RAN}

Due to the envisaged allocation of functions, it seems reasonable to use the well-established AS (Access Stratum) protocols as they are [10] as transport layer for SIP messages, which will be enhanced in order to provide the functionalities of the NAS protocols. As a consequence, SIP messages will be sent via the control plane of UMTS, not the user plane as it is now the case. In order to maintain the full functional range being now covered by NAS, additional IEs will have to be defined in the SIP header which shall only be valid between the terminal and the RASN, thereby the specifics of the Access Technology will be encapsulated from all nodes beyond the RASN, allowing a pure IP based core network.

In order to provide full terminal mobility in the proposed system, application layer mobility schemes will be applied using a SIP re-REGISTRATION procedure in cases of discontinuous inter-RASN terminal mobility and a re-INVITE procedure for continuous terminal mobility between different serving RASNs. In order to achieve maximum seamlessness of the relocation, the UMTS I_{ur} interface will be used, which allows to use the old RASN while already being registered with the new RASN, thereby, for a short period of time, the subscriber will have to accept two IP addresses and has a RAB (Radio Access Bearer) in both RASNs' subnets. This allows to cover the intermediate period in which the re-INVITE messages is processed between the peers.

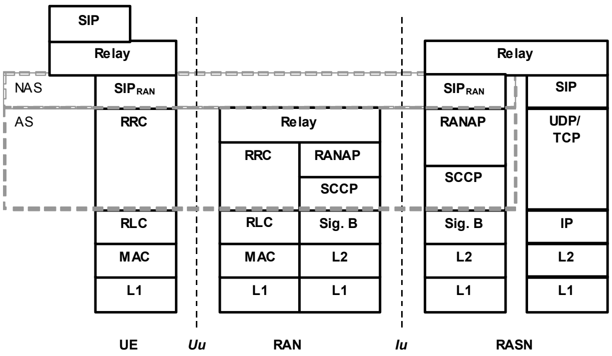

Figure 1 UMTS Control plane in the SIP_{RAN} concept

Figure 1 depicts the protocol stack for the control plane as they are envisaged in the SAILOR project's SIP_{RAN} approach. Except for a very small modification (regarding the notification of the terminal about the necessity to register in a new RASN's access network) the UMTS Access stratum protocols remains as they are currently standardised [10] [11] [12] The SIP-based NAS messages are to be carried in RRC PDUs over the so-called "signalling bearers" at the radio interface. In URAN, the handling of radio signalling bearers is quite similar to that of "radio data bearers" used to transport end-user IP traffic. In fact, the only difference consists of the PDCP layer (whose main function is packet compression), which is absent in the case of signalling bearers. Therefore, the same rates and SDU sizes used for data bearers can be used for signalling bearers conveying SIP messages, and no modifications are needed on the current radio stack to support the rate and format required by the SIP-based NAS concept.

Apart from this, in order to efficiently exploit the combined satellite-terrestrial scenario of SAILOR, a proper mechanism is devised so as to setup (network initiated) PDP contexts either in the satellite AN or in the terrestrial AN without actually existing a signalling connection between such AN and the CN. Thus, SAILOR terminals support parallel data bearers on both ANs having a single connection with the CN through one of the ANs.

The separation of control plane and user plane vanishes in the core network (i.e. upwards the RASN), which is one of the prerequisites for the access unawareness of the latter.

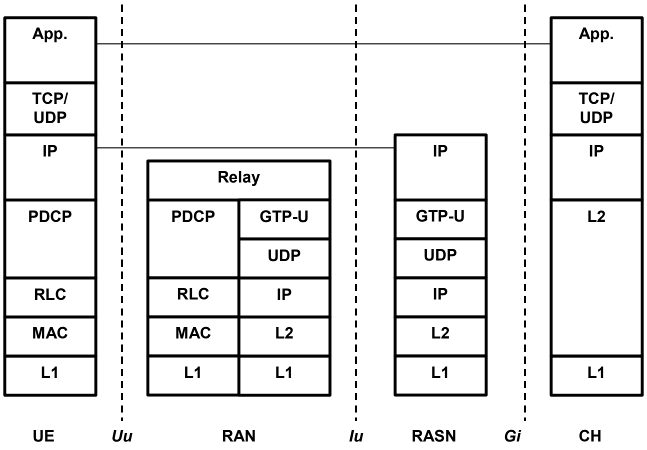

Figure 2 UMTS User plane in the SIP_{RAN} concept

Figure 2 depicts the UMTS user plane as it is envisaged in the SAILOR project's SIP_{RAN} concept. Because of SIP being transported via the control plane, only the user data of SIP sessions and that of other applications (like http) are transported via the user plane. For a bearer establishment for any non-SIP application (like http, ftp etc.), we propose to send a SIP: INVITE message to the RASN, indicating the requested QoS for the following applications. This will result in the assignment of a RAB, which is not specific to the session it is going to be used for.

See the following figures for an illustration of the exchanges of messages required for registering the subscriber with PLMN / IMS and for establishing a SIP based VoIP call. Note that in figures only messages impacted by the new concept are depicted , i.e. RRC signalling is not shown, because it will not change. Also, "ringing" and "trying" responses are not shown because they serve the same purpose and will be sent in the same way in both concepts. The dots at RASN and IMS ins Figure 4 indicate, that the massages are processed in these nodes.

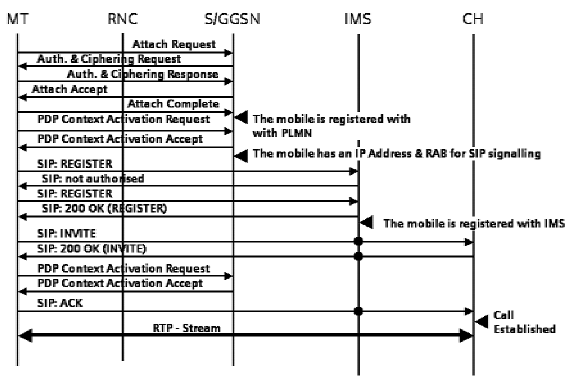

Figure 3 Message Sequence Charts forRegistration & Session initiation in current standardisation

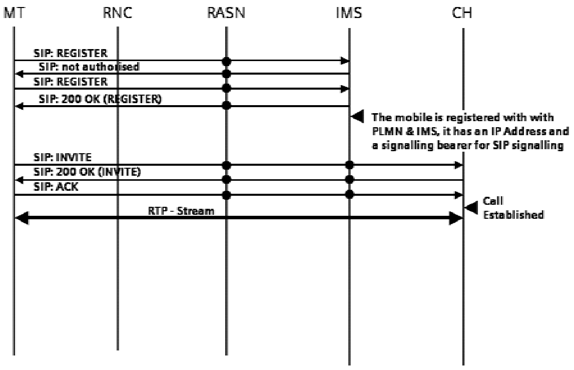

Figure 4 Message Sequence Charts for Registration and Session initiation in SIP_{RAN} concept

IV. COMPARISON OF AMOUNT OF SIGNALLING DATA

The following tables give comparative estimates about the amount of signalling data being transmitted via the air interface in both concepts:

Registration and session initiation	Current UMTS standardisation	SIP_{RAN}
No. of Messages	7	16
Sum of Bytes	2400	3000

Table 1 Comparison of signalling for registration

Discontinuous inter RASN (GGSN) mobility event	Current RAN/IMS standardisation	SIP_{RAN}
No. of Messages	7	3
Sum of Bytes	650	800

Table 2 Comparison of signalling for inter-RASN events

Mobile to mobile VoIP call	Current RAN/IMS standardisation	SIP_{RAN}
No. of Messages	40	18
Sum of Bytes	6900	5000

Table 3 Comparison of signalling for VoIP call between two mobile subscribers

The results in the tables above seem surprisingly because current NAS protocols have been optimised in order to reduce the signalling load at the air interface, while SIP signalling was not. Considering the fact that the SIP_{RAN} concept is based on the assumption, that IMS SIP signalling will in the future be the general control mechanism for fixed and mobile real time communication services, the results become plausible: NAS content was included in application layer signalling, that occurs regardless of the underlying access technology. Furthermore, this concept may be enhanced by the usage of compression methods for SIP [13][14][15]. This would make the new approach even more attractive.

V. CONCLUSION

In this paper, it was shown how the mobility and session management functionalities of SIP can, in principle, be enhanced by those provided by the current NAS protocols and that the replacement of the latter with the so-obtained protocols yields a significant reduction of the signalling load as well as the possibility to use native IP routing in the core network. It can be expected that the functionality being experienced by the user will not suffer from the introduction of the envisaged approach. Moreover, the concept will not only contribute to a more efficient use of network resources, but also open the way for the seamless and simultaneous use of different access technologies on one terminal (e.g. by developing an analogous SIP - enhancement for W-LAN, which could be called SIP_{WLAN} and by introducing a SIP/SIP_{WLAN} relay “below” the SIP application). Such a multi-homing scenario would dramatically increase the average available bandwidth by fully exploiting any available access network as well as allow an application specific use of access technologies (e.g. one might prefer to exchange security – sensitive data like those of an e-/m-commerce application via UMTS while to read web-news via the cheaper W-LAN connection.)

VI. OUTLOOK

The mechanism being proposed in this paper will be exemplarily implemented in a demonstrator system, the Radio access part of which will consider the effects of terrestrial and satellite UMTS radio transmission (the emulation of the radio link will be based on the results obtained from the EC-IST Projects “VIRTUOUS” and ”FUTURE”). The trials using this demonstrator are expected to yield a proof-of-concept for the solution proposed in this paper and will allow to verify the feasibility and the benefits of this solution.

VII. ACKNOWLEDGEMENTS

The work reported in this paper was partially funded by the European Commission under contract no.: IST 2001 – 37266, project acronym: SAILOR.

VIII. REFERENCES

[1] 3GPP TS 23.060 General Packet Radio Service (GPRS); Service description; Stage 2

[2] 3GPP TS 24.008 V5.6.0 (2002-12) “Mobile radio interface Layer 3 specification; Core network protocols

[3] 3GPP TS 23.221 V6.1.0 (2003-09) Technical Specification Group Services and System Aspects; Architectural Requirements

[4] 3GPP TS 23.101 V4.0.0 (2001-04) Technical Specification Group Services and System Aspects General UMTS Architecture

[5] 3GPP TS 23.002 V5.8.0 (2002-09) Technical Specification Group Services and Systems Aspects; Network architecture

[6] F. Vakil, A. Dutta, J-C. Chen, H. Schulzrinne: "Mobility Management in a SIP Environment Requirements, Functions and Issues"draft-itsumo-sipping-mobility-req-00.txt

[7] F. Vakil, A. Dutta, M. Tauil, S. Baba, N. Nakajima, H. Schulzrinne: "Supporting Service Mobility with SIP";draft-itsumo-sipping-mobility-service-00.txt

[8] 3GPP TS 24.228 V5.3.0 (2002-12), "Signalling flows for the IP multimedia call control based on SIP and SDP".

[9] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler: rfc 3261 "SIP: Session Initiation Protocol"

[10] 3GPP TS 25.410: UTRAN Iu Interface: General Aspects and Principles

[11] 3GPP TS 25.301: Radio Interface Protocol Architecture

[12] 3G TS 25.401:UTRAN Overall Description

[13] R. Price et al. rfc3320, "Signaling Compression (SigComp)", January 2003

[14] M. Garcia-Martin et al., rfc4385, "The Session Initiation Protocol (SIP) and Session Description Protocol (SDP) Static Dictionary for Signaling Compression (SigComp)", February 2003

[15] G. Camarillo, rfc 3486, "Compressing the Session Initiation Protocol (SIP)", February 2003