Enabling Dynamic Service Delivery in the 3GPP Evolved Packet Core

Marius Corici, Jens Fiedler, Ancuta Onofrei, Dragos Vingarzan
Fraunhofer FOKUS Institute
Berlin, Germany
{marius-iulian.corici, jens.fiedler, andreea.ancuta.onofrei, dragos.vingarzan}@fokus.fraunhofer.de

Abstract—**With the deployment of new wireless access technologies, like LTE or WiMAX which complement the existing accesses such as UMTS or WiFi, the mobile devices are able to establish services having different requirements for network capabilities. However, the services deployed are lacking an efficient mechanism of adaptation to the momentary connectivity conditions which deters the dynamic service delivery according to the momentary preferences of the user's applications. This paper introduces a new mechanism for the adaptation of the services to the momentary network conditions which includes as part of the vertical handover procedures a resource reservation decision according to the preferences of the mobile device and the capabilities of adapatation of the application as advertised during service establishment, enabling a more dynamic service delivery without requiring resource re-reservation procedures. Furthermore, the novel concept is exemplified on the 3GPP Evolved Packet Core and evaluated using the Fraunhofer OpenEPC toolkit implementation.**

Evolved Packet Core; Heterogeneous Wireless Environment; Dynamic Service Delivery;

I. Introduction

With the development of a multitude of novel wireless access technologies (e.g. LTE, WiMAX, Femtocell technologies etc.) apart of the already existing technologies (e.g. UMTS, WiFi etc.), a growing number of deployments will be able to offer connectivity over more than one access network in specific locations. These dense wireless environments are characterized by an extended variety of connectivity parameters offered by the various accesses, which translates for the mobile users into different service delivery parameters such as various delays and throughput.

On the other side, it is expected that the mobile devices will be able to connect and to establish services over more than one access technology. Therefore, for the same service negotiated between the mobile device and the service provider, depending on the access network through which the connectivity is enabled, the delivery highly varies as momentary operational cost.

The current service delivery mechanisms consider that the services are established with only one level of parameters (e.g. throughput). However, the mobile users, due to the variation in cost, require that the services have different parameters over the different access networks. For example, for a bulk file download the user may prefer to use high resources over a low cost access network such as WiFi and to use minimal or no resources over a high cost one such as UMTS.

However, the current handover and resource reservation mechanisms concentrate on the seamless delivery of the services, thus on maintaining as much as possible the same communication parameters when a handover occurs, not considering that the user may require their change.

The current wireless applications (e.g. video on demand, file transfer, real-time video sessions etc.) are considering transparent the connectivity of the mobile device. Thus, the mobile device has the responsibility to transmit notifications related to the resources that should be reserved for the specific services over the current access network. This implies that after a handover occurred, the mobile device has to re-negotiate the session parameters with the service provider and through this re-negotiation the novel level of resources is reserved. From the network perspective, this implies that the resource reservation procedures are executed twice: once for the handover and once for the service delivery adaptation.

In order to optimize the service delivery in case the mobile user requires a different level of resources to be reserved over the different access network, this paper presents a novel concept in which the user preferences for the service delivery over different access networks are pro-actively shared with the network infrastructure. Additional to this, the concept includes a fast adaptation of the resources reserved for specific applications in case a handover occurs without implying any additional communication with the mobile devices.

The procedures related to this new concept are exemplified using the 3GPP Evolved Packet Core (EPC) architecture as a reference all-IP multi-access core network which integrates various wireless access technologies. The concept here presented integrates in the EPC as an extension of the handover and of the policy based resource reservation functionality.

The remainder of this paper is organized as follows: Section II provides the background of the proposed method. Section III describes the general concept and then exemplified on the 3GPP Evolved Packet Core architecture. Section IV describes the proof-of concept testbed used and in Section V conclusions will be provided.

II. Background

The current state of the art separates the changes of the access network context from the resource reservation modification procedure, thus from the modified service delivery ([1], [2]). The resource reservation modification procedures are now triggered by the mobile devices either at the request of the user or automatically by re-negotiation of the service parameters with the service providers. When the access network context of the mobile device changes (e.g. a vertical handover to another access network with different characteristics), the mobile device, based on its own internal service delivery preferences set by the user has to initiate on its own the service modification procedures including the resource negotiation with the service provider and the resource reservation.

In the meantime the resource reservation infrastructure of the core network transfers the already existing context from the source network to the target access network, thus executes resource reservation procedures considering that the user always requires that the services are delivered with the same parameters.

As depicted in Figure 1, currently, the service modification procedures are following the same path with the ones for the service establishment. When the mobile device executes a handover, its service requirements may change as the target access network operational cost is different from the source access network one. Thus, after the handover occurs, including the reservation of the resources in the target access network as they were previously reserved over the source access network, the mobile device re-negotiates the service parameters with the service provider. The newly agreed level of resources is then transmitted to the wireless core network which then executes a new set of resource reservation procedures according to the new service profile and to the static subscription profile of the mobile endpoint. This automatic processing allows a fast adaptation to the wireless environment of the services. However it presumes that the services are adapted after the event of the network context change and presumes a complete renegotiation of the service parameters.

We proposed another approach on the service delivery adaptation ([13]), considering that the service provider includes all the service profiles in the initial reservation procedure and that upon a handover, the wireless network core selects the most suitable one for the service continuity. Although this solution enables the service adaptation according to the capabilities of the service provider and that the resource reservation procedures can be executed through only one procedure during the handover, it does not include in the decision the preferences of the user. Because of this, the service can be adapted to the parameters supported by the service platform, but not according to the current requirements of the mobile uses. This implies that the user may receive a service adapted to a specific access network for example with a modified throughput although it does not desire such a modification. For example, for a video service having two different delivery qualities, one user may require that when connected over UMTS, the service is delivered with the good quality while another user may require that it is delivered with a worse one, thus requiring that specific users are able to provide their preferences dynamically for each service session and communication context.

For exemplifying the new concept of this paper, the 3GPP Evolved Packet Core (EPC) ([1], [2]) was selected because of its standardized architecture enabling all-IP network convergence for various types of access networks including harmonized mobility management and resource reservation functionality.

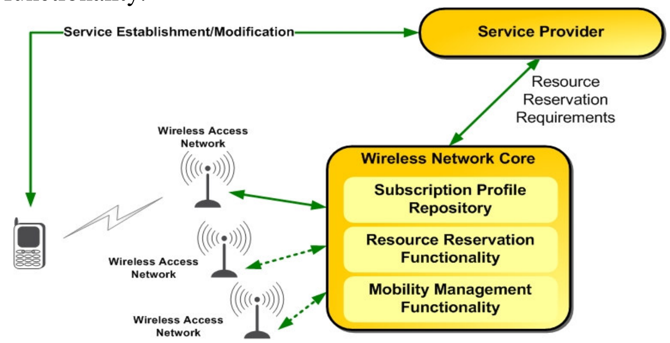

Figure 1 Conceptual Network Architecture

As depicted in Figure 2, EPC includes a counterpart mobile device denominated as User Endpoint (UE). EPC separates the control from the data plane. The data plane includes a set of gateways which enable specific resource reservation procedures for the various access networks such as the Serving GW (S-GW) for 3GPP accesses, the evolved Packet Data Gateway (ePDG) for the untrusted non-3GPP accesses (e.g. WiFi). The Packet Data Network Gateway (PDN GW) is a central gateway terminating the core network towards the IP domain. The PDN GW handles the mobility management anchoring functionality by forwarding the data traffic towards the current location of the UE. All gateways ensure the mobility and the resource reservation on the data path.

For maintaining subscription information on the UEs, EPC includes a Home Subscriber Server (HSS) used also for authentication and authorization and as part of the resource reservation procedures.

As central control entity, EPC uses the Policy and Charging Rules Function (PCRF) ([3], [4]) which makes policy based decisions for access control and for resource reservations based on the subscription profiles and on the information received from the Application Functions (AFs) which represent an abstraction of the service providers ([4]). These decisions are then enforced in the gateways over the Gx and Gxx interfaces ([9]).

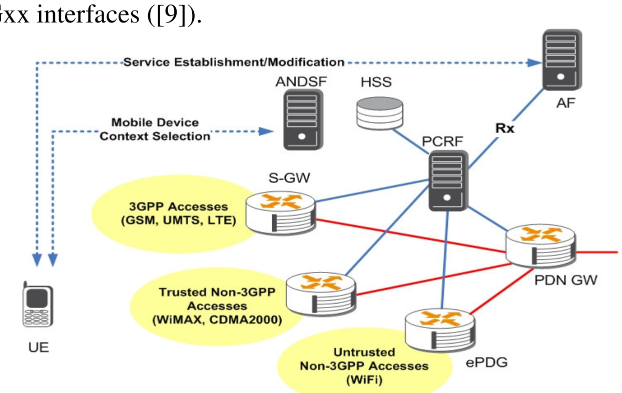

Figure 2 3GPP Evolved Packet Core

EPC recognized the need for communicating with the UEs for a more dynamic and better delivery of the services. For this, as part of the mobility management functionality, it includes an Access Network Discovery and Selection Function

(ANDSF) which is able to receive information on the current location of the UE on and its access network connection preferences and to transmit the operator indications in regard to the accesses that may be selected for a better service delivery.

For the EPC example, this article addresses the problem on how to receive the requirements of the user of the UE for the specific services depending on the access network over which they are delivered, how to integrate them into the research reservation procedures and how to notify the service provider on the new service profile which has to be delivered.

III. CONCEPT

This article proposes a new alternative method for the dynamic service delivery in which the multiple service profiles are received from the UE during service establishment. We will first introduce the approach and then show how to realize this in the EPC.

A. *Approach*

When a handover occurs, the service profile requested by the user of the UE is enforced and the service platform is notified as to be able to deliver the service according to the new parameters. As depicted in Figure 3, the method presumes a functional addition to the interface between the subscription profile storage and the UE and to the interface between the resource reservation functionality and the subscription profile storage.

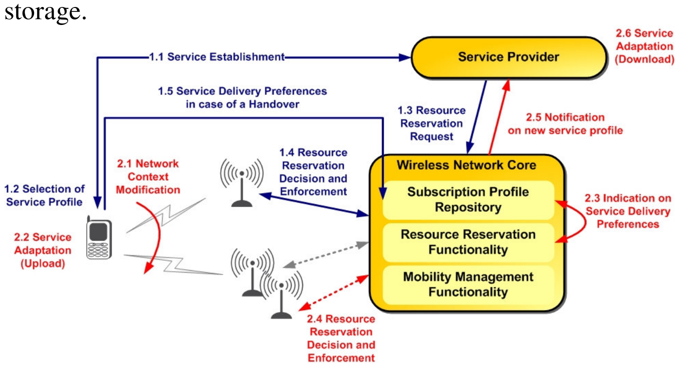

Figure 3 Concept

The state of the art service establishment procedure ([3]) includes the negotiation of the service profile between the UE and the service provider (Step 1.1), the selection by the UE of a service profile which is to be activated (Step 1.2), the request of the service provider for resource reservation (Step 1.3) and the resource reservation decision and enforcement according to the static user profile.

Using the extensions to the architecture, the service establishment procedure was enhanced to contain in Step 1.3, in parallel to the resource reservation request, a subscription for requests for changing the service profile is made to the resource reservation functionality. Compared with the state of the art procedure, in this solution, the services which can be delivered with multiple levels of resources have to subscribe to network related events. This subscription is considered optional in the case of current standards. If this subscription is not executed, then the solution falls back to the standard solution.

In Step 1.5, the UE sends to the subscription profile repository its preferences on the service profile to be enforced on the data path for different possible handover cases. This step represents an addition to the communication with the mobility management. If the UE does not support this step, then the solution also falls back to the standard solution.

Compared with the state of the art service establishment procedure ([3]), a communication with the subscription profile repository in the network is added through which the UE is able to transmit its preferences regarding the active session processing in case of a handover.

This additional operation can be executed in parallel with the other operations, thus it does not influence the duration of the service establishment.

When the context of the UE changes (Step 2.1) – e.g. due to a handover or a network congestion etc. – the service cannot be offered with the same session profile.

When the resource reservation functionality of the core network receives the notification on the context modification event, it uses the service profile requested from the subscription profile repository (Step 2.3) to make a policy based decision and to enforce it on the target access network (Step 2.4). This procedure is using the standardized mechanisms ([9]).

The service provider is notified on the level of resources that are currently reserved (Step 2.5) in order to adapt the active session to it (Step 2.6). This notification may happen before the new session profile is enforced in case the resources allocated are lower than the resources of the previous profile or it may happen after the enforcement in case the resources allocated are higher. This allows that the data lost during the session adaptation to be minimal. No notification has to be transmitted for the adaptation of the upload data traffic as the UE is aware that the network context was changed – i.e. the UE is the one that triggers a vertical handover.

B. *Mapping the Approach to the EPC*

In the EPC, the role of the resource reservation decision function is taken by the PCRF while the role of the service provider is taken by the generic Application Function (AF) functionality and the role of the Subscription Profile Repository by the HSS with a dynamic front-end to the UE.

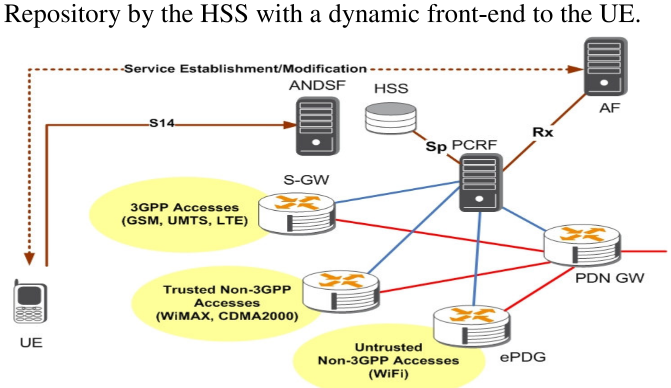

Figure 4: Concept translation to the EPC

In this context, the concept here presented is translated into an extension of the existing Rx interface ([4]) between the

PCRF and the AF, an extension of the decision functionality of the PCRF to request from the HSS the secondary service profiles through the Sp reference point ([3]) and to make new resource reservation decisions based on them when the context of the UE changes. As not to introduce any new interfaces, making the concept easy to be integrated into the overall architecture, the ANDSF is taking the role of the front-end for the communication of the service profile from the UE through an extended S14 interface ([10]) as depicted in Figure 4.

Multiple example procedures may be given for the service establishment and adaptation considering also exception cases. In order to showcase the concept introduced in this paper, the role of the AF is taken by a basic SIP video server similar to the testbed implementation described in the next section. The SIP video server is able to transmit downlink data streams to the UE using multiple levels of throughput e.g. 1Mbit, 500kbit, 300kbit, 50kbit etc.

As depicted in Figure 5, the example EPC setup deploys as mobility protocol Proxy Mobile IPv6 (PMIPv6) ([5], [6]). The User Equipment (UE) is already connected to the service through a 3GPP access network (e.g. UMTS) through the S-GW and through the PDN GW between which a PMIP tunnel is already established.

Step 1-6: A session establishment procedure is executed according to the standard implementation by SIP messages exchange, Diameter Authentication and Authorization Request and Answer (AAR and AAA) resource reservation request conformant to [4] and the policy based decision and enforcement through the IP-CAN session modification procedure according to [1], [2] and [3].

Step 7: Through the S14 reference point, the UE transmits the secondary resource profiles to be reserved in case a change in the network context happens. This operation presumes that session profiles which were not enforced are translated into the S14 Management Object extended format [11] and transmitted to the ANDSF front-end.

When the UE selects a WiFi access network, it executes the attachment procedures including the authentication and authorization and the IPsec tunnel establishment according to [2] and the procedures for establishing the new PMIP tunnel (Proxy Binding Update/Proxy Binding Acknowledgement messages). When the PDN GW receives the request, it initiates an IP-CAN session modification procedure according to [3] in which the PCRF is requested to make a decision on the resources that have to be reserved on the target access network. After this procedure the UE establishes its connectivity on the target access network – the IPsec and the PMIP tunnels are completed. During the decision, the PCRF requests the secondary session profile from the ANDSF containing the resources that are requested by the UE to be reserved according to the parameters of the target access network (Step 9). After the successful reservation of resources, a notification in form of a Diameter Re-Auth Request (RAR) [4] is transmitted to the service provider responded with a Diameter Re-Auth Answer (RAA) to the PCRF acknowledging that the session has to be adapted to the target access network (Step 10). The video server adapts the session to the correspondent parameters for the target access network and transmits the data stream accordingly (Step 11).

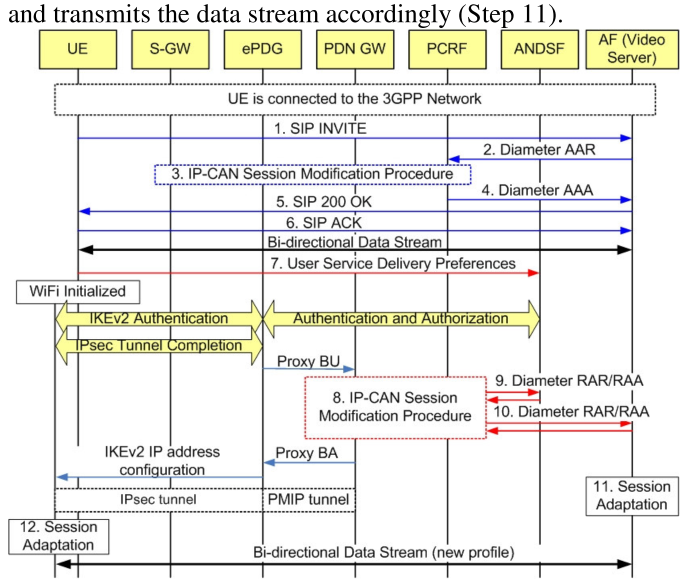

Figure 5 Sequence Diagram for Dynamic Service Delivery

In comparison to the here presented method, in the current 3GPP standards, after the UE connects to the target access network, the service is adapted using the SIP and Diameter signaling having the same messages exchanged as in Steps 1-6 containing the new session profile parameters. This presumes that the resources have to be reserved two times on the target access network, once for the service transfer and once for the service adaptation.

Using the functional extensions described here, the session can be adapted directly during handover without requiring a re-negotiation of the parameters between the UE and the service provider. This reduces the resource reservation procedures required in case of a vertical handover and thus reduces the delay of the service adaptation. As all the interfaces and the functionality here proposed is an addition to the existing standards, the optimization brought by this article can be easily integrated in the real deployments.

IV. Testbed, Results and Evaluation

To be able to evaluate the applicability of the proposed procedures in a real environment a testbed was developed based on the Fraunhofer FOKUS OpenEPC toolkit ([7], [11]). OpenEPC includes the basis functionality for mobility and resource reservations and subscription profile based decisions. It also includes a minimal ANDSF implementation for automated network discovery and selection extended with the information required by the solution of this article.

A mobile device able to connect to both a local WiFi access network and to the public UMTS network was deployed. For the establishment of the connection over the public operator an OpenVPN server is deployed in the S-GW. Through the tunnel it establishes, the data packets are transparently exchanged.

The testbed includes three gateways: a S-GW for UMTS, an ePDG for WiFi and a PDN GW using as mobility management protocol OpenEPC implementation of Proxy Mobile IP and as resource reservation functionality, a Bearer Binding and Event Rules Function in the S-GW and ePDG and an Policy and

Charging Enforcement Function in the PDN GW which enforce the policy decisions of the PCRF on the data path.

As example service in the testbed, the Fraunhofer FOKUS Media Delivery Function ([8]) was deployed. It is able to establish video sessions with different throughput signaled over SIP. A counterpart minimal SIP client was integrated in the mobile device able to deliver adaptable service.

The communication between the different network entities is based on the FOKUS Diameter Stack for Linux and on the FOKUS Proxy Mobile IPv4 implementation.

The proof of concept scenario represents a simplification of the one presented in Section III. It does not include authentication and authorization procedures. The connection to the WiFi access network is done as in the case of trusted non-3GPP access networks. The throughput of the video service was selected to be at 60kbit for the UMTS connectivity and at 300kbit for the WiFi.

Upon service establishment, the PCRF receives from the AF the resource requirements according to the service profile together with a subscription for UE connectivity events. The resources are reserved using the standard procedures. When the service is established, the mobile device transmits the other possible session profiles to the ANDSF. The ANDSF stores these profiles bound to the IMSI of the UE. When a handover occurs, the PCRF requests from the ANDSF the other session profiles for the specified IMSI. If no such profile is present, then the resource reservations are transferred from the source access network to the target access network, as in the standard processing. If the profiles are available, the PCRF makes a policy based decision and enforces the novel profile on the data path entities. Then the SIP server is notified on the new level of resources reserved and it can adapt the service.

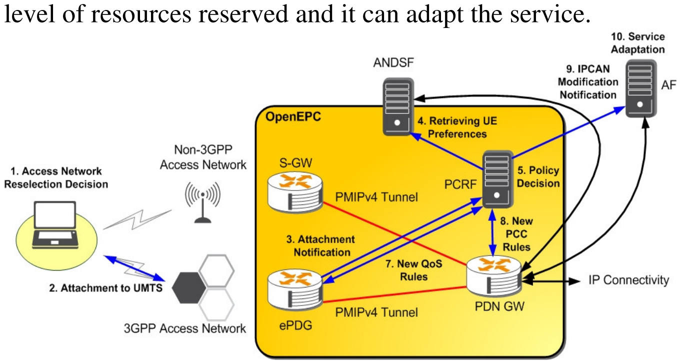

Figure 6 Testbed Architecture

Based on this minimal scenario, the resource reservation procedures are executed only once for both the handover and the service adaptation without requiring new signaling between the mobile device and the video server. This reduction in processing was possible at the cost of the communication between the mobile device and the ANDSF at the service establishment and of the communication between the PCRF and ANDSF on the policy based decision for the resource reservation during handover.

As testbed measurement of the procedures highly depends on the number of mobile devices and as none of the new introduced operations are increasing the handover delay – being executed in parallel with the time critical operations, the solution was evaluated only as proof of concept.

V. Conclusions

This article presented a novel concept for dynamic service delivery which does not require end-to-end re-negotiation of the service parameters and include in the resource reservation procedures during handover the requirements received from the mobile device for service adaptation. This novel approach complements the approach from [13] in which the preferences of the service provider are introduced in the resource reservation decision. These two solutions together enable the development of novel services tailored for mobile communication which can be adapted according to the requirements of the users and to the service capabilities. This concept of service adaptation according to the requirements of the user brings a more flexible approach to the resources that have to be reserved on the data path enabling a more dynamic adaptation which is required by the mass broadband communication.

The concept was exemplified on the 3GPP Evolved Packet Core and evaluated on a proof of concept implementation using the Fraunhofer OpenEPC toolkit, providing an adaptable service delivery. However, the testing was done having only one mobile device connected to the test platform and the solution has to be further developed to be able to perceive its effect in a more real environment in which a multitude of mobile devices are used.

References

[1] 3GPP TS 23.401, “General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access”, December 2009;

[2] 3GPP TS 23.402, “Architecture enhancements for non-3GPP accesses”, December 2009;

[3] 3GPP TS 23.203, “Policy and Charging Control Architecture”, version December 2009;

[4] 3GPP TS 29.214, “Policy and Charging Control over the Rx reference point”, December 2009;

[5] IETF RFC 5213, “Proxy Mobile IPv6”, www.ietf.org, August 2008;

[6] IETF “IPv4 Support for Proxy Mobile IPv6”, draft-ietf-netlmm-pmip6-ipv4-support-17.txt, work in progress, September 2009

[7] Fraunhofer FOKUS OpenEPC toolkit, http://www.openepc.net/

[8] A. Al-Hezmi, J. J. Pallares, C. Riede, T. Magedanz: "Evolving the Convergence of Telecommunication and TV Services over NGN", International Journal of Digital Multimedia Broadcasting Volume(2008);

[9] 3GPP TS 29.212 “Policy and charging control over Gx reference point”, December 2009

[10] 3GPP TS 24.302 “Access to the Evolved Packet Core (EPC) via non-3GPP access networks”, www.3gpp.org, version 9.1.1, December 2009;

[11] 3GPP TS 24.312 “Access Network Discovery and Selection Function (ANDSF) Management Object (MO)”, www.3gpp.org, version 9.0.0, December 2009;

[12] Corici M., Gouveia F., Magedanz T., Vingarzan D., “OpenEPC: A Technical Infrastructure for Early Prototyping of Next Generation Mobile Network Testbeds”, Tridentcom 2010, Berlin, Germany.

[13] Corici M., Magedanz T., Vingarzan D., Weik P., *„Enabling Ambient Aware Service Delivery in Heterogeneous Wireless Environments”,* IEEE Globecom 2010, Miami, USA;