

(51) International Patent Classification:
H04W 28/08 (2009.01)(21) International Application Number:
PCT/EP2011/000101(22) International Filing Date:
12 January 2011 (12.01.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10000221.1 12 January 2010 (12.01.2010) EP(71) Applicant (for all designated States except US): **NEC EUROPE LTD.** [DE/DE]; Kurfürsten-Anlage 36, 69115 Heidelberg (DE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **TALEB, Tarik** [JP/DE]; Konstanzerstrasse 45, 69126 Heidelberg (DE). **LOUREIRO, Paulo Ferrer** [PT/DE]; Bergheimer Strasse 114/3, 69115 Heidelberg (DE). **SCHMID, Stefan** [DE/DE]; In der Neckarhelle 30, 69118 Heidelberg (DE).(74) Agent: **ULLRICH & NAUMANN**; Schneidmühlstrasse 21, 69115 Heidelberg (DE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

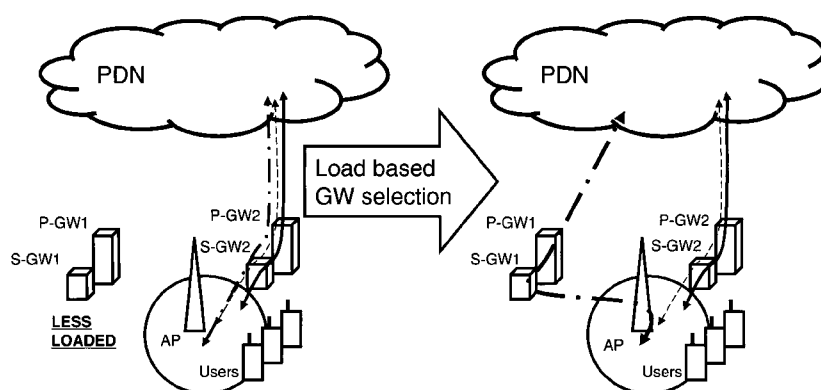
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: METHOD AND SYSTEM FOR SUPPORTING NETWORK-BASED MOBILITY MANAGEMENT OF A MOBILE TERMINAL

**Fig. 1**

(57) **Abstract:** A method for supporting network-based mobility management of a mobile terminal, wherein said mobile terminal is served by a radio access node and wherein a mobility management function is provided that decides on which gateway to handle said mobile terminal and on the respective handover execution decisions, is characterized in the steps of gathering at said mobility management function information on the operational performance conditions and/or technical specifications of radio access nodes and/or network entities involved in or relevant for said mobile terminal's communication performance, and considering said information for gateway selection and/or handover execution decisions. Furthermore, a corresponding system is disclosed.

METHOD AND SYSTEM FOR SUPPORTING NETWORK-BASED MOBILITY MANAGEMENT OF A MOBILE TERMINAL

The present invention relates to a method for supporting network-based mobility management of a mobile terminal, wherein said mobile terminal is served by a radio access node and wherein a mobility management function is provided that decides on which gateway to handle said mobile terminal and on the respective handover execution decisions.

Furthermore, the present invention relates to a system for supporting network-based mobility management of a mobile terminal, including at least one radio access node that serves said mobile terminal, and a mobility management function being configured to decide on which gateway to handle said mobile terminal and on the respective handover execution decisions.

Typically, in network-based mobility networks, a mobility management function decides which gateway shall handle a specific mobile terminal, e.g. a User Equipment (UE) in the terminology of Evolved Packet Systems (EPS). For instance, in EPS networks such mobility management function, which is responsible for handover execution decisions, resides in a network component denoted MME (Mobility Management Entity) or SGSN (Serving GPRS Support Node in case of 3G networks. On the other hand, in PMIPv6 (Proxy Mobile IPv6) the corresponding entity is denoted MAG (Mobile Access Gateway).

Generally, handover execution decisions can be based on a number of parameters, such as loads of involved network entities. Indeed, there are many situations when an operator may want to distribute load among involved network entities. In this context, Fig. 1 illustrates a scenario where geographically nearby P/S-GWs (PDN-Gateways/Serving Gateways) of the EPC (Evolved Packet Core) of an EPS are used for Selective IP Traffic Offload (SIPTO). In Fig. 1 on the left side, without taking into account load balancing, UEs connecting to a particular access point AP may all be registered with the P/S-GW that is geographically nearest. With such naïve selection of P/S-GWs, the number of subscribers to a particular P/S-GW may exceed the maximum number it supports or the LTE (Long

Term Evolution) traffic associated with the UEs may exceed the capacity of the P/S-GW (e.g., P/S-GW2 in Fig. 1). This will lead to excessive queuing delays and packet drops at the selected P/S-GW. To avoid such issue, there is need to also consider load in the selection of P/S-GWs, as shown in the right part of Fig. 1.

A number of solutions have considered loads in their mobility management, using new signaling messages or interfaces, specifically dedicated for that. For instance, in section 20.2.3 of 3GPP Technical Specification 26.300 an inter-cell load management approach in E-UTRAN is described. According to the specifications, the inter-cell load management is performed through the X2 interface. In case of variations in the load conditions, the eNodeB signals the new load condition to its neighbor eNodeBs, e.g. the neighbor eNodeBs for which an X2 interface is configured. Specifically, the LOAD INDICATOR message is used to signal the load conditions between eNodeBs. However, this approach comes along with significant signaling overhead.

According to another approach being applied in the context of EPS, a monitoring component is provided that performs load measurements and reports the results of the measurements to a Domain Name Server (DNS). The MME contacts the DNS server for IP address resolution of a UE and, in return, the MME receives from the DNS an appropriate gateway or a selection of appropriate gateways for that UE. However, this solution suffers from two major disadvantages, the first one being that extensive signaling is required between the network entities and the DNS to report the load as well as between the MME and DNS to check the load, and the second one being that the approach is rather static since the database of the DNS server is updated only in certain intervals.

It is therefore an objective of the present invention to improve and further develop a method and a system of the initially described type in such a way that, by employing means that are readily to implement, the dynamics and efficiency of gateway selection and/or handover execution decisions, in particular with respect to load balancing issues, is improved.

In accordance with the invention, the aforementioned objective is accomplished by a method comprising the features of claim 1. According to this claim such a method is characterized in the steps of gathering at said mobility management function information on the operational performance conditions and/or technical specifications of radio access nodes and/or network entities involved in or relevant for said mobile terminal's communication performance, and considering said information for gateway selection and/or handover execution decisions.

Furthermore, the above mentioned objective is accomplished by a system comprising the features of claim 18. According to this claim such a system is characterized in that said mobility management function further includes collector means, being configured to gather information on the operational performance conditions and/or technical specifications of radio access nodes and/or network entities involved in or relevant for said mobile terminal's communication performance, and decision means, being configured to consider said information for gateway selection and/or handover execution decisions.

According to the present invention it has first been recognized that it is possible to improve the efficiency of gateway selection and/or handover execution decisions by gathering information not only with respect to radio access nodes, but also with respect to network entities that are involved in or, in any regard, relevant for said mobile terminal's communication performance. These network entities may in particular include network mobility anchors and/or user-plane gateways, e.g. a PDN GW or Serving GW, or a LMA (Local Mobility Anchor) or MAG (Mobility Access Gateway). The entities from which information is gathered may include the radio access node the mobile terminal is currently attached to, as well as radio access nodes nearby to which an attachment of the mobile terminal would be possible, and further the network entities along the mobile terminal's current communication path, as well as the network entities which, from a topological point of view, could possibly become part of the communication path. By gathering information on the operational performance conditions and/or technical specifications of all or at least a part of these entities, the mobility management function gets a profound and comprehensive overview of network performance. According to predefined policies the mobility management function performs an

analysis of this information, the results of which constitute the basis for gateway selection and/or handover execution decisions.

According to a preferred embodiment it may be provided that the network entities may be configured to report the information, in particular load and load-related information, by inserting the respective information into existing signaling messages (e.g. a bearer setup response or mobility related messages) towards a mobility management entity, such as MME in EPS networks. Existing signaling messages means that signaling messages that are actually sent between network entities (e.g. a PDN or Serving GW) and a mobility management function (e.g. a MME) for another purpose – e.g. for session or mobility management – are used to piggyback load related information to inform the mobility management function.

In this way optimized handoff and handover execution operations can be achieved with no need for new interfaces or new signaling messages. For instance, in the context of an EPS application, a P-GW (PDN-Gateway) could extend the standard “Modify Bearer Response” message by inserting information on its current load. In addition, an S-GW (Serving-Gateway), upon receiving this message, may generate a “Create Session Response” message in which it inserts the load information received from the P-GW together with information on its own load.

In an alternative embodiment it may be provided that the network entities employ dedicated signaling messages for reporting their information, in particular load and/or load related information, to the mobility management function at the relevant mobility management entity. Although dedicated signaling messages inhibit the disadvantage of increasing the signaling overhead, such embodiments may be useful in terms of achieving enhanced flexibility. Dedicated signaling messages may require the provision of new interfaces on the involved network entities, which are designed to process the dedicated signaling messages.

Advantageously, load and/or load-related thresholds are defined for the network entities, which may be realized through static configuration, e.g. by the operator, or dynamically via the mobility management function. The heights of those thresholds may be adapted to the respective network entities' capacities, i.e. a powerful

network entity with high capacity may be assigned a higher threshold than a network entity with only little capacities. For instance, the thresholds may be set up by the network operator either through static configuration or dynamically via the mobility management function in order to balance the load across different core network entities.

The load and/or load-related thresholds may relate to different information metrics, which may include, but not limited to, instant load, average load, buffer queue length, queuing delay, CPU, inbound traffic rate, outbound traffic rate, and/or overall bandwidth.

In a preferred embodiment, in order to minimize the signaling overhead, network entities may be configured to report information, in particular load or load-related information, to the mobility management function only in case the value of the information exceeds the respective thresholds. In such case the respective entity will report this to the mobility management function, either by using standard signaling messages or by means of dedicated signaling messages. In a preferred embodiment, upon receiving information that the load of a network entity exceeds the respective threshold defined for that network entity, the mobility management function may select another network entity, e.g. another network mobility anchor or another user-plane gateway, for handling consecutive requests from a radio access node. More specifically, the mobility management function will first select any adequate network entity that still runs below the threshold, or if all run above the threshold, it may select the one with the lowest indicated load.

According to another preferred embodiment the mobility management function is configured to send faked signaling messages to radio access nodes and/or network entities. More specifically, the faked signaling messages may include a flag that is associated to a specific kind of information metric or a combination of information metrics. The flag is constructed to provide two different functionalities. On the one hand, it indicates to a recipient of the message that it is just a faked signaling message and that it is not intended that the recipient performs the respective action the signaling messages are related to. On the other hand, by the flag being associated to a specific kind of information metric or combination of

information metrics the faked message advises the recipient what information to include into a response signaling message.

In a specific embodiment it may be provided that the mobility management function sends faked handover requests to a mobile terminal's potential target radio access nodes, in particular upon receiving a message indicating a required handover from the respective mobile terminal's source radio access node. Based on information contained in response signaling messages the mobility management function receives from network entities, it (i.e., mobility management function) may make a profound decision on which target access radio node to choose for handover execution.

In another specific embodiment it may be provided that the mobility management function sends faked request messages for session creation to network entities, in particular to a mobile terminal's potential Serving-Gateways (S-GWs) or Mobility Access Gateway (MAGs). In case a mobile terminal is in active mode, this action may be triggered by a handover request the mobility management function receives from the mobile terminal's target radio access node. On the other hand, in case of a mobile terminal being in idle mode, this action may be triggered by a tracking area update request from the mobile terminal. Based on information contained in response signaling messages the mobility management function receives from network entities, it may make a profound decision on which network entities to choose as gateways for the mobile terminal.

When the mobile terminal is in idle mode it may happen that the mobility management function selects a new PDN Gateway or LMA (Local Mobility Anchor), respectively, for the mobile terminal which is more adequate, e.g. due to load balancing reasons. In such cases the new IP address allocated to the mobile terminal by the new PDN Gateway or LMA may be communicated to the mobility management function by including it into existing signaling messages. For instance, in case of EPS the newly assigned IP address may be included into the "create bearer response" message (for communicating it from P-GW to S-GW), and into the "create session response" message (for communicating it from S-GW to the mobility management function). With respect to the communication of the

newly assigned IP address from the mobility management function to the mobile terminal it may be provided that it is included into the TAU (Tracking Area Update) accept message.

According to specific implementations the mobility management function may be executed/provided by an MME (Mobility Management Entity), for instance in EPS scenarios, or by a MAG (Mobile Access Gateway), for instance in PMIPv6 scenarios, in particular in WiMAX implementations. The radio access node may be an eNodeB, e.g. of an EPS, or a base station, in particular a WiMAX base station. The network entities may include all kinds of mobility anchors and/or gateways that encounter or are in charge of mobility related issues, in particular P-GWs and S-GWs in case of EPS implementations, and LMAs (Local Mobility Anchors) and MAGs (Mobile Access Gateways) in case of PMIPv6 (e.g., in case of WiMAX implementations).

There are several ways how to design and further develop the teaching of the present invention in an advantageous way. To this end, it is to be referred to the patent claims subordinate to patent claims 1 and 18 on the one hand, and to the following explanation of a preferred example of an embodiment of the invention illustrated by the drawing on the other hand. In connection with the explanation of the preferred example of an embodiment of the invention by the aid of the drawing, generally preferred embodiments and further developments of the teaching will be explained. In the drawing

- Fig. 1 is a schematic view illustrating a load rebalancing scenario in the context of SIPTO,
- Fig. 2 is a schematic view illustrating a reference network architecture for applying the present invention,
- Fig. 3 is a schematic view illustrating a 3GPP reference network architecture for applying the present invention,

- Fig. 4 is a diagram illustrating the signaling flow in a threshold based load balancing scenario according to an embodiment of the present invention in case of EPS,
- Fig. 5 is a diagram illustrating the signaling flow of a handoff decision with the UE in active mode according to an embodiment of the present invention in case of EPS,
- Fig. 6 is a diagram illustrating the signaling flow of a handoff decision with the UE in active mode according to another embodiment of the present invention in case of EPS, and
- Fig. 7 is a diagram illustrating the signaling flow of a handoff decision with the UE in idle mode according to an embodiment of the present invention in case of EPS.

Fig. 2 is a schematic view of a reference network architecture, which is the general architecture in which the present invention can be applied. In the lower part of Fig. 2 the backhaul network is depicted, which in the illustrated embodiment includes three radio access nodes denoted Radio Access 1 to 3. A mobile terminal, which in the following will be referred to as User Equipment UE, is located in the overlapping coverage area of Radio Access 1 and Radio Access 3, schematically indicated by the circles.

The connection to the core network is realized by means of mobility gateways. In the architecture of Fig. 2 a total of three mobility gateways are provided, denoted Mobility GW 1 to 3. It is important to note that a radio access node may have established connections to more than one mobility gateway, see for instance Radio Access 2 and 3 in Fig. 2.

Each mobility gateway is connected to one or more mobility anchors, denoted Mobility Anchor 1 to 3, which constitute the gateways to the service network. In this architecture, mobility anchors and mobility gateways, respectively, serve as

LMA and MAG in the context of PMIPv6 or as P-GW and S-GW in the context of 3GPP. The 3GPP case, to which all embodiments that are described in the following refer, is illustrated in Fig. 3. The Serving-Gateways S-GW function as gateways between the backhaul network and the EPC (Evolved Packet Core), and the PDN-Gateways P-GW serve as gateways between the EPC and the service network. For the sake of consistent notation, in Fig. 3 radio access nodes are denoted as eNBs.

Although in the following the main focus will be placed on GTPv2 (GPRS Tunneling Protocol v2), the same concepts that are described in connection with the following embodiments are applicable to PMIPv6, wherein 3GPP's S-GW and P-GW would map to PMIPv6's MAG and LMA, respectively. Furthermore, with respect to existing standard signaling messages, the "Create Session Request" message would have to be replaced by the "Proxy Binding Update" message, and the "Create Session Response" message would have to be replaced by the "Proxy Binding Acknowledgment" message. Additionally, despite the focus of the examples on 3GPP architectures, the same procedures can also be applied to GPRS, 3GPP2, WiFi, or WiMAX reference architectures.

Fig. 4 illustrates a threshold based load balancing scheme according to an embodiment of the present invention. The illustrated embodiment refers to the architecture shown in Fig. 3. The notation is the same as the one employed in connection with Fig. 3.

In the beginning UE1 is attached to a source eNB1, and it performs a handover to target eNB2 (HO1 preparation, HO1 execution). Thereafter, it sends a "Packet Switch Request" message to the MME (Mobility Management Entity). The MME incorporates a mobility management function at which according to the invention information on the operational performance conditions and/or technical specifications of radio access nodes and/or network mobility entities involved in or relevant for a UE's communication performance are gathered, as will be explained in detail below. The MME sends a "Create Session Request" message to S-GW1, which thereupon sends a "Modify Bearer Request" message to P-GW1.

In the embodiment of Fig. 4, the operator either through static configuration or dynamically via the mobility management function of the MME sets up one or more thresholds that are used to load balance the load across different core network entities, i.e. between different S-GWs and P-GWs. Advantageously, the thresholds will be determined for each network entity individually, thereby taking into consideration the specific capabilities and resources of the respective entity. In order to enable the MME to refer to these thresholds when deciding on the acceptance of new service requests, the following steps are executed:

Upon receiving the "Modify Bearer Request" message from S-GW1, P-GW1 generates a response message in form of a "Modify Bearer Response" message and includes information on its current load – Load 1 – into this message. S-GW1, upon receiving this message, sends a "Create Session Response" message to the MME, thereby including into the message the load information received from P-GW1 together with its own load information – Load 2. As a result, by receiving the feedback the MME obtains substantive knowledge of the load conditions in the core network. In the embodiment illustrated in Fig. 4, both S-GW1 and P-GW1 are currently busy and are operating in states in which their load thresholds are exceeded. In a preferred embodiment, an operator may configure the operation of load information signaling at P/S-GWs in a way that load information is signaled by P/S-GW to MME only if their respective loads exceed the relevant thresholds.

Next, UE2 performs a handover from source eNB1 to target eNB2. After preparation and execution of the handover (HO2 preparation, HO2 execution), target eNB2 sends a "Packet Switch Request" to the MME. Owing to the load information that was gathered by collector means of the mobility management function as described above, the MME knows that both S-GW1 and P-GW1 are running with loads exceeding their threshold. By applying decision means of the mobility management function, the MME will take this information into consideration and will choose another S-GW/P-GW to accommodate newly arriving requests. For instance, for the sake of load balancing the handover request of target eNB2 in connection with UE2 is processed by the MME by sending a "Create Session Request" message to S-GW2 instead of S-GW1 and the corresponding "Modify Bearer Request" is sent to P-GW2 instead of P-GW1.

By applying the mechanism as described above, refusals of service requests are efficiently avoided. For example, if an S-GW is running with loads exceeding its threshold, it will still accept the request, but indicate in its response the current load. The MME will learn about this feedback, and for consecutive requests it will choose another S-GW. Here the MME may first select any adequate S-GW that still runs below the threshold, or if all run above the threshold, it may pick the one with the lowest indicated load. It is to be noted that the same concepts may be applied to P-GW selection, regardless of the status of a UE, i.e. regardless of whether the UE is in idle or in active mode.

Referring now to Fig. 5, a scenario is considered whereby a UE is active and is about to perform handoff from source eNB1 to target eNB2 or eNB3 (with reference to Fig. 3). Upon handoff detection, source eNB1 is then requested to prepare the handoff. In Fig. 5, the parts within the circle marks (inside the dotted box) indicate the specific aspects of the present invention with respect to the illustrated embodiment.

Initially, based on measurement reports from UE, source eNB (i.e., eNB1) issues a "handover required" message to MME. MME then sends flagged ("fake") "handover request" messages and "create session request" messages with flags to potential target eNBs (i.e., eNB2 and eNB3) and their corresponding S-GWs, respectively. The flags are intended to provide to the respective receiving entity with two kinds of information: on the one hand it shows the receiving entity that the message is only a faked message that does not require the receiving entity to perform the respective action. On the other hand it indicates to the receiving entity the kind of information metric that is requested.

In response, target eNBs and S-GWs send "handover request ack" messages and "create session response" messages with the required information. Based on the feedback, MME decides which target eNB to contact and then follows the standard procedure to complete the handoff preparation phase. It is to be noted that in the example of Fig. 4, load may be considered as the metric (i.e., required info) for making the handoff decision. However, MME can ask for another metric or a set of

metrics upon which it wants to base its handoff decision. Intuitively, as mentioned already above for each metric a specific flag can be set. These metrics can also consist of hardware capabilities or any other parameters that might assist the handover decision, in particular instant load, average load, buffer queue length, queuing delay, CPU, inbound traffic rate, outbound traffic rate, and/or overall bandwidth.

Admittedly, by having MME issues “fake” handover requests and create session requests to multiple eNBs and S-GWs, a high volume of signaling messages may be generated if many different UEs perform handoff at nearly the same time. As a solution to this signaling overhead, it may be provided that the MME issues the fake requests for a particular handoff request (from a given UE) but then does not repeat the same procedure regarding consecutive handoff requests coming within a predefined period of time. Instead, it may rely on the information gathered in connection with the initial handoff request.

Predominantly, a UE will not be located in the overlapping area of more than one target eNB, so a scenario in which only S-GW selection occurs (instead of selection of both eNB and S-GW as is illustrated in Fig. 5) will be more realistic. Such scenario, which also applies in cases in which there is preference that S-GW selection occurs only after selection of eNB, is illustrated in Fig. 6. Again, the parts within the circle marks (inside the dotted box) indicate the specific aspects of the present invention with respect to the illustrated embodiment.

Fig. 7 refers to a specific embodiment of the present invention with the UE being in idle mode. During idle mode, MME may want to reselect more adequate P/S-GW, mainly upon receiving a Tracking Area Update request from a UE. The mechanism for this selection follows the signaling exchange mechanism as depicted in Fig. 7. Indeed, information about P-GWs (e.g., average load) required by the MME can be obtained by having the MME sending the S-GWs “create session request” messages with flags and having the S-GWs sending P-GWs “modify bearer request” with flags, and then communicating the values to which the flags refer to the MME, using “create session response” messages, after receiving them via “modify bearer response” messages. Alternatively, a PBU (Proxy Binding Update)

and PBA (Proxy Binding Acknowledgment) flagged message exchange could be used taking into account that such an interface exists between the S-GW and the P-GW. In any case, based on the received information the decision means provided at the mobility management function of the MME can make an efficient P/S-GW selection with respect to load balancing.

When a new P-GW is selected, MME may request the UE to detach and then re-attach. In this case a network-initiated PDN connection re-establishment procedure may be applied as shown in the last step of Fig. 6. As a result of a selection of a new PDN connection, MME sends a "create session request" to the selected S-GW, which sends a "create bearer request" to the selected P-GW. These two messages are acknowledged by a "create bearer response" and "create session response", respectively, which carry the UE's newly assigned IP address. Also, the sessions/bearer context to the previous PDN GW is deleted. Finally, in the TAU accept message, MME indicates the new IP address to the UE or (in case dynamic IP address allocation is used) that the P-GW has changed (to trigger the UE to update the IP address). It should be stressed here that P-GW reselection during idle mode mobility (i.e. as part of the TAU procedure) is a completely new concept in the EPS.

Many modifications and other embodiments of the invention set forth herein will come to the mind of the one skilled in the art to which the invention pertains, having the benefit of the teachings presented in the foregoing description and the associated drawings. Therefore, it is to be understood that the invention is not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

Claims

1. Method for supporting network-based mobility management of a mobile terminal, wherein said mobile terminal is served by a radio access node and wherein a mobility management function is provided that decides on which gateway to handle said mobile terminal and on the respective handover execution decisions,

characterized in the steps of

gathering at said mobility management function information on the operational performance conditions and/or technical specifications of radio access nodes and/or network entities involved in or relevant for said mobile terminal's communication performance, and

considering said information for gateway selection and/or handover execution decisions.

2. Method according to claim 1, wherein said network entities are configured to report said information, in particular load and/or load-related information, to said mobility management function by inserting the respective information into existing signaling messages.

3. Method according to claim 1, wherein said network entities are configured to employ dedicated signaling messages for reporting said information, in particular load and/or load-related information, to said mobility management function.

4. Method according to any of claims 1 to 3, wherein load and/or load-related thresholds are defined for said network entities.

5. Method according to claim 4, wherein said load and/or load-related thresholds are defined through static configuration or dynamically via the mobility management function.

6. Method according to claim 4 or 5, wherein said load and/or load-related thresholds relate to specific information metrics, in particular to instant load,

average load, buffer queue length, queuing delay, CPU, inbound traffic rate, outbound traffic rate, and/or overall bandwidth.

7. Method according to any of claims 4 to 6, wherein said network entities report said information to said mobility management function only in case the value of the said information exceeds the respective thresholds.

8. Method according to any of claims 4 to 6, wherein said mobility management function, upon receiving information that the load of a network entity exceeds the respective threshold defined for that network entity, selects another network entity for handling consecutive requests from a radio access node.

9. Method according to claim 8, wherein said mobility management function selects the network entity with the lowest indicated load or the lowest value of the considered information metric.

10. Method according to any of claims 1 to 9, wherein said mobility management function is configured to send faked signaling messages to radio access nodes and/or network entities.

11. Method according to claim 10, wherein said faked signaling messages include a flag that is associated to a specific information metric or a combination of information metrics and that triggers a network entity that receives a faked signaling message to insert information of the respective metric or combination of metrics into a response signaling message.

12. Method according to any of claims 1 to 11, wherein said mobility management function sends faked handover requests to a mobile terminal's potential target radio access nodes, in particular upon receiving a message indicating a required handover from said mobile terminal's source radio access node.

13. Method according to any of claims 1 to 12, wherein said mobility management function decides, based on information contained in response

signaling messages it receives from network entities, which target access radio node to choose for a handover execution.

14. Method according to any of claims 1 to 13, wherein said mobility management function sends faked request messages for session/bearer/tunnel creation to network entities, in particular to a mobile terminal's potential Serving Gateways or Mobility Access Gateways.

15. Method according to any of claims 1 to 14, wherein said mobility management function decides, based on information contained in response signaling messages it receives from network entities, which network entities to choose as gateways for said mobile terminal.

16. Method according to any of claims 1 to 15, wherein, in case said mobility management function selects a new PDN Gateway or LMA, respectively, for said mobile terminal being in idle mode, said mobile terminal's newly assigned IP address is communicated from said new PDN Gateway or LMA, respectively, to said mobility management function by including it into existing signaling messages.

17. Method according to claim 16, wherein said mobility management function communicates to said mobile terminal the newly assigned IP address by including it into the Tracking Area Update accept message.

18. System for supporting network-based mobility management of a mobile terminal, in particular for executing a method according to any of claims 1 to 17, including

at least one radio access node that serves said mobile terminal, and

a mobility management function being configured to decide on which gateway to handle said mobile terminal and on the respective handover execution decisions,

characterized in that said mobility management function further includes

collector means, being configured to gather information on the operational performance conditions and/or technical specifications of radio access nodes

and/or network entities involved in or relevant for said mobile terminal's communication performance, and

decision means, being configured to consider said information for gateway selection and/or handover execution decisions.

19. System according to claim 18, wherein said mobility management function is provided by a MME (Mobility Management Entity) or by a MAG (Mobile Access Gateway).

20. System according to claim 18 or 19, wherein said radio access node is an eNodeB of an EPS (Evolved Packet System) or a base station, in particular a WiMAX base station.

21. System according to any of claims 18 to 20, wherein said network entities include PDN Gateways (P-GWs), Serving Gateways (S-GWs), Local Mobility Anchors (LMAs) and/or Mobile Access Gateways (MAGs).

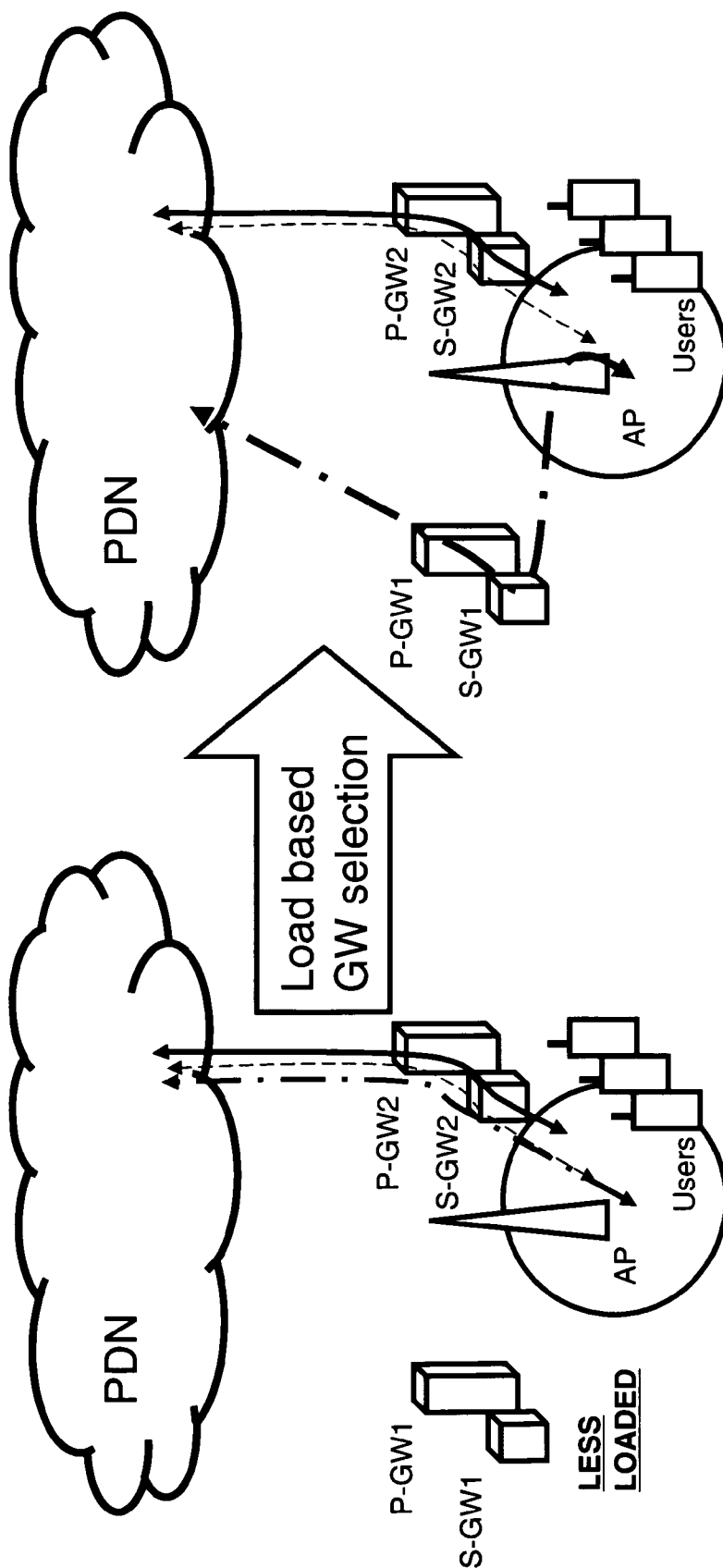


Fig. 1

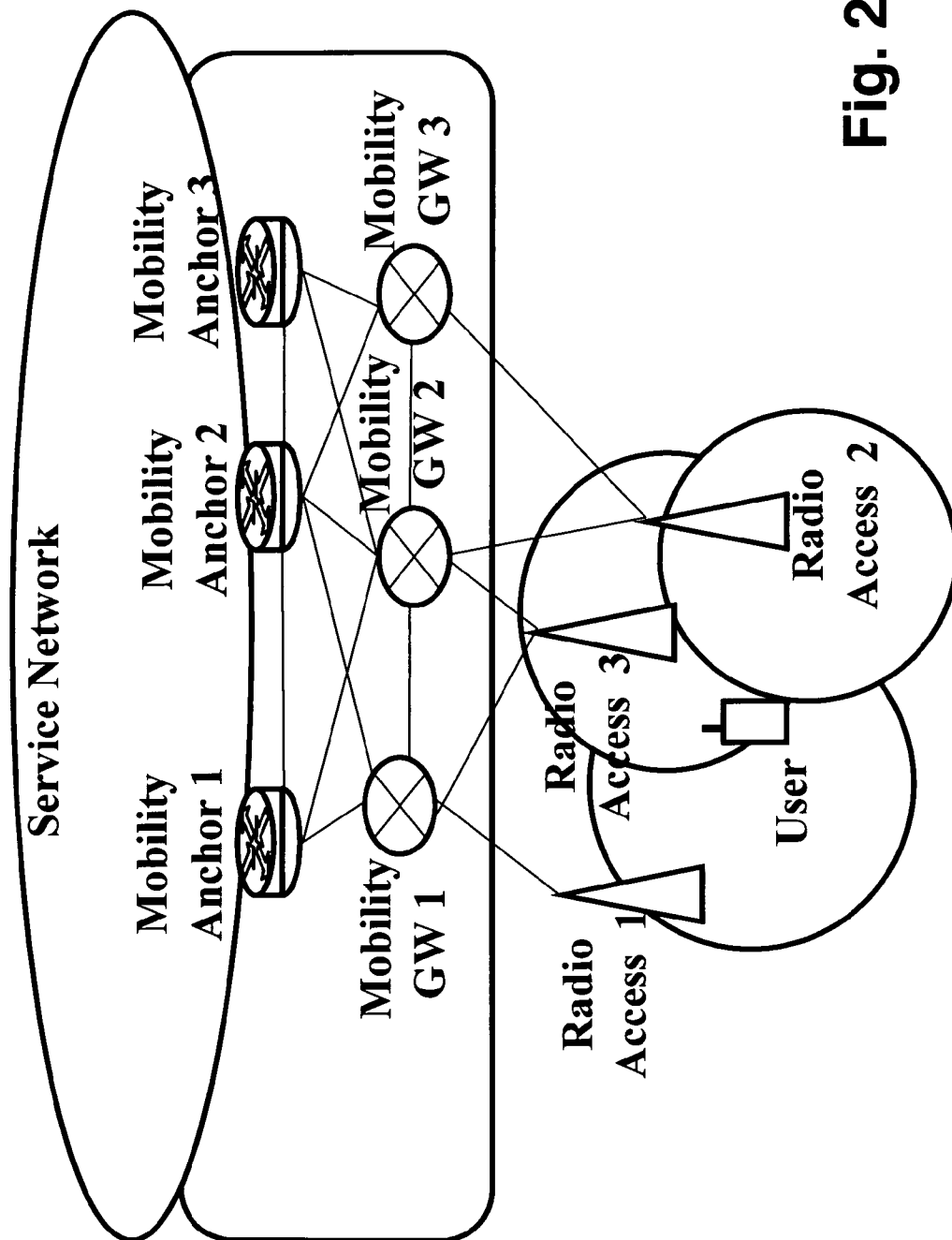


Fig. 2

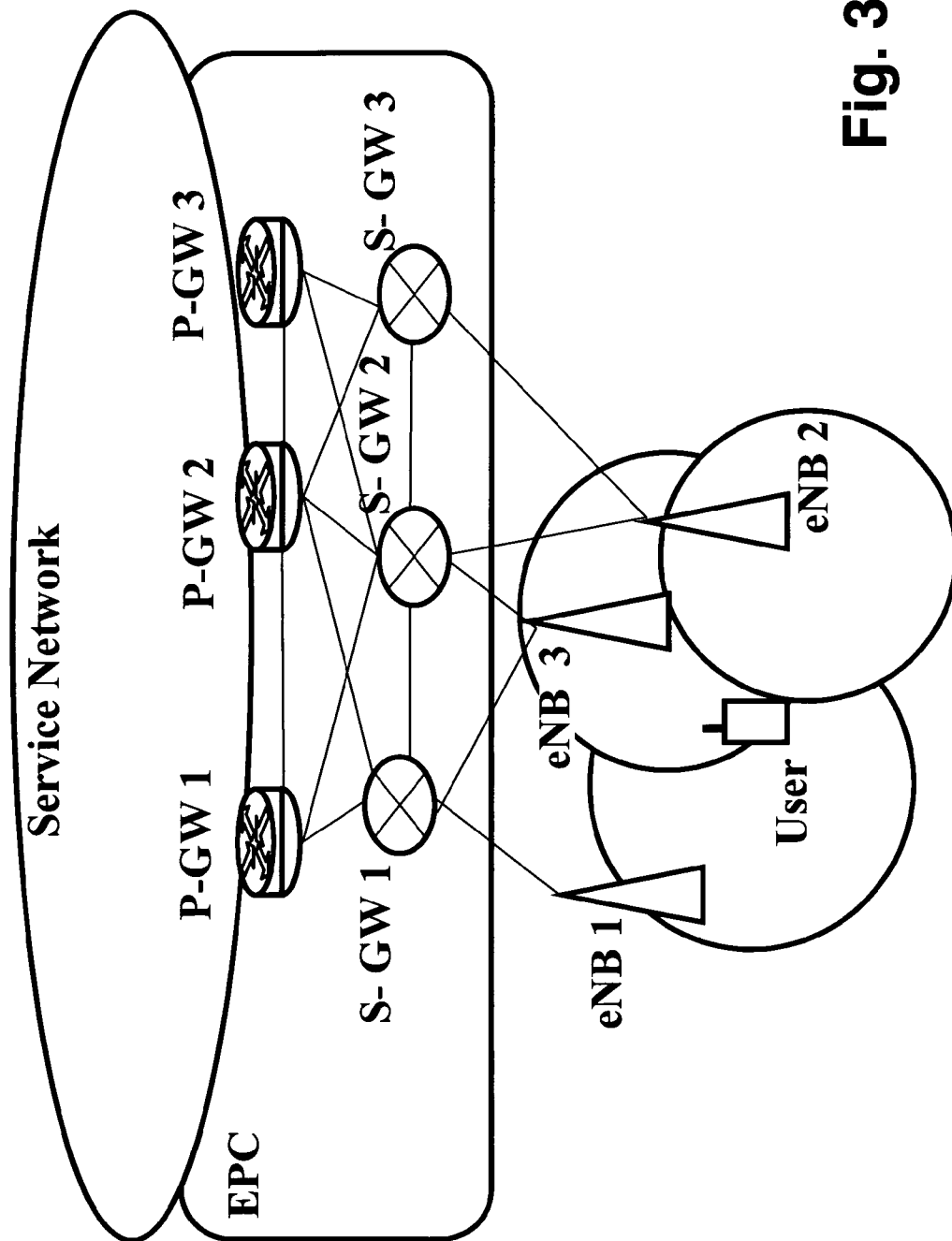


Fig. 3

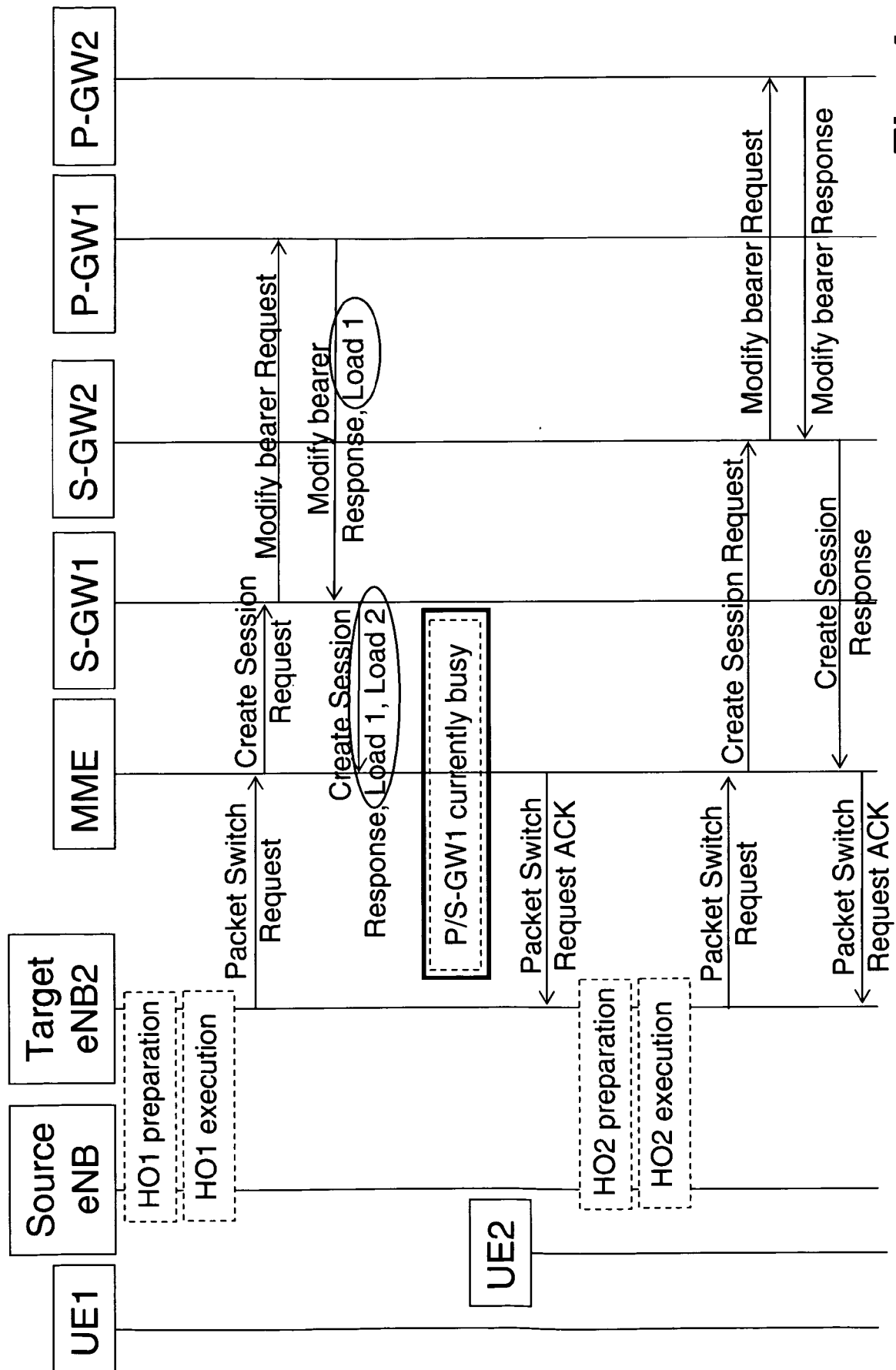


Fig. 4

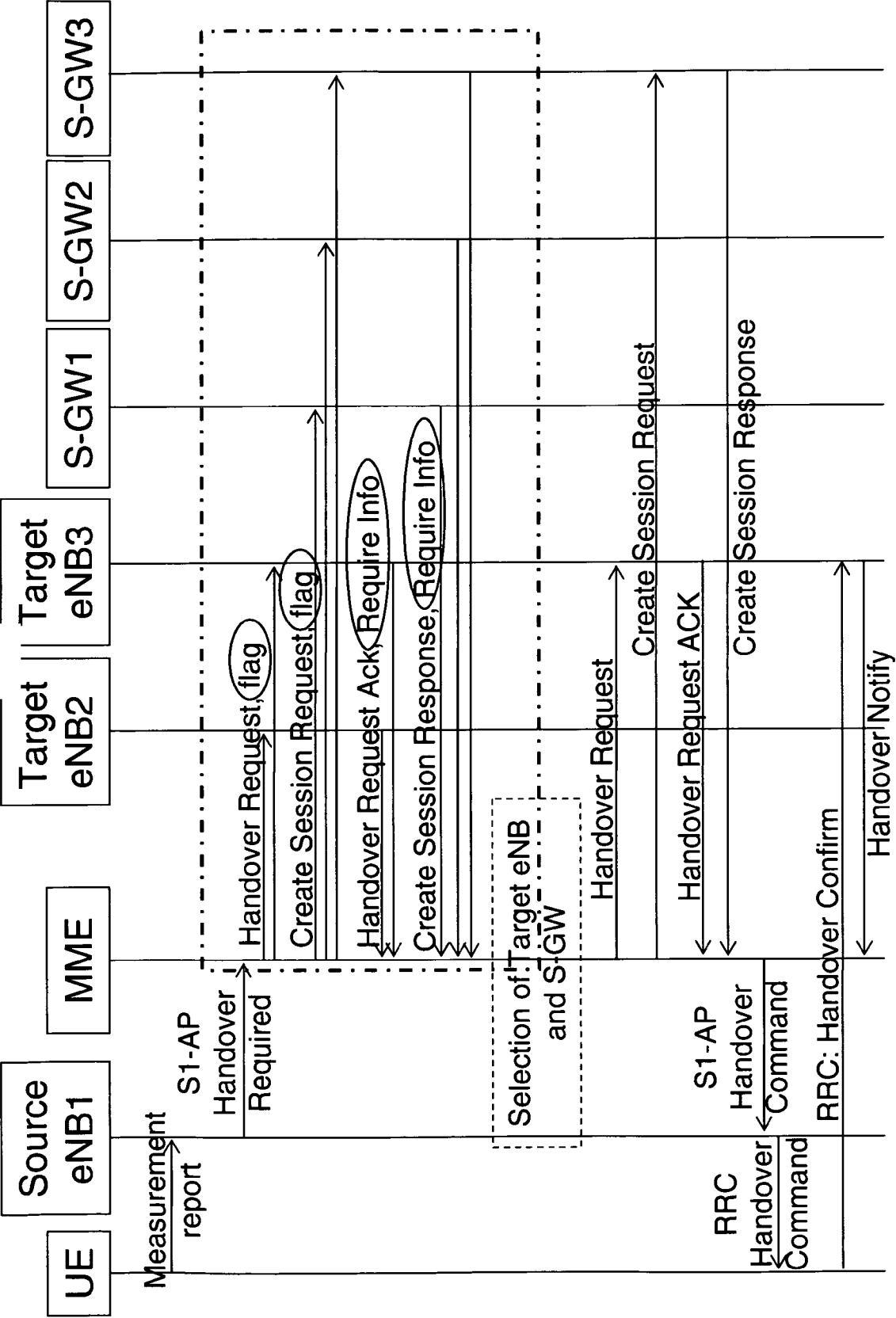


Fig. 5

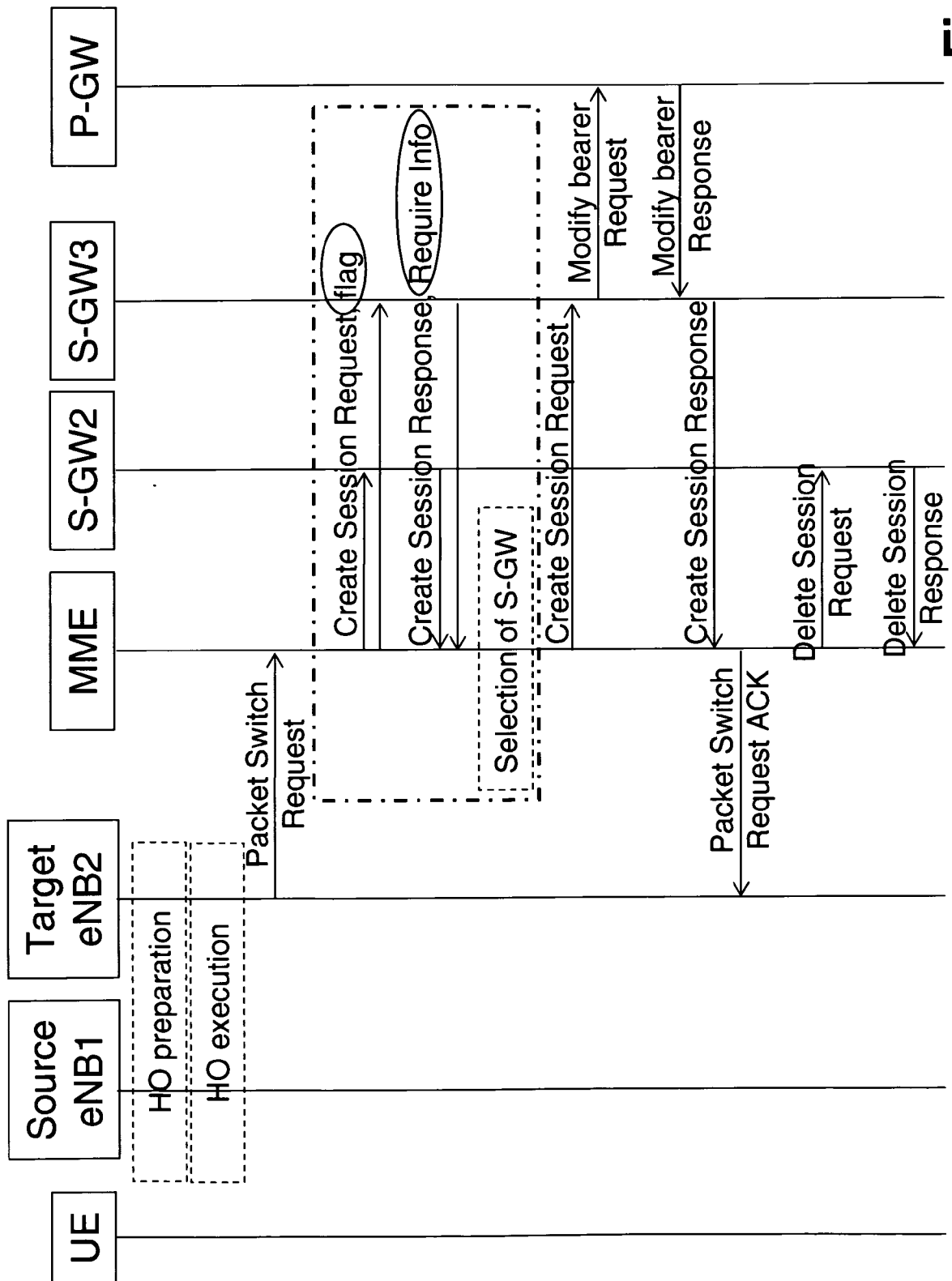


Fig. 6

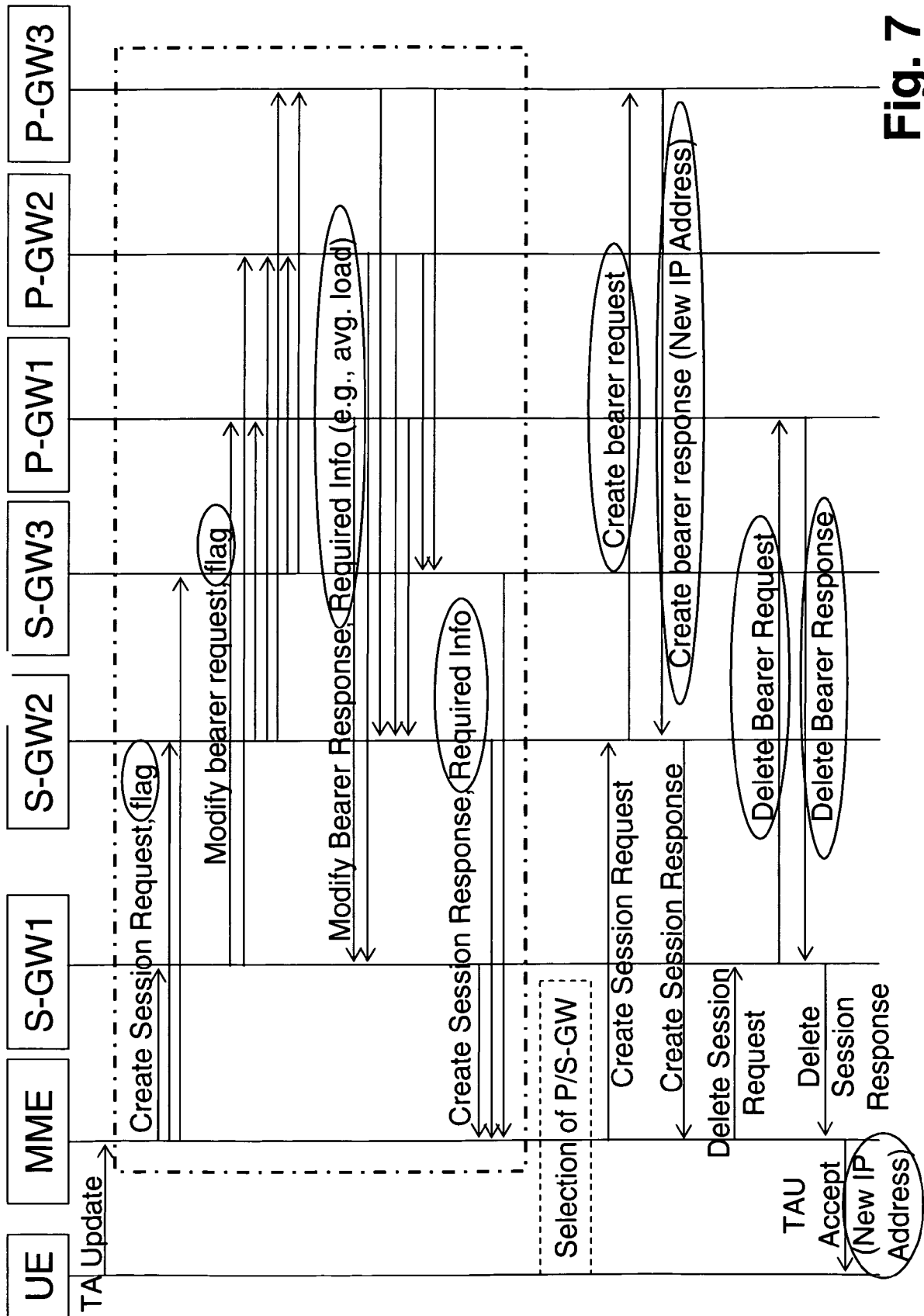


Fig. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/000101

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04W28/08
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04W

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, COMPENDEX, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	NOKIA SIEMENS NETWORKS: "Discussion paper on Load Balancing", 3GPP DRAFT; C4-093440 DISC_LOAD BALANCING, 3RD GENERATION PARTNERSHIP PROJECT (3GPP), MOBILE COMPETENCE CENTRE ; 650, ROUTE DES LUCIOLES ; F-06921 SOPHIA-ANTIPOLIS CEDEX ; FRANCE, no. Beijing; 20091109, 9 November 2009 (2009-11-09), XP050387083, [retrieved on 2009-11-15] the whole document ----- -/-	1-21

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

3 May 2011

Date of mailing of the international search report

11/05/2011

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Rosenauer, Hubert

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2011/000101

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	NOKIA SIEMENS NETWORKS: "Load balancing optimization", 3GPP DRAFT; C4-093439 CR0280R2 29274 LOAD BALANCING_R9 , 3RD GENERATION PARTNERSHIP PROJECT (3GPP), MOBILE COMPETENCE CENTRE ; 650, ROUTE DES LUCIOLES ; F-06921 SOPHIA-ANTIPOLIS CEDEX ; FRANCE, no. Beijing; 20091109, 9 November 2009 (2009-11-09), XP050387082, [retrieved on 2009-11-15] page 2, paragraph 7.2.1 page 7, paragraph 7.2.2 page 13, paragraph 8.xx	1-21
X	----- MUNSUK KIM ET AL: "Load Balancing for Proxy Mobile IPv6 Networks", CONSUMER COMMUNICATIONS AND NETWORKING CONFERENCE, 2009. CCNC 2009. 6TH IEEE, IEEE, PISCATAWAY, NJ, USA, 10 January 2009 (2009-01-10), pages 1-2, XP031425502, ISBN: 978-1-4244-2308-8 the whole document	1,18-21
A	----- NEC: "Load Balancing Signalling and associated SON", 3GPP DRAFT; R2-081175, 3RD GENERATION PARTNERSHIP PROJECT (3GPP), MOBILE COMPETENCE CENTRE ; 650, ROUTE DES LUCIOLES ; F-06921 SOPHIA-ANTIPOLIS CEDEX ; FRANCE, vol. RAN WG2, no. Sorrento, Italy; 20080205, 5 February 2008 (2008-02-05), XP050138949, [retrieved on 2008-02-05] page 2, paragraph 3.1	3
A	----- EP 1 282 321 A2 (ERICSSON TELEFON AB L M [SE]) 5 February 2003 (2003-02-05) paragraph [0002] paragraph [0011] - paragraph [0015] figures 1, 5	1-21

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2011/000101

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1282321	A2	05-02-2003	AT 289149 T 15-02-2005
		DE 60202910 D1	17-03-2005
		DE 60202910 T2	12-01-2006
		US 2003028644 A1	06-02-2003
