



- (51) **International Patent Classification:**  
*H04W 28/02* (2009.01) *H04L 12/823* (2013.01)  
*H04L 12/833* (2013.01)
- (21) **International Application Number:**  
PCT/EP2014/053394
- (22) **International Filing Date:**  
21 February 2014 (21.02.2014)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
13000942.6 25 February 2013 (25.02.2013) EP
- (71) **Applicant:** NOKIA SOLUTIONS AND NETWORKS OY [FI/FI]; Karaportti 3, FIN-02610 Espoo (FI).
- (72) **Inventors:** RUPPELT, Christian; Novalisstr. 6, 81369 München (DE). THEIMER, Thomas; Forstenrieder Weg 1g, 82065 Baierbrunn (DE). STADEMANN, Rainer; Franzikusweg 17, 82335 Berg (DE). WOELKER, Roland Antonius; Sokinmäki 8 A 2, FIN-02760 Espoo (FI).
- (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, BN, 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, IR, IS, JP, KE, KG, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, 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 GATEWAY FOR CONVEYING TRAFFIC ACROSS A PACKET ORIENTED MOBILE SERVICE NETWORK

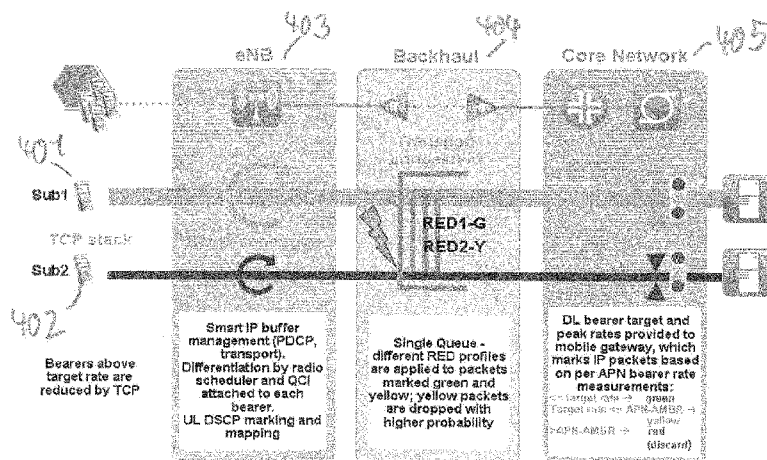


Fig. 4

(57) **Abstract:** A method for conveying traffic across a packet oriented mobile service network is provided, wherein the method comprises assigning a target bit rate and a peak bit rate to each bearer traffic stream; and marking and/or dropping in a gateway and/or radio transceiver data packets of a bearer traffic stream according to their compliance with the target bit rate and/or peak bit rate assigned to the respective bearer before passing them on to a mobile backhaul network for delivery towards the radio transceiver.

METHOD AND GATEWAY FOR CONVEYING TRAFFIC ACROSS A PACKET  
ORIENTED MOBILE SERVICE NETWORK

5

BACKGROUND OF THE INVENTION

Field of the Invention

- 10 The invention relates to the field of mobile service networks. More specifically, the invention relates to a method, mobile service network, and a gateway for conveying traffic across a packet oriented mobile service network.
- 15 As an example some different kinds of mobile service networks are depicted schematically in Figs. 1 and 2. As an example, in a mobile service network 100 at least one radio transceiver (Radio Transceiver) 101, connected via an access and aggregation network (Mobile Backhaul Network; MBH) 102 to
- 20 at least one gateway (Gateway) 103, provides through a radio interface connectivity between a plurality of (in most cases mobile) subscriber devices (user equipment, UE) 104 among each other and with other (mobile) subscriber devices, servers, or other components or devices (not shown),
- 25 reachable from the at least one gateway via the Public Switched Telephone Network (PSTN), through the Internet 105, or using other, potentially dedicated, wired or wireless (fixed or mobile) service networks. Mobile Core Network functions (not shown) may be incorporated in or associated
- 30 with the mobile service network to support mobility management, to implement operator policies, and to perform service control.

User equipment (UE) comprises any fixed or mobile devices, systems, or arrangements in the hands, or at a site, or under control of a subscriber (or user) of the mobile service network and capable of connecting to the network via the radio interface provided by a radio transceiver.

The radio transceiver (RT) may be a base station (BTS), a NodeB, an enhanced NodeB (eNodeB) or any equivalent device providing regional (and preferably cellular) radio access using technologies as specified e.g. in the 2G, 3G, 4G/LTE, or other relevant radio standards.

The gateway may be any device, system, or arrangement capable of providing access to other service networks such as, or through, the PSTN, the Internet, or any other kind of application and/or transport service network. For example the gateway may be a PDN gateway, SAE gateway or any other suitable gateway providing an interface, e.g. a packet data network interface, to a network, e.g. the Internet. Typical applications, among others, could be location based or streaming services. The gateway may to a large extent be implemented in, or comprise, computer program software, which, when loaded into the memory and executed on a computer, causes the computer to implement respective gateway functions. Consequently, a gateway device, system, or arrangement may comprise computer hardware and software and it may be capable to, or actually do, provide and/or share hardware and software resources with other system functions not necessarily specific for the gateway function.

Radio transceivers and gateways may be arranged in redundancy schemes for a better availability and reliability of their respective services.

The mobile backhaul network (MBH) may comprise any components and technologies suitable to interconnect radio transceivers and gateways as described above. More recent systems preferably use packet-oriented data transfer and related MBHs  
5 preferably use packet based transport with protocols and formats as e.g. specified by Ethernet or Internet Protocol (IP) related standards. Reuse of existing infrastructures and use of off-the-shelf routers and switches enables cost efficient solutions. Physical transmission may comprise any  
10 kinds of technologies including microwave radio, optical and electrical systems. Packet Data Network (PDN) Gateways interface to the Internet or dedicated packet oriented service networks using a packet data network interface (PDNI) (Fig. 2).

15

#### Related Art

3GPP TS 23.002 V12.1.0 (2012-12) (as well as other versions of the TS 23.002 document) presents possible architectures of  
20 a packet oriented mobile service network based on various radio access technologies and an Evolved Packet System (EPS) as specified by 3GPP. A short summary of a related architecture is provided by F. Firmin in "The Evolved Packet Core" (retrieved on Jan.31, 2013 at [http://www.3gpp.org/The-](http://www.3gpp.org/The-Evolved-Packet-Core)  
25 [Evolved-Packet-Core](http://www.3gpp.org/The-Evolved-Packet-Core)). As these documents emphasize on the mobile service architecture, they do not show the MBH transport.

A variety of different services is provided to subscribers  
30 and users of such kind of mobile service networks via a user equipment as described above. Services may comprise voice, video and data in various combinations of unidirectional and bidirectional, realtime or non-realtime, interactive, messaging, streaming type, or any other modes of

communication. Accordingly, a variety of different service and Quality of Service (QoS) requirements have to be respected for conveying respective traffic flows across the packet oriented mobile service network.

5

Such requirements are reflected in the concept of bearers as specified e.g. in the standardization document 3GPP TS 23.401. The most recent version of this document with respect to the instant application is 3GPP TS 23.401 V11.4.0 (2012-12), issued on Dec. 18, 2012. A bearer uniquely identifies traffic flows that receive a common treatment between a user equipment and a gateway. Packet filters are associated with the bearers to identify the traffic flows belonging to each bearer. All traffic mapped to the same bearer receives the same bearer level packet forwarding treatment, i.e. routing, queuing, scheduling, rate shaping, etc., in the network and thus exhibits the same QoS behaviour. Actions performed on bearer traffic by individual components of the network may differ according to the different roles of the components in the network (user equipment, radio transceiver, MBH, gateway), but the rules applied to individual packets within a component will always be the same for traffic belonging to the same bearer.

3GPP distinguishes between guaranteed bit rate (GBR) bearers and non-guaranteed bit rate (non-GBR) bearers. A QoS Class Identifier (QCI) is associated with each bearer as a reference to access node-specific parameters that control bearer level packet forwarding treatment (e.g. scheduling weights, admission thresholds, queue management thresholds, etc.) in the mobile service network nodes. Dedicated transmission resources are allocated and blocked for the transfer of GBR traffic. An Aggregate Maximum Bit Rate (AMBR) is assigned to each access point to the network and shared

between all related non-GBR bearers. In uplink (or upstream) direction this access point name AMBR (APN-AMBR) is enforced by the UE and the PDN gateway. In downlink (or downstream) direction it is enforced by the PDN gateway. In a similar way  
5 a further AMBR value is applied across all non-GBR bearers of a user equipment (UE-AMBR) and enforced by the radio transceiver.

QCI, GBR and AMBR values enable a mandatory and fair  
10 allocation of resources assigned to the different bearers on the (by its nature) shared radio interface. Whereas dedicated resources are blocked for exclusive use by admitted GBR traffic, remaining resources are shared according to respective subscription levels (e.g. premium or economy)  
15 between the non-GBR bearers currently active on the interface. A fair allocation of resources, reflecting respective subscription levels, is the target, even in times of congestion. When a premium type bearer has three times more weight than an economy type bearer, then this ratio is  
20 targeted when assigning resources to respective competing bearers independently of the number of bearers currently being served per type. Bearer based traffic shaping and related fair traffic shares can also be applied by the radio transceiver to upstream traffic propagated towards the PDN  
25 gateway via the MBH. Note, that the number of subscription levels is not necessarily limited to two.

Whereas traffic can be treated individually on a per bearer basis in the radio transceiver and the gateway, this is  
30 completely different within the packet based MBH. Class based traffic management is applied instead of bearer based traffic control. Traffic classes are distinguished e.g. by DiffServ Code Points (DSCP) in IP based networks (IETF RFC 2474, RFC 2475, RFC 3260 and others), P-bit values with Carrier

Ethernet (IEEE 802.1Q), or EXP bit values with MPLS (RFC 3270), but no bearer individual information can be used.

3GPP TS 23.401 suggests a potential mapping between QCI  
5 values of EPS bearers and DSCP values. Ekström ("QoS Control  
in the 3GPP Evolved Packet System", IEEE Communications  
Magazine, February 2009) explains that the gateway and the  
LTE RAN implements a QCI to DSCP mapping function to make a  
translation from bearer level QoS (QCI) to transport-level  
10 QoS (DSCP), and he concludes that in the transport network  
the bearer is not visible and hence the traffic forwarding  
treatment of each individual packet is based on the DSCP  
value.

15 For Mobile Operators it becomes increasingly important to  
offer differentiated QoS to their customers. For instance, an  
Operator may want to distinguish between "business users" and  
"economy users" or offer "gold, silver and bronze services".  
The QoS which the end user of an LTE network perceives in  
20 case of network congestion is determined by various QoS  
mechanisms applied in different parts of the network.

Congestion can occur at the air interface and also in the  
Mobile Backhaul Network (MBH). Due to the high peak rates  
25 possible in LTE networks, it is not economical for an  
operator to take maximum cell capacities into account when  
dimensioning the MBH.

The Downlink Packet Scheduler of an eNodeB handles congestion  
30 at the air interface by supporting multiple QoS classes  
identified by the Quality Class Identifier (QCI). The  
Scheduler takes into account the QCI each time, when it  
allocates resources to an individual Radio Bearer. Business  
users and economy users, and gold, silver and bronze services

respectively, each have to be distinguished and individually mapped to different QCIs and respective bearers.

Mobile Backhaul Networks use simpler QoS mechanisms than the  
5 air interface schedulers of an eNodeB. In IP based Mobile Backhaul Networks typically class based traffic management (based on the DiffServ concept) is used to handle congestion. Individual Radio Bearers are not visible in the MBH.

10 In LTE networks dedicated resources are reserved for Guaranteed Bit Rate (GBR) bearers, e.g. using strict priority queues and Admission Control (AC) procedures. This raises the issue of how to distribute the remaining bandwidth among non-GBR bearers. For LTE the problem is currently not solved.

15 The general transport congestion problem is addressed by mechanisms defined within IETF and MEF, but the concrete problem to support any QoS use case under transport congestion situation and align the behavior in transport and  
20 radio scheduling is currently not solved.

In the following it is shown that the state of the art transport congestion solutions would not solve this problem.

25 One transport queue, one DSCP class (PHB):

Non-GBR bearers are not differentiated in the transport network.

Non-GBR bearers are mapped to a single transport queue by eNB (UL) and mobile GW (DL) with common DSCP/RED  
30 profile.

Delay characteristics for all non-GBR bearers are the same.

During transport congestion:



The aggregated non-GBR traffic is buffered in a single transport queue and the same scheduling weight and drop precedence (RED) is applied.

TCP flow control reacts to IP packet discard by reducing the TCP flow bandwidth.

Bandwidth sharing converges to a fair share among TCP flows, e.g. high rate flows will slow down, low rate flows keep their rate.

Bearers containing multiple TCP flows can benefit.

In consequence the solution is prone to induce unwanted/unfair traffic mix.

#### Different transport queues and DSCP classes:

There are multiple transport queues representing defined transport QoS treatments (PHB).

Non-GBR bearers are mapped to respective transport queues according to predefined QoS mapping rules by eNB (UL) and mobile GW (DL).

The transport queues are scheduled according to predefined weights and apply configured RED profiles. Different transport queues allow differentiation in delay characteristics.

#### During transport congestion:

The transport queues will get allocated the bandwidth share representing their configured weights and bandwidth setting. Non-GBR bearers mapped to the same transport queue share the allocated bandwidth and configured RED profile. Resulting bandwidth per bearer thus depends on the number of active bearers per queue.

TCP flow control reacts to IP packet discard by reducing the TCP flow bandwidth. Bearers containing multiple TCP flows can benefit. In consequence the solution is prone to induce unwanted/unfair traffic mix.

There are two major drawbacks of this solution:

On the one hand the availability and e2e manageability of the number of RED profiles sets a limit for differentiation granularity. On the other hand the applicability of the solution is restricted because it increases the number of transport queues.

There are already some queues needed anyway:

- EF for voice, (and eventually another EF queue for ToP),
- high AF for gaming, streaming, GBR services
- AF class for HSPA non-GBR needed,
- at least one non-GBR for LTE
- BE: Best effort,
- and Control and Management plane needs to be queued

As written above, a solution which takes into account an explicit feedback loop as standardized and widely deployed in HSPA between BTS and RNC is currently not standardized for LTE and iHSPA architectures. A congestion feedback loop would solve the problem, but needs standardization effort, because multi-vendor architectures are dominating the market.

#### BRIEF SUMMARY OF THE INVENTION

It may thus be an object of the invention to provide a method, a mobile service network, a gateway and computer program product that improves the Quality of Service (QoS), in particular the downlink QoS, behavior of the Mobile Backhaul (MBH) in case of traffic congestion.

The object may be achieved by a method, a mobile service network, a gateway and a computer program product according

to the independent claims. Further exemplary embodiments are described in the dependent claims.

According to an exemplary aspect a method for conveying  
5 traffic across a packet oriented mobile service network is provided, wherein the method comprises assigning a target bit rate and a peak bit rate to each bearer traffic stream; and marking and/or dropping in a gateway and/or radio transceiver data packets of a bearer traffic stream according to their  
10 compliance with the target bit rate and/or peak bit rate assigned to the respective bearer before passing them on to a mobile backhaul network for delivery towards the radio transceiver.

15 In particular, the bearer may be a non-guaranteed bit rate bearer. In particular, the method for conveying traffic may be suitable for mobile backhaul flat architecture, like LTE or iHSPA. In particular, the target bit rate and/or the peak bit rate may be provided to the gateway. In particular, the  
20 marking may be performed in a similar or identical way than used in existing IP packet color marking concept (trTCM, two rate three color marking) or the so called RED profile concept.

25 According to an exemplary aspect a mobile service network is provided comprising means for executing a method according to an exemplary aspect.

According to an exemplary aspect a gateway of a mobile  
30 service network is provided comprising means to: receive target bit rate and peak bit rate information assigned to bearer traffic streams; and mark and/or drop data packets of a bearer traffic stream according to their compliance with the target bit rate and/or peak bit rate assigned to the

respective bearer before passing them on to the mobile backhaul network for delivery towards a radio transceiver, wherein packets exceeding the peak rate are immediately dropped and packets violating the target bit rate are marked  
5 for preferential dropping in case of mobile backhaul congestion.

In particular, the gateway may be adapted or may comprise means for performing a method according to an exemplary  
10 aspect. For example, the gateway may further comprising means for forwarding the marked packets to the mobile backhaul network.

According to an exemplary aspect a computer program product  
15 is provided comprising software, which when loaded into the memory of a computer enables the computer to execute any of the steps of the method of an exemplary aspect.

According to an exemplary aspect a computer program product  
20 is provided comprising software, which when loaded into the memory of a computer enables the computer to implement any of the means comprised by a gateway according to an exemplary aspect.

25 By marking specific data packets for dropping it may be possible to ensure that the probability of dropping is different for different specific data packets in case of transport congestion. Thus, the possibility to improve the QoS by observing predefined or agreed upon service qualities  
30 like gold, silver or bronze services.

In the following some exemplary embodiments of the method are described. However, the described components and features may

also be used in connection with, the mobile service network, the gateway, and the computer program product.

According to an exemplary embodiment of the method the bearer  
5 traffic stream is a downlink bearer traffic stream and/or an uplink traffic bearer stream.

In particular, the method may be used as well for downlink packets and for uplink packets. It should be noted that in  
10 case of uplink packets the marking and/or dropping may be performed in the radio transceiver eNodeB or base station instead of the gateway which is preferably used in case of downlink packets.

15 According to an exemplary embodiment of the method the packets exceeding the peak rate are immediately dropped and packets violating the target bit rate are marked for preferential dropping in case of mobile backhaul congestion.

20 According to an exemplary embodiment the method further comprises forwarding the marked packets by the mobile backhaul network.

According to an exemplary embodiment the method further  
25 comprises in case of congestion in the mobile backhaul network preferably dropping packets marked for preferential dropping.

According to an exemplary embodiment of the method the target  
30 bit rate and/or the peak bit rate for each bearer is derived from the bandwidth assigned to this bearer by the air interface scheduler of the radio transceiver for transmission of the bearer data stream across the air interface.

According to an exemplary embodiment of the method the target bit rate and/or the peak bit rate for each bearer is set according to an access point name of the bearer and/or a subscriber class of the bearer.

5

In particular, the subscriber class may be set or chosen according to operator policy.

According to an exemplary embodiment of the method an actual  
10 bit rate of at least one bearer traffic stream is measured.

In particular, based on the measured actual bearer traffic rate marked packets may be dropped or not. That is, a decision may be made based on the measured actual bit rate of  
15 the bearer traffic stream whether marked packets are dropped of that specific bearer. By measuring the actual bit rate it may be possible to combine the method according to an exemplary aspect with existing IP packet color marking concepts, like trTCM or two rate three color marking.

20

According to an exemplary embodiment of the method the bearer is a non-guaranteed bit rate bearer.

In particular, only non-guaranteed bit rate bearer may be  
25 handled or may be managed according to the described methods. Guaranteed bit rate bearer may be handled individually with CAC, policing or the like.

According to an exemplary embodiment of the method the  
30 different kinds of markings are used for different drop precedencies.

In particular, the different markings may correspond to the so called RED profiles and represent different drop

probability for the respective marked packet. For example a value green may represent that the respective packet is transferred or transmitted while yellow may represent that the respective packet exceeds the target bit rate and thus  
5 may be dropped in case of a transport congestion and red may represent that the respective packet exceeds the peak bit rate and should be or is dropped anyway.

According to an exemplary embodiment of the method a single  
10 transport queue is used for packets having different kinds of marking.

In particular, a single transport queue with different RED profiles for different drop precedence of different color  
15 marked packets may be used. Thus, existing transport equipment functionality may be reused.

It should be noted, that major parts of the algorithm may be performed outside of the radio transceiver (or eNodeB), e.g.  
20 in any other control instance of the system, which is equipped with a computer. Such instance could e.g. be a network management system or a policy controller.

It should be noted further, that most (or even all) steps of  
25 the method or parts of the related algorithm may be implemented in software, i.e. as a computer executable program code, which when loaded into the instruction memory of a computer, enables the computer to execute the respective method steps or parts of the algorithm. As such the software  
30 may be incorporated with any means capable of storing permanently or temporarily computer program code or related data.

It should be noted that in the same way at least some of the steps and parts may as well be implemented in hardware, e.g. in electronic circuitry and/or logic devices of any kind. Such hardware may especially comprise equipment for packet  
5 classification, queuing and scheduling and their respective control.

Consequently and obviously, any system and device capable of or intended to be used for executing the method, or the  
10 underlying algorithm, or any part of any of these, is preferably equipped, has to be equipped or is at least with respective means, i.e. a computer (processing device with respective memory and input/output capabilities, etc.) and/or other respective hardware means.

15

A main concept of the present invention may be to introduce a flexible traffic management method for MBH flat architecture (e.g. LTE) aligning radio and transport behavior under congestion situations and supporting any QoS use case for the  
20 downlink traffic direction comprising the steps of:

- a) Introducing a target bit rate concept for non-GBR Traffic within mobile GW (GGSN, S-/PGW) for downlink traffic, on a per APN and subscriber class level.
- b) Expand the said target bit rate concept under reuse  
25 of APN-AMBR to a target bit rate / peak bit rate concept.
- c) Combining this concept with bit rate measurement and with existing IP packet color marking concepts (trTCM, two rate three color marking)
- d) Reusing existing transport equipment functionality:  
30 single transport queue with different RED profiles for different drop precedence of the different color marked packets



Thus, it may be possible to align the very flexible bearer based radio QoS behavior with static transport queue behavior.

- 5 In particular, one single transport queue may be used for different DSCP classes (PHB):

Bearers are marked with DSCPs that correspond to the bearer type and imply different RED profiles. They are mapped to a single transport queue serving different  
10 DSCPs and applying respective RED profiles. Each RED profile has a different queue threshold for dropping packets. Multiple RED profiles per bearer class have to be configured consistently across the network. No differentiation of bearers by delay characteristics.  
15 During transport congestion situations the bearers are affected according to the assigned RED profile - higher drop probability for low-priority traffic. Resulting bandwidth per bearer thus depends on the applied RED profile and embedded TCP flows. TCP flow control reacts  
20 to IP packet discard by reducing the TCP flow bandwidth. Bearers containing multiple TCP flows can benefit.

The invention may attain the goal to react within the transport network during transport congestion on a per bearer  
25 granularity. In consequence all non-GBR QoS use cases which needs differentiation on bearer level will be supported under transport congestion conditions as well. The behavior of the transport will be nearly the same as required, signaled and supported by the radio access sites (eNB, NB, iBTS). The  
30 target/peak rate concept is not part of the 3GPP standards and thus is a proprietary concept. It can be used in downlink direction without impacting other network elements, because in general for one mobile operator the mobile gateways are

single vendor equipment. In particular, there is no impact on transport equipment and their state of the art functionality.

5 A person skilled in the art understands that the principles of the invention as disclosed herein and illustrated based on the example of a 3GPP based system architecture are as well applicable to variants of this, or other architectures of mobile and/or fixed service networks employing the same or a similar type of bearer based traffic control in the access  
10 area, and will easily be able to apply these principles accordingly.

The main implementation advantages may be that no standardization work is needed, will work in any multi-vendor  
15 architecture. In addition implementing this functionality may give a unique selling point.

The aspects and exemplary embodiments defined above and further aspects of the invention are apparent from the  
20 examples of embodiment to be described hereinafter and are explained with reference to these examples of embodiment.

#### BRIEF DESCRIPTION OF THE DRAWINGS

25

Fig. 1 schematically shows a mobile service network.

Fig. 2 schematically shows another kind of mobile service network.

30

Fig. 3 shows some main characteristics of the transport network in comparison to the radio network layer.

Fig. 4 schematically depicts a main concept: DL Three-Color Marker based on Bearer Target/Peak rate.

Fig. 5 schematically depicts LTE network architecture (Radio network- and transport network layer QoS)

Fig. 6 shows simulation results of a first solution.

Fig. 7 shows simulation results of second solution.

10

Fig. 8 schematically shows e2e functionality of the invention.

Fig. 9 schematically shows a simple example of the effect of invention.

15

Fig. 10 shows simulation results showing the effect of the invention.

20 Fig. 11 schematically shows an example of the invention.

Fig. 12 schematically illustrates congestion feedback and target/peak-rate per bearer adjustment.

25 Fig. 13 schematically depicts a main concept concerning a first scenario - Single Default Bearer per APN.

Fig. 14 schematically depicts a main concept concerning a second scenario - Application Demotion with Dedicated Bearer.

30

## DETAILED DESCRIPTION OF THE INVENTION

A detailed description of the invention, including background, prior art, drawings and sufficient details for its full understanding is provided below.

In the following some general information concerning general traffic management is given which may be helpful for understanding the invention.

In case of congestion the bottleneck bandwidth should be distributed in such a way that the bandwidth share of a non-GBR bearer does not depend on the root-cause of congestion (air interface overload or MBH overload). To give an example: If a Mobile Operator decides that a Business User should get three times the bandwidth of an Economy User and configures the schedulers for the air interface accordingly, it should not happen that a Business User is constrained to only get the same or even less bandwidth than the Economy User, when the MBH is overloaded. Aligning the bandwidth allocation and the QoS delivered by the air interface and in the MBH is not easy, due to the principal differences between the related mechanisms applied at the air interface and in the MBH. The air interface scheduler acts on individual Radio Bearers while the MBH uses class based traffic management and is not aware of Radio Bearers.

For example, considering an eNodeB serving a number of  $p$  cells typical values of  $p$  could e.g. be 3 or 6. Each of the cells has its own Downlink Packet Scheduler, which allocates bandwidth to non-GBR bearers using weights (one weight per QCI) and further parameters which reflect radio conditions.

To ensure a service differentiation according to the user and service categories agreed upon with subscribers connected through the air interface (e.g. business and economy users and/or gold, silver, and bronze services), traffic management and control mechanism applied to other potential bottlenecks in the system should not interfere with the bandwidth distribution and the respective shares allocated by the air interface schedulers. Especially, the distribution of transport bandwidth in case of a congestion in the MBH should not contradict to or jeopardize the distribution of bandwidth to individual bearers as specified for the air interface.

This, however, cannot be achieved by applying the standard QoS mechanisms available in the MBH. Commercial mobile backhaul services typically offer only a very limited number of service classes (e.g. between two and four, as e.g. defined by the Metro Ethernet Forum in respective implementation agreements such as MEF 22.1). Throughput guarantees can be given for these service classes as a whole (e.g. using weighted fair queuing), but not for individual non-GBR bearers.

The bandwidth allocated to each service class is fixed, which can lead to further problems, when e.g. the traffic mix of different types of users (e.g. Business and Economy) cannot be predicted exactly. Even worse, since the service usage on the air interface frequently changes in time, the weights defining the bandwidth shares of Weighted Fair Queuing (WFQ) schedulers in the MBH will usually not fit with the actual traffic mix. It thus may happen that a Business User, though he should clearly be preferred against an Economy User (and actually receives this preferred service at the air interface), gets less bandwidth in the Mobile Backhaul

Network than the Economy User, because e.g. the traffic mix contains more business users than predicted.

In many cases, only a single service class is available for  
5 non-GBR bearers. Hence, all non-GBR bearers receive equal  
treatment in the Mobile Backhaul Network, even when  
congestion occurs. However, this does not imply that they are  
also receiving the same throughput, because the bandwidth  
allocation is now determined by TCP mechanisms. All TCP  
10 sessions, which use the same traffic class in the MBH, will  
lose packets at the bottleneck in the MBH with a similar (or  
even the same) probability. Therefore, in times of  
congestion, the Mobile Backhaul Network has the tendency to  
equalize the throughput of TCP sessions using the same QoS  
15 class. This is implied by the way, the fairness mechanisms of  
TCP are defined and implemented. As a consequence, the  
throughput of non-GBR bearers, assigned to the same QoS class  
in the MBH, is proportional to the (arbitrary) number of TCP  
sessions contained therein. In other words, in times of MBH  
20 congestion two non-GBR bearers, even if associated with the  
same bandwidth at the air-interface, may come out of the MBH  
with completely different bandwidths, if they contain a  
different number of TCP sessions.

25 In uplink direction the radio transceiver (e.g. an eNodeB of  
an LTE system) can easily do an individual shaping of all  
bearers sharing its resources. By doing so it can control and  
limit the amount and the mix of traffic according to the  
resources available for it in the MBH before the traffic  
30 enters the MBH.

In downlink direction the bearers sharing the air interface  
resources of a radio transceiver may pass through different  
gateways. Even AMBR shaping in the different gateways (as it

has no "common view") cannot avoid potential traffic congestion in the MBH with consequences as described above. The result is a completely different behavior of the network depending on the location of a potential traffic congestion  
5 (air interface or MBH) and a respectively inconsistent service experience for the users.

In the following some general remarks concerning transport congestion in mobile backhaul (MBH) are given which helps to  
10 understand exemplary embodiments of the invention.

Transport congestion in MBH.  
Radio technologies evolve towards continuously increasing peak rates and mobile BB (Broadband) traffic volumes grow at  
15 a high rate. Furthermore, "flat" mobile network architectures for LTE and I-HSPA based on IP technology are becoming deployed. In addition MBH equipment in the access area is mostly (ca. 60%) based on MWR (Microwave Radio) technology. Therefore transport congestion caused by non-GBR (non-  
20 Guaranteed Bit Rate) traffic is the challenge for the MBH. Concepts for controlled handling of MBH transport congestion situations in flat mobile network architecture are needed. They preferably take into account the different approaches for QoS and traffic control in the radio interface and the  
25 transport network. They should be in addition embedded in a multi-radio MBH environment of 2G, 3G/HSPA, I-HSPA and LTE.

In particular, Fig. 3 shows some main characteristics of the transport network in comparison to the radio network layer.  
30

The radio network layer is completely defined by 3GPP For LTE 3GPP has defined different QoS profiles (QCIs) for e.g. different applications or different interactive subscriber classes. Profiles are defined bearer based for GBR and non-

GBR traffic. The profiles are used within radio network control to support different QoS use cases: e.g. fair use policy, subscriber differentiation or application differentiation, see the following references:

- 5           TS 23.203 Technical Specification Group Services and System Aspects; Policy and charging control architecture (Release 9);
- TS 23.401 General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio
- 10          Access Network (E-UTRAN) access (Release 9;

APN definition:

- Access point name (APN) identifies an IP packet data network (PDN), that a mobile data user wants to
- 15          communicate with. In addition to identifying a PDN, an APN may also be used to define the type of service

APN-AMBR is defined by 3GPP see TS 23.401

- The APN-AMBR is a subscription parameter stored per APN
- 20          in the HSS. It limits the aggregate bit rate that can be expected to be provided across all non-GBR bearers and across all PDN connections of the same APN (e.g. excess traffic may get discarded by a rate shaping function).
- The present invention reuses the concept of APN-AMBR as
- 25          defined. The term "Peak Rate" is identical with APN-AMBR.

Transport Network Layer:

- Standards and industry agreements for the transport
- 30          network layer are specified by IETF, MEF, NGMN, BBF and others. However these specifications allow for a large variety of options. In contradiction to the 3GPP defined bearer based approach the QoS in transport network layer



is based on class of service mappings which are not bearer aware.

Within IETF various protocols are defined to cope with transport congestion. The base functionality is active queue management. The main RFCs are:

RFC 2309: Active Queue Management

RFC 3168: Explicit Congestion Notification (ECN)

RFC 5559: Pre Congestion Notification architecture (PCN)

RFC 2698: two-rate three-color marker (trTCM)

RFC 2474, RFC 2475 DiffServ (Differentiated Services) Quality of Service (QoS)-Method for prioritizing IP-packets

TCP behavior references:

- Anthony C.H. Ng, David Malone, Douglas J. Leith, Experimental Evaluation of TCP Performance and Fairness in an 802.11e Test-bed, August 2005
- Junsoo Lee, Stephan Bohacek, Jo~ao P. Hespanha, Katia Obraczka

A Study of TCP Fairness in High-Speed

Within MEF Ethernet bandwidth Profiles are defined, main parameters are here CIR, CBS, EIR, EBS.

The bandwidth profiles in combination with trTCM are defined to cope with transport congestion.

MEF 6.1: Metro Ethernet Services Definitions Phase 2

MEF 10.2: Ethernet Services Attributes Phase 2

MEF 22: Mobile Backhaul Implementation Agreement

MEF 23: Class of Service Phase 1 Implementation Agreement

The above described functionalities and mechanisms may only be reused in exemplary embodiments of the invention and it may be a declared goal not to impact the transport equipment

with additional functionality. An explicit feedback loop as standardized and widely deployed in HSPA between BTS and RNC is currently not standardized for LTE or iHSPA architectures.

5 In the following some general remarks concerning objects and/or possible advantages of exemplary embodiments of the present invention may be given:

The following non-GBR QoS use cases are preferably supported in transport congestion situations in LTE and iHSPA

10 architectures:

Fair Use Policy (FUP)

Fair use policy is applied to mobile data services e.g. reduce priority of non-GBR bearer in case of monthly data cap is reached. Within radio the traffic is scheduled according to radio load, in case of non-  
15 congestion situations no reduction might occur, in case of congestion situations the traffic will be reduced to the configured priority weight, which might be zero bandwidth.

20

Subscriber Differentiation:

Subscription based differentiation of QoE for data services (e.g. premium, economy)-, e.g. subscription classes with different priorities, max bandwidths, and  
25 monthly data caps.

Application Differentiation:

Promotion: Prioritized application treatment e.g. better QoS class, guaranteed bandwidth, support of  
30 latency requirements for pre-defined applications, like video streaming.

Demotion: Lower priority for application traffic to e.g. reduce throughput during high load situations in

favor of higher priority traffic, e.g. for peer-to-peer traffic.

5 A basic problem on which the present invention is based may be that the traffic management in the transport is aligned and coordinated with the traffic management in the radio, especially in transport congestion situations. The desired situation is that the transport behavior during transport congestion  
10 must be in line with the mobile network per bearer QoS behavior supported in the radio interface.

Example:

The radio schedules each bearer based on its weight  
15 (derived from QCI and priority), e.g. if one "premium"-type bearer has three times more weight than an "economy"-type bearer, then this ratio is targeted when assigning resources to two competing bearers, independently from the number of bearers being served  
20 per type. The transport scheduling is in known methods in general not based on individual bearers, but on QoS traffic classes (DIFFSERV). The different traffic e.g. "interactive premium service" or "interactive economy service" is mapped to the different QoS classes.  
25 Resources in the transport network are administratively pre-assigned to the QoS traffic classes and are not adapted when the offered traffic volumes per service classes change. In consequence during transport congestion both mechanisms might  
30 split resources differently. This leads to an inconsistent experience for service users.

In the following some principles of the DL target rate concept implemented in LTE/iHSPA mobile GW (GGSN, S/PGW) are described in more detail:

5 Exemplary embodiments are based on the central concept of a target bit rate/peak bit rate concept. The target rate is used by the operator as a transport network dimensioning parameter. That means that during a busy hour the network can support for all active service users at least the target rate  
10 of a service. The target rate concept enables to react to DL transport congestion on a per bearer granularity, as described in the following:

The DL target rate is set for non-GBR bearers per APN and subscriber class according to operator policy.

15 GBR bearers are not included and are handled individually with CAC, policing etc.

The DL target rate is used together with the Peak Rate (see 1 APN-AMBR) to set the drop precedence for IP packets based on bearer rate measurements.

20 In DL direction the IP packets at an APN are color-marked by the mobile GW:

- Green, if the considered bearer rate is below the target rate;
- Yellow, if the considered bearer rate is between  
25 target rate and APN-AMBR;
- Discarded otherwise (red)

During transport congestion situations "yellow" packets are dropped in transport elements with higher  
30 probability than "green" packets according to configured RED profiles

Fig. 4 schematically depicts the basic concept of an exemplary embodiment. In particular, Fig. 4 depicts a

downlink three-color marker process in a packet oriented mobile service network, wherein the process is based on bearer target/peak rates. The mobile oriented mobile service network comprises a plurality of user equipments (UE) 401 and 402. As depicted in Fig. 4 the two UEs has different priority or belonging to a different quality of service class identifier (QCI) so that the UE 402 has a lower target bit rate than the UE1 associated therewith. Furthermore, a level of base stations or enhanced NodeBs 403 is schematically depicted in Fig. 4 including a radio scheduler. In particular, a differentiation between the UEs is performed by the radio scheduler via the QCI attached to each bearer. Although, the two bearers belonging to two different RED profiles both bearers are included in a single transport queue for a mobile backhaul 404. According to the example of Fig. 4 the RED profile of UE 401 is green while the one of UE2 is labelled yellow. Based on a measured actual bit rate and an available transport bandwidth capacity marked packets may be dropped wherein the one marked yellow have a higher probability to be dropped. The packets are transferred to a gateway or core network 405.

Fig. 5 schematically shows the structure of an LTE mobile network architecture 500 as needed for illustration, comprising radio access 501, backhaul transport section 502 and core 503. In addition it shows the different QoS approaches of 3GPP defined bearer based QoS mechanisms and IETF or MEF bearer unaware QoS transport classes.

This architecture is the base architecture of an exemplary embodiment of an invention. New functionality may be added to the gateway only. The complete transport infra(structure) may be used without any changes. Current state of the art

implementation would be to map 3GPP defined QoS types (i.e. QCI values) to DSCP values, P-bits or EXP bits.

The following two figures show simulation results of state of the art implementations. It is assumed that according to an operator use case there are two traffic classes defined with different charging rules Gold and Bronze. Data rates and radio scheduling weights are defined in 1:4 ratio. The goal of a transport solution would be that in case of transport congestion the e2e solution would fit as best as possible to the defined radio scheduling weights.

In Fig. 6 the results of a first solution is depicted in which the two 3GPP defined traffic classes (Gold and Bronze) are mapped to two different physical transport queues via DSCP values. The simulation shows how this solution works in case of transport congestion under different traffic mix of Gold and Bronze traffic. In particular, Fig. 6 shows simulation results for an experiment with 150 users and the behavior of the data rate in bits per second for different percentage of gold users. Line 601 shows the behavior for the gold user, while line 602 shows the behavior of bronze user. The result shows that in case of an extreme traffic mix (e.g. 80% Bronze traffic and 20% Gold traffic) the solution tend to prefer the Bronze traffic (Bronze traffic data rate higher than Gold traffic data rate). In addition the applicability of this solution is restricted because it increases the number of rare transport queues.

For comparison it is shown what would happen in case of transport congestion, if no differentiation would be implemented in transport 603.

In the solution 2 of Fig. 7 the operator defined traffic classes (Gold and Bronze) are mapped to two different drop precedencies (i.e. RED profiles) within one physical transport queue. The simulation of transport congestion under different traffic mix of Gold and Bronze traffic of this solution shows that in case of a traffic mix 50% Bronze traffic and 50% Gold traffic the solution tend to drop the complete Bronze traffic (unfair behavior). In particular, Fig. 7 shows simulation results for an experiment with 150 users and the behavior of the data rate in bits per second for different percentage of gold users. Line 701 shows the behavior for the gold user, while line 702 shows the behavior of bronze user.

Fig. 8 shows how the invention could be implemented in the GW and how it will work together with the transport in case of transport congestion. According to operator use case it is assumed that there are two traffic classes defined with related charging rules (Gold and Bronze). Data rates and scheduling weights and priorities are defined in 1:4 ratio. A goal of the invention (and expectation of end user) is that in case of transport congestion the transport would as best as possible support those priorities and data rates.

The above described goal may be achieved by the following concept including the following features or characteristics:

New functionality:

A target data rate concept is used (e.g. per subscriber) in the gateway for the non-GBR DL traffic. This target rate is chosen according to the QoS/QoE package sold to end customer and is aligned with the QoS parameters of radio scheduling. Goal is that this data rate should be reached as best as possible even under transport congestion situations.

Reuse:

The APN AMBR (aggregated maximum bit rate per APN) is reused as peak data rate per APN.

5

New functionality: Both values are used to implement a measurement concept for those two bit rates.

According to measurement a color marking concept is implemented.

In DL direction the IP packets are color-marked by the mobile GW:

- Green, if the considered bearer rate is below the target rate;
- 15 - Yellow, if the considered bearer rate is between target rate and APN-AMBR;
- Discarded otherwise (red).

Within transport the different colors are treated as drop precedencies.

20 According to transport defined RED profiles (queue filling status → drop probability) yellow packets are dropped with higher probability than green packets.

The following Fig. 9 explains the functionality in a simple example: For simplicity the APN-AMBR of Gold and Bronze user is the same.

- 4 users (2 Bronze users, 2 Gold users) no congestion
- 6 users (4 Bronze users, 2 Gold users) with congestion
- 6 users (2 Bronze users, 4 Gold users) with congestion

30 The model shows:

- that in case of moderate transport congestion only yellow packets are dropped
- fairness between the different subscriber classes according to target data rates is supported.



A simulation of the invention (solution3) is shown in the following Fig. 10:

In particular, Fig. 10 shows simulation results simulating an  
5 Experiment with 150 users and showing the data rate in bits  
per second for gold user 1001 and bronze user 1002. The plot  
of the simulation shows that a clear and fair separation of  
Bronze and Gold traffic through the whole spectrum of traffic  
mix may be achieved.

10

The next figure (Fig. 11) shows another example (use case) of  
the invention. Besides use case subscriber differentiation  
the invention can be used as well for application  
differentiation. As a sub use case application demotion is  
15 shown in the Fig. 11. In parallel to subscriber  
differentiation this use case can be supported via:

(reused functionality) to be demoted application have to be  
detected

(reuse functionality) a dedicated bearer is established for  
20 that specific data flow

New functionality from invention: the complete DL data flow  
is marked yellow,

all non-GBR traffic is mapped to one physical transport queue  
(reuse transport functionality) in case of transport

25 congestion packets of this data flow will be dropped with  
higher probability

Add-on 1: (for uplink direction)

The add-on is to introduce a flexible traffic management  
30 method for MBH flat architecture (e.g. LTE) aligning radio  
and transport behavior under congestion situations and  
supporting any QoS use case for the uplink traffic direction  
comprising the steps of:

a) Introducing a target bit rate concept for non-GBR Traffic within mobile access equipment (eNB, NB, iBTS) for uplink traffic, on a per APN and subscriber class level.

5

b) Expand the said target bit rate concept under reuse of APN-AMBR to a target bit rate / peak bit rate concept.

10 c) Combining this concept with bit rate measurement and with existing IP packet color marking concepts (trTCM, two rate three color marking).

15 d) Reusing existing transport equipment functionality: single transport queue with different RED profiles for different drop precedence of the different color marked packets

The invention may attain the goal to react within the  
20 transport network during transport congestion on a per bearer granularity.

In consequence all non-GBR QoS use cases which needs differentiation on bearer level will be supported under  
25 transport congestion conditions as well. The behavior of the transport will be nearly the same as required, signaled and supported by the radio access sites (eNB, NB, iBTS). The target/peak rate concept is not part of the 3GPP standards. In uplink direction it is preferably used with  
30 standardization or at least multi-vendor agreements, because in general for one mobile operator the radio access equipment is multi-vendor equipment.

Add-on 2 (see Fig. 12):

The add-on 2 is to expand the main concept by introducing a flexible congestion control concept for downlink non-GBR traffic in the mobile GW (GGSN, S-PGW) comprising the steps of:

a) Completely reuse the methods of the main concept.

b) Expand the said concept in the mobile GW with transport congestion control concept.

That is:

i) detection of transport congestion level in the mobile GW and

ii) intelligent adjustment of the target bit rate according to the detected congestion level

Explanation:

Depending on congestion level the target data rates are reduced for all bearer types by a factor (e.g. 20%).

In the following some specifics of possible implementations and some possible advantages are described.

The implementation is proposed to be done within mobile GW (GGSN S-PGW).

The following outlines an implementation according to an exemplary embodiment:

The target rate concept is a concept which may be implemented in mobile GW;

The concept is preferably implemented for downlink direction non-GRB traffic only;

The concept is preferably completely transparent to all other network elements in particular to the transport network capabilities;

The concept reuses existing transport network capabilities;

The implementation is entirely implementable in the mobile GW (GGSN, S-/PGW)

5       As mobile GWs are in general single vendor environment within one operators network, proprietary features are possible, if not impacting other elements; The principles of the DL target rate concept are the following:

10       The target rate can be used by the operator as a transport network dimensioning parameter and enables to react to DL transport congestion on a per bearer granularity;

      The DL target rate is set for non-GBR bearers per  
15       APN and subscriber class according to operator policy; GBR bearers are not included and are handled individually with CAC, policing etc.;

      The DL target rate is used together with the APN-AMBR to set the drop precedence for IP packets based on  
20       bearer rate measurements;

      In DL direction the IP packets at an APN are color-marked by the mobile GW;

      Green, if the considered bearer rate is below the target rate

25       Yellow, if the considered bearer rate is between target rate and APN-AMBR

      Discarded otherwise (red);

      During transport congestion situations "yellow" packets are dropped in transport elements with higher  
30       probability than "green" packets according to configured RED profiles

In the following two examples are given in context of the Figs. 13 and 14.

In particular, Fig. 13 shows a first example for a single default bearer per APN - Target and peak rate (APN-AMBR) applied to the default bearer. The DL target rate is set for the default bearer of an APN per subscriber class in the mobile GW according to operator policy. Furthermore, the applied peak rate is APN-AMBR. The IP packets of the default bearer are color-marked by the mobile GW, i.e. assigned a drop precedence (RED value). Green is used, if the default bearer rate is below the target rate. Yellow is used, if the default bearer rate is between target rate and APN-AMBR, and the packets are discarded otherwise (red).

In addition no UL color marking is applied and a per bearer differentiation is used during transport congestion in DL. During DL transport congestion situations "yellow" packets are dropped in congested transport elements with higher probability than "green" packets according to configured RED profiles. TCP flow control reacts to IP packet discard by reducing the TCP flow bandwidth. Resulting bandwidth per bearer thus tends to converge to the assigned target rate. Per bearer target rate and coloring, and its use in the transport network can be aligned with radio QoS bearer treatment

In particular, Fig. 14 shows a second example for application demotion with dedicated bearer. The DL target rate is set for the default bearer of an APN per subscriber class in the mobile GW according to operator policy. The demoted application uses a dedicated bearer with corresponding low QoS class settings. The aggregate dedicated and default bearer peak rate applied is APN-AMBR. IP packets of the dedicated bearer are color-marked by the mobile GW (RED value), wherein "yellow" is used, if below APN-AMBR.

IP packets of the default bearer are color-marked by the mobile GW (RED value): green, if the default bearer rate is below the target rate; and yellow, if the default bearer rate is between target rate and APN-AMBR. IP packets are discarded, if aggregated default and dedicated bearer rate is above APN-AMBR.

During transport congestion the following characteristics may be achieved:

Per bearer differentiation during transport congestion in DL.

During DL transport congestion situations "yellow" packets are dropped in congested transport elements with higher probability than "green" packets according to configured RED profiles.

Dedicated bearer (demoted application) is treated with higher drop precedence.

TCP flow control reacts to IP packet discard by reducing the TCP flow bandwidth.

Resulting bandwidth for dedicated bearer is close to 0 in the worst case, default bearer converges to target rate.

In the claims, any reference signs placed in parentheses shall not be construed as limiting the claims. The word "comprising" and "comprises", and the like, does not exclude the presence of elements or steps other than those listed in any claim or the specification as a whole. The singular reference of an element does not exclude the plural reference of such elements and vice-versa. In a device claim enumerating several means, several of these means may be embodied by one and the same item of software or hardware. The mere fact that certain measures are recited in mutually

different dependent claims does not indicate that a combination of these measures cannot be used to advantage.

## List of Abbreviations

|    |          |   |
|----|----------|---|
|    | 3GPP     | 3rd Generation Partnership Project          |
|    | AF       | Assured Forwarding                          |
| 5  | AMBR     | Aggregate Maximum Bit Rate                  |
|    | APN      | Access Point Name                           |
|    | BB       | Broadband                                   |
|    | BBF      | Broadband Forum                             |
|    | BE       | Best Effort                                 |
| 10 | BTS      | Base Transceiver Station                    |
|    | CBS      | Committed Burst Size                        |
|    | CIR      | Committed Information Rate                  |
|    | DIFFSERV | Differentiated Services                     |
|    | DL       | Down Link                                   |
| 15 | DSCP     | DiffServ Code Point                         |
|    | EBS      | Excess Burst Size                           |
|    | ECN      | Explicit Congestion Notification            |
|    | EF       | Expedited Forwarding                        |
|    | EIR      | Excess Information Rate                     |
| 20 | eNB      | Evolved Node B (also abbreviated as eNodeB) |
|    | FUP      | Fair Usage Policy                           |
|    | GBR      | Guaranteed Bit Rate                         |
|    | GGSN     | Gateway GPRS Support Node                   |
|    | GW       | Gateway GPRS Support Node                   |
| 25 | HSPA     | High Speed Packet Access                    |
|    | iBTS     | Internet BTS                                |
|    | IETF     | Internet Engineering Task Force             |
|    | iHSPA    | Internet HSPA                               |
|    | LTE      | Long Term Evolution                         |
| 30 | MBH      | Mobile Back Haul                            |
|    | MEF      | Metro Ethernet Forum                        |
|    | HSS      | Home Subscriber Server                      |
|    | MWR      | Micro Wave Radio                            |
|    | NB       | Node B                                      |
| 35 | NGMN     | Next Generation Mobile Networks             |
|    | PCN      | Pre Congestion Notification                 |
|    | PGW      | Packet Gateway                              |
|    | PHB      | Per Hop Behavior                            |



|    |       |                               |
|----|-------|-------------------------------|
|    | QCI   | QoS Class Identifier          |
|    | QoS   | Quality of Service            |
|    | RED   | Random Early Discard          |
|    | RNC   | Radio Network Controller      |
| 5  | SGW   | Signaling Gateway             |
|    | TCM   | Three Color Marking           |
|    | TCP   | Transmission Control Protocol |
|    | TOP   | Timing over Packet            |
|    | trTCM | Two Rate Three Color Marker   |
| 10 | UE    | User Equipment                |
|    | UL    | Uplink                        |
|    | WFQ   | Weighted Fair Queuing         |

## CLAIMS

1. A method for conveying traffic across a packet oriented  
5 mobile service network, the method comprising:  
    assigning a target bit rate and a peak bit rate to each  
    bearer traffic stream;  
    marking in a gateway and/or radio transceiver data  
    packets of a bearer traffic stream according to their  
10 compliance with the target bit rate and/or peak bit rate  
    assigned to the respective bearer before passing them on to a  
    mobile backhaul network for delivery towards the radio  
    transceiver.
- 15 2. The method according to claim 1, wherein the bearer  
    traffic stream is a downlink bearer traffic stream and/or an  
    uplink traffic bearer stream.
3. The method according to claim 1 or 2, wherein packets  
20 exceeding the peak rate are immediately dropped and packets  
    violating the target bit rate are marked for preferential  
    dropping in case of mobile backhaul congestion.
4. The method according to any one of the claims 1 to 3,  
25 further comprising:  
    forwarding the marked packets by the mobile backhaul  
    network.
5. The method according to any one of the claims 1 to 4,  
30 further comprising:  
    in case of congestion in the mobile backhaul network  
    preferably dropping packets marked for preferential dropping.

6. The method according to any one of the claims 1 to 5,  
wherein the target bit rate and/or the peak bit rate for each  
bearer is derived from the bandwidth assigned to this bearer  
by the air interface scheduler of the radio transceiver for  
5 transmission of the bearer data stream across the air  
interface.

7. The method according to any one of the claims 1 to 6,  
wherein the target bit rate and/or the peak bit rate for each  
10 bearer is set according to an access point name of the bearer  
and/or a subscriber class of the bearer.

8. The method according to any one of the claims 1 to 7,  
wherein an actual bit rate of at least one bearer traffic  
15 stream is measured.

9. The method according to any one of the claims 1 to 8,  
wherein the bearer is a non-guaranteed bit rate bearer.

20 10. The method according to any one of the claims 1 to 9,  
wherein the different kinds of markings are used for  
different drop precedencies.

11. The method according to claim 10, wherein a single  
25 transport queue is used for packets having different kinds of  
marking.

12. A mobile service network comprising means for executing  
a method according to any of the claims 1 to 11.

30 13. A gateway of a mobile service network comprising means  
to:

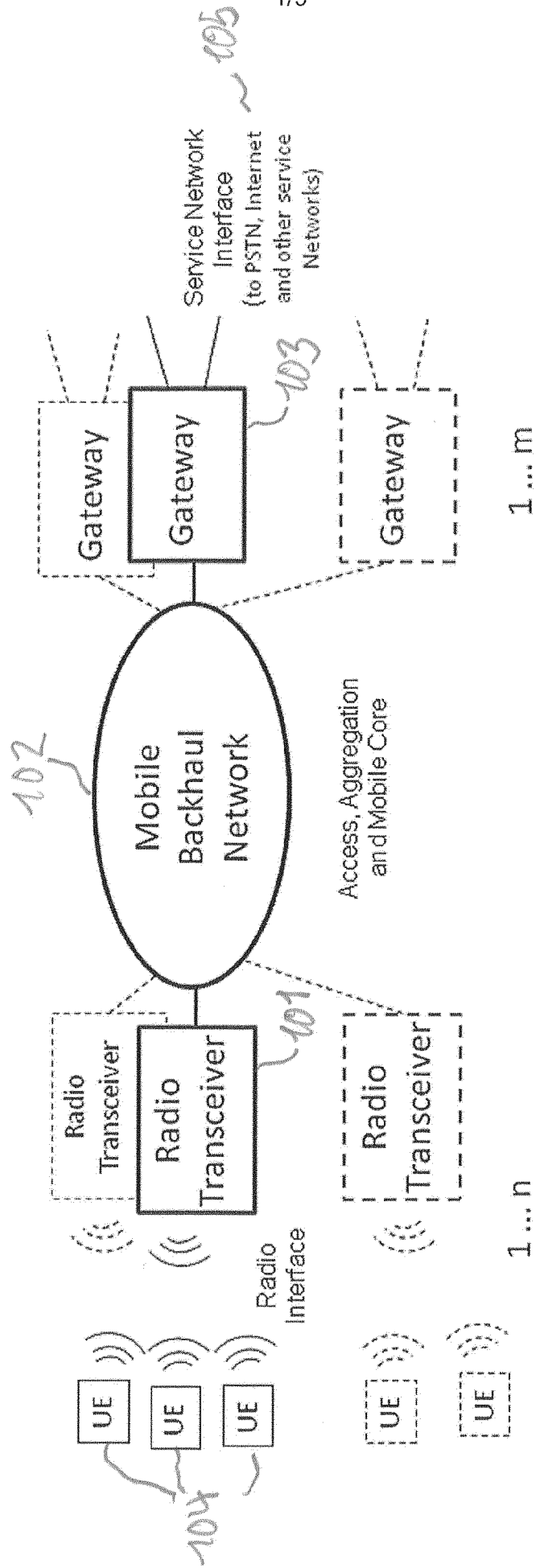
receive target bit rate and peak bit rate information  
assigned to bearer traffic streams; and

mark and/or drop data packets of a bearer traffic stream according to their compliance with the target bit rate and/or peak bit rate assigned to the respective bearer before passing them on to the mobile backhaul network for delivery  
5 towards a radio transceiver, wherein packets exceeding the peak rate are immediately dropped and packets violating the target bit rate are marked for preferential dropping in case of mobile backhaul congestion.

10 14. A computer program product comprising software, which when loaded into the memory of a computer enables the computer to execute any of the steps of the method of any one of the claims 1 to 12.

15 15. A computer program product comprising software, which when loaded into the memory of a computer enables the computer to implement any of the means comprised by a gateway according to claim 13.

100 →



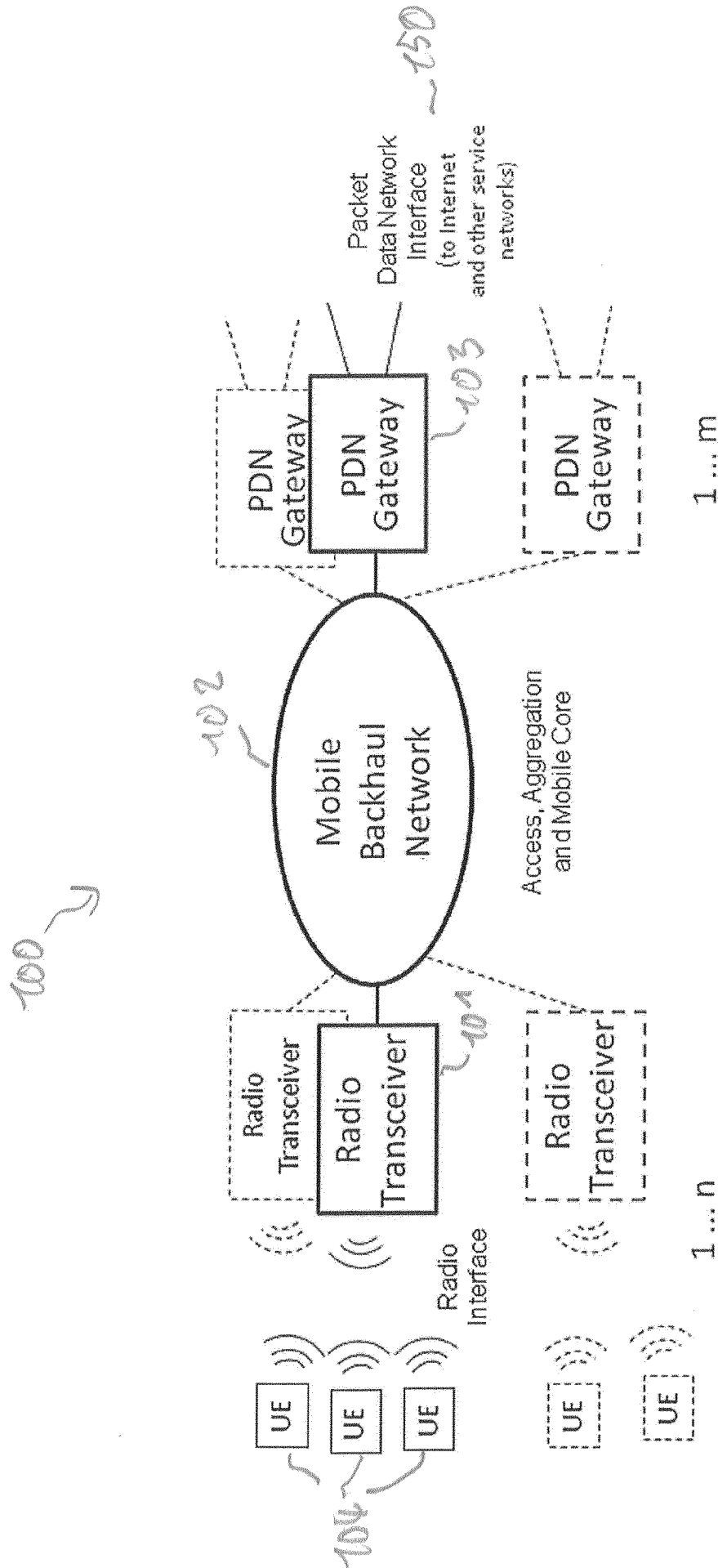


Fig 2

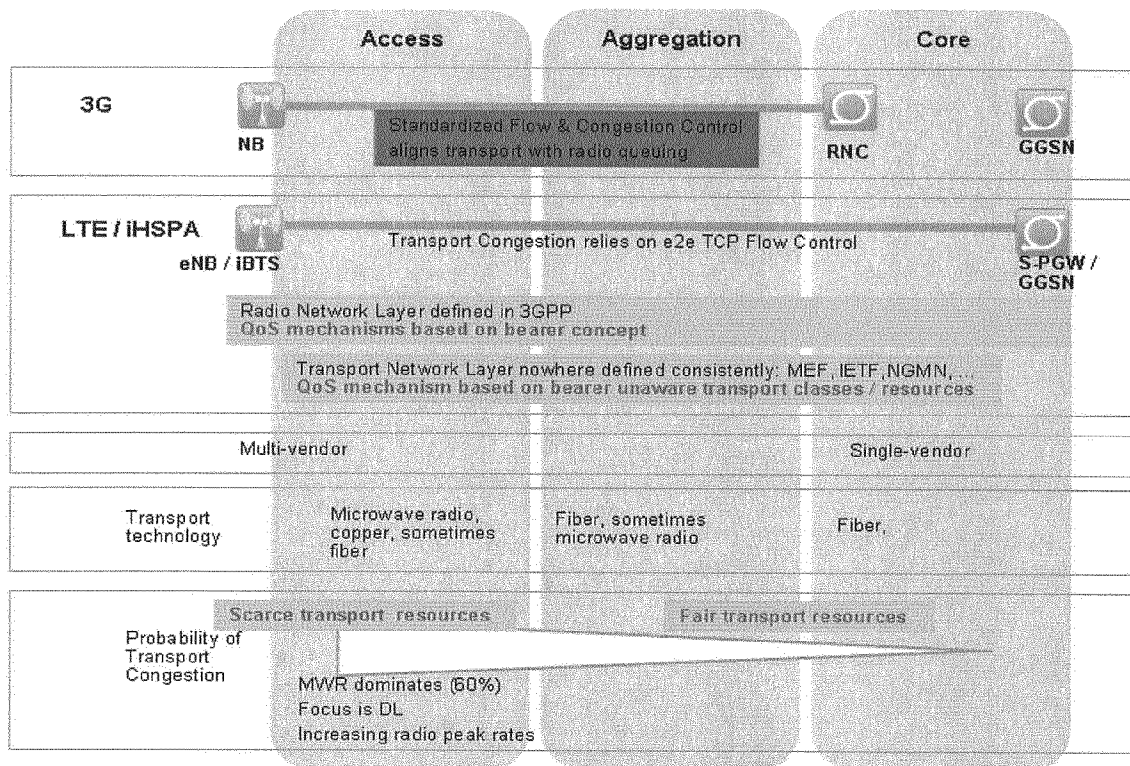


Fig. 3

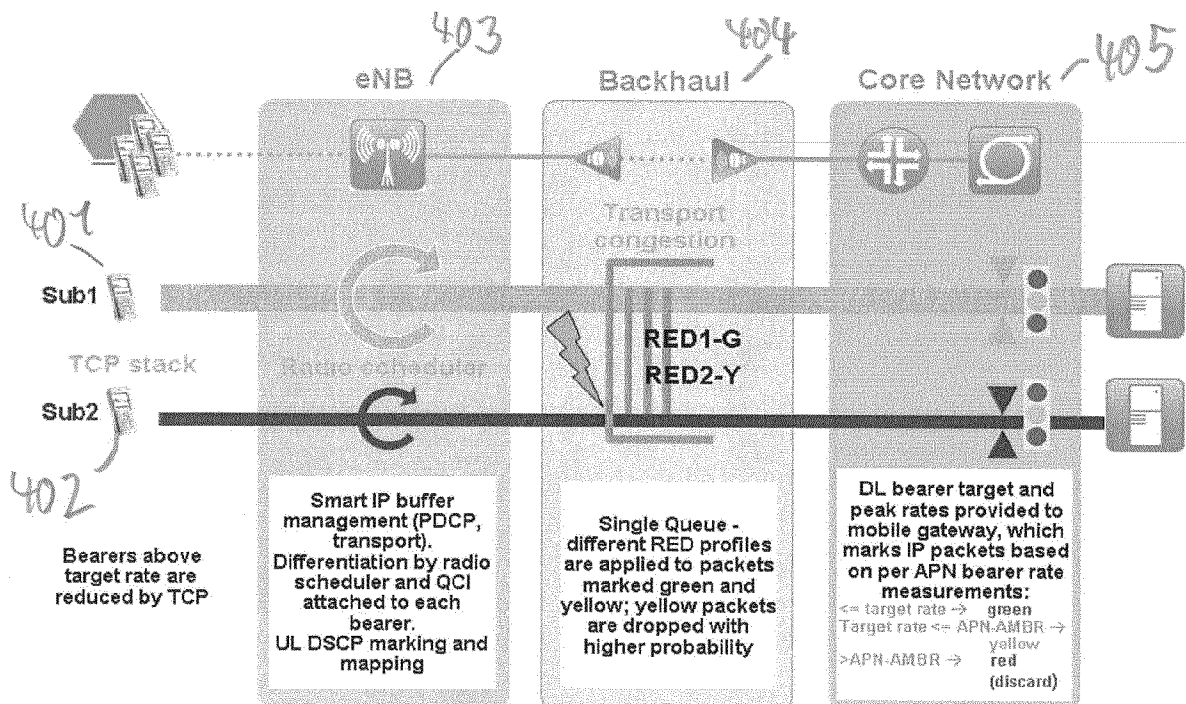


Fig. 4

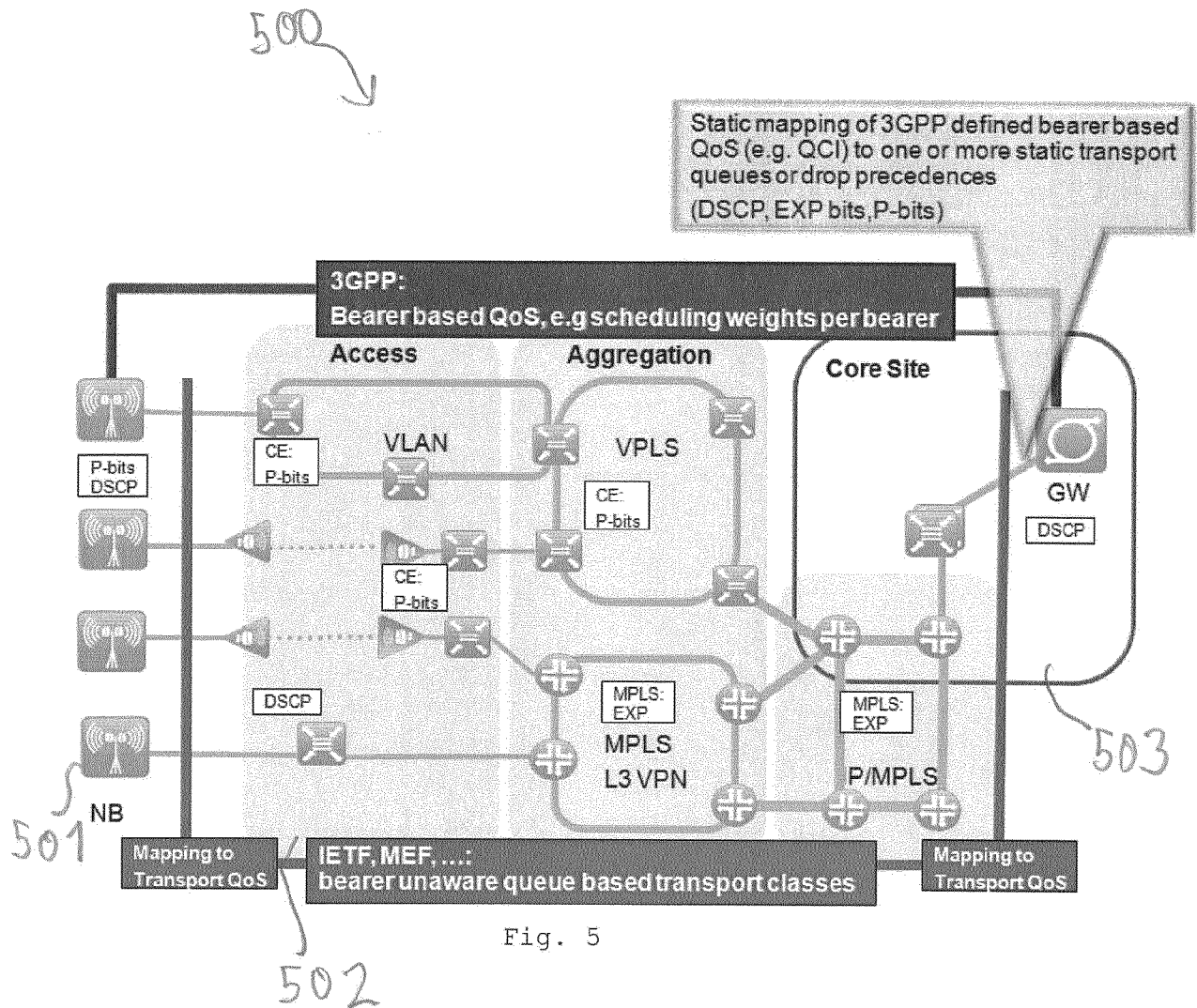


Fig. 5



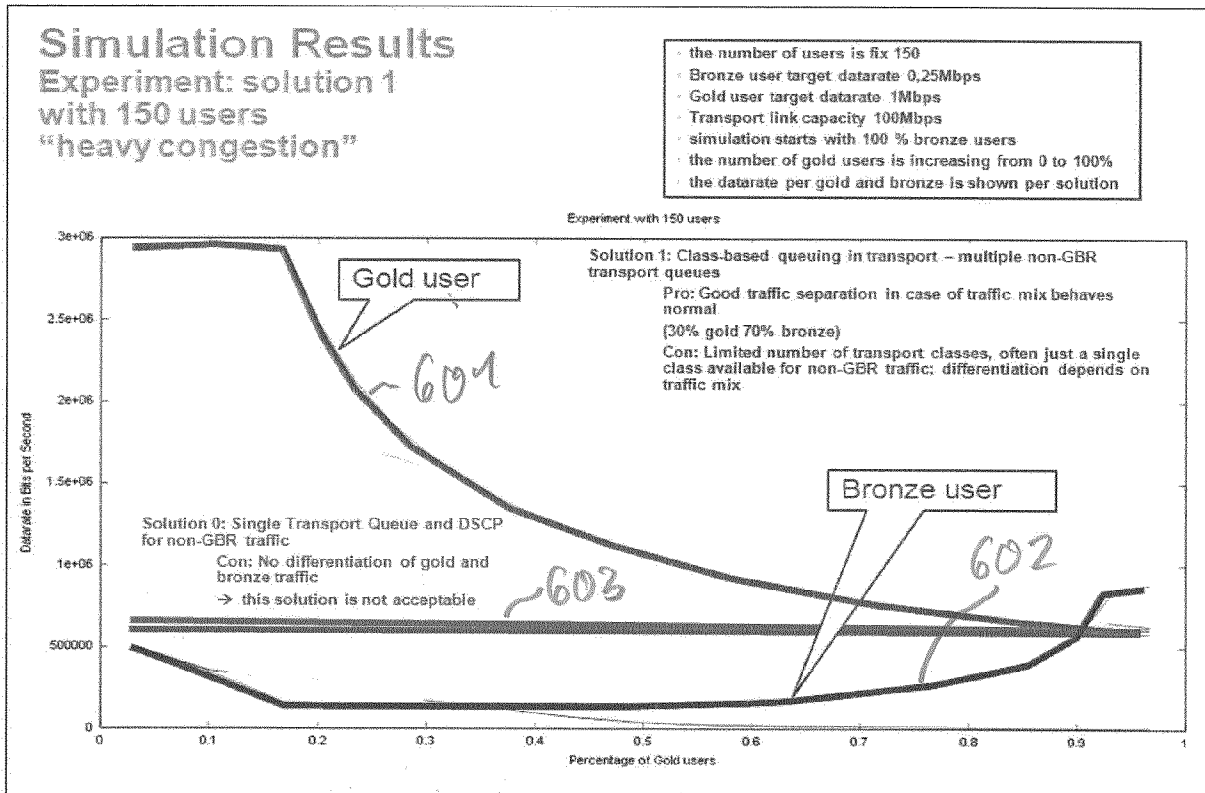


Fig. 6

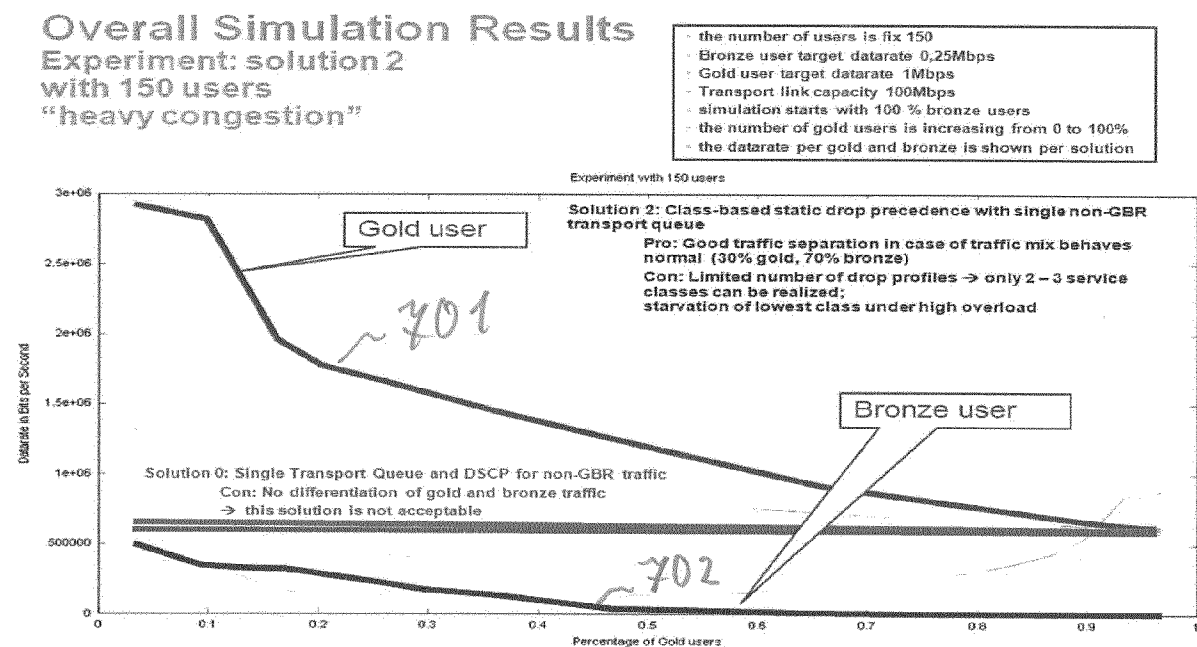


Fig. 7

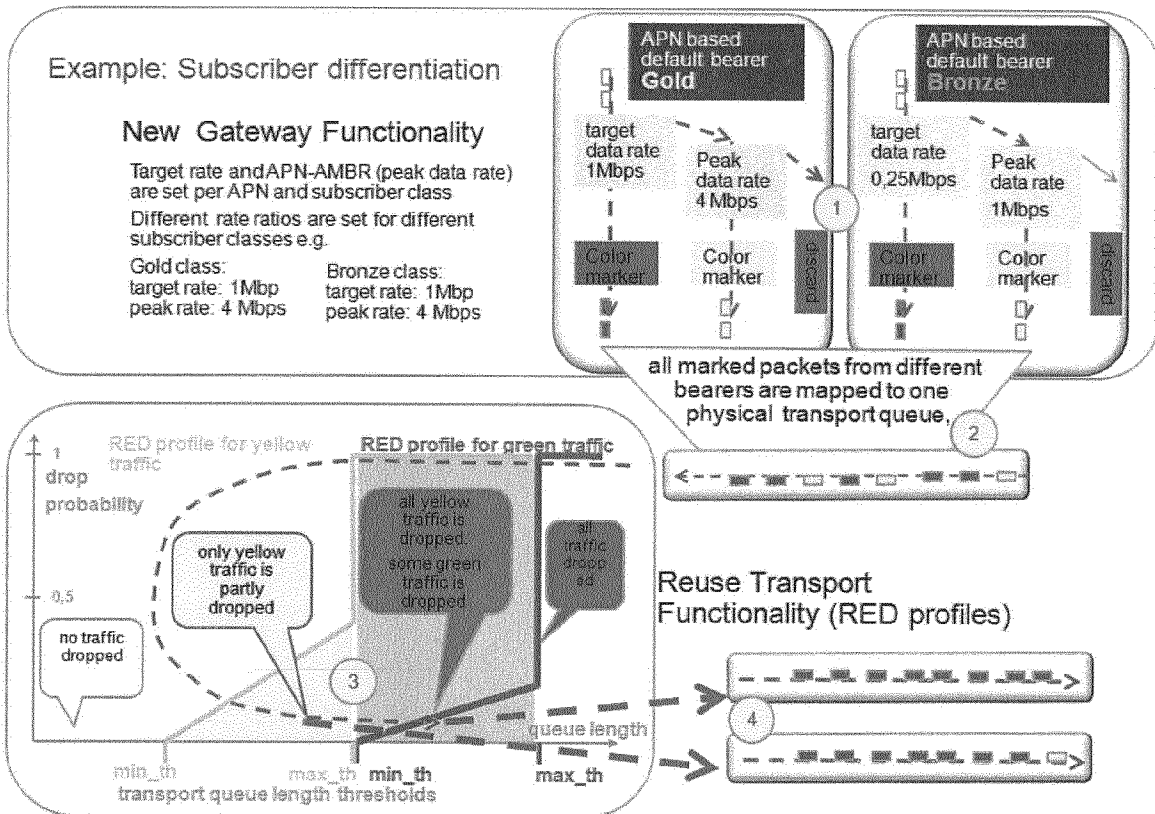


Fig. 8

### Color marking for traffic differentiation: use case examples for two traffic classes (trTCM)

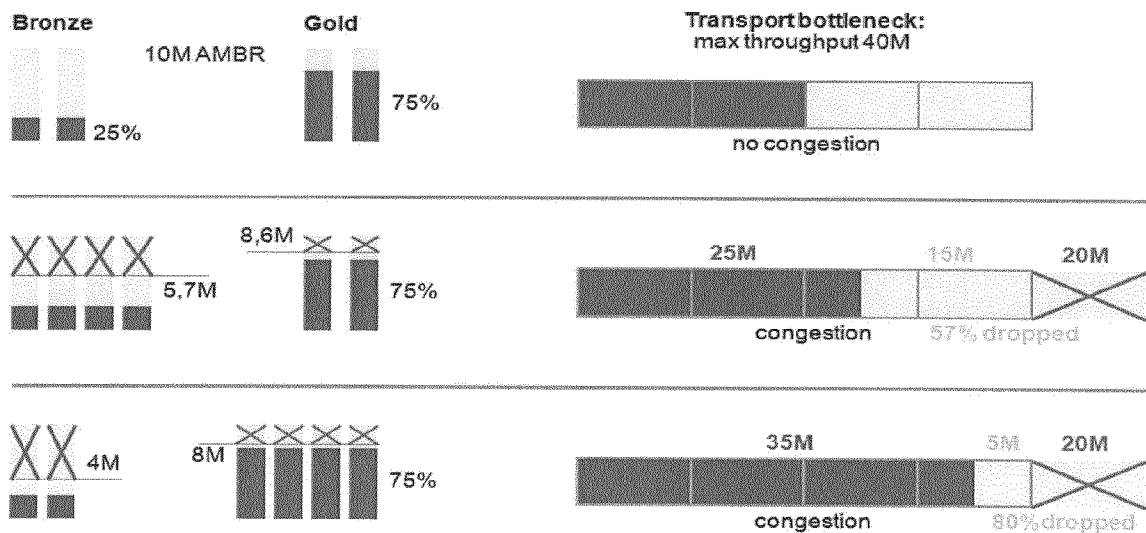


Fig. 9

# Experiment: solution 3 (invention) with 150 users "heavy congestion"

- the number of users is fix 150
- Bronze user target datarate 0,25Mbps
- Gold user target datarate 1Mbps
- Transport link capacity 100Mbps
- simulation starts with 100 % bronze users
- the number of gold users is increasing from 0 to 100%
- the datarate per gold and bronze is shown per solution

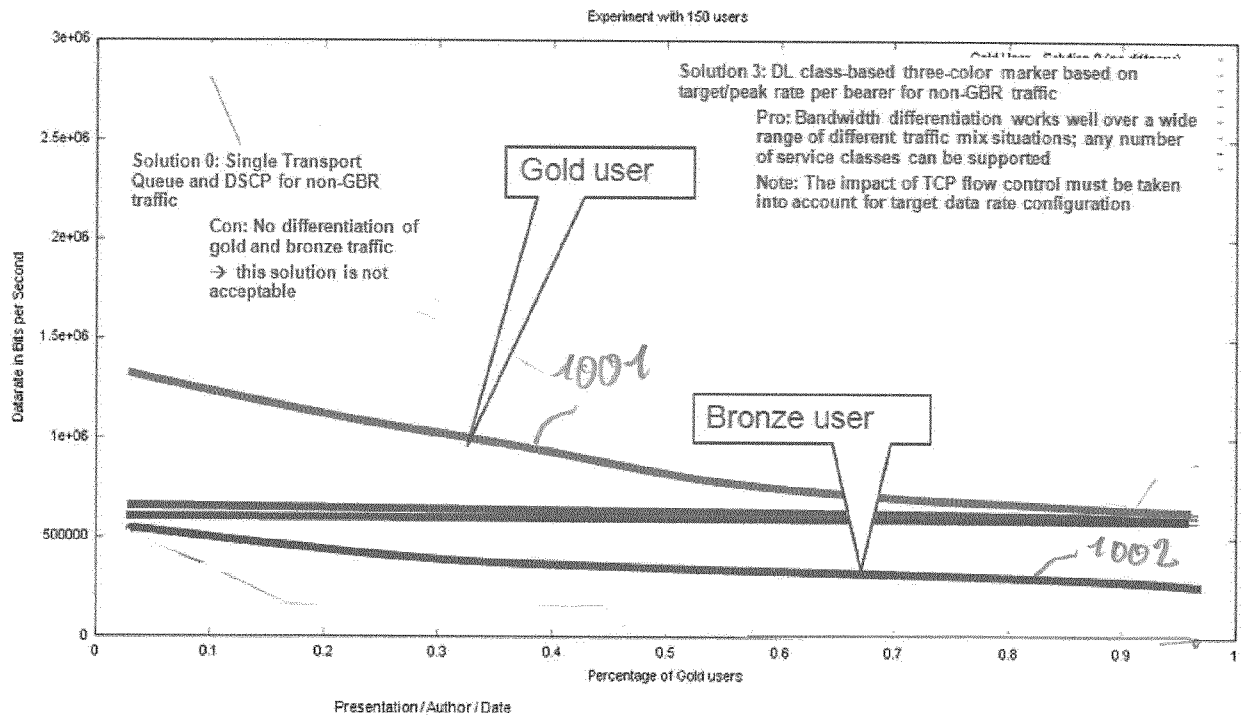


Fig. 10

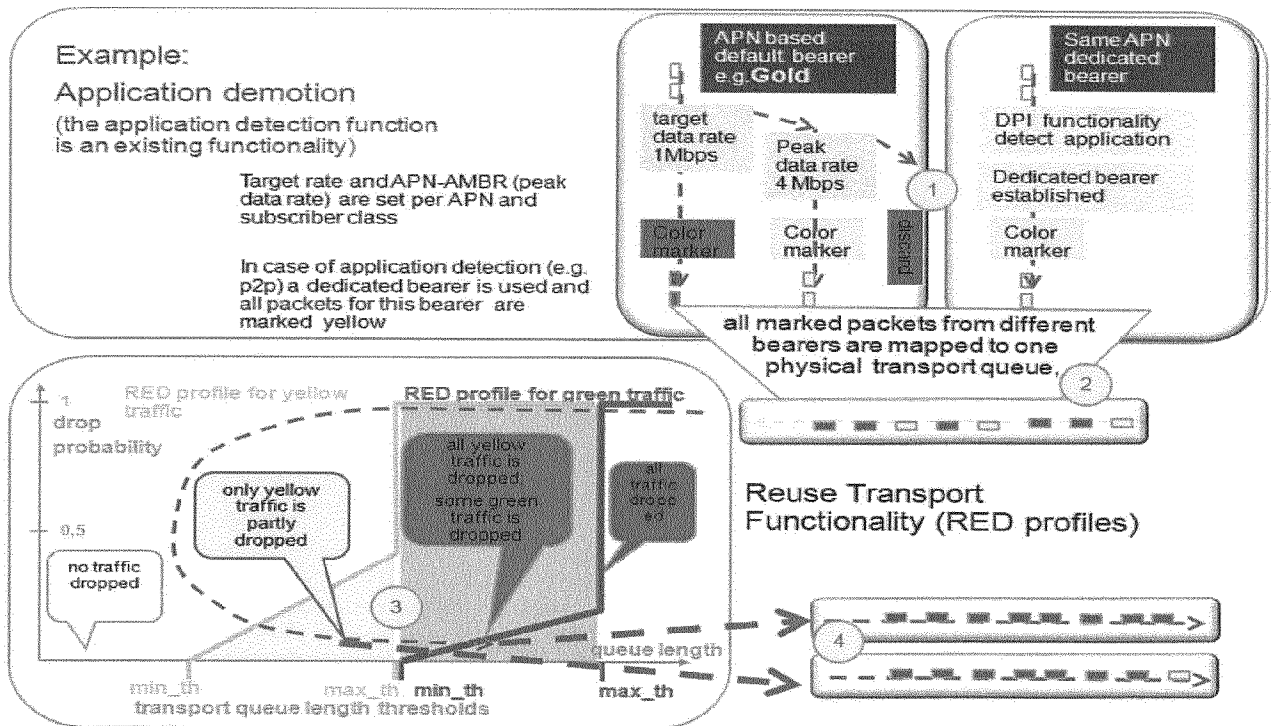
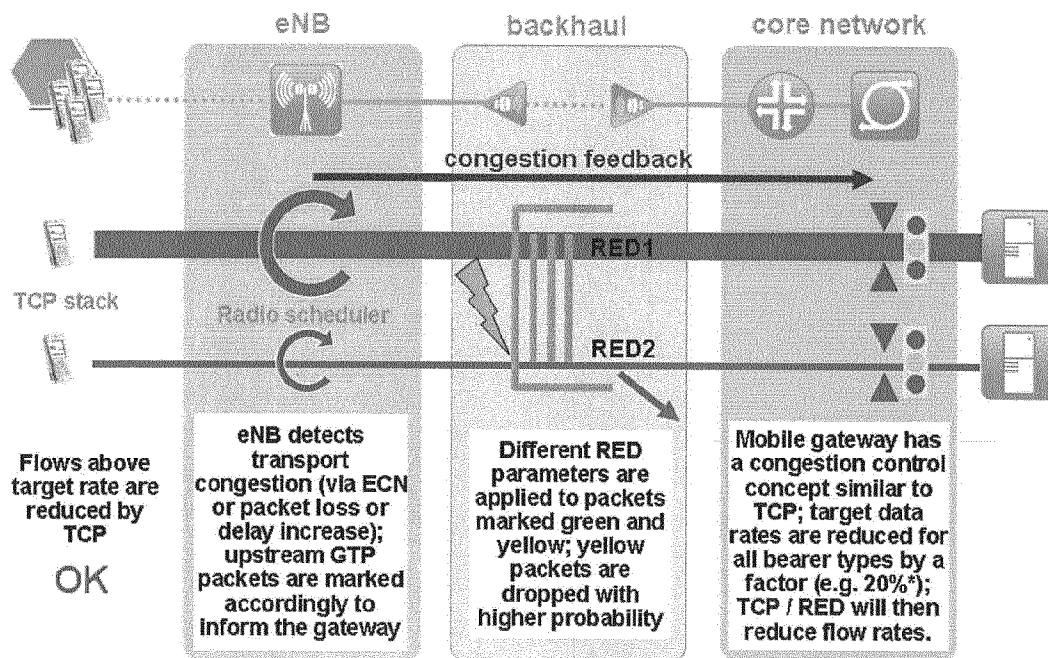


Fig. 11



\* reduction factor may also be dynamically chosen to enforce a bandwidth ratio between classes

Fig. 12

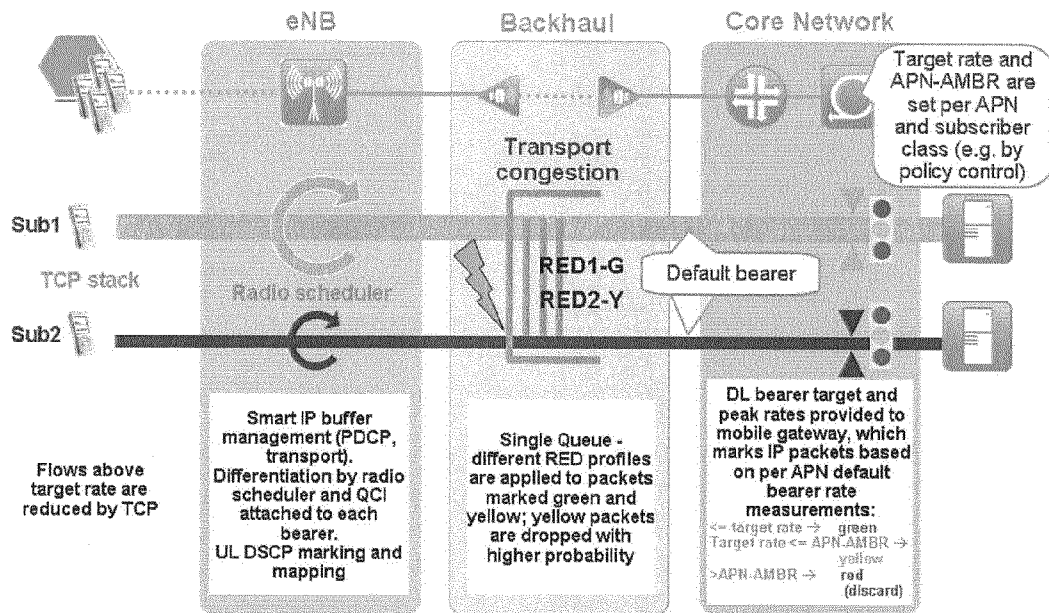


Fig. 13

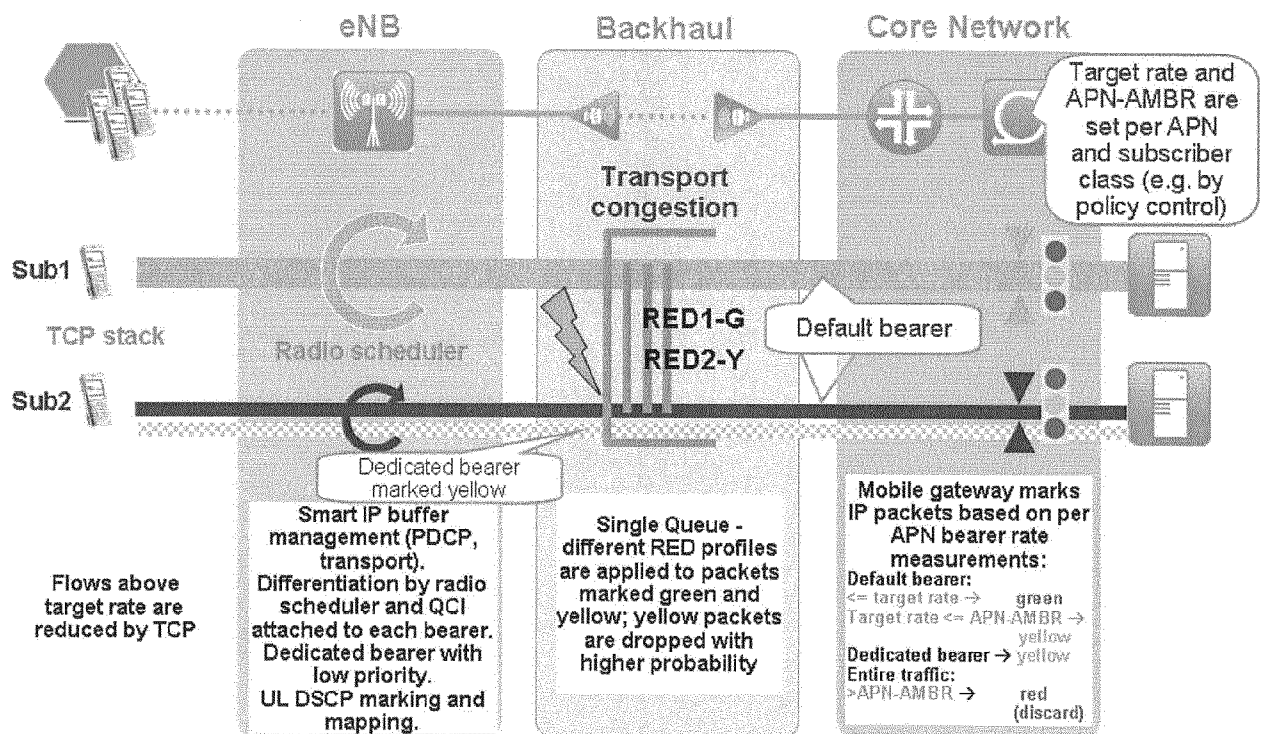


Fig. 14

# INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2014/053394

| <b>A. CLASSIFICATION OF SUBJECT MATTER</b><br>INV. H04W28/02 H04L12/833 H04L12/823<br>ADD.   |   |                       |
|--|---|-----------------------|
| According to International Patent Classification (IPC) or to both national classification and IPC  |   |                       |
| <b>B. FIELDS SEARCHED</b><br>Minimum documentation searched (classification system followed by classification symbols)<br>H04W H04L  |   |                       |
| 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 practicable, search terms used)<br>EPO-Internal, WPI Data   |   |                       |
| <b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>  |   |                       |
| Category*  | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
| Y  | WO 02/052801 A1 (NOKIA CORP [FI]; RUUTU JUSSI [FI]; SUOKNUUTI MARKO [FI])<br>4 July 2002 (2002-07-04)<br>page 4, line 4 - line 21<br>page 16, line 1 - line 7<br>figure 1   | 1-15                  |
| Y  | -----<br>EP 1 585 265 A1 (MCI INC [US])<br>12 October 2005 (2005-10-12)<br>paragraphs [0200], [0201]<br>figure 17   | 1-15                  |
| A  | -----<br>WO 2011/076384 A1 (NEC EUROPE LTD [DE]; KUTSCHER DIRK [DE]; MIR FAISAL-GHIAS [DE]; NUNZI) 30 June 2011 (2011-06-30)<br>page 3, line 28 - page 4, line 3<br>page 5, line 7 - line 23<br>page 8, line 25 - page 9, line 9<br>-----<br>-/-- | 1-15                  |
| <div style="display: flex; justify-content: space-between;"> <span><input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.</span> <span><input checked="" type="checkbox"/> See patent family annex.</span> </div>  |   |                       |
| <div style="display: flex;"> <div style="flex: 1;"> <p>* Special categories of cited documents :</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"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)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="flex: 1;"> <p>"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</p> <p>"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</p> <p>"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</p> <p>"&amp;" document member of the same patent family</p> </div> </div> |   |                       |
| Date of the actual completion of the international search<br><br><div style="text-align: center; font-size: 1.2em;">13 June 2014</div>   | Date of mailing of the international search report<br><br><div style="text-align: center; font-size: 1.2em;">23/06/2014</div>   |                       |
| Name and mailing address of the ISA/<br>European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040,<br>Fax: (+31-70) 340-3016   | Authorized officer<br><br><div style="text-align: center; font-size: 1.2em;">Pérez Pérez, José</div>  |                       |

## INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2014/053394

| C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT |  |                       |
|--|--|-----------------------|
| Category*  | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No. |
| A  | HEINANEN TELIA FINLAND R GUERIN UNIVERSITY<br>OF PENNSYLVANIA J: "A Two Rate Three<br>Color Marker; rfc2698.txt",<br>19990901,<br>1 September 1999 (1999-09-01),<br>XP015008481,<br>ISSN: 0000-0003<br>the whole document<br>----- | 1-5                   |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2014/053394

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s)  | Publication<br>date      |
|---|---------------------|-----------------------------|--------------------------|
| WO 02052801                               | A1                  | 04-07-2002                  | AT 402544 T 15-08-2008   |
|   |                     | EP 1344354 A1 17-09-2003    |                          |
|   |                     | US 2004090917 A1 13-05-2004 |                          |
|   |                     | WO 02052801 A1 04-07-2002   |                          |
| -----                                     |                     |                             |                          |
| EP 1585265                                | A1                  | 12-10-2005                  | EP 1585265 A1 12-10-2005 |
|   |                     | US 2005220014 A1 06-10-2005 |                          |
|   |                     | US 2010040206 A1 18-02-2010 |                          |
| -----                                     |                     |                             |                          |
| WO 2011076384                             | A1                  | 30-06-2011                  | EP 2468036 A1 27-06-2012 |
|   |                     | JP 2013512603 A 11-04-2013  |                          |
|   |                     | US 2012257505 A1 11-10-2012 |                          |
|   |                     | WO 2011076384 A1 30-06-2011 |                          |
| -----                                     |                     |                             |                          |