

(51) International Patent Classification:
H04L 12/70 (2013.01)(21) International Application Number:
PCT/EP2013/063723(22) International Filing Date:
28 June 2013 (28.06.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
13165620.9 26 April 2013 (26.04.2013) EP(71) Applicant: NEC EUROPE LTD. [DE/DE]; Kurfuersten-
Anlage 36, 69115 Heidelberg (DE).(72) Inventors: YOUSAF, Faqir Zarrar; Im Schussgarten 11,
69181 Leimen (DE). MAEDER, Andreas; Randersackerer
Strasse 30, 97072 Wuerzburg (DE). SCHMID, Stefan; 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, 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, 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, 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))

(54) Title: METHOD FOR OPERATING A WIRELESS NETWORK AND A WIRELESS NETWORK

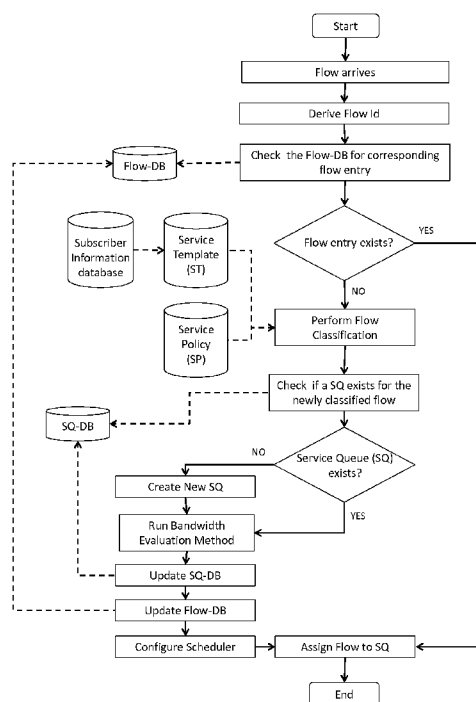


Fig. 1

(57) Abstract: For ensuring a high degree of fairness amongst multiple simultaneous heterogeneous data flows while meeting the end-user QoE expectations a method for operating a wireless network is claimed, wherein data flows aggregating at a traffic aggregation point will be managed for controlling congestion of data flows at said traffic aggregation point, the method comprising the following managing steps: classifying at least one data flow according to a rule used to differentiate traffic of said data flows for providing a service signature of said at least one data flow; assigning said at least one data flow to a Service Queue, SQ, in the form of a Per Flow Queue, PFQ, or an Aggregate Flow Queue, AFQ, based on said at least one data flow's service signature; and allocating of available bandwidth amongst the Service Queues based on a QoE estimation, such that the QoE of said data flow and/or the overall QoE estimation over all data flows traversing said aggregation point is improved. Further an according wireless network is claimed, preferably for carrying out the above mentioned method.

METHOD FOR OPERATING A WIRELESS NETWORK AND A WIRELESS NETWORK

The present invention relates to a method for operating a wireless network, wherein data flows aggregating at a traffic aggregation point will be managed for controlling congestion of data flows at said traffic aggregation point. Further, the present invention relates to a wireless network, wherein data flows aggregating at a traffic aggregation point will be managed by a system for controlling congestion of data flows of said traffic aggregation point.

Mobile networks are experiencing a data explosion that is forecasted to grow exponentially in the near future. This data explosion has been spurred by the proliferation of smart phones and other mobile devices that are capable of handling high bandwidth applications, such as live/progressive video streams and even video conferencing applications.

This data explosion, in view of finite network resources, e.g., bandwidth, is the main cause of congestion in the network with an adverse effect on the QoS (Quality of Service) of different applications, thereby affecting the users' QoE (Quality of Experience). Mobile network operators (MNOs) are under pressure to meet the expected QoE demands of such mobile users in order to extract maximum revenue from their existing infrastructure, while at the same time attract more customers to increase their revenue base. Increasing the physical capacity and resources of existing network infrastructure to handle burgeoning traffic is no longer a feasible option due to high CAPEX/OPEX (Capital Expenditure/Operational Expenditure) and the ever increasing gap between the traffic and ARPU (Average Revenue per User). There is therefore a need to devise a method/system that will be able to dynamically and intelligently manage multiple flows at network traffic aggregation point, in order to prevent the buildup of congestion, while meeting the individual flows' QoS/QoE requirements on a priority basis.

Existing systems utilize standard end-to-end protocols like ECN (Explicit Congestion Notification) and TCP (Transmission Control Protocol) for congestion

- 2 -

notification and control. To prevent against congestion many queuing and scheduling policies have been proposed in order to allocate available bandwidth fairly amongst multiple traffic flows. Some of the more popular of them are Priority Queuing (PQ), Fair Queuing (FQ) Weighted fair Queuing (WFQ), Weighted Round Robin (WRR), Deficit Round Robin (DRR) scheduling, to name a few, where each one has specific advantages and disadvantages. The more advanced scheduling policies such as WFQ and DRR assume that all traffic that belongs to a specific class is aggregated in a separate queue, and that each queue is assigned a specific weight according to the traffic class, corresponding to the amount of packets that are transmitted from these class based queues.

Most of these methods ensure efficient utilization of bandwidth and provide some degree of QoS and fairness for bandwidth allocation at an aggregate level to flows that belong to a specific QoS class. However, none of these methods take into account their effect on the end-user QoE, which may not exactly be up to user expectation, even if the network is meeting all its KPIs (Key Performance Indicator). For example, a high priority queue assigned for video streams may be meeting the QoS thresholds as an aggregate, although a significant number of individual flows are experiencing bad service conditions resulting in poor viewing quality for users of those disadvantaged flows.

It is an object of the present invention to improve and further develop a method for operating a wireless network and an according network for ensuring a higher degree of fairness amongst multiple simultaneous heterogeneous flows while meeting the end-user QoE expectations.

In accordance with the invention, the aforementioned object is accomplished by a method comprising the features of claim 1 and by a network comprising the features of claim 26.

According to claim 1 the method is characterized by the following managing steps:

- classifying at least one data flow according to a rule used to differentiate traffic of said data flows for providing a service signature of said at least one data flow;

- 3 -

- assigning said at least one data flow to a Service Queue, SQ, in the form of a Per Flow Queue, PFQ, or an Aggregate Flow Queue, AFQ, based on said at least one data flow's service signature; and
- allocating of available bandwidth amongst the Service Queues based on a QoE estimation, such that the QoE of said data flow and/or the overall QoE estimation over all data flows traversing said aggregation point is improved.

According to claim 28 the network is characterized by:

- means for classifying at least one data flow according to a rule used to differentiate traffic of said data flows for providing a service signature of said at least one data flow;
- means for assigning said at least one data flow to a Service Queue, SQ, in the form of a Per Flow Queue, PFQ, or an Aggregate Flow Queue, AFQ, based on said at least one data flow's service signature; and
- means for allocating of available bandwidth amongst the Service Queues based on a QoE estimation, such that the QoE of said data flow and/or the overall QoE estimation over all data flows traversing said aggregation point is improved.

According to the invention it has been recognized that it is possible to allow for a more sophisticated and fair queuing concept amongst different data flows. In a first step of the method at least one data flow will be classified according to a rule used to differentiate traffic of said data flows for providing a service signature of said at least one data flow. The provided service signature will be used in the next step wherein at least one data flow will be assigned to a Service Queue in the form of a Per Flow Queue or an Aggregate Flow Queue based on said at least one service signature of the at least one data flow. Finally, available bandwidth will be allocated amongst the Service Queues based on a QoE estimation. Such a QoE estimation will be used for the allocation of available bandwidth amongst the Service Queues for ensuring a higher degree of fairness amongst the multiple data flows.

- 4 -

The invention provides a QoE/context aware priority queuing concept that will ensure a very high degree of fairness amongst multiple simultaneous heterogeneous flows while meeting the end-user QoE expectations. Further, the invention proposes a more dynamic and diverse queuing model that can be used
5 by many weight-based schedulers/traffic-shapers and can take into account a variety of context information and can provide a higher degree of flow isolation to prevent against bandwidth starvation due to misbehaving flows.

Within a preferred embodiment the method can be realized in an access device or
10 network element at said traffic aggregation point or as a middle box solution to be placed at said traffic aggregation point. In other words, the method can be employed at traffic aggregation points, either integrated in access devices at such locations or as a middle box solution to be placed at traffic aggregation points.

With regard to a very effective and simple classifying step the rule can comprise at
15 least one service template, ST. Such a service template is a rule used to differentiate traffic. In this context the rule can use at least one property of said at least one data flow to differentiate traffic. Said at least one property can preferably be, and not limited to, IP-5-tuples and/or application type and/or subscriber
20 information and/or header information and/or payload information and/or type of Service information and/or QoS information, particularly Differentiated Service Code Points, and/or QoS requirements and/or QoE requirements and/or security requirements and/or device capabilities. The individual application for use of one or more of the above mentioned properties will depend on the individual network
25 and/or application type.

Within a further preferred embodiment the rule can be associated with at least one Service Policy, SP, for defining a handling of said at least one classified data flow. On the basis of such a Service Policy a simple and effective classifying step can
30 be realized.

Preferably said at least one Service Policy defines the Service Queue type and/or a priority of handling and/or an associated bandwidth and/or a key performance indicator, KPI, of said at least one data flow.

- 5 -

5 Within the assigning step at least one data flow will be assigned to a Service Queue based on said service signature. If no suitable Service Queue is available or present a creation of at least one new Service Queue can be performed on the basis of a data flow's service signature. Thus, it can be ensured that suitable Service Queues are available for all types of data flows.

10 Within a further preferred embodiment a data flow being sensitive to QoS and/or QoE variations will trigger a creation of a new Per Flow Queue to which the data flow will be assigned to. Concretely, long-lived traffic flows which are sensitive to QoS/QoE variations such as adaptive or non-adaptive video streaming are assigned to Per Flow Queues in order to be able to meet the QoS/QoE requirements of traffic flows individually and to provide efficient isolation from other traffic. A data flow not being sensitive to QoS and/or QoE variations and/or being
15 not suitable for a per-flow scheduling will be assigned to an Aggregate Flow Queue. Traffic which is not suitable for a per-flow scheduling could be short-lived flows or bursty traffic such as web-traffic, instant messaging or signaling messages.

20 Preferably, data flows belonging to a specific application type and/or sharing the same priority will be assigned to a dedicated Aggregate Flow Queue, wherein preferably – in case the Aggregated Flow Queue has not been created – the first data flow associated with said AFQ will trigger the dynamic creation of said AFQ. Thus, an effective handling of such data flows can be ensured.

25 Within a further preferred embodiment data flows that can not be classified, e.g. encrypted data flows, will be assigned to a default Aggregate Flow Queue. Thus, all data flows can be assigned to suitable Service Queues.

30 For realizing a very sensitive and effective classifying of the at least one data flow the classifying step can comprise an analyzing step of at least one packet, preferably the first packet with application data payload, of the at least one data flow. Based on that, the method or network is able to detect the application type and can then instantiate a new Per Flow Queue or Aggregate Flow Queue, if no

- 6 -

adequate queue is available yet, or assign the new flow to an already existing AFQ.

5 Within a further preferred embodiment the analyzing step can comprise a Deep Packet Inspection, DPI. In case of flows whose application can not be detected by analyzing the first data packet only, i.e. if further packets need to be observed, they are assigned to a default queue and then assigned to an appropriate queue once its type has been identified.

10 With regard to a very effective and simple method as a first step – prior to the classifying step – a flow-id of an arriving packet, e.g. arriving at the traffic aggregation point, will be derived, a check of a Flow-Database, flow-db, on the basis of the flow-id will be performed to determine if the packet belongs to an already known or classified data flow, and in case the data flow already exists, the
15 packet or data flow will be directly assigned to its associated Service Queue without the classifying step, and in case the data flow does not exist, a new flow-id will be derived and the classifying step will be performed.

20 Preferably the flow-id can be a hash derived from a packet's header information and/or payload information, preferably depending on the system's DPI capabilities.

Further preferred the flow-db contains a mapping of the flow-ids with its corresponding service signature.

25 For realizing a very sophisticated and effective method, user and/or network context information from at least one control entity of the network can be used during said classifying and/or assigning and/or allocating step. Preferably, said user and/or network context information can be provided – preferably via an access and/or an interface – to said at least one control entity.

30 Generally and within a preferred embodiment the QoE estimation can be based on an application type and/or flow statistics. Such flow statistics can include average throughput over some time windows or the total flow time, jitter, delay etc. The

- 7 -

individual and preferred realization of such a QoE estimation can depend on individual application situations.

5 Within a further preferred embodiment for each Service Queue a target bandwidth can be computed and used as input to parameterize a scheduler within the allocating step. Preferably, the allocating step can provide maximum bandwidth utilization.

10 Within a further preferred embodiment the allocating step targets to maximize the QoE for all data flows under the constraint of limited network resources, which preferably can be bandwidth and/or processing power.

15 For further simplifying a bandwidth management a Service Profile Matrix, SPM, that lists the QoS and/or QoE requirements or QoS and/or QoE profiles of different service types in terms of target bitrates for different load or congestion levels can be used for determining the target bitrate for each Service Queue.

20 With regard to a very effective method QoS and/or QoE profiles and/or the QoE estimation of all the flows or of individual flows with lower priority according to a Service Profile Matrix, SPM, can be decremented, if the sum of all computed target bitrates exceeds the available/measured bandwidth.

25 For realizing a very effective bandwidth utilization the allocating step can be performed, when a new SQ is created or when a number of data flows assigned to an AFQ exceeds a specified number or during each evaluation epoch.

30 The above invention proposes a system's or network's method of dynamic queue management for effective congestion management that will allow for a high degree of demarcation and flow isolation between multiple traffic flows based on application or service requirements. Such a method will enable intelligent and QoE/context aware scheduling of flows in times of congestion, or even prevent against the buildup of congestion. With reference to the proposed method, it is described a novel bandwidth sharing/management method for optimizing per-application utility of the bottleneck capacity.

- 8 -

5 The above invention, which is aimed at managing user plane congestion while ensuring to meet user expected QoE, can either be employed in an existing mobile network entity that aggregates traffic, or it can be deployed as a middle box solution at the traffic aggregation point.

Important aspects of the invention are summarized as follows:

10 The invention proposes a system's or network's method of dynamic queue management for effective congestion management by QoS/QoE/context aware management of individual flows.

15 The proposed system or network method classifies flows and assigns them to a specific Service Queue (SQ), or dynamically instantiates a relevant new SQ if none exists, based on the flow's service signature. Once the flow session terminates, and no flows are associated to a particular queue, then the associated queue will be destroyed.

20 The creation of SQs can thus be driven by the traversing traffic flows and their classification according to service templates (STs), and the associated service policies (SPs). An ST is a rule used to differentiate traffic; it can use a multitude of properties such as IP-5-tuples, application type, subscriber information, etc. for differentiation. The SP then defines the handling of the categorized traffic, such as defining the associated queue type, the priority, associated bandwidth, key
25 performance indicators (KPIs), etc.

With respect to the invention, the system or network can in general have the following basic types of SQs:

- 30
1. Per Flow Queues (PFQ)
 2. Aggregate Flow Queue (AFQ)

Long-lived traffic flows which are sensitive to QoS/QoE variations such as adaptive or non-adaptive video streaming are assigned to PFQs in order to be

- 9 -

able to meet the QoS/QoE requirements of traffic flows individually and to provide efficient isolation from other traffic. Traffic which is either not suitable for per-flow scheduling, e.g. short-lived flows/traffic bursts such as web-traffic, instant messaging or signaling messages, or which is less sensitive to QoS/QoE variations is put to AFQs.

The aggregation can be based on common application type, where all flows belonging to a specific application type and sharing the same priority are aggregated in a dedicated AFQ. For example, all web traffic flows for users with the same subscription level, e.g. "gold", may be aggregated in a dedicated AFQ with certain QoS properties such that on average, the expectations of the users with this subscription level for web traffic are fulfilled. A separate AFQ can be instantiated for users with "bronze" subscription. All other flows that could not be classified, e.g. encrypted flows, can be aggregated in a separate AFQ, and handled accordingly. The latter type can be considered as a default queue which is not application specific, but aggregates all unclassified flows.

Similarly, high priority signaling traffic, for example TCP SYN/ACK messages or IMS signaling messages, can be aggregated in a dedicated AFQ with a high priority. If the application type could not already be detected by the src/dst IP address and port pairs and protocol type, i.e. "five tuple", the system will analyze the first packet with application data payload, e.g. through DPI. Based on that, the system is able to detect the application type and can then instantiate a new PFQ or AFQ, if no adequate queue is available yet, or assign the new flow to an already existing AFQ. In case of flows whose application cannot be detected by analyzing the first data packet only, i.e. if further packets need to be observed, they are assigned to a default queue and then assigned to an appropriate queue once its type has been identified.

A system or network employing the method can have a multiple instances of PFQ and AFQ, the numbers of which will change dynamically, owing to traffic classes.

Important aspects of embodiments of the invention:

- 10 -

- 1) Effective management of user plane congestion at traffic aggregate points
- 2) The system will ensure that the flows satisfy the user expected QoE
- 5 3) Independent of any specific scheduler, and can work with any weight based fair scheduler.
- 4) Highly scalable due to on-demand queue creation and the underlying periodic application aware dynamic bandwidth evaluation/distribution and
10 assignment

Advantages of embodiments of the invention:

15 The advantage is

1. Higher degree of flow isolation on priority basis
2. Fair distribution of total available bandwidth based on flow priority
- 20 3. Low probability of bandwidth starvation
4. Effective and scalable user plane congestion management
- 25 5. Application/QoE aware flow classification and queuing
6. On-demand queue creation and release.

30 The effectiveness of the invention depends on the system's ability to accurately identify the flow's application type.

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 claim 1 on the one hand and to the following

- 11 -

5 explanation of preferred examples of embodiments of the invention, illustrated by the drawing on the other hand. In connection with the explanation of the preferred embodiments of the invention by the aid of the drawing, generally preferred embodiments and further developments of the teaching will be explained. In the drawings

Fig. 1 is a general overview of an embodiment of a method according to the invention in the form of a flow diagram,

10 Fig. 2 is showing important features of an example system or network employing and managing a hybrid queuing system according to an embodiment of the invention,

15 Fig. 3 is showing a generic overview of a bandwidth evaluation/distribution method according to an embodiment of the invention and

Fig. 4 is showing an example of a Service Profile Matrix.

20 Fig. 1 shows a general process overview of how a data flow is classified and assigned to a specific queue type according to an embodiment of the invention. When a packet arrives, the system will do the following:

- 25 1. Derive the flow-id. The flow-id can be a hash derived from the packet's header information, and possible payload information depending on the system's DPI capabilities.
- 30 2. Check the Flow-Database, flow-db, to determine if the packet belongs to an already known/classified flow. The flow-db will contain a mapping of the flow-id with its service signature.
3. In case the flow exists, the packet will be directly assigned to its associated queue.

- 12 -

4. If the flow entry is not found, then the system will derive a new flow-id and perform flow classification based on the header information, and additionally payload information. For flow classification, the supporting system will maintain and/or have access to a Service Template (ST) database. A ST will typically have preconfigured filters that will discriminate between different flows based on, but not limited to, the application type, header information, payload type, QoS/QoE requirements, security requirements etc. The ST will also have access to subscriber profile information, accessed from an external/internal Subscriber Information Database, to derive flow profiles for same application type but for different user profiles. The classifier will derive/assign a service signature for the new flow based on the flow profile match obtained from the ST and the Service Policy (SP) rule defined by the operator defining how the flow should be managed, for example, defining the queue type. Hence the classifier uses the input from the ST and SP to derive a service signature including the specification/definition/characterization of a Service Queue (SQ) for the particular flow.

5. Once the flow has been classified, the system will check if a corresponding Service Queue (SQ) already exists for this flow. In case a SQ exists, typically an AFQ, it will be assigned to that particular queue, and the SQ-Db and the Flow-Db are updated accordingly. If a SQ does not exist, then a new queue will be instantiated and the SQ-Db / Flow-Db are updated with the new entry. An SQ-Database would typically maintain a mapping between a flow-id and the queue-id.

6. With the new flow in the system and assigned to a queue, either existing or new, the system scheduler will be appropriately (re)configured to ensure fair scheduling of packets from the multiple queues within the system.

A functional architecture of an embodiment of a system or network employing our invention is illustrated in Fig. 2, and is assumed to have the following functional capabilities:

- 13 -

1. The system should be able to detect/identify application type of each flow.
2. Based on the application type, the system should be able to classify and prioritize the flows and assign relevant queues.
3. The system should have access/interface to relevant control entities of a network, e.g., PCRF (Policy Charging and Rules Function), HSS (Home Subscriber Service) of a 3GPP infrastructure, to be able to determine relevant user/network context information.
4. The system should be able to make QoE estimate based on the knowledge of application type, flow statistics, such as but not limited to throughput, delay, jitter, PER/BER (Packet Error Rate/Bite Error Rate) and other context information.

It is also assumed that the system knows the maximum bandwidth (β) of this egress link.

Embodiment of a generic scheduler parameterization method

The challenge presented by the proposed dynamic queuing system is the dynamic and optimal parameterization of the scheduler to multiple instances of PFQ and AFQ.

The challenge is to find the best QoE utility for all the application flows. As a result, we propose a scheme whereby the QoE utility is maximized for all the current applications, based on the following optimization problem:

$$\begin{aligned} & \max \sum_k U(r_k) \\ \text{s.t.} \quad & \sum_k r_k \leq \beta \end{aligned}$$

where r_k is the assigned bandwidth for traffic flow k and $U(r_k)$ is the associated utility function.

- 14 -

Based on the assigned application flow rates, the system computes for each queue the target bandwidth, by adding the computed flow rates for each queue. The target queue bandwidth is then used as input to parameterize the scheduler.

5

One embodiment of such a method is defined as following:

The bandwidth distribution method should be such that the system ensures maximum bandwidth utilization of the egress link such that:

10

$$\sum_{i=1}^n x_{SQ[i]} = \beta$$

Where $x_{SQ[i]}$ is the bandwidth x assigned to the i^{th} SQ in the queue pool consisting of a total of n active SQs, while β is the total egress bandwidth.

15

Fig. 3 gives a generic overview of a bandwidth evaluation/distribution method that can be employed with respect to our invention. At the start, the system iterates over all SQs and computes the Target Bandwidth (TBW) for all the flows. The TBW is the sum of the bandwidth (data rate) of all flows that the classifier would have recommended based on the SS (Subscriber Service) of the individual flow to meet the maximum quality requirements of a service. It can be represented as:

20

$$TBW = \sum_{i=1}^n \sum_{m=1}^n x_{SQ[i]}^{SS[m]}$$

25

The system will then compare the computed TBW with β . If it is less than or equal to β , then the SQ-Rates are assigned to the respective SQs and the underlying scheduler is configured accordingly, where the SQ-Rate ($\mathcal{R}_{SQ}$) is represented by

30

$$\mathcal{R}_{SQ[i]} = \sum_{m=1}^n x_{SS[m]}$$

- 15 -

In case the TBW exceeds β then the system will decrement the QoS/QoE profile of all the flows sequentially until it becomes equal to β , and the underlying scheduler is configured corresponding to the new $\mathcal{R}_{sq}$.

5

The bandwidth distribution method is executed when

1. A new SQ is created.
2. Number of flows assigned to AFQ exceeds a specified number.
3. During each evaluation epoch.

10

15 The system periodically monitors the SQs and evaluates flow statistics with reference to compliance to specific QoE/QoS requirements. At each evaluation epoch, the system can then (re)compute the $\mathcal{R}_{sq}$ for all SQs in order to ensure that the QoE/QoS objectives are being met, while at the same time enable maximum bandwidth utilization of the egress link.

20

For finding the TBW rates for bandwidth management and distribution, the system can potentially make use of a Service Profile Matrix (SPM) that will list the QoS/QoE profile of different service types in terms of bitrates. Figure 4 shows an example embodiment of a simplified SPM.

25

The SPM can be manipulated in a variety of ways while determining the TBW for a SS, depending on the service type priority and operator's policy.

30

Our proposed invention is independent of any particular type of scheduler and can be used with any weight based fair scheduler.

Embodiment 1: The invented dynamic queue management and scheduling system can be implemented in any packet switched/routed forwarding node, e.g. a router or switch, that targets to support application/QoE aware traffic differentiation and

- 16 -

bandwidth scheduling for the forwarded traffic. Examples of such nodes are routers, switches, wireless base station, e.g. WLAN APs or LTE eNBs, WiMAX base stations, or other middle boxes or packet gateways, e.g. NAT entities, Firewall nodes.

5

For such standalone entities, the dynamic queue management and scheduling system operates autonomously and is managed based on common network management principles – i.e. operators would provide the policies and parameters that control/drive via the network management system or configure them locally –
10 through a direct control/management interface.

Embodiment 2: The invented dynamic queue management and scheduling system can be implemented in a 3GPP based packet gateway, e.g. the GGSN, SGSN, PDN-GW, Serving-GW, ePDG, TWAG, (e)BNG or TDF (Traffic Detection
15 Function).

In case of a GGSN or PDN-GW based implementation, the operator could – besides the available network management interface or local configuration – also provide the policies for the dynamic queue management and bandwidth
20 scheduling via the Gx interface from the PCRF. The dynamic queue management system could be part of the PCEF function of the GGSN or P-GW or a separate function. The existing Application Detection and Control (ADC) function of the PCEF would be leveraged to detect the application type and the information would be used to assign application flows to the adequate queues. A close cooperation
25 between the ADC and the invented functions would be foreseen. The invented queuing and scheduling functions would be either integrated before the bearer binding function, i.e. to schedule the order by which packets are passed into the related – default/non-GBR (Guaranteed Bite Rate) – bearers, or integrated into and extend the existing, bearer centric queuing and scheduling functions of the
30 gateway entities.

In case of a TDF based implementation, the operator could – besides the available network management interface or local configuration – also provide the policies for the dynamic queue management and bandwidth scheduling via the Sd interface

- 17 -

from the PCRF. The existing traffic detection function of the TDF would be leveraged to detect the application type and the information would be used to assign application flows to the adequate queues. The invented queuing and scheduling functions would be integrated into and extend the TDF's existing packet forward/routing function and its queuing and scheduling functions.

In case of a Serving-GW, ePDG, TWAG, (e)BNG based implementation, the operator could – besides the available network management interface or local configuration – also provide the policies for the dynamic queue management and bandwidth scheduling via the Gxx interface, i.e. Gxc for Serving GW, Gxa for TWAG or (e)BNG, or Gxb for ePDG, from the PCRF. The dynamic queue management system could be part of the bearer binding function of the Serving GW or general packet forwarding function of the TWG, (e)BNG or ePDG. The invented queuing and scheduling functions would be integrated into and extend the entities' existing packet forward/routing function and its queuing and scheduling functions.

Important steps of embodiments of the invention are indicated in the following:

- 1) A method and a system or network that enables congestion management through traffic differentiation based on application type, subscriber class and according to network provider policies, wherein

at aggregate points in the network the system performs real-time application flow classification, dynamic management of per application-class or per-flow queues and dynamic parameterization of the scheduler based on a QoE utility optimization function with reference to the available resources. The invention encompasses:

- a. A novel traffic driven dynamic queue management and scheduling system based on application/QoE aware bandwidth allocation, controlled by policies. The queue management system dynamically creates a mix of Per Flow Queues (PFQ) and – application-specific – Aggregate Flow Queues (AFQ). PFQ/AFQs are dynamically

- 18 -

instantiated and created when a new flow arrives, and destroyed when the last flow session terminates.

5 i. The PFQ/AFQs are created and destroyed dynamically based on the currently detected traffic. Traffic flows are assigned to the PFQ/AFQs based on the application QoE and/or context information, e.g., subscription profile, device capabilities etc.

10 ii. Incoming flows are classified on-the-fly according to service queue templates and assigned to specific type of service queue depending on the policies, e.g. based on flow application type and priority.

15 b. The bandwidth allocated to the different queues dynamically adapts to the current traffic situation, maximizing the overall QoE of the application traffic flows for a given egress bandwidth.

20 c. The invention is aimed at managing user plane congestion at traffic aggregate points in the network by ensuring that the aggregate service queues' (SQ) bandwidth does not exceed the total bandwidth of the egress link, while ensuring maximum utilization of the link bandwidth. To this effect a QoE aware dynamic SQ bandwidth demand optimization method is used to dynamically parameterize the scheduler. Also a particular embodiment of this generic method
25 has been specified as part of this invention. This embodiment utilizes a Service Profile Matrix for bandwidth assignment to different SQs based on the flow priority.

30 Many modifications and other embodiments of the invention set forth herein will come to mind 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

- 19 -

specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

- 20 -

Claims

1. A method for operating a wireless network, wherein data flows aggregating at a traffic aggregation point will be managed for controlling congestion of data flows at said traffic aggregation point,

5 characterized by the following managing steps:

- classifying at least one data flow according to a rule used to differentiate traffic of said data flows for providing a service signature of said at least one data flow;

- 10 - assigning said at least one data flow to a Service Queue, SQ, in the form of a Per Flow Queue, PFQ, or an Aggregate Flow Queue, AFQ, based on said at least one data flow's service signature; and

- 15 - allocating of available bandwidth amongst the Service Queues based on a QoE estimation, such that the QoE of said data flow and/or the overall QoE estimation over all data flows traversing said aggregation point is improved.

2. A method according to claim 1, wherein the method will be realized in an access device or network element at said traffic aggregation point or as a middle box solution to be placed at said traffic aggregation point.

3. A method according to claim 1 or 2, wherein the rule comprises at least one service template, ST.

25 4. A method according to one of claims 1 to 3, wherein the rule uses at least one property of said at least one data flow to differentiate traffic.

5. A method according to claim 4, wherein said at least one property is IP-5-tuples and/or application type and/or subscriber information and/or header information and/or payload information and/or type of Service information and/or QoS information, particularly Differentiated Service Code Points, and/or QoS requirements and/or QoE requirements and/or security requirements and/or device capabilities.

30

- 21 -

6. A method according to one of claims 1 to 5, wherein the rule is associated with at least one service policy, SP, for defining a handling of said at least one classified data flow.

5 7. A method according to claim 6, wherein said at least one service policy defines the Service Queue type and/or a priority of handling and/or an associated bandwidth and/or a key performance indicator, KPI, of said at least one data flow.

10 8. A method according to one of claims 1 to 7, wherein a creation of at least one new Service Queue will be performed on the basis of a data flow's service signature.

15 9. A method according to one of claims 1 to 8, wherein a data flow being sensitive to QoS and/or QoE variations will trigger a creation of a new Per Flow Queue to which the data flow will be assigned to.

20 10. A method according to one of claims 1 to 9, wherein a data flow not being sensitive to QoS and/or QoE variations and/or being not suitable for a per-flow scheduling will be assigned to an Aggregate Flow Queue.

25 11. A method according to one of claims 1 to 10, wherein data flows belonging to a specific application type and/or sharing the same priority will be assigned to a dedicated Aggregate Flow Queue, wherein preferably – in case the Aggregated Flow Queue has not been created – the first data flow associated with said AFQ will trigger the dynamic creation of said AFQ.

30 12. A method according to one of claims 1 to 11, wherein data flows that can not be classified, e.g. encrypted data flows, will be assigned to a default Aggregate Flow Queue.

13. A method according to one of claims 1 to 12, wherein the classifying step comprises an analyzing step of at least one packet, preferably the first packet with application data payload, of the at least one data flow.

- 22 -

14. A method according to one of claims 1 to 13, wherein the analyzing step comprises a Deep Packet Inspection, DPI.

5 15. A method according to one of claims 1 to 14, wherein as a first step a flow-id of an arriving packet, e.g. arriving at the traffic aggregation point, will be derived, a check of a Flow-Database, flow-db, on the basis of the flow-id will be performed to determine if the packet belongs to an already known or classified data flow, and in case the data flow already exists, the packet or data flow will be directly assigned to its associated Service Queue without the classifying step, and in case
10 the data flow does not exist, a new flow-id will be derived and the classifying step will be performed.

16. A method according to claim 15, wherein the flow-id is a hash derived from a packet's header information and/or payload information.

15 17. A method according to claim 15 or 16, wherein the flow-db contains a mapping of the flow-ids with its corresponding service signature.

20 18. A method according to one of claims 1 to 17, wherein user and/or network context information from at least one control entity of the network will be used during said classifying and/or assigning and/or allocating step.

25 19. A method according to claim 18, wherein said user and/or network context information will be provided to said at least one control entity.

20. A method according to one of claims 1 to 19, wherein the QoE estimation will be based on an application type and/or flow statistics.

30 21. A method according to one of claims 1 to 20, wherein for each Service Queue a target bandwidth will be computed and used as input to parameterize a scheduler within the allocating step.

- 23 -

22. A method according to one of claims 1 to 21, wherein the allocating step targets to maximize the QoE for all data flows under the constraint of limited network resources.

5 23. A method according to claim 22, wherein the network resources are bandwidth and/or processing power.

24. A method according to one of claims 1 to 23, wherein the allocating step provides maximum bandwidth utilization.

10

25. A method according to one of claims 1 to 24, wherein a Service Profile Matrix, SPM, that lists the QoS and/or QoE requirements or QoS and/or QoE profiles of different service types in terms of target bitrates for different load or congestion levels will be used for determining the target bitrate for each Service Queue.

15

26. A method according to one of claims 21 to 25, wherein QoS and/or QoE profiles and/or the QoE estimation of all the flows or of individual flows with lower priority according to a Service Profile Matrix, SPM, are decremented, if the sum of all computed target bitrates exceeds the available/measured bandwidth.

20

27. A method according to one of claims 1 to 26, wherein the allocating step will be performed, when a new SQ is created or when a number of data flows assigned to an AFQ exceeds a specified number or during each evaluation epoch.

25

28. A wireless network, preferably for carrying out the method for operating a wireless network according to any one of claims 1 to 27, wherein data flows aggregating at a traffic aggregation point will be managed by a system for controlling congestion of data flows at said traffic aggregation point, wherein the system comprises:

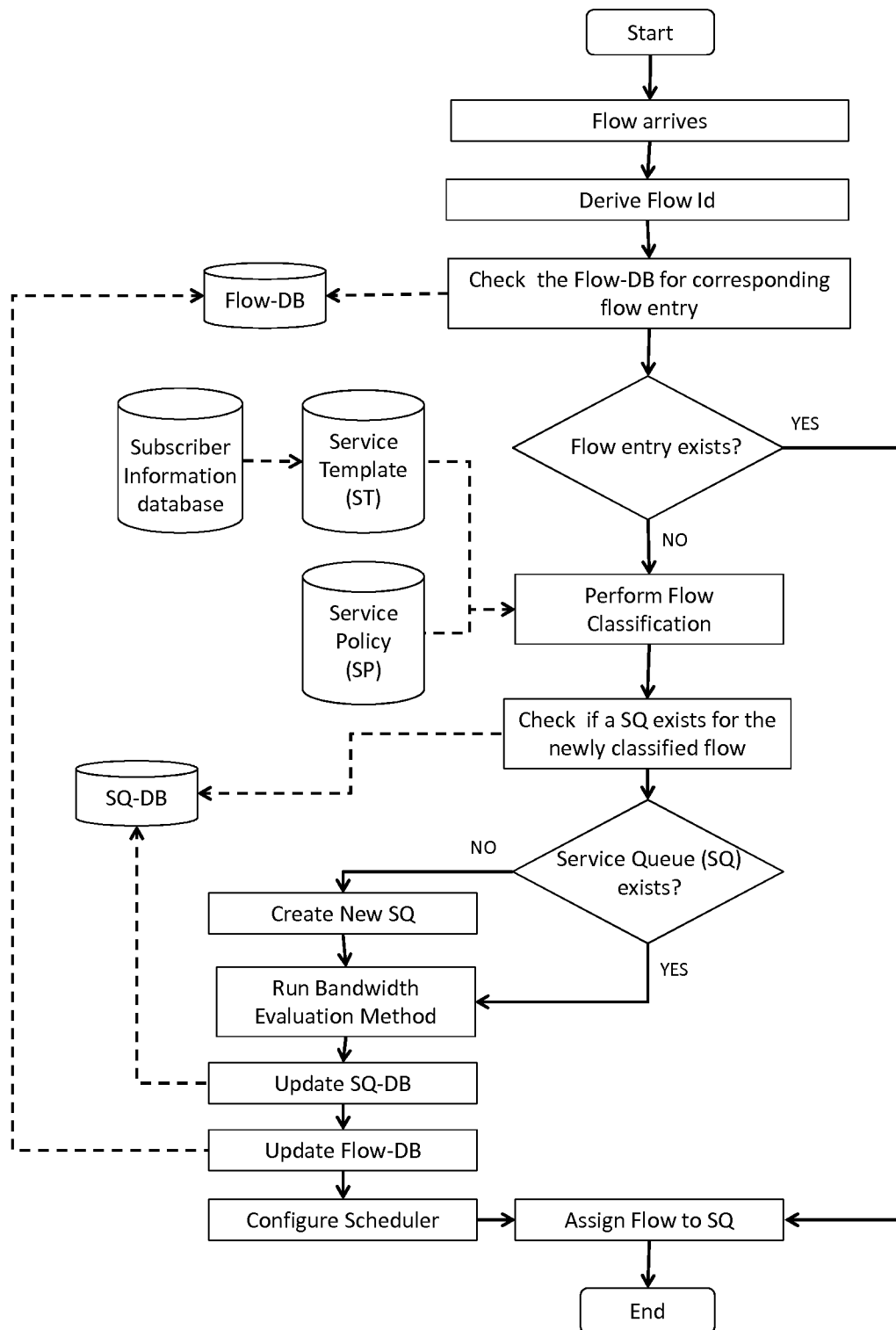
30

- means for classifying at least one data flow according to a rule used to differentiate traffic of said data flows for providing a service signature of said at least one data flow;

- 24 -

- 5
- means for assigning said at least one data flow to a Service Queue, SQ, in the form of a Per Flow Queue, PFQ, or an Aggregate Flow Queue, AFQ, based on said at least one data flow's service signature; and
 - means for allocating of available bandwidth amongst the Service Queues based on a QoE estimation, such that the QoE of said data flow and/or the overall QoE estimation over all data flows traversing said aggregation point is improved.

1/4

**Fig. 1**

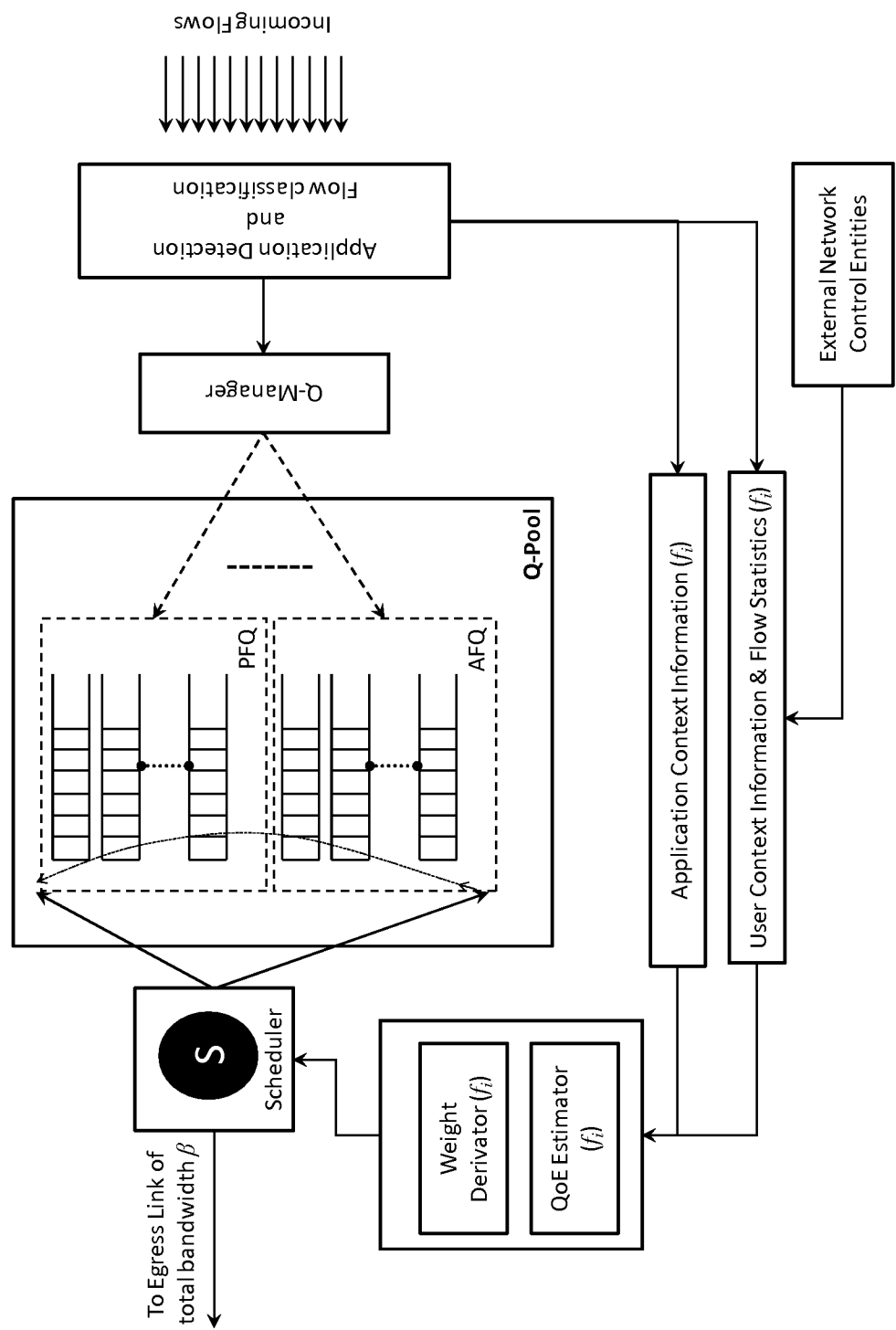
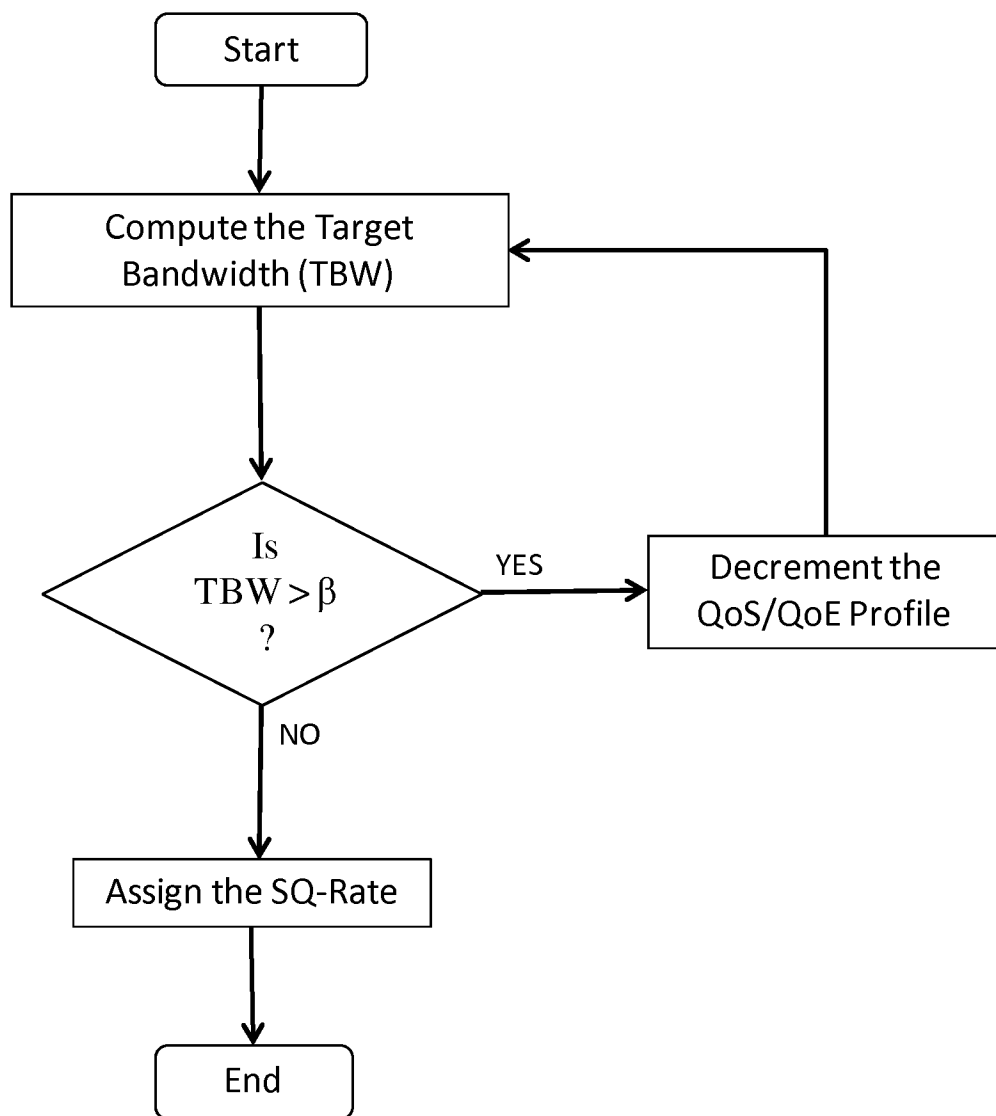


Fig. 2

3/4

**Fig. 3**

Rate index Service Type	1	2	3	4		n
SS_{DASH}	225	300	425	725		1500
$SS_{YouTube}$	225	325	450	750		2000
$SS_{rtVideo}$	275	375	500	800		3000
SS_{FTP1}	180	180	300	500		1500
SS_{FTP2}	150	150	180	325		1000
SS_{web1}	100	125	175	250		1000
SS_{web2}	100	90	150	225		800
SS_{web3}	90	90	100	180		500

Fig. 4

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2013/063723

A. CLASSIFICATION OF SUBJECT MATTER
 INV. H04L12/70
 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)
 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)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2012/327779 A1 (GELL DAVID [US] ET AL) 27 December 2012 (2012-12-27) abstract paragraph [0002] - paragraph [0008] paragraph [0035] paragraph [0042] paragraph [0053] - paragraph [0054] paragraph [0064] - paragraph [0083] paragraph [0134] - paragraph [0138] paragraph [0092] - paragraph [0104] paragraph [0139] - paragraph [0140] paragraph [0146] paragraph [0158] paragraph [0161] paragraph [0183] paragraph [0279] - paragraph [0296] claims 1,24,30,43 figures 5-7,9-11,16-17,21-23 paragraph [0246] - paragraph [0247] -/--	1-28



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 application or patent 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

4 December 2013

Date of mailing of the international search report

11/12/2013

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

Schrembs, Gerd

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2013/063723

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	paragraph [0214] - paragraph [0217] ----- US 2009/034426 A1 (LUFT SIEGFRIED J [CA] ET AL) 5 February 2009 (2009-02-05) abstract paragraph [0001] paragraph [0006] paragraph [0021] - paragraph [0023] paragraph [0025] paragraph [0035] - paragraph [0054] claims 1-6 figures 3,7,9 -----	1-28
A	WO 2013/041587 A1 (NEC EUROPE LTD [DE]; TALEB TARIK [DE]) 28 March 2013 (2013-03-28) abstract page 1, line 4 - page 8, line 25 page 10, line 1 - page 12, line 5 page 16, line 9 - page 17, line 21 claims 1,16 figures 1,3 -----	1-28
A	WO 2009/049676 A1 (ERICSSON HUNGARY LTD [HU]; KOVACS AKOS [HU]; GODOR ISTVAN [HU]; RACZ S) 23 April 2009 (2009-04-23) abstract page 1, line 5 - page 7, line 4 page 9, line 16 - page 10, line 24 page 13, line 20 - page 15, line 4 page 16, line 26 - page 17, line 2 claims 1-4,18-21 figures 1,3,5-7 -----	1-28

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2013/063723

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2012327779	A1	27-12-2012	NONE

US 2009034426	A1	05-02-2009	CA 2691939 A1 05-02-2009
			EP 2171920 A1 07-04-2010
			US 2009034426 A1 05-02-2009
			WO 2009015461 A1 05-02-2009

WO 2013041587	A1	28-03-2013	NONE

WO 2009049676	A1	23-04-2009	NONE
