S2-170138

5G QoS parameters

Ericsson

© 3GPP 2016

Introduction

- In SA2 it has been agreed that there will be two flow types, type A and type B. Type B will be used for most kinds of specialized flows:
  - Current GBR flows
  - MTC and Critical MTC
  - Public service with specialized requirements
  - Over-the-top flows with known service requirements
- Type B flows will be assigned flow ID's dynamically, so it is not possible to specify the flow treatment through a set of AN configured parameters per flow type.
  - All information that the RAN needs in order to optimize service performance and prioritize flows, must be provided through the QoS parameters.
- In current QoS framework, observability of QoS target fulfillment is difficult.
  - Ericsson sees a need for well defined QoS target definitions that map directly to service performance.

© 3GPP 2016

Proposal - summary

It is proposed that the QoS parameters shall be used to describe:

- Service QoS targets
  - KPI's for end-point service performance, useful for network optimization.
    - Bitrate requirements
    - Packet delay and loss requirement
    - Averaging window size
- Priority
  - Priority to fulfill QoS targets
  - Priority to get more resources than needed for QoS targets.
- Service characteristics useful for network optimization
  - Periodicity
- Flow treatment with regards to Admission, Pre-emption and Notification.

© 3GPP 2016

Service QoS Targets

© 3GPP 2016

Service QoS Targets

- A set of QoS target that can describe end-point service performance should be available within the network
  - The QoS Targets can be used for
    - Observability of service performance and app coverage.
    - Network optimization of user satisfaction

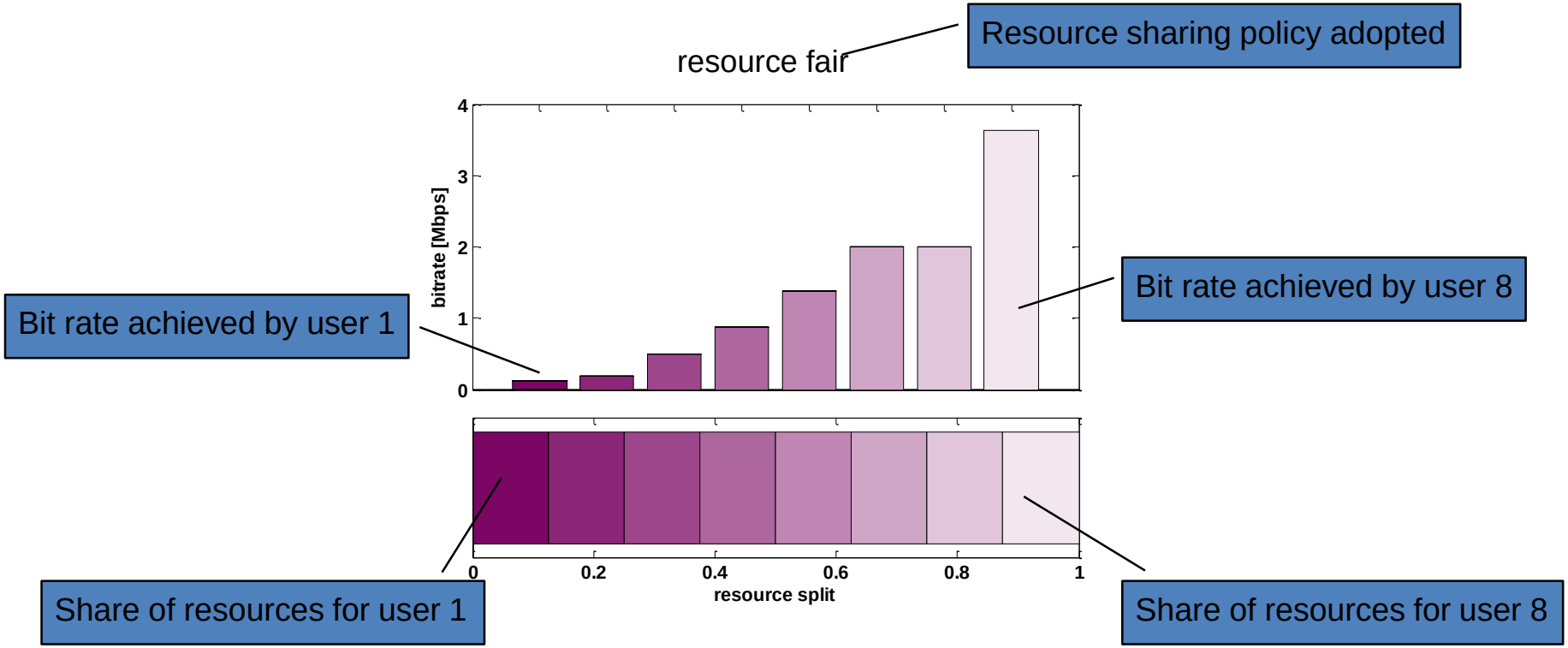
Service QoS Targets:

- Guaranteed or Preferred minimum Flow Bitrate
  - Ericsson proposes a minimum Bitrate to be defined for all type B flows. For non-GBR flows we propose a Preferred minimum Flow Bitrate that is based on service request.
  - All services have minimum bitrate requirements. If this minimum requirements are known to the AN, it is possible to improve app coverage.
- Packet Delay and Loss requirement
  - Improved definitions needed in order to support ultra-reliable services
  - Current standard requires that 98% is within delay requirement. Lower late-loss requirements should be available.
- Averaging window
  - A window size should be defined for measuring fulfilment of all targets.

© 3GPP 2016

A simple model to illustrate the effect of RRM algorithms

- Simulating a situation with user radio conditions for 8 users with different resource cost-per-bit
  - All users are assumed to have much data and not being limited by MBR

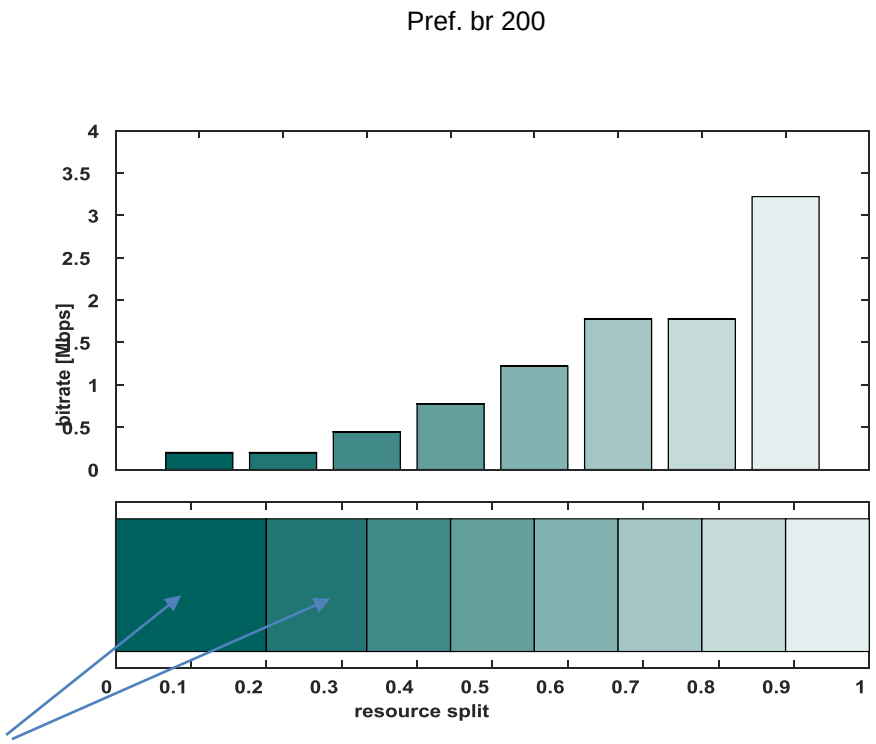


© 3GPP 2016

Preferred bitrate known in scheduler

- 8 users with different radio conditions
  - All users are assumed to have much data and large MBR
  - 200 kbps needed for App coverage
- Scheduler enforces a preferred bitrate = 200 kbps

  Note: in the example, once the preferred bit rate has been satisfied for all users, a resource fair split is used
- Most flows get same amount of resources.
- Extra resources scheduled to flows with bad radio.
- App coverage is 100%



Introducing a “Preferred Minimum flow Bit Rate” makes it possible to optimize App coverage

© 3GPP 2016

Averaging Window

- Ericsson proposes a standardized definition of the averaging window used when observing the fulfilment of QoS targets. (I.e Flow bitrate, delay and packet loss requirement.)
  - Without a definition the interpretation of the parameters is unclear.
  - Different definitions selected by different vendors could result in large variations of the measured bit rates, delay-, and loss-rates among the vendor implementations.

- Ericsson propose that the length of the averaging window is defined through a new QoS parameter.
  - It is possible to optimize network and transmit data ‘just-in-time’ if the time requirements are known.
  - Better observability of QoE if the QoS targets match the need of the application.
    - For streaming video, the bitrate target need to be fulfilled in the timeframe of the application buffer.
    - For signaling flows, the bitrate must be fulfilled on a much shorter timescale

© 3GPP 2016

Priorities

© 3GPP 2016

Priorities

- A set of priorities that can describe the importance of a flow should be available within the network for RRM features like admission, retention and packet forwarding

- The priority for admission and retention of a flow in the network
- The priority of QoS targets fulfillment
- The priority of getting more resources after fulfillment of the QoS targets

© 3GPP 2016

The priority of getting more resources after fulfillment of the QoS targets

- A large part of the flows in the network are dynamic, where the user satisfaction increases with the amount of data transmitted.
- Current framework does not define how to prioritize a flow with a bitrate above GBR.
- In current EPS framework, a non-GBR flow with high QCI priority should be scheduled within the delay requirement, regardless the amount of data transmitted, or level of network congestion. The end-point congestion control can then not detect congestion, and will not back down in congestion.
  - Only way to prevent this is to use MBR, but that will block data even when there is no congestion.

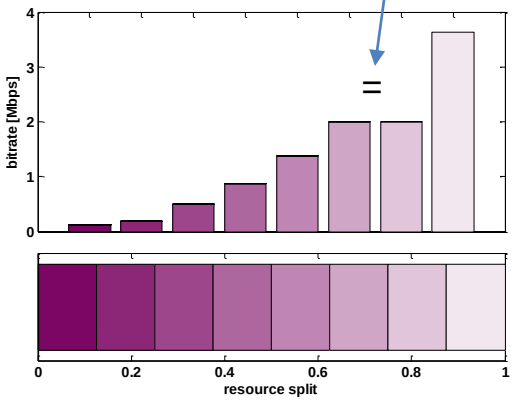
We propose that it should be possible to define the relative amount of resources that should be spent on each flow, after QoS targets are fulfilled.

© 3GPP 2016

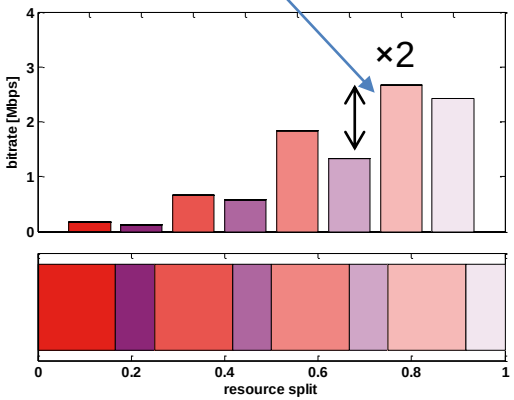
Resource split

- 8 users with different radio conditions
- Resource fair: split resources evenly, best radio quality gets highest rate
- The split can be controlled by applying relative priority
  - Red flows: Relative Priority=1, Purple flows: Relative Priority=2
  - 2 x resources to red flows, compared to purple flows.
  - With same radio quality, red flows have double bitrate.

resource fair



relative priority



A relative priority parameter allows for a differentiated QoE through “a relative resource” sharing above QoS targets

© 3GPP 2016

Priority for admission/retention of a flow and QoS target fulfillment

- Each flow has a set of minimum QoS targets
  - Some flows have more stringent targets on bit rate, delay and reliability, and other flows have more relaxed or no targets
- Each flow also has an importance to get access to the network and to get the minimum QoS targets fulfilled
- Mechanisms that may be used for fulfilling QoS targets and considering the importance/priority of a flow:
  - Admission control
  - Packet forwarding (scheduling)
  - Pre-emption / Retention
  - AQM / ECN
- These mechanisms need to co-work to ensure that QoS targets are fulfilled in the order of importance/priority of flows

It is proposed to use one and the same Priority for all mechanisms

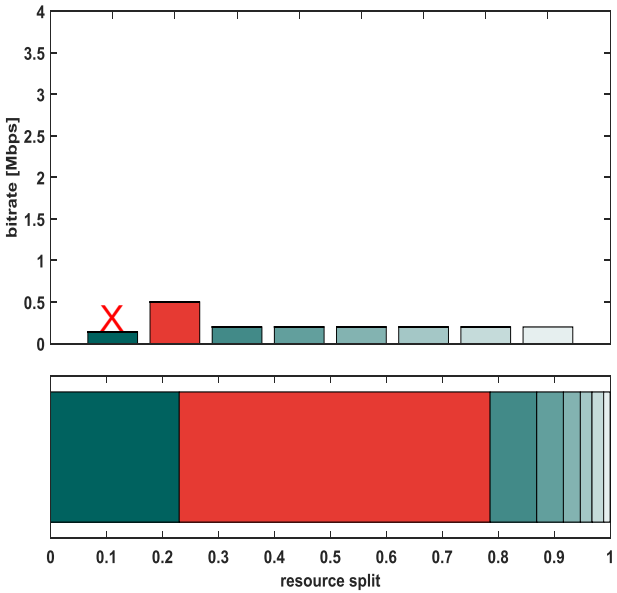
The importance of admitting a flow should be the same as the importance of fulfilling the QoS targets in packet forwarding, as well as for pre-empting/retaining a flow

© 3GPP 2016

Using ARP priority in the scheduler

- 7 flows have GBR=200 kbps
- A red flow belong to public safety UE is admitted.
  - GBR=500 kbps, High ARP priority
- There is congestion
  - not sufficient resources to fulfill all QoS targets.
- GBR and ARP priority is used in scheduling.
  - Public service flow first gets resources up to GBR. →Flow is OK.
  - Then other flows get resources up to GBR. →Flow with worst radio fails.

GBR + ARP



If same priority is used for Admission and Scheduling to fulfill QoS targets, targets fulfillment can be guaranteed in order of importance.

© 3GPP 2016

UL/DL Periodicity

- If the periodicity is known,
  - effective pre-scheduling can be used in the RAN, minimizing delays and signaling overhead.
  - Connected state DRX periods can be set to limit battery consumption without affecting transmission delays.

© 3GPP 2016

Proposed
QOS parameters

© 3GPP 2016

QoS Parameters

Agreements as of SA2 #118:

13.1. QoS parameters may include the following:

a) UL and DL Maximum Flow Bit Rate.

b) UL and DL Guaranteed Flow Bit Rate.

c) Priority level.

d) Packet Delay Budget.

e) Packet Error rate.

f) ARP.

g) Notification Control.

NOTE 12: Parameters c, d), e) apply for both bullets #11 and #12. Parameters a), b), f) and g) apply only to bullets #12.

NOTE 13: Need for other parameters such as packet jitter, size, periodicity and relative priority etc. is FFS and will be determined during normative phase.

NOTE 14: Parameters a) through f) are based on definitions in TS 23.203 [3]. Parameter a), b) and g) are only applicable to GBR QoS flows. Parameter g) controls whether notification should be made if the QoS targets are no longer fulfilled for a QoS flows. Mechanisms related to the notification and how to minimize the notifications to the CN will be determined as part of the normative phase. Whether f) applies to non-GBR QoS flows will be determined as part of the normative work.

NOTE 15: It is to be determined during normative phase whether c) Priority level and ARP Priority in f) may be indicated through a single parameter.

NOTE 16: The network can decide which parameters need to be signalled to the UE and when.

© 3GPP 2016

Proposed Parameter definitions-QoS Targets

Changes from interim agreements and/or 23.203 marked as italic

QoS Targets

- UL and DL Guaranteed flow bit rate
  - The flow bitrate which is required.
  - *GBR is measured over an Averaging window of defined size.*
- *UL and DL Preferred minimum flow bit rate*
  - *May be defined for non-GBR flows.*
  - *Value may be used by AN to optimize user satisfaction.*
  - *PBR is measured over an Averaging window of defined size.*
- Packet delay budget
  - Defines the upper bound for the time a packet may be delayed between the UE and the PCEF (CN_UP?)
  - Confidence level of 98%
  - *Tougher confidence level is needed for some flows. How to signal that is for further study.*
  - *The confidence level should be measured over an averaging window of defined size*
- Packet error rate
  - Defines the amount of packet losses due to radio loss. (Harq failure etc.)
  - *For end-point packet loss, should also AQM drops be included?*
  - *For some real-time flows, late-loss should be included.*
  - *Should be measured over an averaging window of defined size*

© 3GPP 2016

Proposed Parameter definitions – Priorities, Characteristics

Changes from interim agreements and/or 23.203 marked as italic

Priorities

- Priority level.
  - *Used to indicate the relative priority above the minimum QoS targets, and distributing excess resources between flows when the minimum QoS targets* are fulfilled for all flows (having minimum targets) Applicable to GBR and non-GBR flows.*
    - *Flows not having minimum QoS targets will be given resources considering the relative priority*
- *ARP priority*
  - Used to decide whether AC should admit a flow, and during resource limitation, which existing flows to pre-empt, *and which existing flows QoS targets to prioritize*.

Service characteristics useful for network optimization

- *UL/DL Periodicity*
  - *The parameters should be set if the transmission will be periodic.*
- *It is for further studies if more service characteristics parameters are needed.*

*Note: Minimum QoS targets are defined by parameters b), d), e) and more, if included from Note 13.

© 3GPP 2016

Proposed Parameter definitions – Treatment

Changes from interim agreements and/or 23.203 marked as italic

Flow treatment with regards to Admission, Pre-emption and Notification.

- *ARP PVI, PCI*
  - Keep parameters for the sake of backward compatibility.
  - We see no need for it, since there are no default bearers that need to be preserved.
- *GBR/non-GBR*
  - *If the parameter GBR is included for a flow, the flow is a GBR flow*, and AC should be used.
  - GBR flows should only be admitted if it is expected that GBR *and other QoS targets* can be fulfilled.
  - GBR flows should be pre-empted if the QoS targets can no longer be fulfilled due to congestion and/or coverage, *and if notification is not set*.
  - For non-GBR flows, QoS targets should not be taken into account for Admission control.
- *Notification*
  - *Defines whether a flow should be notified if QoS targets can not be fulfilled*
  - *Notification in advance, is indicated with a time period*
  - *Define frequency of notifications is FFS*

*Note: Minimum QoS targets are defined by parameters b), d), e) and if more are included from Note 13.

© 3GPP 2016