Bandwidth, Throughput, Latency & Jitter in mobile networks

@3g4gUK

Bandwidth in mobile networks

The simplest way to understand bandwidth is to think of them as pipes. The fatter the pipe, the more the bandwidth

Channel Bandwidth (BW)

Release 15 3GPP TS 36.104 V15.1.0 (2017-12)

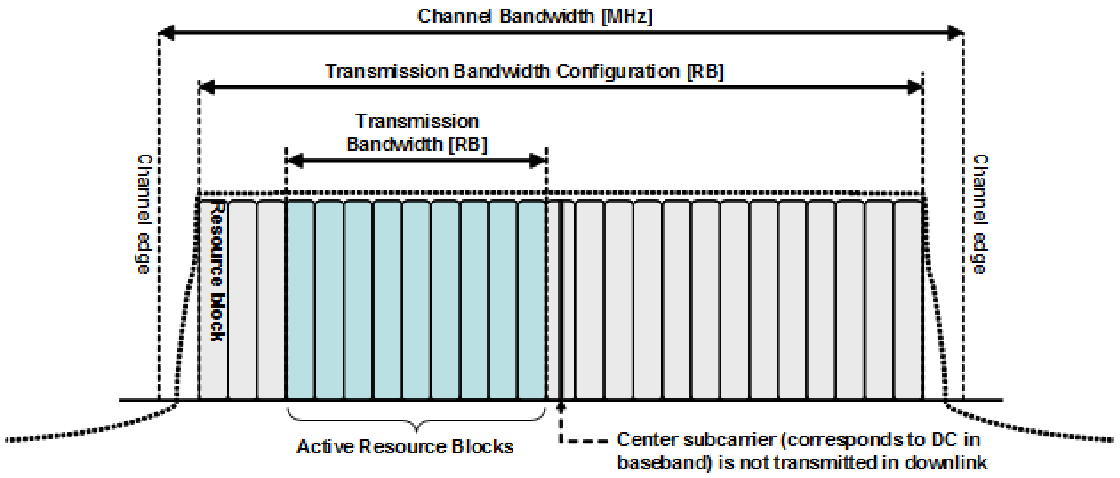


Figure 5.6-1 Definition of Channel Bandwidth and Transmission Bandwidth Configuration for one E-UTRA carrier

Channel Bandwidth (often referred to as just bandwidth) is the RF bandwidth supporting a single RF carrier with the transmission bandwidth configured in the uplink or downlink of a cell. The channel bandwidth is measured in MHz (Hertz = cycle / second) and is used as a reference for transmitter and receiver RF requirements

For example LTE supports scalable channel bandwidths, from 1.4 MHz to 20 MHz. The BW of UMTS channel is 5MHz

Channel Bandwidth (BW)

Release 15 3GPP TS 36.104 V15.1.0 (2017-12)

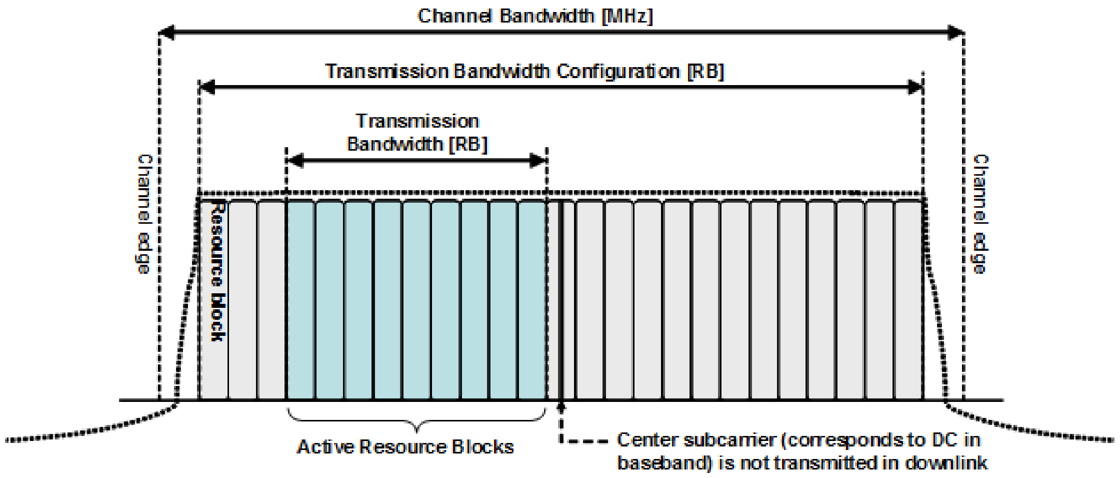
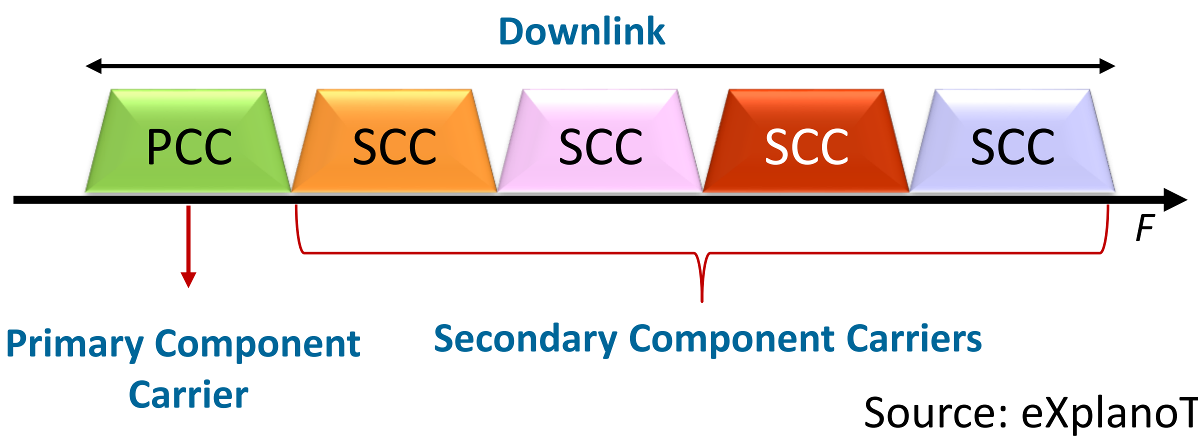


Figure 5.6-1 Definition of Channel Bandwidth and Transmission Bandwidth Configuration for one E-UTRA carrier

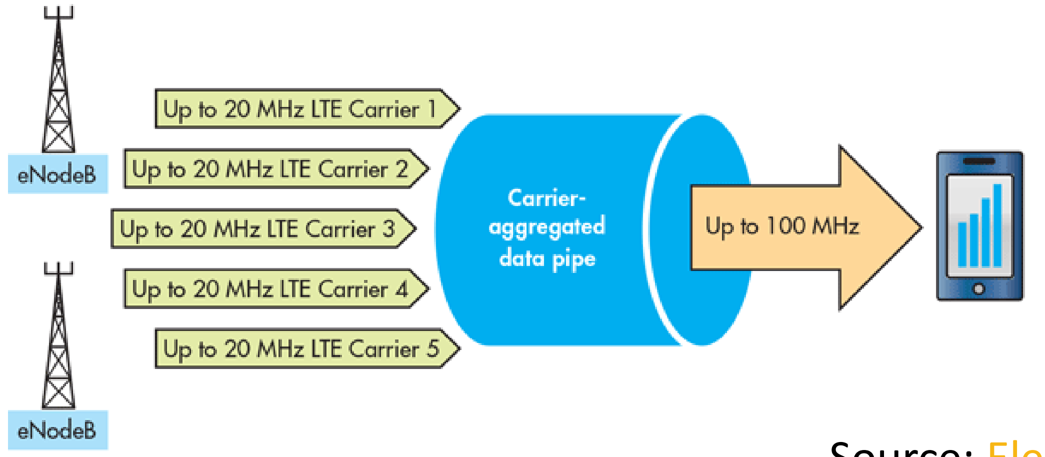
Transmission bandwidth: Bandwidth of an instantaneous transmission from a UE or BS, measured in Resource Block units.

Transmission bandwidth configuration: highest transmission bandwidth allowed for uplink or downlink in a given channel bandwidth, measured in resource block units.

Carrier aggregation



Source: eXplanoTech



Source: ElectronicDesign

Carrier aggregation Bandwidth

Release 15 3GPP TS 36.104 V15.1.0 (2017-12)

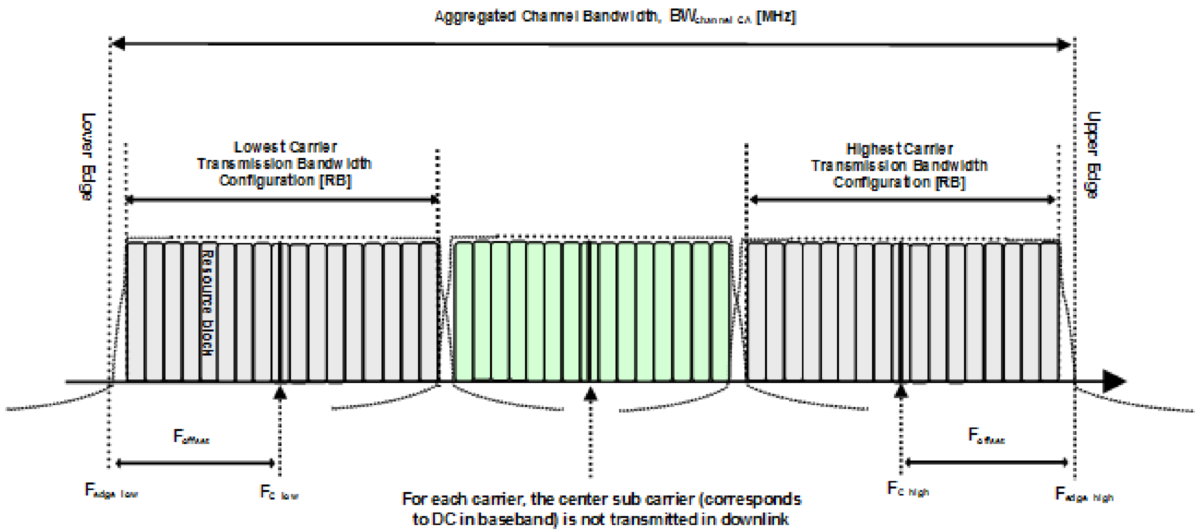


Figure 5.6-2 Definition of Aggregated Channel Bandwidth for intra-band carrier aggregation

Aggregated Channel Bandwidth: RF bandwidth in which a base station transmits and/or receives multiple contiguously aggregated carriers. The Aggregated Channel Bandwidth is measured in MHz.

Aggregated Transmission Bandwidth Configuration (ATBC): total number of aggregated physical resource blocks (PRB).

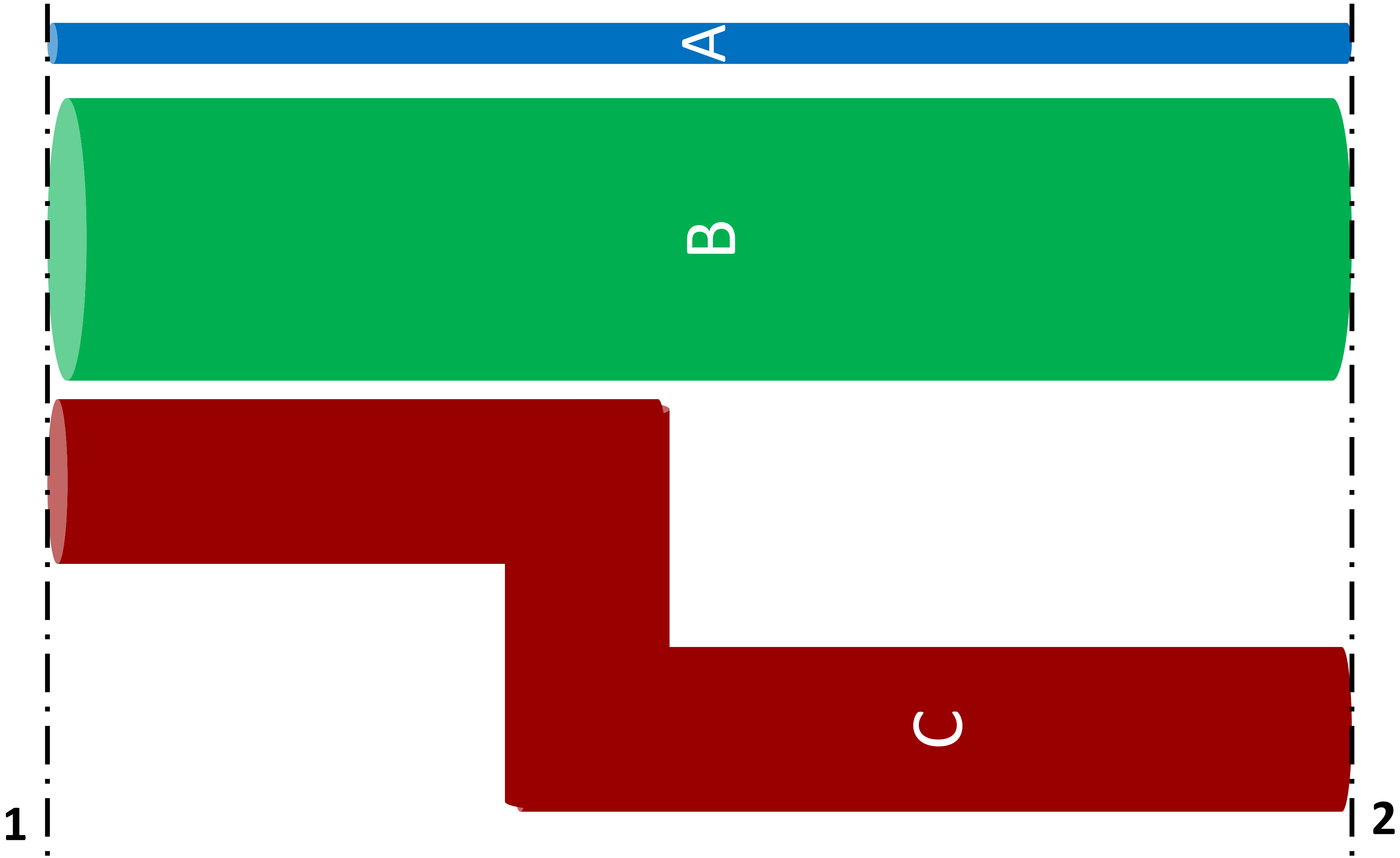
Latency & Jitter

Latency is generally defined as the time it takes for a source to send a packet of data to a receiver. In simple terms, half of Ping time. This is also referred to as **one way latency**.

Sometimes the term **Round trip latency** or round trip time (RTT) is also used to define latency. This is the same as ping time.

Jitter is defined as the variation in the delay (or latency) of received packets. It is also referred to as ‘delay jitter’.

Explaining Latency vs Bandwidth



Bandwidth is often referred to as a measure of capacity.

While **Latency** is a measure of delay.

Speed Test example

Latency in 3GPP & ITU

In 3GPP and ITU, control-plane latency and user-plane latency is discussed for a particular technology

Control-plane latency is defined as the transition time from idle state to connected state.

The **user-plane latency**, also known as transport delay, is defined as the one-way transit time between a packet being available at the IP layer of the origin and the availability of this packet at IP layer of the destination.

Control Plane Latency of LTE-A system

Control plane latency

• LTE fulfills ITU-R requirements on control plane latency for idle to connected transition

ITU-R Requirement: less than 100

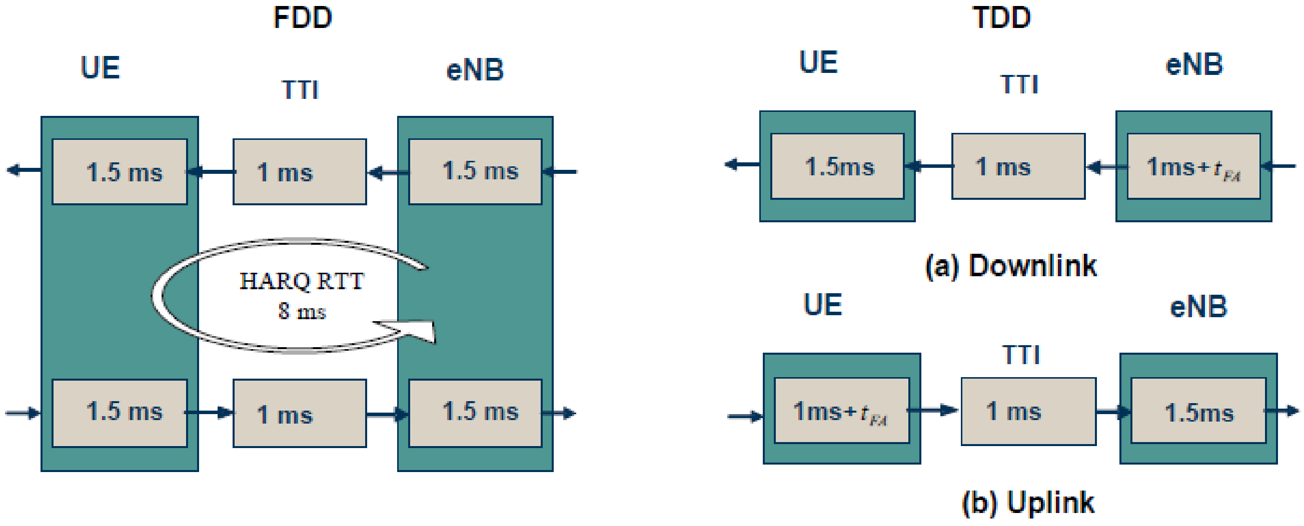
Component	Description	Time (ms)
1	Average delay due to RACH scheduling period (1ms RACH cycle)	0.5
2	RACH Preamble	1
3-4	Preamble detection and transmission of RA response (Time between the end RACH transmission and UE's reception of scheduling grant and timing adjustment)	3
5	UE Processing Delay (decoding of scheduling grant, timing alignment and C-RNTI assignment + L1 encoding of RRC Connection Request)	5
6	Transmission of RRC and NAS Request	1
7	Processing delay in eNB (L2 and RRC)	4
8	Transmission of RRC Connection Set-up (and UL grant)	1
9	Processing delay in the UE (L2 and RRC)	12
10	Transmission of RRC Connection Set-up complete	1
11	*Processing delay in eNB (Uu → S1-C)*	
12	*S1-C Transfer delay*	
13	*MME Processing Delay (including UE context retrieval of 10ms)*	
14	*S1-C Transfer delay*	
15	Processing delay in eNB (S1-C → Uu)	4
16	Transmission of RRC Security Mode Command and Connection Reconfiguration (+TTI alignment)	1.5
17	Processing delay in UE (L2 and RRC)	16
	Total delay	50

Source: Presentation by Takehiro Nakamura, 3GPP TSG-RAN Chairman, ITU-R WP 5D 3rd Workshop on IMT-Advanced, 15 October 2009

User Plane Latency of LTE-A system

User plane latency

- LTE fulfills ITU-R requirements on user plane latency



0 % BLER	4.0 msec
10 % BLER	4.8 msec

0 % BLER	4.9 msec
10 % BLER	6.035 msec

Source: Presentation by Takehiro Nakamura, 3GPP TSG-RAN Chairman, ITU-R WP 5D 3rd Workshop on IMT-Advanced, 15 October 2009

End-to-end (E2E) Latency

End-to-end (E2E) latency: the time that takes to transfer a given piece of information from a source to a destination, measured at the communication interface, from the moment it is transmitted by the source to the moment it is successfully received at the destination.

5G Latency Requirements – Industry Targets

NGMN 5G Requirements

- 5G E2E Latency (eMBB) = **10ms** (i.e. RTT from UE-Application-UE)
- 5G E2E Latency (URLLC) = **1ms** (i.e. RTT from UE-Application-UE – or just UE-UE)

In both cases, the values are defined as capabilities that should be supported by the 5G System.

GSMA 5G Requirements

- 5G E2E Latency = **1ms** (again, defined as a capability target, not as a universal requirement)

ITU-R IMT-2020 Requirements

- eMBB User Plane Latency (one-way) = **4ms** [radio network contribution]
- URLLC User Plane Latency (one-way) = **1ms** [radio network contribution]
- Control Plane Latency = **20ms (10ms target)** [UE transition from Idle to Active via network]

Low Latency Use Case Requirements (various sources)

- Virtual Reality & Augmented Reality: **7-12ms**
- Tactile Internet (e.g. Remote Surgery, Remote Diagnosis, Remote Sales): < **10ms**
- Vehicle-to-Vehicle (Co-operative Driving, Platooning, Collision Avoidance): < **10ms**
- Manufacturing & Robotic Control / Safety Systems: **1-10ms**

Source: Andy Sutton

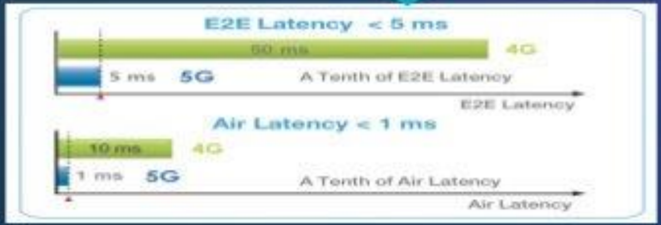
BT

Latency & Reliability Definitions (#URLLC)

Latency and Reliability (definitions)

- End-to-end (E2E) latency: scheduling delay+ queuing delay+ transmission delay+ receiver-side processing and decoding delay+ multiple HARQ RTT
- User plane latency (3GPP) [1]: **one-way** time it takes to successfully deliver a packet…
- Control plane latency (3GPP) [1]: transition time from a most “battery efficient” state (e.g., Idle state) to the start of continuous data transfer (e.g. active state).

- Reliability per node: transmission error probability, queuing delay, violation probability and proactive + dropping probability
- Reliability (3GPP): successfully transmit 32byte message over the 5G **radio** Interface within **1ms** with a success probability of **1-10^-5**
- Availability: probability that a given service is available (i.e., coverage). Higher availability entails lower reliability



Reliability ⟷ NO packet drop
NO delayed packet
NO erroneously decoded packet

- ITU and 3GPP require 5G to successfully transmit 32byte message over the 5G **radio** Interface within **1ms** with a **1-10^-5** success probability ------------------- maximum **BLER of 10^-5**
- 3GPP further requires 5G to be able to achieve an average latency over the 5G radio interface of 0.5ms
- While URLLC are E2E requirements, 3GPP and ITU consider **only one way** latency over 5G RAN

[1] 3GPP, “Service requirements for the 5g system” in 3rd Generation Partnership Project (3GPP), TS 22.261 v16.0.0, 06 2017, 2017.

Source: Mehdi Bennis

Throughput

Throughput is the actual rate that information is transferred. It is defined as the quantity of data being sent/received by unit of time.

In mobile networks, the end user throughput is the amount of information received in bits /second.

Throughput is measured at Layer 1/2 or even at application layer.

In a network, often cell throughput is calculated which is throughput of all simultaneous users in the cell.

Network Throughput & Network Capacity

$$\underbrace{\text{Network Throughput}}_{\text{bit/s/km}^2} = \sum_{\text{band}} \underbrace{\text{Quantity of Spectrum}}_{\text{Hz}} \times \underbrace{\text{Cell Density}}_{\text{Cell/km}^2} \times \underbrace{\text{Spectral Efficiency}}_{\text{bit/s/Hz/Cell}}$$

Source: Emil Björnson

$$\underbrace{\text{Network Capacity}}_{\text{bit/s}} = \sum_{\text{band}} \underbrace{\text{Quantity of Spectrum}}_{\text{Hz}} \times \underbrace{\text{Cell Spectral Efficiency}}_{\text{bit/s/Hz}} \times \text{Number of Cells}$$

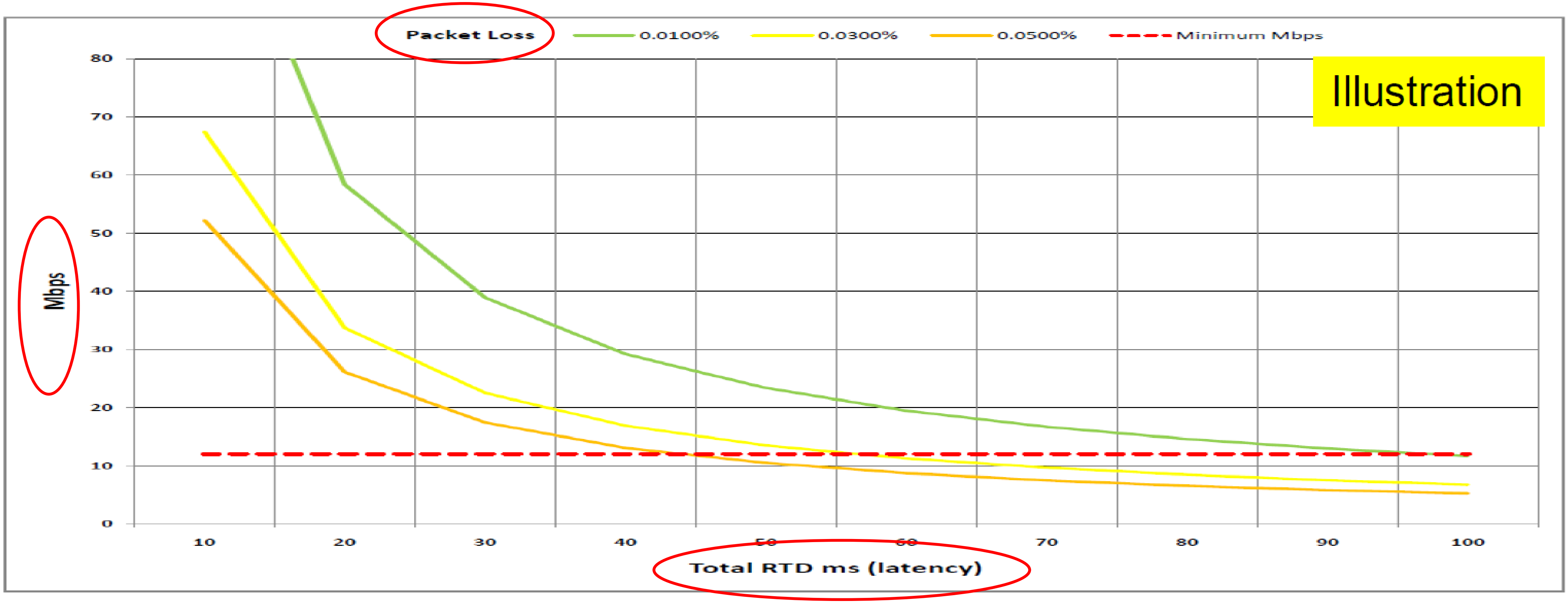
Source: Simon Chapman, Keima

Packet Loss and Packet Error Rate

Packet loss reflects the number of packets lost per 100 of packets sent by a host.

Packet Error Rate (PER) defines an upper bound for a rate of non-congestion related packet losses. The purpose of the PER is to allow for appropriate link layer protocol configurations (e.g. RLC and HARQ in RAN of a 3GPP access).

How Does Latency, Packet Loss Impact LTE?



Source: Cisco

Performance requirements for low-latency and high-reliability scenarios

Table 7.2.2-1 Performance requirements for low-latency and high-reliability scenarios.

Scenario	End-to-end latency (note 3)	Jitter	Survival time	Communication service availability (note 4)	Reliability (note 4)	User experienced data rate	Payload size (note 5)	Traffic density (note 6)	Connection density (note 7)	Service area dimension (note 8)
Discrete automation – motion control (note 1)	1 ms	1 µs	0 ms	99,9999%	99,9999%	1 Mbps up to 10 Mbps	Small	1 Tbps/km^2	100 000/km^2	100 x 100 x 30 m
Discrete automation	10 ms	100 µs	0 ms	99,99%	99,99%	10 Mbps	Small to big	1 Tbps/km^2	100 000/km^2	1000 x 1000 x 30 m
Process automation – remote control	50 ms	20 ms	100 ms	99,9999%	99,9999%	1 Mbps up to 100 Mbps	Small to big	100 Gbps/km^2	1 000/km^2	300 x 300 x 50 m
Process automation – monitoring	50 ms	20 ms	100 ms	99,9%	99,9%	1 Mbps	Small	10 Gbps/km^2	10 000/km^2	300 x 300 x 50
Electricity distribution – medium voltage	25 ms	25 ms	25 ms	99,9%	99,9%	10 Mbps	Small to big	10 Gbps/km^2	1 000/km^2	100 km along power line
Electricity distribution – high voltage (note 2)	5 ms	1 ms	10 ms	99,9999%	99,9999%	10 Mbps	Small	100 Gbps/km^2	1 000/km^2 (note 9)	200 km along power line
Intelligent transport systems – infrastructure backhaul	10 ms	20 ms	100 ms	99,9999%	99,9999%	10 Mbps	Small to big	10 Gbps/km^2	1 000/km^2	2 km along a road
Tactile interaction (note 1)	0,5 ms	TBC	TBC	[99,999%]	[99,999%]	[Low]	[Small]	[Low]	[Low]	TBC
Remote control	[5 ms]	TBC	TBC	[99,999%]	[99,999%]	[From low to 10 Mbps]	[Small to big]	[Low]	[Low]	TBC

NOTE 1: Traffic prioritization and hosting services close to the end-user may be helpful in reaching the lowest latency values.
NOTE 2: Currently realised via wired communication lines.
NOTE 3: This is the end-to-end latency the service requires. The end-to-end latency is not completely allocated to the 5G system in case other networks are in the communication path.
NOTE 4: Communication service availability relates to the service interfaces, reliability relates to a given node. Reliability should be equal or higher than communication service availability.
NOTE 5: Small: payload typically ≤ 256 bytes
NOTE 6: Based on the assumption that all connected applications within the service volume require the user experienced data rate.
NOTE 7: Under the assumption of 100% 5G penetration.
NOTE 8 Estimates of maximum dimensions; the last figure is the vertical dimension.
NOTE 9: In dense urban areas.
NOTE 10: All the values in this table are targeted values and not strict requirements.

3GPP TS 22.261

Thank You

To learn more, visit:

3G4G Website – http://www.3g4g.co.uk/

3G4G Blog – http://blog.3g4g.co.uk/

3G4G Small Cells Blog – http://smallcells.3g4g.co.uk/

Operator Watch - http://operatorwatch.3g4g.co.uk/

Follow us on Twitter: https://twitter.com/3g4gUK

Follow us on Facebook: https://www.facebook.com/3g4gUK/

Follow us on Linkedin: https://www.linkedin.com/company/3g4g

Follow us on Slideshare: https://www.slideshare.net/3G4GLtd

Follow us on Youtube: https://www.youtube.com/3G4G5G

Follow us on Storify: https://storify.com/3g4gUK