arcep
autorité de régulation
des communications électroniques,
des postes et de la distribution de la presse

FRENCH REPUBLIC

2020 code of conduct on Internet quality of service

FOR STAKEHOLDERS INVOLVED IN QOS MEASUREMENT

September 14, 2020

ISSN No. 2258-3106

2020 Code of conduct on quality of service

For stakeholders involved in QoS measurement

The Code of conduct is intended for stakeholders that conduct measurements designed to determine internet quality of service or quality of experience.

This document is an update of the 2018 version of the Code of conduct, whose purpose is to strengthen and increase the accuracy of the fixed and mobile internet quality of service measurements, and the publications on their findings. As with the first version, this update was produced by Arcep based on input from measurement tools, operators, consumer protection organisations and academics, with whom Arcep consulted during multilateral and bilateral meetings over the course of 2020. **This 2020 version will continue to evolve over time, with the implementation of the "Access ID card" API[1],** to strengthen the criteria listed, but also to supplement them by taking into account the information supplied by the API.

The Code of conduct defines a set of best practices whose purpose is to increase the transparency and quality of the tests performed and measurements taken, and of the resulting publications. It is divided into two main parts: Part 1 sets out best practices for the test protocols used to perform measurements, while Part 2 details best practices for the subsequent presentation of findings ("aggregate publications"). **Each part describes the methods that make it possible to guarantee both the transparency of the choices made** – so that any third party will be able to analyse the results produced by the tool – **and the robustness of the practices employed** – i.e. that they are reliable, representative and guarantee that the findings can be compared. These best practices for ensuring the method's robustness seek to avoid questionable practices, while keeping the field open enough to welcome innovation and diversity. As mentioned earlier, these practices will be further fleshed out in future versions of the Code of conduct, with the deployment of an "access ID card" API in the main ISPs' boxes. **The measurement tools wanting to declare their adherence to the Code of conduct are asked to employ the following declaration of compliance:**

"[Company name] declares itself in compliance with the 2020 Code of Conduct established by Arcep, in concert with the ecosystem's stakeholders, for the design of [name of tool]'s test protocols and/or the aggregate publication of the resulting measurements."

Measurement tools wanting to declare their compliance with the 2020 Code of conduct publicly agree to satisfy the transparency and robustness requirements regarding their test protocols (measurement methodologies and test servers) and aggregate publications. The conditions that must be satisfied to be in compliance with the Code of conduct are detailed in the framed paragraphs below.

Any party who uses the "Arcep" brand without the permission of the French Regulatory Authority for Electronic Communications, Postal Affairs and Print Media Distribution may expose themselves to civil liability claims.

[1] API (*Application Programming Interface*); cf. Decision No. 2019-1410 (in French): https://www.arcep.fr/uploads/tx_gsavis/19-1410.pdf.

Data privacy

It is up to the measurement tools to implement internal policies and procedures to maintain their compliance with Regulation (EU) No. 2016/679, commonly known as the General Data Protection Regulation (GDPR), and France's Law No. 2018-493 of 20 June 2018.

1 Test protocols

To comply with this Code of conduct, measurement tools must satisfy different transparency and robustness criteria. These criteria are detailed below for each of the different types of measurement tools: web / installable applications, Android / iOS mobile applications and hardware probes.

1.1 Measurement methodologies

Transparency over methodological choices is vital to ensuring that any third party can analyse the findings delivered by the tool.

If most of the choices made are worthwhile, some practices appear to remain questionable, and warrant being modified.

In order to make the results easier to understand and to increase the measurements' accuracy, tables 1, 2, 3, 4 and 5 set out the methodologies' minimum transparency and robustness requirements when measuring download and upload speeds, latency, web browsing and video streaming.

These parameters will be enhanced in future versions of the Code of conduct. New indicators may also be added.

Transparency and robustness of the measurement methodologies

Measurement tools that declare their compliance with the Code of conduct agree to satisfy transparency and robustness requirements (detailed in tables 1, 2, 3, 4 and 5 below) with regard to the methodologies used to measure download or upload speed, latency, web browsing and video streaming indicators*:*

- the measurement tools agree to complete, publish and update Annex 1 of this Code of conduct every six months;

- should a parameter vary between the tests, the measurement tools also agree to include the information obtained at the end of each unit test;

- the measurement tools agree to meet a minimum level of robustness in their measurement methodologies;

- the measurement tools agree to answer any potential requests from Arcep for additional information on their measurement methodologies.

Table 1: download and upload speeds

	Parameters	Web / installable applications	Android / iOS applications	Hardware probes
Transparency criteria	**Measurement protocol:** *Example (annex 1): HTTP/1.1*	List all of the information on the different parameters in annex 1 of the report. This annex 1 to be published on the measurement tool's website. When the information between measurements varies, this information should be displayed at the end of each test (e.g. in an advanced tab).		List all of the information on the different parameters in annex 1 of the report. This annex 1 to be published on the measurement tool's website.
	TCP or UDP ports used *Example (annex 1): TCP 80, 443 and 8080 ports* *Example (unit test): TCP port 443*			
	Number of TCP connections used simultaneously during the speed test *Example (annex 1): between 1 and 16 TCP connections* *Example (unit test): One TCP connection*			
	Length of each test (provided max. volume has not been reached) *Example (annex 1): 10 seconds*			
	Maximum volume of data exchanged *Example (annex 1): no limit*			
	Speed test stream encryption *Example (annex 1): Encrypted and unencrypted* *Example (unit test): Yes*			
	Information on whether or not slow start has been removed *Example (annex 1): exclusion of the first two seconds of the speed test*			
	Version of the Internet protocol and selection method used *Example (annex 1): IPv4 and IPv6. Test conducted in IPv6 (provided IPv6 is available end to end) and the user can choose to conduct the test in IPv4* *Example (unit test): IPv6*			
	Explanation of the indicators displayed once test is complete: explain the (IP/TCP) throughput displayed and the way in which it was calculated *Example (annex 1): peak TCP throughput: average TCP throughput during the fastest timespan, representing 30% of the total length of the test.* *Average IP throughput: average IP throughput during the entire length of the test*			
Robustness criteria	**Length of the test or volume of data exchanged**	Criterion: default test length ≥ 8 seconds or ≥ 100 MB of data.	Criterion: default test length ≥ 5 seconds or ≥ 50 MB of data.	Criterion: default test length ≥ 8 seconds or ≥ 100 MB of data.
	Propose a single thread test (can be optional) at least for Android/iOS applications	For tools that conduct multi-thread tests by default.		N/A
	Maximum number of speed tests per hour	N/A	N/A	Criterion: maximum 6 tests using an identical protocol per hour.

Table 2: latency

	Parameters	Web / installable applications	Android / iOS applications	Hardware probes
Transparency criteria	**Measurement protocol** *Example (annex 1): HTTP/1.1*	List all of the information on the different parameters in annex 1 of the report. This annex 1 to be published on the measurement tool's website. When the information between measurements varies, this information should be displayed at the end of each test (e.g. in an advanced tab).		List all of the information on the different parameters in annex 1 of the report. This annex 1 to be published on the measurement tool's website.
	TCP or UDP ports used *Example (annex 1): ports TCP 80 and 443* *Example (unit test): TCP port 443*			
	Number of latency unit tests (if overall time-out has not expired) *Example (annex 1): 20 tests*			
	Number of bytes typically exchanged for each latency unit test *Example (annex 1): 100 octets*			
	Length of the time-out in seconds, for each latency unit test *Example (annex 1): 1 second*			
	Length of the time-out in seconds, for all latency unit tests *Example (annex 1): 5 seconds for the 20 latency tests*			
	Latency test stream encryption *Example (annex 1): encrypted or unencrypted* *Example (unit test): Yes*			
	Version of the Internet Protocol (IP) and selection method used *Example (annex 1): IPv4 and IPv6. Test conducted in IPv6 (provided IPv6 is available end to end) and the user can choose to conduct the test in IPv4* *Example (unit test): IPv6*			
	Explanations of the indicators displayed at the end of the test: how latency / jitter / packet loss indicators are calculated *Example (annex 1):* *- Minimum: minimum latency of the 20 tests* *- Median: median latency measured*			
Robustness criteria	**Measurement protocols**	Criterion: do not use ICMP to measure latency.		N/A
	Number of latency unit tests	Criterion: number of tests must be equal to at least 10.		
	Result: display the median latency of the unit tests conducted	Criterion: median latency must be displayed upon completion of the test.		N/A

Table 3: web browsing

	Parameters	Web / installable applications	Android / iOS applications	Hardware probes
Transparency criteria	**List of the URL of the websites used** *Example (annex 1):* *- https://www.google.fr/* *- https://www.qwant.com/* *- etc.*	List all of the information on the different parameters in annex 1 of the report. This annex 1 to be published on the measurement tool's website.		
	Length of time-out in seconds, for each web browsing unit test *Example (annex 1): 10 seconds to load each page*			
	Length of time-out in seconds, for all web browsing tests *Example (annex 1): maximum 30 seconds for the 6 pages tested*			
	Web cache status *Example (annex 1): the cache is emptied after every page load*			
	Explanation of the indicators displayed at the end of the test *Example (annex 1): time it takes to load all of a page's elements, excluding advertisements*			
Robustness criteria	**Selection of the websites tested**	Criterion: do not use operators' portals.		N/A
	Time-out for each website tested	Criterion: time-out of less than 30 seconds.		

Table 4: video streaming

	Parameters	Web / installable applications	Android / iOS applications	Hardware probes
Transparency criteria	**Video platforms tested and resolutions (if the resolution is set in advance)** *Example (annex 1):* *- YouTube 720p* *- YouTube 1080p* *- Dailymotion adaptative resolution*	List all of the information on the different parameters in annex 1 of the report. This annex 1 to be published on the measurement tool's website.		
	Number of videos tested and selection method *Example (annex 1): the most popular video in the country (Number of views)*			
	Length of each video test *Example (annex 1): 30-second test (2 videos of 15 seconds each)*			
	Length of the time-out in seconds, for each video streaming unit test *Example (annex 1): 20 seconds*			
	Explanations of the indicators displayed at the end of the test: what formulas are used to calculate the different indicators listed *Example (annex 1): Average time to fill the two buffers, and total number of pauses while streaming the two videos*			
Robustness criteria	**Traffic stream encryption**	Criterion: employ the same encryption as the one used by default on the tested platform.		

Table 5: other information

Parameters	Web / installable applications	Android / iOS applications	Hardware probes
Generic information given to users on factors that might influence the different measurements, notably speed, web browsing and video streaming	Provide information on how Wi-Fi / OS / browser / device might affect the measurement. Give an example of the minimum hardware and software configuration that makes it possible to achieve a measured speed of 1 Gbit/s.	Provide information on how the device might affect the measurement.	N/A
Type of background tests performed *Example (Annex 1): background measurement of video streaming quality performed each time on YouTube*	List all of the information on the different parameters in annex 1 of the report. This annex 1 to be published on measurement tool's website.		

1.2 Test servers

Transparency over the test servers used (i.e. target servers) is also essential to understand the results. The test servers must also comply with certain conditions to ensure that the measurements are reliable. Table 6 describes these relevant transparency and robustness requirements for test servers.

These parameters are likely to be amended as the Code of conduct evolves.

Test server transparency and robustness

Measurement tools that declare their compliance with the Code of conduct agree to satisfy transparency and robustness requirements (detailed in table 6 below) with regard to test servers:

- the measurement tools agree to complete, publish and update annex 2 of this Code of conduct every six months;

- should a parameter vary between the tests, the measurement tools also agree to include the information obtained at the end of each unit test;

- the measurement tools agree to meet a minimum level of robustness for their test servers;

- the measurement tools agree to answer any potential requests from Arcep for additional information on their test servers and the methodology used to select the default test server.

Table 6: test servers

<table>
<tr><th></th><th>Parameters</th><th>Web / installable applications</th><th>Android / iOS applications</th><th>Hardware probes</th></tr>
<tr><td rowspan="4">Transparency criteria</td><td>Information on the host of the test server used
Example (unit test): Zayo France</td><td colspan="2" rowspan="2">Display information at the end of each test (e.g. in an advanced tab).</td><td rowspan="4">List all of the information on the different parameters in annex 2 of the report. This annex 2 to be published on the measurement tool’s website.</td></tr>
<tr><td>Information on the capacity of the test server used
Example (unit test): 10 Gbit/s</td></tr>
<tr><td>Explanation of how the default test server is selected
Example (Annex 2): Random – each test server is used for one in four tests, without the client being able to choose it manually</td><td colspan="2" rowspan="2">List all of the information on the different parameters in annex 2 of the report. This annex 2 to be published on the measurement tool’s website.</td></tr>
<tr><td>Provide the required detailed information on every test server
Provide the following information for each test server:
Sponsor (optional), city, region, use of IPv4/IPv6, connection capacity, port used, TCP congestion protocol (optional), host name and AS.</td></tr>
<tr><td rowspan="2">Robustness criteria</td><td>Test server capacity</td><td colspan="3">Criterion: do not use test servers with an internet connection of < 1 Gbit/s.</td></tr>
<tr><td>Capacity to conduct tests in IPv6</td><td colspan="3">Criterion: at least 20% of the test servers must be IPv6-enabled.</td></tr>
</table>

2 Aggregate publications

2.1 Data processing

Post-processing of the collected data is a crucial stage for eliminating false, manipulated or irrelevant measurements. It creates the ability to ensure that the results are representative and as widely comparable as possible, and to protect against attempted fraud.

The tools must therefore implement efficient data processing algorithms to deliver the most reliable results possible. It is particularly important that **measurements obtained from a target server that has proved to be limiting factor** (notably when the its capacity is below or equal to that of the line being tested) be excluded.

Arcep will consult with the ecosystem's stakeholders with a view to possibly establishing more detailed transparency and robustness criteria in the coming months.

2.2 Transparency of the published findings

To ensure that any third party can assess the reliability of the published findings, tools need to be transparent about the number of tests performed to obtain the subsequent aggregate publications, and report any bias attributable to the testing method that is likely to distort the representativeness or create comparability issues.

Transparency of the aggregate publications

Measurement tools that declare their compliance with the Code of conduct agree to satisfy the transparency requirements regarding aggregate publications by publishing the data listed in table 7.

Measurement tools also agree to answer any possible requests from Arcep for additional information on the processing of the collected data and their publication, notably regarding the post-processing methods used or the methods used to calculate the overall QoS score given to the connection.

Table 7: aggregate publications

Parameters	Fixed	Mobile
Period covered by the publication	Clearly indicate the period covered by the publication. *Example: from 1 April 2020 to 30 June 2020.*	
Number of tests per published category	Indicate the total number of tests for each published figure that aggregates several tests. *Example:* *xDSL category:* *- Bouygues Telecom 24 236 tests* *- Free 78 225 tests* *- Orange 145 265 tests* *- SFR 45 872 tests* *FTTH category:* *- Bouygues Telecom 85 872 tests* *- Free 125 265 tests* *- Orange 278 245 tests* *- SFR 45 236 tests*	Indicate the total number of tests for each published figure that aggregates several tests. *Example:* *3G category:* *- Bouygues Telecom 458 tests* *- Free 1 452 tests* *- Orange 782 tests* *- SFR 252 tests* *4G category:* *- Bouygues Telecom 2 523 tests* *- Free 7 824 tests* *- Orange 14 526 tests* *- SFR 4 587 tests*
Details on the data processing performed	Provide as many details as possible on the methods used to adjust the results.	
Test location	N/A	Indicate the percentage of customers per region, if publishing findings by region. Indicate the percentage of customers by population density in the tested location (separate by high/medium/low density areas and location unknown) and, if applicable, weighting used to calculate an aggregate indicator. Indicate the percentage of tests conducted while mobile (significant distance travelled between the beginning and end of the test).
Operating system	Indicate the percentage of operating systems per operator taken into account in the results.	Indicate the percentage of Android and of iOS devices used.
Version of the Internet protocol used during the tests	Indicate the percentage of tests conducted in IPv4 and in IPv6.	
Distribution of default test servers	Indicate the breakdown of tests by operator, with respect to the choice of test servers.	
Other factors likely to introduce a significant bias in the analysis of the compared categories	When the measurement tool publishes an inter-operator comparison “all technologies combined” or including a broadband / superfast broadband separation for fixed networks, it must clearly indicate that this combination of technologies used by ISPs introduces significant biases in the results. The tool must also indicate any other possible bias (test server’s limitations, user device, etc.).	Where there are significant differences tied to the devices, the results must be broken down: e.g. by type of device for each operator, or by indicating the percentage of tests per smartphone model, on the most widely used devices.
List simple indicators that are similar for all the tools	Display the results from peak traffic times (6 – 11 pm): - Median upload and download speeds with a slow start - Median latency	

This section could be further improved in the next version of the Code of conduct. The deployment of an “access ID card” API by the main ISPs will help improve the characterisation of the measurements significantly, and to round out the Code of conduct with criteria regarding the relevance and publication of the test results.

Annex 1 – Measurement methodologies

Every measurement tool wanting to declare its compliance with the Code of conduct must complete and publish the Table from annex 1 containing information on its measurement methodologies.

Download and upload speeds

Measurement protocol	
TCP or UDP port used	
Number of connections used simultaneously during the speed test	
Length of each test (provided the volume threshold has not been reached)	
Maximum volume of data exchanged	
Speed test stream encryption	
Information on whether or not slow start has been removed	
Version of the Internet Protocol (IP) and selection method used	
Explanations of the displayed indicators	

Latency

Measurement protocol	
TCP or UDP port used	
Number of latency unit tests (if overall time-out has not expired)	
Number of bytes typically exchanged for each latency unit test	
Length of the time-out in seconds, for each latency unit test	
Length of the time-out in seconds, for all latency test	
Latency test stream encryption	
Version of the Internet Protocol (IP) and selection method used	
Explanations of the displayed indicators	

Web browsing

List of the URLs of the websites used	
Length of the time-out in seconds, for each web browsing unit test	
Length of the time-out in seconds, for all web browsing tests	
Web cache status	
Explanations of the displayed indicators	

Video streaming

Video platforms tested and resolutions (if the resolution is set in advance)	
Number of videos tested and selection method	
Length of each video test	
Length of the time-out in seconds, for each video streaming unit test	
Explanations of the displayed indicators	

Other information

Generic information given to users on factors that might influence the different measurements, notably speed, web browsing and video streaming	
Type of background tests performed	

Annex 2 – Test server

Every measurement tool wanting to declare its compliance with the Code of conduct must complete and publish the table from annex 2 containing information on their test servers.

The three examples given in the table are provided solely for the purpose of illustration.

Test servers

Method for selecting the default test server								
Sponsor (optional)	**City**	**Region**	**IPv4/IPv6 protocol**	**Connection capacity[2]**	**Port used**	**TCP congestion protocol (optional)**	**Host name**	**AS (Autonomous System)**
Orange	Paris	Île-de-France	IPv4 or IPv6	10 Gbit/s	443	TCP Illinois	Orange	AS3215
One Provider	Vitry-sur-Seine	Île-de-France	IPv4 only	1 Gbit/s	443	TCP Cubic	Scaleway	AS12876
Adeli	Saint-Trivier-sur-Moignans	Auvergne-Rhône-Alpes	IPv4 or IPv6	1 Gbit/s	443	TCP BBR	Adeli	AS43142

[2] When a test server is hosted in a CDN, this test server's capacity should not be listed.