# Web QoE: Moving beyond Google's SpeedIndex

Dario Rossi, Enrico Bocchi, Luca De Cicco

## I. Problem statement

The World Wide Web is still among the most prominent Internet applications. While the Web landscape has been in perpetual movement since the very beginning, these last few years have witnessed some noteworthy proposals such as SPDY, HTTP/2 and QUIC which could disrupt the Web status quo and profoundly reshape the protocols family at application layer. Technically solid means are clearly needed to assess whether these new protocols can keep their promises: The risk is that these new protocols could otherwise fail to be adopted. While this investigation is already under way, both the industry and the research community are in our opinion *expressing the right question, to which they however answer using the wrong tools.*

Over the years, webpages have grown to quite complex entities including hundreds of objects of several types, sharded over many domains. Yet, the current practice is to express Web Quality of Experience (QoE) via the document completion time (*onLoad*) despite its known inaccuracy and poor correlation with the actual user experience. At the same time, while better metrics do exist (e.g., the SpeedIndex, proposed by Google in 2012), they are complex to evaluate and require a prohibitive amount of computing resources (i.e., record filmstrips of the visual rendering in the browser). As such, their use is limited to lab experiments, but have to date failed to catch up in larger scale.

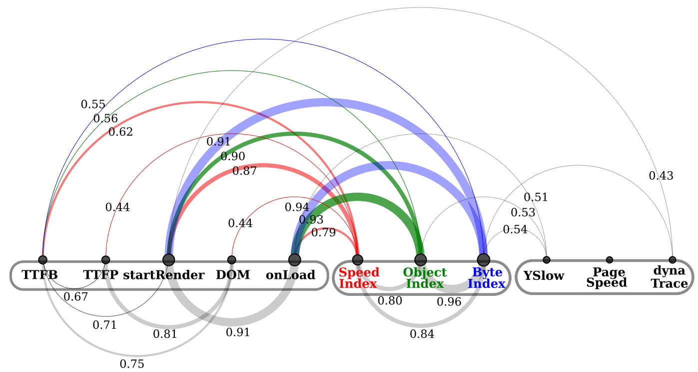
*Fortunately, there is a way out of this impasse: as it often happens, once found, the solution is very simple and elegant.*

## II. Innovation

SpeedIndex goes beyond the typical approach of picking one specific event in a webpage lifetime (e.g., DOM, onLoad, etc.) by using all events and integrating the residual visual completion in a single compound metric. However, the SpeedIndex does so by "by capturing a video of the page loading in the browser and inspecting each video frame (10 frames per second in the current implementation". Of course dumping high-quality videos in real time, albeit at slow rate, is a resource-intensive task to the point that the SpeedIndex cannot be computed in end-users devices without significantly slowing down the user experience (from 2× to 4×).

*Our main innovation is to propose ObjectIndex and ByteIndex metrics that are structurally similar (i.e., have the same integral form, weighting all captured events) but tremendously simpler to compute (by several orders of magnitude).* In a nutshell we argue that, to some extent, the objects (or bytes) that are received by the browser (or the network card) can provide a first approximation of the visual completeness of the rendering process.

We illustrate the relationship among the metrics we propose (i.e., ObjectIndex and ByteIndex), the SpeedIndex, and specific loading events (e.g., the Time to the First Byte – TTFB, the Time to the First Paint – TTFP, onLoad, etc.) on experiments performed downloading the top-100 Alexa webpages (10 repetitions per URL). For completeness, we additionally consider scores that are popular in the industry




such as Yahoo's YSlow, dynaTrace and Google's PageSpeed Insights. These last are compound scores weighting altogether a number of heuristics defined by domain experts.

We express the relationship via Pearson correlation, that we depict as an arc diagram. To reduce visual cluttering, we only report correlations larger than 0.4, with actual correlation values annotated in the plot. To let the strongest correlation emerge, we use a quantized scaling for line width, which doubles every 0.1 correlation steps. Metrics are arranged intro groups: (i) the left group comprises specific individual events from the page load and rendering process; (ii) the middle group reports the original SpeedIndex and our proposed variants; and (iii) the right-group reports compound heuristic scores. Correlations within group are reported in gray below the label name, while intra-group correlations are reported above. Correlations of the time integral group are highlighted in red (SpeedIndex), green (ObjetIndex) and blue (ByteIndex).

From the picture it appears that:

- The SpeedIndex correlates with several times-events.
- As designed, our proposed byte-level and object-level replacements do exhibit similar correlation to the SpeedIndex.
- the YSlow, dynaTrace and PageSpeed heuristics are poorly correlated among them (and with any other QoE metric).

## III. Expected Impact

Expected impact of our proposal is potentially very relevant for all Internet players from ISPs, to content providers, to browsers ecosystems, to end-users.

- ISPs: the byte version of the metric can be computed by routers, and this in spite of HTTPS and encryption-by-default.
- Content providers (CPs): the object-level version of the metrics can easily be computed in the browser, so that CPs can quickly obtain a feedback over the experience of their customers
- Browser developers: part of the critical path of the user experience lays in the browser; having a systematic way to assess users' QoE would simplify regression tests and help in making better and faster browsers experience-wise.
- Users: ultimately, the effort of previous actors to ameliorate the quality of user-experience and the expected impact of such metrics adoption cannot but be beneficial to all end-users.