CHARACTERIZATION OF THE ENERGY CONSUMPTION OF WEBSITES

Impact of website implementation on resource consumption

Olivier Philippot
Green Code Lab
Nantes, France
info@greencodelab.fr

Alain Anglade
ADEME
Sophia-Antipolis, France
alain.anglade@ademe.fr

Thierry Leboucq
Kaliterre Nantes,
France
tleboucq@kaliterre.fr

***Abstract*— This document presents the Web Energy Archive (WEA) project led by the Green Code Lab. This tools aims at measuring the computer resource consumption (energy, memory…) of the loading of a website in a browser. Firstly, considering the measure of 500 websites, website behavior can be characterized in terms of environmental impact. In a second step, best practices (green patterns) can be recommended to developers but also to browser editors.**

***Index Terms*— Energy consumption, web, environmental footprint, green patterns**

I. Introduction

The aim of Web Energy Archive (WEA) is to measure the energy consumption of websites from the user side by actual measurements. This tool measures how websites are proposed to the "client" (in a computing meaning that is to say from the user's point of view). The objective is to use it in real-life situations (market browsers, homepage visualization, etc.) and to inventory the largest number of websites in order to be representative of the Web activity.

WEA is available online at http://webenergyarchive.com and is used to measure more than 500 sites a year. Metrics measured by WEA give trends on websites environmental footprint.

WEA allows showing that the impact of a website is not the same depending on the level of eco-design of the site. The measured energy varies from 10 to 200 Wh per 1000 page views. This impact gets significant when considering the consumption implied by a large number of visitors. Considering the 100 most visited websites in France, the annual energy consumption goes up to 68 GWh or the equivalent of 25,400 household electricity bills. The estimates from the server side show that these 100 sites use 100 times less energy. Consumption in terms of downloaded data is just as important: 171 PiB (equal to 5.7 million Blu-ray Disc). This mass of information not only affects the customer but also the network and servers. Indeed, every byte received by the client goes through network elements (box, switch, router, etc.) and are transmitted by servers.

A browser comparison highlights differences in terms of resource consumption. The measurements show a higher consumption of Chrome by 27 Wh per 1,000 page views compared to Internet Explorer and Firefox browsers. Particularly, mismanagement in caching has been observed for all browsers. The consideration of these criteria by the browser editors would probably improve the resource consumption of user computers. In parallel, application of caching good practices by developers would reduce resource consumption on servers and network.

Finally, this analysis shows that the link between performance and website eco-design is not systematic. As an example, a high performance with a high Google Page Speed score does not imply a low memory or energy consumption. Thus, optimizations of both resources consumption and performance criteria go in the right direction.

II. Findings

A. Software bloat (Bloatware)

Websites that keep on loading, obligation to renew fully functional computers to install a new software version, problems of responsiveness and performance of business applications: software and data obesity (bloatware) handicaps companies while inexorably increasing the cost of information systems operation.

For the general public, this obesity is the cause of a perceived obsolescence of equipment and leads to an early renewal with an impact on the environment.

B. Web Bloatware

The power consumption of Internet infrastructures (data centers, networks, etc.) was estimated at 0.8% of global consumption in 2005. In 2012, it exceeds 2%, the same as civil aviation [1]. The Dresden University has calculated that if no action is taken, in the next 25 years, Internet will consume as much energy as humanity in 2008 (calculated by Gerhard Fettweis of the University of Dresden in 2008) [2]. The development of emerging countries and mobile applications worsen the current situation.

According to ADEME [3], most of the environmental impacts linked to a web page (depletion of non-renewable resources, pollution of soils and air, water eutrophication, etc.)

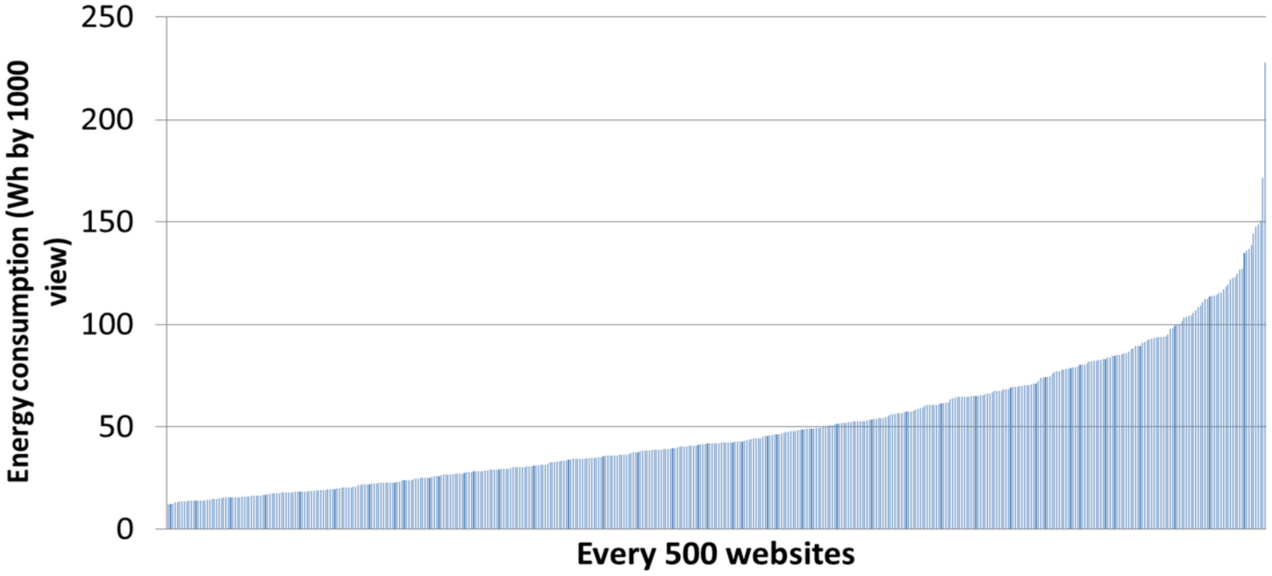

Figure 1- Energy Consumption of 500 sites from the user side

are correlated with time spent by the user in front of the screen and the active life of her/his computer.

III. PROJECT

A. Principles of the project

WEA's objective is to measure the energy consumption of websites from the user side by actual measurements. This measures how websites are proposed to the "client" (in a computing meaning, which is to say from the user's point of view).

The consumption from the server side is not taken into account initially by the measure (unavailability of analysis capacity on actual data) but by an estimate.

The objective is to represent use in real-life situations (market browsers, visualization of a homepage ...) and to inventory the largest number of websites to be representative of the Web activity.

B. Results

The expected results are energy consumption of websites from the customer side as well as the trend over time (per month). A correlation of consumption was made with certain parameters (page size, technologies, browsers...)

These data will allow a research orientation in eco-design and the identification of axes of consumption for pages and websites.

IV. PROJECT DESCRIPTION

A. Architecture

The solution needed to implement these following functional modules:

Module 1: This module allows managing measurement campaigns and databases. This is a robot which centralizes measures and presents data. It is based on the open solution "HTTP archive." [4].

Module 2: This module is in charge of measuring the energy consumption of different client platforms. It concerns home pages visualization of websites identified in the project. It is completely developed. A power meter was used for physical measurement together with a specific library for modeling the energy behavior.

Module 3: This module identifies the technical specifications of web pages of identified websites (technology, performance of web pages...) This data is then crossed with the energy data. The Web Page Test solution [5] is used for acquisition of performance data.

B. Measurement Methodology

Website consumption measurement is based on the PowerAPI framework [6]. The measurement is performed every second. In order to avoid the impact of other software on the studied one, a maximum number of applications are closed. The measurement methodology is then as follows:

1. Browser's cache is cleared,
2. Browser is launched with a blank local page,
3. Detection of a stabilized consumption,
4. Launch of the webpage,
5. Waiting 40 seconds,
6. Closing the browser,
7. Reiteration of the measurement with full cache (step 2).

Step 2 and 3 permit to measure stabilized energy consumption which is due by the system (Operating system, browser…). Stabilization is detected if power goes under a certain value (depending of the platform) and not decreases under a certain power delta. The time of detection depend of the system, of the context… We observed values between 1 minute and 5 minutes.

Measures are done while 40 seconds. 40s is an average time of consultation of webpages.

The measurements are performed three times to average the measurements. The measures deviations show that 3 is sufficient.

C. Measurement perimeter

The measurement is performed on different environments. Two configurations that allow having representative platforms of the market have been chosen:

- Configuration 1: P4HT Intel / Windows XP / 3GHz/ 512 MB,
- Configuration 2: Intel Core 2 / Windows Seven / 2.4GHz / 2 Go

The selected indicators were calculated by weighting the market shares and taking measures on the three major browsers (Chrome, Internet Explorer, and Firefox) with and without cache :

Metric = Measure of Configuration 1 + Measure of Configuration 2 * 3

In order to complement the results, some measurements were carried out on Android mobile platforms. We began the study with PC platform to validate the protocol and the principle of WEA. Moreover, automatization of the platform is more difficult for Android. Widespread measures will be done on smartphone platforms later. The tests were performed on a 5 LG Optimus II (LG- E460) hardware platform and OS Android 4.1.2 software platform. Hardware sensors of the smartphone were used to retrieve information concerning the battery: voltage and current consumed.

The same methodology as for the PC platform is used:

- Chrome's cache cleared,
- Chrome is launched with a blank page,
- Waiting for idle (15 s),
- Reset to 0 in energy consumption,
- Launch of the webpage,
- Waiting 40s,
- Recovery of energy measurement.

V. Results for PC

A. Panel of websites

500 sites of various types have been measured during one year with a one month frequency. In this article, the analysis is based on the measures of November 2013. The sites were selected to be representative of the most visited sites. Many visitors added more sites during the measurement year which improved representativeness.

B. Analysis of measurement

1) Correlations

Several metrics were measured:

- Impact of energy consumed on the client device (Wh),
- Impact of RAM (MiB),
- Transferred Data (MiB),
- Loading time for pages (s),
- Number of requests to display a page.

This metrics were chosen in order to characterize all the impacts of the software on the hardware. the CPU load was excluded because it is take into account in the model of Power API. RAM was taken into account because memory is the principal cause of obsolescence. The transferred datas metric and number of requests were taken into account because there are simple and convenient metrics which characterize the size of a website. The loading time metric permits to measure the performance of website.

The distribution of the measured energy is as in figure 1.
The average is 50.1 Wh. The median is 42.2 Wh.
The correlations between the different measures are:

Energy / RAM	Energy / Google Page speed	Energy / Req Number	Energy / Loading speed	Energy / Web site size	RAM / Google page speed	RAM / Req Number	RAM / Loading speed	RAM / Website size
0.77	0.17	0.69	0.30	0.50	0.23	0.73	0.39	0.66

Table 1 - Correlations

The correlation coefficient for the energy / RAM relationship is 0.77. The relationship between energy consumption and RAM memory is strong. Similarly, the correlation between RAM and the number of requests is also strong. See figure 2. This can be explained with the fact that the more there are request, more treatments need to be done.

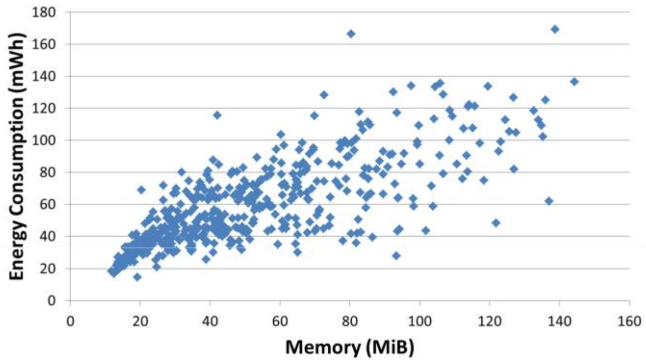

Figure 2 - Energy vs RAM

The impact of websites on the resource consumption of client computers can fluctuate depending on the hardware platform, the browser and the website. However, some trends can be observed: more requests and heavier pages lead to more important resource consumption (in particular in terms of energy and RAM). An optimization of consumed resources is possible on the side of developers and website users.

2) Performance vs Ecodesign

Google Page Speed is a framework of good practices for performance of web pages [7]. The evaluation of the page gives a score out of 100. The distribution of energy consumption of measured sites depending on the Google Page Speed score is as follows (Figure 3)

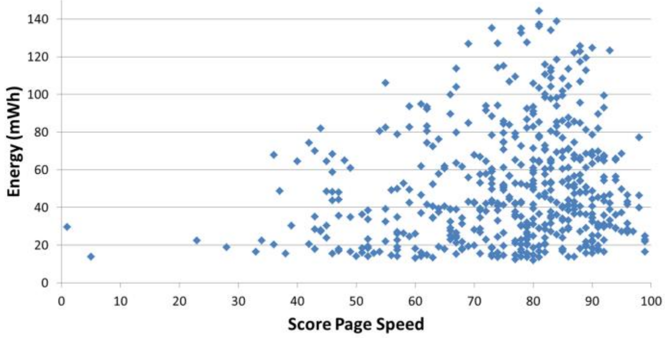

Figure 3- Energy vs Google Page Speed score

The correlation coefficient is 0.03. So there is no correlation between the performance (depending on Google framework) and the consumption from client side. Indeed, one can observe that quite efficient sites (Page Speed Score > 90) have disparate consumptions (spread between the minimum and maximum consumption of the sample).

This can be explained by the fact that the recommendations of performance aim at making available web page as quickly as possible to the user. A good practice would be to move the loading of the scripts at the end of the page for instance. This practice increases the page loading but don't reduce the energy consumption.

3) Browsers

Firefox and Internet Explorer consume the same energy. However, Chrome consumes much more (27 mWh more by page than Firefox). See Figure 5.

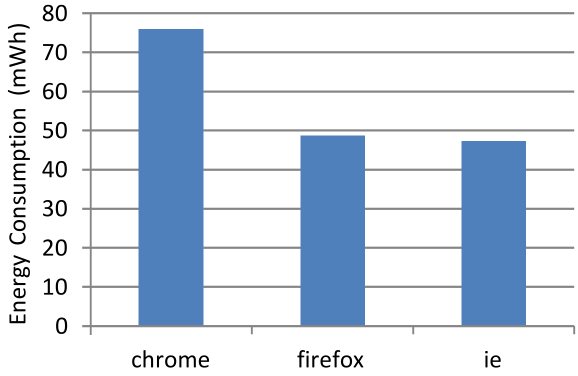

Figure 5- Energy consumption of browsers

However in terms of memory consumption, Chrome is still the largest consumer (61 MiB), but not as significantly. Firefox plays its cards well (13 MiB per site less than Chrome against 3 MB less than IE).

The study of the Fraunhofer institute for Microsoft [8] has studied the consumption of the 3 browsers. The conclusions are that IE is the less consumer and Firefox the more consumer. We have analyzed that this study is not as complete because there are only 10 websites versus 500 for ours study. Moreover Fraunhofer study is focus on power (Watt) and not energy (watt-hours). Power is not sufficient because power can change a lot every second. To manage the cache, Internet Explorer handles better reloading (gain up to 3.3 mWh). See Figure 4. Firefox manages not as well as IE but there are energy savings. But, Chrome uses more energy for the reloading of the page a second time.

To refine the finding, the sites that had a mixed of cached and not cached elements have been excluded.

The gains on Internet Explorer are less important if there are less cached elements. However, the conclusion is reverse for Chrome and Firefox: there are more gains when there are no cached elements.

These results can be explained by the fact that even if a page is not downloaded again (when it is already in the cache), however many tasks must be performed: communicate with the server to retrieve http information, analyze the results, search the cached items, read items on the disc, post the elements.

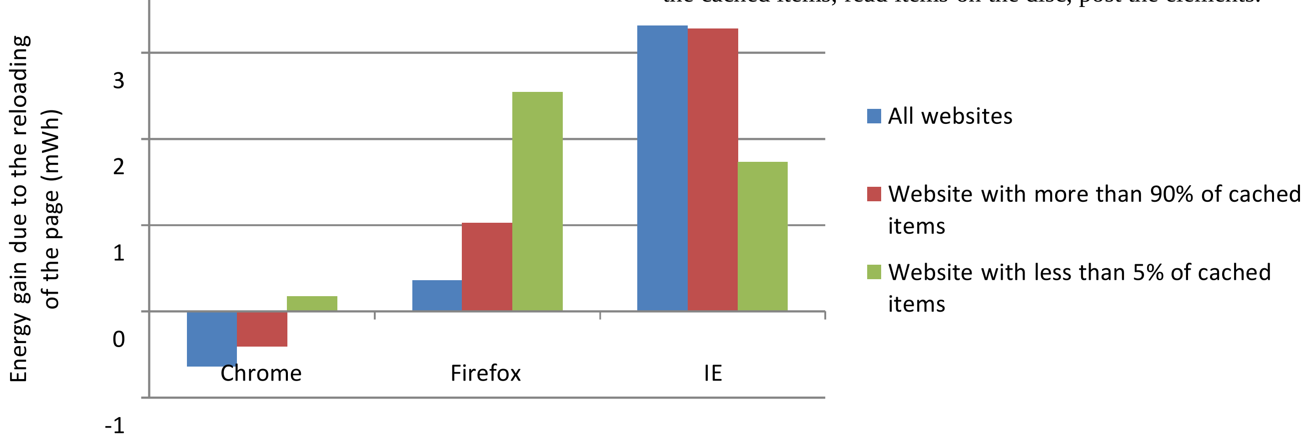

Figure 4 - Management of browser's cache

These tasks can consume more or less compared to direct download of items. The different browsers do not seem to manage similarly this caching. Chrome and Firefox seem to better manage the re-download than the cache recovery. This can be explained by a better management of the cache memory or by file writing to the disk.

So a better cache management can be both done with the improvement of browser on the client side and both with correctly applying HTTP cache control mechanisms on the server side.

C. Energy Label

Energy rating was established with 7 classes (A to G). The calculation method is as follows: division in 7 classes, the equivalent to taking a minimum and maximum. These thresholds are dynamic and change according to the measurements of previous months. The classes are calculated to obtain Gaussian repartition of all websites : There are several websites with E and D. A is difficult to obtain because this need a very light website. This will permit to increase continuously the green level of websites. The dimensionless energy efficiency index (EEI) will be used later in the project.

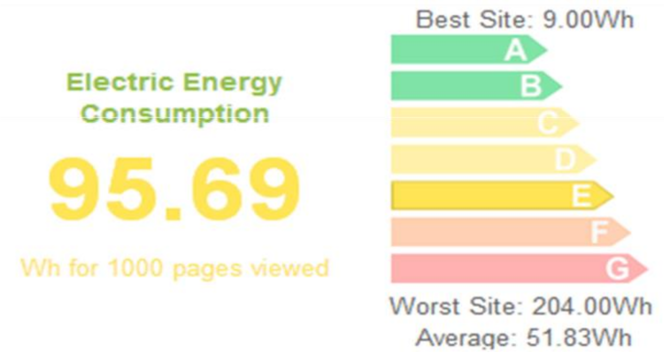

Figure 6 - Energy Label

VI. Android results

The average energy consumption is 9.2 mWh and the median 8.9 mWh, about 5 times less than the PC platforms.

The average consumption of the clear cache is 9.43 mWh and the consumption with a full cache is 9.00 mWh. Savings are very low. These conclusions are similar to the PC platform ones. Indeed, few sites integrate cache management. Moreover, as seen on browsers, the cache is not necessarily well managed and consumed resources for reading cached items can make the loading process as consumer.

VII. Server results

A. Methodology

1) Issue

Unlike the evaluation of the client part, the server part cannot actually be measured. Indeed, it would require the integration of measure probes within servers of websites. However, several elements allow estimating consumption:

- The characteristic of the page (Element size, number of requests),
- The origin of the elements.

An estimate of the server consumption based on the items that are measured by WEA is possible.

2) Methodology

The capacity and sizing of a server can be compared to the metric request per second (ReqS). This is the server's ability to deliver the number of requests during a second. This figure will depend on the server technology.

The server time used by a request can then be deduced; it is the opposite of the number of requests per second (1/ReqS).

Then, the power of a web server has to be taken as a metric. In most modern data centers, web servers are virtualized and hosted on physical machines. The power of the web server will be proportionate share of the power of the physical server hosting virtual machine (PServer / NbServer). Furthermore, it is necessary to take into account the energy consumption of the elements which constitute the data center. The air conditioning consumption, power components, etc., have also to be considered. Thus, the Power Usage Effectiveness (PUE) [9], which characterizes the energy efficiency of the data centers, can be used.

The energy consumption of a request (mWh) can be calculated as follows:

$$Creq = PServer / NbServer * PUE * (1/ReqS) / 3{,}600$$

B. Application

1) Hypotheses

At first, all servers' data cannot be collected so hypotheses have been developed.

ReqS:
A classic server typically has a capacity of about 500 requests per second. Some specific servers such as Content Delivery Networks (CDN, which can deliver content faster) go up to 2,000 requests per second [10] [11]. It is initialized to 500 requests per second.

PServer / NbServer:
According to measurements made during joint operations of ADEME [12], the average power of virtual machines is 20 W.

Average PUE:
According to Digital Reality study, the average PUE of data centers in 2012 is 2.53. [13]

2) Calculations and Analysis

On these assumptions, an average consumption of 28 µWh is thus obtained for a request. To display a full page, it is necessary to perform several requests (1 to several tens)

For all measured sites, an average consumption of 2.7 mWh with a median of 1.7 mWh has been obtained.

This methodology has been applied with the French newspaper Le Monde website lemonde.fr and an impact of

3.3 mWh for the consultation of the page has been calculated (See Figure 7)

One can observe that the energy impact of the consultation of the site on internal servers (as opposite of CDN or ad servers for example) is not the most prominent. This spread use brings complexity in analyzing the impact of a site but brings also a potentially greater impact. The impact of a page view is not concentrated only on the server, but also on the network between the client and the server and on the client. In addition, the impact will be different depending on where the servers are located (even more if the greenhouse gas emissions that may be different between countries are taken into account).

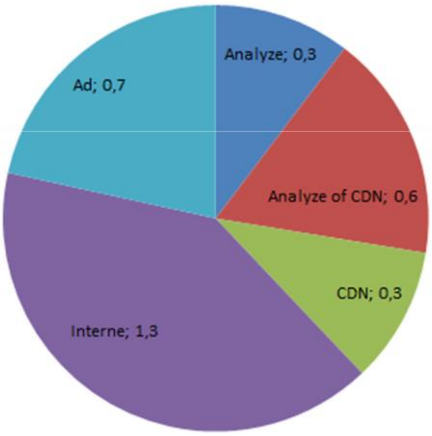

Figure 7- Lemonde.fr server consumption by type of server (in mWh)

VIII. Macroscopic analysis

A. PC and Server

As a confirmed audience, the 100 most visited sites in France according to ODJ / Médiamétrie have been chosen.

The total calculated consumption is the impact of energy consumption from client side of the most visited sites. For each site, the calculation is as follows:

Consumption (MWh/year) = Number of page views per month * Unit Consumption (Wh) * 12/1,000,000

The impact for these websites has been assessed to between 2 MWh and 1,940 MWh per year. The total impact of these sites is 8.3 GWh per year or the equivalent of 3,077 French households electricity consumption.

The “impact” term is used because it is the extra cost of consumption, from the client side, created by the site. To get the total consumption of the view of the website, it is necessary to take the consumption of the user PC in idle mode. If an average power of 42 W [12] and a measurement time of 40 seconds are considered, an energy consumption of 42 * 40 / 3,600 or 466 mWh per page view has to be added to the metric.

This leads to the consumption of 68 GWh or of 25,400 households’ equivalent.

With the calculation methods of the previous chapter from the server side, the following calculation can be done for each site:

Consumption (MWh/year) = Number of page views per month * Unit Consumption (2.7 mWh) * 12/1,000,000,000

The impact of these sites from a server-side is 0.58 GWh. The total consumption is 171 PiB, the equivalent of 5.7 million Blu-ray Disc. This mass of information not only affects the customer but also the network and servers. Indeed, every byte received by the client goes through network elements (box, switch, router ...) and are transmitted by servers.

Taking the hypothesis of a possible optimization of these sites, these 100 sites would consume at maximum the average of all sites (49 mWh and 1.7 MiB for the size of the page). The potential savings would be 2.24 GWh or the power consumption of 833 households’ equivalent and 48.7 PiB.

B. ANDROID

Consumption on mobile platforms is smaller (test on 126 websites): in average 9 against 49 mWh on fixed platforms. However there is a significant disparity: 5 mWh for the lightest sites against 13 mWh for the heaviest sites.

On the tested mobile phone, the battery capacity is estimated to 1,700 mAh by taking the average voltage of the battery of 4 V, an available energy of 6,800 MWh. Taking our hypotheses of consultation (40 s per page or 90 pages per hour), if the user visits the site for 1 hour, the consumption is as follows:

The lightest sites : 90 * 5 or 450 mWh or 6.6% of total available energy,

The heaviest sites: 90 * 13 or 1,170 mWh or 17.2% of total available energy.

It is clear that the resource consumption of web sites on a smart phone can have a significant impact on the battery. Assuming a one hour consultation per day, the difference goes up to 10.6% less in autonomy which corresponds to one more discharge cycle every 10 days.

IX. Conclusions

A. Web sites are not all equal before eco-design

Measuring the resource consumption by WEA shows that a great disparity exists between the 500 websites. There is a correlation between the number of requests, the page size and the consumed RAM on the client and the energy consumed. Similarly, the correlation between RAM and energy exists. However, the causal connection has not been proven. Indeed, for example, does the number of requests affect the RAM or the size of the requests? Does the lower consumption of RAM impacts the energy consumption? However, website frugality in terms of elements (number of requests, size of items, etc.) automatically leads to a reduction of resource consumption (energy and RAM).

Such a finding is even more important that the most visited sites (particularly news sites) are also the heaviest. Given the scale effect due to the millions of visitors of these sites, the impact of the use phase is significant from a user's side point of view. Indeed, the energy consumption from the client side has been proven to be 100 times more important than from the server side.

This disparity is present between browsers. The resources consumed by Google Chrome have been highlighted to be much higher than those of other browsers. Overall, the publishers should control this impact.

The management of cache needs to be optimized. Browsers only bring little benefit during the reloading of the page. This however does not remove the benefit of cache mechanisms from the server and network sides. However, it is unfortunate that the client side, which may be considerable in terms of impacts, will not benefit. This observation is also true on mobile platforms where efficiency is even more important.

On these platforms, there is lower resource consumption but a same disparity in consumption levels. The impact of websites on the autonomy is not negligible between an optimized website and a non-optimized one.

The websites tendency to use external services (e.g. CDN or scan engines) has been noticed. The environmental impact of a website is much more complex. However, it is clear that the impact of a service is highly distributed: server, network, clients but also external servers and associated network. The unit impact of all these elements may be low but the sum of all the services and the scale effect makes the measurement and control of this impact, necessary.

B. A platform that showed interest

Beyond the results on resource consumption of websites, the platform WEA showed many interests. Throughout the project, contributors and members of the Green Code Lab used WEA for workshops, studies and demonstrations at trade shows and conferences. It shows several important things:

The platform allows making sensitive users (not just developers) on the environmental impact of web sites. Simple metrics and the energy label make it possible to concretely understand the impact of a website.

WEA can raise alerts from companies who thought they applied best practices in terms of eco-design but ultimately consume a lot in terms of resources. WEA is a management tool for decision makers.

The common feeling is that performance and eco-design are related. WEA measures show that they are not necessarily correlated. It is necessary to actually measure good practices to outline that performance practice is or is not in agreement with the eco-design.

Even if the assumptions need to be improved, WEA can accurately characterize the environmental impact of a website from the client side and to take concrete actions.

C. Numerous possible developments for WEA

To be representative of the actual use of web sites, WEA must evolve regularly and take into account trends in terms of usage and equipment. That is why it is necessary firstly to increase the number of sites measured. This requires strengthening the measurement capacity (currently 3,000 websites per month) by purchasing measurement servers. This number (fixed and mobile PC) has to be increased. Similarly, it would be necessary to increase the number of operating systems and browsers characterized with WEA.

WEA must also extend the analysis to the entire Internet chain: server but also network. Indeed, as shown, it is necessary to measure to be able to understand and act. In this way, in order to identify points of important consumption and best practices, measurement and the overall estimate are required. One can imagine that the location of the server can greatly impact the consumption of the network part.

It is also interesting to add functionalities that allow going further in the analysis of trends (e.g. measurements of several pages of the site, scrolling the page...) This would allow more developers and users to identify concrete good practices.

D. WEA Observatory for awareness and action

WEA platform allows collecting a lot of information on trends in resource consumption of the websites. A lot of measured data have not yet been analyzed (type of servers that host, website category, characteristics, etc.) Thanks to these data, trends can be better understood and best practices can be identified. The Green Code Lab aims at supporting this work to establish the first observatory for Internet consumption. WEA is the first brick.

Acknowledgment

We thank ADEME for the support they gave us during this project. We also thank Green Code Lab members for their contributions to the project.

References

[1] *GESI*. http://gesi.org/ICT_sustainability_studies_and_reports

[2] Fettweis, G., Zimmermann, E. 2008. *ICT Energy Consumption – Trends and Challenges*. 11th International Symposium on Wireless Personal Multimedia Communications (WPMC 2008). https://mns.ifn.et.tu-dresden.de/Lists/nPublications/Attachments/559/Fettweis_G_WPMC_08.pdf

[3] Bio Intelligence Service For ADEME (2011, Juin). *Etudes d'analyse de cycle de vie de la communication électronique.*

[4] *HTTP Archive.* http://httparchive.org/

[5] *Web Page Test*. http://www.webpagetest.org/

[6] *Power API.* https://github.com/abourdon/powerapi-akka

[7] *Google Page Speed Best practices*. https://developers.google.com/speed/docs/best-practices/rules_intro?hl=fr

[8] Fraunhofer – "THE IMPACT OF INTERNET BROWSERS ON COMPUTER ENERGY CONSUMPTION"2013

[9] *PUE Definition*. http://www.thegreengrid.org/sitecore/content/Global/Content/white-papers/The-Green-Grid-Data-Center-Power-Efficiency-Metrics-PUE-and-DCiE.aspx

[10] *Random drake. Server Test Load* http://randomdrake.com/2009/07/14/benchmark-results-show-400-to-700-percent-increase-in-server-capabilities-with-apc-and-squid-cache/

[11] *Load Storm. CDN test load*. http://loadstorm.com/2013/08/webperflab-impact-cdn/

[12] ADNOuest. *Collective measure operation of IT ADEME*. http://www.blog-adnouest.org/premire-opration-collective-en-france-sur-la-mesure-de-la-consommation-de-lit/

[13] *Reality - PUE*. http://www.digitalrealty.com/us/knowledge-center-us/?cat=Research