Effects of Web Page Contents on Load Time over the Internet

Zhou Munyaradzi[1], Giyane Maxmillan[2], Mutembedza Nyasha Amanda[3]

[1, 2, 3] Midlands State University, Faculty of Science and Technology,
Department of Computer Science and Information, Gweru, Zimbabwe

Abstract: ***This paper focuses on the effects of web page components on web page load time and how they can be modified to reduce load time. Website speed measuring software tools (EValid and YSlow), online website monitoring tools and a mathematical model was used to make quantitative and qualitative analysis on web page load time in relationship to web page content. This research showed that university X's website loads slowly within an average of 9.766 seconds and may even fail to load as indicated by its average downtime of 9.32%. For application performance across the internet to improve, the website load time must be on average 8 seconds which increases the image of the site hence the perceived quality of services or products offered and thus increasing stakeholders' satisfaction. The findings of the study revealed that an increase in total size of web page content components is directly proportional to increase in response time and web page content components have different effects on web page load time. The evidence in this study shows that web page load time is mainly affected by its web page content characteristics. The study recommends web developers to measure web page response time against threshold values such as the 8 second rule during designing and implementation of a web page and adapt techniques identified which reduce load time.***

Keywords: Website optimization, EValid, YSlow, Web page load time, web page content.

1. Introduction

Websites play an important role in academic institutions especially in universities and colleges as they are used as a method of teaching and a means of online internationalization. Internet users demand the similar level of performance as they experience when connecting to applications via the LAN. Typically, websites on LAN load faster mainly because the web page content is stored locally requiring less load time [1]. Most universities have stakeholders globally dispersed who study at campuses or purse their studies online and they rely more on their websites as a tool to convey information. The load time of websites is one of the most important factors affecting its usability, most internet users will just skip a site altogether if it fails to load within an average of 8 seconds [2]. To increase its global market share the institute must standardize its web page load time. Information Communication Technology is a strategic goal of University X and now it relies more on its website and also its web portal. The website has e-learning facilities where students access their learning materials (notes, assignments, continuous assessments marks, module course outlines, and so on), a chat panel between student and lecturer, news and updates through a notice board for both staff and students. The university's degree programmes include conventional, parallel, block-release and visiting students. Block-release and visiting school students usually access their e-learning accounts out of campus since they are employees in different regions of SADC and abroad more so to international students during their holidays/potential international students hence they access the site over the internet so the load time of website needs to be analyzed. A second delay in web page response time leads to a 16% decrease in customer satisfaction [3].

University X's mission statement on exploiting ICT is being implemented and there is need to meet international standards as they use the website as a means of teaching, admission of international students, collaboration, e-learning and so on. The institute is in collaboration and is in the process of collaborating with different universities around the globe and national colleges who access the website over the internet. Various groups of students such as visiting and block-release access the site externally so the issue of web page load time is important since they claim the website response time is slow and it might fail to load completely. Having more diversified web users means there is need to meet more diversified user needs, which challenges the web developers and systems' capacity. This research seeks to determine how websites load time can be sped up, establish whether the number of HTTP requests has a direct relationship with latency, find out how web page objects such as images files affect website load time and how these objects can be optimized to reduce load time, investigate whether web page load time differ in two major browser technologies, Internet Explorer and Firefox and lastly examine how web page weight relate to web page uptime/downtime. This research will help to cement the institute's mission on how it can exploit Information Communication Technology (ICT) on websites. Also the Information Technology Services software development team will benefit from the recommendations to improve the site's performance globally. At many organizations, web-based applications are the backbone of business-critical processes such as ecommerce operations, financial transactions, and media. End users have direct connection to applications to initiate and complete transactions from their web browser, so the success or failure of web-based applications depends on a fast and reliable access conditions for the end users. This is particularly critical for e-commerce applications where even small differences in response times can have a dramatic effect on such metrics as page views, number of searches, and site revenue. Website optimization will help enhance the organization's image and increase stakeholder's loyalty. One of the benefits is helping to protect and even increase online revenue by preventing downtime for critical systems. Another related benefit of

preventing downtime is preserving the reputation of the company.

2. Materials and Methods

Literature study on web optimization and experiments on University X's website were used as a case study in this investigation since it is easily accessed over the intranet. Website online monitoring tools were used to make quantitative and qualitative analysis on web page load time for the locations including Los Angeles and New York, (US), Gloucester (UK), Dortmund (Germany), St. Petersburg (Russia) and Sydney (Australia). Mainly a quantitative research method was used through using page-speed and page-test measurement tools. YSlow, a software tool was used to make an analysis of web page components since it analyzes web page performance and shows why performance is slow. EValid software tool was used since it clearly shows the identifiable characteristics of a web page and their sizes in kilobytes. A mathematical model was developed and was verified using an online measurement tool YSlow.

3. Results

The results obtained from the experiments will be presented in the form of tables and figures.

3.1 YSlow and EValid experiments

Table 1: Object type and size

Object Type	*Size(bytes)*	*Object Load Time (s)*
HTML	6408	0.23
HTML Images	101643	4.34
CSS Images	0	0
Total Images	101643	4.34
JavaScript	23493	0.92
CSS	1281	0.21
Multimedia	0	0
Other	0	0

Table 2: Effects of Objects size on response time

Object	*Increase/(Reduction) in response time as a percentage with respect to change in size of the object*
HTML	(4.72%)
HTML	(45.91%)
JavaScript	552%
CSS	25%

Table 3: External Objects

External Object	*Quantity*
Total HTML	1
Total HTML Images	19
Total CSS images	0
Total Images	19
Total Scripts	4
Total CSS imports	1
Total frames	0
Total Iframes	0
Total objects	**25**

3.2 JavaScripts position

There are three (3) JavaScript scripts found in the head of the document (home page). These JavaScripts need to be put at the bottom of the page to reduce latency which results in an increase in web page load time.

3.3 Expire headers

There are 40 static components without a far-future expiration date. These components need to be specified the day, month and year which they will expire so that the content will become cacheable hence reduce web page load time.

3.4 Static components

There are 23 static components that are not on a Content Delivery Network (CDN). There is need to specify CDN hostnames in different countries within all continents which will dramatically improve the speed of a website since users are artificially nearer to the server.

3.5 Images

HTML images have a total size of 4 305 bytes which is approximately 3.24% of the total web page size and these images have been compressed using Gzip. An experiment carried out using Yahoo! Smush.it tool showed that since these images have been compressed, there are no savings which can be done on these images. The experiment was carried out since a reduction in web page size will correspondingly reduce web page load time.

3.6 Uptime and downtime experiments using Uptrends

An analysis of 10 reports from uptrends was done on the current web page and uptime was 99.32% hence a page size of less than 146 744 bytes results in a decrease in downtime which is also related to load time.

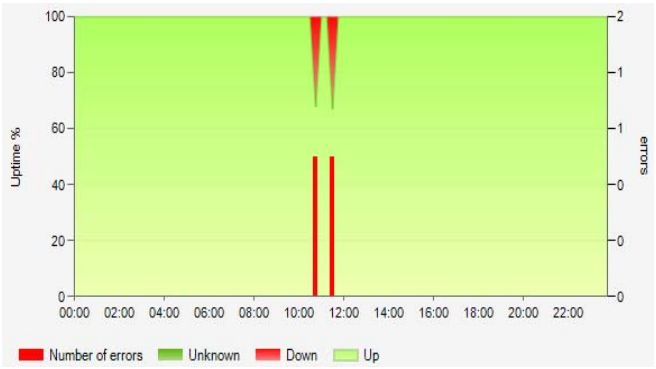

Figure 1: Daily Uptrends Management report for account: **293361**

Table 4: Uptime and downtime sample report of the current web page.

Vital statistics in period	
Uptime	99.32%
Number of errors	2
Number of warnings	2
Number of alerts	4
Number of checks	290

3.7 Website load time (24*7)

Ten online experiments were carried out on the monitoring site https://www.site24x7.com/ and an average response time

of 6.5seconds was recorded, there is a direct relationship between web page size and load time. A sample of a report carried out is as follows:

Figure 2: Website load time

The correlation coefficient of all objects with relation to response time is 0.032286153 which shows that there is no linear pattern in the size of objects in relation to response time and this can be further cemented by the least square formulae derived from these results: Load time (milliseconds) = 0.888255814 (object type size) + 1722.92774 milliseconds. A linear pattern can be achieved when comparing specific object type size in relation to load time. This means that for a specified object size the **Response time(s)** = (object size/4480) – 0.09728±0.7s (packet loss factor)

3.8 Internet supervision website load time

An internet supervision site www.internetsupervision.com was used to carry out ten experiments on different check points supported by the service after upgrading of the site. The average load time was as follows according to check point and its connection speed:

Dortmund, Germany (5 Mbps) 2.65
Beijing, China (5 Mbps) 3.02
Sydney, Australia (5 Mbps) 3.45
Gloucester, UK (5 Mbps) 2.82
Chicago, IL (45 Mbps) 2.25
Washington, DC (3 Mbps) 4.58
Los Angeles, CA (1.5 Mbps) 2.50
Santiago, Chile (760 Kbps) 2.83

Ten experiments were done using the ping and tracert commands for Harare and Johannesburg. The average web page load time in Harare (Zimbabwe) was 1.85 seconds and in Johannesburg (South Africa) it was 2.35 seconds. Generally the web page load time mean is 5.3 seconds using online monitoring tools for the selected check points which were 0.4 seconds less than the mean response time carried out using EValid and YSlow tools.

3.9 Multi-browser website performance

Five experiments were carried out on multi-browser website performance and on average Internet Explorer was 1.427 slower than Firefox using the online tool on www.gomez.com. The results of the experiment are shown in **Table 5.**

Table 5: Browser and Location Response Time Differences

Location	*Internet Explorer (seconds)*	*Firefox (seconds)*
New York, NY	9.766	6.061
Seattle, WA	9.118	8.445
Los Angeles, CA	9.055	8.206
Chicago, IL	6.882	6.321

4. Discussion

The results from this study shows that web page load time is mainly affected by its web page content rather than client and web-server network conditions. There are 23 static components that are not on a CDN. The average load time according to check point and its connection speed shows that an increase in distance between the client and web server increases response time. There is need to introduce CDNs (the server with the fewest network hops or the server with the quickest response time is chosen since it reduces response time as noted by [4], [5], [6].

Web page content components have different effects on web page load time. JavaScripts increase the number of HTTP request which will also increase web page load time. There are 3 JavaScript scripts found in the head of the document. These JavaScripts need to be put at the bottom of the page to reduce latency which results in an increase in web page load time.

There are 23 static components without a far-future expiration date. These components need to be specified the day, month and year which they will expire so that the content will become cacheable hence reduce web page load time. Caching is the most important and most widely used performance improvement technique for web-based systems since retrieving data from these caching locations will reduce transmission time across the internet as noted in the motivational literature [7], [8], [9]. The study revealed that on average Internet Explorer is 1.44725 slower than Firefox. An increase in the number of connections per server a browser accommodates reduces web page response time which agrees with web browser technology literature review. Firefox allows an average of 6 connections per server which increases parallel downloads while Internet Explorer allows an average of 3.

[10] revealed that the larger the page, the longer it takes to load as was noted in the mathematical model and the online experiments. However studies by [8]; [10] does not show the relationship of web page content characteristics to response time which the study had clearly identified. The web page images had either a width or height missing attribute which causes poor response time. Specifying the width and height attribute of an image makes a web page to load the entire objects before the image is fully uploaded hence reducing response time. HTML images have a total size of 101 643 bytes which is approximately 76.524% of the total web page size.

5. Conclusion

Web page content affects the web page response time in relationship to object type and size. A critical analysis of the web page characteristics needs to be carried out by web developers to deploy a website which will comply with the 2 seconds rule by [11]. Load time increases as the geographical distance increases from the server to the client assuming that clients have the same connection speed. The problem of distance can be solved by using CDNs. From the motivational literature Internet Explorer has the largest market share while the web page is faster on Firefox than on Internet Explorer hence the developers must tailor the web page content to reduce the response time when using Internet Explorer.

6. Recommendations

The total number of objects on the web page is 25 which by their number will dominate web page delay and therefore objects must be reduced. Above 20 objects per page the overhead from dealing with the actual objects (description time and wait time) accounts for more than 80% of whole page latency. There is need to combine, refine, and optimize external objects. Graphic rollovers must be replaced with CSS rollovers to speed display and minimize HTTP requests. Using CSS sprites helps in consolidating decorative images. Using CSS techniques such as colored backgrounds, borders, or spacing instead of graphic techniques can reduce HTTP requests. The total number of images on this page is 19 which need to be reduced through combining, replacing, and optimizing graphics and hence focusing on reducing size. Replacement of graphic rollover menus with CSS rollover menus will speed up throughput and minimize HTTP requests. Also optimizing parallel downloads by using different hostnames to reduce object overhead.

The total size of the page is 132411 bytes, which loads in over 20 seconds on a 56Kbps modem or 31.39 seconds on a 56Kbps modem, response time of less than 20 seconds can be achieved by reducing the web page size to not more than 100Kb. The total number of external script files on this page is 4, there is need to reduce this to 1 or 2. Re-factor and minify must be combined to optimize JavaScript files. Ideally University X should have one (or even embed scripts for high-traffic pages) on their web page. Placing external JavaScript files at the bottom of their BODY, and CSS files in the HEAD will enable progressive display in XHTML web pages. Combining JavaScript into one or eliminating them, and using HTTP compression was appropriate for any scripts placed in the HEAD of the document since it helps in reducing response time. As recommendations for further study researchers can also focus on different programming languages such as Delphi, Ruby and Java on how they affect web page load time for a distinct web page to be designed. There is need to research on how to promote websites through Search Engine Optimization (SEO) techniques adoption. In addition a further research on upgrading the normal desktop website version to a mobile site version will be a study of great value in relationship to response time.

References

[1] J. Eland, “Web Stores Do-It-Yourself for Dummies”. Wiley Publishing, Inc., Indianapolis, Indiana, pp 361, 2008.

[2] Zona Research Group, “The Need for Speed”. Research Report: Zona Research Group, 2003.

[3] Aberdeen Research, “Customer Satisfaction on the web”. Aberdeen Research Group Report, USA, 2010

[4] G. Pallis, A. Vakali, “Insight and Perspectives for Content Delivery Networks”. Communications of the ACM, Vol. 49, No. 1, ACM Press, NY, USA, pp. 101-106, 2006.

[5] A. Vakali, G. Pallis, “Content Delivery Networks: Status and Trends”. IEEE Internet Computing, IEEE Computer Society, pp. 68-74, 2003.

[6] G. Peng, “CDN: Content Distribution Network”. Technical Report TR-125, Experimental Computer Systems Lab, Department of Computer Science, State University of New York, Stony Brook, NY, 2003.

[7] G. Barish, K. Obraczka,”World Wide Web caching: Trends and techniques”.USC Information Sciences Institute, 2000.

[8] P. Killelea, “Web Performance Tuning”, 2nd Edition. O'Reilly & Associates, Inc. Sebastopol, CA, USA, 2002

[9] A., K. Manjhi, “Technical report submitted in partial fulfillment of the course CS625: Advanced Computer Networks”.UK, 2008

[10] NetForecast, “Field Guide to Application Delivery Systems Report”. NetForecast, 2006.

[11] Forrester Research and Gartner Group, “Two load time report surveys”. Forrester Research, Inc., 400 Technology Square, Cambridge, MA 02139, USA, 2009

Author Profile

Munyaradzi Zhou holds a Post Graduate Diploma in Tertiary education from Midlands State University, a Bachelor of Science in Information Systems Honours degree from the Midlands State University and a Master of Science degree in Information Systems Management from Midlands State University and is currently a Doctoral Student at University of South Africa (UNISA). He is a lecturer at the Midlands State University, in the Department of Computer Science and Information Systems.

Maxmillan Giyane holds a Post Graduate Diploma in Tertiary education from Midlands State University, a Bachelor of Science in Computer Science Honours

degree from the Midlands State University and a Master of Science degree in Information Systems Management from Midlands State University and is currently a Doctoral Student at University of South Africa (UNISA). He is a lecturer at the Midlands State University, in the Department of Computer Science and Information Systems.

Mutembedza Nyasha Amanda holds a Post Graduate Diploma in Tertiary education from Midlands State University, a Bachelor of Science in Information Systems Honours degree from the Midlands State University and a Master of Science degree in Information Systems Management from Midlands State University and is currently a Doctoral Student at University of South Africa (UNISA). She is a lecturer at the Midlands State University, in the Department of Computer Science and Information Systems.