



Elements of Efficient Websites

A White Paper
By Kevin Lennon
lakeareawebs.com



Elements of Efficient Websites

Table of Contents

Introduction	3
Websites and HTML.....	4
Why Phase Out HTML?	4
Translated and Generated HTML.....	5
Improving Web Site Effectiveness.....	6
Updating Web Page Markup.....	7
Fixing Inefficient HTML.....	8
Upgrading to CSS.....	9
Summary.....	10



Introduction

This white paper was written to share some of the experiences we have encountered in our quest to improve the internal operations of our development and client websites.

In this paper we speak of “elements” in the rather narrow sense of improving those factors that contribute to the handling of a web page. These factors include retrieving the web page from the server sending it to the browser where it is rendered and displayed.

In order to measure “efficiency” under these circumstances, we decided to assign internal weights to the importance of each factor.

We did this by examining the internal elements that make up the web page markup itself. This allowed us to evaluate the appropriateness, implementation, maintainability and other attributes that each factor contributed to the entire process.

To start the investigation we examined a small sample from about 1,000 randomly selected websites.

After analyzing these pages, we upgraded and validated the markup and scripts on the pages to be compliant with the latest recommendations.

We then made up a testing platform to exercise the original and the upgraded pages in order to time those elements

that had an effect on either response time or transfer time.

The pages of the comparison runs were then executed from the server. They were instrumented and delivered through a PHP /CGI interface. PHP scripts also controlled the iterations and supplied the HTML/XHTML markup. In addition, PHP scripts also generated the needed Javascript code on the pages to instrument readings on the client side.

While we recognize that this testing scenario has a number of scientific method shortcomings, it was very effective for our purposes. The method answered many of the questions we had about how to improve website performance and what it would cost.

We’ve found that upgrading the software and markup aspects of a website provides a significant reduction in the bandwidth resources needed to deliver and display these pages.

It also provides significant benefits in terms of maintainability, enhanced displays, and a number of other areas described in later sections.



Websites and HTML

Websites and browsers live on markup languages. The most popular on the internet at this time is the venerable Hyper Text Markup Language (HTML).

HTML is probably the largest single factor underlying the success of the internet as we know it today. However, by today's standards, the language has grown rather long of tooth.

Over the last 15 years of sustained service HTML has undergone 4 generations of revision and now has finally reached the point of being deprecated by the world standards organizations itself.

Since early 2000 the World Wide Web Consortium (W3C) has published two recommendations for extensible hypertext markup language (XHTML), the successor to HTML markup. These recommendations have been received well in the development community and XHTML is rapidly being adopted at this time.

Why Phase Out HTML?

The exponential growth period of the internet over the last decade has provided great incentive to both commercial and private interests to publish more attractive and distinctive web page presentations.

To meet these demands, web designers and browser developers began to use the HTML markup tags which were

really intended to control the page structure, in innovative ways that allowed them to enhance the presentation aspects of the data.

Using the language in this way introduced a great deal of overhead into the page markup. Soon it became common to see web pages whose markup was more verbose than the content itself.

The situation became exacerbated when the designers began to use the HTML table structures to position the content on the page. These structures often required deep nesting. This practice soon made reading and understanding the HTML document an almost impossible job.

These issues forced the web design community to recognize that the markup language must enforce division between structure (HTML) and presentation (Color, Font, Alignment...etc.) within the web page markup.

The solution to these problems began with Cascading Style Sheets (CSS). CSS greatly extended the presentation and positioning aspects of HTML, and left the structuring elements (e.g. Headings, Paragraphs etc.) intact within the existing HTML.

This close relationship of HTML and CSS preserved the existing web page base, and at the same time greatly extended the presentation capabilities.



Translated and Generated HTML

As the markup language was steadily being refined to increase the internet capability, the page content of millions of websites were being developed with seriously impaired markup generated by word processing and graphic web page generation software.

When people found out that they could use their word processor or a What You See Is What You Get screen to generate a web page, that technique became an instant success.

This could be done without any knowledge of HTML. This practice created a cottage industry of web pages on the internet that were produced without any regard to the quality of the underlying markup code.

Unfortunately, the markup quality from most of the pages created in this manner was almost indescribably poor!

While the impact of this poor implementation is mainly anecdotal at this time, it appears likely that many millions of those web sites are currently in operation on the internet.

While they consume a great deal of bandwidth because of their implementation, it is fortunate that, in general, these sites have a low traffic profile.

That is not to say that building a website in this way should be universally

condemned. The ability to create a pleasing web page without any deep understanding of how the page was created opened the door to many millions of personal and small business websites. Collectively this contributed greatly to the internet's popularity.

Nevertheless, we now recognize that the price of convenient web page generation software can result in a significant reduction in the internet resources that must be shared by all.

On the other hand, cleaning up this code and bringing it up to current standards will not only benefit everyone by reducing the bandwidth and maintenance requirements. It will provide extra benefits to the publisher as well.

These issues will be discussed in later sections of this paper.



Improving Web Site Effectiveness

In our view, there are three essential components that characterize the overall *effectiveness* of the web pages that pass between the internet websites and the client machines:

- ✚ The Content
- ✚ The Web Page Markup
- ✚ The Presentation

Content is the body, soul, and foundation of the web page. It, more than any other factor, determines the quality of the page. Content can be arranged and organized in many ways, but like a good storyteller, a professional copywriter can always make the page more effective.

Web Page Markups determine the physical aspects of the way the content is presented. In modern times this is usually done on the client side with HTML/XHTML markup, as well as other presentation tools such as Javascript and Cascading Style Sheets. Usually it takes a professional web developer to achieve the desired presentation aspects of the page with these tools.

The Presentation aspects of a web page are best handled by those design professionals who can maximize the aesthetics of attractive blending of the graphics, content, and presentation layout of the web page.

Trying to devise an effective strategy that will combine these components in a way that will optimize them all and

produce a stunning page, great response, and captivating content is a much more difficult task however.

From a practical point of view, having professionals contribute to every aspect of the web page is a luxury that only a small percentage of web sites can afford. As a consequence, most sites today are conceived, designed, and published by a single individual.

The fact also remains, that there is little that developers, professional or not, can do to improve the use of internet bandwidth by altering the page content. The content is the essence of the web page.

In a similar fashion, the presentation aspects of the page design need to maintain their artistic integrity to be effective. Little bandwidth can be saved by improvements in this area without being counter-productive to the overall design.

That leaves us with the real issue of just what we can do to improve the Web Page Markup, that will compliment and balance both the Presentation and Content aspects of the page.

What we can do in this area is discussed next.



Updating Web Page Markup

We have found the best strategy for enhancing website effectiveness is to begin with a close examination of the web pages that make up the site.

In our recent experiments with a little over a thousand randomly selected pages from the web, we found that the vast majority could be reduced significantly in size without any degradation of the presentation.

The bandwidth savings on some of these pages was increased by as much as 300% when the excessive and redundant markup was removed from these pages.

This size reduction was attributable to improving the markup and scripting embedded in the content. It did not include the savings that can be achieved by managing the image content of the page. It is not uncommon to see an identical .bmp image reduced to 1/10 of the original size by simply converting it to a .jpg form. Both techniques provide a real opportunity to deliver the same page content using much less bandwidth resources.

From a practical point of view however, it's a difficult task to convince a small business or a personal website owner that it's worthwhile for them to do this.

That's especially true when one considers that this area generally requires a high degree of technical

expertise to get the markup part of the job done. This does not include making any improvements to the Presentation and Content components.

The only tangible result that the user usually sees is that the site delivers the pages a fraction of a second more quickly.

On the other hand, doing this same kind of markup optimization on a high volume commercial website can provide significant increases in profitability while reducing the costs involved in maintaining the website.

Under the circumstances, it's encouraging that the internet can still operate well with the existing imperfections, without having to restrict the ways that many low volume personal and small business websites are published.

The internet can display inefficient pages without concern for the bandwidth consumption, primarily because their low traffic profiles have been anticipated.

At the same time significant benefits can accrue to the many high volume commercial website operations, by upgrading the web page markup to the latest international standards.

If this action is taken, then everyone on the internet will benefit.

We'll examine how that can happen next.



Fixing Inefficient HTML

In previous sections of this paper the importance of reducing the effective web page internet traffic by improving the structure and reducing the size of the pages was emphasized.

Nevertheless, it should be recognized that there are many other markup and layout factors that should be examined as well:

1. Presentation elements within the HTML
2. Bloated code generated by graphic and word processing presentation software.
3. Improper structure for Search Engine Analysis.
4. Exaggerated list of keywords included in metatags not appearing in content.
5. Use of tables for control of page layout.
6. Outdated versions of HTML.
7. Missing Doctype tags
8. Non-validated HTML
9. Not employing Cascading Style Sheets.

All of these have a negative impact on not only the size of the code required for

presentation, but also on maintainability, bandwidth consumption, search engine exposure, and browser rendering processes.

To address these inefficiencies in a web page requires a knowledgeable person who is fluent with HTML. That person must examine every line of the content and markup on the page and then replace any deficiencies with upgraded statements.

In terms of time, we've found that bringing up a typical 300 line static web page to the current standards requires about a half day of effort on the part of an experienced technician.

Upgrading dynamically generated web pages, in general, requires a more knowledgeable technician fluent both in the markup languages (e.g. CSS, HTML, XHTML etc.), client side Javascript, and the server side generation language employed (e.g. Perl, PHP, Visual Basic, etc.).

In our experience, the time required to upgrade server side dynamic pages is more than double that of a static page. This is primarily due to the fact that the underlying HTML pages must be re-designed even before the server side software is adapted to handle the new page.

The time estimates mentioned above include the time required to validate the code changes, but do not include verifying the page presentations on more than a single browser.



Upgrading to CSS

While the importance of including CSS in improving the performance of web pages was mentioned in several previous sections, it was not possible to convey the real importance this tool brings to the task of improving the overall website efficiency within in those contexts.

In addition to the previously mentioned merits of reducing the size of an HTML page (reducing the bandwidth requirements) CSS also provides improvements to many other areas of web site operations:

- ✚ CSS enhances the ability to make the web page more accessible to a wide variety of devices (e.g. handheld, printers, aural screen scans).
- ✚ Changing a single CSS file can broadcast the changes to all of the web pages on a site.
- ✚ CSS enhances the ability to produce identical displays on browsers that render the markup differently.
- ✚ CSS eliminates the need to use tables for positioning content in the page layout.
- ✚ CSS does not require all your cells to be the same height and width as is the case when you use tables for positioning.

By eliminating those tables in the markup used for positioning the content the website will be made more efficient in the following ways:

- ✚ CSS displays content much faster than that positioned in tables. That's because browsers need to pass through tables twice before displaying their contents; once to work out their structure and once to determine their content.
- ✚ Browsers cannot display parts of a table. The entire table must be loaded into the client browser before any part is displayed. CSS is rendered and displayed line by line, and is much faster.
- ✚ Without CSS, when the content layout is described in tables, the browser must read in the entire table description *on each page* to render the display. CSS can handle all of those pages with one reading of the description.

For the larger, high traffic websites on the network, upgrading to CSS will be a good investment for these reasons.



Summary

Our overview of the markup used in a large sampling of current internet pages indicates a majority use outdated markup languages, and include bloated and poorly written markup generated by word processing and graphic page generation software.

Eliminating these inefficiencies by re-writing the pages is very labor intensive. So much so that it may be difficult to cost justify undertaking this improvement on low volume websites.

However, it also finds significant economic and performance rewards can be realized for the larger, high traffic websites by bringing the current web page markup up to date with the latest recommendations of the World Wide Web Consortium (W3C) internet standards organization.