LYONSCG

eCommerce Realized!

HOW TO MEASURE SITE SPEED IN ECOMMERCE

Why Page Load Can be Misleading

Stephen Chinn
Director of Hosting Services
January 2015

INTRODUCTION: THE IMPORTANCE OF SITE SPEED FOR ECOMMERCE

Website speed in eCommerce is widely accepted as one of the most important factors in driving revenue. So, it's only common sense that if shoppers can't easily and quickly navigate through a website, search for products, add them to a shopping cart, and complete the transaction, the likelihood of a sale goes down.

Keeping customers happy means providing a satisfying user experience. Just think of your own online shopping behavior. Since speed can make or break a sale, let's take a look at some statistics on web performance for eCommerce sites.

- According to a case study from Radware[1], 51 percent of online shoppers in the U.S claimed if a site is too slow they will not complete a purchase.
- Radware[2] also discovered in another study that the demand for loading speed has increased over time. For example, in 2010 a page that took six seconds to load witnessed a 40 percent conversion hit. In 2014, pages with that same loading time suffered 50% drop in conversions.
- During peak traffic times, 75 percent consumers[3] are willing to visit competitor sites instead of dealing with a slow loading page.

Clearly speed impacts conversion rates and site traffic. However, what it takes to ensure website speed is often misunderstood. In the first part of this paper, we'll take a look at some of the typical metrics for site speed, why those metrics can give you a false sense of security, and alternative ways to measure site speed that are more meaningful.

Site speed is a function not only of web design but also of your hosting infrastructure. Because we offer a complete suite of eCommerce services including hosting, we have a holistic view of how they work together to drive revenue growth. Thus, in the second half of this paper, we'll dig into tactics — the technical and design elements you should focus on to make sure your site works as fast as it can.

Illustration: Digital Trust's Lakeside Technology Center in Chicago where LYONSCG our dedicated hosting facility

_____

[1] Search Engine Journal, http://www.searchenginejournal.com/seo-101-important-site-speed-2014/111924/

[2] Ibid

[3] Ibid

LYONSCG

THE IMPORTANCE OF SPEED

Why does page speed (or load-speed, or time-to-load or any number of interchangeable terms used to describe the speed of your eCommerce website) matter? Faster page speeds are positively correlated with two major benefits, both of which generally lead to more eCommerce revenue: First, faster websites have a *higher conversion rate*, meaning more of the people who come to the store make a purchase. And second, *faster websites increase site traffic*, meaning that a faster site attracts more people into the store — although this may not be for the reason you think. While site speed has some small technical influence with respect to Search Engine Optimization efforts, providing a *satisfying user experience* causes people to return to the site more frequently (as well as link to it and share with their friends and family).

SPEED & INCREASED CONVERSION: SOME PERSPECTIVE

There's a lot of people selling speed (so to speak) but when you drill through the hype, you find that much of what you read is exaggerated or taken out of context. For example, a well-known study conducted by Walmart in 2012 showed that for each 1-second improvement in page load time, conversion rate increased by 2%.[1] That report has been widely repeated, but taken too literally. When you dig deep into the data, you see that the increase in conversions tapers off sharply after reaching 5 seconds. Yes, there is improvement for improving speed, but there's a difference between a 1 second improvement on a baseline of 3-seconds versus a similar improvement from a baseline of 10-seconds.

Another classic example of misleading references to website speed comes in the form of Greg Linden's famous 100ms Amazon quote. In 2006, in a presentation he gave at Stanford University on data mining, Greg shared the assertion that 100ms of increased latency results in a 1% decrease in sales[2]. However, in 2010 Greg was asked to clarify his statement, he said **if your page already takes 2 seconds to load, +-200 ms more isn't going to matter.** Unfortunately, some continue to take the 100ms/-1% sales tradeoff as being as solid as Moore's Law[3].

Even today, eight years after Linden's original statement, and four years after his clarification, you can still find dozens of websites with references to the effect that "Experiments at Amazon show that **every** 100ms increase in the page load time decreased sales by 1 percent." It's not true, and it's not what Linden says.

As these two examples of statistics show, you can't believe everything you read. Sometimes, to get to the truth, you need to dig deeper into the data.

---

[1] http://www.slideshare.net/devonauerswald/walmart-pagespeedslide

[2] http://glinden.blogspot.com/2006/12/slides-from-my-talk-at-stanford.html

[3] The observation that, over the history of computing hardware, the number of transistors in a dense integrated circuit doubles approximately every two years.

LYONSCG

CONVERSION RATE: IMPACT OF RESPONSE DELAY

Serious research (vs. anecdotal evidence) tell us that there **are** fundamental rules that govern the relationship between website speed, conversion rates, and eCommerce revenue. Early findings by Zona Research, way back when most people accessed the Internet with a 56K Modem and sites were far less complicated than they are now, suggested the average web user would wait about eight seconds for a page to download. Referred to as "The 8 Second Rule", it found that if load time of a web page exceed eight seconds, users were unlikely to wait.[4]

Even today, with much faster connection speeds and user expectations of richer content, "The 8 Second Rule" holds true[4]. People still run out of patience at around eight seconds; it's human nature. Here's what research tell us:

- Between 0 and 1 seconds – users are perfectly happy
- Between 1 and 2 seconds – users are quite happy – improved speed will have some impact but it'll be small
- Between 2 and 4 seconds – the user experience is significantly improved as you move closer to 2 seconds
- Over 4 seconds – you've lost the opportunity to impress people with speed. After that, both attitudes and intensions remain fairly flat till you hit around 6 seconds.
- Between 6 and 8 seconds – dissatisfaction with the experience starts climbing fast
- Over 8 seconds – you're seriously losing people

If your website takes more than 10 seconds to load, even when shoppers really want your products, their attention can wander. At this point, you're going to lose even the most avid customers. And while other studies may quibble over the specific impact of the time when shopper decay starts, the pattern is similar. These results have been validated many times over the years.

Other factors can influence tolerance for slow loading sites. User familiarity has a positive impact on buyer behavior because they'll have more patience with sites they know. However, design or operational changes can frustrate even a loyal visitor. To prevent a reduction in shopper familiarity when changes are made to a site, companies need to incorporate speed improvements.

---

[4] "Video Stream Quality Impacts Viewer Behavior: Inferring Causality Using Quasi-Experimental Designs",
http://people.cs.umass.edu/~ramesh/Site/HOME_files/imc208-krishnan.pdf

LYONSCG

## CONVERSION RATE: THE SPEED AND RICH EXPERIENCE DILEMMA

A customer attention range from 2 – 8 seconds is a large window. So, what kind of response time really matters?

Let's start with the premise that website shoppers want to feel that they're in control of the experience. When they do something — visit a webpage, perform a search, and the like — they want rapid feedback. It's as fundamental as this: shoppers want to know that their click worked. Of course, they also want lots of information; and as a merchant you want to give them lots of inspiration to buy your products. For a basic product search or view, that inspiration comes from lots of information, *and* elegant presentation. Multiple images, zoom-in graphics, perhaps even 360° rotations are the kind of rich content that drives a rich user experience. However, all that content on the page inherently takes longer to load.

The dilemma is this: you need to provide both **fast response** and a **rich user experience.** How do you accomplish both? The answer is to set aside an overall page-load as the way to measure site speed, and instead focus on time-to-interact.

CONVERSION RATE: TIME-TO-INTERACT IN CONTEXT

So, what's time-to-interact? Let's compare it to other common ways to measure site speed. Remember to keep the visitor in mind when thinking about meaningful figures.
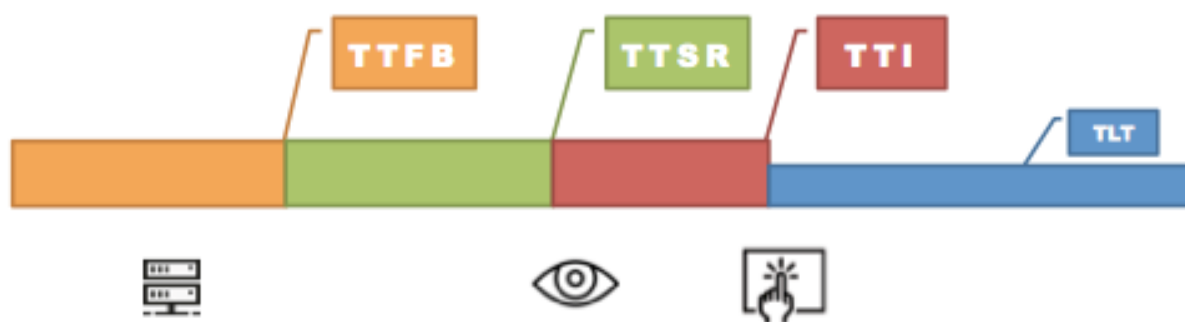

- **Time-to-First-Byte** (TTFB) is a useful data point because it indicates how long it takes the server to create the response page plus the time it takes to transfer the page from the server to the client.
- **Time-to-Start-Render** (TTSR) is important because it's the first point in time the user sees something displayed on the screen; this feedback reassures them their click actually did something.
- **Time-to-Interact** (TTI) is the point at which the user can take their next action such as clicking on a button.
- **Total-Load Time** (TLT) is when everything for the page has completely downloaded and been processed by the browser.

LYONSCG

CONVERSION RATE: TIME-TO-INTERACT TESTING

Web Page Test[5], an open source project, primarily developed and supported by Google, is the tool Radware uses to collect test data for their annual speed and web performance report. Their most recent report[6] shows time-to-interact and time-to-load for the top 12 web stores in the Alexa Retail 500. Measured by the highest combination of unique visitors and page views, the graph below shows the 10 most popular web stores, including Amazon, Kmart, AE.com, and Sears. TTI is depicted in gray; TTL in orange.

What's interesting is that all of eCommerce sites have time-to-interact times considerably shorter than their time-to-load times. Some have an amazingly high time-to-load. It's worth noting that all 10 sites in the Radware report have time-to-interact times less than three seconds.

You can see the home page for AE.com took 19.8 seconds to finish loading completely; but they're still one of the most successful stores (#8) on the Internet. How can that be, if shoppers begin to lose attention when time-to-load exceeds eight seconds? The answer is that time-to-interact is more important than time-to-load. In this case, TTI is 2.1 seconds.
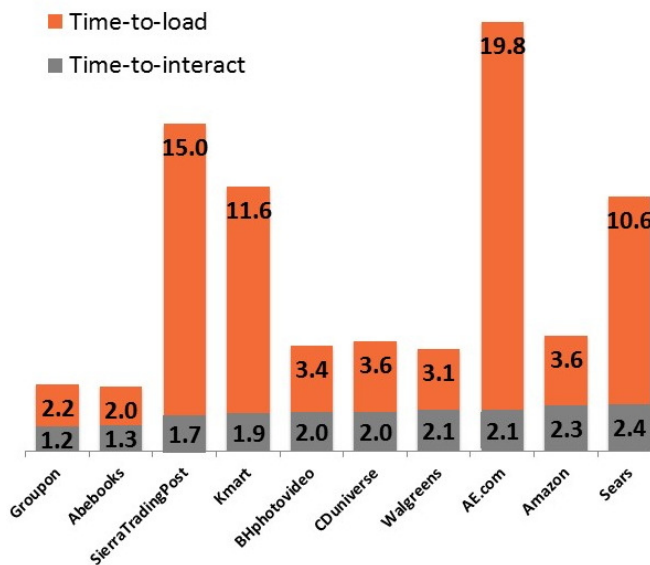
10 Fast eCommerce Sites: Time-to-interact

This significance is that so much focus on page load times is mislaid. What really matters, in terms of eCommerce success, is the customer experience — as measured by time-to-interact.

---

[5] http://www.webpagetest.org

[6] http://blog.radware.com/applicationdelivery/applicationaccelerationoptimization/2014/10/sotu-for-ecommerce-page-speed-web-performance-fall-2014/

**LYONSCG**

CONVERSION RATE: TIME-TO-INTERACT IN PRACTICE

Let's take a look at time-to-interact on a practical basis. The screen shot below is for the Tom Petty and the Heartbreakers site where fans can purchase concert tickets and merchandise.

What you see in this image is the moment the shopper reaches the time-to-interact. The key call-to-action controls are usable, which means the shopper can check out at this point — the ultimate goal for an eCommerce site. Alternatively, they could also enter search term, click on a menu item, visit another page or take other actions. The shopper has received positive feedback that their action to reach this page was successful, and control over next steps.
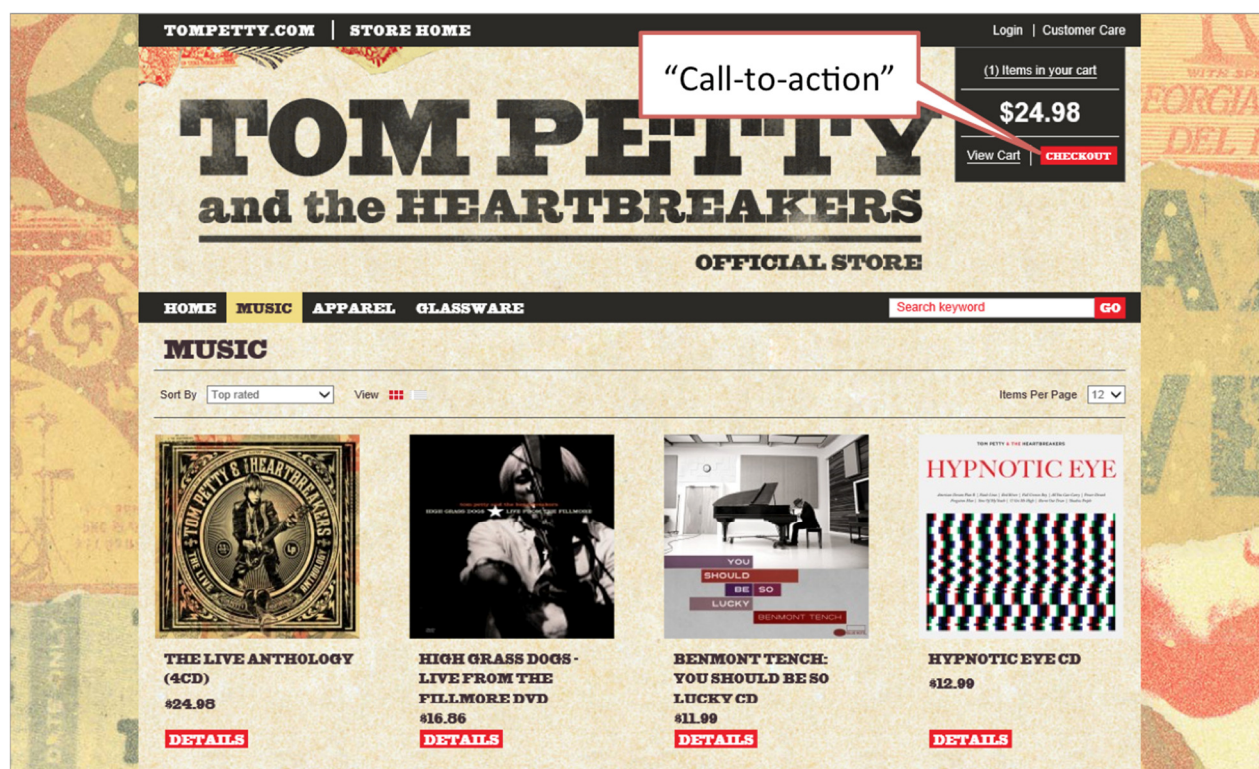
Let's look at the results from using a customized version of Web Page Test across multiple contexts and perspectives. We're able to obtain very precise data on the platforms being tested – not just different browsers, but different hardware configurations.

LYONSCG

Web Page Test provides snapshots of the page as it loads. If the website is built with best design and coding practices, the time-to-interact will quickly follow time-to-start-render. By looking at these snapshots, we can see when the key "call-to-action" controls become available. This marks the time-to-interact; the point at which users can actually do something on the site.
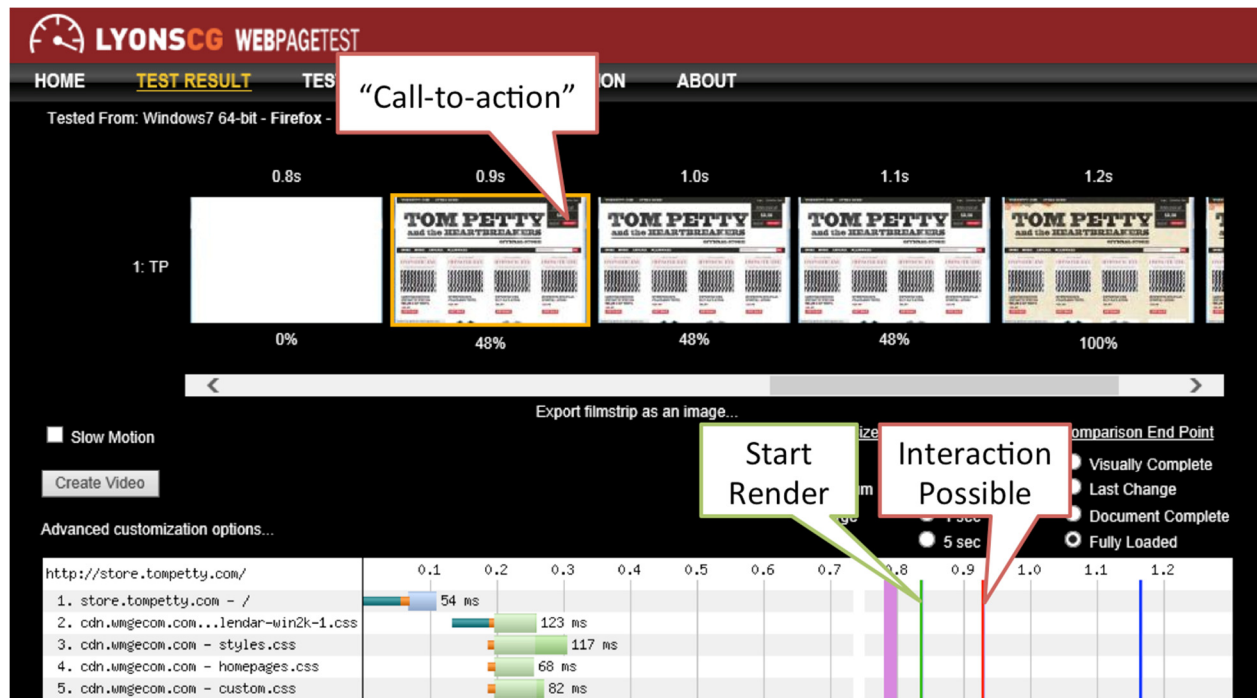
Keep in mind that sometimes it's not always a simple matter to determine when we've reached a useable call-to-action that would meet the definition of time-to-interact. We have to know that the control we see is in fact useable. If a JavaScript file still needs to load before the button will work, then we can't say we've reached the time-to-interact point.

Details show the loading sequence for page elements which can have a profound impact on time-to-interact. For example, background images are loaded by a content style sheet **after** the page has been sent to the browser; **and after** the call-to-action button has been rendered. If the shopper is ready to buy and wants to click on that checkout button, they don't have to wait for the background images to load. Had we built this page to load the background first, it would add precious microseconds to the time between the start of rendering and the time-to-interact (CTA button).

There's lot more that can be done to increase time-to-interact. First, let's look at other Google Analytics and instances of other speed metrics that don't always "measure up."

LYONSCG

SITE TRAFFIC: SEARCH ENGINE MEASUREMENTS

In 2010, Google officially announced that they'd made site speed a ranking factor in their algorithm; and ever since, Google has been strongly advocating for fast web pages. Again, this needs to be put into context. Google and the other search engines have a vested interest in being able to quickly read page content, so their advocacy of speed isn't entirely altruistic. However, site speed is just one of 200 factors considered in their algorithm; and Google has said it will have no impact at all on 99% of sites.

In their quest for a faster Internet, Google Analytics lists some metrics that can be a little misleading. For example, Google Page Speed Score does not give you a measurement of your relative page speed. Instead, it measures how many of Google's suggested page rules the merchant is following. For the most part, these are sensible rules, but there are circumstances when you will want to ignore them. The Tom Petty store referenced previously gets a lowly 43 out of 100. However, it has a time-to-interact of less than one-second and is fully loaded in about 1.3 second. It's fair to say that the Google Speed score does not accurately reflect the consumer experience!

Another example of a potentially misleading metric is Average Document Interactive Time[7], which sounds very much like time-to-interact. It's not. It measures the point at which the Document Object Model is ready to interact. The customer will see certain objects such as the call-to-action button. The Google search bot may report a fast page speed. However, on a poorly designed site, a shopper will still be cooling their heels, waiting to do anything useful.

It's also important to know that Google's default setting for this metric uses only 1% of samples in its calculation – that makes it far less than 1% of the actual page requests. That's much too small a sample size to be accurate. On sites with relatively low traffic it will produce meaningless nonsense. For this metric to be useful, it would need to be bumped-up the sampling rate, perhaps to 10%.

When it comes to how to measure site speed, however, focus your efforts on time-to-interact. Get that right and the rest fall into place.

---

[7] https://support.google.com/analytics/answer/2383341

**LYONSCG**

THE TACTICS OF SPEED

Now that you understand the importance of time-to-interact as in indicator of speed and user experience, we can discuss three general tactics to help you improve it:

- Caching
- Compression
- Execution

CACHING: OPTIMIZE FOR ANTICIPATED USE

Caching is a way to store data so future requests for that data can be served faster. This example shows a common scenario of product information stored in a database. It's very flexible in how it can be accessed and used. If you need that data for a web page, you can query it, process it, and turn it into a block of HTML code. Once the data is made into an HTML block, it can be made into a web page; but cannot be passed to other applications like inventory or order management. This presents a dilemma. You can store raw data, and incur a delay when it needs to be rendered for a web page. Or, you can store it as HTML, which makes it easier to use in other applications.

Caching provides an alternative approach. Let's say 20 people can be expected to request a web page that uses the same data. Then storing (caching) the data as an HTML block *in addition to* storing it as raw data becomes very useful.
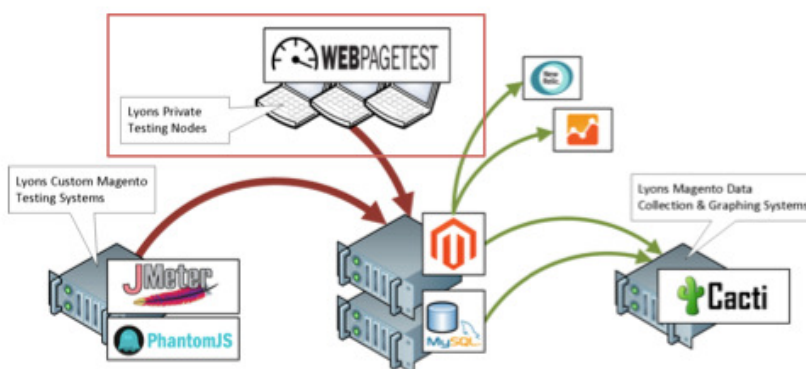
If you cache the data as an HTML block, you'll only need to pay the cost of creating that block once. You save a lot of processing power and at least 19 people will get their requested web page much faster. If you don't cache the block, you'll spend processing time and power to transform the data into HTML every time you get one of those 20 page requests.

In another example of how caching can increase speed, a customer sends a page request, and, once the HTML page has been built, it gets sent to them. If settings allow it, the page will be stored in a browser cache on the customer's computer. If they wants to look at the page again, they can get it from their local browser cache and avoid the time and cost of transporting it from the server again.

But here's the rub. Let's say you update a product attribute. Now the data in the database has changed. All the versions of the data, stored in cached forms, throughout the system, will be wrong. The old HTML block needs to be rebuilt so that it accurately reflects the database.

LYONSCG

CACHING AND FREQUENCY OF CHANGE

Thus, frequency of change is a critical factor for caching. If data changes every few seconds, caching isn't very useful. If your data changes once a month, it can be pre-processed; put it in cache; and served multiple times, providing fantastic speed for your customers.

So, let's look at some of the many levels of caching. The database, the place to start, contains the platform's **system of record**, the data on which everything else builds on. On the database server, there are caching mechanisms, like specialized indexes, de-normalized data in flat table, and results from recent queries in the query cache. On the application servers there are the application, session and page caches. And in many systems there's a proxy-caching layer which serves as an intermediary for requests from clients seeking resources from other servers. Sophisticated use of a proxy-caching layer can add to ongoing development cost, but you get incredibly fast time-to-first-byte times.

It's important to consider the data's frequency-of-change as it relates to the browser cache as well. The frequency for refreshing data within the browser cache can be handled by setting a file expiration time. eCommerce store owners often set this to zero (or just a few minutes) if they think the information might change before the shopper returns. Of course, this approach minimizes the value of the browser cache for improving website speed. So it makes sense to separate data that will change frequently from data that won't, so you can use the browser cache to greater advantage.

Content Data Networks, or CDNs, are distributed content caches which can be very powerful. However, most CDNs don't respond to requests any faster than most eCommerce web sites. The primary value of a CDN comes from the pre-positioning of assets closer to where they're going to be used, which eliminates network latency time. If, however, you're using a good hosting facility and, most of your customers are in the same region as that facility, moving assets to a CDN can actually slow performance. Like any other caching mechanism, CDNs can be extremely useful, if they're used thoughtfully.

LYONSCG

IMAGE COMPRESSION: SMALLER IS FASTER

The Radware report referenced earlier concluded that image compression represents one of the single greatest performance challenges. Our experience readily validates that conclusion. The problem is that even when done perfectly for a new site launch, image compression can degrade over time. Often, the staff responsible for uploads of new images isn't always as rigorous about compressing them as they should be. So it's worth reviewing image compression on a regular basis. There are four types of image compression to consider:

- **Progressive images:** Creative people don't like the idea of blurry images, even for a second, as they load. If the goal is to get something to the customer as quickly as possible, progressive images do work. While there's a delay, the trade-off is you can get away with a larger file size; thus the image will look nicer when it's done.
- **Sprites**: An individually animated element in a larger graphic image or set of images, Sprites are very efficient. But they're also time consuming to maintain. Sprites are best used for something you aren't going to change a lot (in that sense, it's like caching).
- **Server compression:** It's rare to find a site which hasn't ben compressed to improve transfer speed and bandwidth utilization. If you're evaluating your site for some quick speed improvements, reviewing server
- **Minification and consolidation:** These tactics, reduce file sizes and connection costs, can have some impact but that's largely limited to the first page request in a user session. In most cases, browser caching will eliminate these costs after the first page hit. Having said that, the first page a customer receives is also their first impression, so, even if it's just one page request, it's worth considering.

With each of these tactics, employing them needs to be considered within a cost vs. benefit framework. That, in turn, needs to reference your business objectives and the experience you want your customers to have.

EXECUTION: REMOVE UNNECESSARY WAITING

Execution refers to how fast a computer performs instructions which control a sequence of actions. There are three ways to impact page load and site speed.

- **Deferred Execution:** The idea is to get the customer what they want as quickly as possible. Deferred execution deals with organizing how things are loaded. We saw some of the benefits of deferred execution in the Tom Petty example. Loading the background images was deferred until after the key "call to action" controls were made useable.

- **Asynchronous Execution:** The problem addressed here is what happens when you create a synchronous connection to a third-party service to deliver content and the communications path has a problem or gets very busy. The process, such as delivering a web page, will stop and wait until the system receives the "Ok got it" message from the third-party provider. Since you don't have control of the third party service, you can pretty much guarantee, at some point, this will be a problem. Therefore, it's far better to use an asynchronous connection which doesn't require the "Ok got it" message, so there's never a delay in serving up the page for the customer. Better yet, you can batch content delivery and separate it from the page service process entirely.

LYONSCG

- **Separated Execution or Hole Punching**: You can significantly improve the value of caching through hole punching, a computer networking technique for establishing communications between two parties in separate organizations who are both behind restrictive firewalls[8]. It enables you to have different caching strategies for different parts of the web page. Areas that will be the same for everyone and rarely change (such as headers & footers) will always receive cached content. Not only it does happen fast, it won't take much power to deliver. Areas with content unique to a particular user or that changes a lot (search results, personalized content), can be cached differently or not at all.

On the server side this might be a Magento content block, excluded from Full Page Caching, or a Varnish template, using Edge Side Includes. On the client side, AJAX can be used to make page areas with entirely separate web requests. Hole punching is enormously important because it facilitates the most advantageous caching strategies. And good caching improves website speed.

CONCLUSIONS

Site speed in eCommerce is widely accepted as one of the most important factors in driving revenue. That's because speed has a major impact on conversion rates and site traffic.

Unfortunately, many eCommerce merchants don't understand how to measure site speed. They place too much emphasis on page load while not paying enough attention to what makes for a satisfying user experience. The key metric for user experience is time-to-interact, which is the point at which a customer can take action. Providing a *satisfying user experience* increases the chance of making a sale and gives customers a reason to return often to shop as well as tell friends and family about your site.

So how can you improve time-to-interact without sacrificing the quality of the customer experience? Three general tactics you can employ to optimize speed:

- **Caching,** a way to store data so future requests, so pages can be served faster.
- **Compression** of images whereby making them smaller increases the speed by which they're downloaded.
- **Execution** which controls the order in which tasks are performed, placing an emphasis on those which are key to the shopping process and reducing unnecessary wait time.

---

[8] http://en.wikipedia.org/wiki/Hole_punching

LYONSCG

Thank You

LYONSCG appreciates the opportunity to present our solutions for eCommerce to you. We believe that our extensive experience with strategy, creative services and technical implementation of solutions for many leading eCommerce sites will translate into a solution for you that will enable desired growth. We would be happy to review some current live sites we've worked on with you, and provide a detailed eCommerce demonstration.

About LYONSCG

LYONSCG is the industry's premier eCommerce Digital Agency, serving brand, retail, and B2B organizations with tailored eCommerce solutions that maximize online potential. Headquartered in Chicago, the firm offers a full range of services beginning with Digital Strategy and Digital Marketing and extending through Experience Design, Platform Implementation, Application Development, Hosting and Support. The approach is holistic — to provide every client with a creative, robust and increasingly profitable eCommerce website. LYONSCG is eCommerce Realized!

LYONSCG

20 N. Wacker Drive, Suite 1750
Chicago, IL 60606
P: 312.506.2020
F: 312.506.2022
W: www.lyonscg.com

LYONSCG