

更快速网站 (影印版)



Even Faster Web Sites

O'REILLY®

東南大學出版社

Steve Souders 著

更快速网站 (影印版)

Even Faster Web Sites

Steve Souders

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Sebastopol • Taipei • Tokyo

O'Reilly Media, Inc. 授权东南大学出版社出版

东南大学出版社

图书在版编目 (CIP) 数据

更快速网站: 英文 / (美) 苏德尔 (Souders, S.) 著.
影印本. —南京: 东南大学出版社, 2010.1

书名原文: Even Faster Websites

ISBN 978-7-5641-1934-8

I. 更… II. 苏… III. 网络—开发—英文 IV. TP393.092

中国版本图书馆 CIP 数据核字 (2009) 第 205637 号

江苏省版权局著作权合同登记

图字: 10-2009-250 号

©2009 by O'Reilly Media, Inc.

Reprint of the English Edition, jointly published by O'Reilly Media, Inc. and Southeast University Press, 2009. Authorized reprint of the original English edition, 2009 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由 O'Reilly Media, Inc. 出版 2009。

英文影印版由东南大学出版社出版 2009。此影印版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc. 的许可。

版权所有, 未得书面许可, 本书的任何部分和全部不得以任何形式重制。

更快速网站 (影印版)

出版发行: 东南大学出版社

地 址: 南京四牌楼 2 号 邮编: 210096

出 版 人: 江 汉

网 址: <http://press.seu.edu.cn>

电子邮件: press@seu.edu.cn

印 刷: 扬中市印刷有限公司

开 本: 787 毫米 × 980 毫米 16 开本

印 张: 16 印张

字 数: 269 千字

版 次: 2010 年 1 月第 1 版

印 次: 2010 年 1 月第 1 次印刷

书 号: ISBN 978-7-5641-1934-8

印 数: 1~1600 册

定 价: 42.00 元 (册)

本社图书若有印装质量问题, 请直接与读者服务部联系。电话 (传真): 025-83792328

更快速网站 (影印版)

Even Faster Web Sites

Credits

Even Faster Web Sites contains six chapters contributed by the following authors.

Dion Almaer is the cofounder of Ajaxian.com, the leading source of the Ajax community. For his day job, Dion coleads a new group at Mozilla focusing on developer tools for the Web, something he has been passionate about doing for years. He is excited for the opportunity, and he gets to work with Ben Galbraith, his partner in crime on Ajaxian and now at Mozilla. Dion has been writing web applications since Gopher, has been fortunate enough to speak around the world, has published many articles and a book, and, of course, covers life, the universe, and everything else on his blog at <http://almaer.com/blog>.

Douglas Crockford was born in the wilds of Minnesota, but left when he was only six months old because it was just too damn cold. He turned his back on a promising career in television when he discovered computers. He has worked in learning systems, small business systems, office automation, games, interactive music, multimedia, location-based entertainment, social systems, and programming languages. He is the inventor of Tilton, the ugliest programming language that was not specifically designed to be an ugly programming language. He is best known for having discovered that there are good parts in JavaScript. This was an important and unexpected discovery. He discovered the JSON (JavaScript Object Notation) data interchange format (<http://www.json.org/>). He is currently working on making the Web a secure and reliable software-delivery platform. He has his work cut out for him.

Ben Galbraith is the codirector of developer tools at Mozilla and the cofounder of Ajaxian.com. Ben has long juggled interests in both business and tech, having written his first computer program at 6 years old, started his first business at 10, and entered the IT workforce at 12. He has delivered hundreds of technical presentations worldwide, produced several technical conferences, and coauthored more than a half-dozen books. He has enjoyed a variety of business and technical roles throughout his career, including CEO, CIO, CTO, and Chief Software Architect roles in medical, publishing, media, manufacturing, advertising, and software industries. He lives in Palo Alto, California with his wife and five children.

Tony Gentilcore is a software engineer at Google. There, he has helped make the Google home and search results pages lightning fast. He finds that the days seem to fly by while writing web performance tools and techniques. Tony is also the creator of the popular Firefox extension, Fasterfox.

Dylan Schiemann is CEO of SitePen and cofounder of the Dojo Toolkit, an open source JavaScript toolkit for rapidly building web sites and applications, and is an expert in the technologies and opportunities of the Open Web. Under his guidance, SitePen has grown from a small development firm to a leading provider of inventive tools, skilled software engineers, knowledgeable consulting services, and top-notch training and advice. Dylan's commitment to R&D has enabled SitePen to be a major contributor to and creator of pioneering open source web development toolkits and frameworks such as Dojo, cometD, Direct Web Remoting (DWR), and Persevere. Prior to SitePen, Dylan developed web applications for companies such as Renkoo, Informatica, Security FrameWorks, and Vizional Technologies. He is a cofounder of Comet Daily, LLC, a board member at Dojo Foundation, and a member of the advisory board at Aptana. Dylan earned his master's in physical chemistry from UCLA and his B.A. in mathematics from Whittier College.

Stoyan Stefanov is a Yahoo! frontend developer, focusing on web application performance. He is also the architect of the performance extension YSlow 2.0 and cocreator of the Smush.it image optimization tool. Stoyan is a speaker, book author (*Object-Oriented JavaScript* from Packt Publishing), and blogger at <http://phpied.com>, <http://jspatterns.com>, and YUIblog.

Nicole Sullivan is an evangelist, frontend performance consultant, and CSS Ninja. She started the Object-Oriented CSS open source project, which answers the question, How do you scale CSS for millions of visitors or thousands of pages? She also consulted with the W3C for their beta redesign, and she is the cocreator of Smush.it, an image optimization service in the cloud. She is passionate about CSS, web standards, and scalable frontend architecture for large commercial websites. Nicole speaks about performance at conferences around the world, most recently at The Ajax Experience, ParisWeb, and Web Directions North. She blogs at <http://stubbornella.org>.

Nicholas C. Zakas is the author of *Professional JavaScript for Web Developers*, Second Edition (Wrox) and coauthor of *Professional Ajax*, Second Edition (Wrox). Nicholas is principal frontend engineer for the Yahoo! home page and is also a contributor to the Yahoo! User Interface (YUI) library. He blogs regularly at his site, <http://www.nczonline.net>.

Preface

Vigilant: alertly watchful, especially to avoid danger

Anyone browsing this book—or its predecessor, *High Performance Web Sites*—understands the dangers of a slow web site: frustrated users, negative brand perception, increased operating expenses, and loss of revenue. We have to constantly work to make our web sites faster. As we make progress, we also lose ground. We have to be alert for the impact of each bug fix, new feature, and system upgrade on our web site’s speed. We have to be watchful, or the performance improvements made today can easily be lost tomorrow. We have to be vigilant.

Vigil: watch kept on a festival eve

According to the Latin root of *vigil*, our watch ends with celebration. Web sites can indeed be faster—dramatically so—and we can celebrate the outcome of our care and attention. It’s true! Making web sites faster is attainable. Some of the world’s most popular web sites have reduced their load times by 60% using the techniques described in this book. Smaller web properties benefit as well. Ultimately, users benefit.

Vigilante: a self-appointed doer of justice

It’s up to us as developers to guard our users’ interests. At your site, evangelize performance. Implement these techniques. Share this book with a coworker. Fight for a faster user experience. If your company doesn’t have someone focused on performance, appoint yourself to that role. *Performance vigilante*—I like the sound of that.

How This Book Is Organized

This book is a follow-up to my first book, *High Performance Web Sites* (O’Reilly). In that book, I lay out 14 rules for better web performance:

- Rule 1: Make Fewer HTTP Requests
- Rule 2: Use a Content Delivery Network
- Rule 3: Add an Expires Header
- Rule 4: Gzip Components

- Rule 5: Put Stylesheets at the Top
- Rule 6: Put Scripts at the Bottom
- Rule 7: Avoid CSS Expressions
- Rule 8: Make JavaScript and CSS External
- Rule 9: Reduce DNS Lookups
- Rule 10: Minify JavaScript
- Rule 11: Avoid Redirects
- Rule 12: Remove Duplicate Scripts
- Rule 13: Configure ETags
- Rule 14: Make Ajax Cacheable

I call them “rules” because there is little ambiguity about their adoption. Consider these statistics for the top 10 U.S. web sites* for March 2007:

- Two sites used CSS sprites.
- 26% of resources had a future **Expires** header.
- Five sites compressed their HTML, JavaScript, and CSS.
- Four sites minified their JavaScript.

The same statistics for April 2009 show that these rules are gaining traction:

- Nine sites use CSS sprites.
- 93% of resources have a future **Expires** header.
- Ten sites compress their HTML, JavaScript, and CSS.
- Nine sites minify their JavaScript.

The rules from *High Performance Web Sites* still apply and are where most web companies should start. Progress is being made, but there’s still more work to be done on this initial set of rules.

But the Web isn’t standing still, waiting for us to catch up. Although the 14 rules from *High Performance Web Sites* still apply, the growth in web page content and Web 2.0 applications introduces a new set of performance challenges. *Even Faster Web Sites* provides the best practices needed by developers to make these next-generation web sites faster.

The chapters in this book are organized into three areas: JavaScript performance (Chapters 1–7), network performance (Chapters 8–12), and browser performance (Chapters 13 and 14). A roundup of the best tools for analyzing performance comes in the Appendix.

* AOL, eBay, Facebook, Google Search, Live Search, MSN.com, MySpace, Wikipedia, Yahoo!, and YouTube, according to Alexa.

Six of the chapters were written by contributing authors:

- Chapter 1, *Understanding Ajax Performance*, by Douglas Crockford
- Chapter 2, *Creating Responsive Web Applications*, by Ben Galbraith and Dion Almaer
- Chapter 7, *Writing Efficient JavaScript*, by Nicholas C. Zakas
- Chapter 8, *Scaling with Comet*, by Dylan Schiemann
- Chapter 9, *Going Beyond Gzipping*, by Tony Gentilcore
- Chapter 10, *Optimizing Images*, by Stoyan Stefanov and Nicole Sullivan

These authors are experts in each of these areas. I wanted you to hear from them directly, in their own voices. To help identify these chapters, the name(s) of the contributing author(s) are on the chapter's opening page.

JavaScript Performance

In my work analyzing today's web sites, I consistently see that JavaScript is the key to better-performing web applications, so I've started the book with these chapters.

Douglas Crockford wrote Chapter 1, *Understanding Ajax Performance*. Doug describes how Ajax changes the way browsers and servers interact, and how web developers need to understand this new relationship to properly identify opportunities for improving performance.

Chapter 2, *Creating Responsive Web Applications*, by Ben Galbraith and Dion Almaer, ties JavaScript performance back to what really matters: the user experience. Today's web applications invoke complex functions at the click of a button and must be evaluated on the basis of what they're forcing the browser to do. The web applications that succeed will be written by developers who understand the effects of their code on response time.

I wrote the next four chapters. They focus on the mechanics of JavaScript—the best way to package it and load it, and where to insert it in your pages. Chapter 3, *Splitting the Initial Payload*, describes the situation facing many web applications today: a huge JavaScript download at the beginning of the page that blocks rendering as well as further downloads. The key is to break apart this monolithic JavaScript for more efficient loading.

Chapters 4 and 5 go together. In today's most popular browsers, external scripts block everything else in the page. Chapter 4, *Loading Scripts Without Blocking*, explains how to avoid these pitfalls when loading external scripts. Loading scripts asynchronously presents a challenge when inlined code depends on them. Luckily, there are several techniques for coupling inlined code with the asynchronous scripts on which they depend. These techniques are presented in Chapter 5, *Coupling Asynchronous Scripts*.

Chapter 6, *Positioning Inline Scripts*, presents performance best practices that apply to inline scripts, especially the impact they have on blocking parallel downloads.

I think of Chapter 7, *Writing Efficient JavaScript*, written by Nicholas C. Zakas, as the complement to Doug's chapter (Chapter 1). Whereas Doug describes the Ajax landscape, Nicholas zooms in on several specific techniques for speeding up JavaScript.

Network Performance

Web applications aren't desktop applications—they have to be downloaded over the Internet each time they are used. The adoption of Ajax has resulted in a new style of data communication between servers and clients. Some of the biggest opportunities for growth in the web industry are in emerging markets where Internet connectivity is a challenge, to put it mildly. All of these factors highlight the need for improved network performance.

In Chapter 8, *Scaling with Comet*, Dylan Schiemann describes an architecture that goes beyond Ajax to provide high-volume, low-latency communication for real-time applications such as chat and document collaboration.

Chapter 9, *Going Beyond Gzipping*, describes how turning on compression isn't enough to guarantee optimal delivery of your web site's content. Tony Gentilcore reveals a little-known phenomenon that severely hinders the network performance of 15% of the world's Internet users.

Stoyan Stefanov and Nicole Sullivan team up to contribute Chapter 10, *Optimizing Images*. This is a thorough treatment of the topic. This chapter reviews all popular image formats, presents numerous image optimization techniques, and describes the image compression tools of choice.

The remaining chapters were written by me. Chapter 11, *Sharding Dominant Domains*, reminds us of the connection limits in the popular browsers of today, as well as the next generation of browsers. It includes techniques for successfully splitting resources across multiple domains.

Chapter 12, *Flushing the Document Early*, walks through the benefits and many gotchas of using chunked encoding to start rendering the page even before the full HTML document has arrived.

Browser Performance

Iframes are an easy and frequently used technique for embedding third-party content in a web page. But they come with a cost. Chapter 13, *Using Iframes Sparingly*, explains the downsides of iframes and offers a few alternatives.

Chapter 14, *Simplifying CSS Selectors*, presents the theories about how complex selectors can impact performance, and then does an objective analysis to pinpoint the situations that are of most concern.

The Appendix, *Performance Tools*, describes the tools that I recommend for analyzing web sites and discovering the most important performance improvements to work on.

Conventions Used in This Book

The following typographical conventions are used in this book:

Italic

Indicates new terms, URLs, email addresses, filenames, file extensions, pathnames, and directories

Constant width

Indicates commands, options, switches, variables, attributes, keys, functions, types, classes, namespaces, methods, modules, properties, parameters, values, objects, events, event handlers, XML tags, HTML tags, macros, the contents of files, and the output from commands

Constant width bold

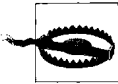
Shows commands or other text that should be typed literally by the user

Constant width italic

Shows text that should be replaced with user-supplied values



This icon signifies a tip, suggestion, or general note.



This icon indicates a warning or caution.

Comments and Questions

Please address comments and questions concerning this book to the publisher:

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472
800-998-9938 (in the United States or Canada)
707-829-0515 (international or local)
707-829-0104 (fax)

We have a web page for this book, where we list errata, examples, and any additional information. You can access this page at:

<http://www.oreilly.com/catalog/9780596522308>

To comment or ask technical questions about this book, send email to:

bookquestions@oreilly.com

For more information about our books, conferences, Resource Centers, and the O'Reilly Network, see our web site at:

<http://www.oreilly.com>

Using Code Examples

You may use the code in this book in your programs and documentation. You do not need to contact us for permission unless you're reproducing a significant portion of the code. For example, writing a program that uses several chunks of code from this book does not require permission. Selling or distributing a CD-ROM of examples from this book *does* require permission. Answering a question by citing this book and quoting example code does not require permission. Incorporating a significant amount of example code from this book into your product's documentation *does* require permission.

We appreciate, but do not require, attribution. An attribution usually includes the title, author, publisher, and ISBN. For example: "*Even Faster Web Sites*, by Steve Souders. Copyright 2009 Steve Souders, 978-0-596-52230-8."

If you feel your use of code examples falls outside fair use or the permission given above, feel free to contact us at permissions@oreilly.com.

Safari® Books Online



When you see a Safari® Books Online icon on the cover of your favorite technology book, that means the book is available online through the O'Reilly Network Safari Bookshelf.

Safari offers a solution that's better than e-books. It's a virtual library that lets you easily search thousands of top tech books, cut and paste code samples, download chapters, and find quick answers when you need the most accurate, current information. Try it for free at <http://my.safaribooksonline.com>.

Acknowledgments

I first want to thank the contributing authors: Dion Almaer, Doug Crockford, Ben Galbraith, Tony Gentilcore, Dylan Schiemann, Stoyan Stefanov, Nicole Sullivan, and Nicholas Zakas. They've made this a special book. Each of them is an expert in his or her own right. Most of them have written their own books. By sharing their expertise, they've helped create something unique.

I want to thank all the reviewers: Julien Lecomte, Matthew Russell, Bill Scott, and Tenni Theurer. I extend an especially strong thank you to Eric Lawrence and Andy Oram. Eric reviewed both this book as well as *High Performance Web Sites*. In both cases, he provided incredibly thorough and knowledgeable feedback. Andy was my editor on *High Performance Web Sites*. More than anyone else, he is responsible for improving how this book reads, making it flow smoothly from line to line, section to section, and chapter to chapter.

A special thank you goes to my editor, Mary Treseler. Coordinating a book with multiple authors is an opportunity that many editors will pass over. I'm glad that she took on this project and helped guide it from a bunch of ideas to what you're holding in your hands now.

I work with many people at Google who have a penchant for web performance. Tony Gentilcore is the creator of Fasterfox (<http://fasterfox.mozdev.org/>) and the author of Chapter 9. He's also my officemate. Several times a day we'll stop to discuss web performance. Steve Lamm, Lindsey Simon, and Annie Sullivan are strong advocates for performance who I work with frequently. Other Googlers who have contributed to what I know about web performance include Jacob Hoffman-Andrews, Kyle Scholz, Steve Krulewitz, Matt Gunderson, Gavin Doughtie, and Bryan McQuade.

Many of the insights in this book come from my friends outside Google. They know that if they tell me about a good performance tip, it's likely to end up in a book or blog post. These performance cohorts include Dion Almaer, Artur Bergman, Doug Crockford, Ben Galbraith, Eric Goldsmith, Jon Jenkins, Eric Lawrence, Mark Nottingham, Simon Perkins, John Resig, Alex Russell, Eric Schurman, Dylan Schiemann, Bill Scott, Jonas Sicking, Joseph Smarr, and Tenni Theurer.

I've inevitably forgotten to mention someone in these lists. I apologize, and want to thank all of you for taking the time to send me email and talk to me at conferences. Hearing your lessons learned and success stories keeps me going. It's important to know there are so many of us who are working to make the Web a faster place.

Thank you to my parents for being proud to have a son who's an author. Most importantly, thank you to my wife and three daughters. I promise to take a break now.

Table of Contents

Credits	xi
Preface	xiii
1. Understanding Ajax Performance	1
Trade-offs	1
Principles of Optimization	1
Ajax	4
Browser	4
Wow!	5
JavaScript	6
Summary	6
2. Creating Responsive Web Applications	7
What Is Fast Enough?	9
Measuring Latency	10
When Latency Goes Bad	12
Threading	12
Ensuring Responsiveness	13
Web Workers	14
Gears	14
Timers	16
Effects of Memory Use on Response Time	17
Virtual Memory	18
Troubleshooting Memory Issues	18
Summary	19
3. Splitting the Initial Payload	21
Kitchen Sink	21
Savings from Splitting	22
Finding the Split	23
Undefined Symbols and Race Conditions	24

4. Loading Scripts Without Blocking	27
Scripts Block	27
Making Scripts Play Nice	29
XHR Eval	29
XHR Injection	31
Script in Iframe	31
Script DOM Element	32
Script Defer	32
document.write Script Tag	33
Browser Busy Indicators	33
Ensuring (or Avoiding) Ordered Execution	35
Summarizing the Results	36
And the Winner Is	38
5. Coupling Asynchronous Scripts	41
Code Example: menu.js	42
Race Conditions	44
Preserving Order Asynchronously	45
Technique 1: Hardcoded Callback	46
Technique 2: Window Onload	47
Technique 3: Timer	48
Technique 4: Script Onload	49
Technique 5: Degrading Script Tags	50
Multiple External Scripts	52
Managed XHR	52
DOM Element and Doc Write	56
General Solution	59
Single Script	59
Multiple Scripts	60
Asynchronicity in the Real World	63
Google Analytics and Dojo	63
YUI Loader Utility	65
6. Positioning Inline Scripts	69
Inline Scripts Block	69
Move Inline Scripts to the Bottom	70
Initiate Execution Asynchronously	71
Use Script Defer	73
Preserving CSS and JavaScript Order	73
Danger: Stylesheet Followed by Inline Script	74
Inline Scripts Aren't Blocked by Most Downloads	74

Inline Scripts Are Blocked by Stylesheets	75
This Does Happen	77
7. Writing Efficient JavaScript	79
Managing Scope	79
Use Local Variables	81
Scope Chain Augmentation	83
Efficient Data Access	85
Flow Control	88
Fast Conditionals	89
Fast Loops	93
String Optimization	99
String Concatenation	99
Trimming Strings	100
Avoid Long-Running Scripts	102
Yielding Using Timers	103
Timer Patterns for Yielding	105
Summary	107
8. Scaling with Comet	109
How Comet Works	109
Transport Techniques	111
Polling	111
Long Polling	112
Forever Frame	113
XHR Streaming	115
Future Transports	116
Cross-Domain	116
Effects of Implementation on Applications	118
Managing Connections	118
Measuring Performance	119
Protocols	119
Summary	120
9. Going Beyond Gzipping	121
Why Does This Matter?	121
What Causes This?	123
Quick Review	123
The Culprit	123
Examples of Popular Turtle Tappers	124
How to Help These Users?	124
Design to Minimize Uncompressed Size	125
Educate Users	129

Direct Detection of Gzip Support	130
10. Optimizing Images	133
Two Steps to Simplify Image Optimization	134
Image Formats	135
Background	135
Characteristics of the Different Formats	137
More About PNG	139
Automated Lossless Image Optimization	141
Crushing PNGs	142
Stripping JPEG Metadata	143
Converting GIF to PNG	144
Optimizing GIF Animations	144
Smush.it	145
Progressive JPEGs for Large Images	145
Alpha Transparency: Avoid AlphaImageLoader	146
Effects of Alpha Transparency	147
AlphaImageLoader	149
Problems with AlphaImageLoader	150
Progressively Enhanced PNG8 Alpha Transparency	151
Optimizing Sprites	153
Über-Sprite Versus Modular Sprite	154
Highly Optimized CSS Sprites	155
Other Image Optimizations	155
Avoid Scaling Images	155
Crush Generated Images	156
Favicons	157
Apple Touch Icon	158
Summary	159
11. Sharding Dominant Domains	161
Critical Path	161
Who's Sharding?	163
Downgrading to HTTP/1.0	165
Rolling Out Sharding	168
IP Address or Hostname	168
How Many Domains	168
How to Split Resources	168
Newer Browsers	169
12. Flushing the Document Early	171
Flush the Head	171
Output Buffering	173