

Andrew B. King: Speed up your site – Web site optimization

„User patience is a time bomb on the web.“, das ist der erste Satz auf dem rückseitigen Bucheinband. Viele Entwerfer von Web-Plätzen kümmern sich nicht darum, wie langsam die Information, nach denen ein Benutzer Ausschau hält, auf dem Bildschirm erscheinen wird. In ihrem Büro haben sie eine *super-high-speed connection*¹ zum Server von ihrem *high-tech computer* mit seinem *GHz processor*. In B2B und B2C Anwendungen, um die es in den Empfehlungen des Autors vor allem geht, müssen Verbesserungen auftreten. Viele Benutzer haben zuhause noch eine langsame *56 kbit/sec audio modem connection* zu ihrem Anbieter (in den USA nach einer Studie von Sept. 2003 um 72 %!) und sind zufrieden mit dem MHz-Rechner zum Aufbereiten, den sie vor einiger Zeit gekauft und bezahlt haben und mit dem sie vertraut sind. Erfährt man (nach einer anderen Studie), dass die Ladezeit von Web-Seiten bei den 40 Spitzenanbietern für Konsumenten von 19.5 sec im August 2001 auf 21.4 sec im August 2003 angestiegen ist, muss unbedingt eine höhere Performance erreicht werden. Selbst wenn dieses Szenarium nicht mehr völlig richtig liegt – viele Benutzer haben modernere Einrichtungen –, das Herunterladen einer 2 Megabyte Graphik ist immer noch ärgerlich.

Dieses Buch erschien zur genau richtigen Zeit auf dem Markt. Es ist kein Lehrbuch für einen Anfänger beim Entwurf und der Programmierung eines Web-Platzes. Es ist ein Kompendium von Vorgehensweisen zum Verbessern eines Entwurfs und Unterstützen des Programmierens von rasch arbeitenden Web-Plätzen. In sechs Teilen mit 19 Kapiteln werden alle aktuellen Problembeispiele angesprochen:

– TEIL I: „*The Psychology of Performance*“ verlangt (im ersten Kapitel beschrieben) eine Systemreaktionszeit von acht

Sekunden, plus oder minus zwei; eingeführt wird (im zweiten Kapitel) ein Konzept, das *flow* genannt wird, ein oft vernachlässigtes Performancekriterium, das als ein integrales Maß die anhaltende, zufriedenstellende Erfolgserfahrung beim Arbeiten und schließlich Erreichen eines Ziels abdeckt, das sich der Benutzer gesetzt hat.

– TEIL II: „*Optimizing Markup: HTML and XHTML*“ stellt in den Kapiteln 3 bis 6 schlechte Programmierpraxis heraus, die zu überlangen Dateien führt. Es ist eines der „überlangen“ Teile – man erlaube mir das Wortspiel; warum nicht, um zu zeigen, worauf der Autor in seiner Argumentation besonders eingeht –, 95 Seiten, mit viel Details über HTML, XHTML, XML.

Das dritte Kapitel, „*HTML Optimization*“, gibt in 11 Schritten mit gut gewählten Beispielen und Illustrationen Ratschläge, was ein Entwerfer für rasch zu ladende Plätze vermeiden sollte; es wird immer nach einem Ausgleich zwischen Seitengröße und Aufbaugeschwindigkeit gesucht.

Im vierten Kapitel, „*Advanced HTML Optimization*“, sind dann mögliche Optimierungsvorgehensweisen erklärt, automatische und/oder manuelle. Das Resultat, das der Autor zeigen möchte, im 6. Kapitel dann später exemplifiziert, ist das Optimieren einer „*bloated page to one sixteenth of its original size*“ (*bloated*, „aufgeblasen“, gemeint ist HTML-Text mit u.a. CSS und Java-Script – diese Worte muss man sich merken! –). Eine Optimierung ist zunächst automatisch mit Hilfe eines Analyseprogramms zu erreichen (als Beispiel bezieht sich der Autor auf *SpaceAgent*), bei dem ein unstrukturierter HTML-Textblock im Beispiel mit bis zu 11.6 % Größeneinsparung entsteht; Verlust von Lesbarkeit gilt nicht als starkes Gegenargument. Die zweite Vorgehensweise, um eine Reduktion auf nur noch 6.2 % der Originalgröße zu erreichen, manuelle Optimie-

rung, erfordert sicherlich Erfahrung, Erfahrung, wie sie in diesem Buch vorgestellt wird (in beiden Fällen sind die Prozentangaben nach *zipping*). Alles ein bemerkenswerter Erfolg, nicht wahr? Um die Herausforderung zu berichten, die der Autor an den HTML-Experten stellt, ein Zitat aus der Einleitung: „...*HTML size is the one web page component that users cannot control. They can turn off JavaScript, images, plug-ins, and Java, but they can't turn off HTML bloat – only designers can.*“ Ich möchte, ohne in kritischer Absicht argumentieren zu wollen, hinzufügen, was King empfiehlt ist, HTML als die Assemblersprache des Web anzusehen, ohne Formatierung für Lesbarkeit und ohne irgendeinen netten, aber Bandbreite erfordernden sonstigen Ballast.

Ein anderes Zitat, aus Kapitel 4 „*Advanced HTML Optimization*“, ist hart: „*Using tables ... is usually a recipe for disaster*“; Das Hauptproblem bzgl. Ladezeit ist hier, dass die Tabellengröße dynamisch berechnet wird und in sich abhängig ist (Vermerk des Compilerentwicklers, der ich bin: das Aufbereiten ist nicht mehr linear in der Textlänge, ein Grauen!); *stylesheets* sind empfohlen.

Noch im Kapitel 4, vom Autor nicht so kritisch gesehen, sind HTML-forms und Menüs die nächsten Punkte, einschließlich Reduktion der Maus-klicks; man findet gute Ratschläge. Sind viele URLs in einer Seite, wird der Gebrauch des *mod:rewrite*-Moduls empfohlen; berichtet werden dabei Ersparnisse der HTML-Größe von bis zu 30 %.

Der letzte behandelte Punkt ist dann dabei „*HTML and Compression Saving*“, typischerweise 80 % bis 85 %, selbst bis zu 90 %.

Für XML-Entwerfer enthält Kapitel 5, „*Extreme XHTML*“, 23 Seiten, ähnliche Ratschläge. Man nehme das Buch zur Hand und studiere es selbst!

Der Inhalt von Kapitel 6, „*Case Study: PopularMechanics.com*“, 12 _ detailierte Seiten wurde schon erwähnt.

¹ Man erlaube mir auch einmal, Denglisch zu schreiben!

– Teil III: „*DHTML Optimization: CSS and JavaScript*“, ist mit 107 Seiten in 5 Kapiteln mit Nummern 7 bis 11 wieder „überlang“. Man glaube es mir, man findet wieder guten Rat. Zwei Kapitel will ich hervorheben.

Kapitel 10, „*Optimizing JavaScript for Execution Speed*“ bringt einen neuen Aspekt, nicht nur Textgröße. Es liest sich wie ein Lehrbuch über das Optimieren der Programmausführung mit Zusätzen im Kontext der Web-Programmierung. Lesen Sie es selbst, die aufgebrachte Leszeit wird für Sie selbst und alle Leser Ihrer Seite gut kompensiert!

Das kurze Kapitel 11, ein weiterer Bericht über eine „*Case Study: DHTML.com*“, beweist die Nützlichkeit der Empfehlungen bzgl. Ladezeit.

– Teil IV: „*Graphics and Multimedia Optimization*“, drei Kapitel, erläutert Techniken, um Laden und Präsentieren von Grafiken zu beschleunigen.

JPEG Fotos und gescannte Bilder, GIF- und PNG-Grafiken werden in Kapitel 12 behandelt, „*Optimizing Web Graphics*“, – mit Koautor Konstantin Balashov –. Minimieren der Anzahl der HTTP-Anforderungen ist der erste Ratschlag. Dies startet bereits beim Fotografieren und Vorbereiten für die Aufnahme in eine Web-Anwendung. Entwurfsentscheidungen gibt es dabei, z. B. Kombinieren von Grafiken und so das Minimieren der HTTP-Anforderungen oder Verwenden von Suchbildern (*thumb-nails*) als Zwischenstopps beim Navigieren. All dies erfordert einen Ausgleich zwischen der Größe der Grafikdatei und der Bildpräsentationsqualität. Illustrationen dazu sind schwarz-weiß; farbige Illustrationen finden sich in einem begleitenden Web-Platz (siehe unten). Es gibt viele Hinweise, wie man mit verfügbaren kommerziellen Grafikbearbeitungsprogrammen umgehen kann.

Kapitel 13, „*Minimizing Multimedia*“, beschäftigt sich mit Multimedia, d. h. (Kombinationen von) Audio, Video, Text, Graphiken, Animation usw.; zum Selbststudium empfohlen! Das kurze Kapitel 14 bietet eine dritte „*Case Study: Apple.com*“ über Multimedia-Minimierung. Diese beiden Kapitel wurden von Jason Wolf verfaßt.

– Teil V: „*Search Engine Optimization*“ besteht aus zwei Kapiteln, wobei ich

vorschlage, bei einer Neuauflage ihn „*Optimization for relevant retrieval by search engines*“ zu nennen. Behandelte Punkte sind u. a., lassen sie es mich Englisch sagen, *keyword selection*, *formulation of good multiword phrases*, und *best placement in your site*. Das zweite Kapitel darin berichtet wieder über Fallstudien.

– Der letzte Teil VI: „*Advanced Optimization Techniques*“, mit drei Kapiteln, behandelt Techniken auf der Server-Seite (Kapitel 17) und Kompression (Kapitel 18); eine weitere Fallstudie über yahoo.com und WebReference.com schließt die Sammlung der Ratschläge ab. Ich empfehle wieder Selbststudium. Nur einige Angaben, dass es wert ist zu lesen: Abkürzung von URL und Verzeichnissen sparte 28 % bzw. 20 % der Dateigröße der Portalseite ein.

Hatten Sie bemerkt, dass die Portalseite eines großen USA-Provider (*Yahoo*), bekannt als eine der geschäftigsten Web-Plätze der Welt, von 1969, 5.998 Byte HTML-Text, auf 30.887 Byte anwuchs (King nennt dies „*Web Page Size Inflation*“). Das ist keine Ausnahme. Es besteht die große Gefahr, dass der Web-Zugriff trotz technologischer Verbesserungen langsamer wird. King war sicherlich durch Beobachtungen dieser Art motiviert, das Buch zu schreiben.

Gute Empfehlungen, wussten Sie, dass:

- Leerzeichen in HTML-Dateien Gift sind für die Performance und sollten vermieden werden; eine überraschende Tatsache, wenn man gewohnt ist, Text einzurücken, um Lesbarkeit (durch einen Menschen) zu erleichtern (ich würde Einrücken nicht als so schlecht ansprechen, man suche nach einem Kompromiss).
- Web-Seiten *graphics-only* ohne ein **alt** sind, so spricht der Autor sarkastisch, „*beautiful slow-loading pages ... not usable with graphics turned off*“.
- Text überträgt Information viel schneller als Grafiken und ist von Behinderten leichter zu lesen.
- Man benutze nur **meta**-Tags **description** und **keywords**.
- Ein abschließendes „/“, das am Ende einer URL fehlt, die auf ein Verzeichnis verweist, erfordert Aufwand und

erhöht die Suchzeit nach dem, was angefordert ist.

- Scripts werden beim Laden ausgeführt. Daher ist es ein guter Ansatz, Java-Skripts – beispielsweise – zu verzögern, *deferred*, wann immer möglich: „*Every byte and HTTP request you put before your viewable content delays its display*“, ein Zitat von Seiten 199 – 201.
- Bilder, die mit einer Digitalkamera aufgenommen sind, erfordern unbedingt spezielle Einstellungen oder einige Nachbearbeitung, um optimale Übertragung über das Web und gute Aufbereitung im Stöberer zu erreichen. Voreinstellungen in der Kamera sind gewöhnlich für Papierbilder getroffen, nicht für Web-Seiten.
- Es ist ratsam, die Paketgröße beim Aufteilen von Dateien bzw. Grafiken zu bedenken (1200 – 1500 Byte), insbesondere wenn eine Hochgeschwindigkeitsverbindung – beispielsweise in einem Intranet – vorliegt, um eine gute, ausgeglichene Mehrfachbearbeitung zu ermöglichen.
- Es gibt eine Möglichkeit, Suchhinweise auf Relevanz zu prüfen: <http://www.wordtracker.com/>.
- Einige *spiders* (d. h. *automatic agents visiting your site for search preparation*) kommen mit *frames* nicht zu recht, man verliert Treffer (ich weiß, dass es noch andere Argumente zur Gebrauchstauglichkeit von *frames* gibt, nämlich Probleme bei der Navigation – schade, ich fand keine Anmerkung dazu im Buch ! –).

Falls Sie das nicht wussten oder dem nicht folgen wollen, das Buch gibt Rat, Fallen für schwache Performance bzw. schlechter Erreichbarkeit Ihrer Web-Seiten zu umgehen. Viele Anregungen sind nicht neu, wieder darüber zu lesen hilft, sich an sie bei praktischer Entwurfstätigkeit zu erinnern.

Puristen im Software-Ingenieurwesen werden nicht alle Empfehlungen schätzen. Hier ist die alte Kontroverse zwischen Wissenschaft und Gewerbe auf letzteres ausgerichtet. Das Web ist (noch) ein technisches Kunstwerk, das hoch optimierte Lösungen für seinen Gebrauch erfordert. Nichtsdestotrotz, die meisten gegebenen Ratschläge sind durch technisch und wissenschaftlich

fundierte Artikel in wissenschaftlichen Veröffentlichungen oder Büchern gestützt. Und es ist durchaus besonders an Vorteile für Benutzer gedacht !

Das Buch ist gut geschrieben, 496 Seiten. Beim Lesen findet man viele, auffällig hervorgehobene Nebenbemerkungen, gut gewählte Illustrationen durch Web-Seiten, Beispiele aus der Realität, die vermieden werden sollten, und eine Vielzahl von Kodierungsbeispielen. Bezug auf Normen, auf hilfreiche Optimierungsprogramme für Web-Seiten usw. sind vorhanden. Für eilige Leser sind gebotene Kapitelzusammenfassun-

gen hilfreich. In Anbetracht der 496 Seiten, wirklich viel, sollte eine kritische Nebenbemerkung erlaubt sein, einige Wiederholungen hätten vermieden werden können.

Es gibt einen begleitenden Web-Platz <http://www.WebSiteOptimization.com> mit zusätzlichem Material (einschl. Errata); und es gibt ein leistungsfähiges Analysewerkzeug unter <http://WebPageAnalyzer.com>

(ich hatte einige Probleme, auf diese Seiten zuzugreifen; seien Sie geduldig, es funktioniert irgendwann!). Vergessen Sie nicht, die dritte Einbandseite mit ei-

ner Reihe anderer (teils kommerzieller) Web-Seiten zur Kenntnis zu nehmen, die das Buch ergänzen.

Prof. Dr. Hans-Jürgen Hoffmann, Technische Universität Darmstadt, HJHoffmann@informatik.tu-darmstadt.de oder HJHoffmann@ACM.org

„Speed up your site – Web site optimization“. Andrew B. King with a Foreword by Jacob Nielsen, Co-writers Jason Wolf (chapters 13 and 14) and Konstantin Balashov (chapter 18). New Riders Publishing, Boston – Indianapolis – London – Munich – New York – San Francisco. 2003. ISBN: 0-7357-1324-3.