

Thomas Stensitzki

# »Geschwindigkeit ist keine Hexerei

**Professionelle Webseiten laden zum Teil immer schneller und Internetsurfer werden dadurch immer ungeduldiger. Welche Konsequenzen langsame Webseiten für den Umsatz haben und mit welchen technischen Möglichkeiten man dem entgegensteuern kann, erklärt Thomas Stensitzki.**

Die Performance einer Website entscheidet darüber, ob sie ihr Ziel erreicht: Image und Bekanntheit, Leser- oder Kundengewinnung, Werbeschaltung oder Bestellung. Ausschlaggebend für die Performance sind im Wesentlichen drei Kriterien:

- » die Funktionalität,
- » die Geschwindigkeit und
- » die Anwenderfreundlichkeit (Usability).

Dieser Artikel greift den wichtigen Aspekt der Geschwindigkeit heraus und beleuchtet ihn von drei Seiten: Toleranz des Anwenders, Konsequenzen für unterschiedliche Geschäftsmodelle und technische Ansätze zur Problemlösung.

## 1. Welche Rolle spielt die Geschwindigkeit?

Während Microsoft und Google Verbesserungen ihrer Webangebote bereits in Millisekunden rechnen, ist vielen Unternehmen offenbar nicht klar, dass Geschwindigkeit einer der entscheidenden Erfolgsfaktoren im Online-Business ist, egal, ob sie das Internet nur als Online-Visitenkarte nutzen oder Geschäfte damit machen wollen. 39 Prozent der Konsumenten halten die Geschwindigkeit einer Seite für wichtiger als ihre Funktionalität, nur die Sicherheit ist noch wichtiger. Für Online-Nutzer ist bei ihrem Banking-Portal sogar die Schnelligkeit das wichtigste Kriterium, noch vor Sicherheit und Anwenderfreundlichkeit (Quelle: Penny Crosman in „Bank Systems & Technology“).

Die wichtigsten messbaren Indikatoren (Key Performance Indicators – **KPI\***) für den Erfolg einer Website sind die Reichweite (Wie viele Anwender finden die Seite?), die Aktivität (Wie häufig interagieren sie mit der Seite?) und die Kon-

version (Tun sie das, was der Betreiber möchte, und wie sehr profitiert er davon?).

Schaubild 1 zeigt vereinfacht, wie diese KPI aufeinander aufbauen und welche Kennzahlen zurate gezogen werden können, um sie zu messen.

Wie wirkt sich nun die Geschwindigkeit einer Seite auf den Anwender aus?

## 2. Webperformance aus Sicht des Anwenders

Niemand steht gerne an. Während wir im realen Leben aber oft nicht umhin können, uns in einer Schlange einzureihen und zu warten, bis wir an der Reihe sind, haben wir im Internet eine Alternative: Wir klicken einfach weg. Oder wir öffnen nebenbei eine andere Seite und verlieren das Interesse an der vorigen – entweder, weil die neue besser und schneller unsere Fragen beantwortet und unsere Wünsche erfüllt, oder weil wir die erste schlicht vergessen haben.

Wir wissen ja vorher nicht, wie lange wir warten müssen. Das „Technion-Israel Institute of Technology“ in Haifa hat in einer Studie die Bereitschaft von Anrufern, in einer Telefonschleife zu warten, untersucht. Dazu wurden Personen, die sich an einem Universitätslabor für Experimente anmelden wollten, in unterschiedliche Warteschleifen geschickt.

Ein Drittel der Anrufer hörte 108 Sekunden lang die „Ballade pour Adeline“ von Richard Clayderman. Beim zweiten Drittel wurde das Klavierstück dreimal von einer Stimme unterbrochen, die sich beim Anrufer entschuldigte und ihn vertröstete. Das letzte Drittel erhielt bei den Unterbrechungen eine konkrete Auskunft über seinen Platz in der Warteschleife („Es sind noch vier Anrufer vor Ihnen“).

TURBO

### DER AUTOR



**Thomas Stensitzki** ist Senior Consultant bei der

iCOMcept GmbH in Aachen, einem IT-Beratungsunternehmen und Microsoft-Certified-Gold-Partner, das sich auf die Konzeption ganzheitlicher IT-Lösungen, Sicherheit und Infrastruktur, Hosting, E-Commerce und Prozessoptimierung fokussiert hat.

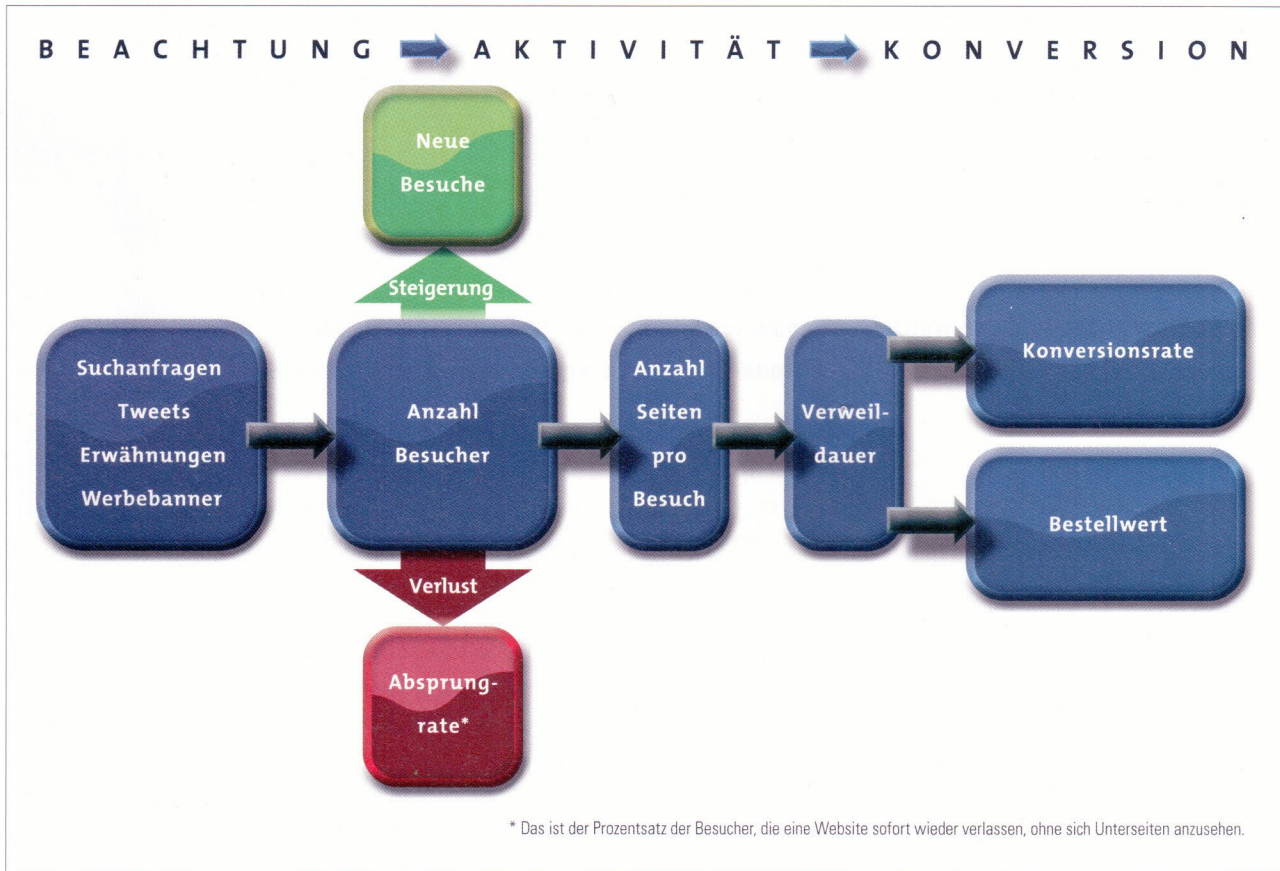


Abbildung 1: Beachtung, Aktivität und Konversion lassen sich anhand diverser Kennzahlen messen (Quelle: Joshua Bixby, „Web Performance Today“)

Das Ergebnis verwundert nicht: Bei den Entschuldigungen verloren die Anrufer am schnellsten die Geduld, gefolgt von der Musik. Beide wirkten anonym, statisch und somit undurchschaubar. Hingegen wurde die Auskunft über den Warteplatz als angenehm und hilfreich beurteilt und sorgte für geduldige, zufriedene Anrufer. Diese Auskunft wurde deshalb so positiv aufgenommen, weil sie den Anrufern ein Gefühl der Kontrolle vermittelte: Sie wussten, wann sie an der Reihe waren (Quellen: Journal of Applied Psychology, 2/2007; <http://einfach.st/jap>).

Softwareentwickler haben sich dieses Prinzip schon lange zu eigen gemacht, indem sie uns auf einem sich füllenden Balken zeigen, wie schnell das Herunterladen oder Installieren eines Programms fortschreitet. Wenn es sehr schnell geht, können wir die Wartezeit gut aushalten. Wenn es länger dauert, können wir die Zeit für andere Dinge nutzen.

Dieses Gefühl der Kontrolle haben wir aber beim Surfen nicht. Und weil wir nicht wissen, wie lange der Aufbau einer Website dauert, reagieren wir mit Abbruch. Dabei ist die Toleranzzeit im Laufe der Zeit immer niedriger geworden:

1999 erwarteten Online-Shopper im Schnitt, dass eine Seite sich in 8 Sekunden aufbaut. 2006 gab man ihr noch 6 Sekunden und heute sind es nur noch 2 Sekunden oder weniger (Quelle: Forrester Consulting).

Hinzu kommt, dass Anwender ein falsches Zeitgefühl haben: Fragt man sie, wie lange der Aufbau einer Seite dauert, liegen ihre Angaben ca. 15 Prozent über dem realen Wert. Und wenn man sie erst später fragt, ist in ihrer Erinnerung die „gefühlte“ Wartezeit sogar 35 Prozent höher (Quelle: Stoyan Stefanov „Psychology of Performance“).

57 Prozent der Online-Nutzer verlassen eine Seite, wenn sie nach 3 Sekunden nicht geladen ist. 80 Prozent davon kehren nicht zurück – und 3 Prozent erzählen es anderen Anwendern (Quelle: PhoCusWright). Auch wenn die Performanceprobleme erst während einer konkreten Transaktion auftauchen, verlassen noch zwischen 37 und 49 Prozent die Seite oder springen zur Konkurrenz. 77 Prozent der Letzteren teilen diese Erfahrung Freunden und Kollegen mit (Quelle: Sean Power „Metrics“).

Überraschende Zahlen liefern die Untersu-

**„ Kurzeitgedächtnis und Angst vor Kontrollverlust beeinflussen unser Verhalten beim Online-Shopping. “**

**„Fragen Sie Ihren Provider, wie viele Domains sich bei ihm einen Server teilen müssen. Je weniger das sind, desto besser für Ihr Suchmaschinenranking!“**

chungen auch über das unterschiedliche Verhalten der Alters- und Nutzergruppen: Die Generation Y<sup>1</sup> ist allgemein bekannt für ihre notorische Ungeduld, aber hätten Sie gedacht, dass die ungeduldigste aller Nutzergruppen mobile Webanwender über 45 sind (Quelle: Equation Research)?

16 Prozent der Konsumenten haben bereits via Mobil- oder Smartphone eingekauft, aber 27 Prozent sind unzufrieden mit dieser Erfahrung, weil es zu langsam ging. Ein Drittel der Konsumenten will in Zukunft via Smartphone shoppen und 5 Prozent davon kündigen an, dass dies ein wichtiges Kriterium für ihre Loyalität zum Online-Verkäufer sein wird (Quelle: Forrester Consulting).

Noch eine Zahl: 37 Prozent der enttäuschten Nutzer kommen nicht wieder, 27 Prozent gehen zur Konkurrenz. Selbst wenn wir einzelne dieser Prozentwerte anzweifeln oder andere Untersuchungen andere Zahlen nennen: Die Tendenz zeigt zweifellos, dass die Geschwindigkeit einen wichtigen Beitrag zur Anwenderzufriedenheit leistet.

### 3. Konsequenzen geringer Webperformance

Bevor Sie als Anbieter eines Online-Shops oder Werbeportals Ihre Besucher durch Geschwindigkeit des Seitenaufbaus überzeugen können, müssen Sie aber erst einmal gefunden werden. Schon an dieser Stelle spielt die Geschwindigkeit eine wichtige Rolle.

#### 3.1. Suchmaschinen

Suchmaschinenoptimierung ist eine Wissenschaft für sich und eine ganze Branche lebt davon, dieses vermeintliche Expertenwissen zu verkaufen. Dass neben der Gestaltung der Seite, der Aktualität und dem Textinhalt (Content) die Geschwindigkeit des Aufbaus eine Rolle spielt, liegt auf der Hand:

Um Rankingergebnisse zu ermitteln, müssen Suchmaschinen die Seiten indizieren. Dies geschieht nach anspruchsvollen Algorithmen, die mehrere Hundert Kriterien berücksichtigen. Google beispielsweise wertet zwischen 200 und 300 „Signale“ aus. Nach einem komplizierten Kriteriensystem „entscheidet“ die Suchmaschine, ob sie eine Seite überhaupt besucht und indiziert, wo sie die Seite in den Ergebnislisten einreicht und wie sie sie klassifiziert.

Bei Google ist beim Indexierungsprozess ein Scheduler für die Geschwindigkeit des Crawlers (Robot) verantwortlich, der den Webserver vor Überlastung schützt. Reagiert ein Server langsam – oder teilen sich viele Domains einen einzigen Server – können innerhalb eines bestimmten Intervalls nur wenige Seiten der einzelnen Domains indexiert werden. Im Umkehrschluss werden bei einer Beschleunigung des Servers mehr Seiten indexiert, was die Chance auf ein höheres Ranking erhöht.

Ein weiteres Kriterium ist das Verhalten der Anwender: Werden Links, die oben im Ranking erscheinen, selten geklickt oder sofort wieder verlassen, zieht die Maschine Rückschlüsse auf den Wert des Links und wird ihn gegebenenfalls abwerten. Die Geschwindigkeit des Seitenaufbaus und der Wiedergabe im Browser, die Toleranz der Besucher bei langsamen Seiten sowie deren Erfahrung mit der Website haben also Konsequenzen für das Ranking in den Suchmaschinen. Kurz gesagt: Schneller Seitenaufbau bedeutet höheres Ranking bedeutet größere Reichweite.

#### 3.2. Bankingportale und E-Shops

Die Performance eines E-Shop ist verantwortlich für Kaufabschluss oder Kaufabbruch. Lange Ladezeiten können dazu führen, dass ein Einkauf gar nicht gestartet wird, gefüllte Einkaufswagen stehenbleiben oder Konsumenten zur Konkurrenz abwandern. Insgesamt macht der Einfluss der Performance auf den Gesamtumsatz bis zu 9 Prozent aus (Quelle: Aberdeen Group).

Ein paar Zahlen zur Erläuterung: Eine Verzögerung von einer Sekunde verursacht 11 Prozent weniger Seitenaufrufe, 16 Prozent weniger Kundenzufriedenheit und 7 Prozent weniger Konversion. Eine Seite, die in drei Sekunden lädt, bedeutet 22 Prozent weniger Seitenaufrufe und eine um 50 Prozent höhere Abbruchrate als eine Seite, die nur eine Sekunde braucht, die Konversionsrate liegt bei minus 22 Prozent. Bei einer Seite, die in fünf Sekunden lädt, ergeben sich 35 Prozent weniger Seitenaufrufe, eine um 105 Prozent höhere Abbruchrate und eine Konversion von minus 38 Prozent (Quelle: Strangeloop). Rechnen Sie sich aus, wie viel das im Jahr bei einem E-Shop macht, der im Schnitt 10.000 € Umsatz am Tag generiert.

Eine Befragung von Online-Bankern und Brokern ergab, dass mehr als die Hälfte (56 Prozent) eine Seitenladezeit von 2 Sekunden oder weniger erwarten. 65 Prozent haben schon schlechte Erfahrungen mit Portalen gemacht, Hauptgrund für ihre Enttäuschung war die Performance.

**„Shopzilla konnte mit einer Verringerung der Seitenladezeit von 6 auf 1,2 Sekunden die Anzahl der Besucher, die über Suchmaschinen kommen, verdoppeln“**  
(Quelle: Shopzilla).

<sup>1</sup> „Generation Y“ oder „Millennials“ nennt die Soziologie die Generation der nach 1980 Geborenen, die in einem Umfeld von Internet und mobiler Kommunikation aufgewachsen ist.

Je dringender und komplexer die Online-Aktionen sind, desto weniger sind die Anwender bereit, Verzögerungen in Kauf zu nehmen, sie wechseln sogar zu teureren Anbietern, um ihre Transaktion zu beenden (65 Prozent), oder nutzen Offline-Kanäle, wenn die Website unerreichbar oder langsam ist (54 Prozent) (Quelle: Forrester Consulting).

Die Beschleunigung der Seitenladezeiten brachte bei einigen großen E-Shops signifikante Ergebnisse: Shopzilla verkürzte die durchschnittliche Seitenladezeit von 6 auf 1,2 Sekunden und erreichte damit ein 12-prozentiges Umsatzwachstum und 25 Prozent mehr Seitenaufrufe. Gleichzeitig verdoppelte sich die Zahl der Besucher, die über Suchmaschinen kamen (Quelle: Shopzilla).

Amazon fand heraus, dass jede Beschleunigung der Seiten um 100 Millisekunden sich in einem Umsatzwachstum von 1 Prozent auszahlt. Mit einer Beschleunigung um 2,2 Sekunden erreichte Mozilla eine Steigerung der Download-Konversion um 15,4 Prozent, das sind schätzungsweise 60 Millionen mehr Firefox-Downloads im Jahr (Quelle: Mozilla). Yahoo steigerte den Datenverkehr auf der Seite um 9 Prozent pro 400 Millisekunden Beschleunigung (Quelle: Yahoo).

### 3.3. Werbeflächen

Für Werbeportale mit bezahlter Werbeschaltung bedeuten schnellere Ladezeiten bares Geld. Bei einer Buchungszeit nach Sekunden können Sie entsprechend mehr Bannerwerbung verkaufen: viermal mehr Seiten bedeuten viermal mehr Advertorials.

Gleichzeitig wird die Aufmerksamkeit der Besucher signifikant erhöht: Ein Besucher, der 8 Sekunden lang auf den Seitenaufbau warten muss, betrachtet nur ein Prozent der gesamten Zeit die Werbung auf der Zielseite. Dagegen sieht sich ein Besucher, dem die Seite sofort angezeigt wird, 20 Prozent seiner

Verweilzeit die Werbefläche an (Quelle: Jakob Nielsen).

### 3.4. Intranet

In barer Münze zahlt es sich auch für große Konzerne aus, wenn sie in die Optimierung des Intranets investieren: Reisekostenabrechnung und Projektmanagement, Terminplanung und Ressourcenverwaltung werden über das Intranet abgerechnet. Langsame Ladezeiten der Webseiten haben Frust und Unlust der Mitarbeiter zur Folge. Das macht sich nicht nur in mangelnder Akzeptanz dieser Anwendungen bemerkbar, sondern senkt auch insgesamt deren Motivation.

Schnelle und funktionale Seiten dagegen erhöhen die Akzeptanz und interne Anwendungen werden zeitnah ausgeführt, statt sie lange hinauszuzögern. Die Mitarbeiter sparen Arbeitszeit, die sie für produktive Verbesserungen oder verkaufsfördernde Maßnahmen nutzen können.

## 4. Webperformance und die Technik

Natürlich gibt es unzählige Möglichkeiten, die Geschwindigkeit der Seiten zu verbessern: von der Komprimierung der Inhalte, Formulare und Bilder über Datenbanken bis hin zu Software-Optimierung. An dieser Stelle möchte ich daher nur einige technische Maßnahmen skizzieren, mit deren Hilfe die Geschwindigkeit des Seitenaufbaus verbessert werden kann, und ihre Grenzen aufzeigen.

### 4.1. Ressourcen vergrößern

Die Erhöhung der Datenübertragungsrate oder Datentransferrate (umgangssprachlich: Bandbreite) ist die naheliegendste Lösung, um die Geschwindigkeit zu erhöhen. Wenn die mögliche Datenmenge, die über einen digitalen Kanal übertragen werden kann, vergrößert wird, fließen mehr Daten gleichzeitig, die Übertragung wird schneller.

Die maximal mögliche Kapazität und

die Latenzzeit (Antwortverzögerung) sind das Maß für die Leistungsfähigkeit des Kanals. Ein Kanal kann beispielsweise eine Verbindung im Rechnernetz oder zum Provider oder eine Schnittstelle zu einem Datenspeicher sein. Während zu Beginn des Datenzeitalters Übertragungsraten von 50 bit/s genügten, rechnet man heute in Gbit/s.

Eine andere Möglichkeit besteht darin, die zur Verfügung stehende Bandbreite aufzuteilen und dynamisch für diverse Verbindungen zu nutzen. Das bedeutet meist eine Priorisierung bestimmter Anwendungen (etwa VoIP) oder Nutzergruppen. Allerdings kann dies zu Engpässen führen, unter denen die nicht priorisierten Anwendungen leiden.

Selbstverständlich können Sie auch immer mehr, immer bessere und immer größere Server aufsetzen. Sie können viele Server zusammenschließen und die Anfragen mittels Lastverteilung (Load Balancing) auf mehrere parallel arbeitende Systeme verteilen.

Eine aktuelle Lösung bieten Content Distribution Networks (CDN): Das sind Netze lokal verteilter und über das Internet verbundener Server, die zusammenarbeiten, um Anfragen (Requests) von End-Nutzern nach Inhalten (Content) möglichst ökonomisch zu befriedigen. Dabei werden die Daten so vorgehalten (Caching), dass die Auslieferung möglichst schnell geht (performanceoptimiert) oder möglichst wenig Bandbreite verbraucht (kostenoptimiert) oder beides.

Gemeinsam ist den hier beschriebenen Lösungen, dass sie sich einen Wettlauf mit dem Datenfluss liefern und wie dieser immer größer, immer breiter, immer schneller werden wollen. Offensichtlich kann man aber die Geschwindigkeit auf diese Weise nicht unendlich erhöhen.

Ein weiteres Problem liegt darin, dass die Webperformance bezogen auf die Geschwindigkeit von mehreren Fak-

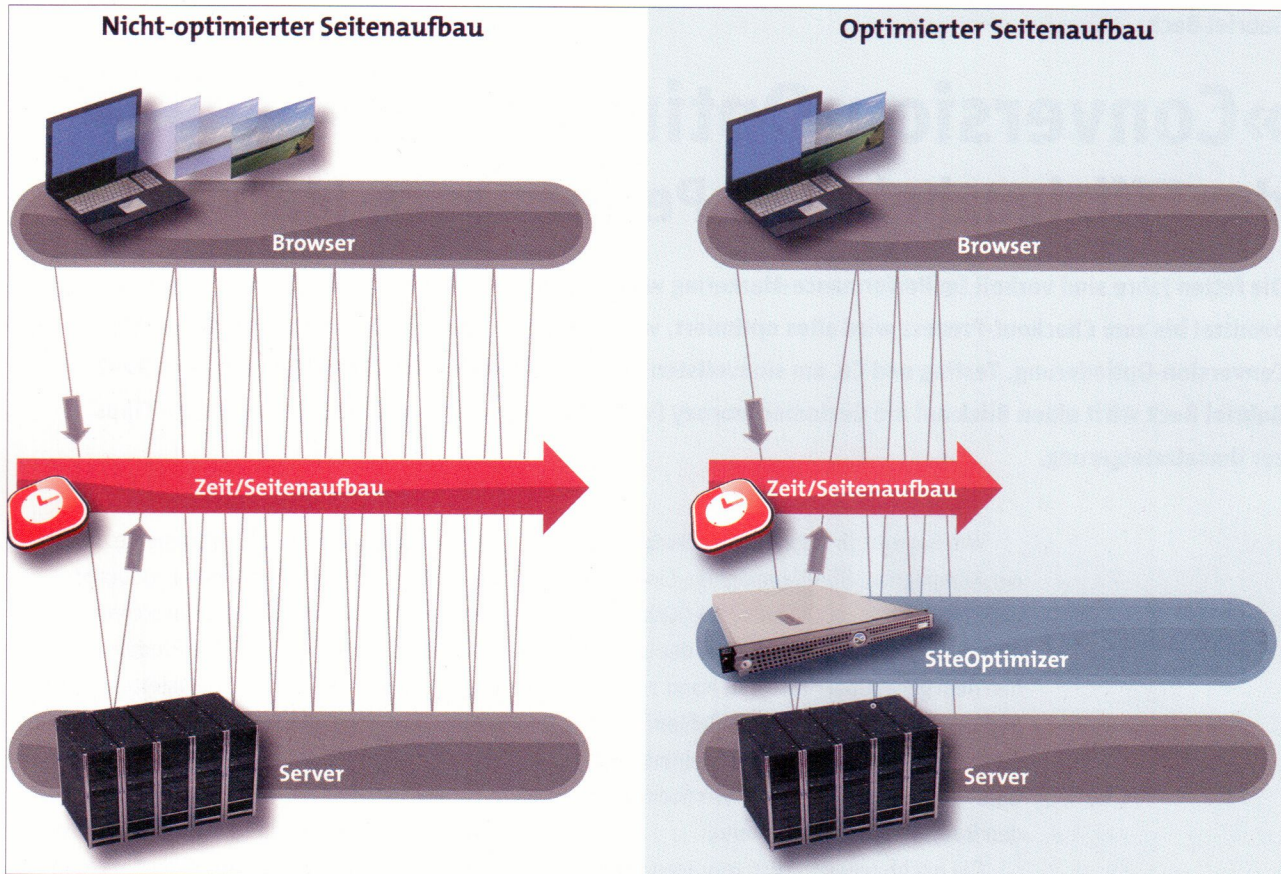


Abbildung 1: Beachtung, Aktivität und Konversion lassen sich anhand diverser Kennzahlen messen (Quelle: Joshua Bixby „Web Performance Today“)

toren abhängt: Außer der Serverleistung und der Anzahl der Anfragen spielen die Heterogenität der Browser und die Nutzererwartung eine Rolle. Die letzteren beiden Faktoren kann der Anbieter aber nicht beeinflussen: Welche Browser nutzen die Anwender? Und wie hoch ist ihre Erwartung an die Antwortzeit?

Normalerweise stellt ein Anwender eine „Anfrage“, diese wandert zum Server, die Antwort kommt zurück. Beim nächsten Klick (der nächsten Anfrage) wird erneut ein solcher „Roundtrip“ gestartet usw. Bei jedem Roundtrip werden 10 bis 20 Prozent der Ladezeit einer Website am Back-End verbraucht (dem Server), aber 80 bis 90 Prozent am Front-End (der Eingabestelle, also dem Computer des Anwenders).

Wie bereits erwähnt, handelt es sich um Millisekunden. Aber diese summieren sich zu einer Wartezeit in Sekunden, die Anwender nicht mehr tolerieren wollen.

#### 4.2. Ein neuer Ansatz: „lernende“ Algorithmen

Das Manko der technischen Lösungsansätze ist, dass sie das eigentliche Problem nicht berühren: die ständige Erhöhung der Datenmenge, deren Übertragung immer schneller bewältigt werden muss, und die Ladezeit am Front-End, der Ausgabestelle. Wie kann man aber die Datenmenge selbst reduzieren, sodass sie mit den vorhandenen Hardware-Ressourcen bewältigt werden kann? Und wie kann die Ausgabe der Daten am Front-End beschleunigt werden?

Ein neuer Ansatz nutzt die Erkenntnis, dass das Front-End die meiste Ladezeit verbraucht, und schaltet einen „Optimizer“ dazwischen, der den Aufbau der Webseite im Browser des Anwenders beschleunigt.

Die wichtigsten Merkmale sind:

» Dynamische Analyse der http-Anfragen und Antworten (Request/Response)

- » Übertragen und Speichern von Ressourcen im Browsercache des Anwenders
- » Für ASP.NET-Seiten: Ersetzen des ASP.Net View-State durch einen Token
- » Konsolidieren von Grafiken gleichen Typs (gif, jpg, png etc.) in eine Bibliotheksdatei sowie von JavaScript-Dateien und CSS-StyleSheet-Dateien in eine Ressource
- » Beantworten der SSL-Anfragen (Secure Socket Layer, https) zur Website

Indem der Optimizer ein bestimmtes Anwenderverhalten für die Website vorausberechnet, kann er Ressourcen für wahrscheinliche Folgeseiten bereits im Voraus senden und wiederkehrende Elemente länger zwischenspeichern. Dadurch reduziert er die Anfragen des Browsers an den Server (Roundtrips) und die Datenmenge der zu sendenden Daten um bis zu 80 Prozent. ¶