

Übungsblatt 2

Aufgabe 2: HTTP

Beim Abruf einer nicht existenten Webseite erscheinen je nach Webbrowser verschiedene Seiten. Erklären Sie dieses Verhalten! Woher stammt der jeweilige Seiteninhalt?

Antwort:

Je nach Einstellung des Servers werden unterschiedliche 404 Seiten angezeigt:

- Server meldet 404 Fehler, ansonsten nichts konfiguriert
 - Browserspezifische Anzeige einer 404-Seite

- Server meldet 404 Fehler, Erklärungsdaten werden übermittelt
 - Eintrag in .htaccess-Datei, z.B.:
ErrorDocument 404 http://www.xyz.de/index.php?p=404
⇒ Die festgelegte Seite wird zusätzlich zum 404 mit übermittelt
 - auch für andere Fehler-Codes möglich
 - Gemäß RFC erlaubt: <http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html>
 - 10.4 Client Error 4xx
The 4xx class of status code is intended for cases in which the client seems to have erred. Except when responding to a HEAD request, the server SHOULD include an entity containing an explanation of the error situation, and whether it is a temporary or permanent condition. These status codes are applicable to any request method. User agents SHOULD display any included entity to the user.
 - Opera zeigt z.B. dies auch alles an, Internet Explorer tanzt aus der Reihe und zeigt eine eigene Fehlerseite an, wenn sie > 512 Bytes groß ist
 - Siehe dazu auch:
 - <http://support.microsoft.com/default.aspx?scid=kb;DE;294807>
 - <http://www.404-error-page.com/404-error-page-too-short-problem-microsoft-ie.shtml>

Aufgabe 3: Cache-Statistiken (1)

Im Jahr 2006 wird aufgrund der Konsolidierungsmaßnahmen der Universität die Nutzung des Internets für die Abteilungen kostenpflichtig. Das KIZ bezahlt pro GB Übertragungsvolumen eine Gebühr von 5 Cent. Um das eigene Budget schonen, wird ein transparent web proxy zum Einsatz gebracht.

Der HTTP-Anteil des Traffics einer Abteilung beträgt 70%. Die Utilisation des Web-Proxys beträgt 25%. Die Abteilungen erhalten ein Freivolumen von 400 GB Traffic pro Monat.

- Welche Kosten entstehen dem KIZ, wenn eine Abteilung das Freivolumen voll ausschöpft?
- Die Abteilungen bezahlen oberhalb des Freivolumens 20 Cent pro GB. Ab welcher Trafficmenge ist dies ein positives Geschäft für das KIZ?

(Sämtliche Umstände und Preisangaben sind frei erfunden und werden – hoffentlich – nicht der Realität entsprechen)

Antwort:

- Nutzung: 400 GB
HTTP-Anteil des Traffics: 70% von 400 GB → 280 GB
Durch Web-Proxy: 25% von 280 GB → 70 GB
⇒ tatsächlicher Traffic: 400 GB – 70 GB = 330 GB
⇒ Kosten: 330 GB · $\frac{0,05 \text{ €}}{\text{GB}}$ = 16,50 €

- Gesuchter Traffic: t [GB]
 $[t - (t \cdot 0,7 \cdot 0,25)] \cdot 0,05 = (t - 400) \cdot 0,2$
⇔ $[t - (0,175 \cdot t)] \cdot 0,05 = 0,2 \cdot t - 80$
⇔ $0,04125 \cdot t = 0,2 \cdot t - 80$
⇔ $0,15875 \cdot t = 80$
⇔ $t = \underline{503,94}$ [GB]

Ab ca. 504 GB ist es ein positives Geschäft für das KIZ

Aufgabe 4: Cache-Statistiken (2)

Die Statistiken des Web-Proxys der Universität Ulm haben sich im Laufe der Zeit stark verändert:

- 1995: Hit rate ca. 75%
- 2003: Hit rate ca. 35%

Geben Sie mögliche Gründe für diese Entwicklung.

Antwort:

- Anstieg dynamischer Web-Inhalte, die nicht oder nur schlecht gecached werden können (z.B. Diskussionsforen, Gästebücher, ...)