

## Web Engineering - 2. Übungsblatt

Name: Oliver Knörzer  
E-Mail: novil@gmx.de

2) - Statt nur den HTTP-Statuscode "404 File Not Found" zurückzugeben, gibt es die Möglichkeit, den Webserver so zu konfigurieren, dass eigene Fehlerseiten aufgerufen werden (auch für andere HTTP-Statuscodes möglich). Bei Apache-Webservern geschieht dies mit der ErrorDocument-Direktive, die in der Form "ErrorDocument <StatusCode> <Action>" aufgebaut ist, z. B. also mit "ErrorDocument 404 /errors/my\_error.html", wobei als Statuscode auch weiterhin 404 zurückgegeben wird.

- Wenn keine eigenen Fehlerseiten verwendet werden, verwendet der Webbrowser des Clients die eigene Standardfehlerseite, die je nach verwendetem Webbrowser unterschiedlich aussehen kann. Laut der Apache-Dokumentation verwendet der Internet Explorer diese auch, wenn die Größe der angegebenen eigenen Fehlerseite "zu klein" sei.

3)  $\text{HTTP-Traffic}_{\text{Gesamt}} = 0,7 * 400 \text{ GByte} = 280 \text{ GByte}$   
 $\text{HTTP-Traffic}_{\text{Cache}} = 0,25 * 280 \text{ GByte} = 70 \text{ GByte}$   
 $\text{Traffic} = 400 \text{ GByte} - 70 \text{ GByte} = 330 \text{ GByte}$   
 $\text{Kosten} = 330 \text{ GByte} * 0,05 \text{ €/GByte} = \mathbf{16,50 \text{ €}}$

$(x - 400 \text{ GByte}) * 0,20 \text{ €/GByte} > (x - 0,25 * 0,7 * x) * 0,05 \text{ €/GByte}$   
 $\Leftrightarrow 0,2 x \text{ €/GByte} - 80 \text{ €} > 0,04125 x \text{ €/GByte}$   
 $\Leftrightarrow 0,15875 x \text{ €/GByte} > 80 \text{ €}$   
 $\Leftrightarrow x > 503,937... \text{ GByte} \Rightarrow \mathbf{504 \text{ GByte}}$

4) Für diese Entwicklung gibt es vor allem zwei wichtige Gründe:

- Das World Wide Web ist derart groß geworden, dass die Wahrscheinlichkeit, dass zwei verschiedene Benutzer zumindest die selbe Website oder gar die selbe Seite aufrufen, immer kleiner wird. Dass sich eine aufgerufene Seite im Cache befindet, wird daher immer kleiner.

- Im World Wide Web sind heute immer mehr dynamische statt statische Inhalte anzutreffen. So werden die Ergebnislisten von Suchmaschinen, die Einträge in Foren und andere Inhalte erst bei der Anfrage dynamisch generiert oder aus Datenbanken geladen. Derartige Inhalte zu cachen lohnt sich aufgrund ihrer kurzen Halbwertszeit nicht oder es wird vom Webserver von vornherein mit `no-cache`-Angaben (o. ä.) unterbunden.