



## 11. Aufgabenblatt zum Internet-Praktikum

### Aufgabe 1: (100 Punkte) *Untersuchung des Effektes von Proxycaches auf Webverkehr*

In dieser Aufgabe soll der Effekt von Proxycaches auf Webverkehr untersucht werden. Zu diesem Zweck sind unter `~inetprak/daten/11.uebung/` zwei Traces namens `A.trace.gz` und `B.trace.gz` abgelegt.

Beide Traces wurden von nicht-persistenten Verbindungen angefertigt. In diesem Falle existieren maximal zwei parallele TCP-Verbindungen zwischen jedem Client und Server.

Die Trace-Datei enthält für jeden Zugriff auf eine URL Informationen zu den folgenden Events: TCP-Verbindungsaufbau, HTTP-Request, HTTP-Response, HTTP-Daten, und TCP-Verbindungsabbau. Zu jedem Event werden auf Client-Seite die Startzeit und die Endzeit in Sekunden geloggt. Daraus ergibt sich das folgende Logformat (jeweils in einer Zeile):

<code>&lt;URL ID&gt;</code>	ID der URL
<code>&lt;size&gt;</code>	Größe der vom Server gelieferten Daten in Bytes
<code>&lt;client&gt;</code>	ID des Clients
<code>&lt;server&gt;</code>	ID des Servers
<code>&lt;SYN start&gt; &lt;SYN end&gt;</code>	Timestamp Anfang & Ende des Verbindungsaufbaus
<code>&lt;request start&gt; &lt;request end&gt;</code>	Timestamp Anfang & Ende des Senden des HTTP Requests
<code>&lt;response start&gt; &lt;response end&gt;</code>	Timestamp Anfang & Ende des Empfangen der HTTP Response
<code>&lt;data start&gt; &lt;data end&gt;</code>	Timestamp Anfang & Ende des Empfangen der HTTP Daten
<code>&lt;FIN start&gt; &lt;FIN end&gt;</code>	Timestamp Anfang & Ende des Verbindungsabbaus

- Informiere Dich über die Funktionsweise eines Caching-Proxys für Webverkehr. (Eine Idee, wie Proxys funktionieren solltest Ihr seit dem zweiten Übungsblatt haben.) Erkläre kurz die Vorteile und Nachteile eines Caching-Proxys sowie den Unterschied zwischen HTTP mit non-persistent, persistent und pipelined Verbindungen.
- Angenommen, diese HTTP-Anfragen würden nun stattdessen über einen Caching-Proxy gestellt. Was wäre die maximal mögliche Cache-Hit-Rate in % Bytes und in % Zugriffe für die beiden Traces? Siehst Du das Verwenden eines Caching-Proxys als gerechtfertigt an? Begründe Deine Aussage.
- Berechne für jeden Request die Zeitdauer, bis die Nutzdaten<sup>1</sup> am Client angekommen sind *ohne* Proxy, d. h. so, wie sie im Trace beobachtet worden sind.
- Simuliere, was passieren würde, falls ein Proxy zwischen den Clienten und den Servern eingefügt wird. Die Round-Trip-Time zwischen Client und Proxy beträgt 100 ms, die RTT zwischen Proxy und Server ist unverändert, wie beim direkten Zugriff. Die Bandbreite zwischen Client und Proxy beträgt 32kBit/s pro TCP-Verbindung; die zwischen Proxy und Server ist die selbe wie in den Tracefiles zwischen Client und Server. Dein Skript/Programm soll pro Anfrage ausgeben:
  - die ID der angefragten URL
  - ob die Seite aus dem Cache bedient wurde
  - die Dauer, bis die Nutzdaten am Client angekommen sind *ohne* Proxy
  - die Dauer, bis die Nutzdaten am Client angekommen sind *mit* Proxy

---

<sup>1</sup>Nutzdaten sind hier nur die Daten, die die Applikation braucht um die Seite darzustellen. Sie hören genau *vor* dem FIN auf.

Hierbei sollen non-persistent Verbindungen verwendet werden. Jeder Client und jeder Proxy darf beliebig viele TCP-Verbindungen verwenden.

- (e) Interpretiere die somit erhaltenen Daten für die beiden Traces.
- (f) Ändere die Simulation so ab, dass zu jedem Zeitpunkt jeder einzelne Client maximal eine *nichtpersistente* TCP-Verbindung zum Proxy und der Proxy zu jedem einzelnen Server maximal eine *nichtpersistente* TCP-Verbindung halten darf. Erläutere die Unterschiede zu den vorherigen Ergebnissen.
- (g) Ändere die Simulation so ab, dass zu jedem Zeitpunkt jeder einzelne Client maximal eine *persistente* TCP-Verbindung zum Proxy und der Proxy zu jedem einzelnen Server maximal eine *persistente* TCP-Verbindung halten darf. Erläutere die Unterschiede zu den vorherigen Ergebnissen.

Beim Modifizieren der Traces darf man annehmen, dass alle Objekte unabhängig voneinander geladen werden, d. h.: Die Startzeit für eine Anfrage des Benutzers ändert sich durch die Simulationen nicht.

Abzugeben sind jeweils:

- Deine Skripte
- ASCII-Plaintextfiles mit den geforderten Erklärungen

**Details zur Abgabe der Aufgaben: siehe FAQ**  
(unterhalb <http://www.net.in.tum.de/teaching/SS05/inetprak/>).  
**Abgabe:** bis Dienstag, den 5. Juli 2005, 23:59h s. t.