

SYR1 Réseau – TP 1

Étude du protocole HTTP

Léo Noël-Baron & Thierry Sampaio

24/11/2015

Résumé

Ce TP a pour objectif d'étudier quelques caractéristiques du protocole HTTP, en s'appuyant sur différentes pages pertinentes stockées sur le serveur `anubis.istic.univ-rennes1.fr/TP_syr/`. Le trafic est analysé grâce au logiciel gratuit WireShark.

Requête d'une page légère sans objets

Le chargement de la page `fichier1.html` dans le navigateur se traduit dans la capture WireShark par une requête GET et une réponse du serveur.

Les informations contenues montrent que notre navigateur utilise le protocole HTTP en version 1.1 et accepte la locale fr-FR. Notre poste a pour IP 148.60.3.133 (réseau local), et le serveur (Apache en version 2.4.10) utilise aussi HTTP 1.1. Il retourne un code 200 OK qui signifie que la page a bien été transmise ; elle a été modifiée pour la dernière fois le 18/06/2015 et pèse 126 octets.

GET conditionnel

Le chargement (avec un cache vide) de l'adresse `fichier2.html` génère une requête GET sans champ If-Modified-Since ; la réponse du serveur contient le texte explicite de la page demandée, visible en dernière ligne (champ Line-based text data).

Le rechargement lance une nouvelle requête GET qui inclut cette fois la date de dernière modification de la page demandée dans le champ If-Modified-Since. Le serveur retourne alors un simple code 304 Not Modified, qui indique au navigateur que sa version en cache de la page est à jour et peut être implicitement rechargée.

Requête d'une page lourde

Une seule requête GET est envoyée pour demander la page `fichier3.html`, mais la réponse du serveur se fait sur deux paquets TCP, la page demandée étant trop lourde pour n'en utiliser qu'un. Le premier paquet TCP contient un message HTTP portant l'en-tête 200 OK ; le deuxième ne contient pas de message HTTP mais seulement les données de continuation du premier.

Requête d'une page référençant des objets

La page à l'adresse `fichier4.html` contient des images stockées sur différents serveurs. Lorsqu'on la charge, le navigateur envoie tout d'abord une requête GET au serveur, puis après sa réponse (qui contient la page sans les images), en renvoie deux en parallèle pour récupérer les images (soit au total trois GET vers les IP respectives 148.60.12.25, 129.20.126.100 et 129.20.134.3).