CS433 Project Presentation

Smart Proxy Server

Group Members :

Aslaliya Juhil : 20110030
Meet Vankar : 20110112
Tejas Parmar : 20110125
Patel Prey   : 20110132

Smart Proxy Server

- A simple proxy server acts as an intermediary with a cache.
- It stores the site data of frequently visited sites into its cache and displays the cached data instead of requesting the server for site data.
- This decreases the page loading time as well as it also decreases the load on the web-server.
- A "smart" proxy server has many features implemented on top of the proxy server functions.
- A smart proxy server can have features like URL filtering, Ad blocking, Firewall, User Authentication, Blacklisting, Whitelisting, and more.

Implemented Features

We have implemented the following features in our smart proxy server:

- Cache and its Cache-Invalidation Mechanism
- URL Blacklisting for Clients (Adding & Removing)
- URL Blocking for Malicious Websites (Security)
- Advertisement URL Blocking (Ad-blocking)
- Threaded Processing for Requests (Performance)
- Responsive Proxy (Sends different response as per situation)
- Logger to Log the Events

Cache

- Cache is maintained to quickly send previously requested responses.
- If requested response is not found in cache, the proxy requests web-server for response, and then stores it in the cache, with timestamp.
- To avoid filling up the entire memory, a timeout-based cache-invalidation policy is implemented.
- It removes the entries that are there in the cache for more time than specified timeout time duration.
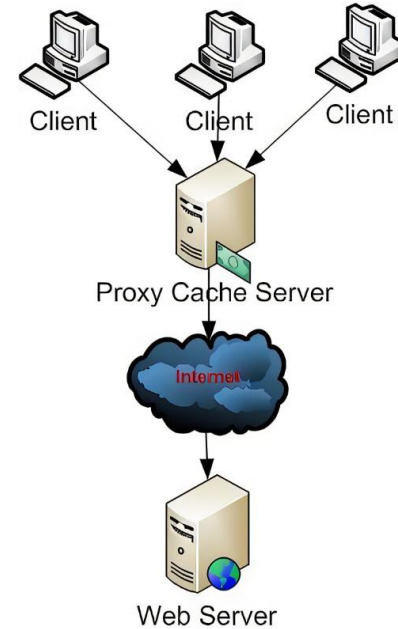
Ad-Blocking

- Proxy server acts as a URL filter and prevents ads from reaching the user's browser.
- For example URL contains text/domains like **/ad/dfp/*** or **.php/ad/** etc**.** would not show up on the webpage.
- Utilizing well-known ad domains containing text file Easylist to cover most of the ads.
- Since proxy server for HTTPs is quite complex to implement, Ads on HTTP only website would be blocked.
- Current method might struggle with dynamically loaded ads, encrypted ad traffic & maintaining up-to-date list of ad domains.

Blacklisting

- Using dictionary to track blacklisted sites by each user.
- Allowing to block or unblock site from search bar.
- For example https://github.com/ can be blocked just by searching https://_block_github.com/ or can be unblocked by https://_unblock_github.com/
- _ (underscores) are not valid character for domain name.
- Using _ to differentiate between command to block/unblock and request to search website.

Thanks!