

Analyse de Logs Apache

Manuel d'utilisation

I. Présentation générale de l'application

Cet outil en ligne de commande permet d'analyser les fichiers de logs Apache afin d'extraire des statistiques sur les requêtes effectuées sur un site web. Il offre également la possibilité de générer des représentations graphiques sous forme de graphes de parcours. L'utilisateur peut filtrer les requêtes selon divers critères, notamment la date et l'extension des fichiers. Les différentes fonctionnalités sont détaillées dans la suite de ce manuel.

A travers ce projet, on a pu développer des fonctionnalités qui consistent en:

- ☒ Lecture et analyse des logs Apache.
- ☒ Affichage des 10 documents les plus consultés.
- ☒ Génération d'un fichier .dot pour une visualisation GraphViz des parcours utilisateurs.
- ☒ Exclusion des fichiers statiques (images, CSS, JS) de l'analyse.
- ☒ Filtrage des requêtes par plage horaire.
- ☒ Tests automatisés avec un framework fourni.
- ☒ Gestion des erreurs et logs de diagnostic.

II. Prérequis pour utiliser l'application

1. Installation

Vous pouvez obtenir le projet avec deux manières:

1> Clonage avec git avec cette commande:

`git clone <URL-du-référentiel>`

`cd <nom-du-projet>`

2. Téléchargement en format ZIP

1. Cliquez sur le bouton "Code" sur la page du dépôt GitHub.
2. Sélectionnez "Download ZIP".
3. Extrayez le fichier dans un répertoire de votre choix.

2. Compilation

Il est recommandé de s'assurer d'avoir un compilateur C++ compatible avec la norme C++11 ou supérieure sous Windows ou Linux.

Il est nécessaire de se mettre dans le dossier src afin de créer l'exécutable en utilisant la commande **make**: l'exécutable sera copié automatiquement dans le dossier bin du framework des tests. Nous avons pensé aux droits d'accès, ainsi, la fonction make du makefile vous accorde tous les droits sur tous les fichiers.

L'exécution du programme nécessite toujours au minimum une option qui correspond au chemin qui référence le fichier log à analyser.

Voici un exemple d'utilisation : **`./analog nomfichier.log`**

III. Utilisation des différentes options de l'application

Les différentes options doivent être précisées avant le nom du fichier log à analyser.

Voici la syntaxe générale: **`./analog [options] nomfichier.log`**

Options disponibles :

-g nomfichier.dot: produit un fichier au format GraphViz du graphe analysé.

-e: permet d'exclure tous les documents qui ont une extension de type image, css ou javascript.

-t heure : permet de ne prendre en compte que les hits qui sont dans le créneau horaire correspondant à l'intervalle [heure,heure+1]

Sans option: affiche les 10 documents les plus consultés.

Exemples de commandes:

1. Afficher les 10 documents les plus consultés : **`./analog logs_apache.log`**
2. Générer un fichier GraphViz : **`./analog -g output.dot logs.txt`**
3. Exclure les fichiers statiques et générer un graphe : **`./analog -e -g parcours.dot logs_apache.log`**
4. Analyser uniquement les hits entre 12h et 13h : **`./analog -t 12 logs_apache.log`**

5. Générer un graphe des parcours : ***[./analog -g parcours.dot logs_apache.log](#)***

Remarques supplémentaires :

Pour ***-g nomfichier.dot***, le nom de fichier doit être indiqué juste après ‘-g’ et le fichier doit pouvoir être ouvert sinon le programme produira une erreur.

Pour ***-t heure***, l’heure doit être indiquée juste après ‘-t’ et ce doit être un nombre entier compris entre 0 et 23 sinon le programme produira une erreur.

Les options sont supportées dans n’importe quel ordre mais il est obligatoire de toujours préciser le chemin du fichier log à analyser après avoir précisé l’ensemble des options.

Le programme est conçu pour fonctionner avec des fichiers de logs Apache au format standard.

Un fichier de log anonyme est disponible dans /tmp/anonyme.log pour les tests.

Un framework de test est fourni pour valider les résultats .

L’outil GraphViz doit être installé pour permettre la génération des fichiers .dot.