INF2132 : Systèmes d'Exploitation

TP4 - Manipulation de Texte avec SED

Nom de l'enseignant :
Pr. Ilias Tougui

Nom de l'assistant :
Pr. Yasser Aderghal

7 octobre 2025

Table des matières

Lisez attentivement cette page

1 Contexte et Probélmatique

Vous êtes administrateur système dans une entreprise de services informatiques, "Tech Consult Morocco". L'entreprise héberge plusieurs serveurs web pour différents clients et génère quotidiennement des logs volumineux qui doivent être traités et analysés.

Le serveur principal a subi une panne, et vous devez agir rapidement :

* Analyser les logs d'accès pour identifier les causes de la panne
* Nettoyer et formater les fichiers de logs pour les équipes de sécurité

Cette situation nécessite une manipulation rapide et efficace de texte, impossible à faire manuellement sur des fichiers de plusieurs "milliers" de lignes.

2 Objectifs pédagogiques

Ce travail pratique vise à maîtriser l'utilisation de la commande `sed` (Stream EDitor) pour manipuler et analyser des fichiers de logs issus d'un serveur web. Vous apprendrez à :

* Rechercher, modifier, extraire et supprimer des lignes ou motifs spécifiques
* Filtrer des informations pertinentes dans des logs (adresses IP, dates, erreurs, statuts HTTP)
* Appliquer des scénarios réalistes d'administration système
* Maîtriser les expressions régulières dans un contexte professionnel

3 Pré-requis et matériel

* Système Linux/Unix avec la commande `sed` installée
* Fichiers fournis : `access.log` et `error.log`
* Connaissance de base des expressions régulières
* **Aide-mémoire `sed` fourni en annexe**

Pour les étudiants qui n'ont pas installé WSL ni de machine virtuelle Ubuntu, il est possible d'utiliser temporairement `https://cocalc.com/features/terminal` pour réaliser les exercices de TPs. CoCalc offre un terminal Linux complet accessible directement dans le navigateur.

Lisez attentivement cette page

4 Annexe : Aide-mémoire `sed`

Syntaxe de base

```
sed [options] 'commande' fichier
```

* Flag `g` : remplacement global (toutes les occurrences)
* Flag `i` : insensible à la casse
* Option `-i` : modification directe du fichier
* Suppression : `sed 'Xd' fichier` (ligne X) ou `sed '/motif/d' fichier`
* Insertion : `sed 'Xi\texte' fichier` (avant la ligne X)
* Ajout : `sed 'Xa\texte' fichier` (après la ligne X)
* Intervalle : `sed 'X,Ys/motif/remplacement/' fichier`

Conseils

1. Testez d'abord vos commandes sans l'option `-i` pour vérifier le résultat
2. Faites des copies de sauvegarde de vos fichiers avant de les modifier
3. Utilisez des guillemets simples pour éviter l'interprétation par le shell
4. Pour les caractères spéciaux, utilisez le backslash \ pour les échapper

5 Préparation de l'Environnement Linux

Durée : 20 min, Note : 5 points

En utilisant les commandes : `pwd`, `ls`, `cd`, `mkdir`, `touch`, `nano` et `cat`, suivez les consignes ci dessous pour créer un environnement de travail structuré et générer les fichiers de logs nécessaires pour preparer votre environnement de travail.

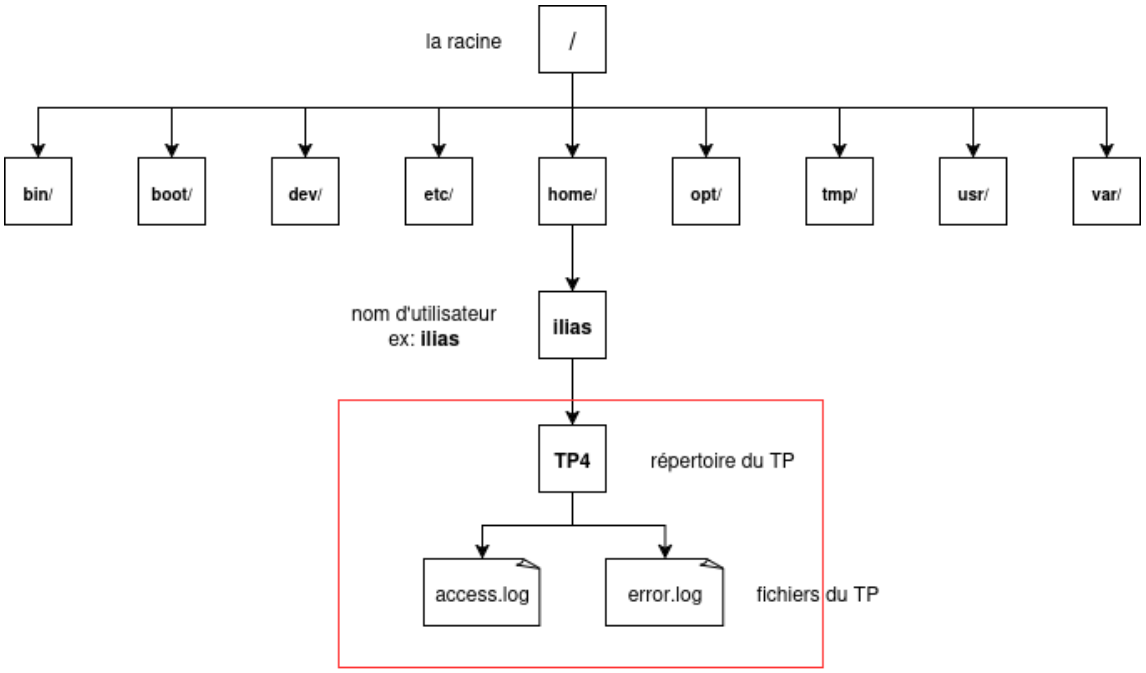

FIGURE 1 – Structure de l'environement de travail.

Fichier : access.log

Le fichier `access.log` contient les enregistrements d'accès au serveur web. Chaque ligne représente une requête HTTP avec les informations suivantes : adresse IP, nom d'utilisateur, date/heure, méthode HTTP, chemin de la ressource, code de statut, taille de la réponse et user-agent.

```
# access.log
192.168.1.100 - admin [25/Sep/2025:10:15:32] "GET /index.php HTTP/1.1" 200 2834 "Mozilla/5.0"
10.0.0.50 - user1 [25/Sep/2025:10:16:45] "POST /login.php HTTP/1.1" 302 156 "Mozilla/5.0"
172.16.0.25 - guest [25/Sep/2025:10:17:22] "GET /error.php HTTP/1.1" 404 287 "Chrome/91.0"
192.168.1.100 - admin [25/Sep/2025:10:18:33] "GET /config.php HTTP/1.1" 403 98 "Mozilla/5.0"
203.45.67.89 - hacker [25/Sep/2025:10:19:44] "GET /admin.php HTTP/1.1" 401 156 "Wget/1.20"
10.0.0.51 - user2 [25/Sep/2025:10:20:15] "GET /dashboard.php HTTP/1.1" 200 5432 "Firefox/89.0"
192.168.1.200 - operator [25/Sep/2025:10:21:30] "POST /update.php HTTP/1.1" 500 234 "Mozilla/5.0"
172.16.0.30 - visitor [25/Sep/2025:10:22:18] "GET /contact.php HTTP/1.1" 200 1876 "Safari/14.0"
```

Fichier : error.log

Le fichier `error.log` contient les messages d'erreur, d'avertissement et d'information du serveur. Chaque ligne commence par une date/heure au format `[25/Sep/2025:10:15:00]`

suivie d'un niveau de gravité (ERROR, WARNING, INFO) et d'un message descriptif.

```
# error.log
[25/Sep/2025:10:15:00] ERROR: Database connection failed - host: db.techconsult.ma
[25/Sep/2025:10:16:30] WARNING: High memory usage detected - 85% utilization
[25/Sep/2025:10:17:45] ERROR: Authentication failure for user: hacker@malicious.org
[25/Sep/2025:10:18:12] INFO: Backup process completed successfully
[25/Sep/2025:10:19:33] ERROR: File permission denied - /var/www/config/database.ini
[25/Sep/2025:10:20:55] WARNING: Disk space low on /var partition - 92% full
[25/Sep/2025:10:21:20] ERROR: SQL injection attempt detected from IP: 203.45.67.89
[25/Sep/2025:10:22:40] INFO: System restart initiated by admin user
```

Exercice

1. Déplacez-vous dans le répertoire racine (/) à l'aide de la commande `cd`.
2. Changez ensuite votre répertoire de travail pour votre répertoire personnel (~).
3. Créez un nouveau répertoire nommé `TP4` en utilisant la commande `mkdir`.
4. Accédez au répertoire `TP4`.
5. Créez deux fichiers de logs (`access.log` et `error.log`) à l'aide de la commande `touch`.
6. Vérifiez la création des fichiers en utilisant la commande `ls`.
7. Remplissez chaque fichier avec le contenu indiqué en haut en utilisant l'éditeur `nano`.
8. Enregistrez vos modifications dans `nano` à l'aide des touches `Ctrl + O`, puis `Entrée`, et fermez l'éditeur avec `Ctrl + X`.
9. Affichez le contenu de chaque fichier dans le terminal avec la commande `cat`.
10. Comment afficher les trois premières lignes de chaque fichier ? Quelle commande faut-il utiliser ?
11. Comment afficher les dernières lignes de chaque fichier ? Quelle commande faut-il utiliser ?
12. Lisez attentivement le contenu de chaque fichier.
13. Assurez-vous de bien comprendre les manipulations effectuées avant de demander au professeur de valider votre travail.

6 Manipulation du fichier error.log

Durée : 30 min, Note : 5 points

1. Écrivez la commande `sed` pour remplacer la première occurrence du mot `ERROR` par `ERREUR` sur chaque ligne (sans modifier le fichier).

2. Écrivez la commande pour remplacer **toutes** les occurrences de `ERROR` par `ERREUR` dans le fichier (affichage uniquement, sans modification).

3. Modifiez directement le fichier `error.log` pour remplacer tous les `WARNING` par `ALERTE` (option `-i`).

4. Remplacez le mot `INFO` par `INFORMATION` uniquement sur la ligne 4 du fichier (avec modification directe).

5. Remplacez tous les `ERROR` par `CRITICAL` uniquement pour les lignes 1 à 5 (avec modification directe).

6. Remplacez de manière insensible à la casse tous les mots commençant par `error` (ERROR, error, Error) par `ERREUR` dans tout le fichier.

7. Supprimez toutes les lignes contenant le mot `INFO` du fichier (avec modification directe).

8. Supprimez les lignes 3 à 5 du fichier `error.log` (avec modification directe).

9. Insérez la ligne `[25/Sep/2025:10:14:00] INFO: System startup initiated` avant la première ligne du fichier.

10. Ajoutez la ligne `-- End of error log --` après la dernière ligne du fichier.

7 Manipulation du fichier access.log

Durée : 30 min, Note : 5 points

1. Remplacez toutes les occurrences de `GET` par `OBTENIR` dans le fichier (affichage uniquement).

2. Remplacez toutes les occurrences de `POST` par `ENVOYER` avec modification directe du fichier.

3. Remplacez l'adresse IP `192.168.1.100` par `10.10.10.100` dans tout le fichier (avec modification directe).

4. Modifiez uniquement la ligne 5 pour remplacer `hacker` par `suspicious_user` (avec modification directe).

5. Remplacez le code de statut HTTP `404` par `NOT_FOUND` uniquement sur la ligne 3.

6. Remplacez tous les `.php` par `.html` dans les lignes 1 à 4 (avec modification directe).

7. Supprimez toutes les lignes contenant l'utilisateur `admin` du fichier (avec modification directe).

8. Supprimez la première ligne du fichier `access.log` (avec modification directe).

9. Insérez une ligne de commentaire `# Access log for September 25, 2025` au début du fichier (avant la ligne 1).

10. Ajoutez la ligne `# Total requests: 8` après chaque ligne contenant le code de statut `200`.