

Proiect ITBI 2024-2025

Echipa: Queens Of The Stone Age

Studenti: Drăghici Bianca-Elena și Roibu Amelia-Maria

Problema aleasă (cerință)

File&DiskUsageMonitor (1-2 studenti) Comanda script permite înregistrarea unei sesiuni de lucru în shell într-un fișier numit implicit typescript. Scrieți un program (script) shell care folosește una sau mai multe înregistrări typescript pentru a analiza și monitoriza evoluția în timp a structurii de fișiere și directoare de interes pentru utilizator, precum și a spațiului utilizat pe disc. Concret, programul poate analiza output-ul unor comenzi de tip `ls -l` la momente diferite de timp (i.e., între diferite fișiere typescript, sau între un anumit fișier typescript și momentul curent) și să genereze un raport cu privire la fișierele/directoarele aparute/disparute în timp. Diferența între rezultatele diferitelor comenzi `ls` se poate face cu comanda `diff`, dar programul trebuie să acceseze aceste diferențe într-o formă accesibilă utilizatorului neprofesionist. Similar, programul trebuie să facă o analiză a utilizării spațiului de disc așa cum e raportată de comanda `df`.

Descrierea proiectului

În gestionarea sistemelor de fișiere, utilizatorii și administratorii întâmpină adesea dificultăți în a monitoriza schimbările survenite în structura fișierelor și directoarelor, precum și în utilizarea spațiului pe disc. Aceste schimbări, cum ar fi adăugarea sau ștergerea fișierelor, și variațiile în utilizarea spațiului de stocare pot fi dificil de urmărit fără un instrument dedicat. Necesitatea unei soluții se accentuează în medii unde schimbările pot afecta performanța sau accesibilitatea resurselor.

Proiectul își propune să ofere un script shell care să faciliteze monitorizarea acestor schimbări în timp, utilizând date generate cu ajutorul comenzii `script`. Acest instrument permite analiza și compararea rezultatelor comenzilor `ls -l` și `df -h` la momente diferite, generând rapoarte clare și accesibile pentru utilizatori neprofioniști.

Descrierea soluției adoptate

Obiective principale

1. **Monitorizarea structurii de fișiere și directoare:** Identificarea fișierelor și directoarelor nou adăugate sau șterse la diferite momente în timp.

2. **Monitorizarea utilizării spațiului pe disc:** Analiza variațiilor în spațiul utilizat pe disc, așa cum este raportat de comanda `df -h`.
3. **Generarea rapoartelor:** Crearea unor rapoarte ușor de înțeles, care să evidențieze schimbările survenite între stări diferite ale sistemului.

Structura soluției

Proiectul este compus din 3 scripturi shell, ce pot fi împartite în mai multe etape principale:

1. Script principal (`main.sh`):

- Verifică numărul și validitatea parametrilor de intrare.
- Determină tipul de comparație: între un singur fișier `typescript` și starea curentă sau între două fișiere `typescript`.
- Apelează scripturile dedicate pentru comparații, în funcție de parametri.

2. Script pentru comparația cu un singur fișier (`single_file_comparison.sh`):

a. Curățarea și pregătirea datelor din fișierul `typescript`

-Extrage doar liniile relevante din fișierul `typescript`, eliminând codurile de culoare și caracterele speciale

-Preia informații despre: fișiere și directoare (`ls -l`) și despre utilizarea discului (`df -h`)

-Exemplu din cod:

```
clean_typescript() {
    local file="$1"

    grep -E '^[^d]' "$file" | sed 's/\x1B\[[0-9;]*[a-zA-Z]//g' | awk '{print $9}' | sed 's/\r//g' | sort > /tmp/ls_data

    grep '/dev/sda2' "$file" | sed 's/\x1B\[[0-9;]*[a-zA-Z]//g' | awk '{print $3}' | sed 's/\r//g' > /tmp/df_data
}
```

b. Obținerea stării curente a sistemului

-Colectează date actuale despre fișiere/directoare și utilizarea discului

-Stochează aceste date în fișiere temporare

-Exemplu din cod:

```
get_current_state() {
    ls -l | grep -E '^[^d]' | awk '{print $9}' | sed 's/\r//g' | sort > /tmp/ls_current

    df -h | grep '/dev/sda2' | awk '{print $3}' | sed 's/\r//g' > /tmp/df_current
}
```

c. Compararea datelor dintre `typescript` și starea curentă

-Compară informațiile din fișierul **typescript** cu cele din starea curentă pentru a identifica modificările

-Utilizează comanda **comm** pentru a detecta: fișiere și directoare adăugate și șterse

-Exemplu din cod:

```
44 added_files=$(comm -13 $(echo "$ls_data1" | grep -v 'main.sh' | grep -v 'single_file_comparison.sh' | grep -v 'double_file_comparison.sh' | sort) $(echo "$ls_data2" | grep -v 'main.sh' | grep -v 'single_file_comparison.sh' | grep -v 'double_file_comparison.sh' | sort))
45 removed_files=$(comm -13 $(echo "$ls_data1" | grep -v 'main.sh' | grep -v 'single_file_comparison.sh' | grep -v 'double_file_comparison.sh' | sort) $(echo "$ls_data2" | grep -v 'main.sh' | grep -v 'single_file_comparison.sh' | grep -v 'double_file_comparison.sh' | sort))
```

d. Compararea utilizării discului

-Curăță valorile de utilizare a discului pentru a elimina unitățile de măsură și caracterele speciale

-Compară utilizarea discului între fișierul **typescript** și starea curentă

-Exemplu din cod:

```
disk_used1=$(clean_disk_value "$(cat /tmp/df_data)")
disk_used2=$(clean_disk_value "$(cat /tmp/df_current)")

disk_diff=$(echo "$disk_used2 - $disk_used1" | bc)
```

e. Generarea raportului

```
echo "Fișiere și directoare adăugate:"
echo "$added_files"
```

```
echo "Fișiere și directoare șterse:"
echo "$removed_files"
```

```
if (( $(echo "$disk_diff > 0.01" | bc -l) )); then
|   echo "Diferență în utilizarea spațiului pe disc: +${disk_diff}GB"
elif (( $(echo "$disk_diff < -0.01" | bc -l) )); then
|   echo "Diferență în utilizarea spațiului pe disc: ${disk_diff}GB"
else
|   echo "Nu există modificări semnificative în utilizarea spațiului pe disc."
fi
```

3. Script pentru comparația între două fișiere (**double_file_comparison.sh**):

a. Curățarea fișierelor **typescript**

-Extrage informațiile relevante despre fișiere/directoare și utilizarea discului din fiecare fișier `typescript`

-Exemplu din cod:

```
clean_typescript() {  
    local file="$1"  
  
    sed 's/\x1B\[ [0-9;]*[a-zA-Z]//g' "$file" > /tmp/cleaned_typescript  
  
    grep -E '^-d' /tmp/cleaned_typescript > /tmp/ls_data  
  
    grep '/dev/sda2' /tmp/cleaned_typescript > /tmp/df_data  
}
```

b. Curățarea valorilor de utilizare a discului

- Proceesează valorile utilizării discului pentru a elimina caracterele speciale și unitățile (cum ar fi G, M, etc.)

-Exemplu din cod:

```
clean_disk_value() {  
    local value="$1"  
  
    value=$(echo "$value" | sed -e 's/[A-Za-z%]//g')  
  
    value=$(echo "$value" | tr -d -c '[:digit:].')  
  
    if ! [[ "$value" =~ ^[0-9]+(\.[0-9]+)?$ ]]; then  
        echo "Valoare invalidă pentru utilizarea discului: $value"  
        exit 1  
    fi  
    echo "$value"  
}
```

c. Compararea fișierelor și directoarelor

-Scriptul compară fișierele și directoarele din cele două fișiere `typescript`, identificând: fișiere/directoare adăugate (existente în al doilea fișier, dar nu în primul) și fișiere/directoare șterse (existente în primul fișier, dar nu în al doilea)

d. Compararea utilizării discului

-Exemplu din cod:

```
disk_used1=$(echo "$df_data1" | awk '{print $3}' | sed 's/G//')
disk_used2=$(echo "$df_data2" | awk '{print $3}' | sed 's/G//')

disk_diff=$(echo "$disk_used2 - $disk_used1" | bc)
```

e. Generarea raportului

Probleme întâlnite

În cadrul proiectului, cea mai semnificativă provocare a fost procesarea fișierelor **typescript**, mai exact eliminarea liniilor irelevante și izolarea informațiilor utile generate de comenzile **ls -l** și **df -h**. În codul implementat, acest lucru s-a realizat prin utilizarea comenzii **grep** pentru a identifica liniile care conțin caracteristici specifice ale ieșirilor acestor comenzi (de exemplu, linii care încep cu **-** sau **d** pentru **ls -l**, respectiv linii ce includ **/dev/sda2** pentru **df -h**). Deși abordarea a funcționat, complexitatea parsării a dus la numeroase provocări, mai ales în eliminarea caracterelor speciale și manipularea datelor pentru generarea rapoartelor.

Funcționalități cheie

1. **Curățarea datelor:** Eliminarea codurilor de culoare și a caracterelor speciale din fișierele **typescript** pentru a obține informații utilizabile.
2. **Compararea stărilor:** Utilizarea comenzilor **comm** și **awk** pentru a identifica diferențele între fișiere și directoare.
3. **Analiza spațiului pe disc:** Transformarea și compararea valorilor utilizării discului pentru a evidenția variațiile semnificative.

Avantaje

- **Automatizare:** Scripturile permit automatizarea procesului de monitorizare, reducând efortul utilizatorului.
- **Claritate:** Rapoartele generate sunt clare și adaptate utilizatorilor fără cunoștințe avansate de tehnologie.
- **Flexibilitate:** Soluția funcționează atât cu un singur fișier **typescript**, cât și cu două, oferind flexibilitate în monitorizare.

Limitări

Vă rugăm să țineți cont de faptul că programul nu oferă o examinare detaliată și graduală a schimbărilor, însă prezintă utilizatorului o privire de ansamblu asupra elementelor analizate.

- Numele executabilelor NU trebuie modificate

- Numele fișierelor **typescript** trebuie să fie de forma “typescript[cifră]”, adică numele trebuie să fie urmat de o cifră (de la 0 la 9)
- În interiorul fișierelor **typescript** trebuie să existe o singură comandă **ls -l**, respectiv o singură comandă **df -h**
- În cazul în care numele fișierelor din memorie conțin spații, la generarea raportului va fi afișat pe ecran doar primul cuvânt din nume

Rezultate experimentale

Context experiment:

```

amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ cat typescript1
Script started on 2025-01-15 01:35:16+02:00 [TERM="xterm-256color" TTY="/dev/pts/0" COLUMNS="111" LINES="24"]
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ ls -l
total 1048600
drwxrwxr-x 2 amelia amelia      4096 Jan 15 01:33 director1
drwxrwxr-x 2 amelia amelia      4096 Jan 15 01:33 director2
-rwxrwxr-x 1 amelia amelia     2456 Jan 14 22:45 double_file_comparison.sh
-rw-rw-r-- 1 amelia amelia         0 Jan 15 01:33 file1
-rw-rw-r-- 1 amelia amelia 1073741824 Jan 15 01:35 file2
-rwxrwxr-x 1 amelia amelia       553 Jan 14 19:49 main.sh
-rwxrwxr-x 1 amelia amelia     2873 Jan 14 22:45 single_file_comparison.sh
-rw-rw-r-- 1 amelia amelia         0 Jan 15 01:35 typescript1
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ df -h
Filesystem      Size  Used Avail Use% Mounted on
tmpfs            387M  2.1M  385M   1% /run
/dev/sda2        79G   16G   59G  21% /
tmpfs            1.9G   27M   1.9G   2% /dev/shm
tmpfs            5.0M   8.0K   5.0M   1% /run/lock
tmpfs            387M  140K  387M   1% /run/user/1000
/dev/sr0         88M   88M    0 100% /media/amelia/CDROM
/dev/sr1         5.8G   5.8G    0 100% /media/amelia/Ubuntu 24.04.1 LTS amd64
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ exit
exit

amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ cat typescript2
Script started on 2025-01-15 01:36:09+02:00 [TERM="xterm-256color" TTY="/dev/pts/0" COLUMNS="111" LINES="24"]
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ ls -l
total 20
drwxrwxr-x 2 amelia amelia 4096 Jan 15 01:33 director2
-rwxrwxr-x 1 amelia amelia 2456 Jan 14 22:45 double_file_comparison.sh
-rwxrwxr-x 1 amelia amelia  553 Jan 14 19:49 main.sh
-rwxrwxr-x 1 amelia amelia 2873 Jan 14 22:45 single_file_comparison.sh
-rw-rw-r-- 1 amelia amelia 1850 Jan 15 01:35 typescript1
-rw-rw-r-- 1 amelia amelia   0 Jan 15 01:36 typescript2
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ df -h
Filesystem      Size  Used Avail Use% Mounted on
tmpfs            387M  2.1M  385M   1% /run
/dev/sda2        79G   15G   60G  20% /
tmpfs            1.9G   27M   1.9G   2% /dev/shm
tmpfs            5.0M   8.0K   5.0M   1% /run/lock
tmpfs            387M  140K  387M   1% /run/user/1000
/dev/sr0         88M   88M    0 100% /media/amelia/CDROM
/dev/sr1         5.8G   5.8G    0 100% /media/amelia/Ubuntu 24.04.1 LTS amd64
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ exit
exit

Script done on 2025-01-15 01:36:16+02:00 [COMMAND_EXIT_CODE="0"]

```

Exemplul 1:

```

amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ ./main.sh typescript1 typescript2
Fișiere și directoare adăugate:

Fișiere și directoare șterse:
file1
file2
director1
Diferență în utilizarea spațiului pe disc: -1GB

```


Exemplul 2:

```
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ ./main.sh typescript2
Fișiere și directoare adăugate:
director3
file1
file2
file3
Fișiere și directoare șterse:

Diferență în utilizarea spațiului pe disc: +3GB
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ ls -l
total 3145768
drwxrwxr-x 2 amelia amelia      4096 Jan 15 01:33 director2
drwxrwxr-x 2 amelia amelia      4096 Jan 15 01:36 director3
-rwxrwxr-x 1 amelia amelia      2456 Jan 15 01:43 double_file_comparison.sh
-rw-rw-r-- 1 amelia amelia 1073741824 Jan 15 01:36 file1
-rw-rw-r-- 1 amelia amelia 1073741824 Jan 15 01:37 file2
-rw-rw-r-- 1 amelia amelia 1073741824 Jan 15 01:37 file3
-rwxrwxr-x 1 amelia amelia       553 Jan 14 19:49 main.sh
-rwxrwxr-x 1 amelia amelia      2873 Jan 14 22:45 single_file_comparison.sh
-rw-rw-r-- 1 amelia amelia      1850 Jan 15 01:35 typescript1
-rw-rw-r-- 1 amelia amelia      1678 Jan 15 01:36 typescript2
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ df -h
Filesystem      Size  Used Avail Use% Mounted on
tmpfs           387M  2.1M  385M   1% /run
/dev/sda2       79G   18G   57G  24% /
tmpfs           1.9G   27M   1.9G   2% /dev/shm
tmpfs           5.0M   8.0K   5.0M   1% /run/lock
tmpfs           387M  140K  387M   1% /run/user/1000
/dev/sr0        88M   88M    0 100% /media/amelia/CDROM
/dev/sr1        5.8G   5.8G    0 100% /media/amelia/Ubuntu 24.04.1 LTS amd64
```

Exemplul 3:

```
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ ./main.sh typescript2
Fișiere și directoare adăugate:

Fișiere și directoare șterse:

Nu există modificări semnificative în utilizarea spațiului pe disc.
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ ls -l
total 24
drwxrwxr-x 2 amelia amelia 4096 Jan 15 01:33 director2
-rwxrwxr-x 1 amelia amelia 2689 Jan 15 01:46 double_file_comparison.sh
-rwxrwxr-x 1 amelia amelia  553 Jan 14 19:49 main.sh
-rwxrwxr-x 1 amelia amelia 2873 Jan 14 22:45 single_file_comparison.sh
-rw-rw-r-- 1 amelia amelia 1850 Jan 15 01:35 typescript1
-rw-rw-r-- 1 amelia amelia 1678 Jan 15 01:36 typescript2
amelia@amelia-VMware-Virtual-Platform:~/basic-tools-and-techniques-in-comp-sci$ df -h
Filesystem      Size  Used Avail Use% Mounted on
tmpfs           387M  2.1M  385M   1% /run
/dev/sda2       79G   15G   60G  20% /
tmpfs           1.9G   27M   1.9G   2% /dev/shm
tmpfs           5.0M   8.0K   5.0M   1% /run/lock
tmpfs           387M  140K  387M   1% /run/user/1000
/dev/sr0        88M   88M    0 100% /media/amelia/CDROM
/dev/sr1        5.8G   5.8G    0 100% /media/amelia/Ubuntu 24.04.1 LTS amd64
```

Concluzie

Proiectul oferă o soluție eficientă și accesibilă pentru monitorizarea schimbărilor în structura fișierelor și utilizarea discului. Prin utilizarea comenzilor standard și a fișierelor `typescript`, scripturile automate generează rapoarte clare și ușor de interpretat, adresând nevoile

utilizatorilor neprofesioniști. În ciuda unor limitări tehnice, soluția propusă reprezintă un instrument valoros pentru managementul resurselor sistemului.