Lektion 11 – Laboration 2

Analys av Apache loggfiler

1: Generera loggfiler

- Kör flog (fake log generator) via podman eller docker. Behöver du installera välj podman. https://github.com/mingrammer/flog

```
$ podman run -it --rm docker.io/mingrammer/flog --help
```

Du kan byta ut podman mot docker om redan har det installerat.

- Testa att köra log-generatorn:

```
$ podman run -it --rm docker.io/mingrammer/flog -n 1
```

Det borde skrivits ut en logg-rad i följande format:

```
211.224.122.1 - - [20/Mar/2021:15:51:20 +0000] "PUT /wp-admin HTTP/1.0" 404 5023
```

https://en.wikipedia.org/wiki/Common_Log_Format

- **Generera en gzippad loggfil i CLF – format med 10000 rader för analys.**
  - **Du behöver inte gzippa för hand – log-generatorn klarar det mer rätt argument!**
- **Använd ”zcat” för att läsa in den gzippade filen i skripten nedan.**

```
$ podman run -it --rm -v $PWD:/mnt/log docker.io/mingrammer/flog -o
/mnt/log/filnamn -t log
```

Du måste lägga till parametern **`-v $PWD:/mnt/log`** *för att komma åt filen som sparas. Den kommer att sparas i foldern du befinner dig i. OBS! Kommandot ska skrivas på en rad.*

2: Status-statistik

- Skapa ett skript som med hjälp av ”while read” summerar det totala antalet anrop som genererat statuskoder som börjar på 2 och 5. Skriptet ska sen rapportera hur vanlig varje kodtyp är i procent av total trafik.
- Skriptet tar en gzippad loggfil som argument och skriver ut rapporten på **stdout**.

Exempelutskrift:

```
Total: 10000
2XX: 9023 (90%)
5XX: 184 (2%)
```

Tips:

- Ange argument till ”read” för varje fält i loggfilen. Tänk på att fälten avdelas med ”whitespace”.
- Eftersom bash använder heltalsaritmetik, multiplicera upp med 100 innan du dividerar för att räkna ut procent.
  - Om du är klurig kan du få till avrundning också, men det är inte viktigt!

3: URL-statistik

- Utöka ditt skript så det också rapporterar de 5 vanligaste URL:arna.

Tips:

- Använd temp-fil för att spara URL:ar, så att du sen kan använda kommandon för att få fram data till din rapport.
- Använd ”sort” och ”uniq” som i tidigare labbar.

Exempelutskrift:

```
Top URLs:
   1286 /app/main/posts
   1269 /wp-admin
   1265 /explore
   1263 /search/tag/list
   1262 /list
```

4: Nedladdad datamängd

- Utöka skriptet så det rapporterar total nedladdad datamängd. Presentera mängden i ett för människor trevligt format.
- Gör konverteringen från bytes till trevligt format i en funktion. Använd gärna lokala variabler i funktionen. Du behöver bara hantera kB, MB och GB.

Exempelutskrift:

```
Downloaded:
50 MB
```

Tips:

- Eftersom funktioner är kommandon kan utskrifter från en funktion också fångas i en variabel:

```
printStuff() {
    echo ”hello”
}
output=$(printStuff)
```