

## Programação Shell

1. O Squid é um famoso software livre usado para criar servidores Proxy Web. Para cada acesso feito a um site na internet por um cliente, é gerada uma entrada no arquivo de log com um formato padrão descrito no exemplo abaixo:

```
1347542919.400 53 10.2.10.26 TCP_MISS/200 4596 GET
http://l.yimg.com/zz/combo? - DIRECT/200.152.174.74
application/javascript
```

Os campos deste arquivo são separados por espaços. São eles:

1. Data no formato de marca de tempo do UNIX (número de segundos transcorridos desde 1/1/1970). Para converter esta data para um formato legível, você pode utilizar o comando: `date -d @<timestamp>`
2. Tempo gasto em milisegundos para atender a requisição.
3. Endereço IP do cliente
4. Código resultante do acesso. Alguns deles são:
  - a. TCP\_MISS – Objeto obtido diretamente da internet;
  - b. TCP\_HIT – Objeto obtido da cache do squid;
  - c. TCP\_DENIED – Acesso negado pelo controle de acessoVeja todos os possíveis códigos em: [http://wiki.squid-cache.org/SquidFaq/SquidLogs#Squid\\_result\\_codes](http://wiki.squid-cache.org/SquidFaq/SquidLogs#Squid_result_codes)
5. Quantidade de bytes entregue ao cliente;
6. Método de requisição do objeto pelo cliente: GET ou POST;
7. Endereço acessado(URL)
8. hífen
9. Indica o endereço IP da máquina contactada para a obtenção da url solicitada. Caso a página seja obtida do cache, este campo aparece com a palavra NONE
10. O tipo de conteúdo retornado pela solicitação da URL;

De posse destas informações, escreva um script Shell que gere um relatório com as seguintes informações:

1. Número de sites diferentes acessados (considere apenas a parte do nome do host na url);
  - a. Ex: [http://www.google-analytics.com/\\_\\_utm.gif?](http://www.google-analytics.com/__utm.gif?)
2. Número total de dados transferidos para cada tipo de conteúdo;
3. Número de sites encontrados na cache;
4. Número de sites trazidos diretamente da internet;
5. Número total de sites com acesso negado;
6. Número total de requisições de revalidação (TCP\_REFRESH\_UNMODIFIED).
7. Número total de acessos entre 07:44:38hs e 07:44:50hs do dia 16 de julho de 2022.

A saída deve ser apresentada na ordem colocada acima, imprimindo na tela apenas as quantidades, sem qualquer texto.