

Uma das decisões de projeto em sistemas de arquivos é a escolha do tamanho dos blocos lógicos. Como discutido em aula, blocos maiores oferecem maior eficiência na transferência de dados de/para o dispositivo de armazenamento, mas geram maior desperdício de espaço devido a fragmentação interna. Outra questão diz respeito à necessidade de escolher estruturas de metadados que comportem arquivos de diferentes tamanhos. Neste exercício você irá coletar dados do seu sistema de arquivos para raciocinar sobre essas questões.

Primeiro, gere uma lista de arquivos em `/usr/bin` e `/usr/share` e seus respectivos tamanhos, usando o comando abaixo:

```
$ find /usr/bin /usr/share -type f -exec ls -l {} \; 2>/dev/null | \
  awk '{print "\"" $9 "\"", $5}' | tee /tmp/lista-arqs
```

Ao final da execução, o arquivo `/tmp/lista-arqs` conterá duas colunas, a primeira com os nomes de arquivos e a segunda com seus tamanhos (em bytes).

Com base nos dados coletados, responda às seguintes perguntas:

- (a) Quantos arquivos foram encontrados?
- (b) Qual o maior tamanho de arquivo observado? A que arquivo corresponde?
- (c) Quantos arquivos com tamanho zero existem? A que porcentagem do total eles correspondem?
- (d) Qual o tamanho médio de arquivo? Qual a porcentagem de arquivos com tamanho igual ou menor do que a média?
- (e) Qual a mediana do tamanho de arquivo?
- (f) Qual o menor tamanho de bloco necessário para que pelo menos 50% dos arquivos ocupem apenas um bloco? Qual a porcentagem de arquivos que ocuparia um bloco se esse tamanho fosse adotado? Quantos blocos (com o tamanho encontrado) ocuparia o maior arquivo?
NOTA: o tamanho de bloco deve ser uma potência de 2 (512 bytes, 1 KB, 2 KB, 4 KB, 8 KB, ...)
- (g) Sabendo que um i-node no Linux possui 12 ponteiros diretos e indireção simples, dupla e tripla, e supondo que os ponteiros de disco (endereços de bloco) sejam de 32 bits, determine:
 - a porcentagem de arquivos que não precisam de indireção (i.e., que usam apenas os ponteiros diretos);
 - a porcentagem de arquivos que precisam de indireção simples;
 - a porcentagem de arquivos que precisam de indireção dupla;
 - a porcentagem de arquivos que precisam de indireção tripla.
- (h) Considere que o sistema de arquivos aloca um número integral de blocos para cada arquivo. Por exemplo, se um arquivo tem 1337 bytes e o tamanho de bloco é de 1 KB, serão alocados dois blocos para esse arquivo, o que significa que ele ocupará 2 KB no disco, dos quais $2048 - 1337 = 711$ bytes serão desperdiçados por fragmentação interna. Sabendo disso, calcule o espaço desperdiçado por fragmentação interna (total e porcentagem) considerando os arquivos analisados.
DICA: você precisará saber o tamanho de bloco usado pelo seu sistema de arquivos. Para isso, use o comando `stat arq` (onde `arq` é um arquivo contido em `/tmp/lista-arqs`) e veja o valor de `IO Block`.