

Sujet du Projet

Université de Bordeaux, M2 Génie Logiciel

David AUBER, Francieli ZANON BOITO

2019-2020

Dans un supercalculateur pour le calcul hautes performances (HPC), plusieurs applications (ou *jobs*) s'exécutent de manière concurrentielle, sur des noeuds de calcul différents. Les accès aux fichiers est fait via un système de fichiers parallèle (PFS), qui est partagé par les différentes applications. On appelle << motif d'accès >> la façon dont un accès est fait : les fichiers sont lus ou écrits, de façon séquentielle ou aléatoire, avec une certaine API.

On utilisera pour ce projet un jeux de données sur les activités d'accès au PFS du supercalculateur Intrepid, situé aux États-Unis (au Argonne National Laboratory), pendant l'année 2012¹.

Le jeu de données est disponible dans le cluster hdfs sur */raw_data/ALCF_repo/*, et est composé de trois fichiers, dont le plus intéressant (et aussi le plus volumineux) est *phases.csv*. Il contient la séquence des phases d'E/S observées pendant cette année. Chaque phase est représentée dans une ligne, par 9 colonnes séparées par des ;.

```
start;end;duration;patterns;npatterns;jobs;njobs;days;ndays
```

```
220152308402;220152707824;399422;2;1;11;1;3;1
```

```
220152707824;220153139869;432045;-1;0;-1;0;-1;0
```

```
220153139869;220165067403;11927534;7;1;11;1;3;1
```

- Les deux premières colonnes (0 et 1, *start* et *end*) donnent le *timestamp* (en microsecondes depuis un certain moment) de début et de fin de chaque phase. La période couverte par ces données est d'un an. Chaque phase démarre à la fin de la phase précédente. Le fichier est trié par timestamp de début.
- La colonne 2 (*duration*) donne la durée de la phase (*timestamp* de fin – *timestamp* de début).
- La colonne 3 (*patterns*) est une liste des motifs d'accès observés pendant cette phase, séparés par des , . Il est possible d'en avoir plusieurs parce que le système de fichiers parallèle est partagé par plusieurs applications, qui peuvent donc lui accéder en même temps, et aussi par plusieurs processus de chaque application.
- La colonne 4 (*npatterns*) donne le nombre de motifs d'accès observés pendant la phase (la taille de la liste de la colonne 3).
- La colonne 5 (*jobs*) est une liste des applications (*jobs*) qui ont accédé au PFS pendant la phase, et la colonne 6 (*njobs*) donne le nombre de jobs (la taille de la liste).

¹ Ces données viennent du ALCF I/O Data Repository (<https://www.mcs.anl.gov/research/projects/darshan/data/>) en forme d'un ensemble de fichiers, un par job, obtenus avec l'outil de monitoring Darshan. Des chercheurs du LaBRI (Francieli ZANON BOITO) et de l'Université Fédérale du Rio Grande do Sul (Pablo PAVAN, Jean BEZ et Philippe NAVAUX) ont manipulé les données d'origine pour générer la version utilisée dans ce projet.

- La colonne 7 (*days*) donne une liste des jours correspondant à la période de temps de la phase, et la colonne 8 (*ndays*) contient le nombre de jours (la taille de la liste).

Les motifs d'accès dans la colonne 3 et les *jobs* dans la colonne 5 sont des codes, utilisés pour diminuer la taille du fichier *phases.csv*. Pour les traduire, le fichier *jobs.csv* donne le nom de chaque job selon son code, et le fichier *patterns.csv* donne la description de chaque motif d'accès selon son code.

Il y a 22 motifs d'accès dans le jeu de données (codes 0 à 21). Le code -1 est utilisé pour représenter les phases *idle*, c'est-à-dire quand le système de fichier parallèle n'a pas été utilisé.

On voudrait savoir :

1. Des informations sur la durée des phases
 - a. Pour les phases qui ne sont pas *idle*, la distribution² de leur durée.
 - b. La distribution de la durée des phases *idle*.
 - c. Pour chaque motif d'accès (parmi les 22), la distribution de la durée des phases où ce motif apparaît seul (pas dans une liste avec des autres motifs).
2. Nombre de motifs d'accès par phase
 - a. La distribution de *npatterns*. Attention : il ne faut pas y inclure les phases *idle*.
3. Nombre de jobs par phase
 - a. La distribution de *njobs*. Attention : il ne faut pas y inclure les phases *idle*.
4. Temps total d'accès au système de fichier parallèle par job
 - a. Distribution du temps total d'accès au PFS par job.
 - b. Top 10 jobs en temps total d'accès au PFS.
5. Temps total en *idle* du système.
6. Représentativité des 22 motifs d'accès
 - a. Pourcentage du temps total de phases où chaque motif a été observé seul (qu'un seul motif pendant la phase), ou concurrent à des autres.
 - b. Top 10 patterns en représentativité.
7. Proposez et implémentez une solution permettant d'obtenir toutes les plages horaires comportant 4 patterns donnés. Votre solution devra prendre un temps linéaire en fonction du nombre de résultats et non du nombre d'éléments de la base.
8. En considérant que nous recevons ce jeu de données tous les jours, estimer et mesurer l'évolution du temps de calcul pour chacune des questions au cours de l'année.

Rendu :

Le projet sera en binômes, à rendre sur Moodle. On vous demande de soumettre du code, mais également un rapport qui répond aux questions posées et explique en détails comment les résultats ont été obtenus.

² Pour décrire une distribution, on voudra : minimum, maximum, moyenne, médiane, premier et troisième quadrants, histogramme.