

Résumé :

Avec l'augmentation massive de données produites, l'accès aux données et le stockage de celles-ci devient une problématique critique. Certaines données nécessitent d'être accédées avec une latence minimum alors que d'autres types de données, typiquement des données rarement accédées, peuvent être stockées sur des supports à latence élevée.

L'utilisation du SSD et HDD ensemble sur une même machine est devenu assez commun mais l'ajout de bande magnétique ou *Cloud* n'est pas parfait, le premier est relativement lent d'accès et bien que le deuxième soit de plus en plus répandu, dans la plupart des cas les fichiers sur le *Cloud* sont juste une copie des fichiers locaux ou doivent être téléchargés pour être lu.

Plus le nombre de support est grand plus il devient compliqué d'organiser les fichiers pour les répartir en fonction de leurs caractéristiques (fréquence d'accès, taille, etc.). De plus, il faut se rappeler où les fichiers ont été placés pour ne pas perdre du temps à les chercher. Pour finir dans le cas de fichiers volumineux où seulement une partie est accédée régulièrement, c'est tout le fichier qui doit se situer sur le même espace de stockage.

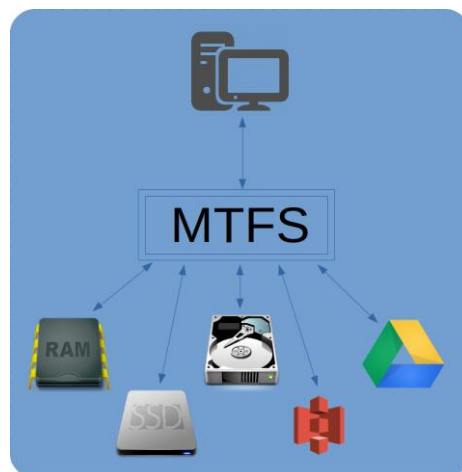
Dans le but de résoudre ces problèmes, la solution proposée est un système de fichiers, capable de gérer plusieurs espaces de stockage et de migrer automatiquement les données d'un espace à l'autre en fonction de règles spécifiées lors de sa création. Les fichiers ne sont plus obligatoirement sur un seul espace de stockage mais peuvent être répartis en plusieurs morceaux sur l'ensemble des espaces de stockage ce qui permet dans le cas du *Cloud* de pouvoir accéder au fichier sans que la totalité ne soit nécessairement téléchargée. Ce système est appelé MTFS (*Multi Tier Filesystem*).

Les services de stockage utilisés par MTFS sont gérés avec des *plugins*. L'interface nécessaire à leur implémentation étant fournie, tout développeur tiers peut intégrer un nouvel espace de stockage au système.

Deux *plugins* ont été réalisés pour les besoins de ce projet. Le premier utilise les espaces de stockage comme le SSD ou le HDD, l'autre utilise l'espace de stockage *Cloud* Amazon S3. Ce dernier utilise le SDK fourni par Amazon.

L'intégralité du projet est sous licence GNU GPLv3 et est disponible sur GitHub à l'adresse suivante :

<https://github.com/Dawen18/mtfs>



Candidat :

M. WITTWER DAVID

Filière d'études : ITI

Professeur(s) responsable(s) :

Glück Florent

En collaboration avec :

Travail de bachelor soumis à une convention
de stage en entreprise : non

Travail de bachelor soumis à un contrat de
confidentialité : non