

INGÉNIERIE DES TECHNOLOGIES DE L'INFORMATION

ORIENTATION – LOGICIELS ET SYSTÈMES COMPLEXES

MOTEUR DE TAGGING ET D'INDEXATION DES FICHIERS AVEC RUST

Descriptif :

Les systèmes de *tagging* de fichiers existants que l'on peut trouver dans le monde Linux aujourd'hui utilisent des bases de données et/ou des fichiers additionnels pour stocker les méta données des fichiers taggés. La problématique liée au fait que les fichiers et leurs méta-données ne sont pas stockés dans la même entité (i.e. le fichier), complique les implémentations et les rend sub-optimales en terme de complexité, performance, latence et robustesses.

Une solution alternative à ces approches traditionnelles serait de tirer parti des *extended attributes* présents dans la plupart des systèmes de fichiers disponibles dans le noyau Linux (ext4, xfs, etc.). Les *extended attributes* permettent, pour chaque fichier, de stocker ou extraire un ensemble de méta-données, le tout de manière efficace.

Le but premier de ce projet est de concevoir et développer un « moteur de gestion de tags » pouvant aisément et efficacement gérer des dizaines ou des centaines de milliers de fichiers et *tags* associés.

Le but secondaire du projet, est d'en réaliser l'implémentation dans le langage Rust. En effet, le langage Rust est un langage système moderne, digne successeur du langage C, et conçu pour être extrêmement fiable, robuste et performant.

Les mécanismes d'indexation seront étudiés et les systèmes de surveillance de fichiers seront investigués. En effet, le système devra s'assurer de surveiller les fichiers modifiés, créés, ou supprimés afin d'indexer les tags avec un minimum de latence, le but étant d'offrir des performances aussi proches du temps réel que possible.

Enfin, si le temps le permet, le système développé sera intégré à un gestionnaire de fichiers populaire dans le monde Linux (Nautilus, Thunar, Konqueror, etc.).

Travail demandé :

- Etat de l'art.
- Etude et prise en main du langage Rust et familiarisation avec celui-ci dans le contexte système du projet.
- Analyse des *extended attributes* ; en particulier, étude de leur comportement lors des opérations d'accès courantes (cp, mv, rsync, copies depuis un gestionnaire de fichier, etc.) afin d'éviter la perte involontaire de tags au cas où certaines opérations ne maintiendraient pas certains attributs.
- Analyse et sélection des systèmes de notification du système de fichiers.
- Analyse et sélection des algorithmes d'indexation et de recherche.
- Conception et implémentation du système.
- Analyse du système implémenté (mesures de performances, etc.).
- Discussion de l'utilisation du langage Rust par rapport à une implémentation traditionnelle en C.
- Démonstrateur.
- Si le temps le permet : intégration à un gestionnaire de fichiers.

Candidat :

M. LIATTI STEVEN

Filière d'études : ITI

Professeur(s) responsable(s) :

GLUCK 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**