

Séances 6 et 7 : encodage et décodage d'un bloc, programmes "utilitaires"

Cette étape a pour objectif d'utiliser les fonctions d'encodage/décodage pour écrire/lire le contenu d'un bloc d'encodage, et ce dans le contexte de la compression d'un fichier de *log* structuré en ligne de texte.

Encodage L'encodage du journal se fait ligne par ligne. Une nouvelle fonction du module *block* reçoit une nouvelle ligne de texte, encode cette ligne et tente de l'ajouter **entièrement** dans le bloc. Si la ligne "ne tient pas" dans le bloc, l'ajout ne peut avoir lieu, et le bloc sera remis dans son état précédant. Il convient aussi de restaurer l'état du registre d'encodage (commande `ENCODE_CMD_ROOLBACK`).

Les lignes du *log* sont obtenues, une par une, grâce à une lecture sur un tube nommé (c'est à dire un producteur-consommateur). N'importe quelle commande UNIX peut être utilisée pour alimenter l'entrée du tube. Lorsque bloc est plein, il est stocké dans le fichier *log*, puis un nouveau bloc est créé avec le nouvel encodage défini selon le nombre d'occurrences des lettres dans le bloc précédant.

Décodage Contrairement à l'encodage, le décodage du *log* se fait bloc par bloc. La fonction de décodage s'occupe d'allouer un tableau de taille adaptée au stockage des données décodées. Le bloc à décoder est lu en une seule opération d'entrée/sortie depuis le fichier de *log*.

Travail demandé

1. Écrire les fonctions permettant d'encoder ligne par ligne un *log*, et de stocker les blocs devenus pleins dans un fichier.
2. Écrire les fonctions permettant de décoder un bloc compressé.
3. Écrire deux programmes utilitaires pour compresser les données arrivant sur l'entrée standard, et décompresser le contenu d'un fichier pour l'afficher sur la sortie standard.