

# METEO PROJET

L'objectif du projet est d'avoir un programme principal en script shell qui va selon nos paramètres lire une base de données de plus de 2M de lignes afin de récupérer uniquement les informations nécessaires et les écrire dans un fichier temporaire.

Ensuite le script shell va appeler le second programme cette fois-ci en C qui va trier ce fichier temporaire avec l'une des 3 méthode de tri disponible

Le script shell va finalement récupérer le fichier trié par le programme en C et envoyer ces informations à gnuplot pour générer une image contenant le graphique demandé

## Répartition des tâches:

Romain Pombourcq (Programme C):

- Mise en place de la structure du projet via le Makefile
- Création de l'algorithme permettant de trier les données du fichier petit à petit pour ne pas saturer la RAM en divisant le tri dans plusieurs fichiers avec une taille maximal de lecture donnée avant de passer au prochain fichier.
- Création de la structure globale du programme via la structure option et la liste de pointeur sur fonction afin d'avoir un code générique et maintenable
- Définir le format du fichier temporaire qui sera prit en entrée par le programme ( key;value) et par conséquent la structure data qui permet de stocker ces informations
- Création des fonctions de tri utiles pour une liste chaînée ( tri par insertion ) et d'arbre binaire ( tri par insertion )
- Ajout d'une fonction permettant d'équilibrer l'arbre binaire pour permettre d'utiliser la méthode AVL
- Après avoir trier le fichier input en plusieurs fichiers temporaires, rassembler les fichiers en un seul fichier output avec l'algorithme d'external merge sort
- Gestion des erreurs et vérifications des leak et free de la mémoire

Roshanth Ruban (Script Shell):

- Récupération des arguments
- Gestion des erreurs ( mauvais argument / argument manquant )
- Utilisation de awk pour lire le fichier input et créer le fichier temporaire avec les informations nécessaires

- Vérification du retour du programme de tri
- Création du graphique avec la librairie gnuplot

## Problèmes:

- Les options de localisation du script shell sont limité à aucun ou celui de la France car des problèmes de conditions qui n'étaient pas vérifier ont été rencontrées dans l'expression awk, par défaut nous avons essayé de définir des variables sous la forme d'un rectangle xmin xmax ymin ymax pour chacune des localisations afin de vérifier ces variables avec celles des coordonnées GPS de la ligne qu'on vérifie mais les condition n'étaient pas vérifier pour les variables ymin et ymax pourtant les variables existaient et si on remplace les variables par leur valeur en brut cela fonctionnait, c'est ce qu'on a donc fait pour la France. Cependant si on avait plusieurs conditions de location aucune d'entre elles n'étaient testés c'est pourquoi nous avons été contraints par l'expression awk. Nous avons décidé de rester avec cette fonction car elle permet une très grande optimisation lors de la lecture du fichier input et les résultats de performances sont sans égales avec une méthode plus basique avec une lecture ligne par ligne.
- Les méthodes -p et -t n'ont pas été implémentées car nous avons rencontré des erreurs lors de l'appel à la librairie gnuplot pour créer un graphique d'erreur, nous n'avons pas réussi à régler ce problème.