

# Optimisation de l'exécution d'une requête analytique dans une base de données

Eliott Paquet

2025

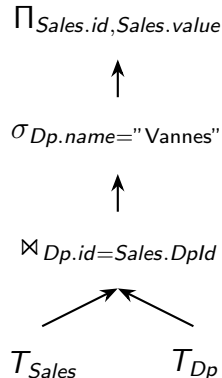
# Edgar F. Codd 1970

Calcul Relationnel Algèbre Relationnel

# Edgar F. Codd 1970

## Calcul Relationnel Algèbre Relationnel

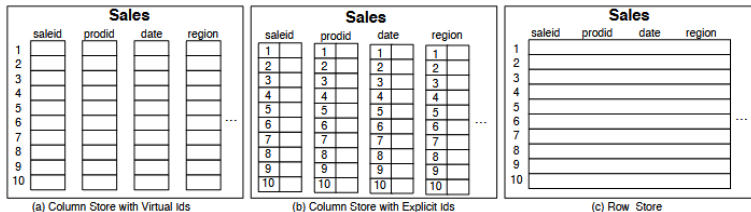
```
1 Select Sales.id, Sales.value
2 From Sales
3 Join Dp
4 On Dp.id = Sales.DpId
5 Where Dp.name = "Vannes";
```



# Problématique

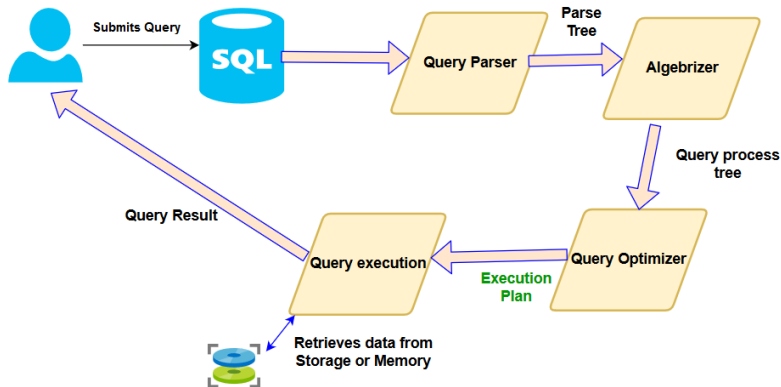
**Comment optimiser des  
requêtes agrégatives itérées  
sur une base de donnée?**

# Les bases de données orientées colonnes



**Figure:** Rendu physique des BDD orientées colonnes et BDD orientées lignes

# Description du traitement d'une requête



**Figure:** Voici (dans les grandes lignes) à quoi ressemble l'exécution d'une requête par une base de données.

# Procédés d'optimisation

## La Vectorisation

Consiste à traiter un vecteur de donnée pour les étapes de Sélection

## Décompression tardive

Pour certaine étape "facile" il est possible de traiter les données compressées

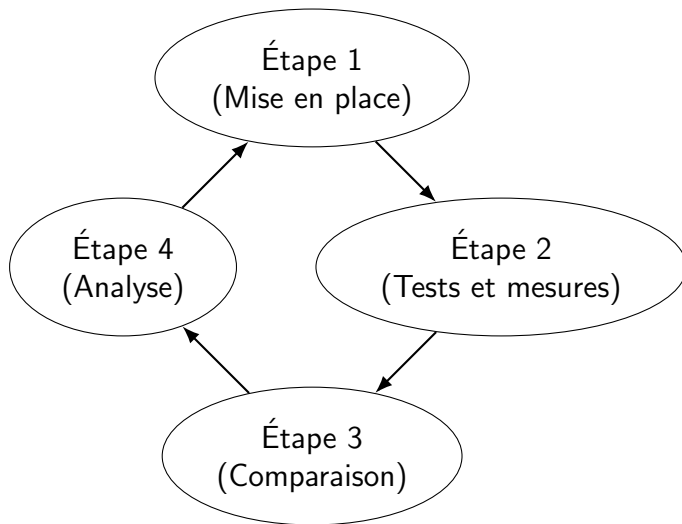
## Matérialisation tardive

Regrouper les couples le plus tard possible pour accélérer l'exécution

## Recherche de plan par estimation du coût

Chercher un plans optimisé en estimant le temps d'exécution

# Procédé Expérimental



# L'avancement de notre implémentation

## Ce que nous avons réussi à implémenter

- ▶ Parseur/Lexer de SQL
- ▶ Stockage et lecture dans la mémoire
- ▶ L'Exécution de la requête

## Ce qu'il nous reste à faire

- ▶ L'Algebrizer
- ▶ L'Optimisateur de requêtes

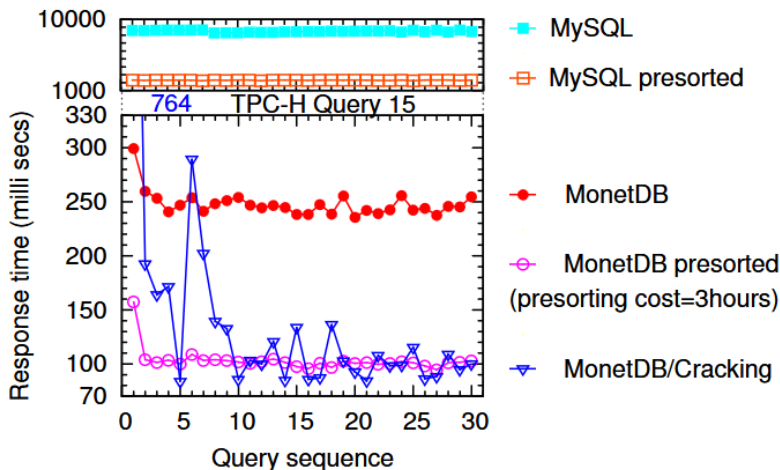


Figure: Notre objectif est de comparer avec ce genre de graphique notre implémentation et ce qui a déjà été fait.

# Nos ressources et contacts

## Nos ressources:

- ▶ *The Design and Implementation of Modern Column-Oriented Database Systems*
- ▶ La documentation de SQLite
- ▶ Les cours du Prof. Andy Pavlo sur les SGBD et leur implémentation au CMU

## Nos contacts:

- ▶ Clément Rouvoy, étudiant ayant fait plusieurs stages sur les BDD en labo de recherche