

UE 4IN803 – SAM – Devoir Maison mars 2020

Ex1 :	Ex2 :	Ex3 :	
--------------	--------------	--------------	--

Devoir à rendre sur le Moodle de l'UE.

Exercice 1 : Exécution de requête

L'objectif de cet exercice est de comprendre le modèle d'exécution basé sur la notion d'itérateur, vu en cours.

Un opérateur est relié à 0, 1 ou 2 opérateurs (fils) de manière à former un arbre d'opérateurs. Tout opérateur dispose des méthodes : open() pour initialiser l'itération, next() retournant soit tuple, soit une valeur indiquant qu'il n'y a pas de tuple suivant, close() pour clore l'itération.

Proposer une implémentation de préférence en python pour les opérateurs suivants

- parcours séquentiel d'une table (table access full en TME)
- sélection (filter en TME)
- jointure par hachage entre 2 opérateurs (Hash join en TME)

L'exécution d'une requête consiste à effectuer une itération sur l'itérateur racine de l'arbre d'opérateurs afin d'afficher le résultat.

Répondre en commençant par décrire en français la solution qui vous proposez (au moins 1 paragraphe). Puis ajouter des extraits de code avec des commentaires.

Exercice 2 : Conception d'une base de données répartie

En vous inspirant des particularités de la situation courante de confinement et de télétravail, proposer un cas de BD répartie: Le cas peut concerner par exemple des informations épidémiologiques hospitalières, celles de l'agence française de santé qui a un site par département. Un autre exemple peut être des données qui avant le confinement étaient produites en manipulées en un seul site et qui sont actuellement produites et manipulées dans de nombreux lieux de confinement, tout en pouvant toujours être manipulées de manière globale. N'hésitez pas à suggérer d'autres exemples.

A) Définir un schéma global et sa fragmentation sur plusieurs sites.

Votre solution doit contenir au moins

- une fragmentation horizontale dont la condition de sélection porte sur **plusieurs** attributs.
- une fragmentation horizontale dérivée complète et disjointe
- une fragmentation horizontale dérivée complète et **non** disjointe
- une fragmentation verticale
- une fragmentation définie par une **combinaison** d'au moins deux fragmentations

B) Proposer des requêtes telles que, après substitution et simplification :

- elle est exécutée entièrement sur un seul site
- elle est exécutée sur au moins deux sites
- la distributivité entre deux opérateurs a permis de supprimer des opérations inutiles

Proposer au moins une requête contenant un group by

Exercice 3 : Optimisation de requêtes

Lire l'article *Exact Cardinality Query Optimization with Bounded Execution Cost* publié à la conférence SIGMOD 2019

Proposer un exemple détaillé de jointure entre 4 relations qui illustre la solution décrite.

Expliquer quelles sont les principales différences entre cette solution et l'optimisation de requêtes présentée en cours.