Sujet ouvert : 29-03-2023

Mécanismes Internes des SGBD

Durée : Environ 2h

L'utilisation de documents et de ressources internet est autorisée, à condition de citer vos sources lorsque vous les utilisez. Le travail est à rendre par groupe de projet « SGBD »

Préliminaires :

On considère le schéma de la base de données suivant :

```
CREATE TABLE Client (
  idCli integer,
  nom varchar(30),
  prenom varchar(30),
  ville varchar(50),
  primary key (id),
  foreign key (codePays) references Pays(codePays)
  )

CREATE TABLE Compte (
  idComp integer,
  idProp integer,
  idBanque integer,
  primary key (idp),
  foreign key (idProp) references Client(idCli),
  foreign key (idBanque) references Banque(idBanque)
  )

CREATE TABLE Banque (
  idBanque integer,
  nomType varchar(30),
  primary key (idBanque)
  )

CREATE TABLE Operations (
  idOp int8,
  idComp integer,
  montant decimal(10, 2),
  timestamp date,
  primary key (idOp)
  foreign key (idCli) references Compte(idComp)
  )
```

Informations sur le contenu de la base :

La base de données est supposée contenir l'ensemble des transactions des cartes bancaires en France. Il y a 400 transactions CB par seconde. On considère qu'il y environ 50 millions de clients domiciliés en Frnace pour 73 millions de comptes.

On considère que les noms des clients sont en moyenne de 7 lettres, et que les prénoms sont en moyenne de longueur 6 lettres, chaque client ayant simplement rempli 1 seul prénom. On considère que la longueur moyenne des noms de villes est de 8 lettres. On considère qu'il y a 10 banques différentes.

Information sur les requêtes :

On considère que les requêtes posées sont essentiellement de plusieurs types :

- R1 : Retourner l'ensemble des opérations d'un compte client sur une période données du type : SELECT idOp, idComp, montant FROM Operations WHERE timestamp BETWEEN <t1> AND <t2>
- R2 : La même requête mais sur l'ensemble des comptes d'un client : SELECT O.idOp, O.idComp, O.montant FROM Operations O, Client Cli, Compte Comp WHERE timestamp BETWEEN <t1> AND <t2> AND Cli.idCli = Comp.idCli AND Comp.idComp = O.idComp
- R3 : Pour chaque banque B_1, et pour chaque période de 24h, le montant d'argent total qu'elle doit verser à une autre banque B_2, correspondant aux transactions des clients de B_1 (débit + crédit) avec des clients de B_2

Q1 : Requêtes

Donnez le code SQL de la requête R3.

Donnez les arbres d'exécution possible des requêtes R1 et R2. Est-ce que rajouter des index pourrait améliorer le temps d'exécution des requêtes R1 et R2 ? Quels types d'index utiliser ? Expliquez.

Q2 : Structure et dimensionnement

Discutez des types utilisés dans la base. Proposez si vous le souhaitez d'autres types plus optimaux.

Combien de place disque va occuper une base de données contenant 1 an de données ? Combien de place vont occuper les index ?

Peut-on héberger l'ensemble de la base de données en mémoire sur un serveur avec 128Gb de RAM ? Quelle quantité de RAM optimale suggérez vous de mettre sur le serveur (par incréments de 16Gb) ?

Q3 : Sécurité

Identifiez plusieurs risques de sécurité liés à l'utilisation de la base de données « en clair ».

On propose l'approche de chiffrement suivante : Dans le client permettant de communiquer avec la base, on va chiffrer le champ « montant » de la table opérations. Discutez de l'impact de cette approche pour les requêtes R1, R2 et R3 et sur votre dimensionnement précédent. Discutez de la sécurité de cette approche par rapport aux risques que vous avez identifiés.

On propose une 2^{e} approche, qui consiste à chiffrer le montant idComp directement avec une méthode de la base de données (on peut utiliser par exemple la fonction DECRYPT(champ) dans une requête SQL pour obtenir la valeur de ce champ). Discutez de la sécurité et des avantages/inconvénients de cette approche, en particulier par rapport aux risques que vous avez identifiés. Précisez comment se déroulera l'exécution des requêtes R1, R2 et R3 dans ce contexte.

Présentez si vous le souhaitez une autre approche de sécurisation des données et discutez de son implémentation.