

Travaux Pratiques

Bases de données distribuées

2020-2021

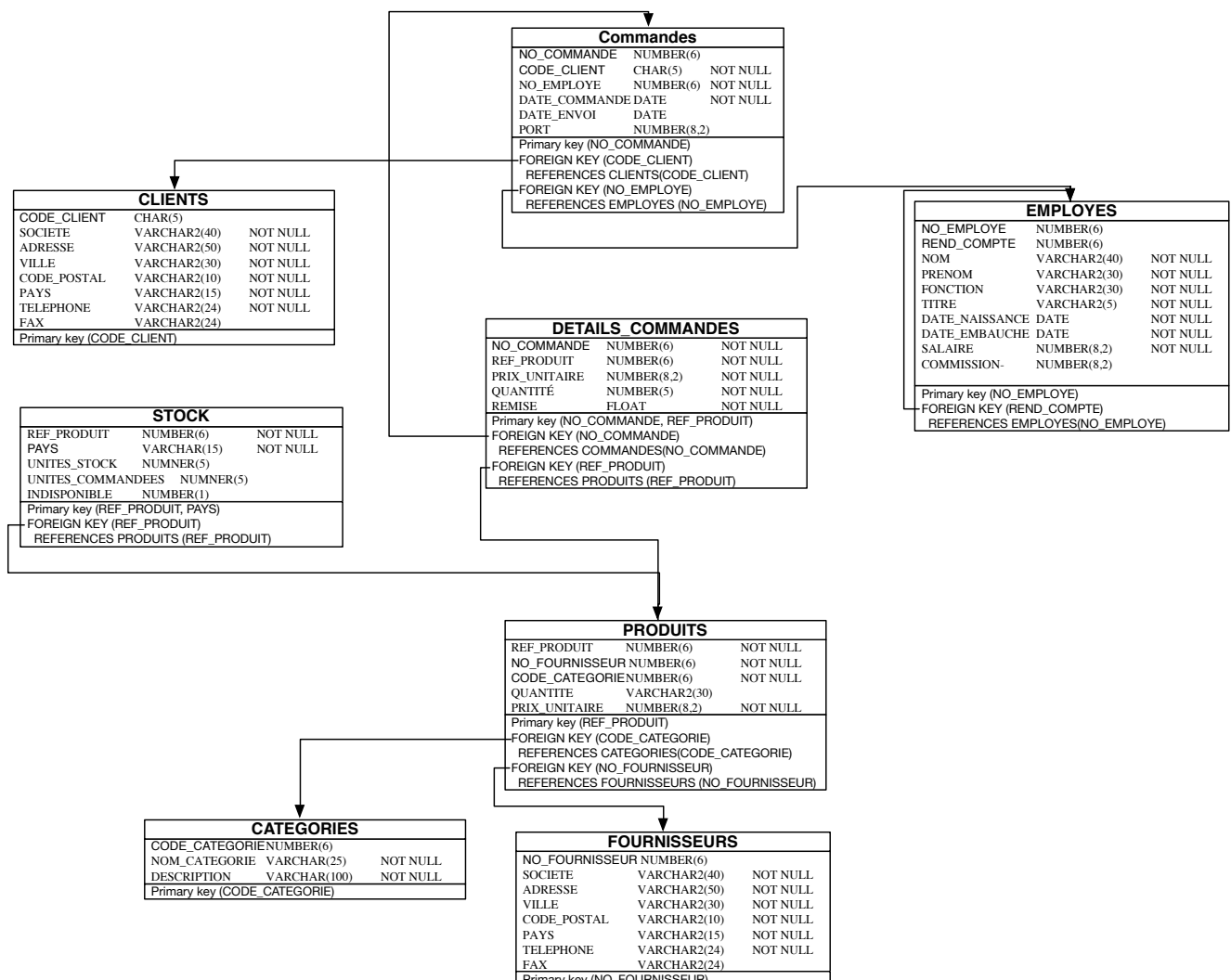
(2 séances)

Introduction.

Nous considérons une entreprise qui, bien qu'elle soit implantée sur plusieurs continents, gère sa base de données de manière centralisée. L'objectif du TP est de lui proposer une solution distribuée.

Nous appellerons cette entreprise «Ryori».

Le schéma simplifié de sa base de données est le suivant.



L'entreprise est largement implantée en Europe mais aussi aux Amériques. Il est donc prévu de répartir la base entre trois sites : Europe du Nord, Europe du Sud et Amériques.

Le siège social est en Allemagne (site Europe du Nord). Si l'entreprise se développe sur d'autres continents, c'est ce site qui sera chargé de la gestion en attendant la création d'un nouveau site.

L'entreprise utilise plusieurs applications :

- L'application **MakeIt**.
Elle gère la fabrication des produits 'maison'. Cela la conduit à accéder très fréquemment à la table des fournisseurs. Elle accède aussi au stock allemand pour mettre à jour au fur et à mesure de la production (consommation de matériels et produits fabriqués). Cette mise-à-jour ne porte évidemment que sur le stock du site de fabrication.
Cette application n'est déployée que sur le site d'Europe du Nord qui est en charge de la production.
- L'application **DesignIt**.
C'est le bureau d'étude qui l'utilise pour concevoir de nouveaux produits. C'est via cette application que la liste des produits fabriqués ou simplement vendus est gérée, ainsi que leurs catégories. C'est via cette application et uniquement celle-ci qu'il est possible de modifier les catégories de produits ainsi que les produits vendus.
Cette application n'est déployée que sur le site d'Europe du Sud.
- L'application **SellIt**.
Elle gère tout ce qui est lié à la vente. En particulier, les commandes, mais aussi la gestion du stock local. Elle permet aussi d'accéder à la table des fournisseurs pour demander un réapprovisionnement.
Cette application est déployée sur tous les sites. Sur chaque site, l'usage le plus fréquent est de consulter le stock local, très rarement les autres stocks, en travaillant exclusivement sur les clients locaux. Bien évidemment, elle accède régulièrement aux informations cruciales à la vente, comme les produits, les catégories, etc.
Tous les sites vendent de tout. Il n'y a aucune différenciation possible à ce niveau.
- L'application **RH**.
Elle est utilisée pour gérer les ressources humaines.
Cette application est déployée sur le site Amériques. En effet, c'est une particularité de l'entreprise, que de n'avoir qu'un seul site de gestion des ressources humaines.

Principes généraux de TP.

Vous devez réaliser ce TP en groupes de trois binômes - un seul compte rendu pour les trois binômes.

La question 1 doit être répondue en mode collégial (tout le monde participe).

Pour les questions suivantes, vous devez vous répartir les rôles. Un binôme aura en charge de gérer un site de production (Europe du Nord, Europe du Sud, Amérique).

Environnement.

Vous disposez de quatre instances d'oracle correspondant aux différents sites du scénario.

| Nom d'hôte | Port | SID | Commentaire |
|-------------|------|------|--|
| if-oracle11 | 1521 | DB11 | Site d'exploitation de la base en mode centralisé (Ryori). |
| if-oracle12 | 1521 | DB12 | Site Europe du Nord pour le déploiement en distribué |
| if-oracle13 | 1521 | DB13 | Site Europe du Sud pour le déploiement en distribué |
| if-oracle14 | 1521 | DB14 | Site Amérique pour le déploiement en distribué |

Pour chacune de ces instances, les autres ont été déclarées dans le fichier `TNSNAMES.ORA` par l'administrateur sous leurs SID respectifs.

D'un point de vue pratique, vous disposez tous de comptes sur toutes ces machines, avec les mêmes logins et les mêmes mots de passe (à changer le plus vite possible).

Sur le compte Moodle du module, le script qui a été utilisé pour la création des contraintes associées aux tables en version centralisée est disponible dans la rubrique TP.

Les données qui peuplent la base ne sont pas disponibles sous forme de script. Elles sont sur le compte Ryori de l'hôte `if-oracle11` qui est utilisé pour l'exploitation centralisée du site. Vous disposez tous des droits en lectures sur les tables de ce compte à partir des comptes sur cette instance.

Listes des pays couverts :

Europe du Nord : Norvège, Suède, Danemark, Islande, Finlande, Royaume-Uni, Irlande, Belgique, Luxembourg, Pays-Bas, Allemagne, Pologne

Europe du Sud : Espagne, Portugal, Andorre, France, Gibraltar, Italie, Saint-Marin, Vatican, Malte, Albanie, Bosnie-Herzégovine, Croatie, Grèce, Macédoine, Monténégro, Serbie, Slovénie, Bulgarie

Amérique : Antigua-et-Barbuda, Argentine, Bahamas, Barbade, Belize, Bolivie, Brésil, Canada, Chili, Colombie, Costa Rica, Cuba, République dominicaine, Dominique, Équateur, États-Unis, Grenade, Guatemala, Guyana, Haïti, Honduras, Jamaïque, Mexique, Nicaragua, Panama, Paraguay, Pérou, Saint-Christophe-et-Niévès, Sainte-Lucie, Saint-Vincent-et-les Grenadines, Salvador, Suriname, Trinité-et-Tobago, Uruguay, Venezuela.

Attention, vérifiez l'orthographe des noms de pays contenus dans la base. Ils peuvent différer sur certains points. Vérifiez aussi que tous les pays présents dans la base sont bien mentionnés ici...

Question 1.

En suivant les principes suivants

- Offrir à chaque site une gestion locale de ses clients.
- Ne prendre en compte que les informations disponibles dans le sujet.

Proposez sur papier un schéma de fragmentation concernant cette base : fragmentation verticale, fragmentation horizontale (directe et dérivée), fragmentation mixte.

Proposer sur papier un schéma de répartition (sans réplication) des fragments sur les trois sites Europe du Sud, Europe du Nord, Amérique.

Question 2.

Répartissez les sites par binômes et implantez les bases sur chaque site : création des tables, peuplement des données, contraintes d'intégrité, droit d'accès, etc.

Chaque site devra permettre d'interroger globalement la base, exactement comme si elle était centralisée (vues, synonymes...).

A la fin de cette étape, pensez à nettoyer votre travail en supprimant tout ce que vous avez mis en place et qui n'est pas utile à l'exploitation distribuée de la base.

Question 3

Réalisez les requêtes suivantes, visualisez leur plan d'exécution, commentez le.

- select * from clients;
- select * from clients where pays in (- - quelques pays du continent local);
- select * from clients_A; -- pour le continent Amériques, respectivement EuN et EuS.

Proposer d'autres requêtes qui vous permettent de vérifier que tout fonctionne bien ou qui vous semblent intéressantes à analyser. Pour chacune, suivez la même démarche.

Question 4

Pour des raisons d'optimisation, il est demandé de répliquer sur tous les sites les tables Employes, Produits, Categories, Fournisseurs.

Les répliquions seront effectuées en mono maître.

Effectuez collégialement une analyse globale des questions que cela soulève. Donnez les solutions par sites. Ces solutions devront préciser la méthode de répliquion choisie ainsi que la fréquence des mises à jour.

Chaque site devra répliquer au moins une table en read only - refresh fast et une autre en read only - refresh complete.

Pour la mise en œuvre et l'organisation générale, chaque site mettant en place une répliquion, devra communiquer par écrit au site maître du répliquat la mise en place du répliquat et, la cas échéant, il précisera ce que le site maître doit mettre en place pour le bon fonctionnement de la répliquion. Cette communication, ainsi que la réponse du site maître, sera versée au rapport final.

Enfin, si nécessaire, pensez à modifier ce qui a été réalisé à la question 2 pour que les répliquats soient pris en compte dans vos vues et autre alias.

Question 5

Testez les requêtes précédemment. Analysez les différences de plan d'exécution. Testez en de nouvelles.

Compte rendu.

Vous devez rendre votre compte rendu par groupes (trois binômes).

Ce document devra être au format pdf (lisible sur n'importe quel système).

Le nom du fichier doit strictement respecter le format suivant :

TP3IF-BDDI-Bxxxx-Byyyy-Bzzzz.pdf

où x, y et z correspondent aux numéros de binômes ayant travaillé ensemble sur le TP.

Il sera déposé sur le serveur Moodle, en utilisant la rubrique prévue à cet effet, la semaine suivant le dernier TP (pour chaque groupe de TP), le même jour de la semaine à 18h00. Aucune autre forme de rendu ne sera acceptée.

Respectez les consignes et les délais. Des pénalités de retard ainsi que des pénalités de non respect de la forme demandée seront appliquées.

Annexe 1 - Exemple de plan général de compte rendu.

- I. Rappel rapide du contexte et des objectifs du TP.
- II. Présentation du groupe de travail et des rôles de chacun
 - A. Groupe de binôme avec numéros des binômes, noms de chacun.
 - B. Identification des tâches à réaliser et répartition
 - Un binôme responsable par site à gérer.
 - Une personne responsable de la coordination générale.
 - Une personne responsable de la documentation et du rapport à rendre.
 - etc.
- III. Fragmentation

(Analyse globale et commune à tous les sites)
- A. Détermination des fragments
 1. Fragmentation horizontale
 2. Fragmentation Verticale
 3. Fragmentation mixte
 4. Bilan de la fragmentation
 - B. Placement des fragments sur les sites (sans réplication)
 1. Analyse
 2. Bilan
 - C. Mise en œuvre de la base sans réplication
 1. Site Europe du Nord
 - a) Binôme responsable
 - b) Création des liens entre les bases
 - c) Création des tables
 - d) Peuplement des tables (insertion des données provenant de la partie centralisée)
 - e) Contraintes d'intégrité
 - f) Droits d'accès
 - g) Définition de synonymes et de vues pour interrogation de la base comme si elle était en centralisé.
 - h) Nettoyages éventuels
 - i) Tests de vérification du bon fonctionnement
 2. Site Europe du Sud (réponse à la question 2bis)
 - a) ... *idem précédent*
 3. Site Amérique (réponse à la question 2ter)
 - a) ... *Idem précédent*
- IV. Tests de requête distribuées et optimisations

A. Site Europe du Sud

1. Requête : SFW...

- a) résultat d'exécution
- b) Analyse du plan d'exécution
- c) Autre écriture possible
 - (1) Résultat d'exécution
 - (2) Analyse du plan d'exécution

2. Requête : SFW...

- a) ...*Idem précédente*

B. Site Europe du Nord

1. *Idem précédent...*

C. Site Amérique

1. *Idem précédent...*

V. Répliques

A. Mise en œuvre des répliques sur le site Europe du Nord

- 1. Rappel binôme responsable
- 2. Objectifs
- 3. Liste des répliques prévues
- 4. Analyse.
(volumétrie des données, dépendances...)
- 5. Mise en œuvre des répliques pour les besoins locaux
 - a) Mise en place du réplicat du fragment AAA
 - (1) Opérations réalisées localement
 - (2) Message émis au site maître
 - (3) Réponse du site maître
 - (4) Tests de vérification de bon fonctionnement de la réplique
 - (5) Évolutions éventuelles des contraintes d'intégrité.
 - (6) Évolutions éventuelles des vues et des synonymes.
 - b) Mise en place du réplicat du fragment BBB
 - (1) *idem précédent...*
- 6. Demandes d'autres sites portant sur des fragments gérés localement
 - a) Fragment DDD
 - (1) Demande du site XXX
 - (a) Description de la demande (Texte, Date, Personne...) et suivi (messages échangés, évolution de la demande...)
 - (b) Analyse de la demande
 - (c) Opérations réalisées en local

(d) Tests de bon fonctionnement (Description, déroulement, résultats).

(2) Demande du site YYY

(a) *idem précédent...*

b) Fragment EEE

(1) *Idem précédent...*

B. Mise en œuvre des répliques sur le site Europe du Sud

C. Mise en œuvre des répliques sur le site Europe du Nord

D. Bilan global des répliques mises en œuvre sur les différents sites.
(Tableau de synthèse, etc.)

VI. Requêtes distribuées : tests et optimisations

A. Site Europe du Sud

B. Site Europe du Nord

C. Site Amérique