

Base de données distribuées



Babacar Diop

Dpt. d'Informatique

UFR des Sciences Appliquées et de Technologies

Université Gaston Berger de Saint-Louis

2018/2019

Base de données distribuées

Définition

- Plusieurs bases de données réparties géographiquement
- Une BD distribuée est un ensemble de plusieurs BDs interconnectées, réparties physiquement sur divers emplacements et communiquant via un réseau informatique
- Un SGBD distribué gère les BD distribuées de sorte qu'elle apparaisse sous la forme d'une base de données unique pour les utilisateurs

Base de données distribuées

Caractéristiques

- Les BDD sont logiquement liées les unes aux autres pour représenter une seule base de données logique
- Les données sont stockées physiquement sur plusieurs sites et gérées par un SGBD indépendant des autres sites
- Incorpore un traitement de transaction, mais ce n'est pas synonyme d'un système de traitement de transaction

Base de données distribuées

SGBDD

- Crée, récupère, met à jour et supprime les BDD
- Synchronise périodiquement la BD et fournit des mécanismes d'accès grâce auxquels la distribution devient transparente pour les utilisateurs
- Met à jour les données modifiées sur n'importe quel site universellement
- Il maintient la confidentialité et l'intégrité des données des bases de données

Base de données distribuées

Motivations

- **Nature distribuée des unités organisationnelles** - À l'heure actuelle, la plupart des organisations sont subdivisées en plusieurs unités physiquement réparties à travers le monde. Chaque unité nécessite son propre ensemble de données locales. Ainsi, la BD globale de l'organisation devient distribuée
- **Besoin de partage de données** - Les multiples unités organisationnelles doivent souvent communiquer entre elles et partager leurs données et leurs ressources. Cela nécessite des BDs communes ou répliquées qui doivent être utilisées de manière synchronisée

Base de données distribuées

Motivations

- **Prise en charge d'OLTP et OLAP** - le traitement des transactions en ligne (OLTP) et le traitement analytique en ligne (OLAP) fonctionnent sur des systèmes diversifiés pouvant contenir des données communes. Les systèmes de BDs distribués facilitent ces deux traitements en fournissant des données synchronisées
- **Récupération de base de données** - L'une des techniques couramment utilisées est la réplication de données sur différents sites. Ainsi, une défaillance de la BD peut devenir presque invisible pour les utilisateurs

Base de données distribuées

Motivations

- **Prise en charge de plusieurs logiciels d'application** –
 - La plupart des organisations utilisent divers logiciels d'application, chacun avec son support de base de données spécifique. Un SGBDD fournit une fonctionnalité uniforme pour utiliser les mêmes données sur différentes plates-formes

Base de données distribuées

Avantages

- **Développement modulaire** – l'extension à de nouvelles unités dans des systèmes de BDs centralisés nécessite des efforts importants et une perturbation du fonctionnement existant
- Dans les BDD, le travail consiste simplement à ajouter de nouveaux composants et des données locales au nouveau site et à les connecter au SD, sans interruption des fonctions actuelles

Base de données distribuées

Avantages

- **Plus fiable** - En cas de défaillance de la base de données, l'ensemble du système de BD centralisées s'arrête. Toutefois, dans les systèmes distribués, en cas de défaillance d'un composant, les performances du système peuvent continuer à diminuer. Par conséquent, un SGBDD est plus fiable.

Base de données distribuées

Avantages

- **Meilleure réponse** - Si les données sont distribuées de manière efficace, les demandes des utilisateurs peuvent être satisfaites à partir des données locales elles-mêmes, permettant ainsi une réponse plus rapide. En revanche, dans les systèmes centralisés, toutes les requêtes doivent passer par l'ordinateur central pour être traitées, ce qui augmente le temps de réponse.

Base de données distribuées

Avantages

- **Coût de communication réduit** - Dans les systèmes de BD distribués, si les données sont locales et utilisées, les coûts de communication pour la manipulation des données peuvent être minimisés. Ce n'est pas faisable dans les systèmes centralisés

Base de données distribuées

Possibles inconvénients

- **Besoin de logiciels complexes et coûteux**
- **Traitement supplémentaire** - Même les opérations les plus simples peuvent nécessiter un grand nombre de communications et des calculs supplémentaires pour uniformiser les données sur les sites

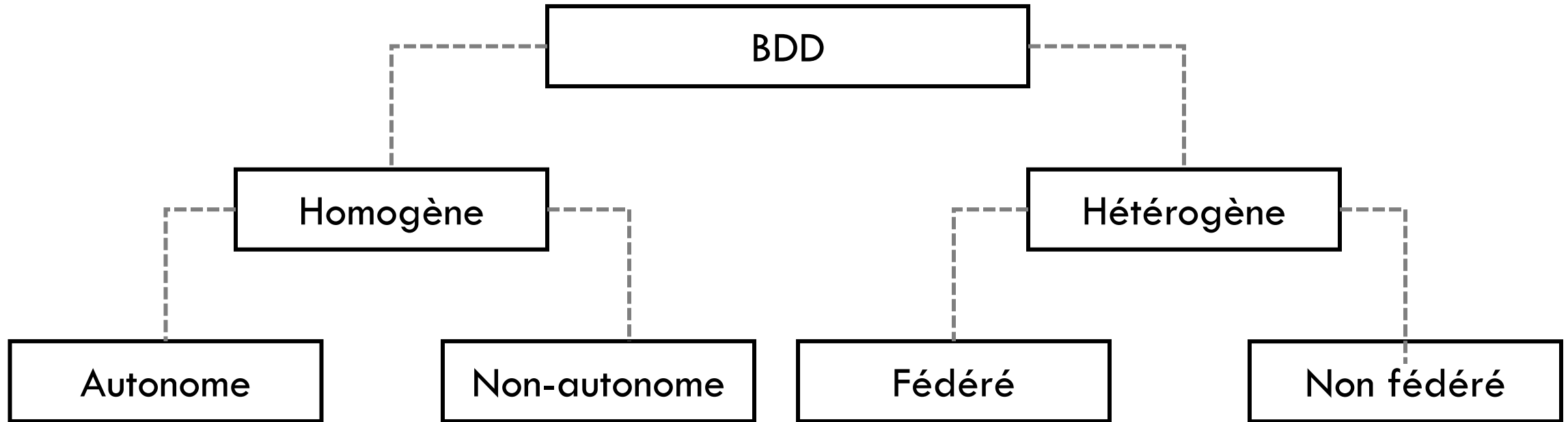
Base de données distribuées

Possibles inconvénients

- **Intégrité des données** - La nécessité de mettre à jour les données sur plusieurs sites pose des problèmes d'intégrité des données
- **Autres problèmes en cas de mauvaise distribution des données** - La réactivité des requêtes dépend en grande partie de la bonne répartition des données. Une mauvaise distribution des données entraîne souvent une réponse très lente

Base de données distribuées

Catégories de BDD



BDD homogènes

Concepts

- Tous les sites utilisent :
 - un SGDB et un OS identiques
 - des logiciels très similaires
 - des bases de données identiques
- Chaque site est informé de tous les autres sites et coopère avec d'autres sites pour traiter les demandes des utilisateurs
- La BD est accessible via une interface unique, comme s'il s'agissait d'une BD unique

BDD homogènes

Catégories

- Il existe deux types de BDD homogène
 - **Autonome** - Chaque BD est indépendante et fonctionne de manière autonome. Elles sont intégrées par une application de contrôle et utilisent la transmission de messages pour partager les mises à jour des données
 - **Non autonome** - Les données sont réparties sur les nœuds homogènes et un SGBD central ou principal coordonne les mises à jour des données sur les sites

BDD hétérogènes

Concepts

- Dans une BDD hétérogène, différents sites ont des systèmes d'exploitation, des produits de SGBD et des modèles de données différents
- Différents sites utilisent des schémas et des logiciels différents (relationnel, réseau, hiérarchique, orienté objet, etc.)
- Traitement complexe de requêtes en raison de schémas différents
- Traitement complexe des transactions en raison de logiciels différents
- Un site peut ne pas être au courant d'autres sites et la coopération dans le traitement des demandes des utilisateurs est donc limitée

BDD hétérogènes

Catégories

- **Fédéré** - Les systèmes de BD hétérogènes sont de nature indépendante et intégrés ensemble, de sorte qu'ils fonctionnent comme un système de BD unique
- **Non fédéré** - Les systèmes de BD utilisent un module de coordination central permettant d'accéder aux BDs