

420-C42

Langages d'exploitation des bases de données

Partie 1

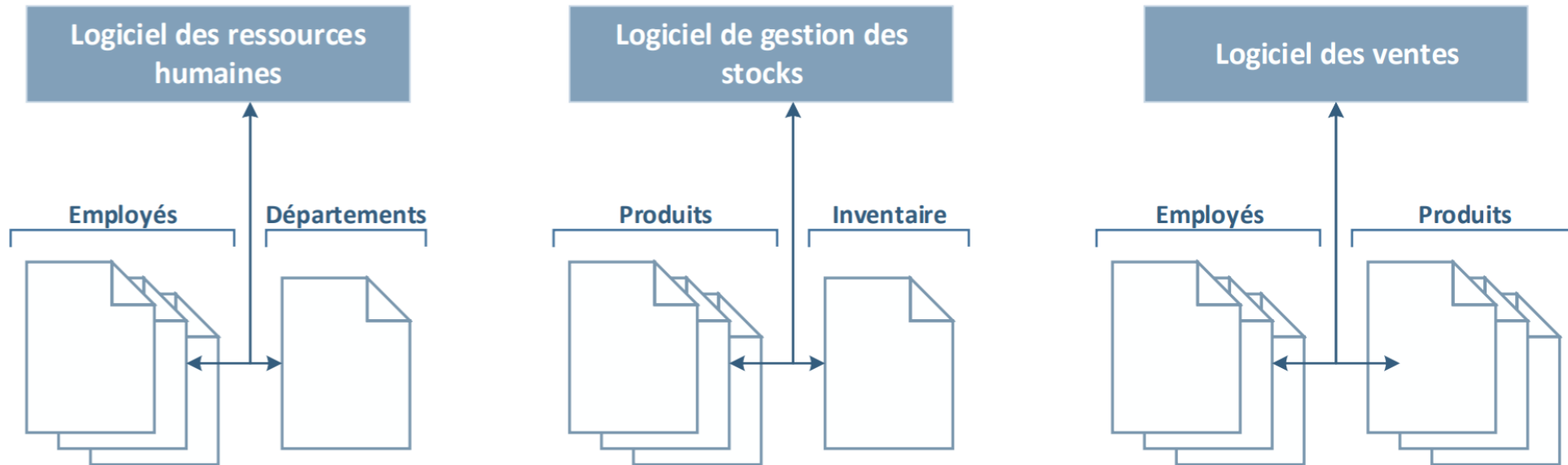
Introduction aux bases de données relationnelles

La gestion d'information

- Un ordinateur ne fait que manipuler de l'information
- La gestion de données est omniprésente :
 - dans tous les projets
 - depuis le début de l'ère informatique
- Exemples où la gestion de données existe

Une démarche inefficace

- Une démarche naïve consiste à développer un programme d'application par groupe informationnel ou par tâche.



Une démarche inefficace

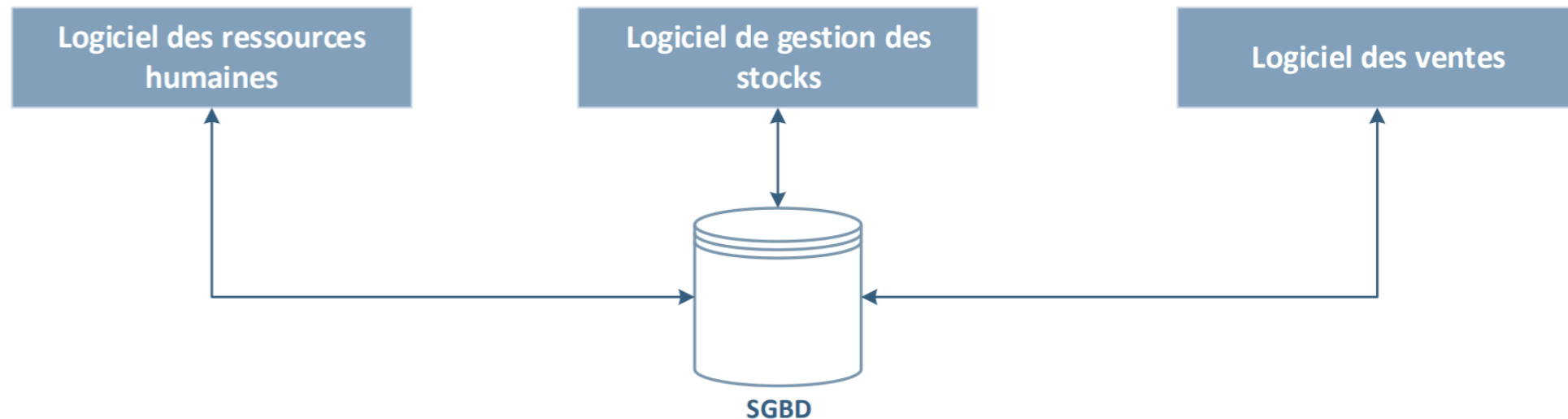
- Cette approche décentralisée apporte plusieurs problématiques :
 - redondance de certaines informations engendrant :
 - inconsistance potentielle des données entre les sous-systèmes
 - intégrité des données
 - augmentation de la charge de traitement pour garantir la synchronicité
 - un accroissement inutile de :
 - nombre de fichiers
 - taille des fichiers
 - temps d'accès
 - charge de calcul
- outils logiciels développés par différentes équipes utilisant :
 - différentes technologies de programmation
 - différents format de données et de fichiers

Une démarche inefficace

- Cette approche décentralisée apporte plusieurs problématiques :
 - complexifie le développement de tâche connexes et récurrentes
 - sécurité des données
 - protection des données
 - accès concurrent
 - ...
 - coûts de développement et de maintenance très élevés (exponentiel dans le temps)
 - maintenance des données par des fichiers gérés par le système d'exploitation

Les bases de données

- Une base de données, ou plus précisément un système de gestion de base de données (SGBD - *DBMS*), est un ensemble de logiciels donnant accès à :
 - un conteneur de données
 - des services reliés



Les bases de données

- Les SGBD modernes permettent une gestion efficace des données en visant la mise en place d'au moins sept objectifs complémentaires :
 1. Centralisation
 2. Indépendance
 3. Liaison
 4. Intégrité
 5. Sécurité
 6. Confidentialité
 7. Partage

Les bases de données

- La **centralisation** consiste à concentrer l'information en un seul endroit logique (pas nécessairement physique) :
 - élimine les redondances d'information
 - permet l'unicité de l'existence d'une donnée (saisie, stockage et consultation)
 - facilite tous les services reliés au SGBD

Les bases de données

- L'**indépendance** entre les données et le traitement consiste à appliquer une notion fondamentale en informatique : la séparation entre les conteneurs et le traitement des données.
- Cette dissociation, si importante en génie logiciel, apporte plusieurs avantages importants :
 - permet à plusieurs systèmes d'accéder aux mêmes données afin d'en faire différentes tâches : affichage, mise à jour, liaison, calcul, ...
 - permet différentes vues selon les logiciels en action (seulement les données requises sont utilisées)
 - offre une flexibilité sans comparable
- Cette indépendance réduit les coûts de développement et de maintenance : à travers divers projets et pendant l'évolution de ces derniers.

Les bases de données

- La **liaison** consiste à définir les liens logiques existant entre les données.
- Permet ainsi de créer des structures complexes indépendantes des données.
- Comme on le verra, la mise en œuvre de liaisons permet de garantir la non redondance, l'intégrité et l'efficacité des données liées.

Les bases de données

- L'**intégrité** consiste à assurer la fiabilité et la cohérence de toutes les données en les considérant unitairement ou en relation.
- Préserver l'intégrité des données est certainement le premier rôle du SGBD.
- L'intégrité est assuré par plusieurs mécanismes internes au SGBD tels que :
 - le format (entier, point flottant, date, ...)
 - le domaine de validité (intervalle, appartenance à une liste, ...)
 - l'unicité d'une donnée
 - l'existence d'une donnée
 - règles diverses (existence d'un lien, valeur par défaut, ...)

Les bases de données

- La **sécurité** des données consiste à assurer l'intégrité du système malgré les pannes physiques éventuelles (inévitables).
- Si l'intégrité des données est la première tâche du SGBD, la sécurité de ces dernières en est certainement la deuxième.
- Les causes de pannes potentielles sont nombreuses.
- C'est la mise en place des mécanismes de récupération et de reprise qui permet de garantir la sécurité des données peu importe les avaries.
- À ces mécanismes sont liés plusieurs stratégies de gestion (transaction, ...) et infrastructures matériels (RAID, clusters, ...)

Les bases de données

- La **confidentialité** des données consiste à protéger le contenu des données sensibles.
- Plusieurs données ne doivent pas être disponible à tous les usagers du SGBD.
- Les notions d'utilisateur et de droits d'accès permettent de limiter les accès.

Les bases de données

- Le **partage** des données consiste à permettre les accès concurrents au SGBD par plusieurs usagers simultanément.
- Par exemple, deux individus accèdent au même compte et font une transaction.
- La notion de transaction permet d'assurer l'intégrité par la mise en place du mécanisme de verrouillage.

Les bases de données

- Les SGBD permettent de :
 - créer les structures de données inhérentes à un projet;
 - créer des liens entre les données;
 - manipuler les données (insertion, modification, suppression et consultation);
 - gérer les droits d'accès aux données en fonction des autorisations données aux utilisateurs;
 - administrer le SGBD afin d'atteindre les performances optimales en fonction des particularités du projet.
- Les SGBD peuvent faire tout ça
 - en respectant les sept critères énoncés
 - tout en masquant la complexité sous-jacente!

Acteurs

- On retrouve principalement six groupes d'acteurs interagissant avec les SGBD :
 - Les usager
 - Les analystes
 - Les développeurs d'application
 - Les concepteurs de base de données
 - Les administrateurs de base de données (DBA)
 - Les développeurs des SGBD

Attention!

- L'utilisation d'un SGBD
 - n'est pas toujours une solution facile à mettre en œuvre, à déployer et à maintenir;
 - peut engendrer des coûts substantiels (matériel, logiciel, main d'œuvre spécialisée, développement, formation, maintien, ...);
 - les performances du système peuvent être inadéquates pour certains types d'application (surcharge de tous les services) et il peut être nécessaire d'embaucher des spécialistes avec des compétences difficiles à trouver;
 - peut être mésadaptée pour certains projets simples, bien définis, n'évoluant pas dans le temps et surtout ne requérant pas les services offerts.