**Notion sur les Bases de données :**

**1. INTRODUCTION :**

 Tout traitement informatique consiste à appliquer des opérations sur des données plus ou moins complexes. Ainsi, au sein d’un programme écrit dans un langage informatique quelconque (Fortran, Ada, C++ ...), les instructions manipulent des variables, c’est-à-dire des objets électroniques : nombres, chaînes de caractères, fichiers, etc. Les fichiers correspondent à une catégorie de variables permettant de stocker de gros volumes de données résidant dans les mémoires auxiliaires (magnétiques ou optiques) d’un ordinateur.

 Dès le début de l’informatique, les fichiers ont été largement utilisés dans les applications scientifiques et de gestion : fichiers de données statistiques, fichiers de produits, fichiers de clients... De nombreux problèmes apparurent dans la gestion des fichiers tels que la lourdeur d’exploitation, les incohérences, la redondance. Ils ont conduit à l’avènement des bases de données regroupant l’ensemble des fichiers nécessaires à une ou plusieurs applications et gérées par un Système de Gestion de Bases de Données.

**2. QU’EST-CE QU’UNE BASE DE DONNEE?**

 Une Base de Données (BD), en anglais Data Base (DB), est une collection d’informations structurées, cohérentes et persistantes.
 L'un des avantages d'une base de données est que les informations qui la constituent peuvent être accessibles aisément par plusieurs programmes qui les utilisent simultanément avec des objectifs différents.
 Une base de données peut être locale ou répartie. Elle est dite locale quand c'est utilisée sur une machine par un utilisateur et repartie quand les informations sont stockées sur des machines distantes (serveur) et accessibles par réseau.

**3. HISTORIQUE DES BASES DE DONNES :**

**Années 60 :** Systèmes de gestion de **fichiers** plus ou moins sophistiqués,

reliés par des pointeurs.

• programmes de la conquête spatiale (APOLLO)

Systèmes de gestion de base de données **réseaux** (travaux de CODASYL) et **hiérarchiques** proches des systèmes de gestion de fichiers.

• TOTAL, IDMS, IDS.2 et IMS.2, etc.

**Années 70:** Codd, mathématicien et chercheur chez IBM, propose le modèle

**Relationnel** (basé sur le concept de Relation de la théorie des ensembles)

Codd définit l’**Algèbre Relationnelle** (base du SQL)

**Années 80: Systèmes de gestion de base de données relationnels**

• SGBD commercialisés à partir de 1980

• ORACLE, SYSBASE, DB2, etc.

**Années 90:** Les systèmes de gestion de base de données relationnels dominent le marché.

• Extension du relationnel (Oracle 8, DB2 Universal Database, etc.)

Début des systèmes de gestion de bases de données orientés objet.

• ObjectStore, GemStone, O2, etc.

**4ème génération:** Internet, les informations non structurées, le multimédia, etc.…

**4. LIMITE DES SYSTEMES DE FICHIERS :**

 Un **système de fichiers** (file system ou file system en anglais) ou **système de gestion de fichiers** (SGF) est une façon de stocker les informations et de les organiser dans des fichiers sur ce que l'on appelle des mémoires secondaires ([disque dur](http://fr.wikipedia.org/wiki/Disque_dur), [SSD](http://fr.wikipedia.org/wiki/Solid_State_Drive), [CD-ROM](http://fr.wikipedia.org/wiki/CD-ROM), [clé USB](http://fr.wikipedia.org/wiki/Cl%C3%A9_USB), [disquette](http://fr.wikipedia.org/wiki/Disquette), etc.). Une telle gestion des [fichiers](http://fr.wikipedia.org/wiki/Fichier_informatique) permet de traiter, de conserver des quantités importantes de [données](http://fr.wikipedia.org/wiki/Donn%C3%A9e_%28informatique%29) ainsi que de les partager entre plusieurs [programmes informatiques](http://fr.wikipedia.org/wiki/Programme_informatique). Il offre à l'[utilisateur](http://fr.wikipedia.org/wiki/Utilisateur_%28informatique%29) une vue abstraite sur ses [données](http://fr.wikipedia.org/wiki/Donn%C3%A9e_%28informatique%29) et permet de les localiser à partir d'un chemin d'accès.

Il existe d'autres façons d'organiser les [données](http://fr.wikipedia.org/wiki/Donn%C3%A9e_%28informatique%29), par exemple les [bases de données](http://fr.wikipedia.org/wiki/Base_de_donn%C3%A9es) (notamment [base de données relationnelle](http://fr.wikipedia.org/wiki/Base_de_donn%C3%A9es_relationnelle)) et les [fichiers indexés](http://fr.wikipedia.org/wiki/Fichier_index%C3%A9).

**5. ORGANISATION D’UNE BASE DE DONNEES:**

 L'organisation physique sous-jacente du médium utilisé (blocs, linéaire) et les mécanismes d'entrée/sortie de bas-niveau sont masqués. L'utilisateur peut donc organiser ses données permanentes en les distribuant dans différents fichiers. Le contenu des fichiers est déterminé par leur [format](http://fr.wikipedia.org/wiki/Format_de_donn%C3%A9es), qui dépend de l'application utilisée.

 En plus de cette organisation abstraite, les systèmes de fichiers peuvent inclure la compression ou le chiffrement automatique des données, une gestion plus ou moins fine des droits d'accès aux fichiers, et une [journalisation des écritures](http://fr.wikipedia.org/wiki/Journal_%28syst%C3%A8me_de_fichiers%29) (pour la robustesse, en cas de défaillance du système). De plus, certains systèmes de fichiers peuvent s'étendre sur un réseau entier, comme NFS. Certains de ces systèmes de fichiers en réseau peuvent être distribués ou répartis, comme PVFS2.

**6. PROPRIETES D’UNE BASE DE DONNEES :**

• Usage multiple des données.

• Accès facile, rapide, protégé, souple, puissant.

• Coût réduit de stockage, de mise à jour et de saisie.

• Disponibilité, exactitude, cohérence et protection des données ; non redondance.

• Évolution aisée et protection de l'investissement de programmation.

• Indépendance des données et des programmes.

• Conception *a priori.*

**7. OBJECTIFS D’UNE BASE DE DONNÉES :**

 Les objectifs que l’on assigne généralement aux bases de données et aux systèmes qui les supportent sont les suivants :

* la centralisation.
* l’indépendance entre les données et les traitements.
* la structuration de données complexes.
* le partage des données.
* la confidentialité.
* la sécurité.

 **La centralisation** des données pour objet de limiter la redondance, c’est-à-dire d’éviter la présence de duplicata de données. L’information n’étant pas dupliquée, il s’ensuit une unicité de la saisie ainsi qu’une centralisation des contrôles qui peuvent être réalisés par le SGBD.

 **L’indépendance** entre les données et les traitements vise à permettre une évolution des structures de données sans répercussion sur les programmes d’application utilisant ces données. Par exemple, il s’agit de garantir le fait de ne pas avoir à recompiler un programme à la suite de la modification d’un format d’une donnée qui ne le concerne pas. Pour atteindre cet objectif, il a été.

 **La structuration de données** la nécessite de séparer, dans la description d'une structure de données, les propriétés abstraites des objets et leur représentation concrète, est reconnue depuis longtemps sur I.e. plan des idées.

Dans [7], par exemple, d'Imperio faisait déjà la distinction entre "structures de données" et Structures de mémoire". En pratique, cependant, cette distinction reste souvent un vœu pieux.

Et la manipulation concrète des structures de données mélange caractéristiques logiques et problèmes

 **Le partage de fichiers** est une technique consistant à distribuer ou à donner accès, à distance, à des [données numériques](http://fr.wikipedia.org/wiki/Donn%C3%A9es_num%C3%A9riques). Il peut s'agir de fichiers de toutes sortes : logiciels, livres, [vidéo](http://fr.wikipedia.org/wiki/Vid%C3%A9o), [audio](http://fr.wikipedia.org/wiki/Format_de_fichier_audio) …etc.

Deux techniques de partage de fichiers existent actuellement : l'hébergement centralisé (modèle [client-serveur](http://fr.wikipedia.org/wiki/Client-serveur)) permet de stocker les données sur un [serveur de fichiers](http://fr.wikipedia.org/wiki/Serveur_de_fichiers) unique et d'y accéder sur celui-ci depuis un autre ordinateur (dit le client).

La technique [pair-à-pair](http://fr.wikipedia.org/wiki/Pair-%C3%A0-pair) qui consiste à mettre des données en partage suivant un modèle de réseau informatique où chaque ordinateur [client](http://fr.wikipedia.org/wiki/Client_%28informatique%29) est aussi un [serveur](http://fr.wikipedia.org/wiki/Serveur_informatique).

 **La confidentialité** est le fait de s'assurer que l'information n'est seulement accessible qu'à ceux dont l'accès est autorisé selon l'ISO (Organisation Internationale de Normalisation). C'est un principe éthique associé à l'informatique, mais aussi à d'autres professions telles que la vente, le droit, la médecine ou le journalisme.

 **La sécurité** informatique est l'ensemble des techniques qui assurent que les ressources du système d'information (matérielles ou logicielles) d'une organisation sont utilisées uniquement dans le cadre où il est prévu qu'elles le soient.

**8. Processus de conception d’une base de données :**

Le processus de conception se compose des étapes suivantes :

* **Déterminer les objectifs de la base de données**

Cette étape vous aide à vous préparer pour les étapes restantes.

* **Rechercher et organiser les informations requises**

Rassembler tous les types d'informations que vous souhaitez stocker dans la base de données, tels que les noms de produit et les numéros de commande.

* **Répartir les informations dans des tables**

Répartissez vos éléments d'information en entités ou sujets principaux, tels que Produits ou Commandes. Chaque sujet devient ensuite une table.

* **Convertir des éléments d'information en colonnes**

Choisissez les informations que vous souhaitez stocker dans chaque table. Chaque élément d'information devient un champ et est affiché sous la forme d'une colonne dans la table. Par exemple, une table Employés peut contenir des champs du type Nom et Date d'embauche.

* **Définir des clés primaires**

Choisissez la clé primaire de chaque table. La clé primaire est une colonne qui sert à identifier de façon unique chaque ligne par exemple, une colonne Numéro de produit ou Numéro de commande.

* **Définir les relations entre tables**

Étudiez chaque table et déterminez de quelle façon les données d'une table sont liées aux données des autres tables. Ajoutez des champs aux tables ou créez de nouvelles tables pour clarifier les relations, le cas échéant.

* **Affiner la structure**

Analysez votre conception et recherchez les erreurs qu'elle peut contenir. Créez les tables et ajoutes des enregistrements contenant des données exemple. Vérifiez si vous obtenez les résultats attendus des tables créées. Apportez des modifications en conséquence.

* **Appliquer les règles de normalisation**

Appliquez les règles de normalisation des données pour vérifier si vos tables sont structurées correctement. Si nécessaire, apportez des modifications.

# 9. Système de gestion de base de données :

**9.1. Définition de SGBD :**

 Est une suite de programmes qui manipule la structure de la base de données et dirige l'accès aux données qui y sont stockées. Une base de données est composée d'une collection de fichiers; le seul moyen d'accéder aux données est par le SGBD, qui sert alors d'intermédiaire entre la base de données et ses usagers. Celui-ci reçoit des demandes de manipulation du contenu et effectue les opérations nécessaires sur les fichiers. Il cache la complexité des opérations et offre une vue synthétique sur le contenu. Le SGBD permet en outre à plusieurs usagers de manipuler simultanément le contenu, et peut offrir différentes vues sur un même ensemble de données.

**9.2. Fonctions des SGBD :**

***a) Description des données***

Aux niveaux externes, internes, conceptuels par les administrateurs grâce à un *Langage de Définition de Données* (LDD)

***b) Recherche de données***

***c) Mise à jour des données***

Insertion, modification, suppression grâce à un *Langage de Manipulation de Données* (LMD)

***d) Transformation des données***

Ex. Changement de format : date sur 2 chiffres ®date sur 4 chiffres

*Bases de données – Introduction* 6

***e) Contrôle de l’intégrité des données***

Exactitude des données stockées dans la base (respect des contraintes d’intégrité)

Maintien de la cohérence entre les copies multiples d’une même donnée

***f) Gestion de transactions et sécurité***

Gestion des transactions = s’assurer qu’un groupe de mises à jour est totalement

exécuté ou pas du tout (*atomicité* des transactions)

Sécurité = personnalisation des accès à la base, mots de passe

***g) Autres fonctions***

Ex. *Déclencheurs* (triggers)

# 9.3. Types de SGBD :

#  On site les trois types de SGBD suivantes :

# DBase :

# *DBASE* est un [SGBD](http://fr.wikipedia.org/wiki/SGBD) destiné à faire partager des fichiers de données par plusieurs utilisateurs. [Ashton-Tate](http://fr.wikipedia.org/wiki/Ashton-Tate) a été l'éditeur d'origine pour les versions dBASE II, dBASE III et dBASEIII Plus. En [1991](http://fr.wikipedia.org/wiki/1991), il est racheté par [Borland](http://fr.wikipedia.org/wiki/Borland) qui sort la version dBASE IV et dBASE V sous Windows. Aujourd'hui c'est une société américaine, dataBased Intelligence Inc. qui développe le logiciel sous le nom DBASE PLUS.

# Access de Microsoft :

#  (Officiellement *Microsoft Office Access*) est un [SGBD](http://fr.wikipedia.org/wiki/Syst%C3%A8me_de_gestion_de_base_de_donn%C3%A9es) relationnel édité par [Microsoft](http://fr.wikipedia.org/wiki/Microsoft). Il fait partie de la suite bureautique [MS Office Pro](http://fr.wikipedia.org/wiki/Microsoft_Office). Il est simple à mettre en œuvre mais n'est adaptée que pour de très petites quantités d'informations et peu d'utilisateurs.

# MSQL et MYSQL :

#  Ces systèmes sont tous les deux disponibles pour des plates-formes Unix-Linux. Leur mise en œuvre est rapide et demande un minimum de compétences en SGBD.