Base de données : qu'est-ce que c'est ? Définition et présentation

Sommaire

- 1.Qu'est-ce qu'une base de données ? Définition
- 2.Comment fonctionnent les bases de données ?
- 3.L'histoire des bases de données
- 4.Quels sont les différents types de bases de données
  - 4.1.Base de données hiérarchique
  - 4.2.Base de données réseau
  - 4.3.Base de données orientée texte
  - 4.4.Base de données SQL (relationnelle)
  - 4.5.Base de données distribuée
  - 4.6.Base de données cloud
  - 4.7.Base de données NoSQL
  - 4.8.Base de données orientée objets
  - 4.9.Base de données orientée graph
- 5.Les bases de données relationnelles et non relationnelles peuvent-elles coexister ?
- 6.Comment le cloud transforme les bases de données
- 7.SGBD (système de gestion de base de données) ou DBMS et RDBMS : les logiciels qui permettent d'accéder aux bases de données
- 8.Qu'est-ce que le SQL ou Structured Query Language ?
- 9.Qu'est-ce qu'un administrateur de base de données (DataBase Administrator) ?
- 10.Quelles sont les meilleures bases de données à l'heure actuelle ?
  - 10.1.Microsoft SQL Server, la base de données Windows
  - 10.2.Oracle, la base de données la plus populaire pour Linux/Unix
  - 10.3.IBM DB2, la base de données Mainframe la plus populaire
  - 10.4.Teradata, la meilleure base de données pour le Big Data
  - 10.5.SAP Sybase, l'ancien leader du marché
  - 10.6.Informix, une base de données rachetée par IBM
- 11.MySQL : la meilleure base de données relationnelle Open Source
- 12.Logiciel de base de données gratuit

- 13. Quelles sont les meilleures bases de données NoSQL ?
  - 13.1. Les bases de données NoSQL orientées clé-valeur
  - 13.2. Les bases de données NoSQL orientées document
  - 13.3. Les bases de données NoSQL orientées colonne
  - 13.4. Les bases de données NoSQL orientées graphe

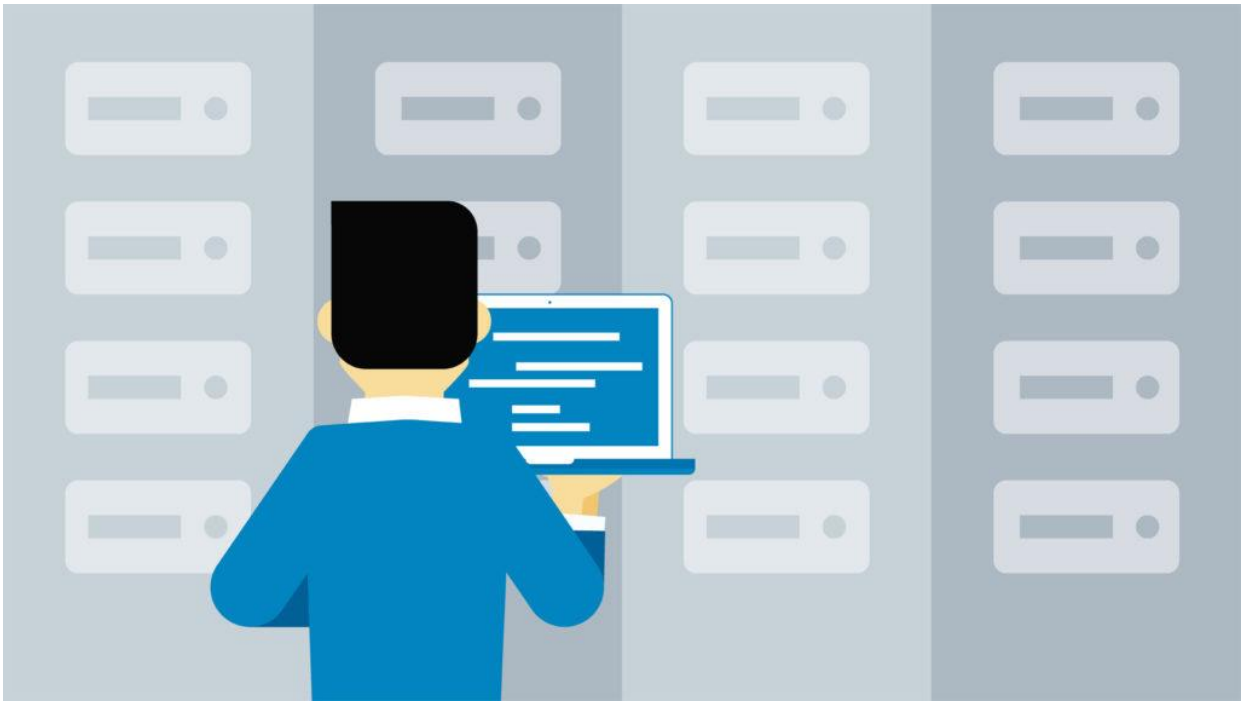
Les bases de données informatiques sont utilisées dans un grand nombre d'entreprises pour stocker, organiser et analyser les données. Découvrez tout ce que vous devez savoir à ce sujet : qu'est-ce qu'une base de données, à quoi sert-elle, comment fonctionne-t-elle, quelles sont les différentes catégories, et quelles sont les meilleures.

Une base de données est une **collection d'informations organisées afin d'être facilement consultables, gérables et mises à jour**. Au sein d'une base de données, les données sont organisées en lignes, colonnes et tableaux. Elles sont indexées afin de pouvoir facilement trouver les informations recherchées à l'aide d'un logiciel informatique. Chaque fois que de nouvelles informations sont ajoutées, les données sont mises à jour, et éventuellement supprimées.

Qu'est-ce qu'une base de données ? Définition

Les bases de données se chargent elles-mêmes de créer, de mettre à jour ou de supprimer des données. Elles se chargent également **d'effectuer des recherches parmi les données qu'elles contiennent sur demande de l'utilisateur**, et de lancer des applications à partir des données.
Les bases de données sont utilisées par de nombreuses entreprises dans toutes les industries. Elles sont notamment utilisées par les **compagnies aériennes pour gérer les réservations**. Elles sont utilisées pour la gestion de production. Pour **les enregistrements médicaux dans les hôpitaux**, ou encore pour les **enregistrements légaux dans les compagnies d'assurances**. Les bases de données les plus larges sont généralement utilisées **par les agences gouvernementales**, les grandes entreprises ou les universités.

Comment fonctionnent les bases de données ?



Les bases de données sont stockées sous forme de fichiers ou d'ensemble de fichiers sur un disque magnétique, une cassette, un disque optique ou tout autre type d'appareil de stockage. Les bases de données traditionnelles (hiérarchiques) sont organisées par **champs (fields), enregistrements et fichiers**. Un champ est une seule pièce d'information. Un enregistrement est un ensemble de champs. Un fichier est une collection d'enregistrements.
Par exemple, un **répertoire téléphonique est l'équivalent d'un fichier**. Il contient un ensemble d'enregistrements, et chaque enregistrement regroupe trois champs : nom, adresse et numéro de téléphone. En guise d'exemple, on peut aussi citer les catalogues de produits ou les inventaires.
La faculté de consulter ou de modifier une base de données (lire ou écrire) est conférée aux divers utilisateurs par un **database manager (gestionnaire de base de données)**. Les bases de données sont principalement présentes au sein des systèmes mainframe les plus larges, mais on les retrouve aussi au sein de workstations distribuées de moindre envergure et autres systèmes midrange comme les IBM AS/400 ou même les ordinateurs personnels.

L'histoire des bases de données

L'histoire des bases de données **remonte aux années 1960**, avec l'apparition des bases de données réseau et des bases de données hiérarchiques. Dans les années 1980, ce sont les bases de données object-oriented qui ont fait leur apparition. Aujourd'hui, les bases de données prédominantes sont les SQL, NoSQL et bases de données cloud.

Il est aussi possible de classer les bases de données en fonction de leur contenu : bibliographique, textes, nombres ou images. Toutefois, en informatique, on classe généralement les bases de données en fonction de leur approche organisationnelle. Il existe de nombreux types de bases de données différentes : **relationnelle, distribuée, cloud, NoSQL**… Voici les différents types de bases de données.

Quels sont les différents types de bases de données

Dans le cas d'une grande base de données, les multiples utilisateurs doivent être en mesure de manipuler les informations qu'elle contient rapidement et n'importe quand. De plus, les grandes entreprises ont tendance à cumuler de nombreux fichiers indépendants comprenant des fichiers liés ou même des données se superposant. Dans le cadre d'une analyse de données, il est nécessaire que les données en provenance de plusieurs fichiers puissent être liées. C'est pourquoi **différents types de bases de données ont été développé pour répondre à ces exigences** : orientée texte, hiérarchique, réseau, relationnelle, orientée objet…

Base de données hiérarchique

Les bases de données hiérarchiques comptent parmi **les plus anciennes bases de données.** Au sein de cette catégorie, les enregistrements sont organisés dans une structure d'arborescence. Chaque niveau d'enregistrements découle sur un ensemble de catégories plus petites.

Base de données réseau

Les bases de données réseau sont également parmi les plus anciennes. Plutôt que de proposer des liens uniques entre différents ensembles de données à divers niveaux, les bases de données réseaux créent **des liens multiples entre les ensembles** en plaçant des liens, ou des pointeurs, sur un ensemble d'enregistrements ou un autre. La vitesse et la polyvalence des bases de données réseau ont conduit à une adoption massive de ce type de databases au sein des entreprises ou dans le domaine du e-commerce.

Base de données orientée texte

Une base de données orientée texte, ou flat file database, se présente sous la forme d'un **fichier (une table) au format .txt ou .ini**. Un fichier plat est un fichier texte, ou un fichier combinant du texte avec un fichier binaire. En général, dans ces bases de données, chaque ligne ne comporte qu'un enregistrement. La plupart des bases de données pour PC sont des bases de données orientées texte.

Base de données SQL (relationnelle)

Les bases de données relationnelles ont été inventées en 1970 par E.F. Codd de IBM. Une base de données relationnelle est une base de données tabulaire dans laquelle**les données sont définies afin d'être accessibles** et de pouvoir être réorganisées de différentes manières.
Les bases de données relationnelles sont constituées d'un ensemble de tableaux. Au sein de ces tableaux, les **données sont classées par catégorie**. Chaque tableau comporte au moins une colonne correspondant à une catégorie. Chaque colonne comporte un certain nombre de données correspondant à cette catégorie.
L'API standard pour les bases de données relationnelles est **le Structured Query Language (SQL)**. Les bases de données relationnelles sont facilement extensibles, et de nouvelles catégories de données peuvent être ajoutées après la création de la base de données originale sans avoir besoin de modifier toutes les applications existantes.

Base de données distribuée

Une base de données distribuée est une base de données dont **certaines portions sont stockées à plusieurs endroits** physiques. Au sein de cette base de données, le traitement est réparti ou répliqué entre différents points d'un réseau.
Les bases de données distribuées **peuvent être homogènes ou hétérogènes**. Dans le cas d'un système de base de données distribuée homogène, tous les emplacements physiques fonctionnent avec le même hardware et tournent sous le même système d'exploitation et les mêmes applications de bases de données. Au contraire, dans le cas d'une base de données distribuée hétérogène, le hardware, les systèmes d'exploitation et les applications de bases de données peuvent varier entre les différents endroits physiques.

Base de données cloud

Une base de données cloud est une base de données **optimisée ou directement créée pour les environnements virtualisés**. Il peut s'agir d'un cloud privé, d'un cloud public ou d'un cloud hybride.
Les **bases de données cloud offrent plusieurs avantages** comme la possibilité de payer pour la capacité de stockage et la bande passante en fonction de l'usage. Par ailleurs, il est possible de changer l'échelle sur demande. Ces bases de données offrent aussi une disponibilité plus élevée.

Base de données NoSQL

Les bases de données NoSQL sont utiles **pour les larges ensembles de données distribuées**. En effet, les bases de données relationnelles ne sont pas conçues pour le Big Data, et les ensembles de données trop larges peuvent poser des problèmes de performances.
Si une entreprise doit **analyser d'importantes quantités de données non structurées**, ou des données stockées sur plusieurs serveurs cloud virtuels, la base de données NoSQL est idéale. Avec l'essor du Big Data, les bases de données NoSQL sont de plus en plus utilisées.

Base de données orientée objets

Les **objets créés à l'aide de langage de programmation orientés objets** sont généralement stockés sur des bases de données relationnelles. Toutefois, en réalité, les bases de données orientées objets sont plus adaptées pour stocker ce type de contenu.
Plutôt que d'être organisée autour d'actions, les bases de données orientées objets sont **organisées autour d'objets**. De même, au lieu d'être organisées autour d'une logique, elles sont organisées autour des données. Par exemple, un enregistrement multimédia au sein d'une base de données relationnelle peut être défini comme un objet de données plutôt que comme une valeur alphanumérique.

Base de données orientée graph

Une base de données orientée graphe, ou base de données graphe, est un **type de base de données NoSQL utilisant la théorie des graphes** pour stocker, cartographier et effectuer des requêtes sur les relations entre les données. Les bases de données graphe sont constituées de noeuds et de bords.
Chaque noeud représente une entité, et chaque bord représente une connexion entre les noeuds. Les bases de données graphes gagnent en popularité dans le domaine des analyses d'interconnexions. Par exemple, les entreprises peuvent utiliser une base de données graphe pour **miner des données sur ses clients à partir des réseaux sociaux**.
De plus en plus souvent, des bases de données jadis séparées sont combinées électroniquement sous forme de collections plus larges **que l'on appelle les Data Warehouses**. Les entreprises et les gouvernements utilisent ensuite des logiciels de Data

Mining pour analyser les différents aspects des données. Par exemple, une agence gouvernementale peut procéder ainsi pour enquêter sur une entreprise ou une personne qui ont acheté une grande quantité d'équipement, même si les achats sont disséminés dans tout le pays ou répartis entre plusieurs subsidiaires.

Les bases de données relationnelles et non relationnelles peuvent-elles coexister ?

Il peut être tentant de voir **les bases de données NoSQL comme un remplacement**pour les bases de données relationnelles. Toutefois, en réalité, il peut y avoir de la place pour ces deux types de technologies dans la plupart des entreprises.
Les bases de données SQL et NoSQL **prennent en charge les informations d'une façon différente** et prennent en charge différents types de workloads. Plutôt que de prendre la place des bases de données relationnelles, les bases de données NoSQL permettent aux entreprises de viser de nouveaux objectifs, de relever de nouveaux défis.
Il est important de bien savoir **dans quelle situation utiliser une base de données NoSQL**, et dans quelle situation il est préférable d'utiliser une base de données relationnelle. En fonction de la nature du workload et des données sous-jacentes, il convient d'utiliser une base de données relationnelle ou une base de données NoSQL.

Ainsi, les bases de données **NoSQL conviennent davantage aux grands ensembles**de données fréquemment exposés à de nouvelles informations, au sein desquels les enregistrements ont des structures variables ne se mariant pas bien aux modèles relationnels.
Une autre raison pour laquelle les bases de données non relationnelles ne remplaceront pas les bases de données relationnelles est que les vendeurs de bases de données relationnelles ont amélioré leurs produits pour les adapter aux applications Big Data. Non seulement **ces deux technologies ont encore leur place en entreprise**, mais il y a aussi de la place pour qu'elles poursuivent toutes les deux leur croissance de leur côté.

Comment le cloud transforme les bases de données

Les bases de données cloud se distinguent des bases de données traditionnelles par le fait qu'elles **résident sur le cloud public, privé ou hybride** et non sur des serveurs internes. On

distingue deux environnements de bases de données cloud distincts : les bases de données traditionnelles, et les bases de données en tant que service (DBaaS).
Une database cloud traditionnelle peut être lancée sur l'infrastructure informatique d'une entreprise **par le biais d'une machine virtuelle**. La maintenance et la gestion de la base de données sont assurées par le personnel informatique de l'entreprise.
Dans le cas d'une DBaaS, l'entreprise s'abonne à un service, et la base de données tourne sur l'infrastructure physique du fournisseur de services cloud. Le fournisseur se charge de la maintenance de l'infrastructure physique et de la base de données, tandis que le client est chargé de gérer le contenu et les opérations de la base de données. Certains fournisseurs proposent également de prendre en charge la gestion, une option très utile pour les petites entreprises qui ont besoin d'une base de données, mais n'ont pas l'expertise informatique.
Les bases de données cloud **apportent de nombreux avantages** par rapport à une base de données traditionnelle. Elles permettent tout d'abord d'éliminer les infrastructures physiques. L'infrastructure physique est en effet fournie à distance par le fournisseur de cloud qui se charge également de la maintenance et assure la disponibilité.
Le second avantage est économique puisque les bases de données cloud **permettent de réduire les coûts**. L'élimination de l'infrastructure physique permet de réduire les dépenses, et implique une réduction de l'effectif nécessaire et de la facture d'électricité. Le cloud permet aussi d'économiser de l'espace physique.
Un DBaaS présente encore plus d'avantages. L'utilisateur jouit d'une **scalabilité instantanée**, qui lui permet d'augmenter ou de réduire la capacité de sa base de données en fonction de ses besoins. Le fournisseur de services se charge rapidement d'augmenter la capacité allouée au client. Une base de données sur site prendrait plusieurs mois pour parvenir au même résultat, avec la nécessité d'installer de nouveaux serveurs et autres ressources de communication.
Par ailleurs, les **accords de niveaux de service** obligent le fournisseur à garantir les performances de la base de données, la disponibilité et le temps de réponse. Si ce niveau de service n'est pas respecté, le client est remboursé. L'utilisateur d'une DBaaS profite aussi de l'expertise du fournisseur de services en matière de base de données, chose qu'il serait très difficile de trouver auprès d'employés en raison des coûts élevés d'une telle qualification. Le fournisseur de DBaaS a des milliers de clients, et ce n'est donc pas un problème pour lui de recruter l'élite des administrateurs de bases de données.
Les fournisseurs de DBaaS s'assurent aussi que leurs logiciels de bases de données, leurs systèmes d'exploitation et toutes les technologies liées à leurs infrastructures soient à la pointe de la technologie. Ceci leur permet de rester compétitifs. Les fonctionnalités et la sécurité sont régulièrement mises à jour. Pour le client, c'est l'assurance de **profiter des dernières technologies en date**.

Pour éviter l'interruption des opérations en cas de panne, les fournisseurs de DBaaS utilisent **des images serveur miroir** et ont plusieurs complexes de stockage de données à disposition. En cas de panne, les opérations sont transférées vers un data center de backup et le client ne se rend compte de rien. Dernier avantage, grâce aux avancées technologiques et à la féroce concurrence sur le marché, le prix des services cloud diminue progressivement. Les DBaaS sont donc de plus en plus accessibles.

Tout comme les bases de données traditionnelles, les **bases de données cloud se divisent en deux catégories** : relationnelles et non relationnelles. Les différences entre ces deux catégories et leurs avantages respectifs, évoquées plus haut, sont valables pour les bases de données cloud.

Il est tout à fait possible de **migrer les bases de données sur site vers le cloud**. Cette démarche présente plusieurs bénéfices. Ceci permet de se débarrasser des serveurs physiques et des infrastructures de stockage, et évite à l'entreprise de devoir trouver des experts en bases

de données. La migration peut aussi améliorer l'efficience du traitement de données, notamment lorsque les outils analytiques et autres applications sont également sur le cloud.
Bien entendu, migrer une base de données vers le cloud permet aussi **d'économiser de l'argent en réduisant les coûts**. L'équipe informatique s'en trouve réduite, et donc moins d'employés doivent être payés. Par ailleurs, les prix du cloud diminuent de plus en plus ce qui permet de profiter d'offres toujours plus aguichantes. En outre, les fournisseurs de services cloud proposent à leurs clients de ne payer que pour les ressources qu'ils consomment. Il s'agit du modèle pay-as-you-go.
Une migration vers le cloud peut permettre par la suite d'adopter davantage de solutions SaaS (logiciel en tant que service). Ces services cloud permettent simplifier l'accès aux informations par le biais d'internet. La consolidation du stockage peut être un autre avantage d'une migration vers le cloud. Par exemple, les bases de données des différents départements d'une grande entreprise peuvent être combinées sur le cloud en un seul système de gestion de base de données.
En termes de fonctionnement structurel et de design, une base de données cloud est en tous points similaire à une base de données sur site. La principale différence est l'endroit où réside la base de données. Une base de données sur site est connectée aux différents utilisateurs locaux de l'entreprise **par le biais d'un réseau local de type LAN**. Une base de données cloud quant à elle réside sur les serveurs d'un fournisseur cloud ou DBaaS. Elle est accessible uniquement via internet.
Dans les deux cas, on accède à la base de données par le biais de requêtes directes ou d'appels API, et la base de données réagit de la même manière. Cependant, **le temps de réponse peut varier**. Une base de données sur site est logiquement un peu plus rapide qu'une base de données cloud, car chaque interaction sur une base cloud doit être transférée sur internet. La différence est toutefois infime.

SGBD (système de gestion de base de données) ou DBMS et RDBMS : les logiciels qui permettent d'accéder aux bases de données

Pour accéder aux bases de données, on utilise un SGBD (système de gestion de base de données), à savoir un **logiciel de type DBMS ou RDBMS**. Un DBMS est un database management system (système de gestion de base de données). Il permet de définir, manipuler, récupérer et gérer les données stockées au sein de la base de données. Le DBMS extrait des informations de la base de données sur demande de l'utilisateur, en fonction des requêtes effectuées. Pour effectuer une requête, l'utilisateur peut entrer un mot-clé, ou effectuer une commande de tri.

La puissance d'un DBMS est sa capacité à **définir de nouvelles relations** à partir de relations basiques données par les tableaux pour répondre aux requêtes. Généralement, l'utilisateur entre une série de caractères, et l'ordinateur cherche les séquences correspondantes pour fournir à l'utilisateur les matériaux sources dans lesquels ces caractères apparaissent. Par exemple, un utilisateur peut rechercher tous les enregistrements contenant un champ lié à une personne portant le nom de famille Dupont.

Le terme **Database est de plus en plus utilisé comme abréviation** pour Database Management System. Il existe beaucoup de DBMS différents. Certains sont des petits systèmes pouvant être lancés sur un ordinateur personnel, d'autres sont d'énormes systèmes nécessitant un mainframe.

Les **DBMS ont été inventés dans les années 1960** pour prendre en charge les bases de données hiérarchiques. Les premiers systèmes étaient organisés de façon séquentielle (alphabétiquement,, numériquement, ou chronologiquement). Il a fallu attendre l'apparition des appareils de stockage à accès direct pour pouvoir accéder aux données aléatoirement par le biais d'index. Parmi les DBMS les plus connus, on compte le IBM Information Management System et le CA Integrated Database Management System.

Un RDBMS est un relational database management system (système de gestion de base de données relationnel). Ce type de logiciel a été **développé dans les années 70**en se basant sur

le modèle relationnel. Encore aujourd'hui, il demeure la façon la plus populaire de gérer une base de données. Les RDBMS les plus connus sont Microsoft SQL Server, Oracle Database, IBM DB2 et MySQL.

Qu'est-ce que le SQL ou Structured Query Language ?

Le SQL (Structured Query Language) est un **langage de programmation standardisé**utilisé pour gérer les bases de données relationnelles et effectuer différentes opérations sur les données qu'elles contiennent. Ce langage a été créé dans les années 70, et continue d'être utilisé régulièrement par les administrateurs de bases de données. Les développeurs qui rédgient des scripts d'intégration de données et les analystes de données l'utilisent aussi pour lancer des requêtes analytiques.
Un standard SQL officiel a été adopté par l'ANSI (Americain National Standards Institute) en 1986. En 1987, l'ISO (International Organization for Standardization) l'a adopté à son tour. Ce standard a été mis à jour plus de six fois depuis lors. La **version la plus récente est le SQL:2011**.
L'utilisation du SQL permet d**e modifier les structures des tableaux et des index des bases de données**. L'utilisateur peut ajouter, mettre à jour et supprimer des rangées de données et retrouver des sous-ensembles d'informations. Ces informations peuvnet ensuite être utilisées pour les applications analytiques ou les traitements de transactions. Une r**equête SQL prend la forme d'une commande écrite**. Les commandes les plus fréquemment utilisées sont la sélection, l'ajout, l'insertion, la mise à jour, la suppression, la création, l'altération et le tronquement.

Les **commandes SQL sont divisées en plusieurs types différents**. Le vocabulaire DML (langage de manipulation des données) permet de retrouver et de manipuler des données. Le DDL (langage de définition des données) est utilisé pour définir et modifier les structures des bases de données. Le contrôle de transactions permet de gérer les transactions pour s'assurer qu'elles soeint complétées ou anulées en cas de problème ou d'erreur. Enfin, les security statements sont utilisés pour contrôler l'accès aux bases de données et créer un système de permissions pour les différents utilisateurs.

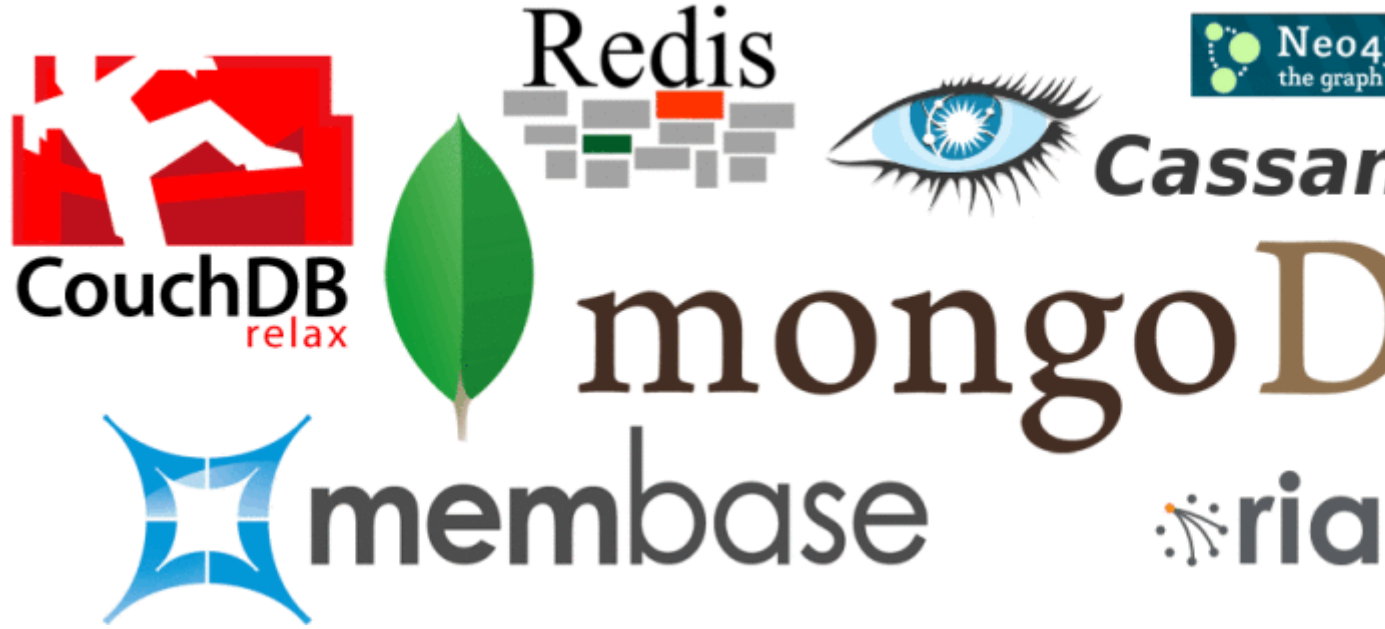
Les entreprises utilisent des RDBS ou systèmes de gestion de base de données développés autour de SQL propriétaires ou open source. Parmi les plus connus, on compte Microsoft SQL Server, Oracle Database, IBM DB2, SAP HANA, SAP Adaptive Server, Oracle MySQL et PostgreSQL. Malgré la standardisation du SQL, la plupart des vendeurs utilisent des **extensions propriétaires** pour la programmation procédurale et les autres fonctions. Par exemple, Microsoft propose un set d'extensions Transact-SQL (T-SQL) et Oracle a sa version étendue (PL/SQL). De fait, les variantes des divers vendeurs ne sont pas compatibles entre elles.

Depuis peu, les moteurs de requêtes SQL-on-Hadoop sont utilisés pour les architectures Big Data développées autour des systèmes Hadoop. Ces moteurs permettent d'exploiter le Big Data sans avoir à utiliser de langage plus complexe. On peut citer l'environnement de programmation MapReduce pour le développement d'applications de traitement batch. Au total, **on dénombre plus d'une douzaine d'outils SQL-on-Hadoop** proposés par les différents vendeurs de distributions Hadoop et autres vendeurs. Notons que le moteur de traitement Apache Spark intègre son propre module Spark SQL compatible avec la programmation SQL.

Qu'est-ce qu'un administrateur de base de données (DataBase Administrator) ?

Un Administrateur de base de données ou DataBase Administrator est la **personne chargée de maintenir un environnement de base de données**. La conception, l'implémentation, la maintenance du système et la mise en place de règles. Il doit aussi former les employés de l'entreprise à la gestion et à l'utilisation de la base de données.
En règle générale, un **DBA dispose d'une formation dans le domaine des sciences informatiques** et d'une expérience professionnelle avec une base de données spécifique ou diverses bases de données. Il doit aussi avoir une expérience avec les principaux produits de gestion de base de données comme SQL, SAP ou Oracle.

Quelles sont les meilleures bases de données à l'heure actuelle ?



Le choix d'une bonne base de données est très important pour votre entreprise. Une base de données facile à utiliser, sécurisée, avec un bon suivi de développement peut permettre d'augmenter la productivité. Il est donc indispensable de bien étudier **les avantages et les inconvénients de chaque base de données**. La base de données que vous choisirez doit s'adapter à l'écosystème de votre entreprise.

Plusieurs **questions essentielles doivent être posées** : quelle quantité de données devrez-vous gérer, quel est le temps de réaction attendu par vos clients, combien de clients ave-vous, comment la base de données s'adaptera si votre nombre de clients et de transactions augmente, comment surveillerez-vous la base de données pour éviter les down times, avez-vous besoin d'une base de données relationnelle ou NoSQL, et comment se comportera la base de données en cas de plantage ou de problème ?

Actuellement, le **marché est dominé par DB2, SQL Server, Oracle et IBM**. Sur Windows, SQL Server est généralement la base de données de prédilection, tandis que Oracle et DB2 sont les plus populaires sur les écosystèmes Mainframe/Unix ou Linux. Pour vous aider celle qui vous convient, voici notre sélection des meilleures bases de données.

Microsoft SQL Server, la base de données Windows

Développée par Microsoft, SQL Server, elle est **compatible exclusivement avec Windows**. Cette database est simple à maîtriser, et de nombreuses personnes y sont formées. L'intégration avec Microsoft Azure a permis d'augmenter sa flexibilité et ses performances. De plus, le cloud permet désormais d'administrer des informations en provenance d'autres serveurs, ce qui la rend plus utile.

Oracle, la base de données la plus populaire pour Linux/Unix

La base de données Oracle peut fonctionner sur presque tous les systèmes. Elle est très populaire, et de **nombreuses personnes sont formées à la maîtriser**. Par ailleurs, elle propose de nombreux outils orientés vers le monitoring et l'administration.

IBM DB2, la base de données Mainframe la plus populaire

Après Oracle, IBM DB2 est la deuxième base de données la plus utilisée sur les écosystèmes Unix/Linux. Pour **Mainframe, il s'agit du choix le plus populaire**. Là encore, de nombreuses personnes sont formées à l'utiliser, même si elle compte moins d'adeptes qu'Oracle.

Teradata, la meilleure base de données pour le Big Data

Teradata est spécialement conçue pour le Big Data. De fait ses capacités de stockage et d'analyse de données sont colossales. Pour une stratégie Big Data, il s'agit sans conteste de la meilleure option à votre disposition.

SAP Sybase, l'ancien leader du marché

Par le passé, cette database était très populaire et dominait largement le marché. Aujourd'hui, elle n'est plus aussi utilisée, mais **reste une solution très pertinente** en termes de scalabilité et de performances.

Informix, une base de données rachetée par IBM

Tout comme SAP Sybase, **Informix a perdu de sa superbe** à la fin des années 90. Suite à une série de mauvaises décisions en termes de management, elle a fini par être rachetée par IBM.
Cette base de données n'existe plus aujourd'hui sous sa forme initiale. Malgré tout,**ses fondations restent utilisées par certains outils et applications IBM**.

MySQL : la meilleure base de données relationnelle Open Source

Dans la catégorie bien particulière des bases de données relationnelles Open Source, on compte trois références principales : **MySQL, MariaDB, et PostgreSQL**. Ces trois systèmes de gestion de bases de données ont plusieurs points communs : une communauté de support très active, un code ouvert permettant aux utilisateurs de les modifier à leur guise en fonction de leurs besoins, et une gratuité (pour les développeurs).

MySQL est un système de gestion de base de données relationnelle (RDBMS) basé sur le SQL (Structured Query Language). Ce SGBDR est compatible avec toutes les plateformes : **Linux, UNIX, et Windows**. Il peut être utilisé avec de nombreuses applications, mais on l'associe le plus souvent aux applications web.

C'est à l'origine l'entreprise suédoise MySQL AB qui a créé MySQL. Toutefois, **en 2008, Sun Microsystems l'a racheté** avant d'être lui-même racheté par Oracle en 2010. Les développeurs peuvent toujours utiliser MySQL sous la licence GPL, mais les entreprises quant à elle doivent acheter une licence commerciale à Oracle.

MySQL est également un **composant de la suite logicielle d'entreprise LAMP**, une plateforme de développement web exploitant Linux comme système d'exploitation, Apache comme serveur Web, MySQL comme SGBDR et le PHP comme langage de script orienté objet.

PHP est en effet le langage le plus populaire pour le développement web. Il s'agit d'un langage gratuit, open source et server-side, ce qui signifie que le code est exécuté sur le serveur. La **combinaison entre le PHP et MySQL** permet de créer n'importe quel type de site, du simple formulaire de contact ou portail d'entreprise.

Le PHP permet aussi de **connecter un script web à une base de données MySQL**. Il s'agit d'une étape essentielle, indispensable pour pouvoir effectuer des requêtes sur une base de données.

Logiciel de base de données gratuit

Il existe différents logiciels de gestion de base de données gratuits. Certains sont tout simplement des doubles gratuits du fameux Excel de Microsoft (Open Office Base), d'autres sont plus complexes à utiliser et prennent le rôle de SGBD. C'est le cas de SQLite, un moteur de database, MariaDB, MySQL, MaxDB, Ingres ou encore PostgreSQL.

Quelles sont les meilleures bases de données NoSQL ?

Parmi les bases de données NoSQL, on dénombre de nombreuses sous-catégories. Chacune de ces catégories se distingue par des caractéristiques spécifiques. Voici les **quatre principaux groupes de bases de données NoSQL**, et les meilleures de chacune de ces catégories.

Les bases de données NoSQL orientées clé-valeur

Ces bases de données sont idéales pour **accéder aux données par le biais d'une clé**. La spécificité est que les données peuvent être stockées sans définir de schéma spécifique. Ces bases de données sont très efficaces pour la lecture et l'écriture, et conçues pour s'adapter massivement et proposer un temps de réaction extrêmement rapide.
Les données sont généralement stockées au sein de structures complexes comme le BLOB. Les références dans cette catégorie de bases de données sont **Redis, Riak, Oracle NoSQL et Microsoft Azure Table Storage**. Radis est une base de données gratuite et open source, Riak est entièrement dédiées aux clés-valeur et idéale pour le stockage de document et les fonctionnalités de recherche.

Les bases de données NoSQL orientées document

Cette catégorie de bases de données **repose sur divers formats (JSON, XML)** et offre la capacité de changer de schéma sans avoir à arrêter la database. Les développeurs peuvent télécharger des documents indexés et y accéder par le biais de moteur de stockage de la base de données. La flexibilité de ces bases de données les rend très polyvalentes.
Parmi les meilleures bases de données NoSQL orientées document, on peut citer **Mongo DB et Couchbase Serve**r. Mongo DB est l'une des bases de données les plus populaires à l'heure actuelle, toutes catégories confondues. Elle permet de prendre en charge aussi bien des données structurées que des données non structurées, et ses performances et sa scalabilité sont excellentes. De nombreuses personnes sont formées pour la maîtriser.
Couchbase Server est une base de données Open Source licenciée sous Apache. Son avantage principal est sa **console d'administration très intuitive** permettant d'accéder facilement à de grandes quantités de données. En revanche, elle ne permet pas de garantir l'intégralité des données à 100%.

Parmi les meilleures bases de données NoSQL orientées document, **on peut aussi citer Mark Logic Server**. Son intégrité des données et sa compatibilité XML, JSON et RDF en font une référence. Mark Logic Server est compatible avec les OS Windows, Solaris, Red Hat, Suse, CentOS, Amazon Linux et Mac OS. On peut enfin citer Elastic Sarche, RavendDB, Apache Jena et Pivotal GemFire.

Les bases de données NoSQL orientées colonne

Les bases de données NoSQL orientées colonne r**eprésentent la valeur des données sous forme de colonne**, ce qui permet à l'utilisateur de cartographier les clés et les valeurs et de les regrouper en structures. Ces bases de données sont principalement utilisées dans les environnements où il est nécessaire d'accéder à de nombreuses colonnes elles-mêmes dotées de nombreuses rangées. Elles sont surtout utiles pour le traitement et l'analyse d'événements, la gestion de contenu et l'analyse de données.

Parmi les meilleures bases de données NoSQL orientées colonne, on peut citer **Apache Cassandra, une base de données créée par Facebook** et désormais distribuée gratuitement. Cassandra est recommandé pour les basées de données regroupant d'immenses quantités de données.

Précisons qu'une **version entreprise intitulée Datastax Enterprise** est également disponible. Cassandra est compatible ASCII, bigint, BLOB, Boolean, counter, decimal, double, float, int, text, timestamp, UUID, VARCHAR et varient. L'autre référence dans cette catégorie est Apache Hbase, conçue pour prendre en charge de multiples accès lecture et écriture en temps réel à d'immenses quantités de données.

Les bases de données NoSQL orientées graphe

Les bases de données NoSQL orientées graphe sont **focalisées sur les propriétés et les relations** qui les unissent. Elles utilisent la théorie des graphes pour connecter les bases de données. Chaque élément est connecté à son élément adjacent. Ces bases de données sont recommandées si vos données sont interrelationnelles, comme sur les réseaux sociaux, dans la détection de fraude ou encore les mises à jour en temps réel.