



كلية العلوم ببنزرت
FACULTÉ DES SCIENCES DE BIZERTE



Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique
Université de Carthage

Administration des Bases de Données

Architecture générique d'un SGBD

Plan

- ▶ Familles de SGBD et cas d'usage
- ▶ Administration vs Programmation
- ▶ Missions de l'administration

Familles de SGBD

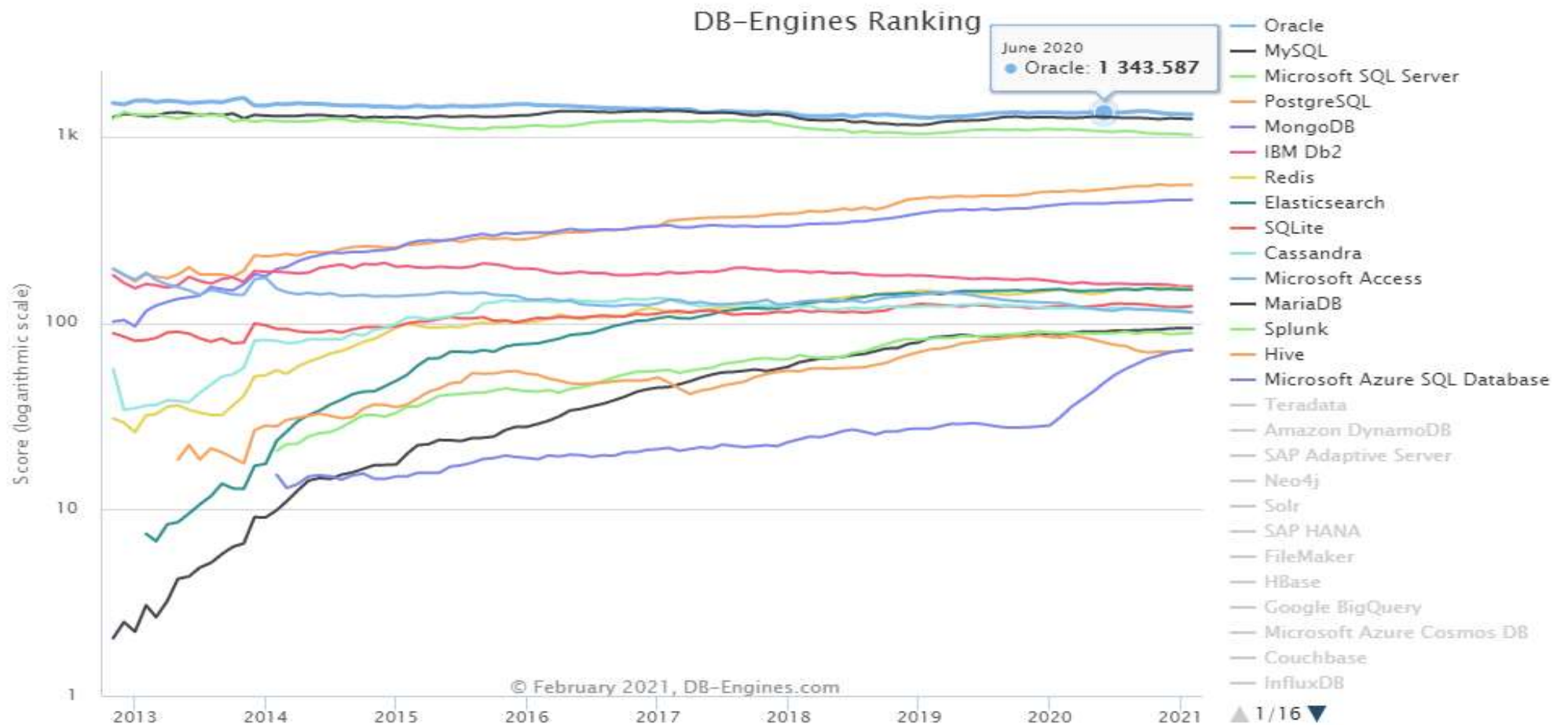
- ▶ Relationnel : Oracle / MySQL / SQLServer / PostgreSQL
 - ▶ Moteur/structure de stockage basés sur le modèle « entité / relation »
 - ▶ Respecte les Formes Normales
 - ▶ Données structurées « à priori »
- ▶ Analytique : Oracle / SQLServer / Teradata
 - ▶ Moteur/structure de stockage basés sur le modèle « fait / dimension »
 - ▶ Ne respecte pas les formes normales (mais issue de données qui les respectent)
 - ▶ Données structurées « à priori »
- ▶ Non Relationnel
 - ▶ Moteur/structure de stockage de données non structurées (à priori)
 - ▶ Documents (JSON) : MongoDB (Oracle / MySQL / SQLServer / PostgreSQL)
 - ▶ Clé/Valeur (token) : Redis
 - ▶ Multicolonnes (tableau) : Cassandra
 - ▶ Recherche/indexation : Elasticsearch / Splunk / Solr
 - ▶ Graphes (vectoriel) : Neo4j (Oracle / SQLServer)

Popularité des SGBD (<https://db-engines.com/>)

364 systems in ranking, February 2021

Rank			DBMS	Database Model	Score		
Feb 2021	Jan 2021	Feb 2020			Feb 2021	Jan 2021	Feb 2020
1.	1.	1.	Oracle +	Relational, Multi-model i	1316.67	-6.26	-28.08
2.	2.	2.	MySQL +	Relational, Multi-model i	1243.37	-8.69	-24.28
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model i	1022.93	-8.30	-70.81
4.	4.	4.	PostgreSQL +	Relational, Multi-model i	550.96	-1.27	+44.02
5.	5.	5.	MongoDB +	Document, Multi-model i	458.95	+1.73	+25.62
6.	6.	6.	IBM Db2 +	Relational, Multi-model i	157.61	+0.44	-7.94
7.	7.	↑ 8.	Redis +	Key-value, Multi-model i	152.57	-2.44	+1.15
8.	8.	↓ 7.	Elasticsearch +	Search engine, Multi-model i	151.00	-0.25	-1.16
9.	9.	↑ 10.	SQLite +	Relational	123.17	+1.28	-0.19
10.	10.	↑ 11.	Cassandra +	Wide column	114.62	-3.46	-5.74
11.	11.	↓ 9.	Microsoft Access	Relational	114.17	-1.16	-13.89
12.	12.	↑ 13.	MariaDB +	Relational, Multi-model i	93.89	+0.10	+6.56
13.	13.	↓ 12.	Splunk	Search engine	88.54	+0.88	-0.23
14.	↑ 16.	14.	Hive	Relational	72.32	+1.89	-11.21
15.	15.	↑ 25.	Microsoft Azure SQL Database	Relational, Multi-model i	71.29	-0.07	+39.88
16.	↓ 14.	↓ 15.	Teradata +	Relational, Multi-model i	70.90	-1.69	-5.91
17.	17.	↓ 16.	Amazon DynamoDB +	Multi-model i	69.14	+0.01	+7.01
18.	18.	↑ 21.	SAP Adaptive Server	Relational	52.25	-2.36	-0.48
19.	19.	↑ 22.	Neo4j +	Graph	52.16	-1.62	+0.96

Popularité des SGBD (<https://db-engines.com/>)



Administration vs Programmation

- ▶ Composants BD :
 - ▶ Modèles de données, données, applications connectées, utilisateurs, requêtes, indexes
- ▶ Composants SG :
 - ▶ Moteur, réseau, CPU, RAM, stockage, OS, progiciel SGBD, ...
- ▶ Les composants DB sont liés à la famille de SGBD et relativement indépendants des éditeurs de SG.
- ▶ L'administrateur gère les impacts sur les composants de l'application (DB) des choix des systèmes de gestion des BD.

Sécurité des bases de données

- ▶ Définie selon un contrat de niveau de service (SLA) sur
 - ▶ (D)isponibilité
 - ▶ Mesurée selon l'accessibilité et la performance
 - ▶ (I)ntégrité
 - ▶ Mesurée selon la qualité des données stockées
 - ▶ (C)onfidentialité
 - ▶ Mesurée selon le respect du moindre privilège et du droit d'en savoir
 - ▶ (T)raçabilité
 - ▶ Mesurée selon la capacité à identifier précisément « qui à fait quoi où quand »

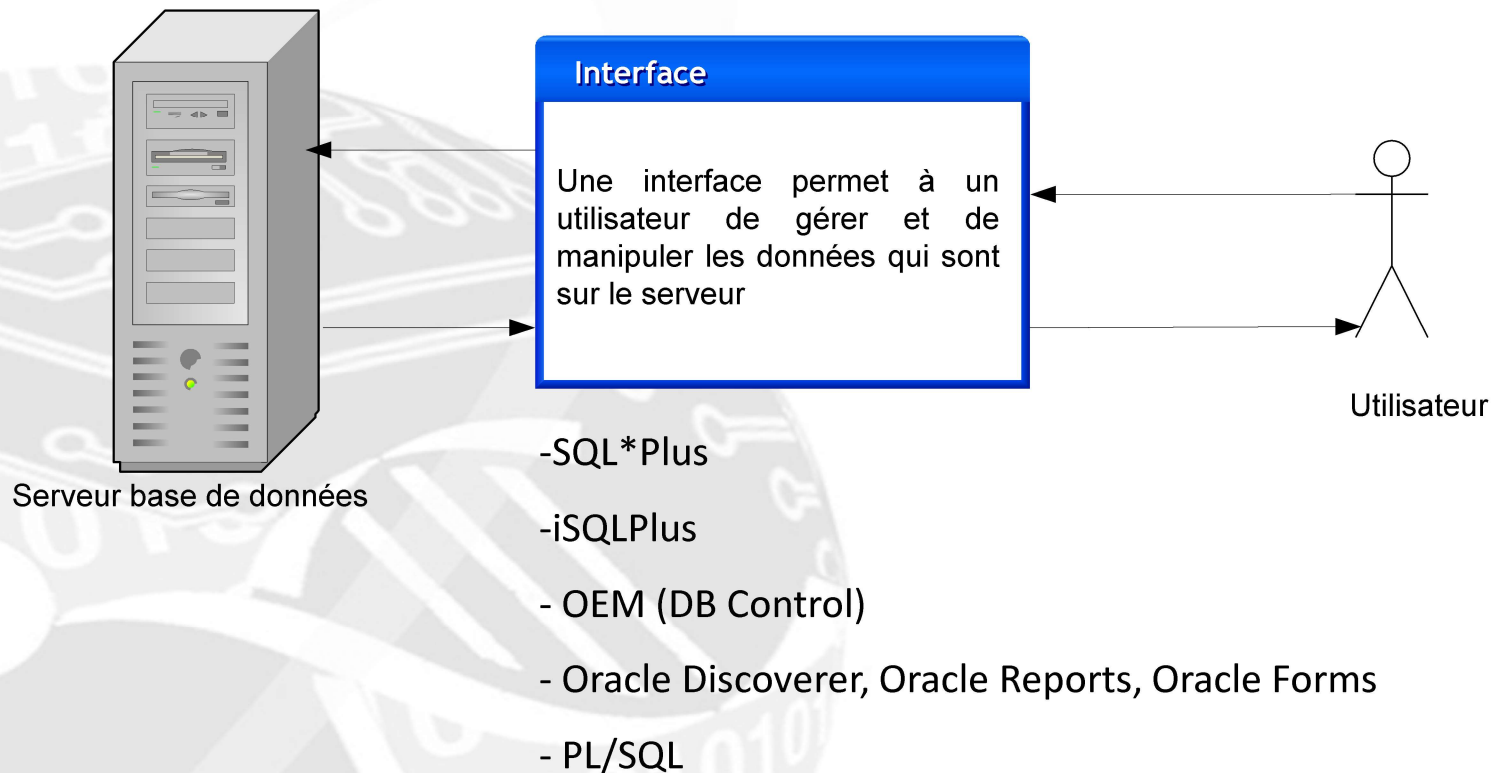
Introduction

► Oracle: Vue d'ensemble

- Oracle est le leader du marché des SGBDR, avec une part de marché allant jusqu'à 48.8% en 2011 (Gartner Group).
- Oracle commercialise différents produits :
 - Oracle Database,
 - Oracle Developer Suite,
 - Oracle Application Server,
 - Oracle Applications,
 - Oracle Collaboration Suite,
 - Oracle Services,
 - ...

Introduction

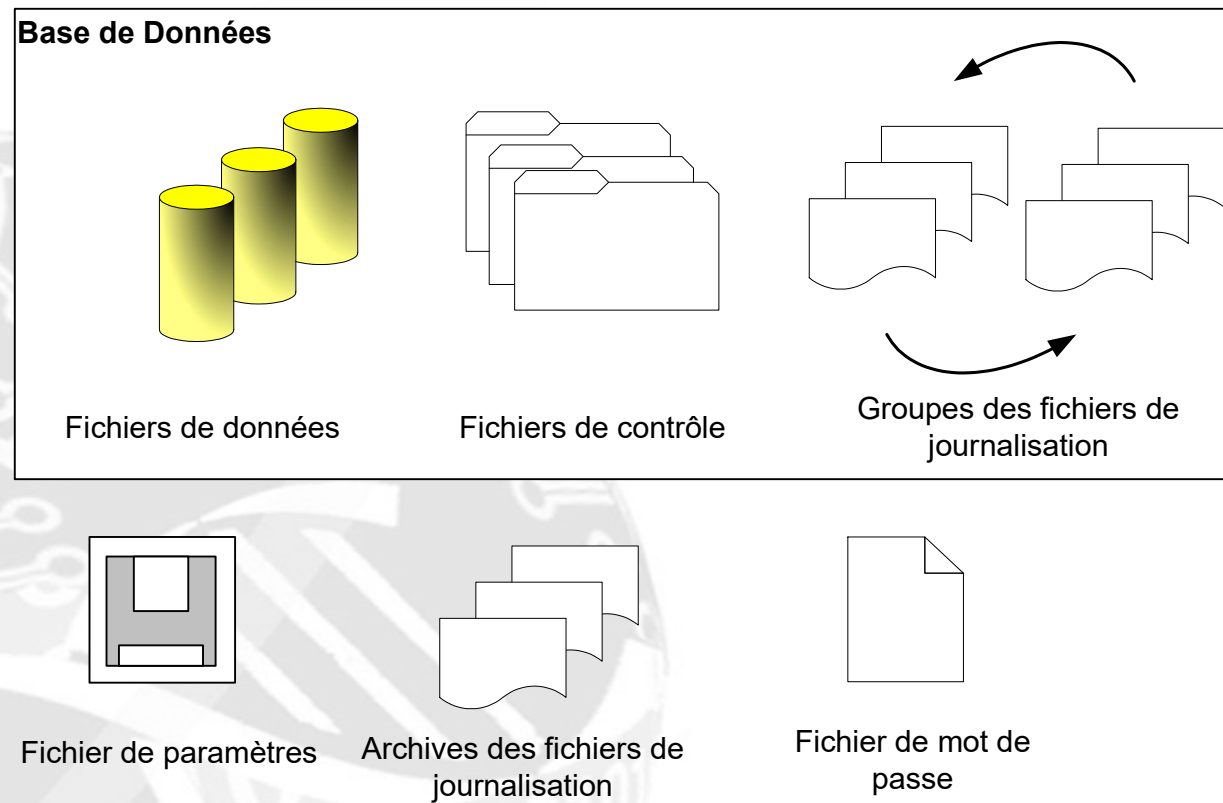
Interaction avec Oracle Database



Architecture d'Oracle Database

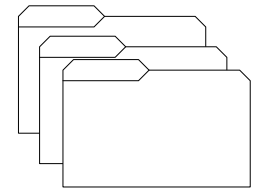
- ▶ Composantes principales d'un serveur Oracle
 1. Base de données : Ensemble de fichiers (sur disque) contenant les données, les informations sur les données, le journal de modifications
 2. Instance : Ensemble de processus et de zones mémoires (mémoire centrale) qui permettent de gérer la base de données.

Base de données



Fichier de contrôle

► Fichier de contrôle (Control File)



Fichier de contrôle

- Contient des informations de contrôle sur la base:
 - Nom de la base, noms et chemins des fichiers de données et de journalisation, informations de restaurations etc...
 - Un des premiers fichiers lu par l'instance lors du démarrage.
 - La vue **V\$CONTROLFILE** nous renseigne sur le contenu du fichier de contrôle.

Fichier de données

► Fichier de données (Data Files)



- Stockent les données sous un format spécial à Oracle.
- Physiquement, un fichier de données est un ensemble de blocs SE. Un bloc SE constitue l'unité d'E/S (écriture/lecture) des fichiers.
- Les fichiers de données sont logiquement regroupés sous forme de **tablespaces**.
- Une BD Oracle 10g inclut au moins deux tablespaces, **SYSTEM** et **SYSAUX**.

Les tablespaces

- ▶ Un tablespace est une unité logique qui correspond physiquement à un ou plusieurs fichiers de données.
- ▶ L'administrateur (DBA) agit sur les tablespaces et non sur les fichiers de données.
- ▶ Une BD est organisée sous forme de plusieurs tablespaces, chacun correspondant à un contexte (thème).
 - ▶ EX: On peut créer plusieurs tablespaces dans une BD d'une ESE commerciale qui gère la FACTURATION, la GRH, et le PARC INFO.

Les tablespaces

- ▶ Le tablespace FACTURATION : fact01.dbf et fact02.dbf
- ▶ Les tables FACTURE, LIGNE_FACTURE, PRODUIT, CLIENT, REGLEMENT.
- ▶ Le tablespace GRH : grh01.dbf
- ▶ Les tables PERSONNEL, PAIE, SANCTION etc.
- ▶ Le tablespace PARC : parc01.dbf
- ▶ Les tables EQUIPEMENT, CATEGORIE, REPARATION, MAINTENANCE etc.
- ▶ Avantage: On peut administrer par tablespace, et donc par partie (par rapport à la BD).
- ▶ Par exemple, pour maintenir certaines tables relatives à la facturation, on peut mettre uniquement le tablespace FACTURATION en offline, au lieu de rendre toute la base indisponible ce qui touchera des centaines d'utilisateurs qui ont besoin de manipuler les données de la GRH, PARC INFO etc.

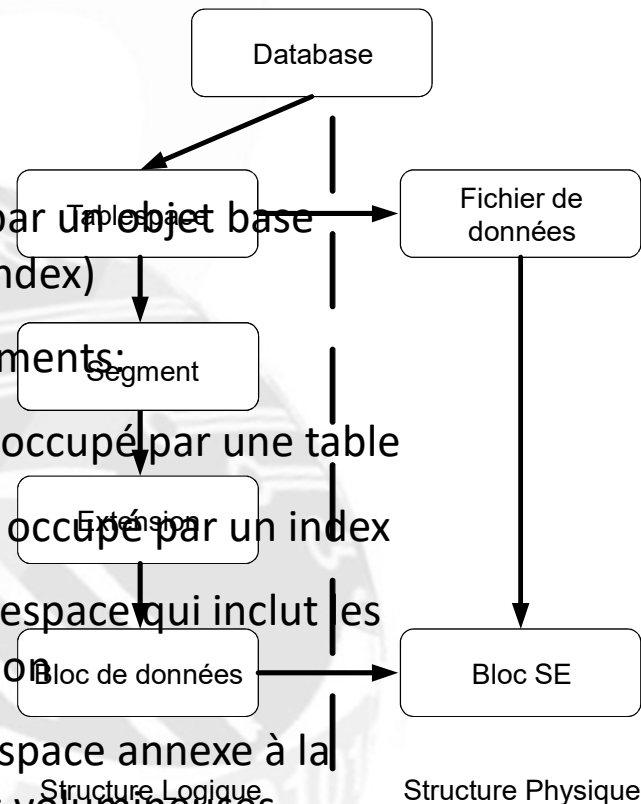
Les tablespaces

- ▶ Les vues **DBA_TABLESPACES** et **DBA_DATA_FILES** incluent toutes les informations relatives aux tablespaces et aux fichiers de données de la base.
- ▶ Pour afficher les noms des fichiers de données ainsi que les tablespaces auxquelles ils correspondent:

```
SELECT tablespace_name, file_name  
FROM DBA_DATA_FILES  
ORDER BY tablespace_name;
```

Structure de stockage

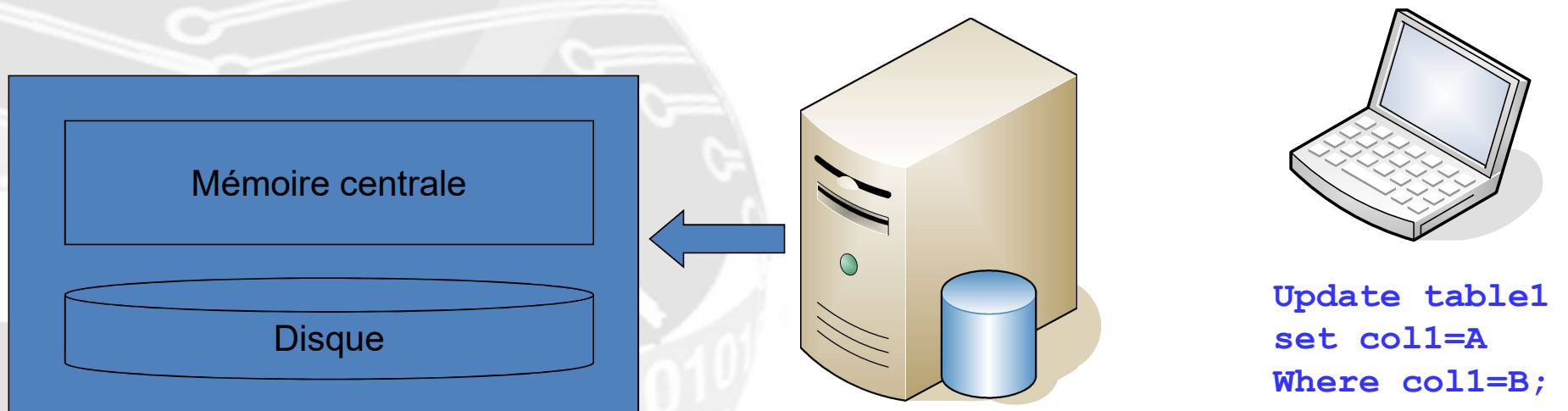
- C'est l'espace occupé par un objet base de données (Table ou Index)
- Il existe 4 types de segments:
 1. Segment table: espace occupé par une table
 2. Segment index: espace occupé par un index
 3. Segment d'annulation: espace qui inclut les informations d'annulation
 4. Segment temporaire: espace annexe à la MC pour les opérations volumineuses.



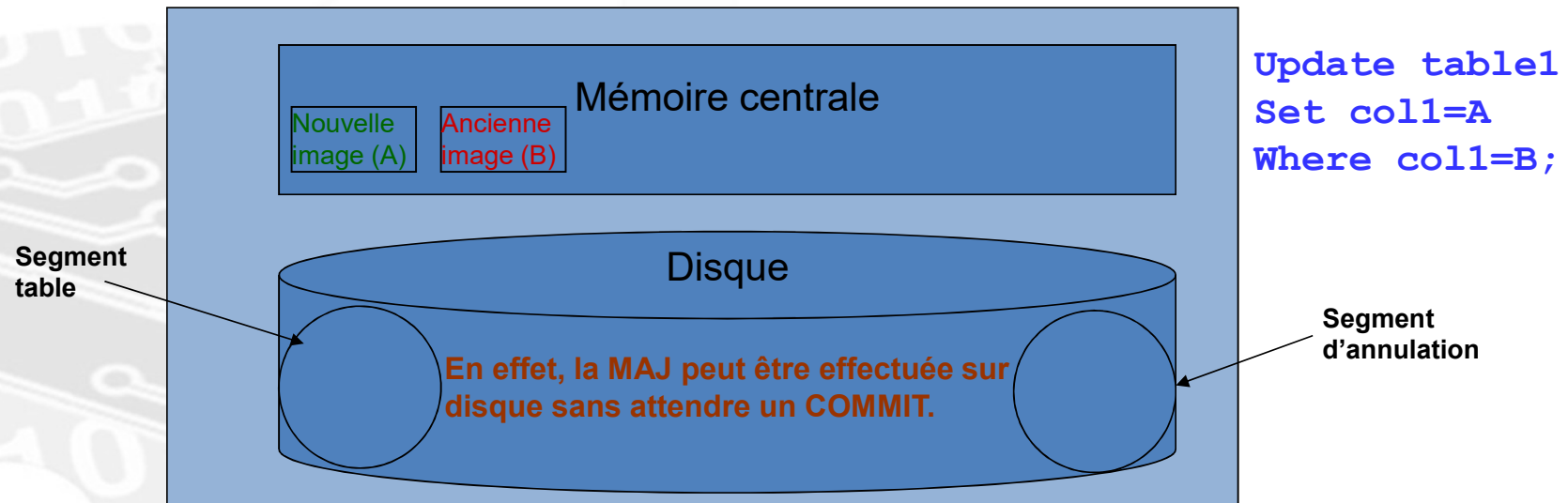
Le segment d'annulation

Le segment d'annulation sert à stocker les données nécessaires:

1. A l'annulation



Le segment d'annulation

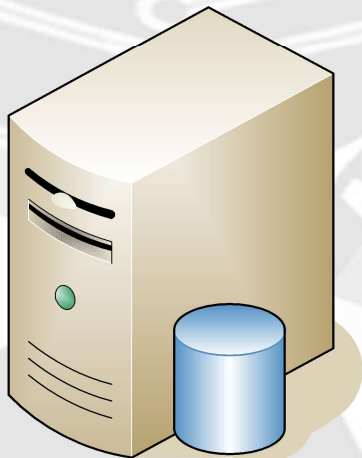


La nouvelle image est enregistrée dans un segment table, l'ancienne dans un segment d'annulation.
Au cas où un ROLLBACK est effectué, c'est cette image qu'on utilisera pour rétablir les données.

Le segment d'annulation

❑ **Le segment d'annulation** sert à stocker les données nécessaires:

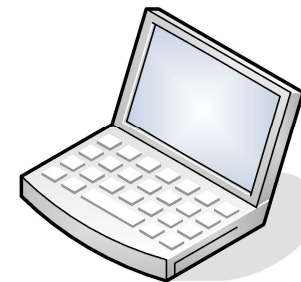
1. A l'annulation
2. A la lecture cohérente (ou dite aussi consistante).



Andiez la valeur
ou ancienne?

12:01

Select col1 From table1
Where col2=C;



11:58
Update table1
set col1=A
Where col2=C;



Le segment d'annulation

❑ **Le segment d'annulation** sert à stocker les données nécessaires:

1. A l'annulation
2. A la lecture cohérente (ou dite aussi consistante).
 - a. Si **pas d'anticipation d'enregistrement des MAJ**, Alors la lecture se fait directement à partir du segment table.
 - b. Si **anticipation d'enregistrement des MAJ**, Alors la lecture se fait à partir du segment d'annulation.

Le segment temporaire

- ▶ Le segment temporaire sert à stocker les données relatives à des opérations volumineuses si la mémoire centrale ne suffit pas à les exécuter.
 - ▶ Exemple d'opérations volumineuses:
 - ▶ Certains tris.
 - ▶ Certaines jointures.
 - ▶ Création d'index.
 - ▶ Etc.

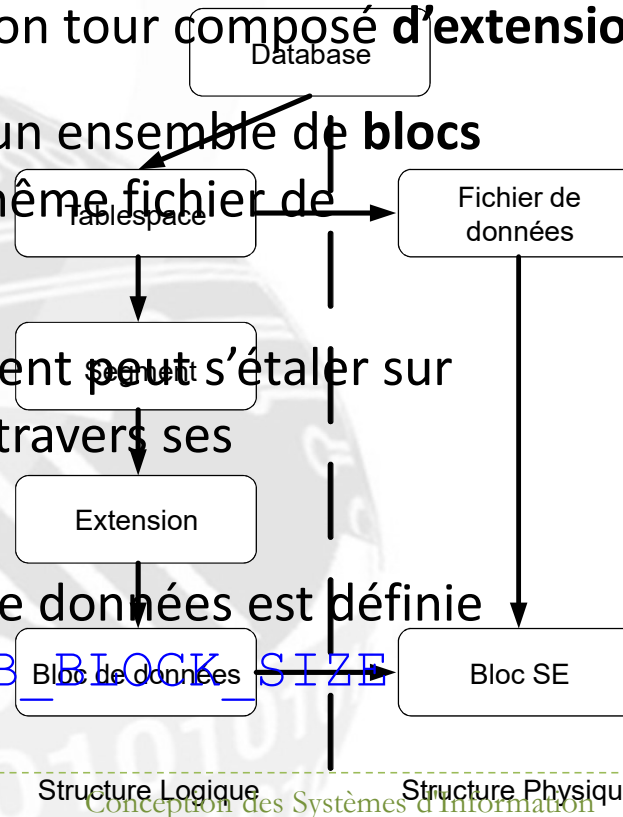
Extension

- ❑ Un **segment** est à son tour composé d'**extensions**

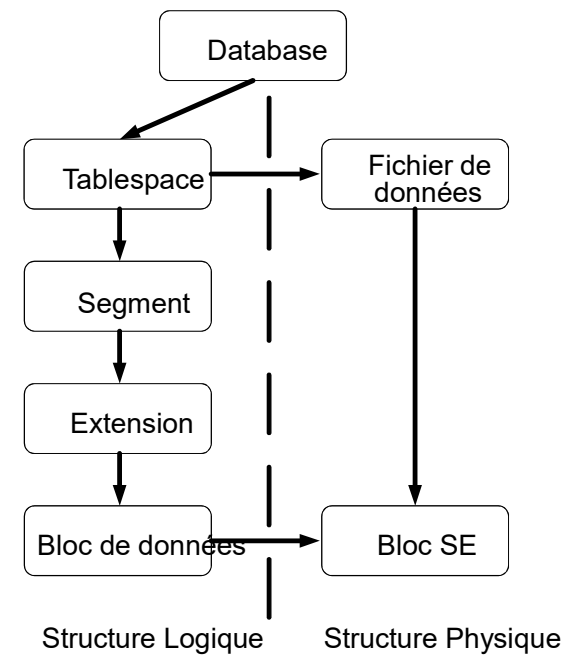
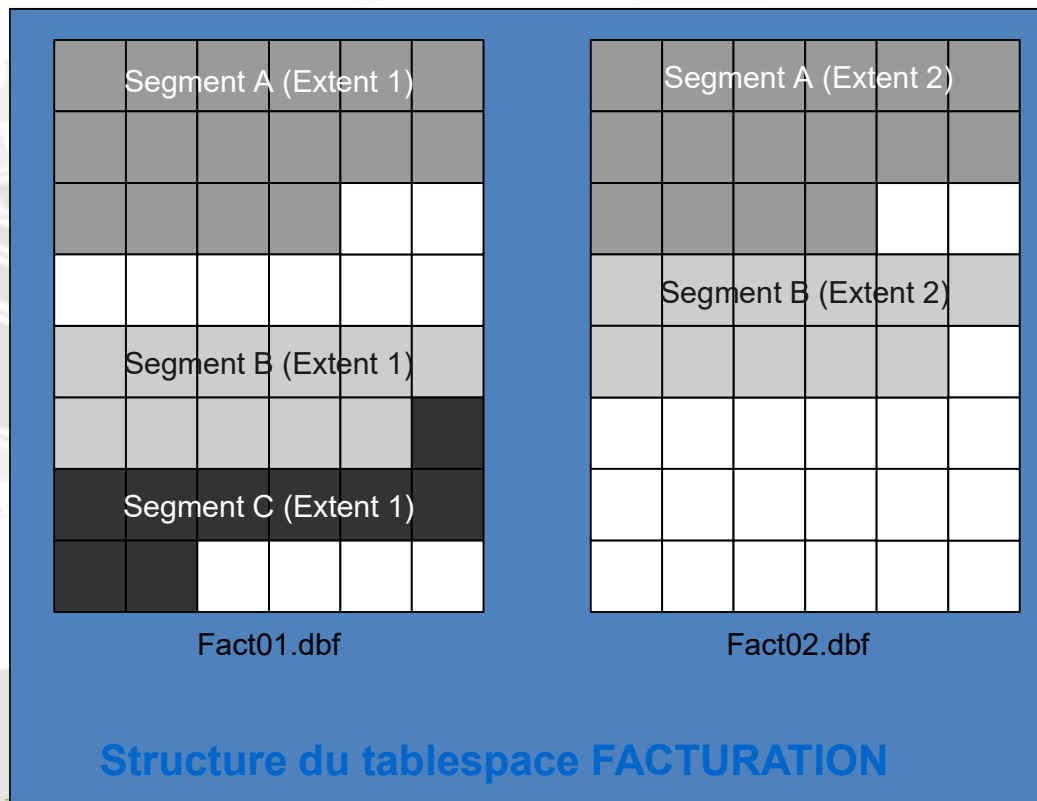
- ❑ Une **extension** est un ensemble de **blocs** appartenant à un même **fichier de données**

- ❑ Par contre un segment peut s'étaler sur plusieurs fichiers à travers ses extensions.

- ❑ La taille d'un bloc de données est définie par le paramètre **DB_BLOCK_SIZE**



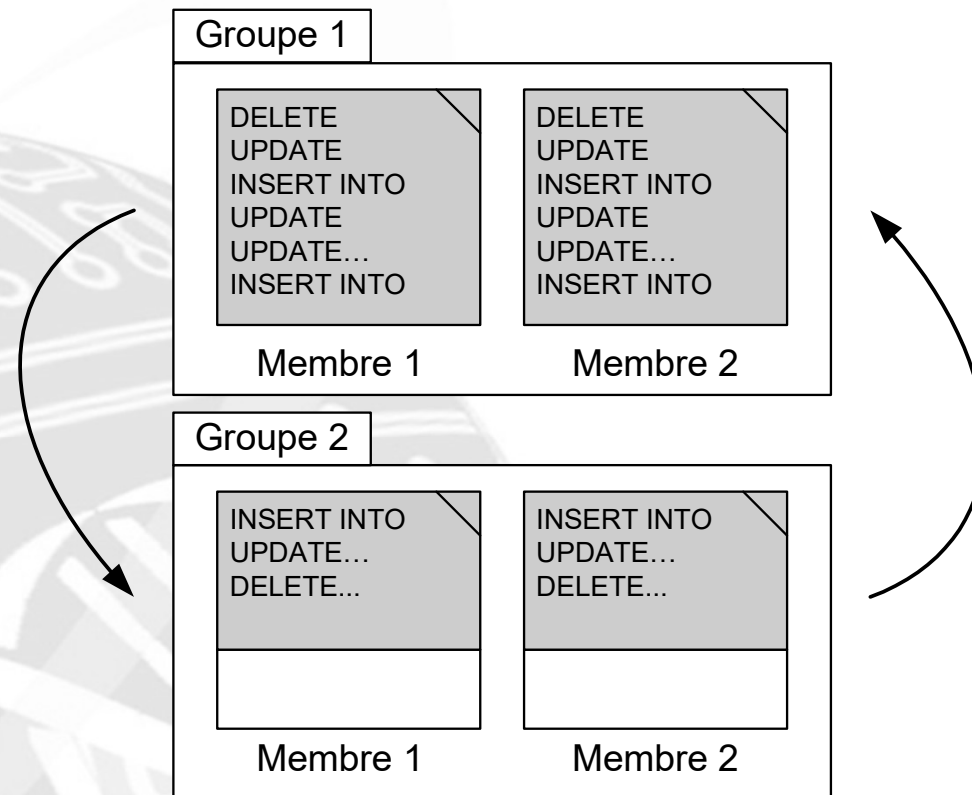
Exemple



Fichiers de journalisation

- ▶ Les fichiers de journalisation contiennent toutes les modifications effectuées sur les données depuis une certaine durée.
- ▶ En cas de crash du système, ou d'altération des fichiers de données, on peut reconstituer les données à partir des fichiers journaux.
- ▶ L'écriture sur les fichiers journaux est multiplexée et cyclique.
- ▶ L'ensemble des fichiers multiplexés (qui contiennent donc les mêmes informations) sont appelés membres et forment un groupe.
- ▶ L'écriture est donc multiplexée à l'intérieur d'un groupe, et cyclique entre les groupes.
- ▶ La vue **V\$LOGFILE** contient les informations des fichiers journaux.

Fichiers de journalisation



Archives des fichiers journaux

- ▶ L'écriture des fichiers journaux est cyclique, ce qui fait qu'à un certain moment, certaines transactions seront écrasées.
- ▶ La solution, c'est d'archiver les fichiers journaux avant de les écraser.
- ▶ Les fichiers journaux archivés peuvent être stockés sur des disques (serveurs) distants, ce qui optimisera la sécurité de la BD.
- ▶ Une base de données n'est pas forcément en mode ARCHIVELOG, si elle ne l'est pas, les fichiers journaux ne sont pas archivés.