

Bases de données multimédias  
Bases de données multidimensionnelles

# Contenu

- BD Multimédia :
  - Caractéristiques
  - Modélisation
  - Interrogation
  - Architectures des SGBD multimédias
- BD Multidimensionnelles
  - Motivations
  - Modèle du cube
  - Architectures

# Volume des données

- Quelques chiffres...
  - 100 lignes de fax : 6,4 MO
  - 100 images couleur 500 MO
  - 1 heure de vidéo : 1 GO
- Stockage
  - Compression indispensable, mais cela a un coût (temps)
  - Hiérarchies de mémoires, plusieurs supports (complique le fonctionnement du SGBD)
- Efficacité de la recherche
  - Index adaptés

# Types de données

- Texte
  - Les modèles classiques ne permettent pas le traitement efficace de documents fortement structurés
- Image
  - Nombreuses applications
  - Description externe + description du contenu
  - Recherches sur le contenu
  - Différents codages, différentes normes
  - Visualisation des résultats
- Sons
  - Pbs: recherche et traitement
- Vidéo
  - Séquences d'images fixes ?
  - Interrogation des vidéos (séquences où Belmondo se bat en duel)
  - Opérations de type magnétoscope

# Texte

- Difficile de définir un schéma, prenant en compte les spécificités de chaque document : structure hiérarchique de profondeur arbitraire, structure hypertextuelle
- Problème de l'interrogation : parcourir, filtrer, extraire et reconstruire, transformer, présenter
- Solutions : SGML, XML pour la représentation, XQuery pour la manipulation.

# Image

- L'image est décrite par des caractéristiques propres et par des relations (sémantiques, fonctionnelles, spatiales) avec d'autres objets.
- Modélisation : matrice de pixels (stockée à part parfois)
- L'interrogation se fait sur
  - Les attributs descriptifs,
  - La similarité avec d'autres images (couleur, texture, forme)
  - Les propriétés spatiales
  - La texture,
  - La détection de contour

# Son

- Donnée continue (prise en compte du temps)
- Deux façons principales de décrire le son :
  - techniques de traitement du signal (séquence de cellules codant la fréquence ou l'amplitude du signal).
  - Métadonnées décrivant le contenu (souvent fait manuellement)
- Extraction d'informations (intensité, volume, rythme...) pour indexer le contenu d'un son.

# Vidéo

- Problèmes
  - Stockage (avec ou sans compression, différentes normes)
  - Modélisation
  - Indexation
  - Recherche et interrogation
  - Vidéo à la demande
- Modélisation
  - Contenu et structure sémantique
  - Structuration hiérarchique : Film, scènes, séquences, plans, images
  - Annotation manuelle (ou semi-automatique) des segments vidéos



# Interrogation de vidéos

- Indexation :
  - Description du contenu visuel : couleur, texture, mouvement des objets, mouvement de caméra, description des objets, relations spatiales entre objet, changement de scènes
- Interrogation :
  - Langage déclaratif (SQL, OQL) + techniques spécifiques pour interroger sur la nature visuelle, auditive, temporelle et spatiale des données
    - recherche d'images (coucher de soleil, contenant un personnage habillé de rouge, image de la tour Eiffel, ...)
    - Recherche dans des séquences vidéo
    - Recherche de séquences (audio, vidéo), d'objets, de mouvements, de propriétés

# SGBD Multimédias

- Stocker les objets de grande taille, en très grande quantité ( $10^{12}$ ,  $10^{15}$ ,  $10^{24}$  octets)
- Modèles et langages spécifiques
- Interfaces spécifiques
- Index
- Performances acceptables
- Gérer les aspects temporels et spatiaux

# Architecture fonctionnelle

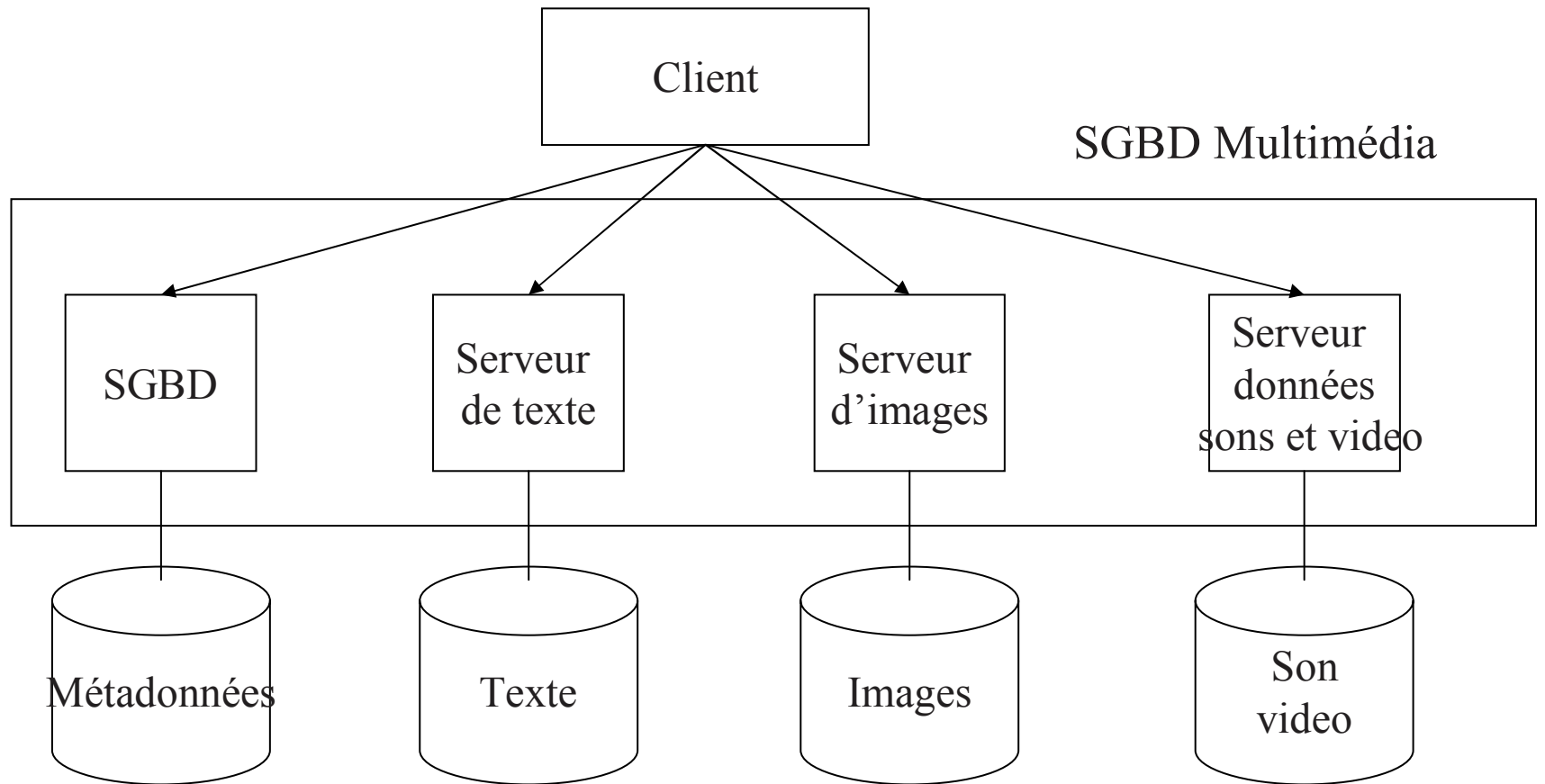
Interface utilisateurs

Présentations multimédias

Serveur d'objets

Schéma, stockage, objets complexes, index,  
transactions

# Architecture physique

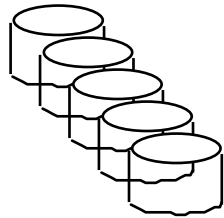


# BD Multidimensionnelles

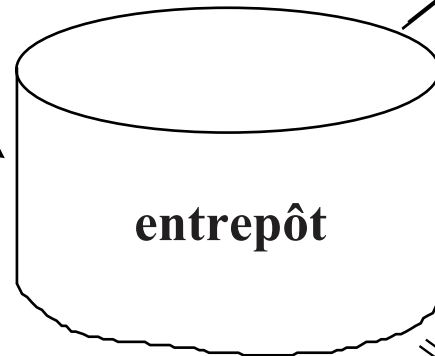
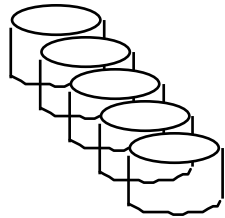
- Les applications d'aide à la décision (OLAP, Decision Support System, Datamining) s'appuient sur des **entrepôts de données**, qui sont de grandes collections d'informations provenant des diverses BD de l'entreprise.
- Les objectifs principaux des entrepôts de données sont
  - **regrouper, organiser, coordonner** des informations provenant de sources **diverses**,
  - les **intégrer** et les **stocker** pour donner à l'utilisateur une vue orientée métier,
  - **retrouver** et **analyser** l'information facilement et rapidement.
- Question typique : *Quels sont les produits qui se vendent le mieux dans chaque région, et quel est l'impact des données démographiques sur ces résultats de vente ?*
- Un entrepôt de données est un **système informationnel** construit à partir des systèmes opérationnels. Pour des raisons de performance, cette BD pour l'aide à la décision est **séparée** des bases de données opérationnelles.

# Architecture

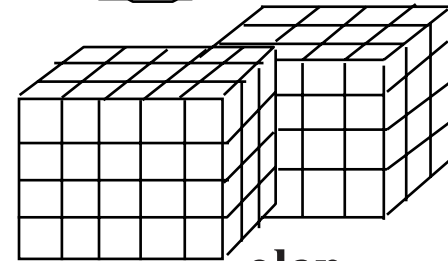
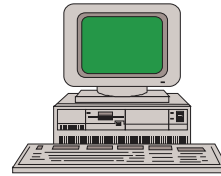
**données externes**  
(connaissances, règles)



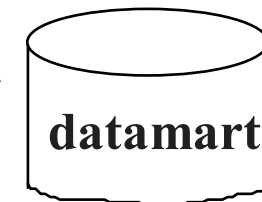
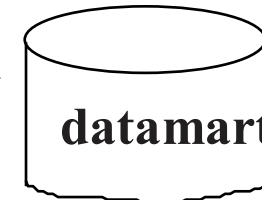
**données de production**  
(y.c. Systèmes légués)



**META-MODELES**



**olap**



# On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

- Comment modéliser les informations pour ce type d'applications ?
- Evolution des structures de données stockées : fichiers, bases de données hiérarchique, **bases de données relationnelles**, bd. distribuées, bd. objet (multimedia).  
⇒ de + en + près de la façon dont les gens *visualisent* et *manipulent* les données
- Bases de données relationnelles [Codd70] :
  - Modèle simple : tables bi-dimensionnelles. Facile à mettre en oeuvre.
  - Le plus répandu. La plupart des entrepôts ont leurs données venant des BD relationnelles.
  - BD relationnelles adaptées à l'OLTP (On-Line Transaction Processing) grâce au standard SQL. Beaucoup de transactions et de requêtes simples, sur peu de données à la fois. Gestion, production.
  - BD relationnelles peu adaptées à l'OLAP (On-Line Analytical Processing).. Requetes moins fréquentes mais plus complexes, longues, nécessitant une reformulation (agrégation) des données de masse. Analyse de données massives (giga, tera octets) stockées dans l'entrepôt pour l'aide à la décision.

# On-Line Analytical Processing & BD multi-dimensionnelles

## Les besoins

- Transformer les données brutes (ex. ventes pour chaque produit, jour, fournisseur) en de l'info. stratégique pour les analyses ultérieures
- Décideurs : finances, ventes, budget, marketing...
- Calculs complexes (analyse de tendance, moyennes mobiles, équations algébriques..) mais requêtes exprimées simplement

*"Total des ventes pour chaque produit par trimestre"*

*"les 5 meilleurs fournisseurs pour chaque catégorie de produit l'an dernier"*

*"pour chaque produit, sa part de marché dans sa catégorie par rapport à celle de 1994"*

*"les fournisseurs dont les ventes ont augmenté dans chaque catégorie de produit ces 5 dernières années"*

...



# On-Line Analytical Processing & BD multi-dimensionnelles

## Les besoins (suite)

- Avoir des résultats à temps, pouvoir raffiner, prolonger ou reprendre les analyses → besoin de garder et visualiser les résultats intermédiaires.
- Appréhender (visualiser) les données dans les dimensions naturelles pour les décideurs

Total des ventes pour chaque produit par trimestre				
	Trim1	Trim2	Trim3	Trim4
Prod1	12	14	32	22
Prod2	15	26	34	7
...				



Total des ventes par catégorie de produit par fournisseur  
par trimestre (+ de détail sur les fournisseurs, - sur les produits)

# On-Line Analytical Processing & BD multi-dimensionnelles

## BD relationnelle et OLTP

Table (relation) *VentesVoiture*

**Attributs**

Modele	Couleur	Ventes
Clio	Bleu	180
Clio	Rouge	244
Jaguar	Bleu	318
Jaguar	Rouge	204
Espace	Bleu	131
Espace	Blanc	153

On veut des détails sur les vendeurs...

## On-Line Analytical Processing & BD multi-dimensionnelles BD relationnelle et OLTP

### *VentesVoiture*

Modele	Couleur	Vendeur	Ventes
Clio	Bleu	Toto	12
Clio	Rouge	Toto	23
Clio	Bleu	Titi	34
Clio	Rouge	Titi	45
Clio	Bleu	Tata	56
Clio	Rouge	Tata	67
Clio	Bleu	Tutu	78
Clio	Rouge	Tutu	89
Jaguar	Bleu	Toto	90
Jaguar	Rouge	Toto	09
Jaguar	Bleu	Titi	98
Jaguar	Rouge	Titi	87
Jaguar	Bleu	Tata	76
Jaguar	Rouge	Tata	65
Jaguar	Bleu	Tutu	54
Jaguar	Rouge	Tutu	43
Espace	Bleu	Toto	32
Espace	Blanc	Toto	21
Espace	Bleu	Titi	11
Espace	Blanc	Titi	22
Espace	Bleu	Tata	33
Espace	Blanc	Tata	44
Espace	Bleu	Tutu	55
Espace	Blanc	Tutu	66

### *Vendeurs*

Vendeur	Succursale
Toto	Auto+
Titi	Auto+
Tata	BelleCaisse
Tutu	BelleCaisse

- **BD relationnelle** : présentation peu conviviale. Données brutes. Normalisation

- **OLTP (SQL)** : mises à jour. Requêtes simples "qui, quoi" (ex. les ventes de Toto).

- Pour l'analyse, besoin de données agrégées, synthétisées

- Possibilité d'agréger (fonctions *de base* uniquement) les données sur une seule table (ex. : somme des ventes par modèle), mais très coûteux (parcourir toutes les tables) et recalcul à chaque étape, à chaque utilisation → *tables de résumés*

- Sur plusieurs tables (ex : somme des ventes par succursale), nécessité de faire des jointures coûteuses (ici  $24 * 4 = 96$  comparaisons) → *dénormalisation*

## On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

### Tables de résumé

- **Idée** : stocker les résultats des fonctions agrégées (ex: table des ventes par vendeur, table de la moyenne des ventes par vendeur de chaque succursale,...) les plus fréquemment utilisés.
- **Problèmes** :
  - tables de résumés pré-définies. Ce qui n'est *pas prévu* n'est *pas disponible*.
  - prolifération des tables de résumés → environnement de décision complexe et confusant

Utilisateurs doivent savoir *quelles tables de résumé* sont disponibles et ce à *quoi elles correspondent*.

- Il faut *rafraîchir* par rapport aux données brutes → ne pas interférer la production (concurrence)

le calcul des agrégats nécessite de lire toutes les données d'une table et donc impose le *verrouillage de tous ses n-uplets*.

## On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

### Dénormalisation

- **Idée** : Stocker toutes les informations dans une seule table pour éviter les jointures.

Modele	Couleur	Vendeur	Ventes	Succursale	Enfants
Clio	Bleu	Toto	12	Auto+	Nono
Clio	Bleu	Toto	12	<i>Auto+</i>	Nana
Clio	Rouge	Toto	23	<i>Auto+</i>	<i>Nono</i>
Clio	Rouge	Toto	23	<i>Auto+</i>	<i>Nana</i>
Clio	Bleu	Titi	34	Auto+	--
Clio	Rouge	Titi	45	<i>Auto+</i>	--
Clio	Bleu	Tata	56	BelleCaisse	--
Clio	Rouge	Tata	67	<i>BelleCaisse</i>	--
Clio	Bleu	Tutu	78	BelleCaisse	--
Clio	Rouge	Tutu	89	<i>BelleCaisse</i>	--
Jaguar	Bleu	Toto	90	<i>Auto+</i>	<i>Nono</i>
Jaguar	Bleu	Toto	90	<i>Auto+</i>	<i>Nana</i>
Jaguar	Rouge	Toto	09	<i>Auto+</i>	<i>Nono</i>
...	...	...	...	...	...
Espace	Blanc	Tutu	66	<i>BelleCaisse</i>	--

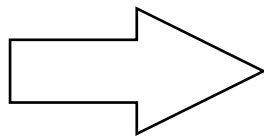
### • Problèmes :

- Grande redondance (ex. succursale et surtout enfants de Toto) → place disque ↑ performances ↓
- Diminue la densité du stockage des données (colonne Enfants *creuse*) → place disque gaspillée
- Augmente aussi la taille et le nombre des index
- Seule alternative → "plus d'acier", i.e.. augmenter les capacités matérielles

## On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

### Vers une autre représentation des données

- *Les bases de données relationnelles ne sont pas adaptées à l'OLAP car les tables représentent une vue aplatie de structures naturellement multi-dimensionnelles.*
- Non seulement perte de performances mais aussi nécessité pour les utilisateurs de savoir comment trouver les liens entre les tables pour recréer la vue multi-dimensionnelle.
- Il est donc nécessaire de disposer d'une *structure de stockage adaptée à l'OLAP*, c'est-à-dire permettant de
  - *visualiser* les données dans *plusieurs dimensions* naturelles,
  - de pouvoir *définir et ajouter des dimensions* facilement,
  - de *manipuler* les données ainsi représentées facilement et efficacement.

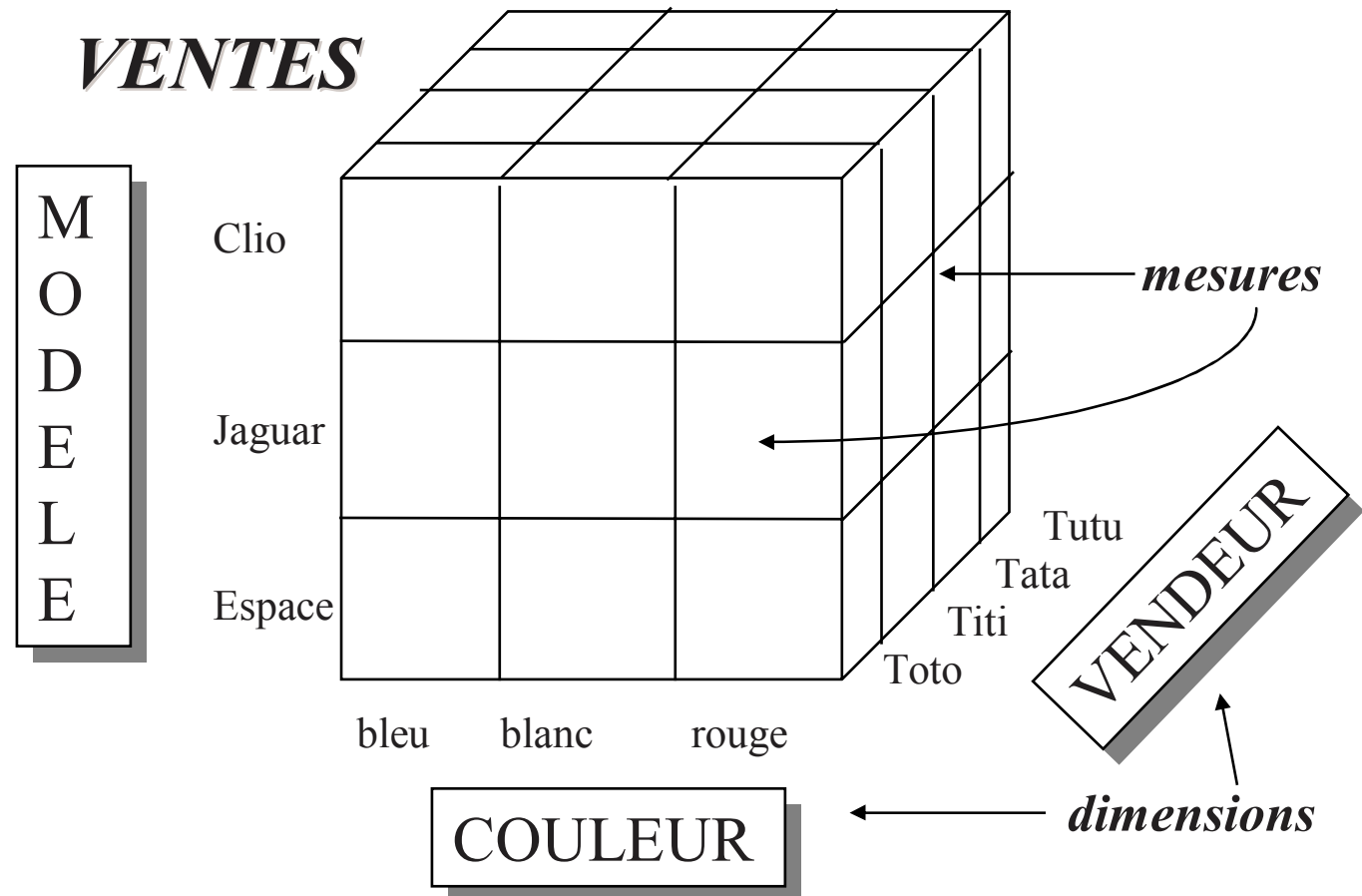


Bases de données multi-dimensionnelles ("Cube")

# On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

## Bases de données multidimensionnelles (1)

- Une *BD multidimensionnelle* [Gray & al. - VLDB'96 ] est un hyper-cube :
  - Les axes sont appelés *dimensions* définies par l'utilisateur
  - Les points dans l'espace (cellules) contiennent des *mesures* calculées à partir de formules plus ou moins complexes.
  - Les *opérateurs* sur le cube sont *algébriques* (retournent un cube) et peuvent ainsi être *combinés*



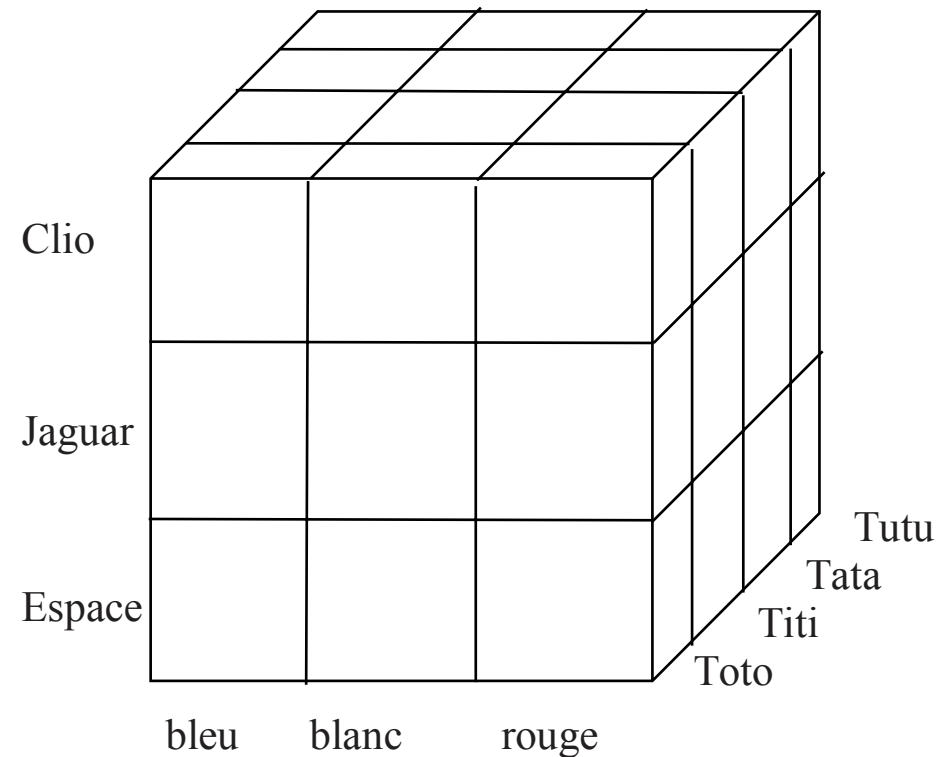
➡ Base de données multi-dimensionnelle = "super-tableur"

## On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

### Bases de données multidimensionnelles (2)

- Représenter ce cube nécessiterait en relationnel une table de  $4 \times 3 \times 3 = 36$  nuplets à "balayer"
- Ici, pour retrouver une valeur dans une cellule, il faut faire  $4 + 3 + 3 = 10$  recherches seulement

Modele	Couleur	Vendeur	Ventes	nuplet
Clio	Bleu	Toto	12	1
Clio	Rouge	Toto	23	2
Clio	Blanc	Toto	22	3
Clio	Bleu	Titi	34	4
Clio	Rouge	Titi	45	5
Clio	Blanc	Titi	48	6
Espace	Bleu	Tutu	55	7
Espace	Blanc	Tutu	66	8
...	...	...	...	...
Espace	Bleu	Tutu	55	35
Espace	Blanc	Tutu	66	36

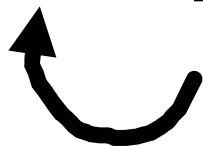
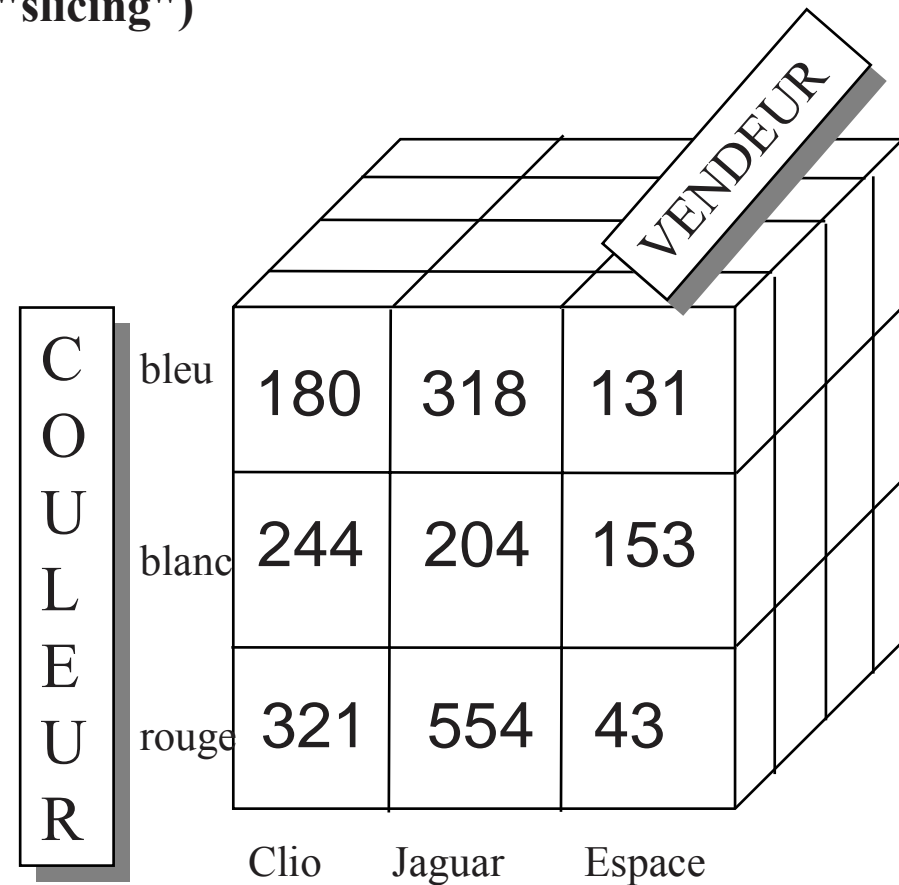
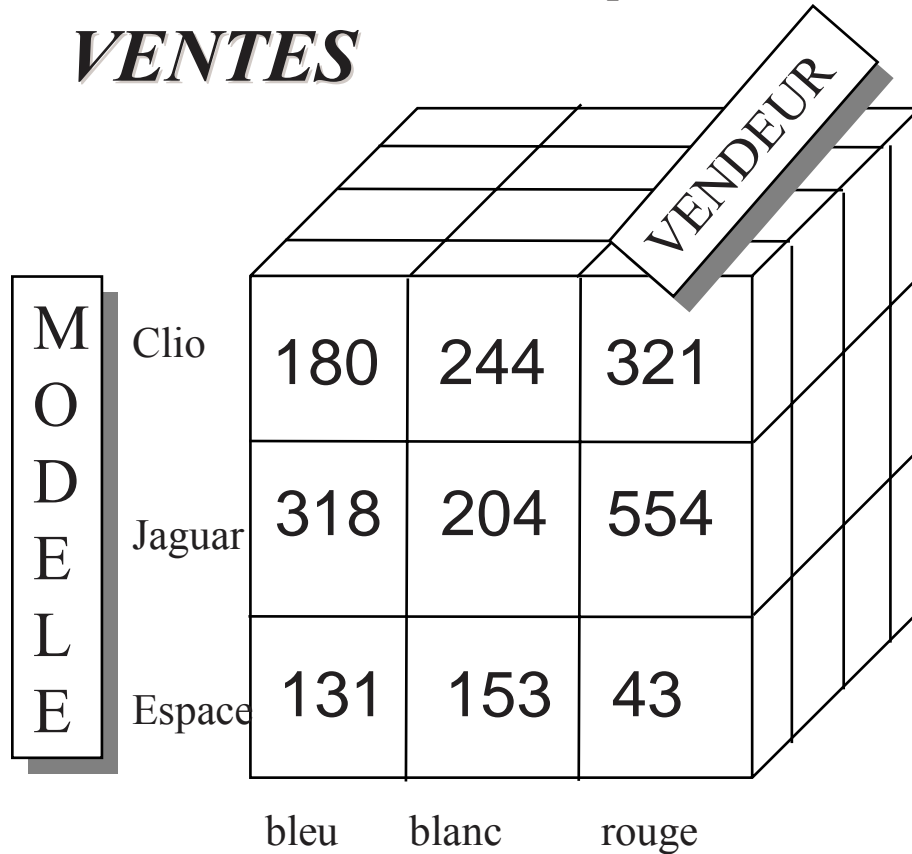




# On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

Opérateurs : **rotation** ("slicing")

**VENTES**



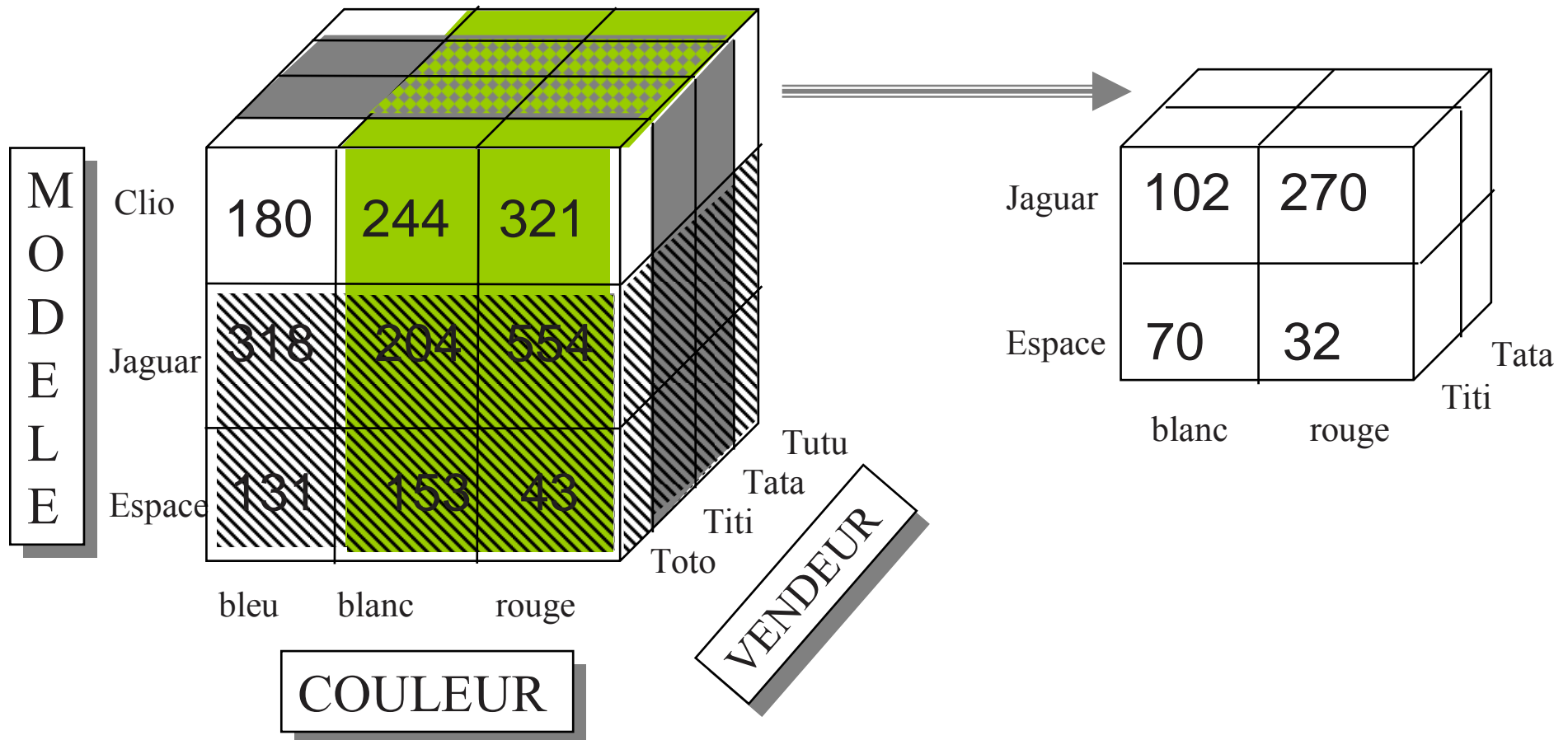
Rotation 90



*Aucun réarrangement des données. 6 vues possibles directement*

On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles  
Opérateurs : **extraire** ("dicing")

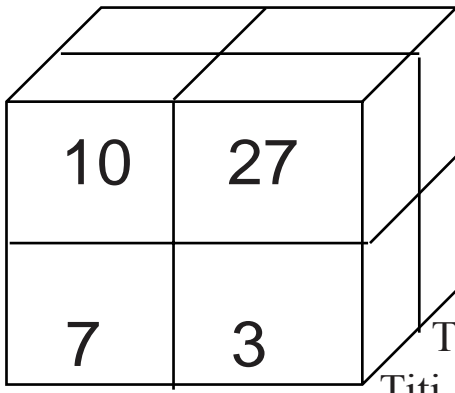
**VENTES**



*Recherche des cellules rapide. Recalcul plus rapide sur les faces*

On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles  
Opérateurs : **ajouter une dimension**

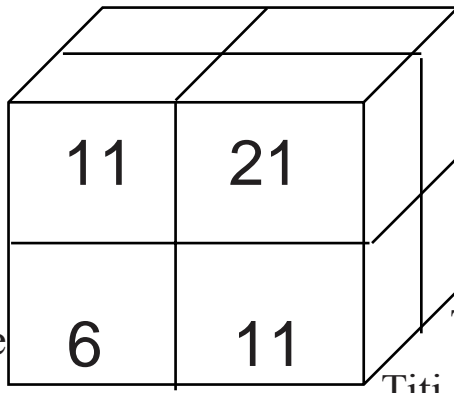
***VENTES***



A 3D cube diagram representing sales data for January. The vertical axis is labeled 'Jaguar' and 'Espace'. The horizontal axis is labeled 'blanc' and 'rouge'. The depth axis is labeled 'Tata' and 'Titi'. The data values are: Jaguar-blanc = 10, Jaguar-rouge = 27, Espace-blanc = 7, Espace-rouge = 3.

	blanc	rouge
Jaguar	10	27
Espace	7	3

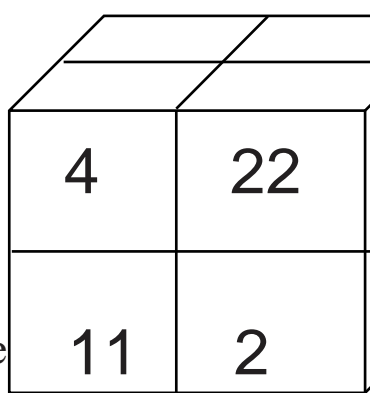
janvier



A 3D cube diagram representing sales data for February. The vertical axis is labeled 'Jaguar' and 'Espace'. The horizontal axis is labeled 'blanc' and 'rouge'. The depth axis is labeled 'Tata' and 'Titi'. The data values are: Jaguar-blanc = 11, Jaguar-rouge = 21, Espace-blanc = 6, Espace-rouge = 11.

	blanc	rouge
Jaguar	11	21
Espace	6	11

février



A 3D cube diagram representing sales data for March. The vertical axis is labeled 'Jaguar' and 'Espace'. The horizontal axis is labeled 'blanc' and 'rouge'. The depth axis is labeled 'Tata' and 'Titi'. The data values are: Jaguar-blanc = 4, Jaguar-rouge = 22, Espace-blanc = 11, Espace-rouge = 2.

	blanc	rouge
Jaguar	4	22
Espace	11	2

mars

*Possibilité d'ajouter une dimension à tout moment*

# On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

## Hiérarchies de dimensions

- Les dimensions peuvent être vues à différents niveaux de granularité, qui correspondent à différents niveaux de consolidation des données.

- Exemples :

  - Vendeur → Succursale → Région → Entreprise

  - Jour → Mois → Trimestre → Année → Décennie

- Ces "chemins de consolidation de données" correspondent aux hiérarchies naturelles de l'entreprise.

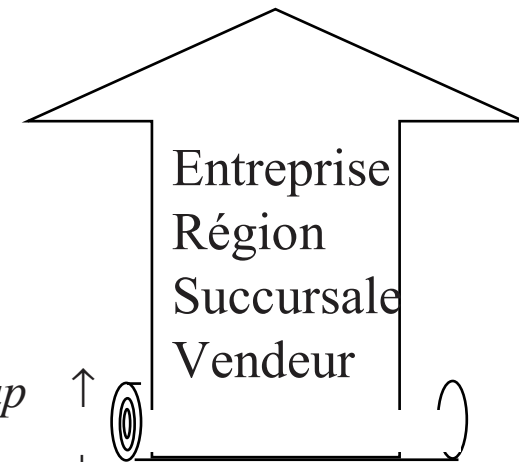
- Il peut en exister plusieurs pour la même dimension

  - Jour → Semaine → Année → Décennie

- Faire passer d'une vue à une vue moins détaillée : opérateur *roll-up*

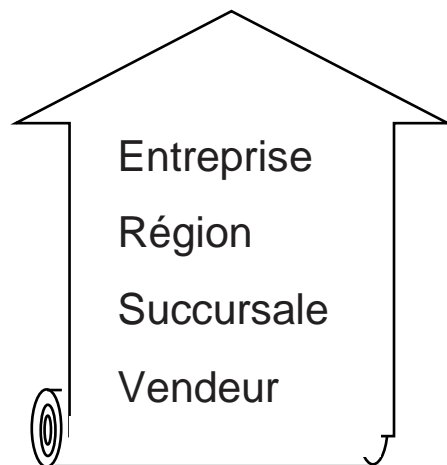
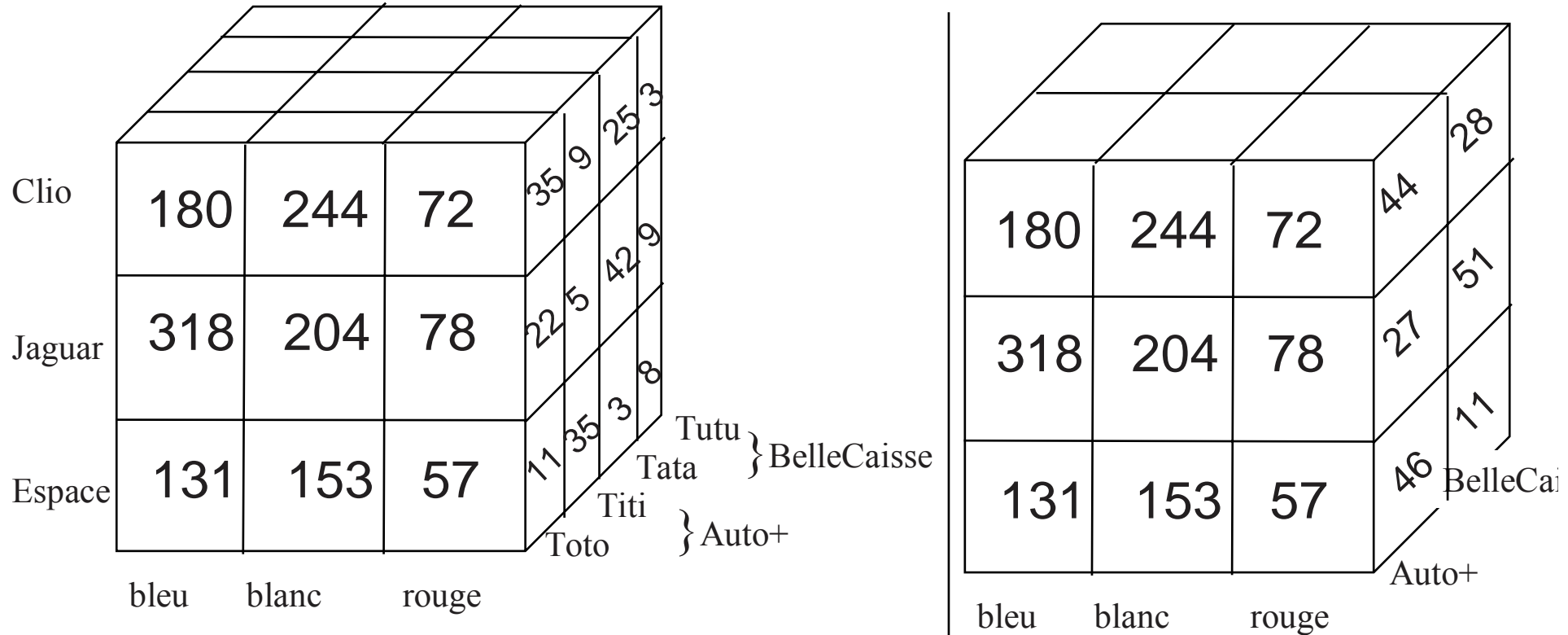
- " " plus " " *drill-down*

- Ces opérateurs doivent être quasi-instantanés



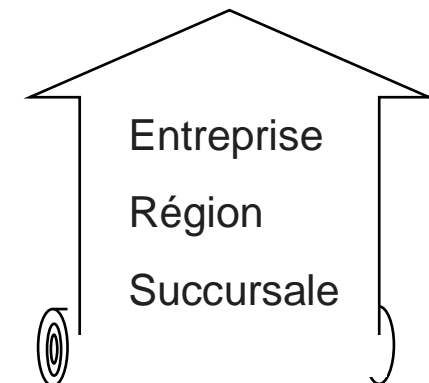
# On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

## Opérateurs : **Roll-up** et **Drill-down**



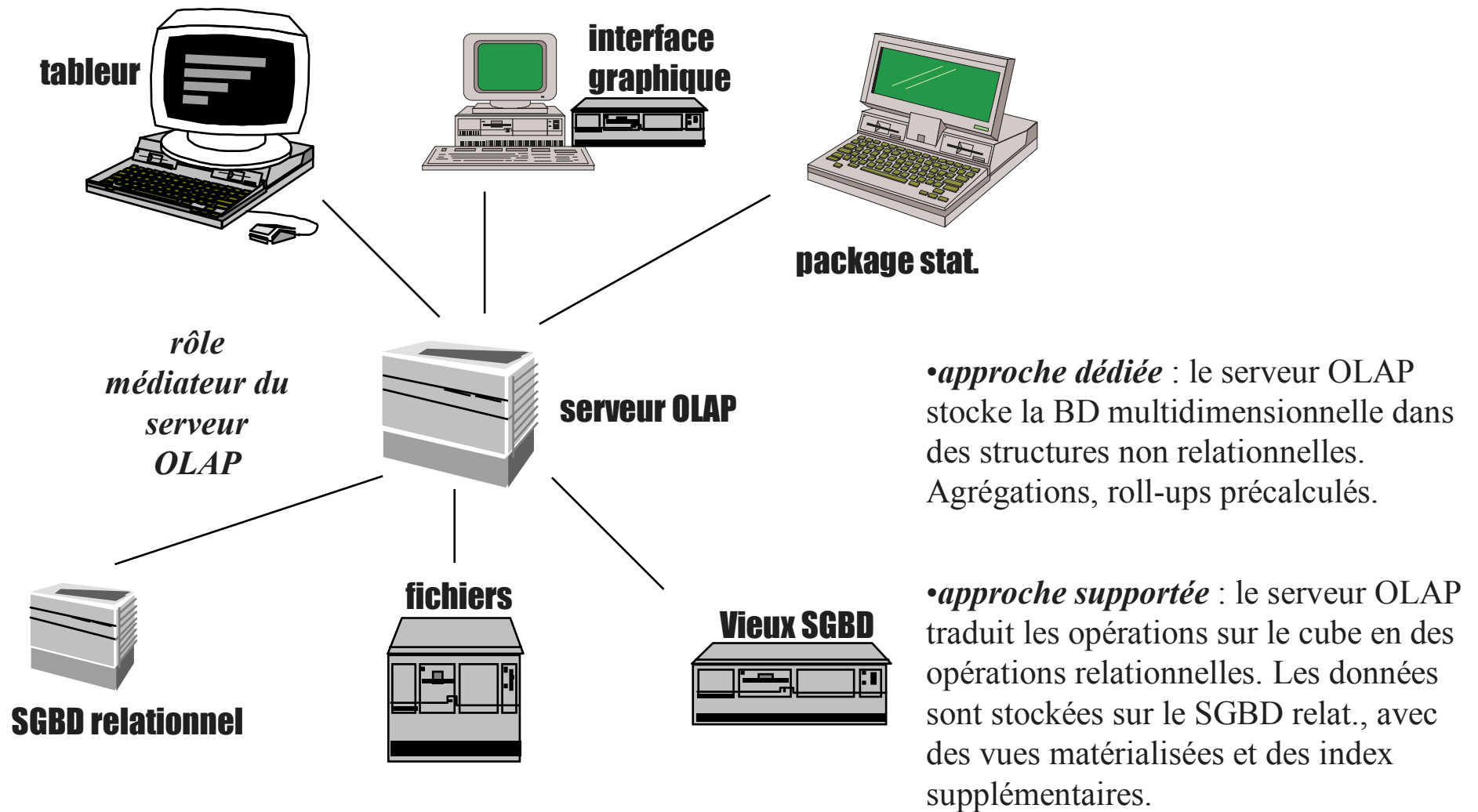
Roll-up

Drill-down



## On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

### Architectures possibles

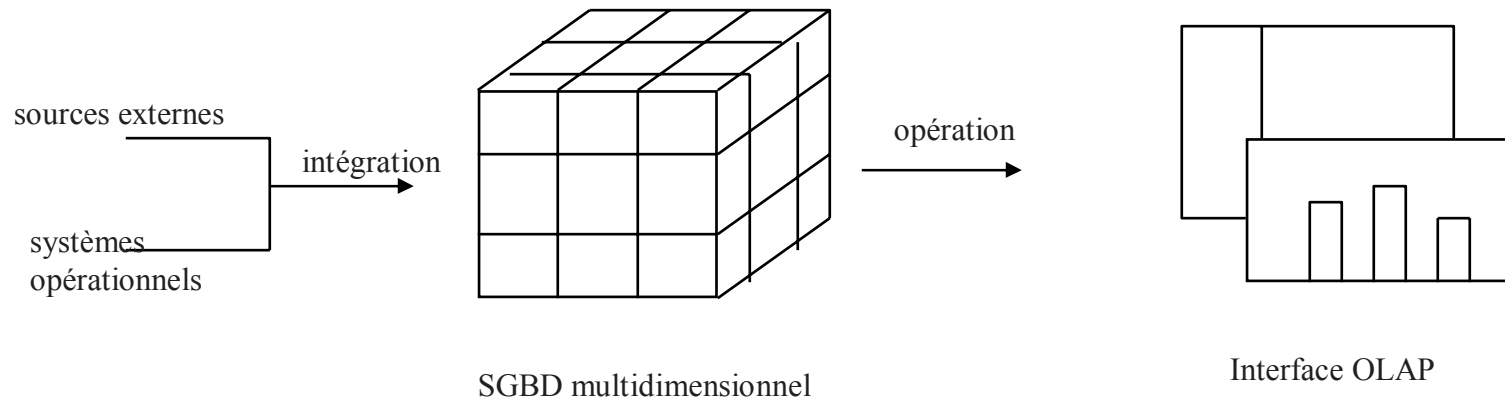


## On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles

### Représentation des données

#### Les systèmes spécialisés : **MOLAP** (OLAP multidimensionnel)

Les données multidimensionnelles sont stockées dans un SGBD dont les structures sont optimisées pour le stockage et le traitement des données. Accès rapide en lecture/écriture pour de grandes quantités de données. Les données sont séparées en plusieurs cubes denses de petite taille, pour traiter la dispersion des données (un petit nombre de cellules possibles a une valeur).

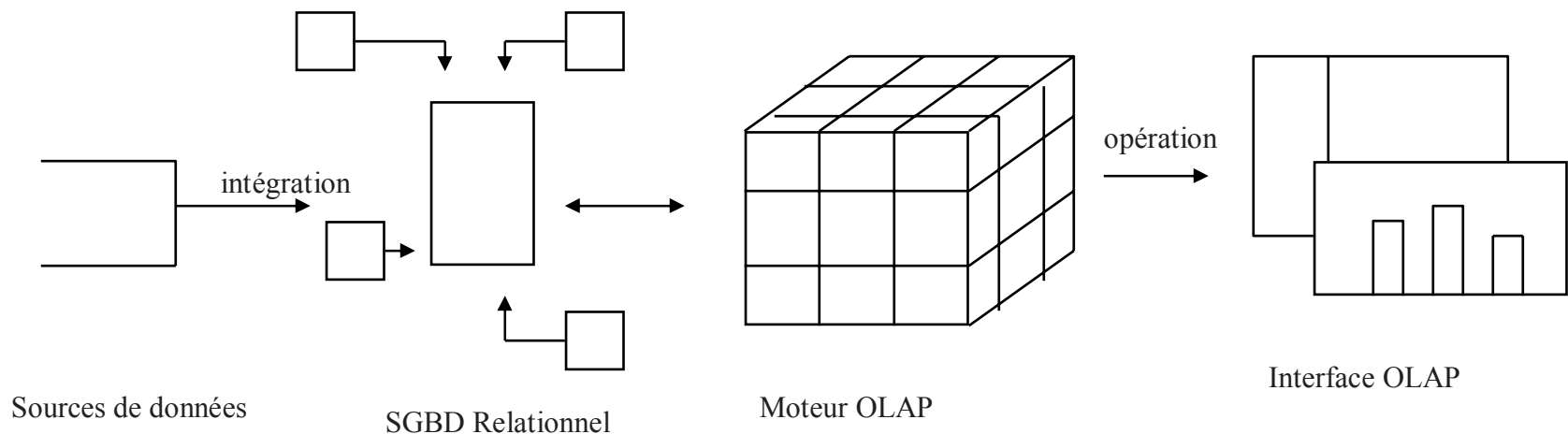


• *Arbor Essbase, IRI Express, Pilot (Pilot Software), ...*

On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles  
**Représentation des données**

**Les systèmes relationnels : ROLAP**

Les données multidimensionnelles sont stockées dans un SGBD relationnel. Elles sont organisées en schémas en forme d'étoiles ou de flocon. Accès en mode lecture.



•*Redbrick, Microstrategy, MetaCube (Informix)...*



On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles  
**Architectures possibles (bilan)**

•**Dans les 2 approches** : les efforts de recherche portent à la fois sur le *calcul du cube* à partir de tables relationnelles (alimentation du serveur), et sur le *calcul d'opérations complexes à partir du cube* (principalement ré-agrégation de données agrégées)

•**Calcul du cube** : par exemple, on veut que l'agrégation sur 3 dimensions soit la somme.

Equivaut a  $2^3 = 8$  "group-by" de SQL (toutes les combinaisons de dimensions)

2 types d'optimisation :

- trier les données pour calculer plusieurs agrégats à partir du même tri
- enchaîner les calculs de manière à utiliser certains résultats intermédiaires à la volée (pipe-line)

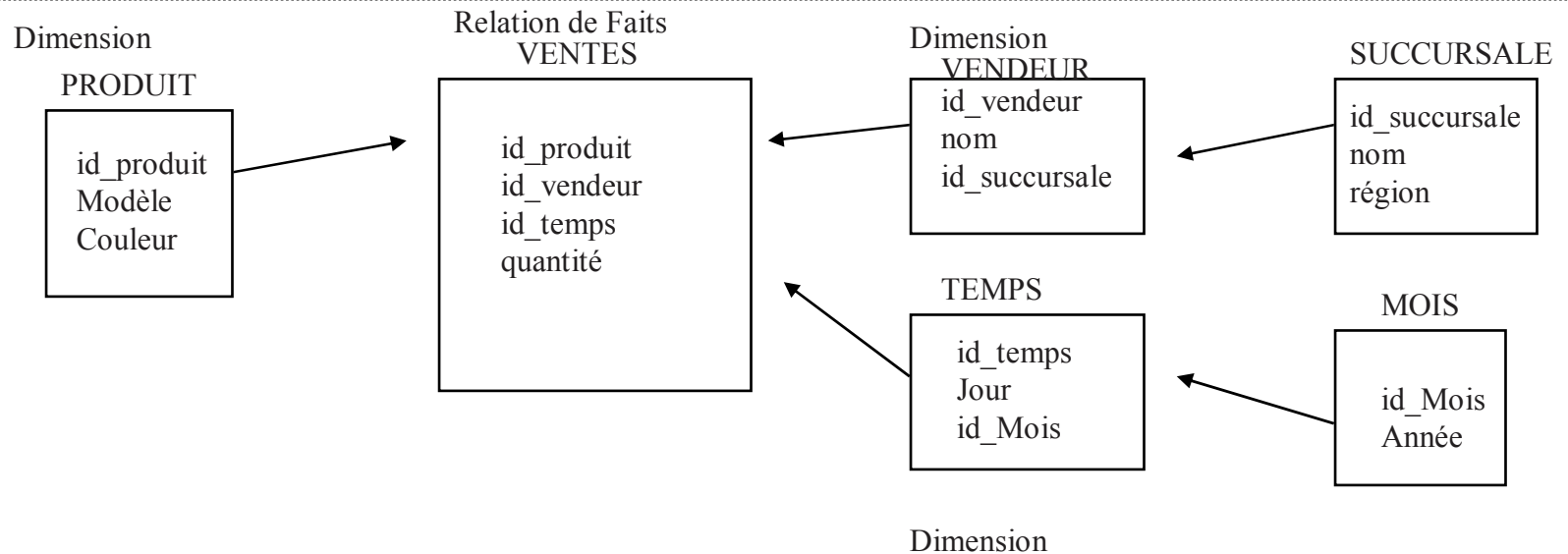
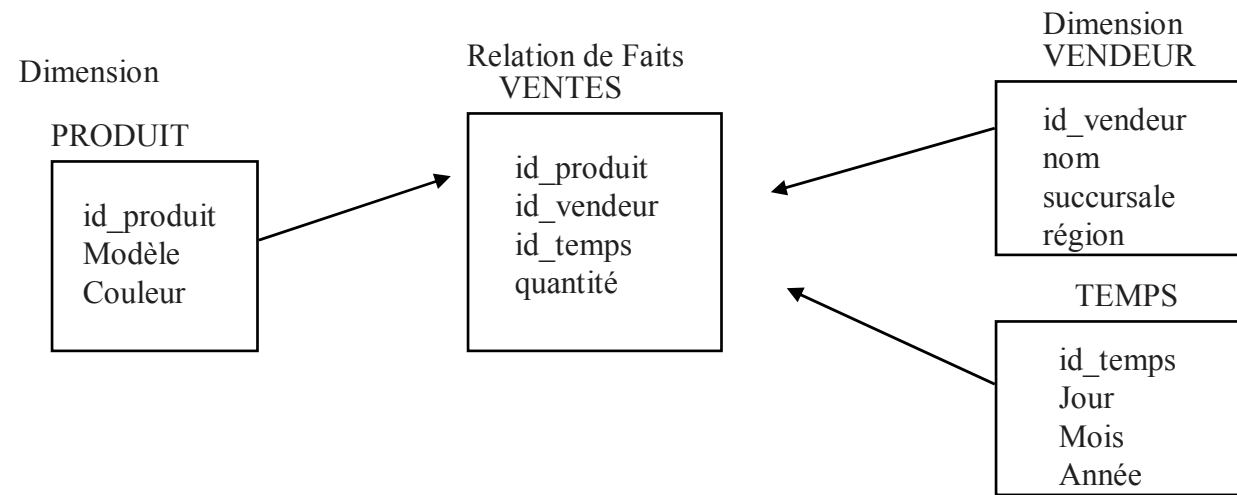
•**calcul d'opérations complexes** :

- précalculer certaines informations auxiliaires (pas plus que la taille du cube)
- mettre à jour ces informations incrémentalement en batch
- utilisation de techniques de codages couvrants...

# Schémas

- Schéma en étoile :
  - 1 table de faits centrale et plusieurs tables de dimensions dénormalisées
  - Les mesures sont stockées dans la table de faits
  - Il existe une table de dimension pour chaque dimension avec tous les niveaux d'agrégation
- Schéma en flocon
  - Version normalisée du schéma en étoile
  - Traitement explicite des hiérarchies de dimension (chaque niveau est représenté dans une table différente)
  - Plus facile à maintenir, plus lent lors de l'interrogation.

On-Line Analytical Processing (OLAP) & BD multi-dimensionnelles  
**Schémas en étoile et en flocon**



BD multidimensionnelles  $\neq$  panacée.

- La technologie multidimensionnelle n'est pas toujours adaptée et ne doit pas être utilisée "à toutes les sauces"
- Exemple :

**Personnel**

Employé	#employé	Age
Toto	01	21
Titi	12	19
Tata	31	63
Tutu	14	31
Tyty	54	27
Bobo	03	56
Bibi	41	45
Baba	33	41
Bubu	23	19



**Personnel**

E M P L O Y E	Toto			21					
	Titi							19	
	Tata	63							
	Tutu				31				
	Tyty					27			
	Bobo						56		
	Baba		45						
	Bibi								41
	Bubu			19					
		31	41	23	01	14	54	03	12

# EMPLOYE

Les données sur le personnel **ne sont pas multidimensionnelles** : pas de relation entre les éléments des différents nuplets

**Les précurseurs des BD multidimensionnelles :  
BD temporelles, BD spatiales, BD statistiques**

- **Bases de données temporelles** : stockage et interrogation des valeurs des différents item de la base selon une ou plusieurs dimensions temporelles (les plus utiles : temps de validité, temps de transaction)

*seul le temps est traité comme une dimension*

- **Bases de données spatiales** : représentation des item selon leur forme et position dans l'espace.

*opérateurs développés géométriques, loin des opérateurs OLAP*

*possibilité d'utiliser les travaux sur l'indexation de telles bases*

- **Bases de données statistiques** : bcp de points communs, mais sans construire un nouveau modèle. Extension du modèle relationnel pour supporter les tables de résumé et les *traitements* statistiques.

*Traitement non-uniforme des dimensions et des mesures.*

*Récupérer les techniques d'implémentation (notamment vues agrégées)*

# Règles d'or de Codd

- 1993 : E.F. Codd formule 12 *règles d'or* (à la demande de Arbor soft. !!)
- 1995 : 18 règles en 4 groupes :

## **Basiques :**

1. vue multidimensionnelle
2. manipulation directe
3. médiation (accessibilité)
4. intégration d'approche dédiée **et** d'approche supportée.
5. support de tous les modèles d'analyse des entreprises  
(seuls les plus simples sont habituellement supportés)
6. Client/serveur
7. Transparence (ne pas avoir à savoir d'où viennent les données,  
même si elles viennent de sources externes).
8. Multi-utilisateurs (lecture seule ?)

# Règles d'or de Codd

## **Caractéristiques spéciales**

- 9. Traitement des données dénormalisées
- 10. Stockage des résultats à part (ne pas interférer avec les mise à jour des transactions de production)
- 11. Représentation des valeurs manquantes
- 12. Traitement des valeurs manquantes.

## **Présentation des rapports:**

- 13. Flexibilité (ajout de dimension...)
- 14. Performances non dégradées si nb. dim. ou taille BD augmente.
- 15. Ajustement de la représentation physique

## **Contrôle des dimensions :**

- 16. Généricité : traitement équivalent de chaque dimension.
- 17. Nombre et profondeur illimités (actuellement, max. = 10 et 6)
- 18. Calculs à travers n'importe quelles dimensions.

## Test FASMI

(Fast Analysis of Shared Multidimensional Information)

- Voulu plus simple, réaliste et général que les règles de Codd
- **Fast** : 1 seconde pour les analyses de bases, - de 5 secondes pour la plupart, très peu au dessus de 20 secondes (au-delà de 30 secondes, CTRL-ALT-DEL!). Même si on a maintenant en 5 minutes ce qui durait des heures, l'utilisateur perd le fil de son raisonnement...
- **Analysis** : doit servir pour n'importe quelle analyse logique ou statistique assez facilement (sans programmer), que ce soit par des outils internes ou des appels a des outils externes (ex. tableur).
- **Shared** : les bonnes propriétés habituelles "multi-utilisateurs" des SGBD : concurrence d'accès en écriture, confidentialité, sécurité.
- **Multidimensional** (view)
- **Information** : toute l'information nécessaire doit pouvoir être produite (rapport de 1 à 1000 entre le produit le moins puissant et le plus puissant). Benchmarking...



# Olap Council (<http://www.olapcouncil.org>)

- Fondé pour le développement et la standardisation de l'OLAP
- Regroupe la plupart des vendeurs d'OLAP (mais pas tous!)
- OLAP MDAPI : Interface standard que doivent fournir les serveurs OLAP, de manière à ce que différents outils d'analyse puissent se développer par rapport à ces spécifications.
- Interopérabilité : le même outil d'analyse pourra alors utiliser simultanément des données provenant de différents serveurs OLAP.
- MDAPI V.5 disponible sur WWW pour commentaires. Uniquement pour des analyses de consultation. Prochaines versions inclueront les mises à jour rétroactives.
- Même stratégie pour le benchmarking : package complet APB-1

## Problèmes ouverts en OLAP

- Langage et optimisation de requêtes
- Stockage et indexation des BD multidimensionnelles
- Utilisation pour le Data Mining
- Mises à jour directe sur le cube. Rétro-action sur les données brutes.