

<http://www-adele.imag.fr/users/Didier.Donsez/cours>

Les SGBDs Décisionnels

Didier DONSEZ

Université Joseph Fourier - Grenoble 1

PolyTech'Grenoble LIG/ADELE

`Didier.Donsez@imag.fr`, `Didier.Donsez@ieee.fr`

Thierry CRUANES

Oracle Corp., US

`thierry.cruanes@oracle.com`

Plan

- Limites du SQL en décisionnel
- Nouvelles fonctions et types de données
- Aspect Systèmes
 - Index Bitmap
 - RAID
- Benchmarks

Limites du SQL et des SGBDs « Transactionnels »

orienté vers le OLTP

- Interface

- Manque d'expressivité des requêtes SQL pour l'OLAP

- Système

- Structures et Algorithmes inadaptées
à la charge de travail de l'OLAP

Winter 2003 VLDB Survey

Largest Database Size, Decision Support

1998 Survey

| | | | |
|-----|-------------------|---------------|-------------|
| 1. | Sears | Teradata | 4.63 |
| 2. | HCIA | Informix | 4.50 |
| 3. | Wal-Mart | Teradata | 4.42 |
| 4. | Tele Danmark | DB2 | 2.84 |
| 5. | CitiCorp | DB2 | 2.47 |
| 6. | MCI | Informix | 1.88 |
| 7. | NDC Health | Oracle | 1.85 |
| 8. | Sprint | Teradata | 1.30 |
| 9. | Ford | Oracle | 1.20 |
| 10. | Acxiom | Oracle | 1.13 |

2001 Survey

| | | |
|-------------------|---------------|-------------|
| SBC | Teradata | 10.50 |
| First Union | Informix | 4.50 |
| Dialog | Proprietary | 4.25 |
| Telecom Italia | DB2 | 3.71 |
| FedEx | Teradata | 3.70 |
| Office Depot | Teradata | 3.08 |
| AT & T | Teradata | 2.83 |
| SK C&C | Oracle | 2.54 |
| NetZero | Oracle | 2.47 |
| Telecom Italia | Informix | 2.32 |

2003 Survey

| | | |
|-----------------------|---------------|--------------|
| France Telecom | Oracle | 29.23 |
| AT&T | Daytona | 26.27 |
| SBC | Teradata | 24.81 |
| Anonymous | DB2 | 16.19 |
| Amazon.com | Oracle | 13.00 |
| Kmart | Teradata | 12.59 |
| Claria | Oracle | 12.10 |
| HIRA | Sybase IQ | 11.94 |
| FedEx | Teradata | 9.98 |
| Vodafone | Teradata | 9.91 |

Exemple

- France Telecom
 - ajoute 1To de données brutes par mois
 - DW de 65 To actuellement.
- Comment
 - Gérer ce volume
 - Maintenir les index, les partitions, les vues matérialisées, les stats (plusieurs jours de calculs *from scratch*)

Amazon.com

- TODO

Systemes

- MOLAP (obsolète)
 - SGBD Spécialisé
- ROLAP (inefficace)
 - SGBD Relationnel
- H-OLAP (*H comme Hybrid*) (l'actualité)
 - SGBD Relationnel avec des extensions OLAP
 - Nouveaux opérateurs
 - GROUP BY CUBE, ...
 - Algorithmes adéquats
 - StarJoin, Index Bitmap, ...

Nouvelles Fonctions pour SQL

- BREAK BY (SAS)
- RANK
 - Rang d'une ligne par rapport à un agrégat
- TOP / BOTTOM
 - Requête de type « Top Ten »
 - Les 10 meilleurs / Les 10 moins bons
- Echantillonnage (Informix Online XPS)
 - Requête effectuée sur un échantillon de données tiré aléatoirement (preview)
 - Limite de temps de calcul pour une approximation du résultat
- Extension du Group By (SQL99)
 - Grouping Sets, Rollup, Cube
- Fenêtre mouvante pour les agrégat : Windows
 - Exemple : moyenne et déviance sur le jour courant et les 3 jours précédents
- SQL:2003 OLAP

Ranking

- `select rank('01-Jan-1990','12-Dec-1989')`
within group (order by `#client`, `#produit`) montant
from `Ventes`;

GROUP BY GROUPING SETS

- Partitionnement selon plusieurs dimensions

- Exemple

```
SELECT #Client, #Produit, SUM(Montant) FROM Ventes
```

```
GROUP BY GROUPING SETS ((#Client), (#Produit))
```

- est équivalent à

```
(SELECT #Client, NULL, SUM(Montant) FROM Ventes GROUP BY #Client)
```

```
UNION
```

```
(SELECT NULL, #Produit, SUM(Montant) FROM Ventes GROUP BY #Produit)
```

GROUP BY ROLLUP

- Réduire progressivement

- Exemple

```
SELECT #Client, #Produit, SUM(Montant) FROM Ventes  
GROUP BY ROLLUP (#Client, #Produit)
```

- est équivalent à

```
(SELECT #Client, #Produit, SUM(Montant) FROM Ventes GROUP BY #Client,  
#Produit)
```

```
UNION
```

```
(SELECT #Client, NULL, SUM(Montant) FROM Ventes GROUP BY #Client)
```

```
UNION
```

```
(SELECT NULL, NULL, SUM(Montant) FROM Ventes)
```

GROUP BY ROLLUP

```
SELECT MONTH(SALES_DATE), REGION, SALES_MGR, SUM(SALES)
FROM SALES WHERE YEAR(SALES_DATE) = 1996
GROUP BY ROLLUP (MONTH(SALES_DATE), REGION, SALES_MGR)
```

| ■ MONTH | REGION | SALES_MANAGER | SUM(SALES) |
|---------|---------|---------------|------------|
| ■ April | Central | Chow | 25000 |
| ■ April | Central | Smith | 15000 |
| ■ April | Central | - | 40000 |
| ■ April | North | - | 15000 |
| ■ April | North | - | 15000 |
| ■ April | - | - | 55000 |
| ■ May | Central | Chow | 25000 |
| ■ May | Central | - | 25000 |
| ■ May | North | Smith | 15000 |
| ■ May | North | - | 15000 |
| ■ May | - | - | 40000 |
| ■ - | - | - | 95000 |

GROUP BY CUBE

- Partitionnement selon tous les sous-ensembles possibles de Grouping Sets

```
SELECT #Client, #Produit, #Fournisseur, SUM(Montant) FROM Ventes
GROUP BY CUBE (#Client, #Produit, #Fournisseur)
```

- est équivalent à

```
SELECT #Client, #Produit, SUM(Montant) FROM Ventes
```

```
GROUP BY GROUPING SETS (
```

```
(), -- total des ventes
(#Client), -- total des ventes par Client
(#Fournisseur), -- total des ventes par Fournisseur
(#Produit), -- total des ventes par Produit
(#Client, #Fournisseur) -- total des ventes par Client et par Fournisseur
(#Client, #Produit), -- total des ventes par Client et par Produit
(#Produit, #Fournisseur), -- total des ventes par Produit et par Fournisseur
(#Client, #Produit, #Fournisseur) -- total des ventes par Client, Produit et Fournisseur
)
```

Fenêtre glissante

- But: cumul, moyenne, dérivation sur une fenêtre glissante du temps

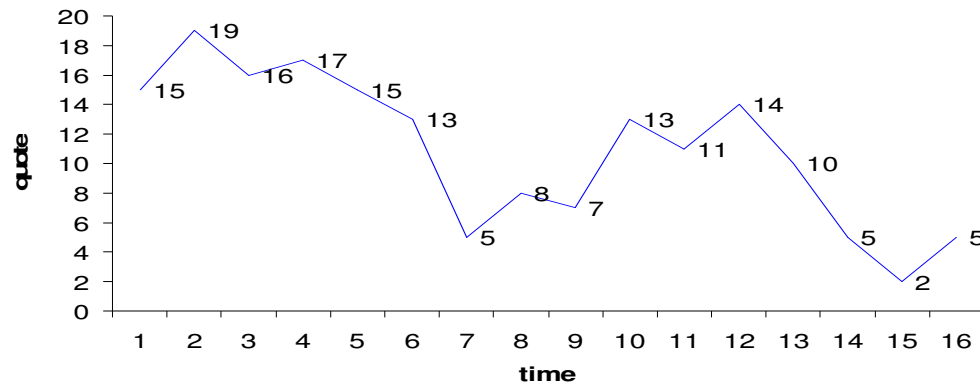
```
SELECT Sf.Region, Sf.Month, Sf.Sales, AVG (Sf.Sales) OVER (
    PARTITION BY Sf.Region ORDER BY Sf.Month ASC ROWS 1 PRECEDING
) AS Moving_avg FROM SalesFact AS Sf ORDER BY Sf.Month ASC;
```

| Region | Month | Sales | Moving_Avg |
|--------------|-----------|-------|------------|
| Eastern | Oct, 1994 | 27497 | 27497 |
| Eastern | Nov, 1994 | 24168 | 25832 |
| Eastern | Dec, 1994 | 27801 | 25984 |
| Eastern | Jan, 1995 | 25991 | 26896 |
| Eastern | Feb, 1995 | 25968 | 25979 |
| Eastern | Mar, 1995 | 23610 | 24789 |
| Mid Atlantic | Oct, 1994 | 7150 | 7150 |
| Mid Atlantic | Nov, 1994 | 7586 | 7368 |
| Mid Atlantic | Dec, 1994 | 6164 | 6875 |
| Mid Atlantic | Jan, 1995 | 6051 | 6108 |
| Mid Atlantic | Feb, 1995 | 4299 | 5175 |
| Mid Atlantic | Mar, 1995 | 6283 | 5291 |

Exercice : Best-profit Query

D'après Alberto Lerner and Dennis Shasha

Find the best profit one could make by buying a stock and selling it later in the same day



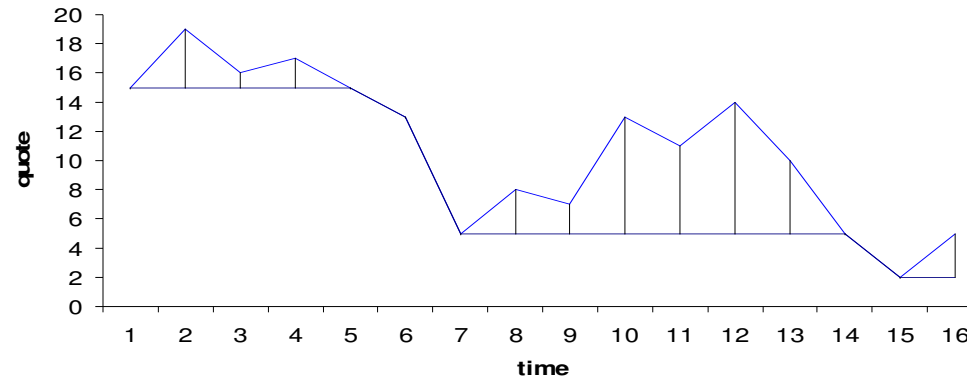
price

15 19 16 17 15 13 5 8 7 13 11 14 10 5 2 5

Exercice : Best-profit Query

D'après Alberto Lerner and Dennis Shasha

Find the best profit one could make by buying a stock and selling it later in the same day



| | | | | | | | | | | | | | | | | |
|-------------|----|----|----|----|----|----|---|---|---|----|----|----|----|---|---|---|
| price | 15 | 19 | 16 | 17 | 15 | 13 | 5 | 8 | 7 | 13 | 11 | 14 | 10 | 5 | 2 | 5 |
| mins(price) | 15 | 15 | 15 | 15 | 15 | 13 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 2 |
| | 0 | 4 | 1 | 2 | 0 | 0 | 0 | 3 | 2 | 8 | 6 | 9 | 0 | 0 | 0 | 3 |

Solution : Best-profit Query

D'après Alberto Lerner and Dennis Shasha

```
[SQL:1999]
SELECT max(rdif)
FROM    (SELECT ID,tradeDate,
             price - min(price)
             OVER
             (PARTITION BY ID,tradeDate
             ORDER BY timestamp
             ROWS UNBOUNDED
             PRECEDING) AS rdif
          FROM    Ticks ) AS t1
WHERE   ID="S"
        AND   tradeDate='1/10/03'
```

MS MDX (*Multidimensional Expression*)

- Langage d'expression OLAP pour MS SQL Server
- Exemples
 - SELECT
NON EMPTY {[Time].[1997], [Time].[1998]} ON COLUMNS,
[Promotion Media].[Media Type].Members ON ROWS
FROM Sales

 - SELECT
{[Measures].[Unit Sales]} ON COLUMNS,
ORDER(EXCEPT([Promotion Media].[Media Type].members, {
[Promotion Media].[Media Type].[No Media]}), [Measures].[Unit
Sales], DESC) ON ROWS
FROM Sales

ADT Séries Temporelles (Time Series)

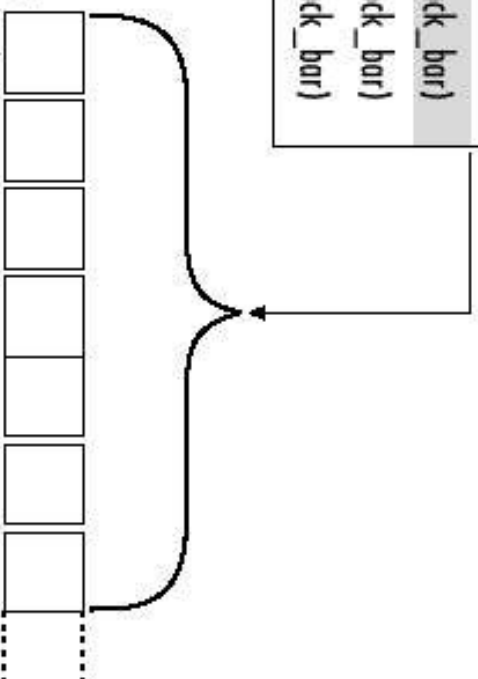
- Définition
 - Suite de couple (Valeur, estampille de temps)
- Applications
 - Finance (stock value), Santé (épidémiologie), ...
- Type
 - calendar, ...
- Opérations
- Index
 - par rapport au temps

ADT Séries Temporelles

Database Table

| Stock_id | Stock_data |
|----------|-----------------------|
| IFMX | Timeseries(stock_bar) |
| IBM | Timeseries(stock_bar) |
| HWP | Timeseries(stock_bar) |

Collection of elements of
IFMX times-series data entry

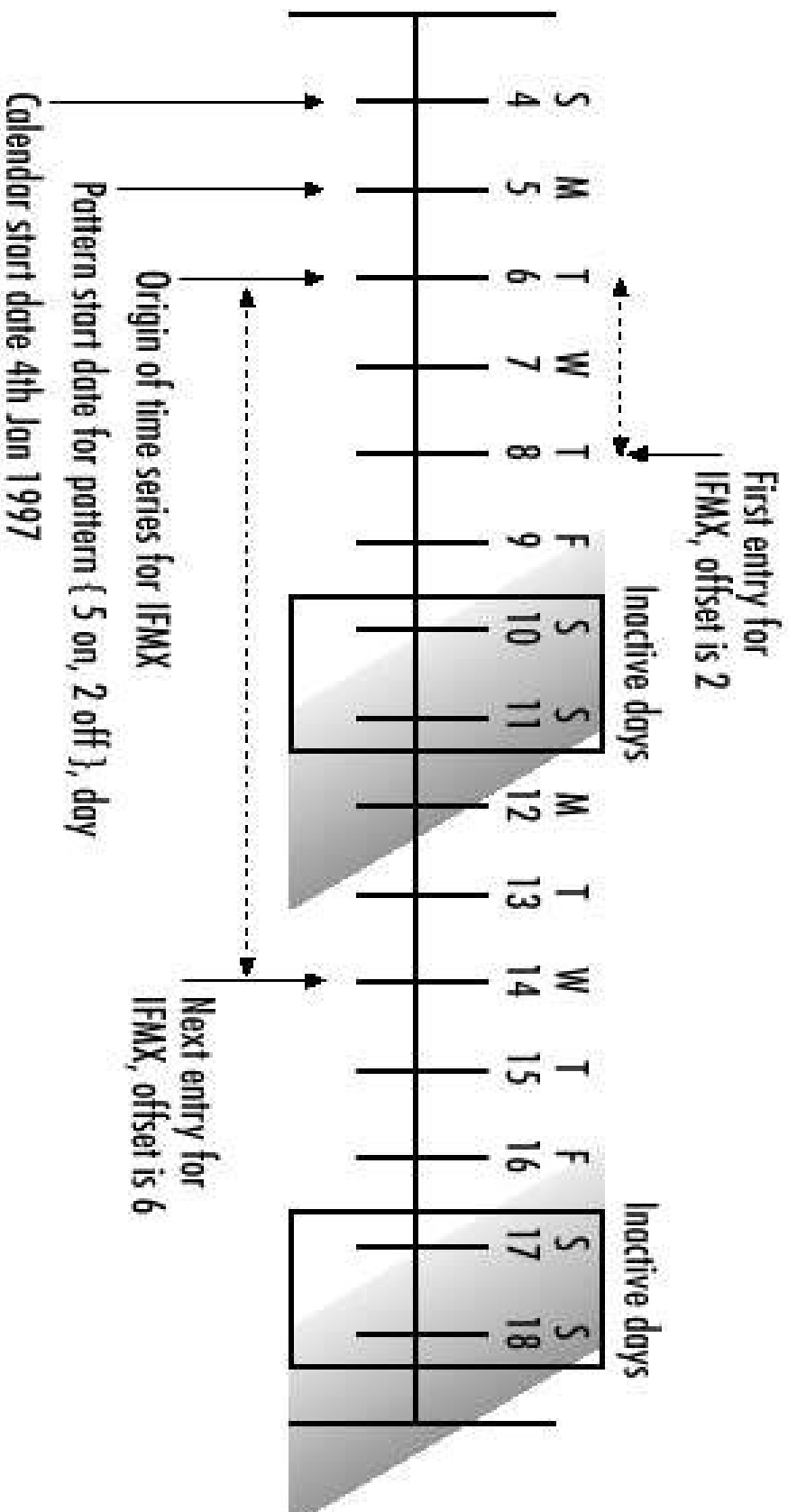


Columns and datatypes in each
element of the row subtype stock_bar

| Timestamp | High | Low | Final | Volume |
|---------------------|------|-----|-------|--------|
| 1997-12-12 00:00:00 | 9.7 | 9.1 | 9.4 | 226400 |

ADT Séries Temporelles

Illustration of a time-series calendar



Architecture des SGBD décisionnels et des infocentres

- Charge de travail
 - Requêtes complexes (nombreuses jointures + agrégats)
 - Très gros volumes de données

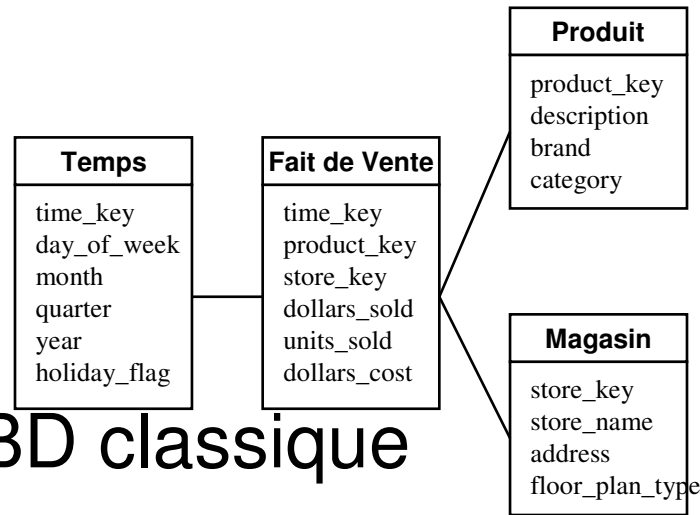
- Réponses
 - SGBD Parallèles sur machines parallèles (SMP, Cluster, ...)
 - RAID et SAN (Storage Area Networks)
 - Index Bitmaps, algorithmes
 - Placement sur disque (Partitionnement, ...)

Aspects Systèmes

- Stockage
 - Tables de fait
 - Append only (RID $\approx\approx$ Timekey)
 - Ligne de fait
 - clés et mesures de type de taille fixe
 - enregistrement de taille fixe (accès aléatoire)
- Jointure en Etoiles
 - Star Query
- Index Binaire (Bitmap)

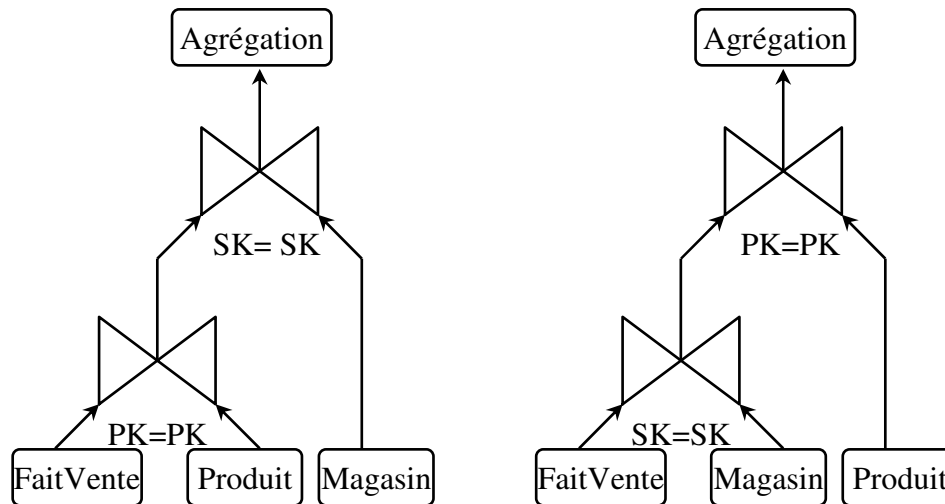
Jointure en Etoile (Star Join)

- Exemple d'entrepôt



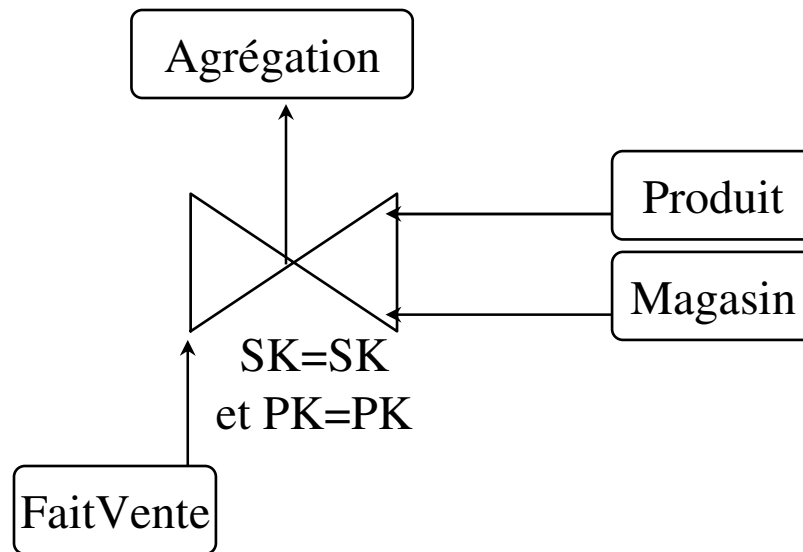
- Optimiseur d'un SGBD classique

- 2 stratégies



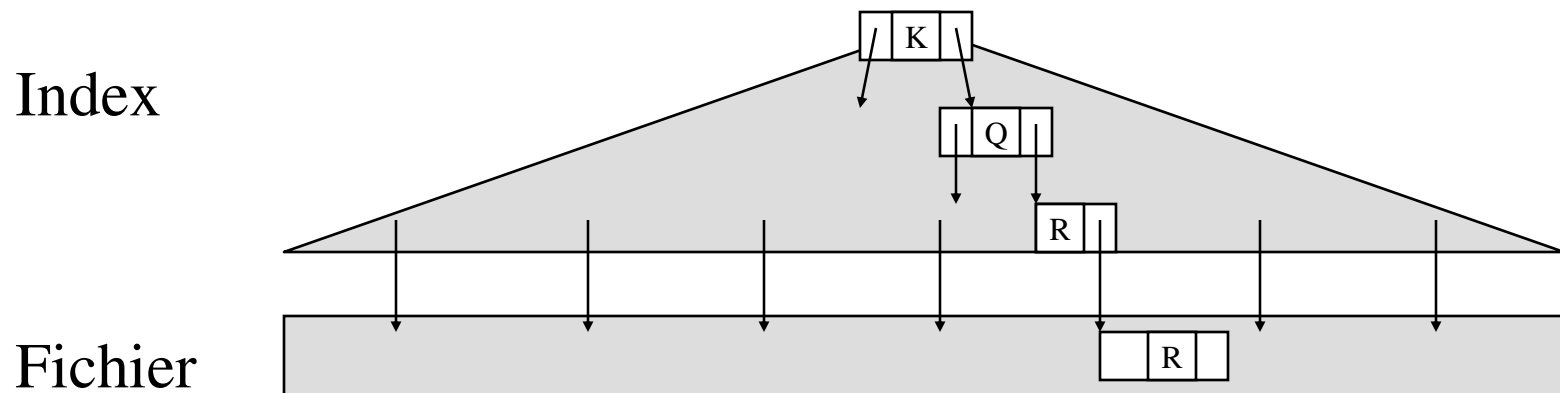
Jointure en Etoile (Star Join)

- Opérateur de jointure n-aire
 - Star Join
 - 1 source : la table de fait
 - 1 à N sources : les tables de dimension



Index Binaire (Bitmap)

- Rappel sur les index B-Tree
 - couple <Valeur Attribut, Pointeur d'enregistrement>



Index Binaire (Bitmap)

- Une chaîne de bits par valeur ou groupe de Valeur ou sur un prédicat

| | | |
|-------------------|------------------|----------|
| Index des Régions | (E) urope | 11101000 |
| | (A) sie | 00010100 |
| | (P) roche Orient | 00000011 |

Fichier



| | | |
|--------------------|----------|----------|
| Index des Couleurs | (r) ouge | 00011000 |
| | (b) leu | 00100101 |

- Opérations logiques OR, AND, NOT

| | | | | | |
|-----------------------|----------|----------|--|-------------------------|----------|
| | E | 11101000 | | E | 11101000 |
| | A | 00010100 | | Not (R OR B) | 11000010 |
| Asie OR Europe | | 11111100 | | Jaune AND Europe | 11000000 |

Optimisation des Index Binaires

- Compression
 - doit permettre les comparaisons
- Hiérarchisation multi-niveau
 - ratio entre niveaux (ex 1:32)

| | (E) urope | (A) sie | (P) roche Orient |
|---------------------------------------|---|---|---|
| Index R non hiérarchisé | 11101000 | 00010100 | 00000011 |
| Index R hiérarchisé (ratio 1:2) | N0 1 1 N1 1 1 1 0 N2 111010-- <i>12 bits stockés</i> | N0 1 1 N1 0 1 1 0 N2 --0101-- <i>10 bits stockés</i> | N0 0 1 N1 - - 0 1 N2 -----11 <i>6 bits stockés</i> |

Index Bitmap

- Indexation de la table de fait
 - sur les valeurs des tables de dimension associées
 - Cardinalité faible des attributs de la dimension
 - Tables de dimension à faible évolution
 - Exemple
 - `CREATE BITMAP INDEX sales_region_ix ON sales(region);`
- Jointure en étoile (Star query)
 - 2 étapes :
 - Sélection des faits et Sélection des dimensions
 - index bitmap pour les faits
 - Jointure en étoile
- Parallélisation
 - balayages // dans les étapes 1 et 2

Index Bitmap

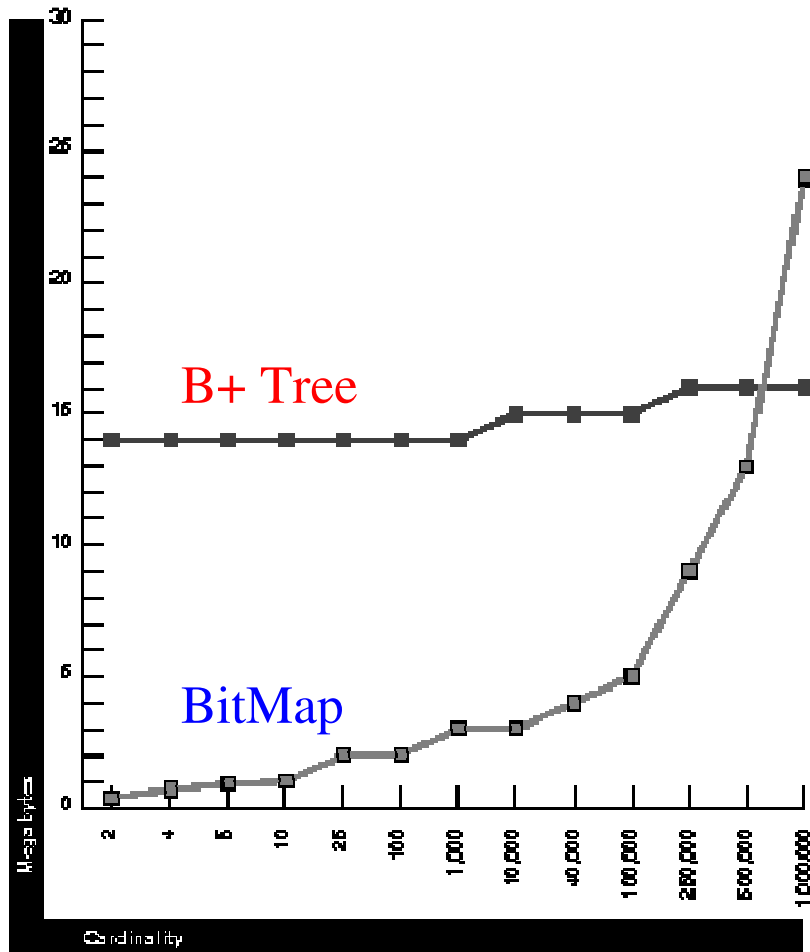
- Compression du bitmap
 - Informix Online XPS, Oracle 8, ...
- Indexation bitmap partielle
 - Informix Online XPS
- Indexation bitmap de colonnes virtuelles
 - Colonne calculée
 - $Vente.Profit = Vente.Prix_Vente - Produit.Prix_Conseillé$
- Bibliographie
 - "Star Query Processing in the Oracle7 Server," Oracle White Paper, March 1996.
 - O'Neil, P. E., Graefe, G., "Multiple-table joins through bitmapped join indexes," SIGMOD Record, Vol. 24, No. 3, September 1995.

Performance des Index BitMap

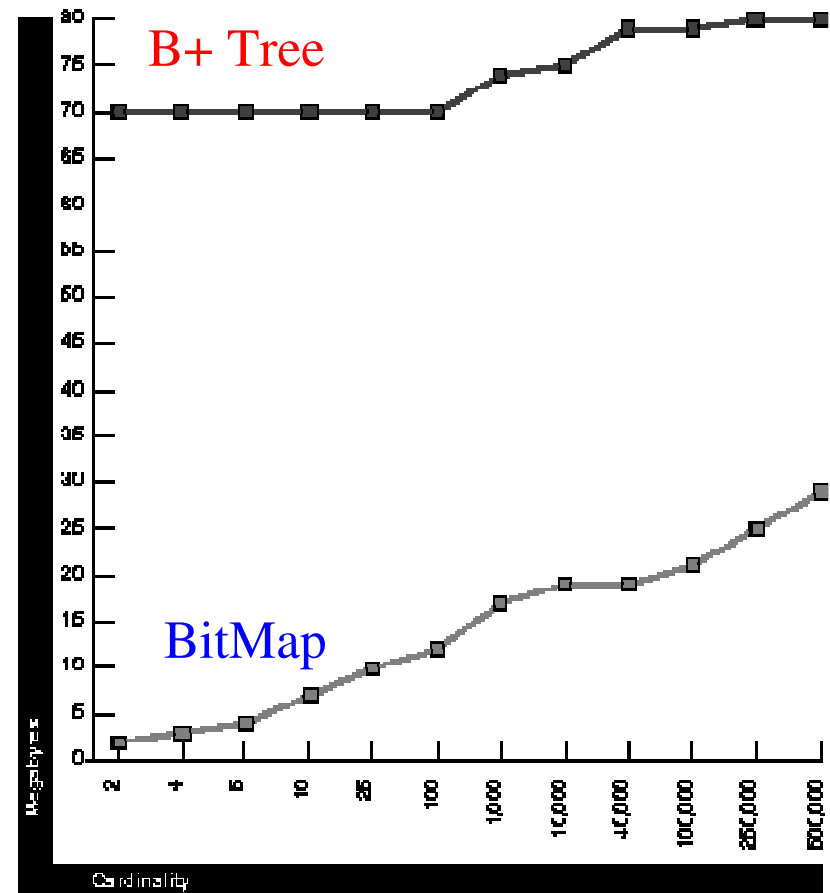
- Pour les tables Append Only
 - Éviter l'OLTP (modification des attributs)
- Domaine de cardinalité faible
 - faible nombre de valeurs différentes

Occupation des Index BitMap

1 000 000 tuples



5 000 000 tuples



Didier Donsez, 1997-2006, Les SGBDs décisionnels

■ B+ Tree Index Size
 ■ BitMap Index Size

■ B+ Tree Index Size
 ■ BitMap Index Size

Partitionnement (*Partitioning*) de la base

« *Pas de DW sans partitionnement.* »

- Partitionnement
 - Placement de tables multi disques / multi contrôleurs (RAID ou non)
- Partitionnement composite
 - Couplé avec d'autres chemins d'accès
 - Exemple de partitionnement composite classique
 - Range sur le temps et Hash sur une autre dimension
 - Remarque : Permet d'obsoleter les vieilles données (les partitions de plus d'1 an) et de bénéficier des performances pour les join sur la dimension H
- Si partitionner, comment maintenir les index

Autres techniques

- Cache (de calcul) d'agrégats
 - Conserve et réutilise les tables temporaires issues de calculs d'agrégat précédents

- Vue matérialisée (SNAPSHOT)
 - Recalcul incrémental de la vue
 - Pré-calcul d'agrégat

Chargement de l'entrepôt (i)

Extract Transform Load

- Outils ETL (Extract Transform Load)
 - Extracteurs à partir des bases de productions
 - Transformateur
 - Chargeur
- Tendances
 - Au départ, outils vendus par des tiers
 - Maintenant, intégration au noyau HOLAP

Chargement de l'entrepôt (ii)

Extract Transform Load

- Extraction depuis les bases de production
 - loader, external tables
 - snapshot log, refresh trigger
 - quelle fréquence d'extraction
- Transformation des données extraites,
 - transformer (table functions)
 - nettoyer (data cleansing)
 - valider (calcul de contraintes en parallèle, trigger).
- Chargement dans l'entrepôt
 - Muti table insert
 - insert d'un flot de tuple dans plusieurs tables
 - Upsert (Merge Update ou Insert)
 - Merge SQL:2003

Merge SQL:2003

```
MERGE INTO INVENTORY AS INV
USING (SELECT PARTNUM, DESCRIPTION, QUANTITY FROM SHIPMENT) AS SH
ON (INV.PARTNUM = SH.PARTNUM)
WHEN MATCHED THEN UPDATE
    SET QUANTITY = INV.QUANTITY + SH.QUANTITY
WHEN NOT MATCHED THEN INSERT
    (PARTNUM, DESCRIPTION, QUANTITY)
VALUES (SH.PARTNUM,SH.DESCRPTION,SH.QUANTITY)
```

RAID Redundant Array of Inexpensive Disks

■ Principe

- « Disque virtuel » : formé d'un groupe de plusieurs disques
 - Redondance, Performance (débit)

■ Niveaux de fonctionnement des RAID

■ Sans redondance

- Niveau 0 : Stripping

- répartition des blocs contigus d'un fichier entre les disques mais pas de redondance

■ Redondance

- Niveau 1 : Disques Miroirs

- Niveau 5 : Contrôle réparti sur les disques du Groupe

- Niveau 0/1 ou « 10 » : Stripping sur des disques miroirs

- Niveau 0/5 ou « 50 » : Stripping sur des groupes 5

- Niveau 6 : Contrôle redondant

- Support plus d'une panne par groupe

Balayage (scan) parallèle

■ RAID 0

- répartit les données sur plusieurs disques pour améliorer les performances.

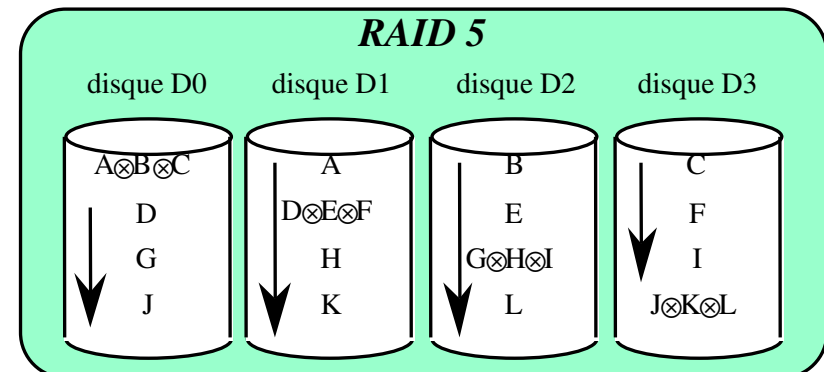
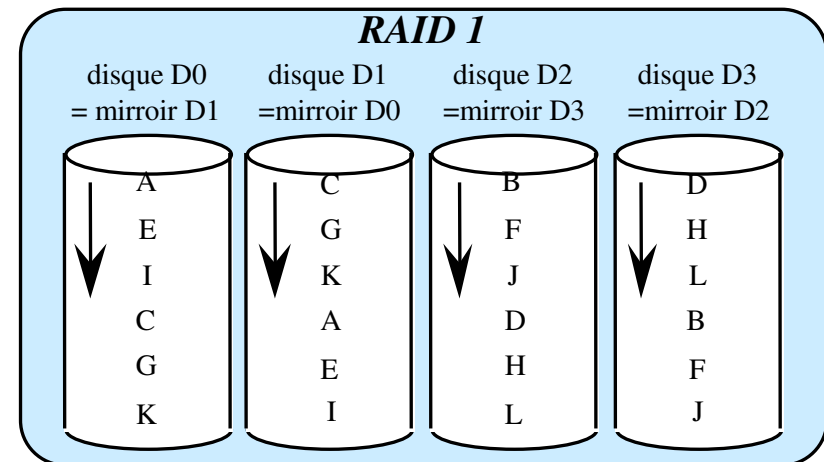
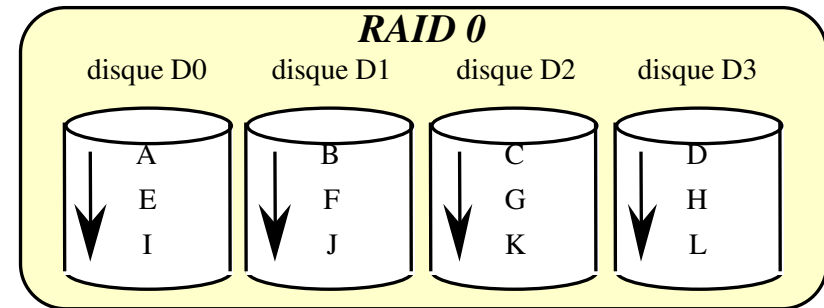
■ RAID 1

- effectue des copies miroirs de disques pour survivre aux pannes.

■ RAID 5

- utilise la correction d'erreur et la répartition des données pour fournir performance et sécurité de manière économique et efficace.
- La correction est basée sur la propriété du XOR (\otimes) :

$$(X \otimes Y) \otimes Y = X$$



SAN (Storage Area Network)

- Réseau de stockage et d'archivage haute performance
 - Clusters de machines partageant des clusters de RAID

Standardisation

- Trop souvent propriétaire
- Standardisation du Référentiel
 - Echec de AD/Cycle et Metadata Coalition
 - OIM Open Information Model
 - MS avec TI et Platinum Technology
 - OLE for OLAP
 - PLATO: Surcouche multidimensionnelle d 'SQL Server

Benchmarks (Banc de Performances)

- Mesurer les performances d'un système (matériel / logiciel) sous une charge de travail caractérisant une application type d'infocentre.
- Objectif
 - comparer les produits entre eux (avec d'acheter)
 - dimensionner son système en fonction de ses besoins
- Les benchmarks DW du TPC
 - <http://www.tpc.org/>
 - TPC-D : BD Décisionnelles (InfoCentre) *Obsolète*
 - TPC Benchmark H (TPC-H)
ad-hoc, decision support benchmark
 - TPC Benchmark R (TPC-R)
business reporting, decision support benchmark

TPC Benchmark TPC-H & TPC-R

■ Benchmarks

- Orienté sur des requêtes décisionnelles (*business report, decision support*)
- Définit une schéma de base
- Définit une charge de travail
 - suite de requêtes ad-hoc (orienté business) et de modifications concurrentes

■ Mesures

- TPC-H Composite Query-per-Hour Performance Metric (QphH@Size)
- TPC-R Composite Query-per-Hour Performance Metric (QphR@Size)

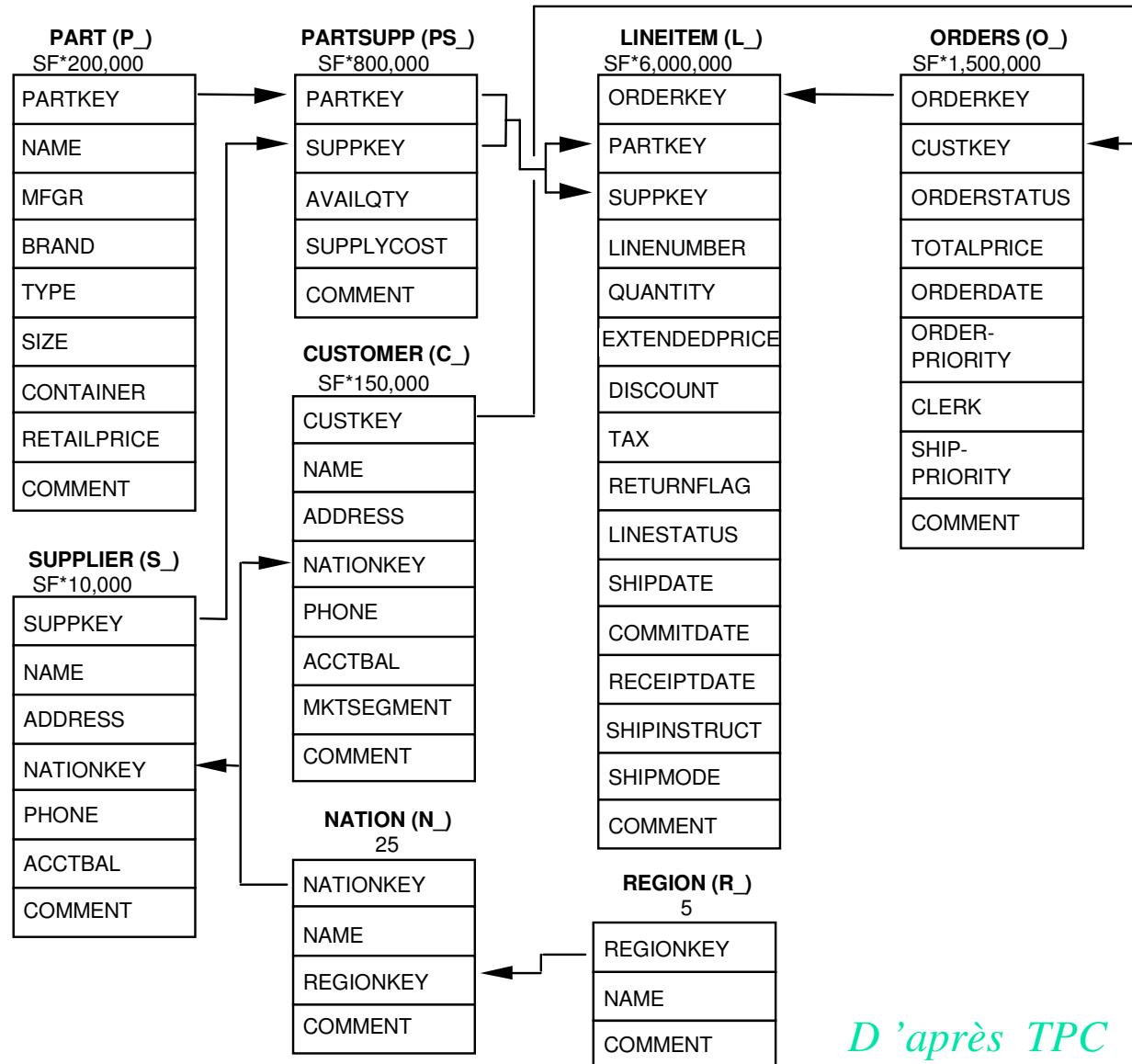
■ Taille des bases

- SF (Scaling Factor)=1 (~1Go), 10, 30, 100, 300, 1,000, 3,000, 10,000

■ Remarque

Modèle de données du TPC-D H R

- 8 tables



Suite de Requêtes du TPC-H & R

- Pricing Summary Report Query (Q1)
- Minimum Cost Supplier Query (Q2)
- Shipping Priority Query (Q3)
- Order Priority Checking Query (Q4)
- Local Supplier Volume Query (Q5)
- Forecasting Revenue Change Query (Q6)
- Volume Shipping Query (Q7)
- National Market Share Query (Q8)
- Product Type Profit Measure Query (Q9)
- Returned Item Reporting Query (Q10)
- Important Stock Identification Query (Q11)
- Shipping Modes and Order Priority Query (Q12)
- Customer Distribution Query (Q13)
- Promotion Effect Query (Q14)
- Top Supplier Query (Q15)
- Parts/Supplier Relationship Query (Q16)
- Small-Quantity-Order Revenue Query (Q17)
- Large Volume Customer Query (Q18)
- Discounted Revenue Query (Q19)
- Potential Part Promotion Query (Q20)
- Suppliers Who Kept Orders Waiting Query (Q21)
- Global Sales Opportunity Query (Q22)
- New Sales Refresh Function (RF1)
- Old Sales Refresh Function (RF2)

Exemple de résultats

**Exercice : quelle configuration
choisir pour un niveau de perf
=20000 QphH**

- TPC-H Results (01/2001)- Revision 1.X - 300GB Scale Factor

| Company | System | QphH | Price Perf. (\$/QphH) | Total Sys. Cost | Currency | Database Software | Operating System | CPU Type |
|-------------|-----------------------|-------|-----------------------|-----------------|-----------|----------------------|------------------------|---------------------|
| HP | HP ProLiant DL585 Cl | 35141 | 60 | 2106123 | US \$ | Oracle 10g RAC with | Red Hat Enterprise Lin | AMD Opteron Mod |
| FujSiemens | PRIMEPOWER 2500 | 34493 | 156 | 5380411 | Euros | Oracle Database 10g | Sun Solaris 9 | Fujitsu SPARC64 '1 |
| Fujitsu/ICL | PRIMEPOWER 2500 | 34493 | 141 | 4861985 | US \$ | Oracle Database 10g | Sun Solaris 9 | Fujitsu SPARC64 '1 |
| IBM | IBM eServer p5 570 wi | 26156 | 53 | 1397456 | US \$ | IBM DB2 UDB 8.2 | IBM AIX 5L V5.3 | IBM POWER5 1.9 |
| IBM | IBM eServer p655 with | 20221 | 69 | 1403446 | US \$ | IBM DB2 UDB 8.1 | IBM AIX 5L V5.2 | IBM Power 4 1700 |
| Legend Ltd. | Legend DeepComp 68 | 9951 | 1321 | 13145726 | China Yua | Oracle Database 10g | Red Hat Linux Advanc | Intel Itanium2 1.3G |
| Unisys | Unisys ES7000 Aries | 9853 | 68 | 669239 | US \$ | Oracle Database 10g | Microsoft Windows Se | Intel Itanium2 1500 |
| Unisys | Unisys ES7000 Aries | 5199 | 119 | 619308 | US \$ | Microsoft SQL Server | Microsoft Windows Se | Intel Itanium2 1.5G |
| HP | HP ProLiant DL760 G2 | 3385 | 59 | 199475 | US \$ | Microsoft SQL Server | Microsoft Windows Se | Intel Xeon MP 2.8C |
| Sun | SunFire V880 | 3247 | 56 | 183080 | US \$ | Sybase IQ 12.5 | Sun Solaris 10 | Sun UltraSPARC II |
| Sun | SunFire V490 | 2959 | 52 | 155152 | US \$ | Sybase IQ 12.5 | Sun Solaris 10 | Sun UltraSPARC I |

Offre du marché

- Microsoft SQL Server BI
 - <http://www.microsoft.com/sql/solutions/bi/default.mspx>
- IBM DB2 BI
 - <http://www-306.ibm.com/software/data/db2bi/>
- Oracle BI 10g
 - <http://www.oracle.com/technology/products/bi/index.html>
- *Informix (IBM maintenant)*

- Teradata (NCR)
 - http://www.ncr.com/en/solutions/data_warehousing/dataware.htm
- ...

Oracle 10g BI

<http://www.oracle.com/technology/products/bi/index.html>

- BI = Business Intelligence
- Composants BI
 - Discoverer
 - Requêtage et analyse pour environnements relationnels et multidimensionnels
 - Report Services
 - Fonction de restitution
 - Spreadsheet add-in
 - Visualisation des analyses OLAP pour Excel
 - BI Beans
 - Composants JavaBeans pour le développement d'applications décisionnelles personnalisées
 - Warehouse Builder
 - Atelier de développement pour paramétrer l'extraction et les chaînes de transformation (GUI + PL/SQL)
- Origine
 - Express, moteur multidimensionnel d'Iri Software

Microsoft SQL Server

- Composants
 - SQL Server 2000 Analysis Services
- TODO

IBM DB2 Universal Database Data Warehouse

- Composants
 - DB2 Alphablox
 - rapid assembly and broad deployment of integrated analytics
 - DB2 Universal Database Database Partitioning Feature
 - large clustered server support
 - DB2 Cube Views
 - OLAP acceleration
 - DB2 Intelligent Miner Modeling, Visualization, and Scoring
 - powerful data mining and integration of mining into OLTP applications
 - DB2 Office Connect Enterprise Web Edition
 - Spreadsheet integration for the desktop
 - DB2 Query Patroller
 - rule-based predictive query monitoring and control
 - DB2 Warehouse Manager Standard Edition
 - enhanced extract/transform/load services supporting multiple Agents
 - WebSphere Information Integrator Standard Edition
 - provide native connectors for accessing data from Oracle databases, Teradata databases, Sybase databases, and Microsoft SQL server databases

Bibliographie - Livre

- Rob Mattison, Data Warehousing -Strategies, Technologies and Technics, IEEE Computer Society 1996, ISBN 0-07-041034-8
- Michael J. Corey, Michael Abbey, Ian Abramson and Ben Taub, « Oracle8 Data Warehousing », Ed Mc Graw Hill, ISBN: 0-07-882511-3, 686 pages
- Chris Date, « Introduction aux Bases de Données », 7ème édition, Chapitre 21.

Bibliographie - Articles

- Morfonios, K., Konakas, S., Ioannidis, Y., and Kotsis, N. 2007. ROLAP implementations of the data cube. *ACM Comput. Surv.* 39, 4 (Nov. 2007), 12. DOI= <http://doi.acm.org/10.1145/1287620.1287623>
- CACM, «Industrial-Strength Data Warehousing » Vol. 41, No. 9 September, 1998
 - <http://www.acm.org/cacm/0998/0998toc.html>
- Computer , Issue on « Data warehouses and decision support systems », Volume: 34 Issue: 12 , Dec. 2001
 - Robert Armstrong: Seven Steps to Optimizing Data Warehouse Performance. *IEEE Computer* 34 (12): 76-79 (2001)
- SQL99 On-Line Analytical Processing (SQL/OLAP)
 - ISO/IEC 9075-1/Amd1:2001 <http://www.ansi.org>

Bibliographie - Articles

- Tips OLAP de SQL Server Magazine
 - <http://www.sqlmag.com>
- Intelligent Enterprise.
 - http://www.intelligententerprise.com/info_centers/data_warehousing/
- Parts de marché des outils OLAP
 - <http://www.olapreport.com/Market.htm>