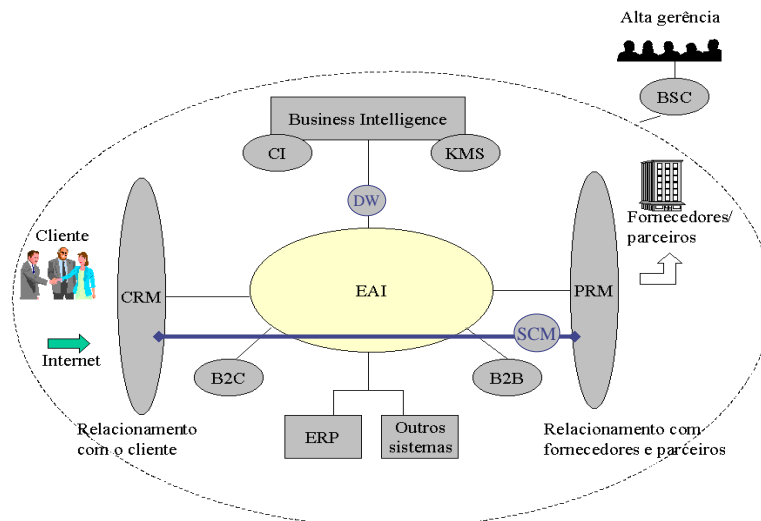


# Data Warehousing

## **Agenda**

- Suporte à Decisão
- Data Warehousing: conceitos básicos
- Arquitetura dos ambientes de DW
- Principais componentes das soluções
- Visão multidimensional
- Ferramentas e aplicações OLAP
- DW 2.0
- Outras tendências

## Novo ambiente de negócios



Fonte: BI: Business Intelligence - Modelagem e Tecnologia  
Carlos Barbieri

## A chave para o sucesso

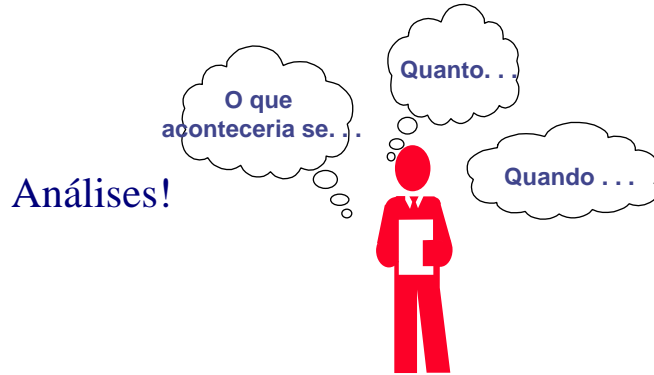
- ⇒ Alta qualidade
- ⇒ Baixo Custo
- ⇒ Flexibilidade
- ⇒ Agilidade

A INFORMAÇÃO está no centro  
de toda  
organização inteligente !

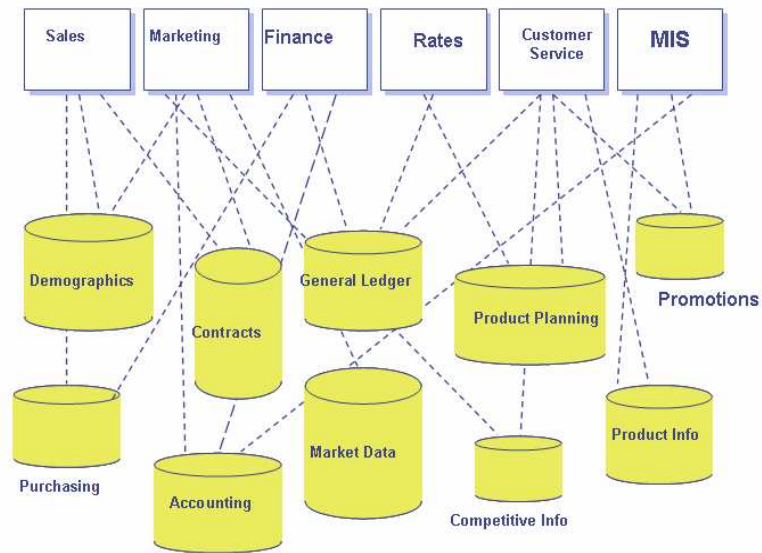
  
  
Organizações Inteligentes

## Suporte à Decisão

Produção e distribuição de **informação útil** para  
*gerentes, executivos e analistas do conhecimento*



## Realidade dos BD corporativos



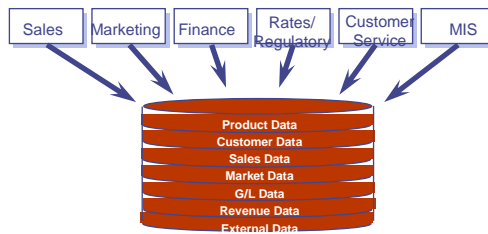
## O que queremos no suporte à tomada de decisão?

### Informação!

Mas, de que forma?

Com :

- ✓ Agilidade
- ✓ Flexibilidade
- ✓ Uma única versão da verdade



Marketing Finance Cliente Produto

DW  
OLAP  
Mining



CRM  
DBM  
Profissionais

## Ambiente de Aplicações

### Operacionais

- Dão suporte às funções associadas à *execução* do negócio da empresa:
  - ⇒ sistemas administrativos
  - ⇒ controle de estoque
  - ⇒ sistemas de expedição etc

### Suporte à Decisão

- Dão suporte às funções associadas à *concepção* do negócio da empresa

**Operação do Negócio**      **Planejamento do Negócio**

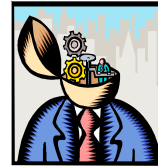
## Ambiente Operacional

- Tipo de processamento: OLTP
  - ⇒ Baseado em transações
  - ⇒ Voltado para velocidade e automação de funções "repetitivas"
  - ⇒ Mantém usualmente situação corrente
  - ⇒ Atualizações e consultas em grande número
  - ⇒ Trabalha com alto nível de detalhe



## **Ambiente de Suporte à Decisão (Analítico)**

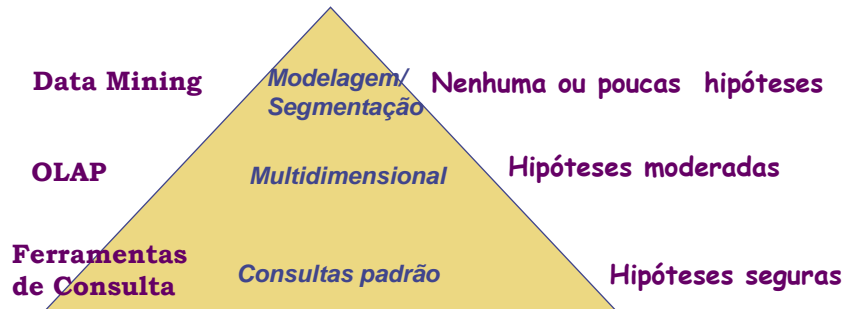
- Tipo de processamento: OLAP
  - ⇒ "Pequeno" número de consultas "variáveis"
  - ⇒ *Necessidade de ver o dado sob diferentes perspectivas: aplicações dinâmicas*
  - ⇒ Operações de agregação e cruzamentos
  - ⇒ *Atualização quase inexistente, apenas novas inserções*
  - ⇒ *Dados Históricos são relevantes*
  - ⇒ *Consistência é fundamental*



## **Requisitos de ambientes operacionais e analíticos**

- Sistemas Operacionais
  - ⇒ Tempo de Resposta
  - ⇒ Segurança
  - ⇒ Recuperação de Falhas
  - ⇒ Muitos usuários concorrentes
- Sistemas Analíticos ou "Informacionais"
  - ⇒ Flexibilidade, facilidade de navegação
  - ⇒ Consultas complexas, não antecipadas
  - ⇒ Gerenciamento de enormes volumes de dados
  - ⇒ Necessidade de examinar o dado em diferentes níveis de detalhe
  - ⇒ Necessidade de acesso a dados de fontes de dados diversas

## Ambientes analíticos

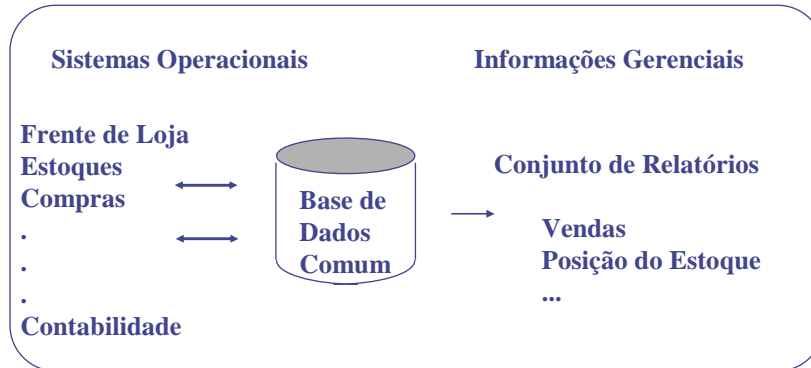


## Análises

- Qual o desempenho dos nossos representantes em cada região?
- Para cada produto, qual o total de vendas no último ano?
- Como tem variado o índice de participação de cada produto em nossas vendas (Product Share) ao longo dos três últimos anos?
- Existe alguma relação entre o desempenho dos representantes e sua faixa de salário?

## Suporte à Decisão “Tradicional”

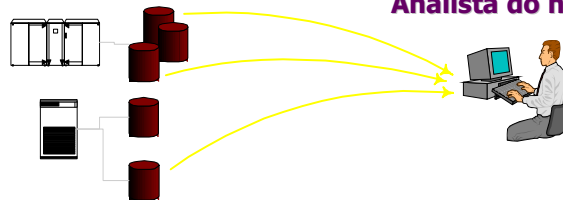
Forte Acoplamento entre Sistemas do Nível Operacional e Gerencial



## Suporte à Decisão “Tradicional”

Ambiente Operacional

Analista do negócio

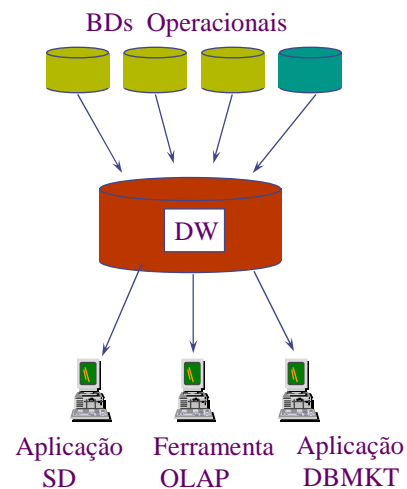


- Dados não estão adequados para suporte à decisão
- Sistemas de suporte à decisão desenvolvidos ad-hoc
- Longo tempo de espera



## Porque um ambiente de Data Warehouse?

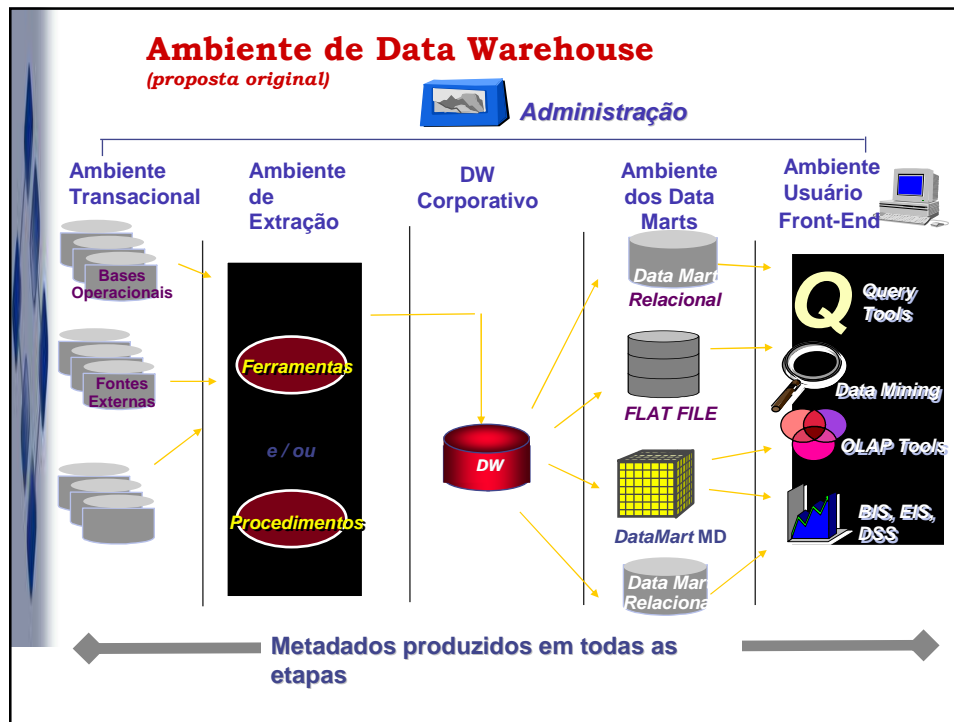
- ✓ Integrar dados de múltiplas fontes
- ✓ Facilitar o processo de análise sem impacto para o ambiente de dados operacionais
- ✓ Obter informação de qualidade
- ✓ Atender diferentes tipos de usuários finais
- ✓ Flexibilidade e agilidade para atender novas análises



## Data Warehouse - Definição

Processos, ferramentas e recursos para gerenciar e disponibilizar informações de negócios precisas e inteligíveis para que indivíduos possam tomar decisões efetivas. (IBM)

Um ambiente para adequadamente organizar, gerenciar e disponibilizar informações oriundas de fontes diversas, fornecendo uma visão única de parte ou de todo o negócio com o objetivo de dar suporte a operações analíticas.



- ### Estratégias de projeto de DW
- **Data Warehouses Corporativos**
    - de grande abrangência
    - complexos
    - alta probabilidade de insucesso
  - **Data Marts Setoriais**
    - Marketing, Financeiro, Administrativo, etc.
    - Projetos evolutivos
    - Enfoque inicial nos aspectos mais críticos
    - Aproveitamento da estrutura operacional disponível
    - Retorno mais rápido
    - Acúmulo de experiência : menor risco e menor custo

## Data Mart - Conceito

“Um subconjunto lógico do Data Warehouse, geralmente visto como um data warehouse setorial.”  
(Kimball)

Uma perspectiva **top-down** considera que um DW completo, centralizado deva ser desenvolvido antes que partes dele, sumarizadas, possam ser derivadas na forma de Data Marts.

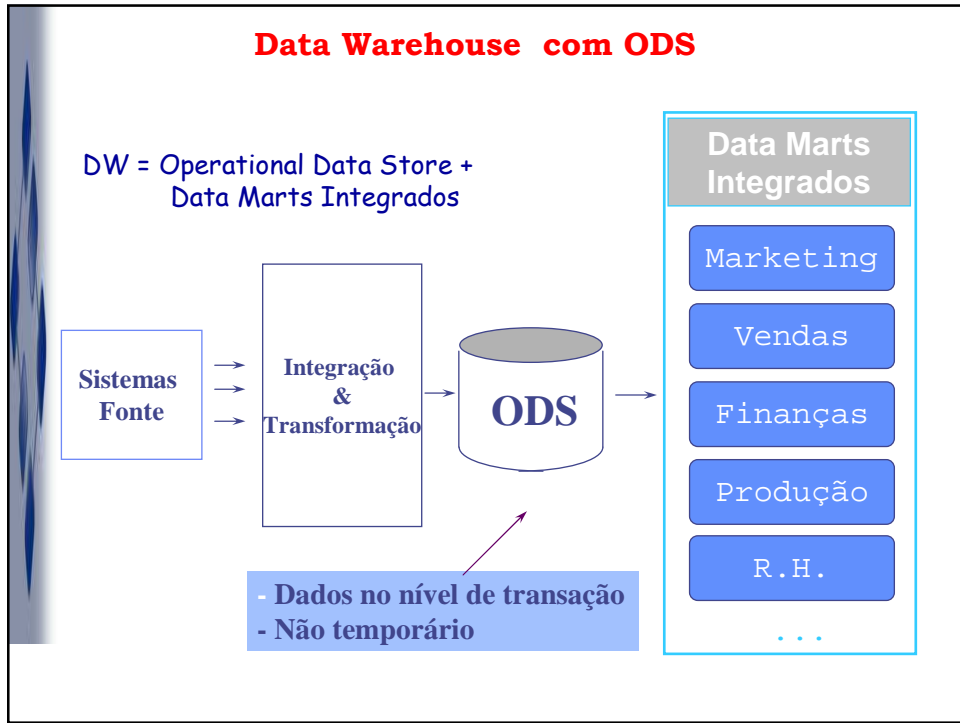
Uma perspectiva **bottom-up**, considera que um DW possa ser composto a partir de Data Marts previamente desenvolvidos.

## Abordagem corrente

Estratégia:

- ⇒ Visão Integrada
- ⇒ Dividir para conquistar
- ⇒ Errar pequeno
- Desenvolver Incrementalmente
  - ⇒ Planejamento Top-Down
  - ⇒ Desenvolvimento Bottom-Up, um DM de cada vez, resultados devem ser atingidos em pequenos ciclos (ex.: a cada 3 meses)
  - ⇒ Cada Data Mart deve ser encarado de forma evolutiva :
    - ✓ complexidade do modelo, volume de dados, investimentos
- Desafio
  - ⇒ Garantir a coerência entre os vários Data Marts

## Data Warehouse com ODS



## Etapas de um projeto de DW



## Transporte de Dados (Data Staging)

- Extração
  - Coleta de dados nos sistemas existentes
  - Operação demorada e complexa
  - Muitas vezes, desenvolvimento ad-hoc
- Transformação
  - Fundamental para clareza e integração
  - Recodificação de categorias: (m/f, male/female to M/F)
  - Alterações e uniformização de unidades de medida, nomes de campos, datas...
- Limpeza
  - Fundamental para qualidade da informação extraída
- Carga e Realimentação
  - Trade-off (muito frequente tem alto custo, pouco significa dados "velhos")

## Ferramentas de ETL

- Deve-se considerar “desenvolver versus comprar”:  
inicialmente, muitas empresas escrevem seus próprios programas;
- Produtos incluem geradores de código ou “transformadores proprietários”;
- Muitas ferramentas são voltadas para áreas específicas, embora com funcionalidades em comum;
- A maioria dos produtos são bastante novos, mas evoluíram rapidamente
- As ferramentas são geralmente muito caras, embora um novo modelo de preços esteja surgindo;
- Estas ferramentas são mais adequadas para ambientes complexos (múltiplas fontes e destinos, muitas transformações, muita limpeza necessária).

## Gerência de Metadados

- O grande desafio na construção e manutenção de um DW
  - Formatos de dados inconsistentes
  - Dados inexistentes ou inválidos
  - Diferentes níveis de agregação
  - Inconsistências semânticas
  - Qualidade de dados e janela de tempo
  - Acesso global (distribuído e replicado)
  - Administração e controle
- Integração do DW com outras ferramentas aumenta o problema

## Componentes Potenciais do Ambiente de DW

1. Repositório de Metadados
2. Ferramentas de Projeto CASE
3. Ferramentas de Extração, Transformação e Carga (ETL)
4. Ferramentas para Qualidade e Limpeza
5. Ferramentas para Replicação
6. Provedores de Interfaces de BD ODBC/OLE
7. Ferramentas de Gateway para BD Legados
8. Bancos de Dados Relacionais
9. (Bancos de Dados Não-Relacionais Legados)
10. Bancos de Dados Multidimensionais

## Componentes Potenciais do Ambiente de DW

11. Ferramentas OLAP
12. Ferramentas de Relatório e Consulta
13. Ferramentas de Data Mining
14. Cross-Platform Batch Schedulers
15. Ferramentas de Monitoramento e Controle
16. Pacotes de Aplicação para Data Warehouse

**Todos estes  
componentes  
manipulam/  
geram  
metadados**

## Diferentes tipos de metadados

- **Metadado Técnico e Administrativo**
  - altamente estruturado
  - informações com definições, transformações, gerência e operação
  - geralmente tratável através de uma ferramenta de repositório
- **Metadado de Negócio**
  - tanto não-estruturado quanto estruturado
  - mais difícil de ser tratado e integrado por uma ferramenta altamente estruturada tipo um repositório
  - necessidade de integrá-lo para o usuário final

### **Importância de um Repositório**

- ⇒ Repositório de Metadados
  - ✓ ferramentas que provêm armazenamento e funcionalidade de gerência e acesso a metadados
- ⇒ Visão global e integrada de metadados
- ⇒ Gerenciamento do ciclo de vida dos metadados
- ⇒ Integração com ferramentas de outros fornecedores

### **Ferramentas de Administração**

- Diferentes ferramentas com diferentes funcionalidades
- Mais produtos surgindo no Mercado
- Muitas voltadas para automação e scheduling
- Algumas oferecem monitoramento e tuning



## ERP e Suporte à Decisão

- Soluções ERP tem seu foco para o nível operacional dos ambientes corporativos
- Solução de suporte à decisão dos fornecedores de ERP apresentam vantagem na integração
- Fornecedores tradicionais em suporte à decisão em geral ofereciam produtos mais potentes e especializados do que as soluções apresentadas de suporte à decisão pelos fornecedores de ERP, mas estas evoluíram rapidamente
  - exemplo: BW

## Modelagem Dimensional para DW



- Visão Multidimensional
- Hierarquias
- Agregados
- Esquema Estrela e Snowflake



## **Modelagem para Data Warehouse**

- Requisitos distintos das aplicações do ambiente transacional:
  - ⇒ flexibilidade quanto às análises a suportar
  - ⇒ medidas a analisar precisam ser vistas sob diferentes perspectivas
- Abordagem utilizada:
  - ▶ **MODELAGEM MULTIDIMENSIONAL**

## **Visão multidimensional**

- Facilita o entendimento e visualização de problemas típicos de suporte à decisão
- Mais intuitiva para o processamento analítico
- Utilizada pelas ferramentas OLAP
- Qual a diferença da visão multidimensional para a visão tabular do ambiente relacional?

## Estrutura Relacional

### Volume de vendas para a concessionária CLYDE

MODEL	COLOR	SALES VOLUME
MINI VAN	BLUE	6
MINI VAN	RED	5
MINI VAN	WHITE	4
SPORTS COUPE	BLUE	3
SPORTS COUPE	RED	5
SPORTS COUPE	WHITE	5
SEDAN	BLUE	4
SEDAN	RED	3
SEDAN	WHITE	2

## Visão matricial ou multidimensional

### Volume de Vendas

<b>M</b> <b>O</b> <b>D</b> <b>E</b> <b>L</b>	Mini Van	6	5	4
	Coupe	3	5	5
	Sedan	4	3	2
		Blue	Red	White
		<b>COLOR</b>		

- ✓ Um array multidimensional tem um número fixo de dimensões e os valores são armazenados nas células
- ✓ Cada dimensão consiste de um número de elementos

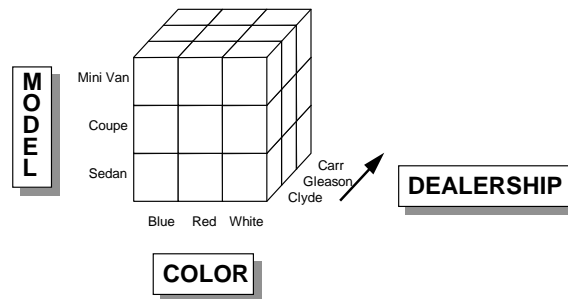
## Acrescentando mais uma coluna...

Volume de Vendas  
para todos os  
fornecedores

MODEL	COLOR	DEALERSHIP	VOLUME
MINI VAN	BLUE	CLYDE	6
MINI VAN	BLUE	GLEASON	6
MINI VAN	BLUE	CARR	2
MINI VAN	RED	CLYDE	3
MINI VAN	RED	GLEASON	5
MINI VAN	RED	CARR	5
MINI VAN	WHITE	CLYDE	2
MINI VAN	WHITE	GLEASON	4
MINI VAN	WHITE	CARR	3
SPORTS COUPE	BLUE	CLYDE	2
SPORTS COUPE	BLUE	GLEASON	3
SPORTS COUPE	BLUE	CARR	2
SPORTS COUPE	RED	CLYDE	7
SPORTS COUPE	RED	GLEASON	5
SPORTS COUPE	RED	CARR	2
SPORTS COUPE	WHITE	CLYDE	4
SPORTS COUPE	WHITE	GLEASON	5
SPORTS COUPE	WHITE	CARR	1
SEDAN	BLUE	CLYDE	6
SEDAN	BLUE	GLEASON	4
SEDAN	BLUE	CARR	2
SEDAN	RED	CLYDE	1
SEDAN	RED	GLEASON	3
SEDAN	RED	CARR	4
SEDAN	WHITE	CLYDE	2
SEDAN	WHITE	GLEASON	2
SEDAN	WHITE	CARR	3

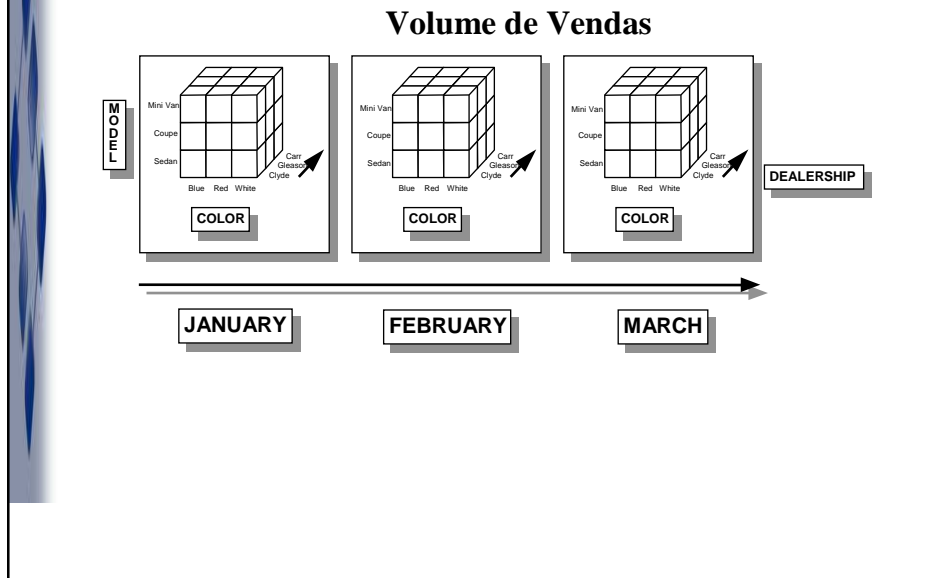
## Visão multidimensional

Volume de Vendas

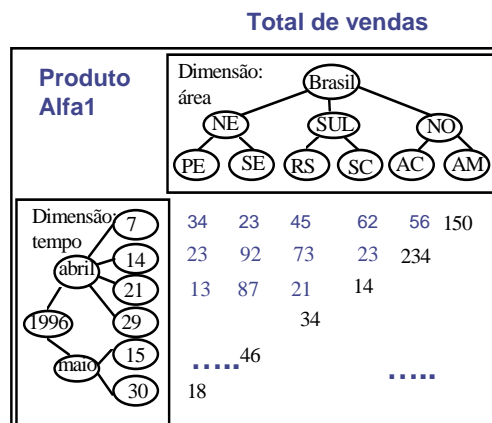


- O cubo é, de fato, apenas uma metáfora visual.
- É uma representação intuitiva do evento porque todas as dimensões coexistem para todo ponto no cubo e são independentes umas das outras.

## Acrescentando Dimensões - Hiper-cubos

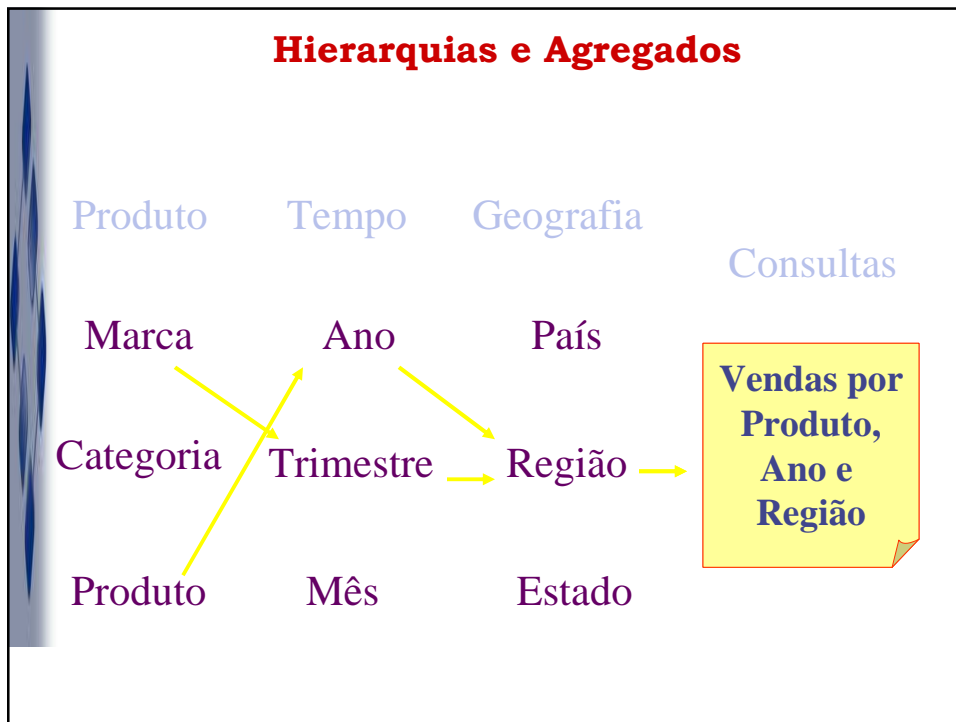


## Níveis nas dimensões ou Hierarquias

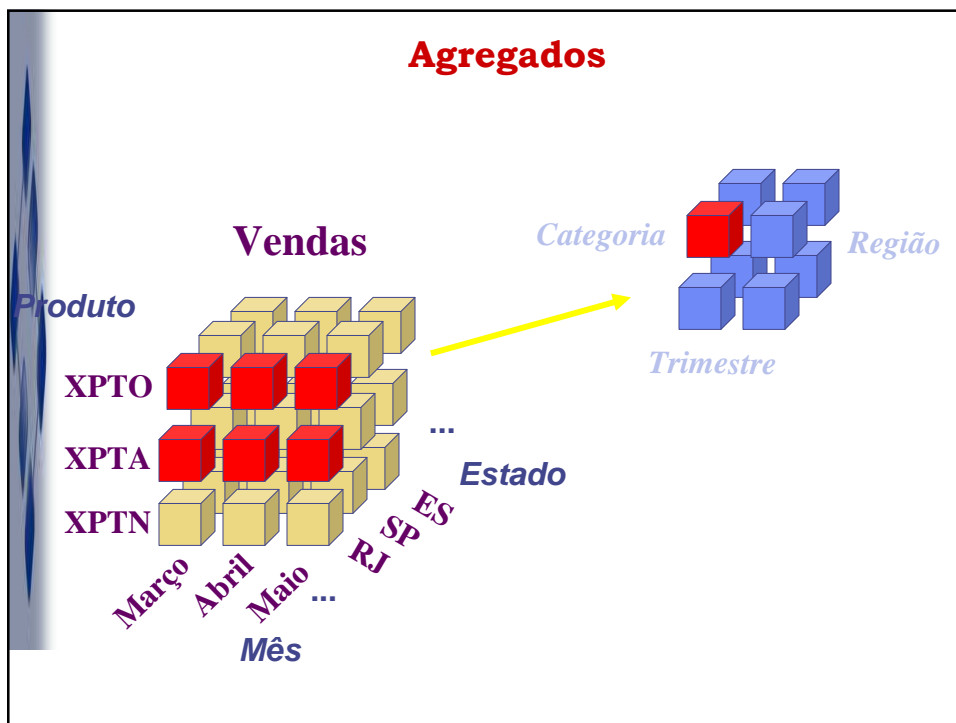


- Hierarquias são a base das agregações

## Hierarquias e Agregados

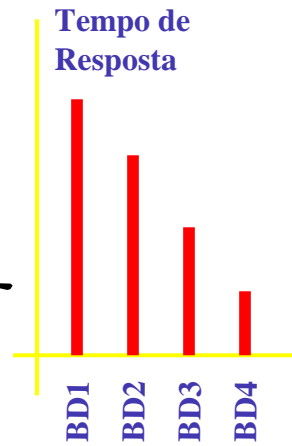


## Agregados



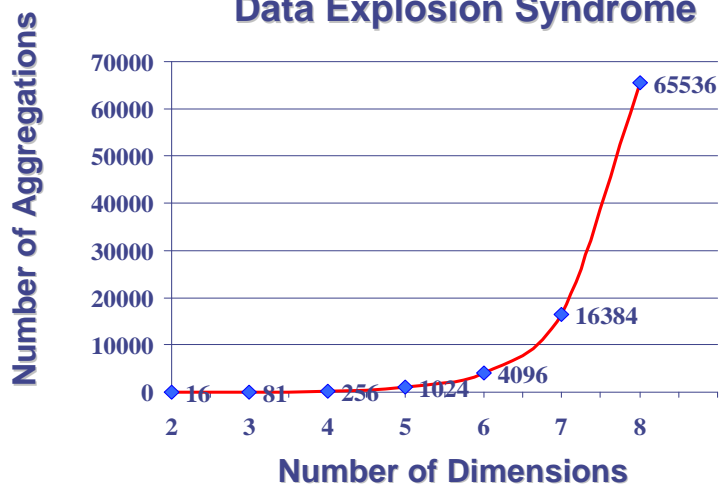
## Problemas

Calcular os agregados no momento da recuperação ou armazená-los?



## A Síndrome da Explosão no Volume de Dados

### Data Explosion Syndrome



(4 levels in each dimension)

## Armazenamento de Agregados

- Como é tratado o armazenamento de agregados pelas ferramentas OLAP?
  - ⇒ Geralmente existe a opção de armazenar parte dos agregados
  - ⇒ Tratamento para dados esparsos

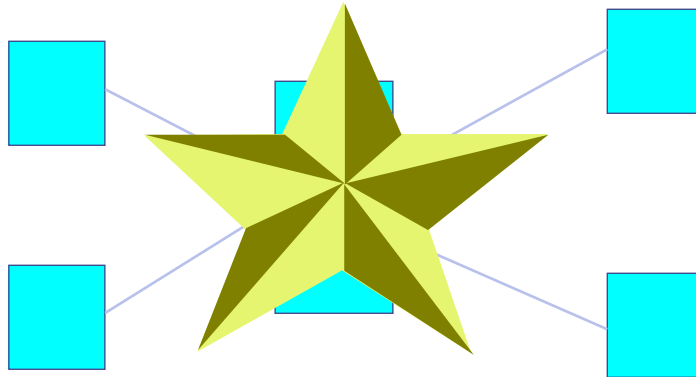
## Modelagem Multidimensional: Esquema Estrela

- Dominante no projeto de DW
- Características:
  - ⇒ Distingue melhor as dimensões dos fatos medidos
  - ⇒ Simplifica a visualização dimensional
  - ⇒ Na verdade é uma mistura de modelagem conceitual com modelagem lógica, pois já é bastante voltada para a abordagem relacional (a literatura fala sempre em *tabelas*)



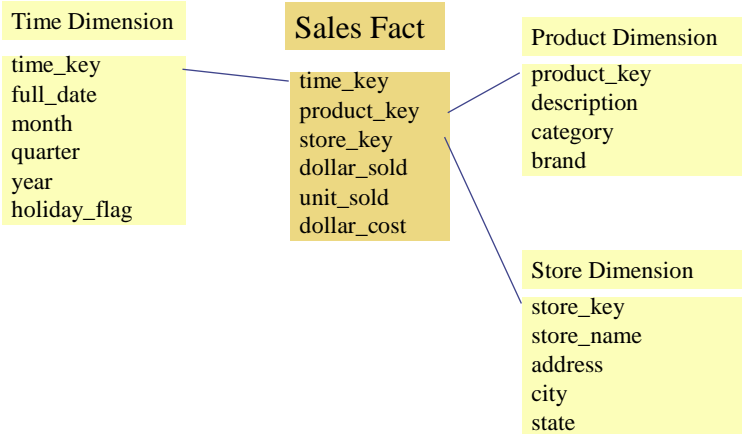
## Esquema Estrela

Uma tabela de fatos cercada de tabelas de dimensões



## Esquema Estrela (Star Schema)

*Um exemplo:*



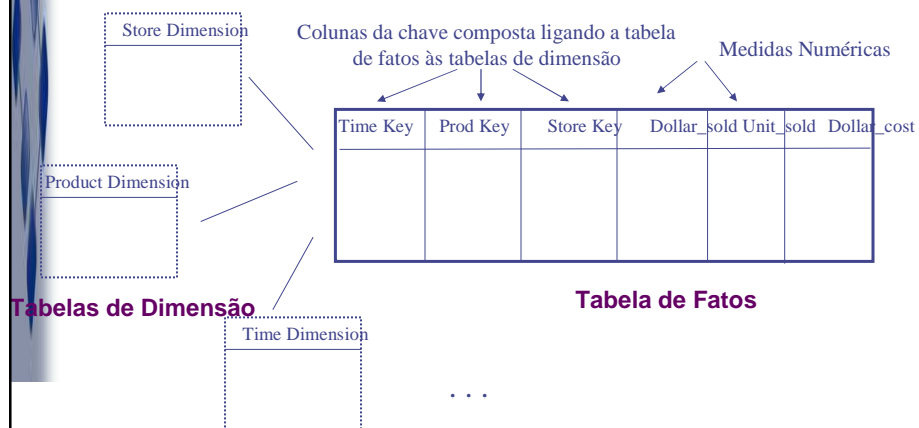
## Exemplo de tabela Tempo

date key	full date	day of week	day num in month	day num overall	day name	day abbre v	weekday flag	week num in year	week num overall	week begin date	week begin date key	month num overall	month name	month abbrev
1	1/1/96	1	1	1	Monday	Mon	y	1	1	1/1/96	1	1	1 January	Jan
2	1/2/96	2	2	2	Tuesday	Tue	y	1	1	1/1/96	1	1	1 January	Jan
3	1/3/96	3	3	3	Wednesd	Wed	y	1	1	1/1/96	1	1	1 January	Jan
4	1/4/96	4	4	4	Thursday	Thu	y	1	1	1/1/96	1	1	1 January	Jan
5	1/5/96	5	5	5	Friday	Fri	y	1	1	1/1/96	1	1	1 January	Jan
6	1/6/96	6	6	6	Saturday	Sat	n	1	1	1/1/96	1	1	1 January	Jan
7	1/7/96	7	7	7	Sunday	Sun	n	1	1	1/1/96	1	1	1 January	Jan
8	1/8/96	1	8	8	Monday	Mon	y	2	2	1/8/96	8	1	1 January	Jan
9	1/9/96	2	9	9	Tuesday	Tue	y	2	2	1/8/96	8	1	1 January	Jan

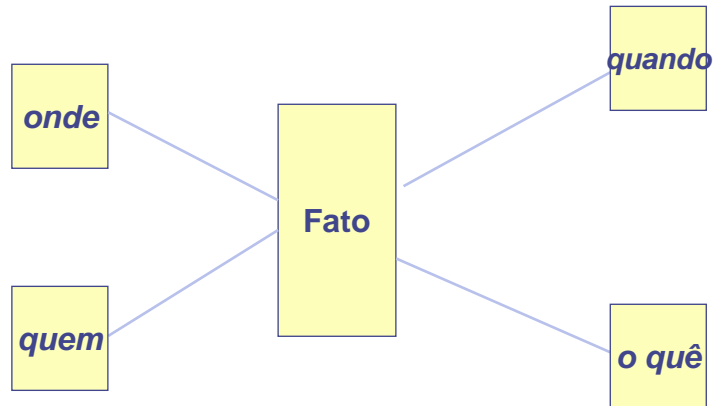
## Exemplo

"Vendas por categoria de produto sobre os últimos seis meses"

"Vendas por marca entre 1990 e 1995"

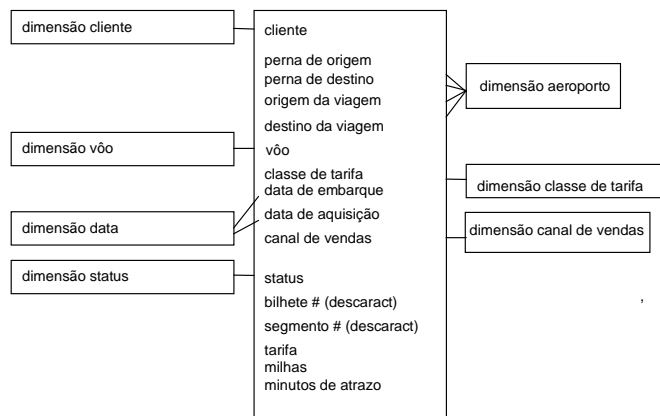


## Tipos de Dimensões mais comuns



## Mais um exemplo

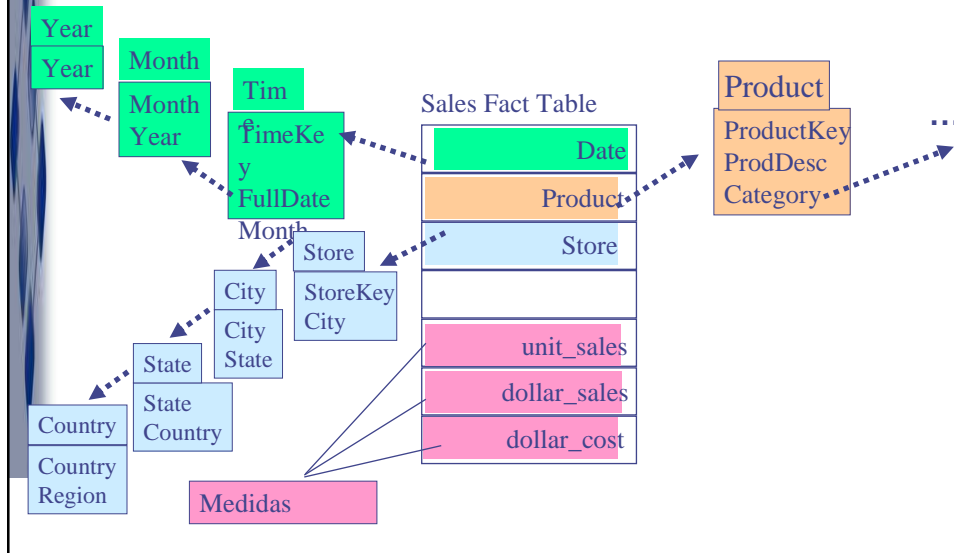
### Viagens aéreas



c

## Esquema Floco de Neve

Refinamento do esquema estrela onde as hierarquias das dimensões são representadas explicitamente através da normalização das tabelas de dimensões

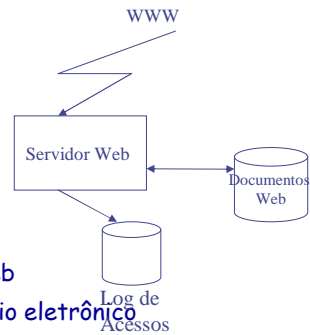


## Exploração dos dados de log do servidor Web

- Criação de um Data Mart para exploração de Weblog visando:

- Melhoria de desempenho
- Melhoria da navegação
  - ✓ Ajuste de ligações
  - ✓ Personalização
  - ✓ Web caching

- Melhoria do projeto de aplicações Web
- Classificação de Clientes para comércio eletrônico
- Identificação de locais para propaganda



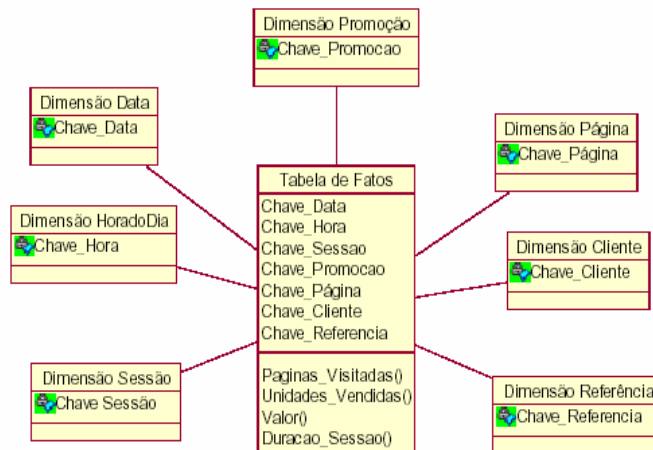
## Web Server Logs

### What's in a Typical Server Log?

<ip\_addr><base\_url> - <date><method><file><protocol><code><bytes><referrer><user\_agent>

```
203.30.5.145 www.acr-news.org - [01/Jun/1999:03:09:21 -0600] "GET /Calls/OWOM.html
HTTP/1.0" 200 3942 "http://www.lycos.com/cgi-
bin/pursuit?query=advertising+psychology&maxhits=20&cat=dir" "Mozilla/4.5 (en) (Win98; I) "
203.30.5.145 www.acr-news.org - [01/Jun/1999:03:09:23 -0600] "GET
/Calls/Images/earthani.gif HTTP/1.0" 200 10689 "http://www.acr-news.org/Calls/OWOM.html"
"Mozilla/4.5 (en) (Win98; I) "
203.30.5.145 www.acr-news.org - [01/Jun/1999:03:09:24 -0600] "GET /Calls/Images/line.gif
HTTP/1.0" 200 190 "http://www.acr-news.org/Calls/OWOM.html" "Mozilla/4.5 (en) (Win98; I) "
203.30.5.145 www.acr-news.org - [01/Jun/1999:03:09:25 -0600] "GET /Calls/Images/red.gif
HTTP/1.0" 200 164 "http://www.acr-news.org/Calls/OWOM.html" "Mozilla/4.5 (en) (Win98; I) "
203.252.234.33 www.acr-news.org - [01/Jun/1999:03:32:31 -0600] "GET / HTTP/1.0" 200 4980 **
"Mozilla/4.06 (en) (Win95; I) "
203.252.234.33 www.acr-news.org - [01/Jun/1999:03:32:35 -0600] "GET /Images/line.gif
HTTP/1.0" 200 190 "http://www.acr-news.org/" "Mozilla/4.06 (en) (Win95; I) "
203.252.234.33 www.acr-news.org - [01/Jun/1999:03:32:35 -0600] "GET /Images/red.gif
HTTP/1.0" 200 164 "http://www.acr-news.org/" "Mozilla/4.06 (en) (Win95; I) "
203.252.234.33 www.acr-news.org - [01/Jun/1999:03:32:35 -0600] "GET /Images/earthani.gif
HTTP/1.0" 200 10689 "http://www.acr-news.org/" "Mozilla/4.06 (en) (Win95; I) "
203.252.234.33 www.acr-news.org - [01/Jun/1999:03:33:11 -0600] "GET /CP.html HTTP/1.0" 200
3218 "http://www.acr-news.org/" "Mozilla/4.06 (en) (Win95; I) "
```

## Webhouse de Loja Virtual



## Cadeias de Valor Clickstream

- Muitas consultas necessitam dados além do clickstream
- Algumas dimensões são típicas de dados Web (Página, Sessão)
- DM de Transações de Vendas
  - ⇒ Uma linha para cada venda concretizada
  - ⇒ Dimensão CANAL acompanha o canal de vendas! (loja, web, telefone,...)
- DM de Contatos com o Cliente
  - ⇒ Uma linha para cada contato com o cliente (loja, web, telefone,...)
  - ⇒ Tabela de Fatos "Factless", acompanha um evento
- DM de Lucratividade na Web
  - ⇒ Uma linha para cada item vendido
  - ⇒ Pode responder muitas perguntas interessantes!

## OLAP

- On-line Analytical Processing
- Refere-se ao conjunto de processos para criação, gerência e manipulação de dados multidimensionais para análise e visualização pelo usuário em busca de uma maior compreensão destes dados.
- É usual a expressão "ferramenta" OLAP, referindo-se aos sistemas com estas funcionalidades e que são, juntamente com o SGBD, a base do ambiente de DW.

## Funcionalidade de Ferramentas OLAP

- Natural habilidade de definir uma estrutura de dados em termos de múltiplas dimensões
- Habilidade de apresentar a multidimensionalidade em telas através de uma configuração tridimensional de linhas, colunas e páginas
- Habilidade de visualizar hierarquias e navegar pelas dimensões
- Habilidade de definir fórmulas e associá-las a membros de dimensões



Facilidade para fazer análises, definindo agregações e cruzamentos, permitindo visualizar os dados através de múltiplos níveis de hierarquias e diferentes perspectivas.

## Operações usuais em ferramentas OLAP

- **Navegação pelas hierarquias e seus elementos:** permite seleccionar as perspectivas sob as quais se deseja visualizar as variáveis ou medidas
- **Cruzamentos:** permitem sumariar fatos segundo diferentes combinações das dimensões
- **Drill down:** navegação ao longo das dimensões na direcção de maior detalhe
- **Roll up (Drill up):** navegação ao longo das dimensões na direcção de menor detalhe
- **Rotação:** capacidade de inverter colunas e linhas navegação ao longo da dimensões na direcção de maior detalhe
- **Slice:** seleção definindo um subcubo  
(Ex: vendas onde cidade = 'Porto Alegre' e data = '1/15/90')
- **Cálculo e ranking** (Ex: top 3% das cidades por média de rendimentos)

## Alternativas para a multidimensionalidade

### ● MOLAP

- ⊖ MD Real
- ⊖ Armazena os dados em formato multidimensional
- ⊖ Não usa SQL como linguagem de acesso aos dados

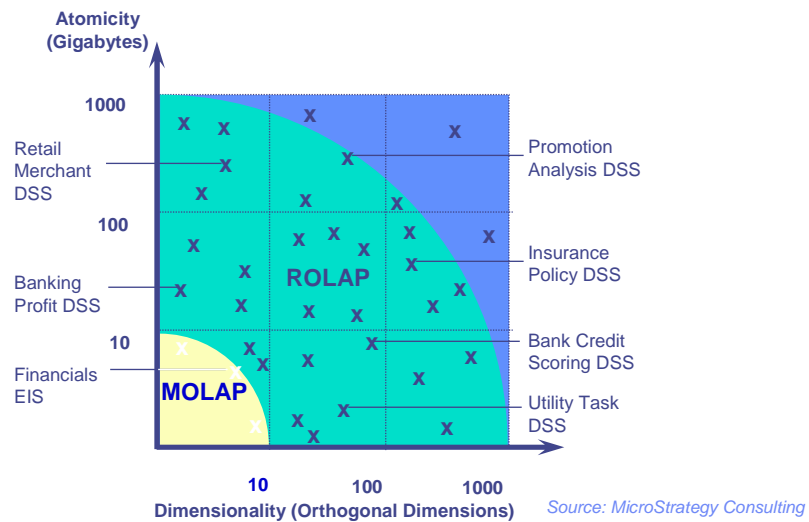
### ● ROLAP

- ⊖ MD Virtual
- ⊖ Armazena os dados em formato relacional
- ⊖ Comandos SQL são gerados para acesso aos dados

### ● HOLAP

- Híbrida
- Mais usual atualmente

## MOLAP vs. ROLAP





### **Alternativas para Ferramentas OLAP em ambientes de DW**

- Soluções proprietárias ou de um único fornecedor:
  - ⇒ menos flexíveis e abertas
  - ⇒ mais fáceis de implantar e manter
  - ⇒ geralmente menos potentes que uma solução mista
- Soluções envolvendo múltiplos fornecedores:
  - ⇒ maior funcionalidade como resultado final, usa-se o melhor para cada necessidade
  - ⇒ maior complexidade e fragilidade por causa de incompatibilidades entre os sistemas
  - ⇒ maior necessidade de tuning acurado

### **Riscos em projetos de DW**

- Descasamento dos objetivos da organização
- Qualidade dos dados e dos metadados desconhecida
- Falta de capacidade técnica para lidar com outra abordagem ou tecnologias
- Falta de software de apoio, questões com fornecedores
- Patrocinador pouco comprometido ou não existente
- Questões políticas ou culturais
- Falta de suporte por parte do usuário, expectativas pouco realistas
- Projeto e arquitetura da solução inadequados
- Escopo amplo demais e mudanças de requisitos

## Estratégia de Projetos de DW

- DW deve ser construído iterativamente - não como "um big bang"
- A primeira iteração é a mais crítica, por isso é importante restringir seu escopo
- Fundamental seguir trilhas de desenvolvimento paralelas:
  - ETL
  - Acesso
  - Metadados
  - Administração

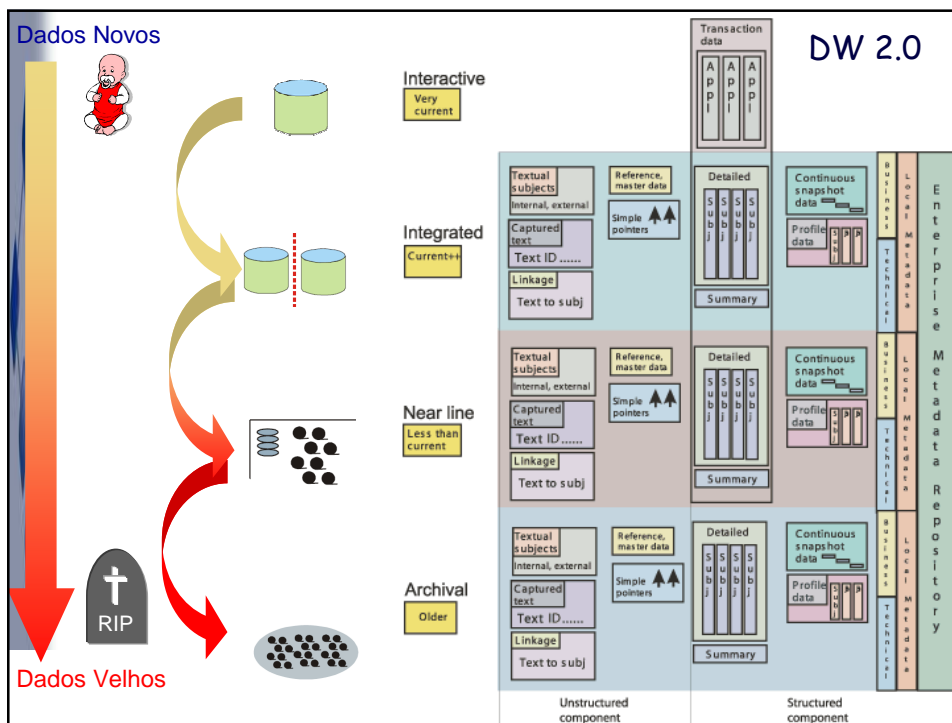
## Plano de Projeto para o DW

- Mantém atividades de um projeto de sistemas tradicional, mas inclui:
  - Identificação e definição de fontes de dados
  - Integração de metadados
  - Definição e gerência de agregados
  - Desenvolvimento de estratégias de carga
  - Seleção de produtos e foco em interoperabilidade
  - Preparar especificações de extração e transformação

## Bill Inmon e o DW 2.0..



- DW 2.0 é uma marca registrada
- Empresas que desenvolvem um DW 2.0 ganham um certificado
- Ênfase:
  - Integração de dados não estruturados
  - Metadados fortemente acoplados ao ambiente do DW
  - Diferentes níveis de dados, desde o mais corrente até o menos utilizado



## Tendências

- SGBDs com maiores facilidades para DW
  - ⇒ Estruturas de armazenamento e acesso especializadas
  - ⇒ Extensões ao SQL
  - ⇒ Maior exploração de paralelismo
- OLAP Híbrido (HOLAP) se tornando o mais utilizado
- DWs/OLAP associados a ambientes ERPs
- Maiores facilidades para integração de informações heterogêneas
  - ⇒ Tratamento de dados não estruturados
- Ambientes de múltiplas camadas
- Forte ênfase em gerência de metadados
  - ⇒ Uso de Repositórios
- XML como intermediário para troca de informações
- Uso de taxonomias e ontologias para conhecimento do domínio e como suporte à gerência de recursos

## Bibliografia

### Livros de referência (entre muitos outros):

- Barbieri, Carlos, *BI- Business Intelligence: Modelagem e Tecnologia*, Axcel Books, 2001
- Kimball, R. Reeves, L., Ross, M., Thornthwaite, W., *The Data Warehouse Lifecycle Toolkit*, John Wiley & Sons, 1998.
- W. H. Inmon, Claudia Imhoff, and Ryan Sousa, *Corporate Information Factory*, Wiley, 2000.
- Marco, D. *Building and Managing the Meta Data Repository*, Wiley, 2000.
- Carvalho, L.A., *Data Mining: A Mineração de Dados no Marketing, Medicina, Economia, Engenharia e Administração*, Ed. Érica, 2001.

### Revistas Técnicas

- Intelligent Enterprise, Data Management Review, B2B, SQL Magazine