

OLTP

- OLTP – On-line Transaction Process
 - Sistemas que suportam o dia-a-dia da empresa
 - Características:
 - Muitas transações
 - Transações pequenas
 - Registradas em tempo real
 - Pequena quantidade de dados envolvidas nas transações

OLAP

- OLAP – On-Line Analytical Processing
 - Permite uma análise de diferentes aspectos da atividade/desempenho da empresa
 - Características:
 - Poucas transações (quando comparados com OLTP)
 - Grandes volumes de dados
 - Permite apenas consultas e exploração de dados
 - Dados provêm de diferentes fontes de informação

OLTP x OLAP

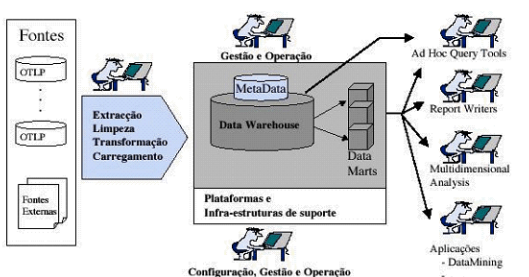
- Ambiente Operacional se diferencia do Analítico sob vários aspectos
 - Utilizadores são diferentes
 - Conteúdo diferenciado
 - Estruturas de dados são diferentes
 - A operação é diferente (forma e ritmo)
 - Gestão do sistema é diferente
 - Software é diferente

[Kimball]

Data Warehouse

Conjunto de dados baseados em assuntos, integrados, não-volátil, variável em relação ao tempo, de apoio as decisões gerenciais.

Arquitetura do Ambiente DW



Características

- Baseado em assuntos
- Integrados
- Não-Volátil
- Variável em relação ao tempo

Baseado em Assuntos

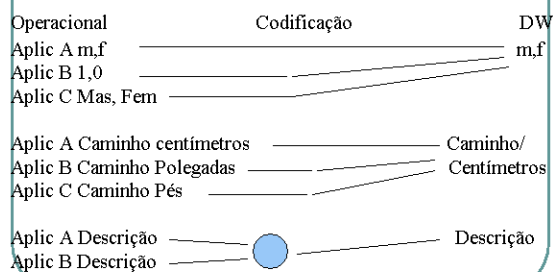
- OPERACIONAL
 - Automóvel
 - Vida
 - Saúde
 - Perdas
- DATA WAREHOUSE
 - Clientes
 - Apólices
 - Prêmio
 - Indenização

Integrado

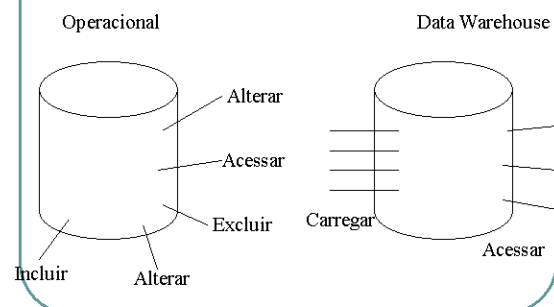
O processo de introdução dos dados no Data Warehouse é conduzido de forma que as muitas inconsistências das aplicações sejam desfeitas.

Codificações, unidades de medida, fontes múltiplas, chaves discrepantes, são algumas das inconsistências.

Integração



Não-Volátil



Variável em Relação ao Tempo

- A estrutura de chaves de um DW sempre contém algum elemento de tempo.
- 5 a 10 anos de dados em um DW é um horizonte de tempo normal.
- Cada carga no DW mostra com exatidão um momento dos sistemas operacionais

Popular um Data Warehouse

- Um grande tempo é aplicado na carga dos dados operacionais para o Data Warehouse. Entre os procedimentos realizados nesta atividade estão:
 - **Limpeza**
 - Erros no operacional implicam erros no data warehouse
 - **Integração**
 - Vários sistemas fonte implica integração
 - **Carga**
 - Cópia inicial para o Data Warehouse
 - **Atualização**
 - Cópia incremental

Granularidade

A granularidade diz respeito ao nível de detalhe ou resumo contido nas unidades de dados existentes no Data Warehouse. Quando mais detalhe, mais baixo o nível de granularidade. Quanto menos detalhe, mais alto nível de granularidade.

Granularidade

nível de detalhamento das unidades de dados
tempo: segundos
minutos
horas
dias, meses, anos, séculos, milênios,...

maior nível de detalhe
menor granularidade
maior volume de dados
maior tempo de processamento
maior complexidade de consultas

Granularidade(Exemplo)

Registro de vendas de uma base de supermercado(10 Lojas)

Mensal: Sumarização de Dados e cargas a cada 30 dias no banco de dados

Diária: Sumarização de vendas e carga diária no banco de dados

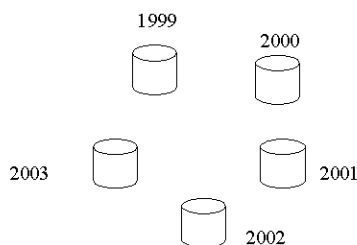
Consideração Importante:

95% das consultas é realizado sobre dados sumarizados

Níveis Duais de Granularidade

Para conseguir uma eficiência de armazenamento e acesso aos dados e ao mesmo tempo a possibilidade de analisar dados em maior detalhe, usa-se diferentes níveis de granularidade.

Particionamento dos dados



Metadados

- São dados sobre os dados ou sobre a estrutura dos dados
 - Funcionam também como documentação
 - Para cada atributo mantido no DW há uma entrada do dicionário de dados
 - Os dados são processados, atualizados e consultados partindo dos metadados
 - Usuários ficam conhecendo a estrutura e o significado dos dados

Camadas de Metadados

- Operacional
 - Definem a estrutura dos dados operacionais
- Data Warehouse
 - Orientados por assunto
 - Informam como os dados do DW foram calculados e como devem ser interpretados
- Usuário
 - Organizam os metadados no DW com base em conceitos familiares aos usuários

Metadados

- Os metadados englobam o DW e mantêm as informações sobre o que e onde estão os dados.
[Inmon]
- Os metadados devem relatar
 - A estrutura dos dados segundo a visão do programador;
 - A estrutura dos dados segundo a visão dos analistas de SAD;
 - A fonte de dados que alimenta o DW;
 - A transformação sofrida pelos dados no momento de sua migração para o DW;
 - O modelo de dados;
 - O relacionamento entre o modelo de dados e o DW;
 - O histórico das extrações de dados;
