

Curso de Tecnologia em Produção Industrial
Disciplina: Tecnologia da Informação Aplicada à Produção
Professor: José Maurício S. Pinheiro

AULA 5: O Armazenamento de Dados nas Grandes Empresas

O aumento do volume de informações que circula nas empresas passou a ser crítico em qualquer corporação e com ele a importância de contar com ferramentas de armazenamento de dados cresceu. Atualmente, o desafio não é só dimensionar a necessidade de espaço, em função dos terabytes a serem preservados, mas gerenciar esse processo de forma eficiente para garantir acesso rápido e inteligente aos dados armazenados.

Assim, a importância dos sistemas de Armazenamento de Dados é hoje incontestável, refletindo-se essa importância numa evolução muito acentuada, quer na diversidade de soluções oferecida quer na inovação técnica, quer ainda no volume de negócios associados. Sistemas de Armazenamento de Dados tratam a transformação dos dados em conhecimento útil para o negócio transformando-os numa vantagem competitiva.

1. Sistema de Apoio à Decisão

Sistema de Apoio à Decisão é um sistema de informação interativo, flexível e adaptável, especialmente desenvolvido para apoiar a resolução de problemas de gestão não estruturados, melhorando o processo de tomada de decisão associado. Tal sistema utiliza os dados disponíveis, fornece uma interface com o usuário e pode adaptar-se às opções de quem decide. Os Sistemas de Apoio à Decisão requerem dados e a informação neles contida serve para apoiar os agentes de decisão na resolução de problemas não estruturados a fim de melhorar a tomada de decisão.

Um Sistema de Apoio à Decisão pode ainda utilizar modelos, sendo construído por um procedimento interativo (muitas vezes pelo usuário final), apoia todas as fases de tomada de decisão e pode incluir ainda componentes de conhecimento.

2. Armazém de Dados (Data Warehouse)

O Armazém de Dados de uma empresa é o repositório de dados único e integrado, formando a infraestrutura base para aplicações de software de informação da empresa, ou seja, o Armazém de Dados é o portão que separa o mundo do processamento operacional do mundo dos Sistemas de Apoio à Decisão (Fig. 1). O conceito surgiu da necessidade de integrar dados corporativos espalhados em diferentes máquinas e sistemas operacionais, para torná-los acessíveis a todos os usuários dos níveis decisórios.

Outro fator que contribuiu para o estabelecimento desse conceito foi a evolução da Tecnologia da Informação, particularmente os Sistemas de Apoio à Decisão (DSS).

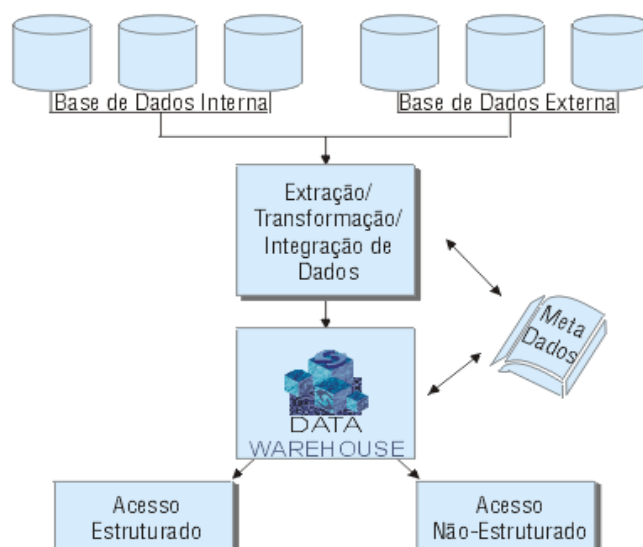


Figura 1 - Modelo de Data Warehouse

O Armazém de Dados compreende apenas os dados, não havendo qualquer referência a todos os processos envolvidos na sua gestão. Assim, passou a utilizar-se o termo Armazenamento de Dados para focar a definição nos processos e não nos dados. A Tabela 1 apresenta de forma detalhada as diferenças entre os dados operacionais e os dados de suporte a tomada de decisão.

Tabela 1- Comparação entre dados no nível operacional e dados no nível de suporte

Tópico ou Função	Nível Operacional	Nível Suporte a Decisão
Conteúdo dos Dados	Valores correntes	Dados históricos, consolidados e trabalhados
Organização dos Dados	Orientada à aplicação	Orientada à informação
Natureza dos Dados	Dinâmica Dados normalizados	Estática, Dados desnormalizados
Estrutura e Formato dos Dados	Complexos, desejáveis para computação operacional	Simple, desejáveis para análises de negócios
Possibilidade de Acesso	Alta	Moderada a baixa
Atualização	Continua	Periódica
Aplicação	Estruturada, processamento repetitivo	Não-estruturada processamento analítica
Tempo de Resposta	Entre 2 e 30 segundos	Segundos a minutos

2.1. Granulosidade

Granulosidade refere-se ao grau de detalhamento com que os dados são guardados no Armazém de Dados. Quanto maior é o detalhe, menor é a granulosidade. Quanto menor é o detalhe, maior é a granulosidade.

Os dados de informação são agregados, enquanto os dados operacionais são detalhados. Assim, num Armazém de Dados a granulosidade dos dados é uma característica fundamental a definir, dado que afeta simultaneamente o tamanho do Armazém de Dados e o potencial de informação que lhe está associado. Quanto menor a granulosidade dos dados, maior será o Armazém de Dados (aumentando a complexidade técnica e, consequentemente, os custos de implementação e manutenção) permitindo, no entanto, análises mais detalhadas por parte dos agentes de decisão.

2.2. Arquitetura do Armazenamento de Dados

A arquitetura do Armazenamento de Dados pode ser vista como constituída por três partes distintas (Fig. 2):

- A aquisição dos dados,
- Administração das estruturas de informação
- Ferramentas de apoio à decisão

A última parte (ferramentas de apoio à decisão) é composta por três grandes famílias de produtos: as soluções de análise multi-dimensional; as aplicações de utilização amigável e produção de relatórios; e as aplicações verticais e de mineração de dados. Um componente importante de um Armazém de Dados constitui-se dos Metadados¹ que têm um papel relevante na documentação e na recuperação de falhas.

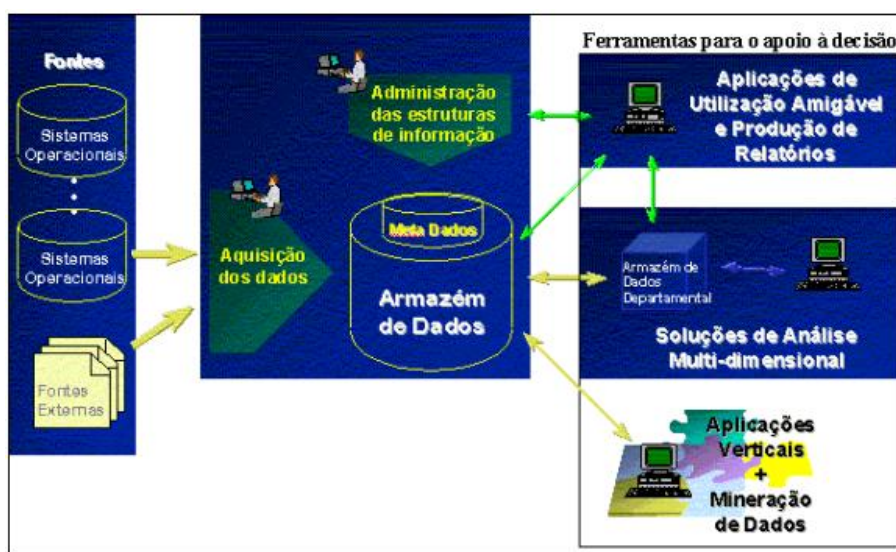


Figura 2- Modelo de Arquitetura de Armazém de Dados

¹ Metadados são dados sobre outros dados. Um item de um metadado pode dizer do que se trata aquele dado, geralmente uma informação inteligível por um computador. Os metadados facilitam o entendimento dos relacionamentos e a utilidade das informações dos dados.

2.3. Administração das Estruturas de Informação

A administração das estruturas de informação é feita recorrendo aos Sistemas de Gestão de Bases de Dados (SGBD). Tais sistemas evoluíram nos últimos anos, em parte como resposta à necessidade de gerir grandes quantidades de informação (Bases de Dados com vários tera bytes), permitindo baixos tempos de resposta a operações de leitura (Fig. 3).

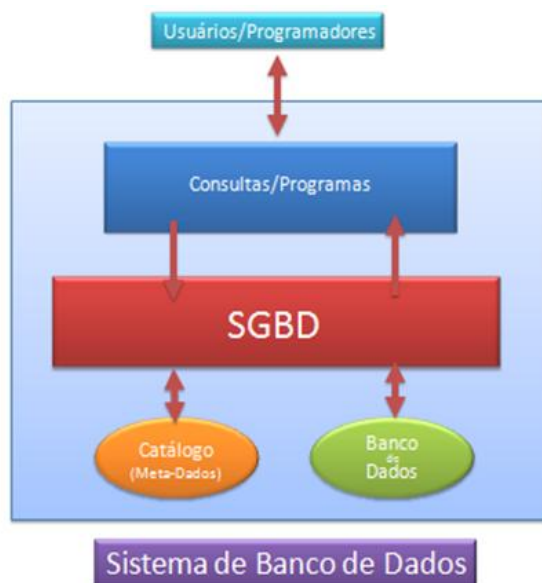


Figura 3 - Modelo de SGBD

Alguns dos Sistemas de Gestão de Bases de Dados utilizados para a gestão dos Armazéns de Dados são os sistemas de bancos de dados aos quais foram acrescentadas funcionalidades que lhes permitem gerir eficazmente grandes quantidades de informação.

3. Virtualização

A virtualização é um tema que tem sido atualmente destaque no mundo da tecnologia da informação (TI), apesar de não ser exatamente uma novidade.

Virtualização é a técnica que permite particionar um único sistema computacional em vários outros denominados de máquinas virtuais. Cada máquina virtual oferece um ambiente completo muito similar a uma máquina física. Com isso, cada máquina virtual pode ter seu próprio sistema operacional, aplicativos e serviços de rede (Internet). É possível ainda interconectar (virtualmente) cada uma dessas máquinas através de interfaces de redes, switches, roteadores e firewalls virtuais, além do uso já bastante difundido de VPN² (*Virtual Private Networks*).

Inicialmente, a virtualização pode auxiliar no trabalho em ambientes onde haja uma diversidade de plataformas de software (sistemas operacionais)

² Virtual Private Network ou Rede Privada Virtual, é uma rede privada construída sobre a infraestrutura de uma rede pública, normalmente a Internet.

sem ter um aumento no número de plataformas de hardware (máquinas físicas). Assim, cada aplicação pode executar em uma máquina virtual própria, possivelmente incluindo suas bibliotecas e seu sistema operacional que, por sua vez, executam em uma plataforma de hardware comum (Fig. 4). Assim, a virtualização proporciona um alto grau de portabilidade e de flexibilidade permitindo que várias aplicações de sistemas operacionais diferentes executem em um mesmo hardware.



Figura 4 - Virtualização de Sistemas Operacionais