

MODELOS DE ARMAZENAMENTO

N-ary Storage Model vs Decomposition Storage Model

Diego Gomes Tomé - MSc. Informática

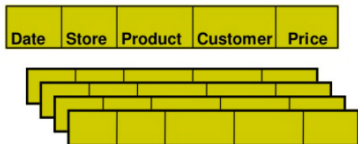
Orientador: Prof. Dr. Eduardo Almeida

October 11, 2016

Universidade Federal do Paraná

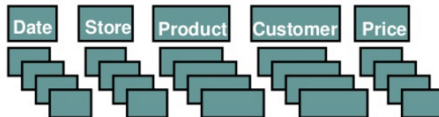
N-ary Storage Model (NSM)

- Orientado a linha



Decomposition Storage Model (DSM)

- Orientado a coluna



N-ARY STORAGE MODEL (NSM)

Em um modelo de armazenamento do tipo **NSM**:

- O SGBD armazena **todos os atributos** para uma **única tupla** de forma **contígua**, ou seja, da mesma forma que é representado no modelo lógico
- Ideal para cargas de trabalho de **OLTP**, onde transações tendem a operar apenas em uma entidade individual e cuja a carga de trabalho é formada por operações de **(INSERT, UPDATE, DELETE)**
- Processamento do tipo **"tuple-at-a-time"**

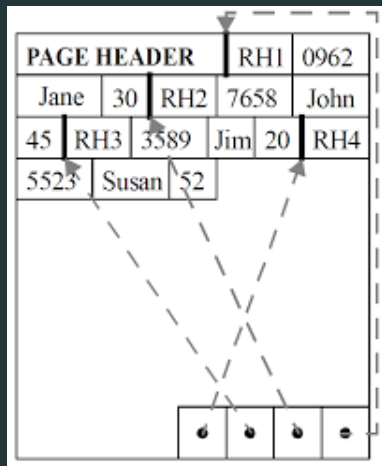
Vantagens:

- (+) Operações de (INSERT, UPDATE, DELETE) são mais rápidas
- (+) Bom para consultas que precisam recuperar toda a tupla

Desvantagens:

- (-) Não é bom para operações que fazem scan de uma grande parte da tabela para um subconjunto de seus atributos

PÁGINA DE MEMÓRIA NO MODELO NSM



DECOMPOSITION STORAGE MODEL(DSM)

Em um modelo de armazenamento do tipo **DSM**:

- O SGBD armazena para **cada atributo** todas as **tuplas** que o correspondem de forma **contígua** em um bloco de dados
- Também chamado de particionamento vertical
- Ideal para cargas de trabalho **OLAP**, onde as consultas são predominantemente de **leitura** e realizam **grandes scans** ao longo de um **subconjunto de atributos** da tabela.
- Processamento do tipo **"vector-at-a-time"**

HISTÓRICO DO MODELO DSM

1970s: Cantor DBMS

1980s: [DSM Proposal](#)

1990s: SybaseIQ (in-memory only)

2000s: Vertica, Vectorwise, MonetDB

2010s: “The Big Three”

Cloudera Impala, Amazon Redshift,
SAP HANA, MemSQL

Vantagens:

- (+) **Reduz** a quantidade de **desperdício** de trabalho, pois o SGBD lê somente os **dados necessários**
- (+) **Bom** para **compressão** de dados
- (+) Elimina valores nulos

Desvantagens:

- (-) **Lento** para consultas pontuais (**INSERT, UPDATE, DELETE**), por conta da separação dos atributos de uma entidade.
- (-) Utiliza até **4 vezes** mais espaço que o modelo NSM

Modificar um atributo

- NSM requer 2 escritas no disco, 1 para registro e 1 para índice
- DSM requer 3 escritas no disco, 2 para registro e 1 para índice

Inserir e/ou deletar um registro

- NSM requer 2 escritas no disco, 1 para registro e 1 para índice
- DSM requer 2 escritas no disco por atributo

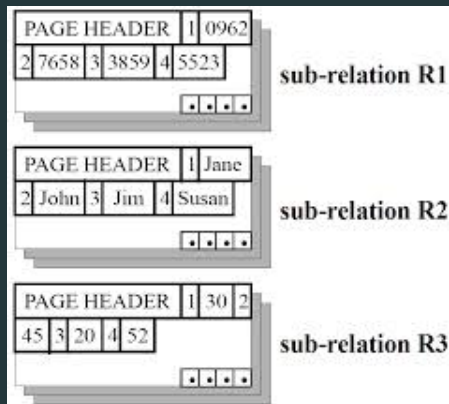
Chave Implícita

- Identificador da tupla corresponde a sua posição
- **Merge Join** para reconstrução de tupla

Chave Explícita

- Armazena em cada bloco o valor da tupla e seu respectivo id.
- (+) Não precisa manter a ordem
- (-) **Espaço extra** para as chaves de entrada

PÁGINA DE MEMÓRIA NO MODELO DSM



EXEMPLOS