



unioeste

Universidade Estadual do Oeste do Paraná

Bancos de Dados Paralelos

**Alunos: ALLAN ROGER BELLO
WILLIAN DIAS TAMAGI**



INTRODUÇÃO

- (SGBD) *Sistema Gerenciador de Banco de Dados* é um software que permite criar, manter e manipular bancos de dados para diversas aplicações.

INTRODUÇÃO

- ❑ Necessidade de bancos de dados distribuídos
- ❑ Os dados possam ser consultados como uma unidade

INTRODUÇÃO

- ❑ O processamento de grandes bases de dados é um problema complexo! Então, por que não utilizar técnicas de processamento paralelo

INTRODUÇÃO

- ❑ Crescimento do número de aplicações de suporte à decisão (DSS).
- ❑ Sistemas de Banco de Dados Paralelos (SBDP).

SBDPs / Sistema relacional

- ❑ O sucesso dos sistemas do SBDP é a adoção do sistema relacional de dados, onde as buscas relacionadas são criadas pensando no paralelismo e consistem de operações uniformes aplicadas a um conjunto uniforme de dados.

Principais características dos SBDPs:

- ❑ Velocidade linear
- ❑ Escalabilidade linear

Dificuldades para garantir essas propriedades do SBDPs

- ❑ Iniciação
- ❑ Interferência
- ❑ Inclinação

Arquiteturas de paralelismo em BD

Desafio:

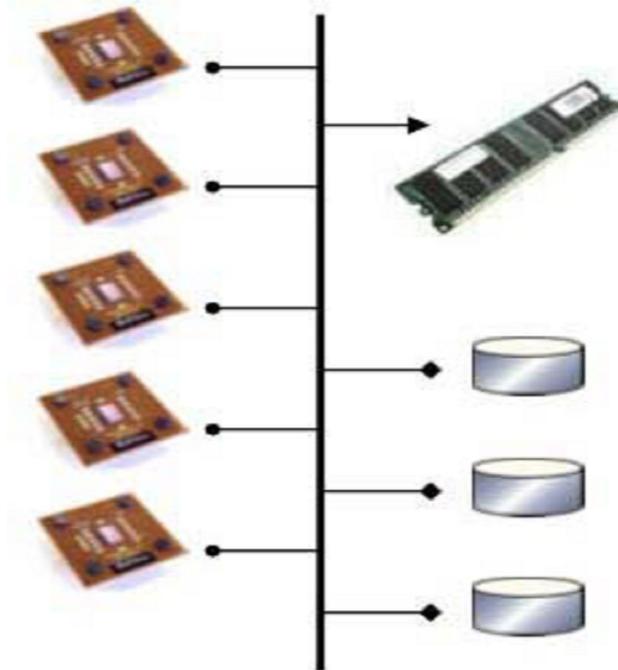
- ❑ Criar um processador muito rápido por meio de diversos processadores de alta velocidade.
- ❑ Criar uma memória muito grande com vasta área de troca de memórias.

Arquiteturas de paralelismo em BD

- ❑ Memória compartilhada
- ❑ Discos compartilhados
- ❑ Sem compartilhamento
- ❑ Hierárquico

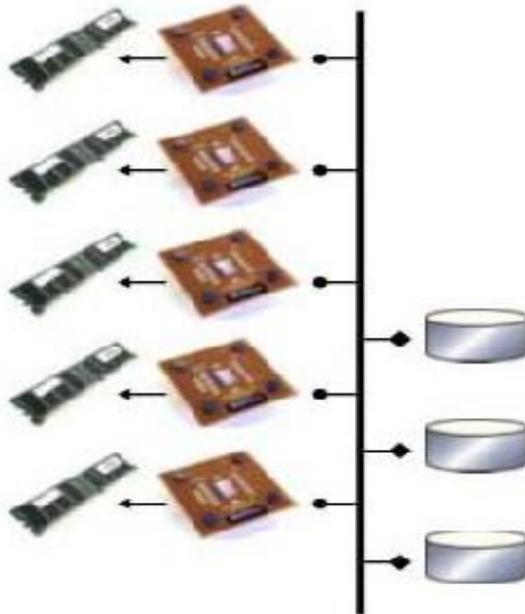
Memória compartilhada

- Vantagens: extrema eficiência na comunicação entre processadores
- Desvantagens: a arquitetura não é adequada ao uso de mais de 32 ou 64 processadores



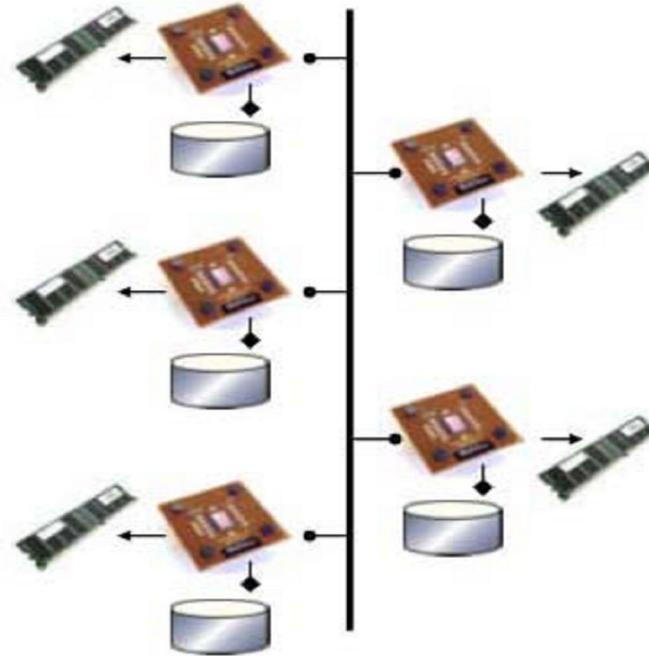
Discos compartilhados

- ❑ Vantagens: o acesso à memória não representa um gargalo; é um modo barato de aumentar a tolerância a falhas
- ❑ Desvantagem: é novamente o grau de crescimento



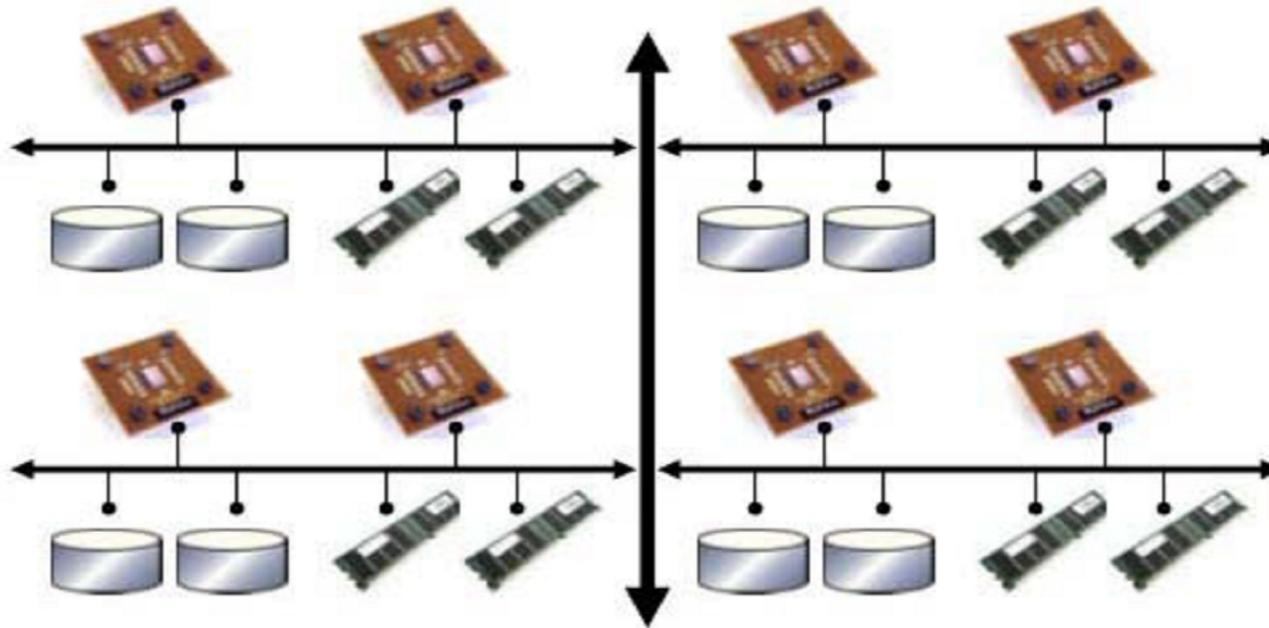
Sem Compartilhamento

- ❑ Vantagem: suporte a um grande número de processadores
- ❑ Desvantagem: comunicação entre processadores é o fator limitante, devido a necessidade



Hierárquico

- Este modelo combina características de várias arquiteturas anteriores, reduzindo a necessidade e complexidade da comunicação entre processadores.



Implementação de paralelismo em BD

Paralelismo na entrada e saída de dados (E/S)

Existem três tipos de particionamento de dados para obtenção de paralelismo de E/S:

- Horizontal
- Vertical
- Misto



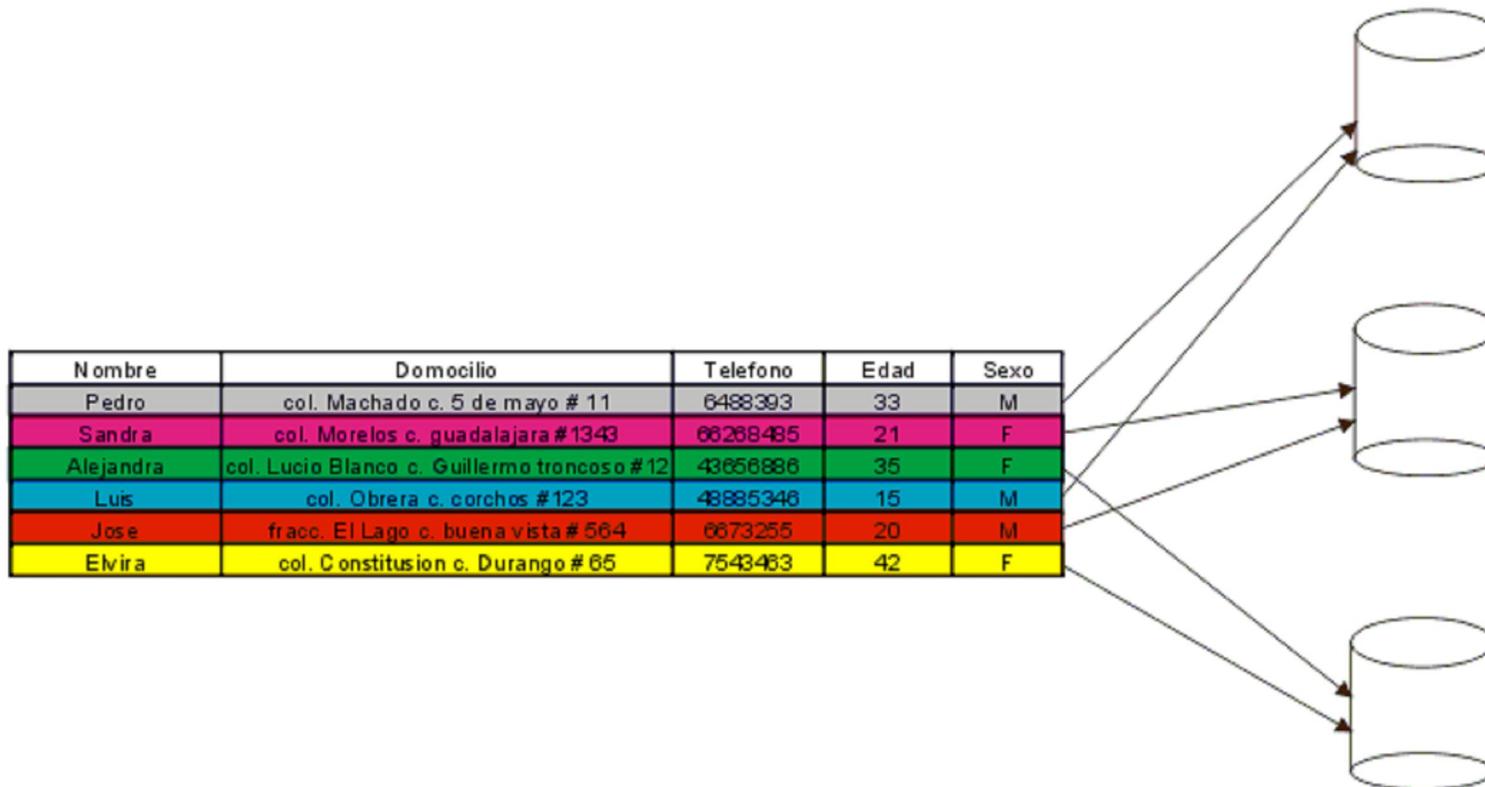
Implementação de paralelismo em BD

Particionamento horizontal

É o tipo mais comum de particionamento. As tuplas de uma relação são divididas entre muitos discos, e cada tupla reside em um disco diferente.

Implementação de paralelismo em BD

Particionamento horizontal



Implementação de paralelismo em BD

Particionamento horizontal

Existem três técnicas de particionamento horizontal:

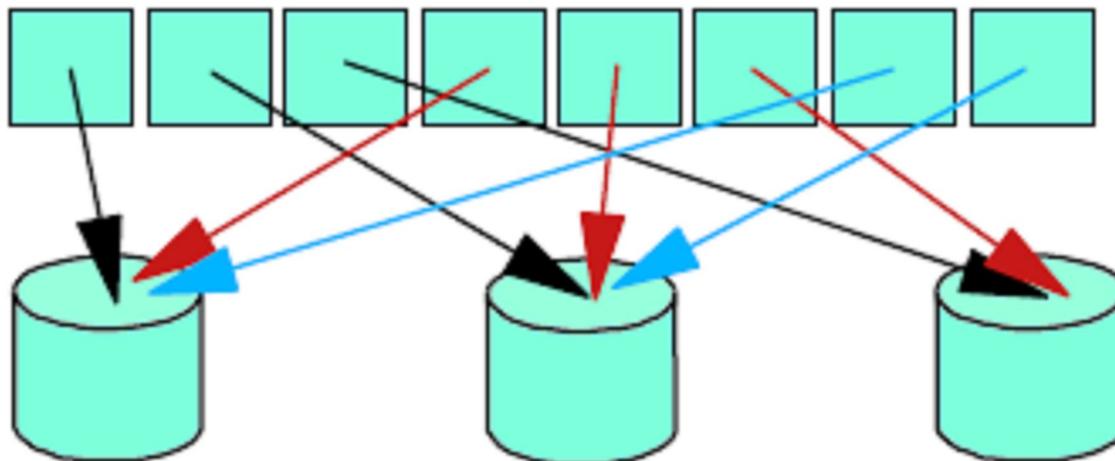
- ❑ Particionamento Round-Robin (Circular)
- ❑ Particionamento Hashing
- ❑ Particionamento por Faixa



Implementação de paralelismo em BD

Particionamento horizontal

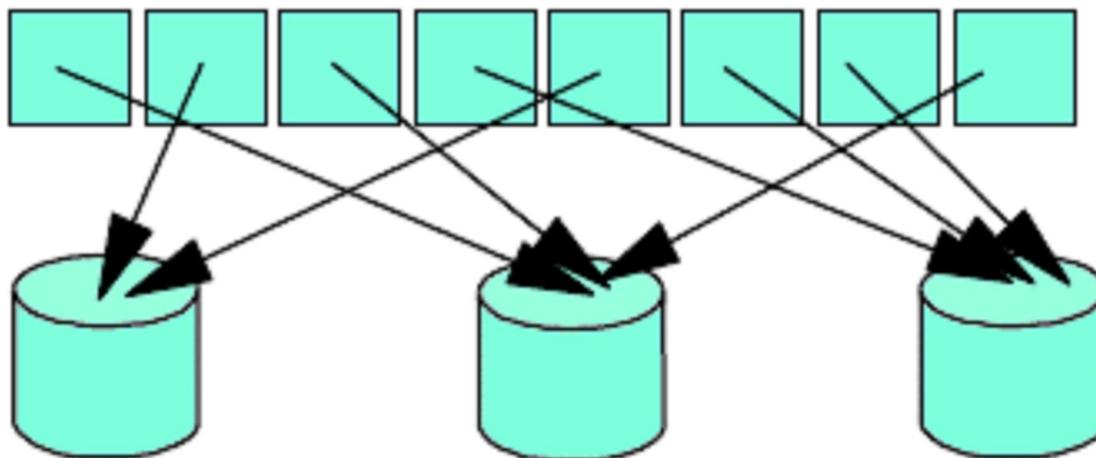
□ Particionamento Round-Robin (Circular)



Implementação de paralelismo em BD

Particionamento horizontal

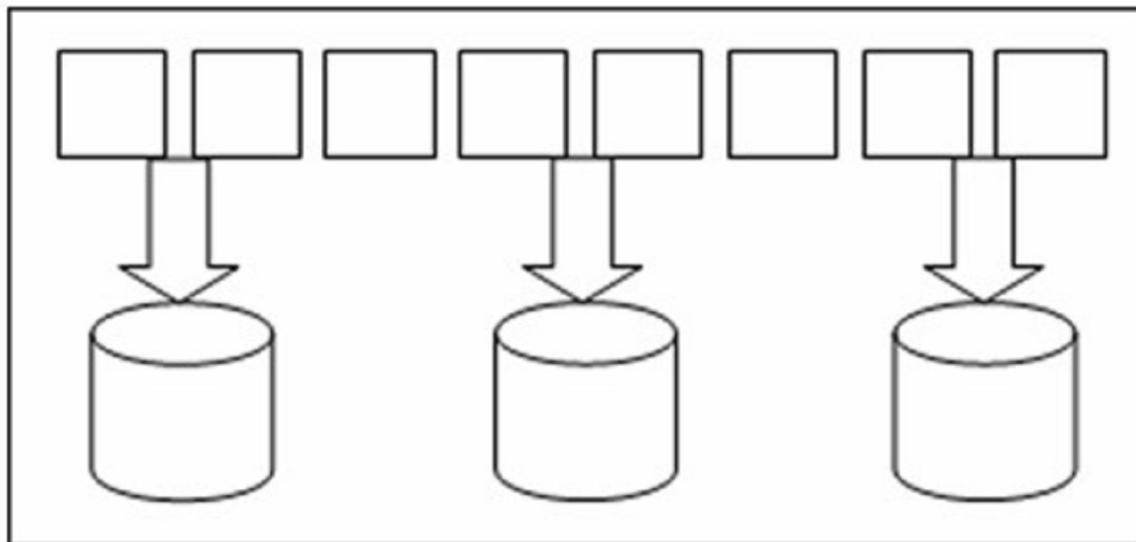
□ Particionamento Hashing



Implementação de paralelismo em BD

Particionamento horizontal

- Particionamento por Faixa



Implementação de paralelismo em BD

Particionamento vertical

Os campos de uma relação se particionam entre os discos, onde cada campo reside em um ou mais discos.

Implementação de paralelismo em BD

Particionamento vertical

| Nombre | Domocilio | Telefono | Edad | Sexo |
|-----------|---------------------------------------------|----------|------|------|
| Pedro | col. Machado c. 5 de mayo # 11 | 6488393 | 33 | M |
| Sandra | col. Morelos c. guadalajara #1343 | 66268485 | 21 | F |
| Alejandra | col. Lucio Blanco c. Guillermo troncoso #12 | 43656898 | 35 | F |
| Luis | col. Obrera c. corchos #123 | 48885348 | 15 | M |
| Jose | fracc. El Lago c. buena vista # 564 | 6673255 | 20 | M |
| Elvira | col. Constitucion c. Durango # 65 | 7543463 | 42 | F |



Implementação de paralelismo em BD

Particionamento misto

Existem dois tipos de particionamento misto:

- HV
- VH

Paralelismo no processamento de consultas

Há dois tipos de paralelismo possível para o processamento de consultas em um bando de dados paralelos:

- ❑ Inter-consultas
- ❑ Intra-consultas



Paralelismo no processamento de operações individuais

A avaliação paralela de operações apresenta alguns custos.

Existem duas formas de paralelismo para o processamento de operações em um banco de dados, as quais podem ser utilizadas simultaneamente em um SGBDP.



Paralelismo no processamento de operações individuais

□ Paralelismo inter-operação

Este método pode ser implementado de duas formas:

- Paralelismo Independente
- Paralelismo Pipeline



Paralelismo no processamento de operações individuais

□ Paralelismo intra-operação

Algumas implementações utilizadas no método de paralelismo intra-operação são:

- Seção paralela
- Classificação em paralelo
- Projeção paralela
- Eliminação de duplicatas em paralelo
- Agregação

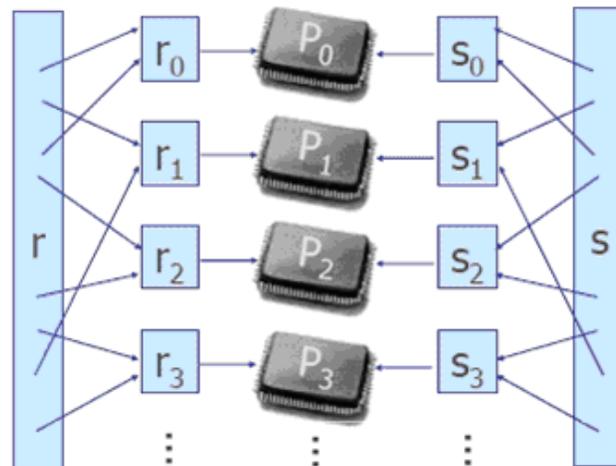


Paralelismo no processamento de operações individuais

- Paralelismo intra-operação

Junção paralela

- Algoritmo de junção particionada

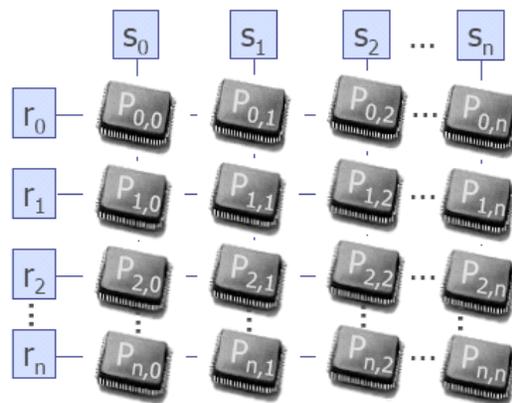


Paralelismo no processamento de operações individuais

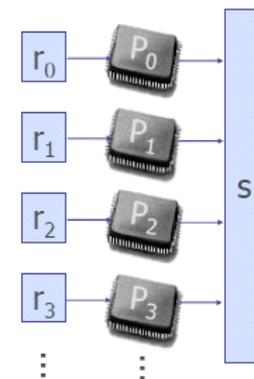
□ Paralelismo intra-operação

Junção paralela

- Algoritmo de junção por fragmentação e replicação



Junção Simétrica



Junção Assimétrica



Paralelismo no processamento de operações individuais

- Paralelismo intra-operação

Junção paralela

- Algoritmo de junção paralela de laço aninhado



Conclusão

Descrevemos as técnicas de fragmentação e alocação de dados, de processamento e otimização de consultas em um banco de dados paralelo, os tipos de paralelismos, bem como detalhes de arquitetura. Sistemas de banco de dados facilitam a exploração de inúmeros hardwares interligados em uma ou várias máquinas, compartilhando, assim, os dados e processamento entre eles.



Referências Bibliográficas

http://www.shammas.eng.br/acad/sitesalunos0606/bdpar/sis_arq.html

<http://www.jsoares.net/CEFET/BD/sbdp.pdf>

<http://www.cos.ufrj.br/~marta/papers/TeseMauroS.pdf>

<ftp://ftp.dca.fee.unicamp.br/pub/docs/ricarte/apostilas/spbdaa.pdf>

www.cos.ufrj.br/~marta/tuttext.ps + banco + de + dados + paralelos&hl=pt-BR&ct=clnk&cd=5&gl=br

FORNARI, Miguel Rodrigues. Sistemas Gerenciadores de Bancos de Dados Geográficos Distribuídos e Paralelos. – Porto Alegre: PGCC da UFRGS, 2002.

SILBERSCHATZ, A.; KORTH, H.F. SUDARSHAN, S. Sistemas de Bancos de Dados, 3ª ed. São Paulo: Makron Books. 1999. 779 pp.

