

# Processamento de Consultas - Materialização e Modelo Volcano

UFPR

Simone Dominico

Laboratório de Banco de Dados

[sdominico@inf.ufpr.br](mailto:sdominico@inf.ufpr.br)

# Modelo Relacional

Nº do cliente	Nome	Endereço
3434	João da Silva	Rua da Esperança
3435	Maria dos Santos	Rua dos Padres
3436	José Abreu	Rua Conselheiro

Tabela Cliente

Referência	Valor	Nº do cliente
Abril/2018	230,45	3436
Jan/2018	126,54	3434
Dez/2017	450,86	3436

Tabela Compras

# Modelo Relacional

Nº do cliente	Nome	Endereço
3434	João da Silva	Rua da Esperança
3435	Maria dos Santos	Rua dos Padres
3436	José Abreu	Rua Conselheiro

Tabela Cliente

Referência	Valor	Nº do cliente
Abril/2018	230,45	3436
Jan/2018	126,54	3434
Dez/2017	450,86	3436

Tabela Compras

# Modelo Relacional

Nº do cliente	Nome	Endereço
3434	João da Silva	Rua da Esperança
3435	Maria dos Santos	Rua dos Padres
3436	José Abreu	Rua Conselheiro

Tabela Cliente

Referência	Valor	Nº do cliente
Abril/2018	230,45	3436
Jan/2018	126,54	3434
Dez/2017	450,86	3436

Tabela Compras

# Modelo Relacional



Como consultar o nome de todos os clientes que o consumo é maior que R\$ 150,00?

Nº c
343
343
343

do cliente
36
34
36



# Processamento de Consultas



# Processamento de Consulta



- Consulta → Extrair informações de um Banco de Dados.
- SQL (Structured Query Language) → Escrevemos **O QUE** queremos dos dados.
- O sistema de banco de dados precisa descobrir **COMO** obter os dados que queremos.
  - Estratégia
- O QUE para COMO → Álgebra Relacional



# Processamento de Consultas

**SQL**

**SELECT** C.Nome

**FROM** Cliente C, Compras Co

**WHERE** C.Numero=Co.Numero

**AND** Co.Valor > 100;



Resultado



# Processamento de Consulta - Álgebra Relacional

## SQL

## Álgebra Relacional

SELECT C.Nome	→	Projeção $\pi$
FROM Cliente C, Compras Co	→	Produto
WHERE C.Numero=Co.Numero	→	Join $\bowtie$
AND Co.Valor > 150;	→	Seleção $\sigma$

$\pi_{C.nome} (\sigma_{co.valor > 150} (Compras) \bowtie Cliente)$

# Processamento de Consulta - PLano de Consulta

## SQL

SELECT C.Nome

FROM Cliente C, Compras Co

WHERE C.Numero=Co.Numero

AND Co.Valor > 150;

## Álgebra Relacional

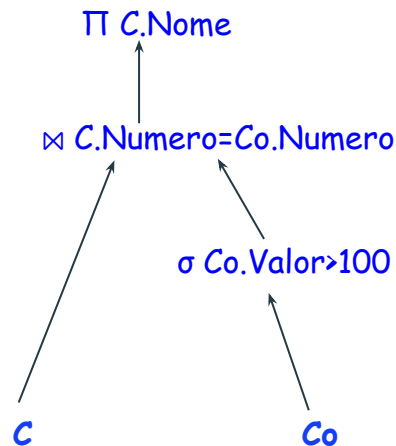
Projeção  $\Pi$


Produto

Join  $\bowtie$


Seleção  $\sigma$

$\Pi_{C.Nome} (\sigma_{Co.Valor > 150} (Compras) \bowtie Cliente)$

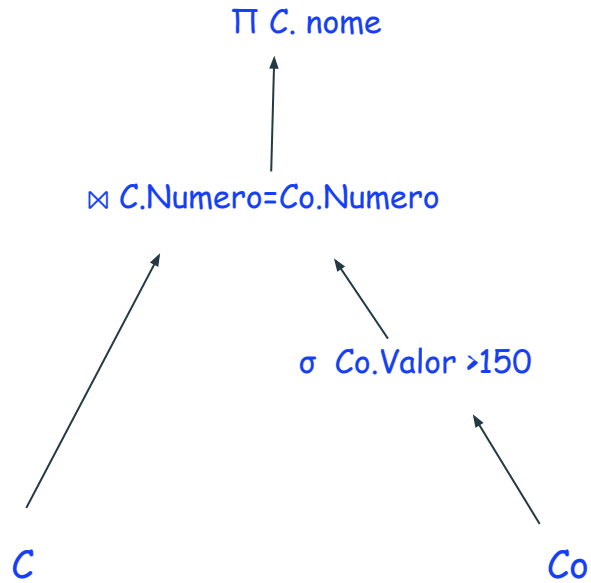




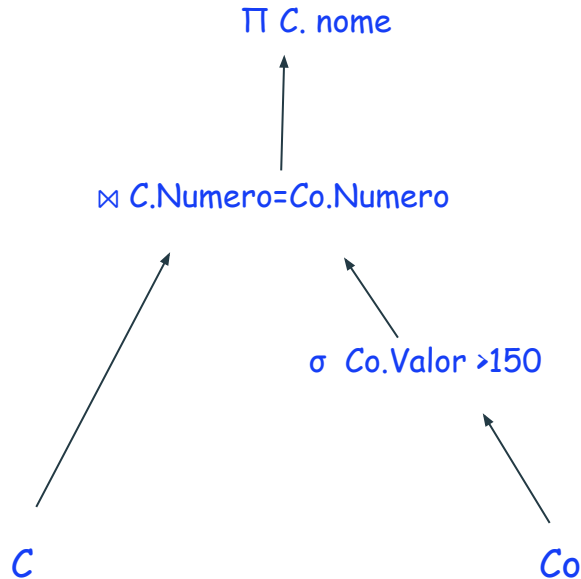
# Modelos de Processamento -Materialização



# Materialização

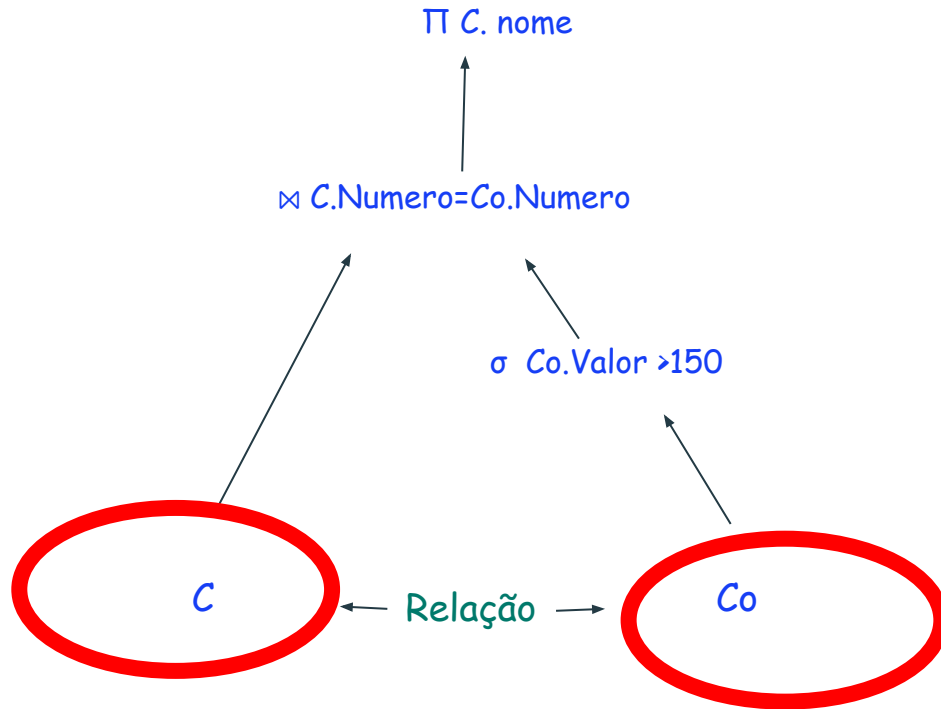


# Materialização



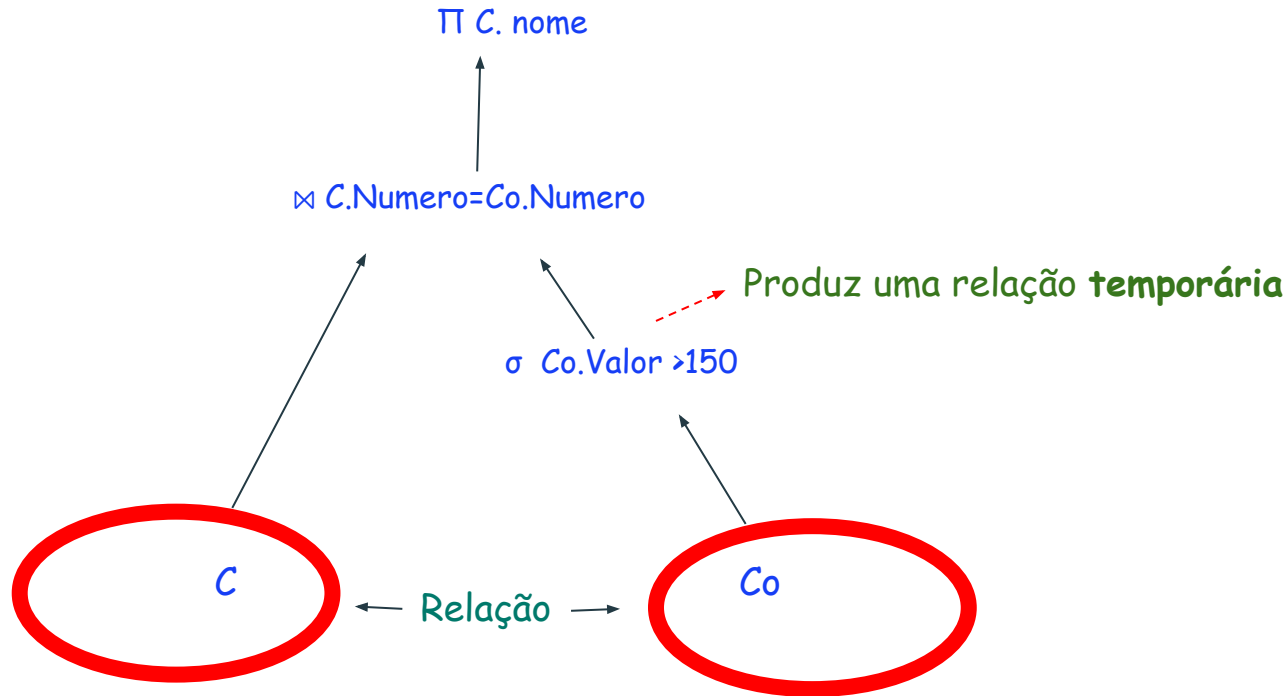
- Os resultados de cada operação intermediária são criados (materializados) e depois usados para avaliação das operações de próximo nível.
- O operador "materializa" a saída como resultado único;
- Processamento de baixo para cima (bottom-Up);

# Materialização

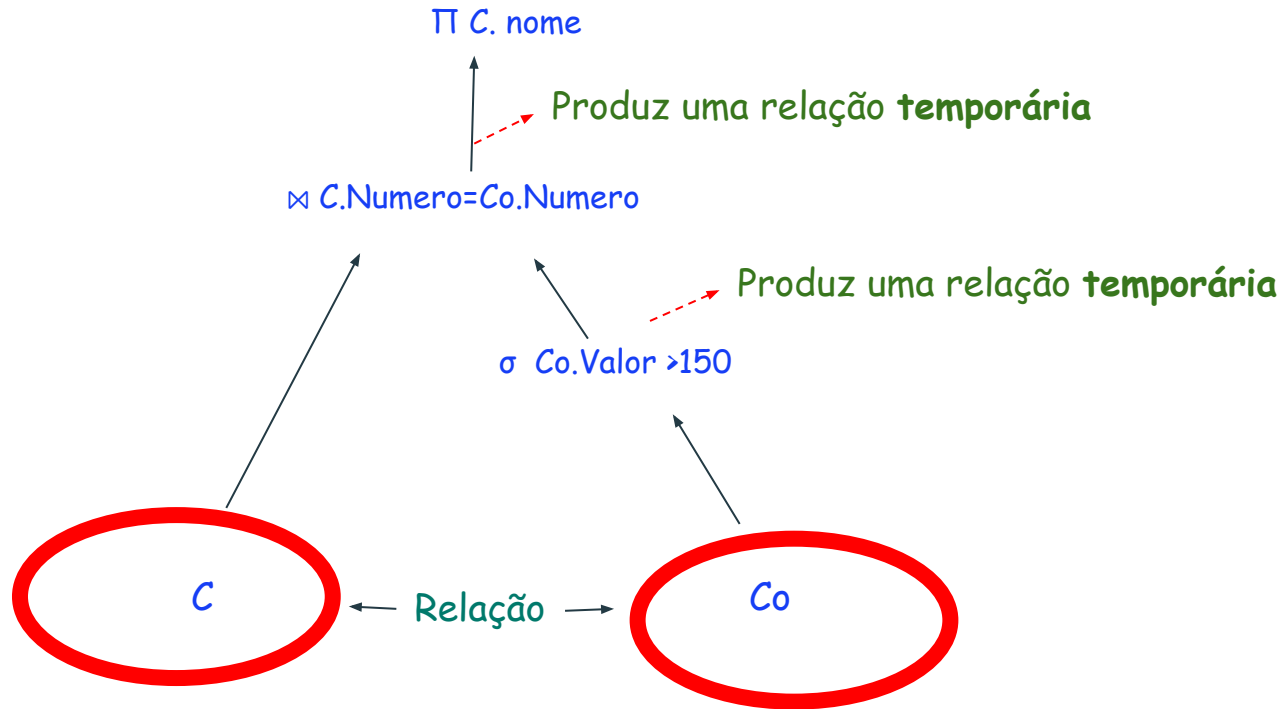


- Uma vez que os resultados de cada operação intermediária são criados (materializados) e depois usados para avaliação das operações de próximo nível.
- O operador "materializa" a saída como resultado único;
- Processamento de baixo para cima (bottom-Up);

# Materialização




# Materialização






# Materialização



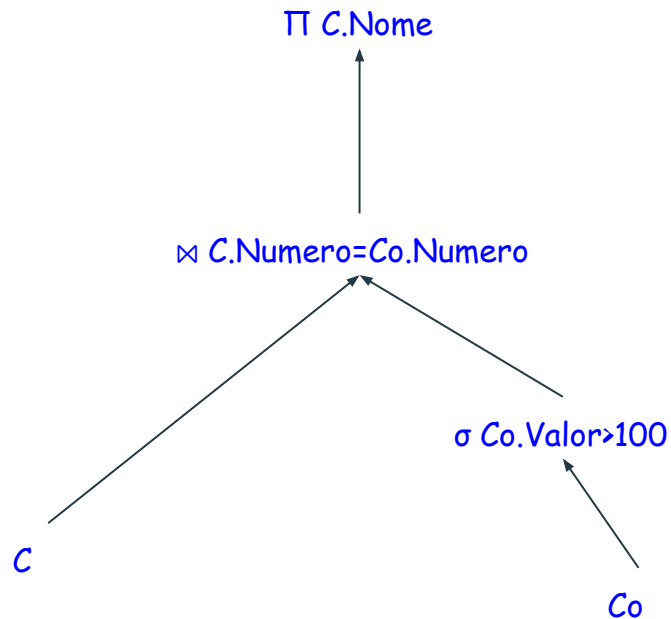


# Modelo Pipeline/Volcano



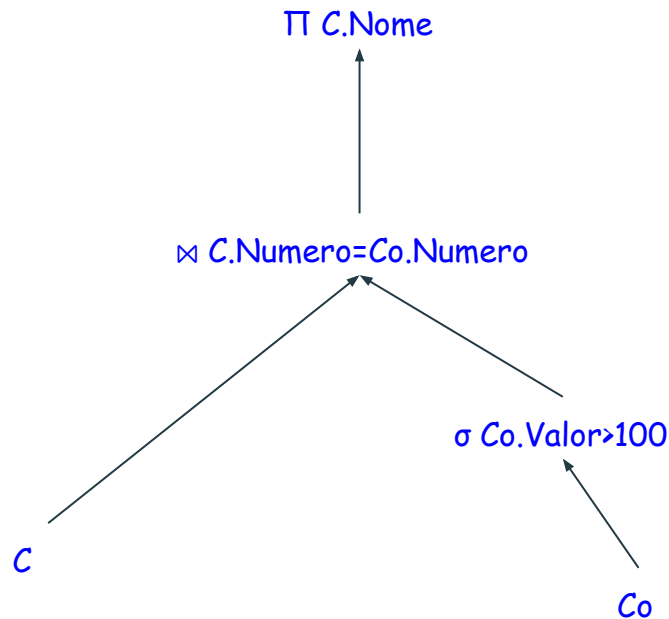
# Modelo Pipeline - Modelo de Iteração

- Cada operador implementa uma função "next";
- Processamento de cima para baixo (Top-down);
- Volcano ou Pipeline;
- Orientado por demanda: as tuplas são produzidas e passadas para outra operação baseado na solicitação;



# Modelo Pipeline/Volcano

- Cada operador implementa uma função "next";
- Processamento de cima para baixo (Top-down);
- Volcano ou Pipeline;
- Orientado por demanda: as tuplas são produzidas e passadas para outra operação baseado na solicitação;
- Produtor-Consumidor: as tuplas são produzidas ansiosamente, não esperam pela solicitação.



# Volcano

Utilizado por quase todos os Sistemas Gerenciadores de Banco de Dados;

ORACLE®



PostgreSQL



Greenplum



cloudera®  
IMPALA

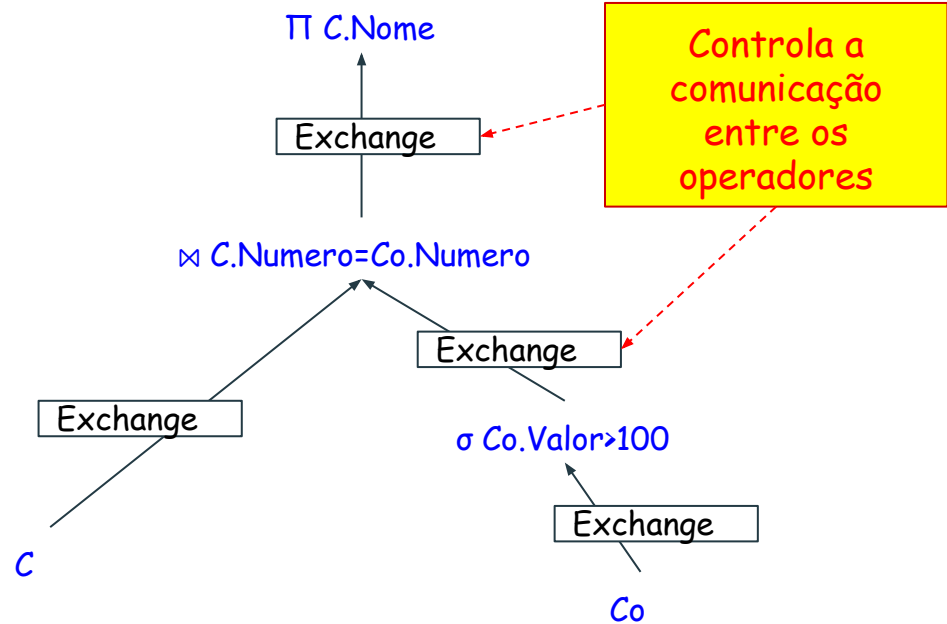


# Modelo Volcano



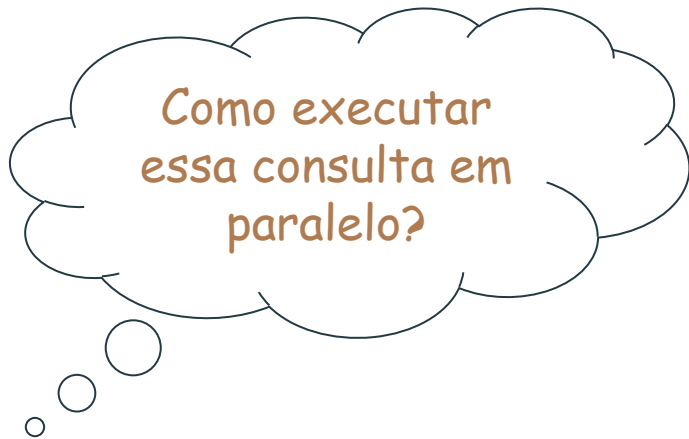
# Modelo Volcano

- Produtor-Consumidor: as tuplas são produzidas ansiosamente, não esperam pela solicitação.



G. Graefe. Volcano - an extensible and parallel query evaluation system. IEEE Trans. Knowl. Data Eng., 6(1):120-135, 1994.

# Modelo Volcano - Paralelismo



Como consultar o nome de todos os clientes que o consumo é maior que R\$ 150,00?

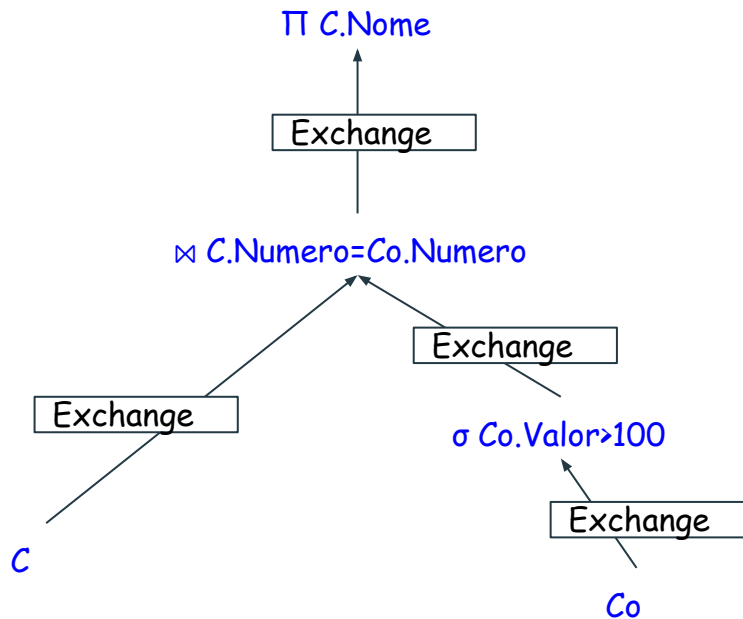


do clien
36
34
36



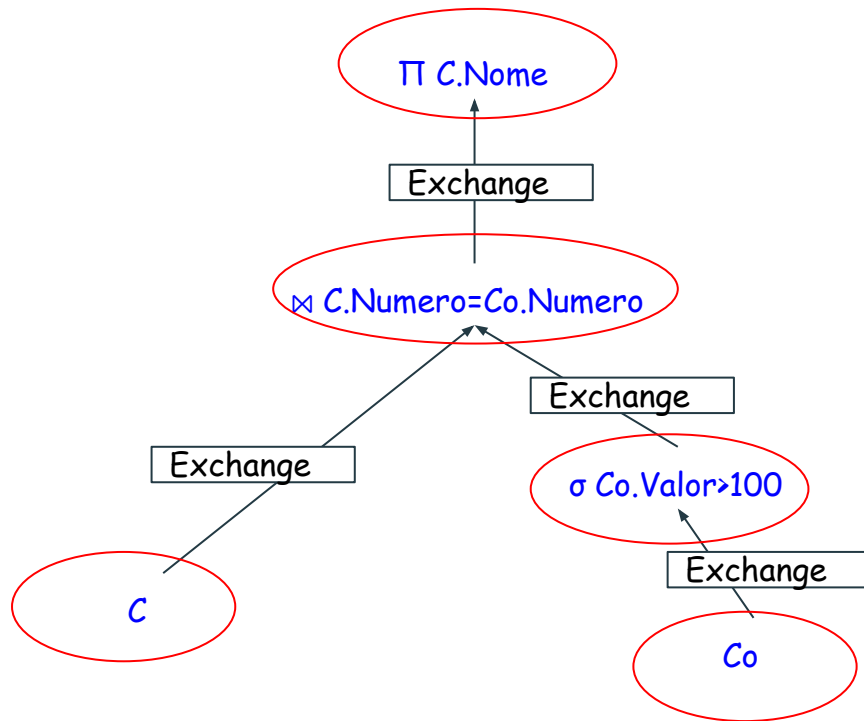
# Modelo Volcano - Paralelismo Vertical

- Cada operação é executada por um processo em paralelo;



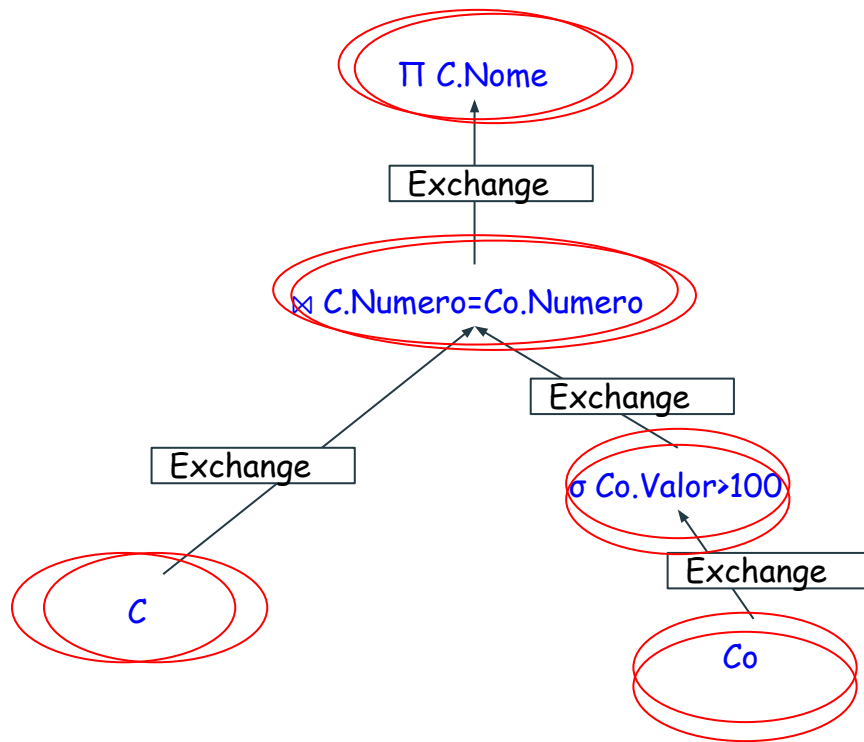
# Modelo Volcano - Paralelismo Vertical

- Cada operação é executada por um processo em paralelo;



# Modelo Volcano - Paralelismo Vertical

- Cada operação pode ser executada por mais de um processo em paralelo;
- O Operador de troca comanda a comunicação entre os processos;





# Conclusão



# Conclusão

- A consulta realiza a extração de informações do banco de dados;
- Consultas SQL são traduzidas em álgebra relacional e representadas em forma de árvore;
- Um plano de consulta pode ser executado de diferentes formas;
- O plano de consulta pode ser executado paralelamente;

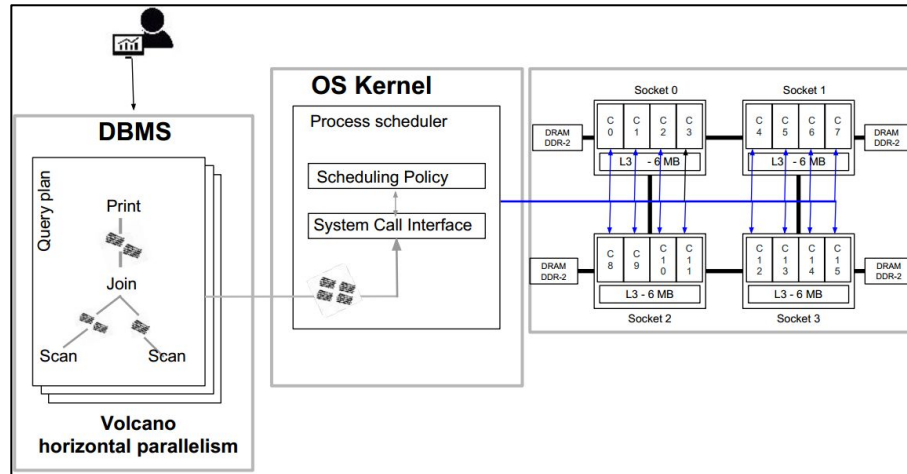


# Pesquisa



# Pesquisa - Assunto

- Execução de carga de trabalho OLAP (Online Analytical Processing) em arquiteturas multi-core (NUMA - Non-uniform Memory access);
- Gerenciamento de alocação de núcleos multi-core através de um modelo abstrato;



International Workshop on  
Data Management on New  
Hardware  
(DaMoN 2017)



# Referências e Material de Apoio

Pavlo, Andy. Lecture 10 - Query Processing. Computer Science Dept. Carnegie Mellon Univ. Disponível em:  
<http://15445.courses.cs.cmu.edu/fall2017/slides/10-queryprocessing.pdf>

Silberschatz, A. Sistema de Banco de Dados. 6 ed., Elsevier, 2012.  
Capítulo 13 – Processamento da Consulta;

G. Graefe. Volcano - an extensible and parallel query evaluation system. IEEE Trans. Knowl. Data Eng., 6(1):120–135, 1994.

# Processamento de Consultas - Materialização e Modelo Volcano

UFPR

Simone Dominico

Laboratório de Banco de Dados

[sdominico@inf.ufpr.br](mailto:sdominico@inf.ufpr.br)