

# Sistemas Gerenciadores de Banco de Dados Distribuídos



**Marcio Belo Rodrigues da Silva**

27 de maio de 1996

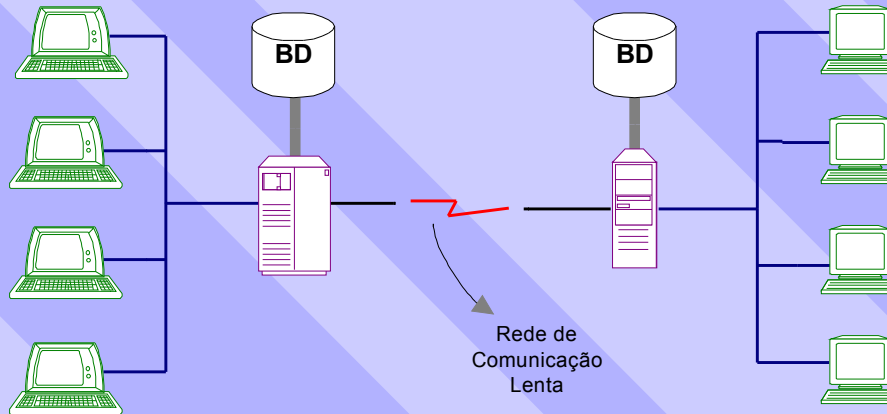
**Seminários**

# Sistemas Gerenciadores de Banco de Dados Distribuídos

## Tópicos da exposição:

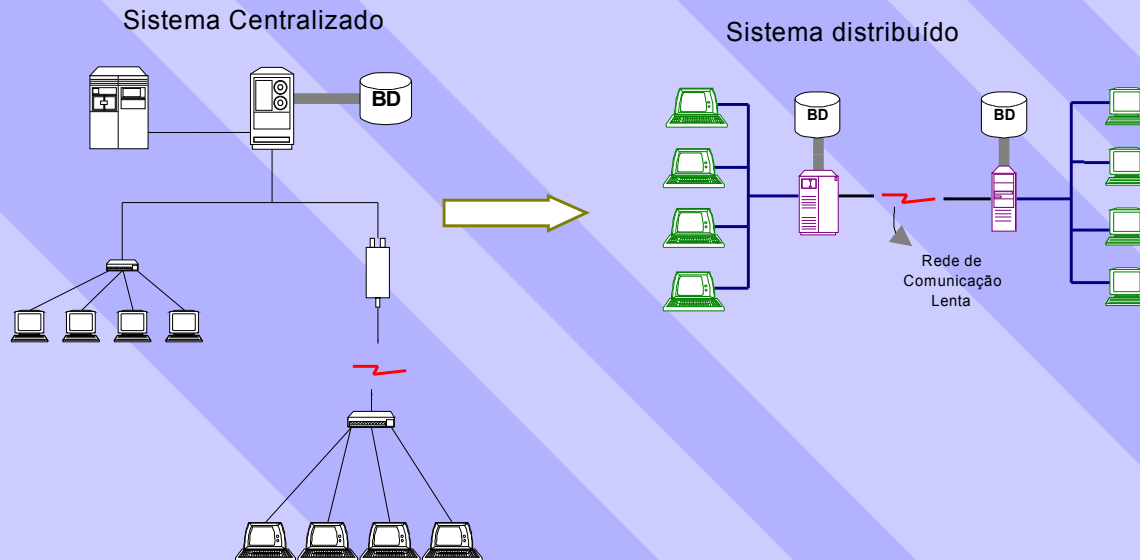
- O que é um SGBD Distribuído ?
- Características
- Vantagens da tecnologia
- Problemas relacionados
- Exemplos de SGBD distribuídos

# O que é um SGBD distribuído ?



- É um software que gerencia Banco de Dados permitindo que este seja distribuído em dois ou mais sistemas de computação geograficamente distantes.

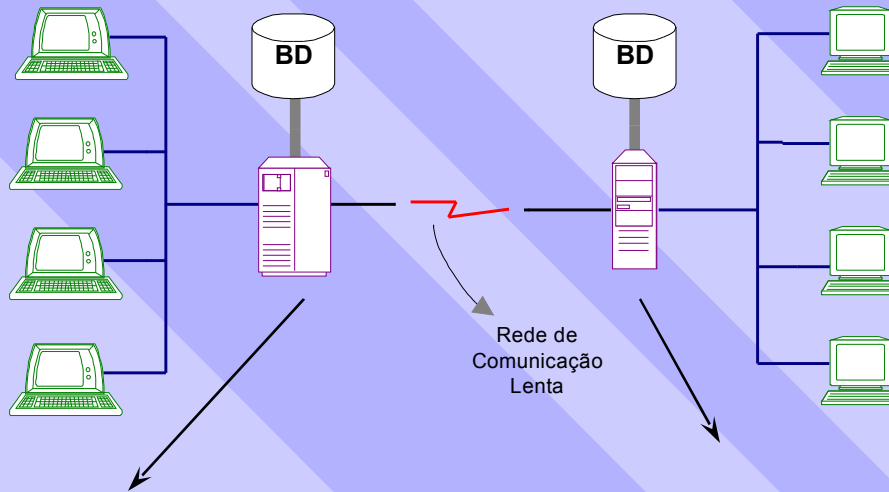
- Disponibiliza um mesmo banco de dados em duas ou mais localidades ligadas à rede



- Descentraliza um sistema de banco de dados, distribuindo o processamento entre duas ou mais estações.
- Permite o particionamento físico dos dados baseado em sua demanda, que pode ser determinada logicamente.

# Características

- O Sistema é homogêneo



SGBD X

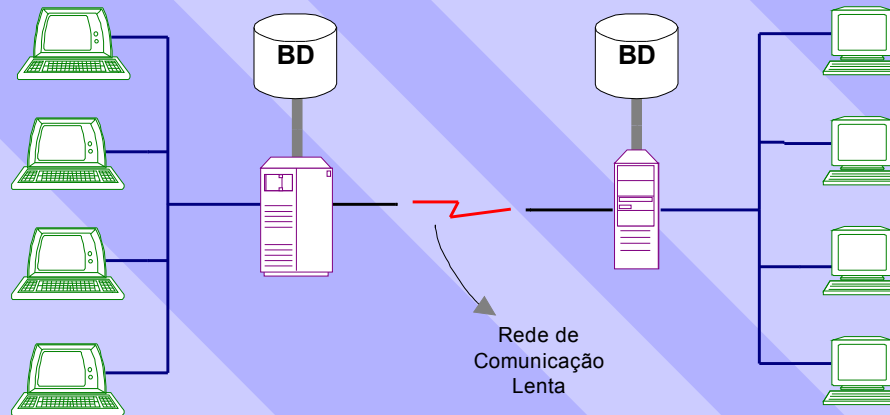
Base de Dados Y

SGBD X

Base de Dados Y

- O SGBD é o mesmo em todas as estações que manipulam o BD distribuído

- Cada localidade (estação) não depende de nenhum serviço ou controle centralizado. AUTONOMIA



- Geralmente, pelo fato das distâncias serem muito grandes, utiliza-se meios de comunicação lentos, com processamento em torno de 50K-100K por segundo.

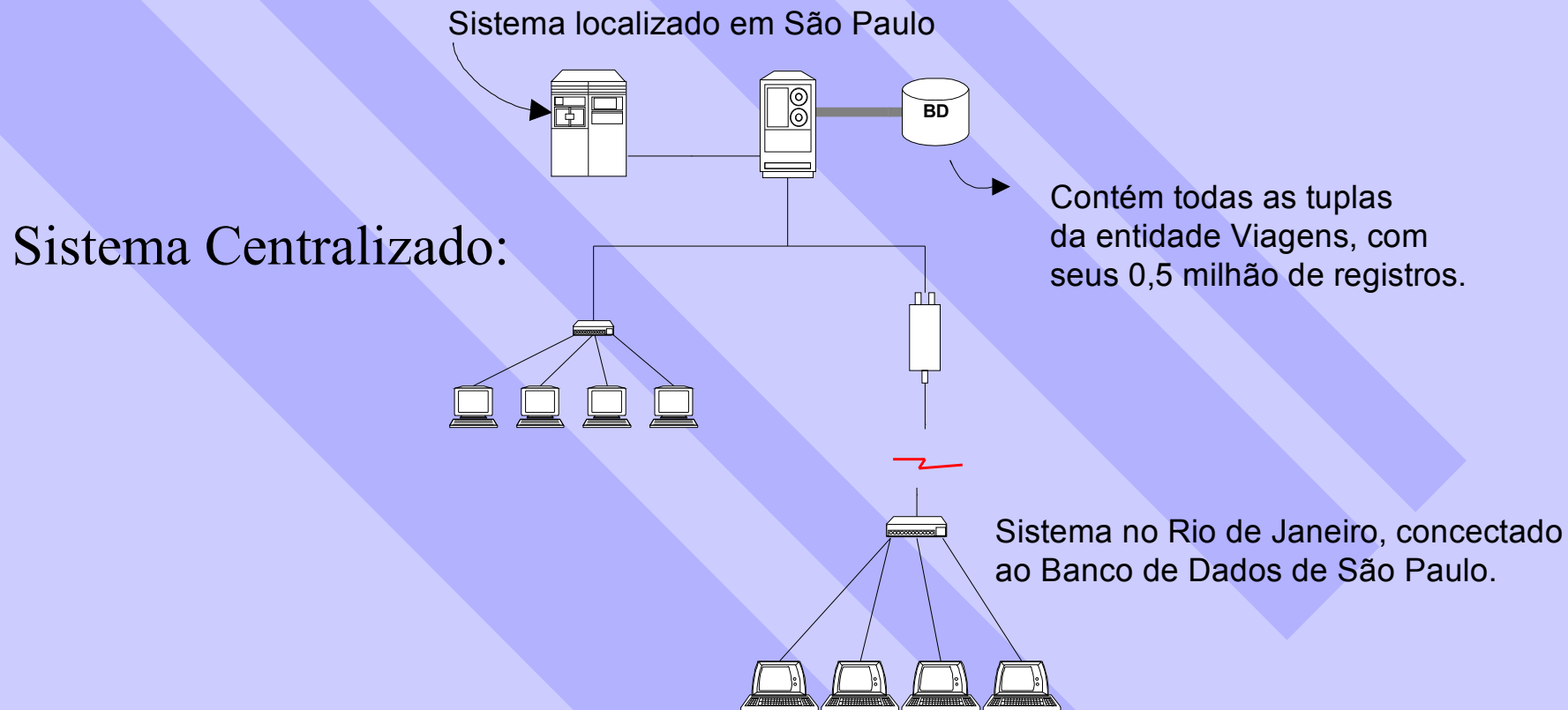
- Permite a fragmentação de dados

Ex: Sistema de Reversas de passagens da Companhia X.

Entidade (tabela) de viagens:

<b>Código</b>	<b>Origem</b>	<b>Destino</b>	<b>Preço</b>	<b>Data</b>	<b>Hora</b>	<b>Capacidade</b>	<b>Reservados</b>
(.....)							
007643	SP	RJ	200,00	22/04/96	00:00h	40	5
007644	RJ	SP	220,00	30/04/96	12:30h	20	8
007645	MG	RS	110,00	27/05/96	03:05h	25	0
007646	SE	PA	120,00	10/04/96	11:00h	30	0
(.....)							

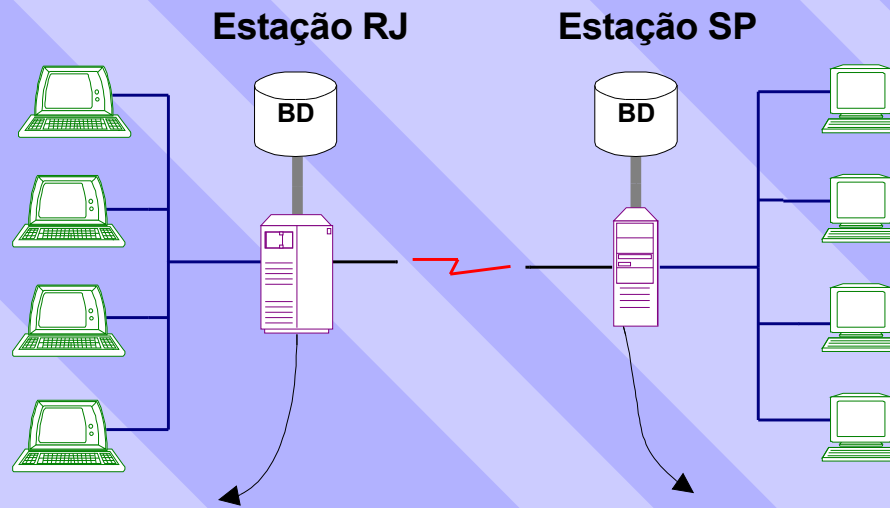
500 mil registros



- Tempo de Resposta alto
- Por mais rápido que seja o sistema computacional central, haverá sempre gargalo nas linhas de tele-processamento.



## Sistema descentralizado:



As tuplas que contém o campo Origem com o valor 'RJ', ficam nesta estação.

As tuplas que contém o campo Origem com o valor 'SP', ficam nesta estação.

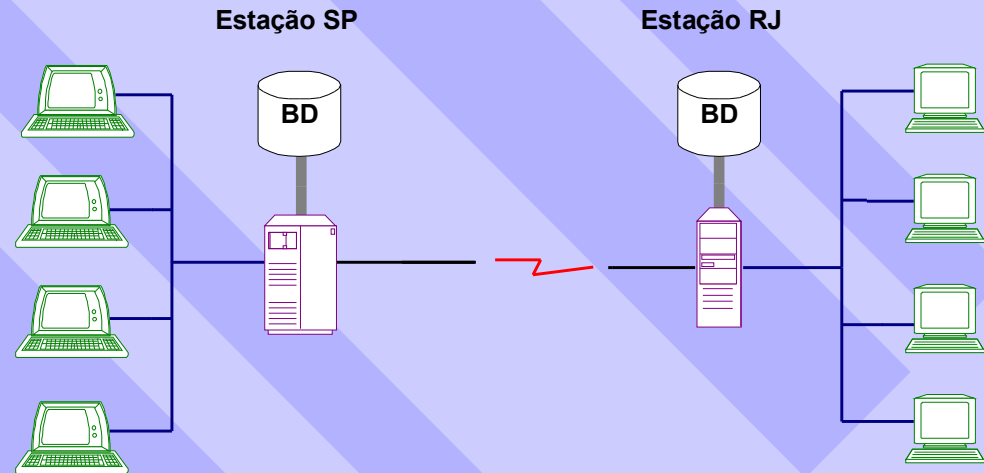
- Melhora radicalmente o tempo de resposta
- Introduce a necessidade de analisar a fragmentação pela sua demanda. -> Fazer com que o maior número de acessos seja local.

- Transparência de localização

- Qualquer terminal logado ao BD distribuído “Enxerga” ele como se fosse único

- Não importa para o usuário/programador onde fisicamente está localizadas as informações

- O SGBD deve dar ao usuário/programador a idéia de que o Banco de Dados é centralizado.



pergunta

PL/SQL: select nome,end,estado from funcs  
where idade >= 65;

resposta

Nome	Endereco	Estado
Fulano	Rua Cabral, 60	RJ
Sicrano	Av. Monsenhor, 20/201	SP
Beltrano	Rua da Lapa, 500/101	SP

## Exemplo de um SGBD não distribuído:

```
DECLARE ESTADO CHAR(2);  
BEGIN  
ESTADO = '<SIGLA DO ESTADO>'  
IF ESTADO = 'RJ' THEN  
    CONNECT DATABASE RJ-SERVER;  
ELSIF ESTADO='SP' THEN  
    CONNECT DATABASE SP-SERVER;  
ENDIF  
SELECT NOME FROM FUNCS;  
END;
```

- Os problemas que envolvem sistemas distribuídos são problemas do SGBD, e não problemas do usuário-programador

- Tipos de Fragmentação

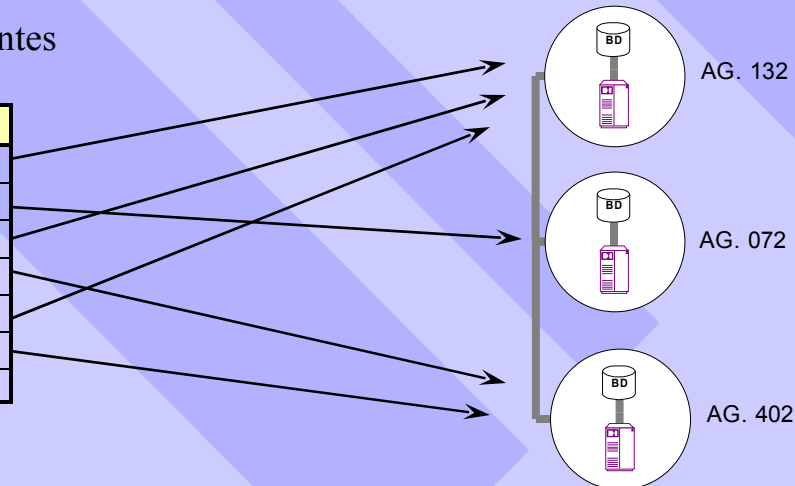
Existem basicamente 3 tipos de fragmentação:

- Fragmentação por restrição

É feita a separação de tuplas de uma tabela de acordo com o conteúdo de um (ou mais) de seus campos.

Tabela de Contas Correntes

Código	Agência	Saldo
00001	132	20
00002	072	15
00003	132	7
00004	402	5
00005	132	0,2
00006	402	8
(.....)		

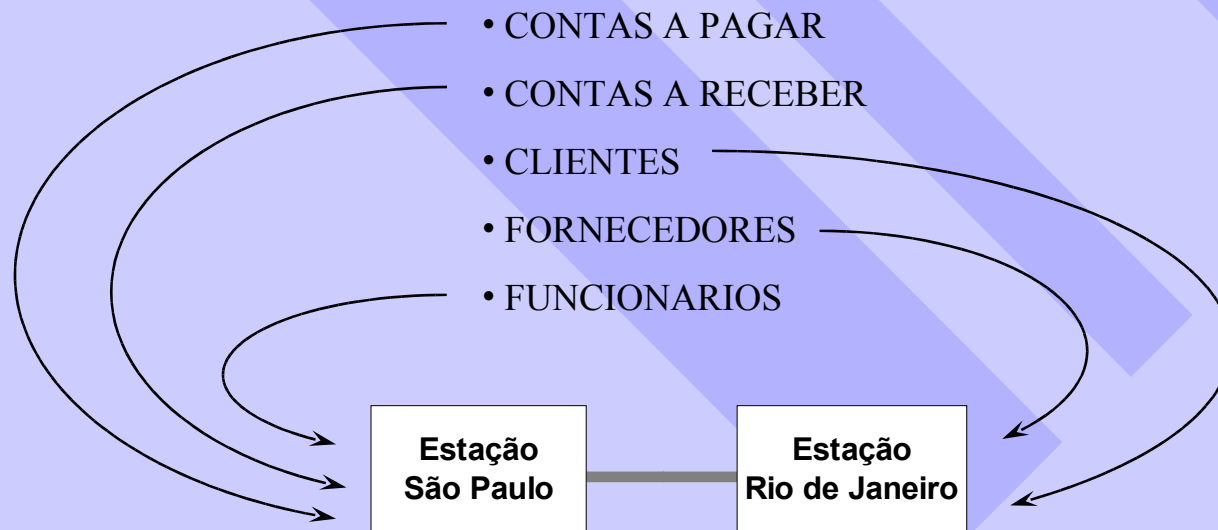


- Fragmentação por entidade de interesse

A entidade (tabela) fica toda localizada em uma das estações.

Ex: A empresa X possui a administração Financeira e de R.H. em São Paulo e a área de Vendas no Rio.

Banco de Dados com as seguintes tabelas



- Fragmentação por projeção

Código	Nome	Total Compra	Situação Financeira	Endereço	Idade	Sexo
(.....)						
0072	Fulano	75,00	OK	Rua X,12/201	30	Masculino
0074	Sicrano	60,00	DÉBITO	Rua Y, 23	40	Feminino
0075	Beltrano	45,00	OK	Av. Brasil, Km4	25	Feminino
(.....)						

Código	Total Compra	Situação Financeira
(.....)		
0072	75,00	OK
0074	60,00	DÉBITO
0075	45,00	OK
(.....)		

**Departamento  
Crédito  
São paulo**

**Departamento  
Marketing  
Rio de Janeiro**

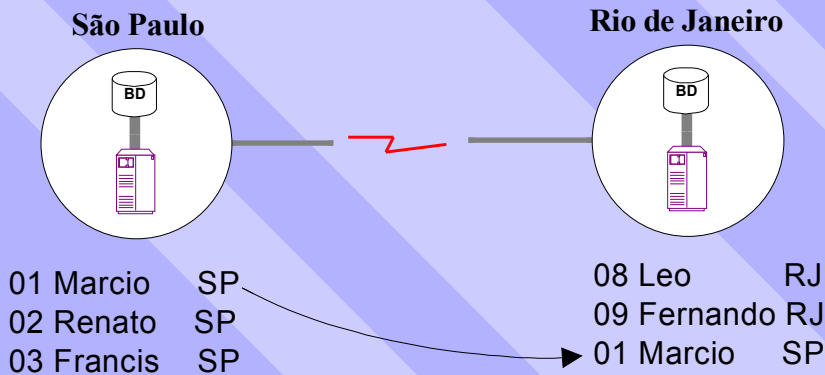
Código	Nome	Endereço	Idade	Sexo
(.....)				
0072	Fulano	Rua X,12/201	30	Masculino
0074	Sicrano	Rua Y, 23	40	Feminino
0075	Beltrano	Av. Brasil, Km4	25	Feminino
(.....)				

## ● Replicação de dados

- consiste em colocar mais de uma cópia de dados do BD em mais de uma localidade (estação).

Repetição de dados

Ex: tabela empregados:

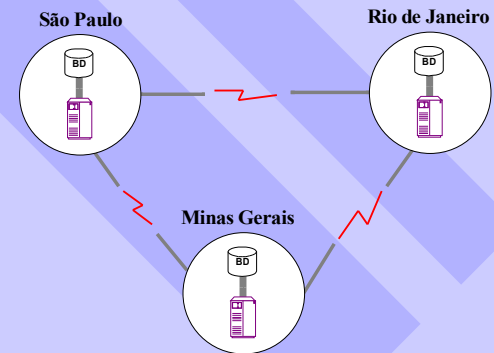


- as consultas são direcionadas para a réplica mais próxima.
- as atualizações devem ser direcionadas à todas as réplicas.
- os dados são replicados quando há muita demanda dele em uma das estações.

# Vantagens da tecnologia

- autonomia local, confiança e disponibilidade

Sistemas centralizados x Sistemas. distribuídos

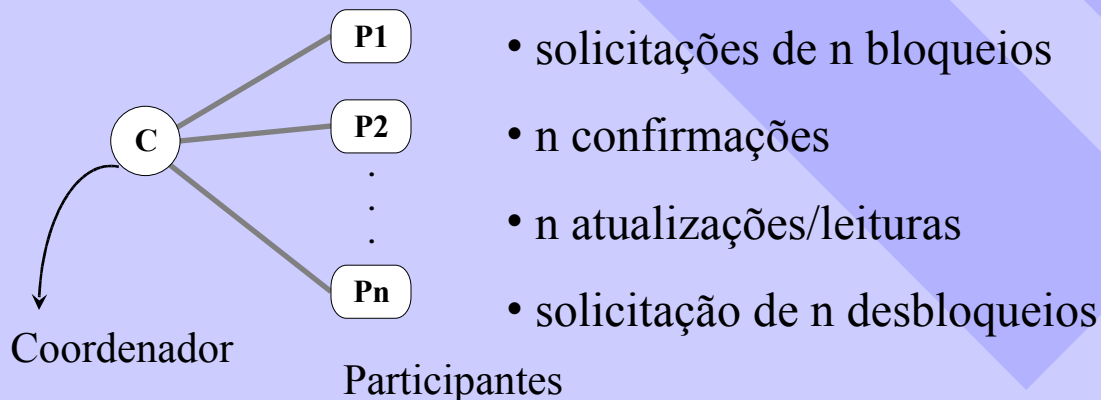


- Capacidade e crescimento incremental
- Eficiência e flexibilidade



# Problemas relacionados

- Velocidade de tele-processamento (50K-100K bits por segundo)
- Processamento de consulta  
Junção de tabelas fisicamente distantes
- Propagação de atualização  
Replicação -> objeto primário
- Concorrência



# Exemplos de SGBD's distribuídos

- SDD-1 (primeiro sistema distribuído)
  - Computer Corporation of America
  - DEC PDP-10, utilizando como rede a ARPANET
  - Possui um otimizador de consulta baseado no operador chamado “semi-junção”. O objetivo é de reduzir a quantidade transmissão.
- R\*
  - Cópias cooperativas múltiplas do System R
  - Conjunto IBM
  - Comunicação via CICS/ISC
- INGRES Distribuído