

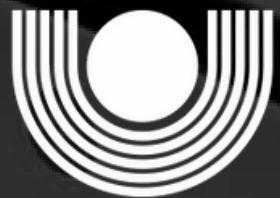


UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ  
CENTRO DE ENGENHARIAS E CIÊNCIAS EXATAS  
DISCIPLINA DE SISTEMAS DISTRIBUÍDOS  
DOCENTE: RENATO BOBSIN

# Banco de Dados Distribuídos

DISCENTES:

Isabela Pimentel Loebel  
Ítalo Gavassi  
Nickolas Crema



# Tópicos

- Introdução
- Caracterização
- Demonstração
- Referencias Bibliográficas



# Introdução

**Banco de dados distribuídos** é um sistema de gerenciamento de dados no qual os dados são armazenados em múltiplos computadores interconectados, em vez de um único servidor [ÖZSU, 1991].

- Projetar um banco consiste em:
  - Análise de Dados;
  - Projeto Conceitual;
  - Projeto Lógico;
  - Projeto Físico.

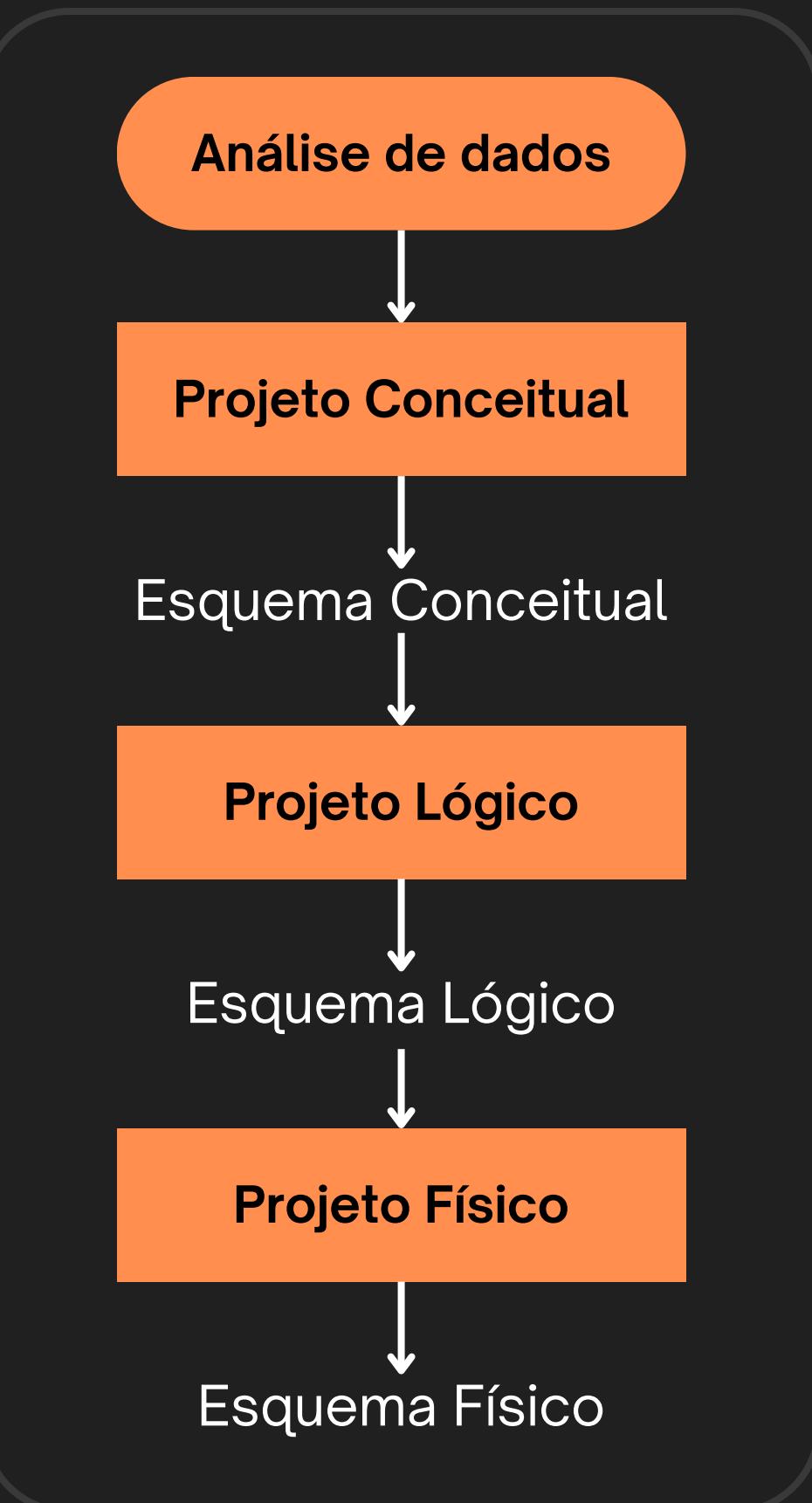


Figura 1: Fases de Projeto de um Banco de Dados.

Fonte: [MESQUITA, 2009].



# Caracterização

## Conceitos

- **Nós:** Dispositivos individuais, cada um com sua própria capacidade de armazenamento e processamento;
- **Rede:** Meio de comunicação que conecta os diversos nós do banco de dados distribuído;
- **Gerenciador de Dados Distribuídos:** Coordena as operações de acesso aos dados distribuídos, gerenciando a distribuição dos dados entre os nós;
- **Distribuição de Dados:** Distribuição eficiente de dados em nós de um banco de dados distribuído para melhorar desempenho e escalabilidade;
- **Replicação de Dados:** Duplicação de dados em múltiplos nós para aumentar disponibilidade e tolerância a falhas;
- **Balanceamento de Carga:** Distribuição equitativa de consultas e atualizações entre nós para usar recursos eficientemente;
- **Tolerância a Falhas:** Implementação de mecanismos para detecção e recuperação automática de falhas nos nós ou na rede.



# Caracterização

## Arquitetura

### Garantir o **ACID**

- **Atomicidade:** Transações são executadas completamente ou revertidas por completo.
- **Consistência:** Transações individuais mantêm a integridade do banco de dados.
- **Isolamento:** Transações são independentes e não interferem entre si.
- **Durabilidade:** Alterações de transações persistem mesmo após o término.





# Caracterização

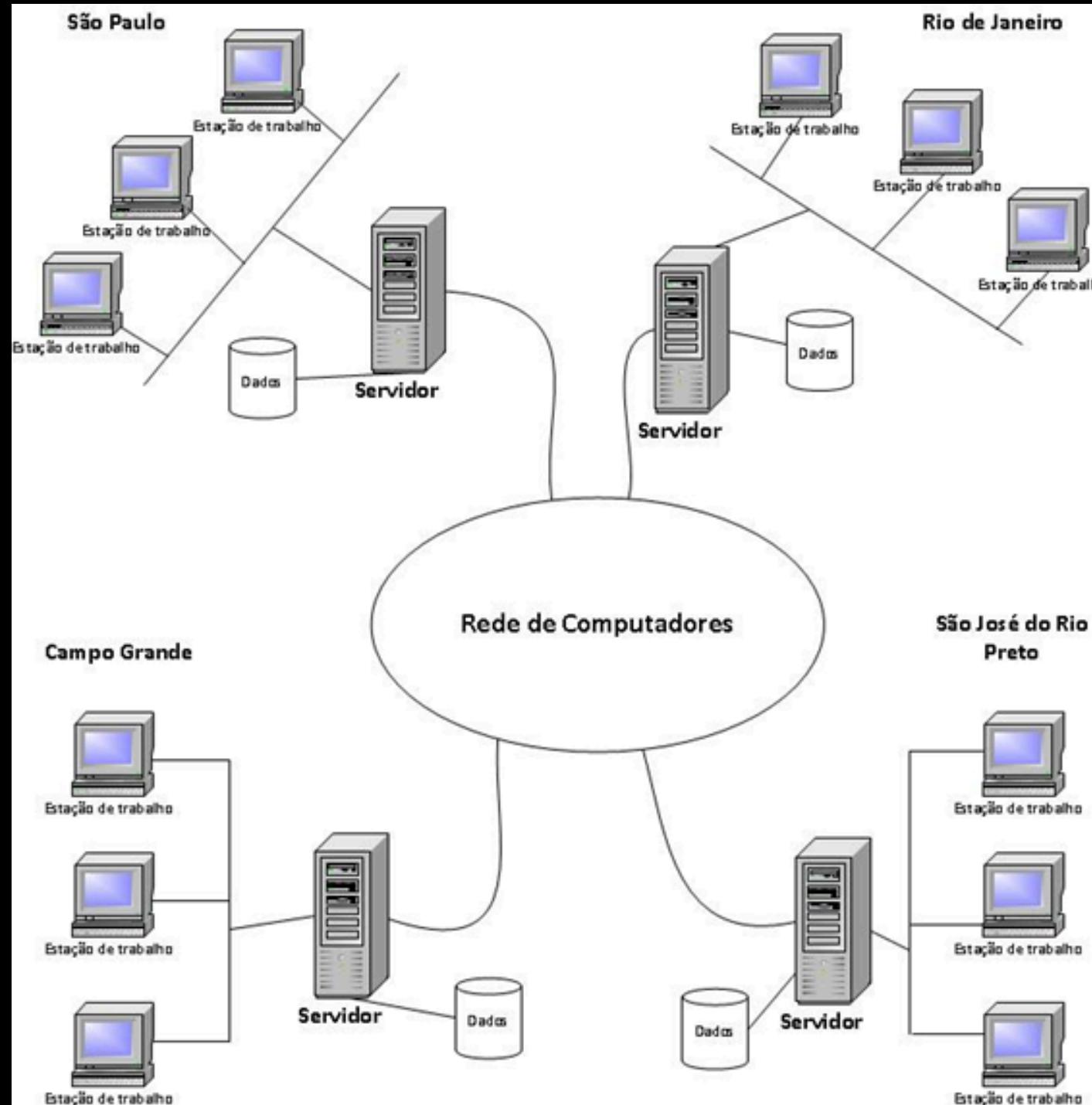


Figura 2: Arquitetura de Banco de Dados Distribuída do exemplo. Fonte: [ROSÁRIO, 2013].

## Exemplo

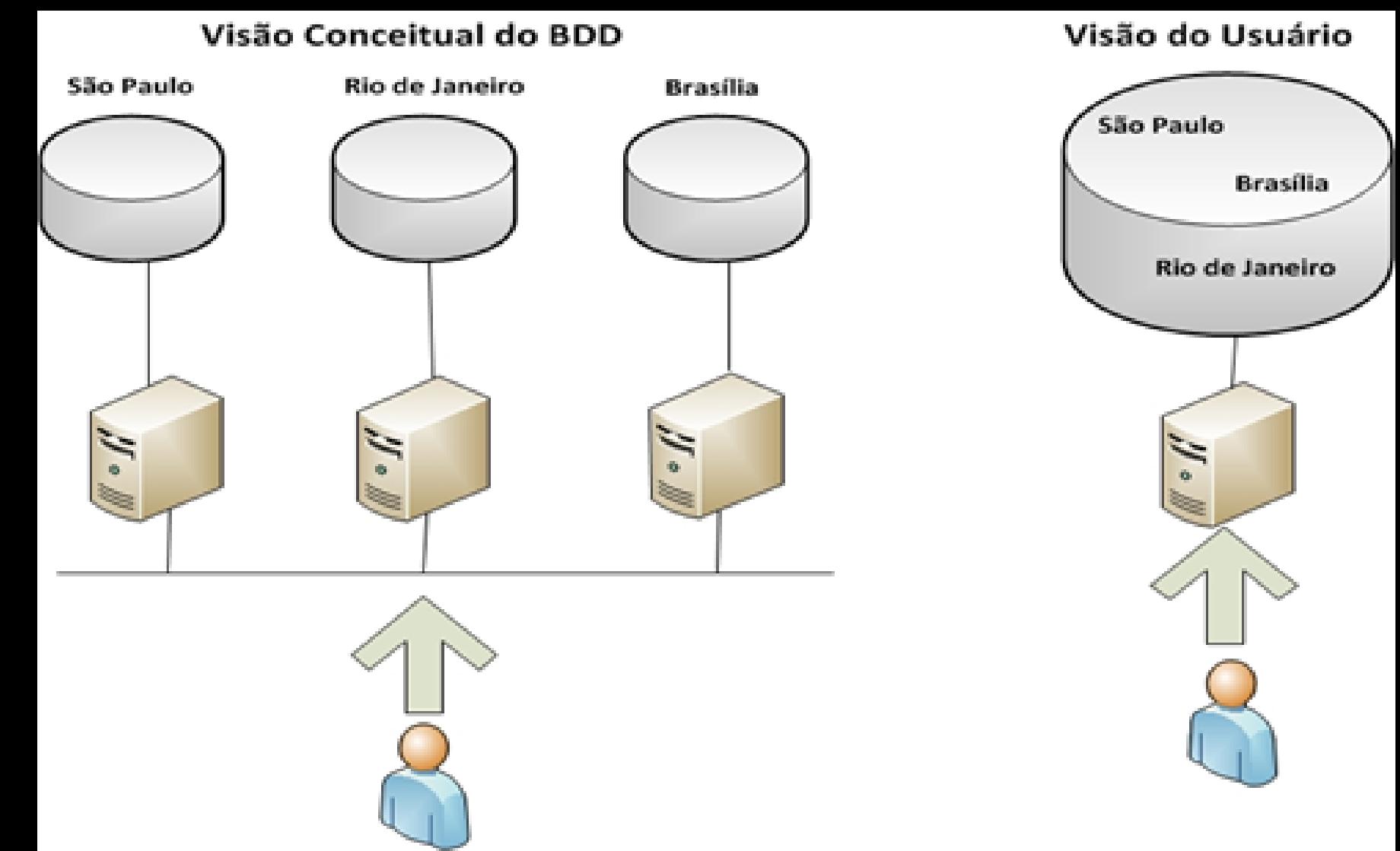


Figura 3: Visão conceitual do Banco de Dados do exemplo. Fonte: [ROSÁRIO, 2013].



# Caracterização

## Vantagens

- **Transparência na Gerência dos Dados Distribuídos e Replicados:**
  - A transparência em sistemas distribuídos simplifica acesso e manipulação de dados, garantindo disponibilidade, confiabilidade e uma experiência de usuário unificada.
- **Confiabilidade através de Transações Distribuídas:**
  - Transações distribuídas em SBDD coordenam operações entre nós, evitando pontos únicos de falha, mantendo a integridade dos dados e garantindo sua disponibilidade.
- **Aumento de Desempenho:**
  - A distribuição de dados entre servidores reduz a latência e o paralelismo no processamento de consultas otimiza o desempenho dos SBDD.
- **Facilidade de Expansão:**
  - Escalabilidade horizontal: permite adicionar nós conforme a demanda cresce, mantendo desempenho e disponibilidade, ideal para ambientes dinâmicos e em evolução.



# Caracterização

## Processamento Distribuído de Consultas

### Abordagens:

- Fragmentação de Consultas;
- Execução de Consultas Distribuídas;
- Otimização de Consultas Distribuídas.

### Fases:

- Análise de Consultas;
- Planejamento de Execução;
- Execução Distribuída;
- Combinação de Resultados;
- Desafios no Processamento de Consultas Distribuídas;
- Técnicas para Otimização de Consultas Distribuídas.

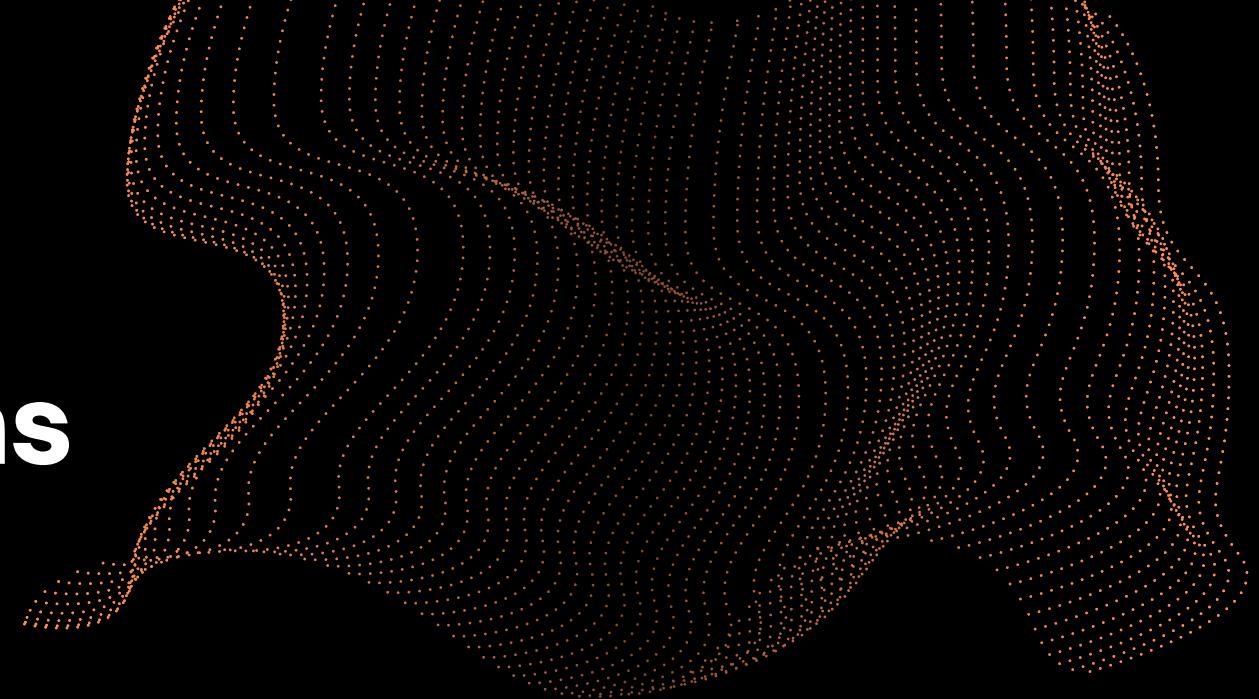




# Caracterização

## Processamento Distribuído de Consultas

### Abordagens



- Fragmentação de Consultas:
  - Divide a consulta em subconsultas executáveis em paralelo, usando algoritmos como Fragmentação por Hash, Intervalo ou Lista.
- Execução Distribuída:
  - Planeja a execução com seleção de operadores eficientes, estimativa de custo e criação de plano de execução.
- Execução Paralela:
  - Realiza subconsultas em paralelo, de forma assíncrona ou síncrona.
- Combinação de Resultados:
  - Utiliza técnicas como Merge-Join, Hash Join ou Nested Loop Join para combinar resultados.
- Otimização:
  - Replicação de Dados, Cache, Balanceamento de Carga, Materialização de Vistas e Reescrita de Consultas visam melhorar desempenho e disponibilidade.

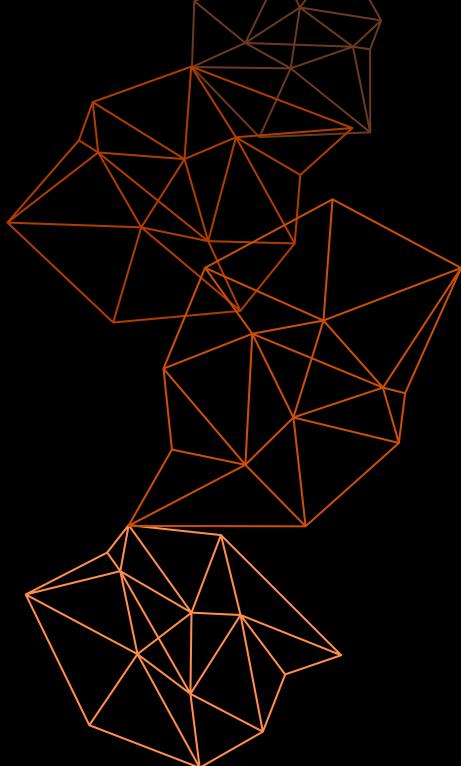


# Caracterização

## Processamento Distribuído de Consultas

### Fases

- Análise e Planejamento de Consultas:
  - Avaliação da sintaxe, identificação de partes para execução paralela e seleção de algoritmos de fragmentação e replicação.
- Execução Distribuída e Combinação de Resultados:
  - Subconsultas executadas em paralelo, transferência de dados entre nós, e combinação dos resultados para retorno ao cliente.
- Desafios e Otimização:
  - Considerações sobre comunicação, falhas de nó, consistência dos dados e segurança, além de técnicas como replicação para otimização.

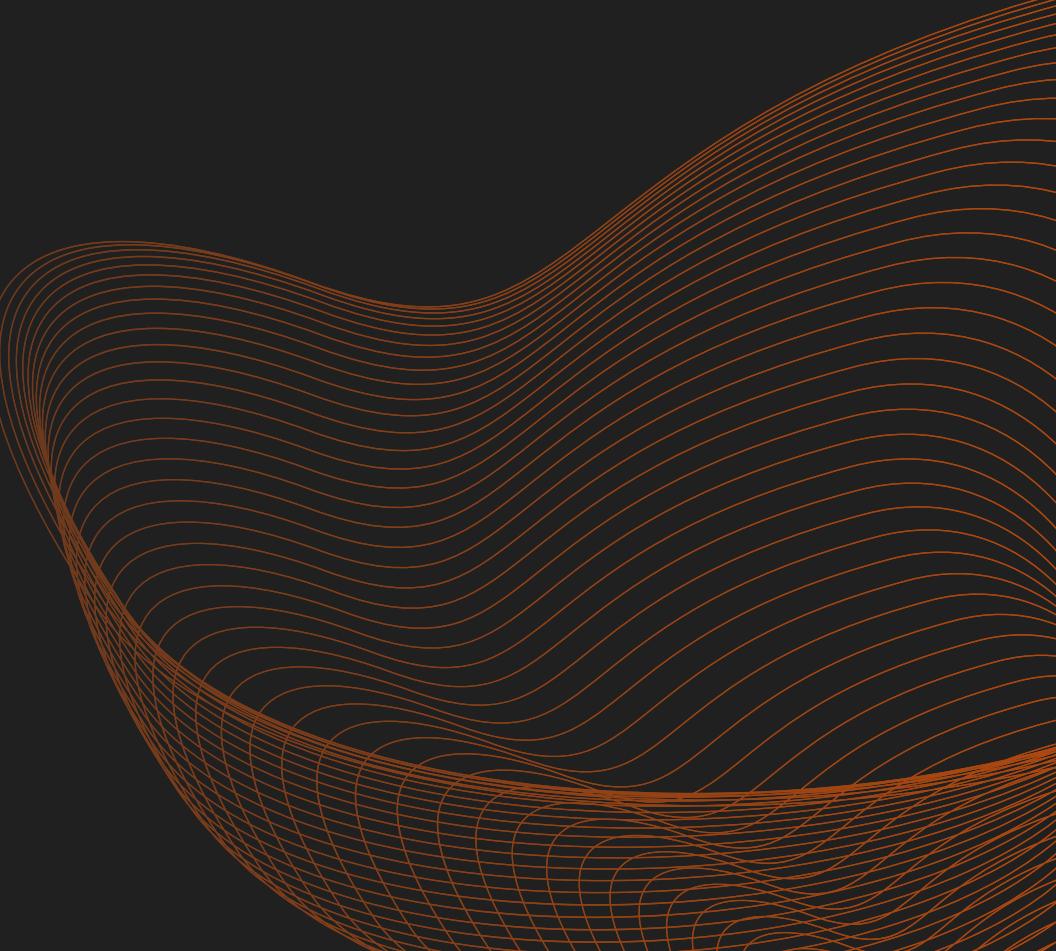




# Caracterização

## Controle de Concorrência

- Necessário para manter a integridade dos dados em transações distribuídas, trata da dispersão de dados e coordenação entre instâncias, com algoritmos de bloqueio, marcação temporal e otimista.
  - Protocolo de bloqueio de duas fases;
  - Algoritmo de bloqueio *would-wait*;
  - Sequência de ordenação por data/hora;
  - Protocolo Otimista.





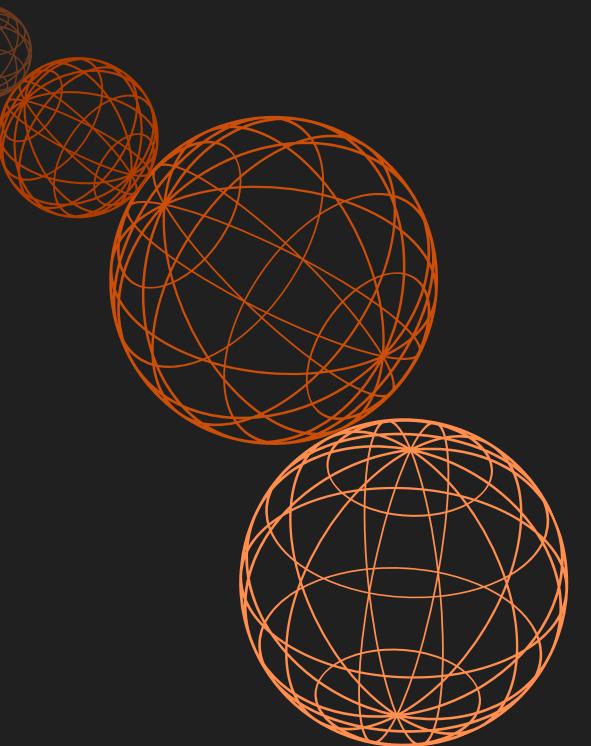
# Caracterização

## Controle de Concorrência

- Protocolo de bloqueio de duas fases:
  - Garante consistência ao converter bloqueios de leitura em escrita, porém não previne deadlocks, demandando reinicialização de transações afetadas.
- Algoritmo de bloqueio *would-wait*:
  - Baseado na política “leia tudo, escreva tudo”, : Prioriza transações antigas, evitando bloqueio de transações recentes para reduzir inconsistências e reinícios.
- Sequência de ordenação por data/hora:
  - Usa carimbos de data/hora para acessar dados, reiniciando transações não sequenciais e descartando atualizações com carimbos anteriores.
- Protocolo Otimista:
  - Utiliza carimbos de data/hora para controle de concorrência, permitindo operações livres até o COMMIT, e validação local de leituras e escritas.



# Demonstração





## Referências Bibliográficas

- MESQUITA, E. J. S.; FINGER, M. **Projeto de Dados em Bancos de dados Distribuídos**. 1998. Tese de Doutorado. Instituto de Matemática e Estatística da Universidade de São Paulo, 27 de mar de 1998.
- CERI, S.; PERNICI, B.; WIEDERHOLD, G. **Distributed database design methodologies**. Proceedings of the IEEE, v. 75, n. 5, p. 533-546, 1987.
- WILDEMBERG, M. et al. **Alocação de Dados em Bancos de Dados Distribuídos**. In: SBBD. 2003. p. 215-228.
- ÖZSU, M. Tamer et al. Principles of distributed database systems. Englewood Cliffs: Prentice Hall, 1999.
- MATTOSO, M. **Sistemas de Bancos de Dados**. Universidade Federal do Rio de Janeiro (UFRJ). 2009. Disponível em: <<https://www.cos.ufrj.br/~marta/BdDistribuido.pdf>>. Acesso em 20 de mar de 2024.
- VIOTTI, P.; KAUFMANN, M. **Modern Distributed Database Systems: The Definitive Guide**. 2015. Disponível em: <<https://docs.aws.amazon.com/whitepapers/latest/microservices-on-aws/distributed-data-management.html>>. Acesso em 20 de mar de 2024.
- CELKO, J.; KAUFMANN, M. **Distributed SQL Databases**. 2017. Disponível em: <[https://en.wikipedia.org/wiki/Distributed\\_SQL](https://en.wikipedia.org/wiki/Distributed_SQL)>. Acesso em 20 de mar de 2024.
- KLEPPMANN, M.; O'REILLY, M. **Designing Data-Intensive Applications**. 2017.
- ÖZSU, M. T.; VALDURIEZ, P. **Distributed database systems: where are we now?**. Computer. 24(8), 68–78. doi:10.1109/2.84879. 1991.
- ROSÁRIO, L. G. **O que é Banco de Dados Distribuído?**. 2013. Disponível em: <<https://imasters.com.br/banco-de-dados/o-que-e-banco-de-dados-distribuido>>. Acesso em 21 de mar de 2024.
- NASSERI, M et al. **Concurrency Control Methods in Distributed Database: A Review and Comparison**. 2017. Disponível em: <[https://www.researchgate.net/publication/319367720\\_Concurrency\\_control\\_methods\\_in\\_distributed\\_database\\_A\\_review\\_and\\_comparison](https://www.researchgate.net/publication/319367720_Concurrency_control_methods_in_distributed_database_A_review_and_comparison)>. Acesso em 21 de mar de 2024.
- RANA, S et al. **Distributed Database Problems, Approaches and Solutions - A Study**. 2018. Disponivel em: <[https://www.researchgate.net/publication/328280121\\_Distributed\\_Database\\_Problems\\_Approaches\\_and\\_Solutions\\_-\\_A\\_Study](https://www.researchgate.net/publication/328280121_Distributed_Database_Problems_Approaches_and_Solutions_-_A_Study)>. Acesso em 21 de mar de 2024.



# Obrigada!

Dúvidas?

Contatos:



- Isabela → [isabelaloebel@gmail.com](mailto:isabelaloebel@gmail.com)
- Ítalo → [italo.gavassi@gmail.com](mailto:italo.gavassi@gmail.com)
- Nickolas → [nick.cremaa@gmail.com](mailto:nick.cremaa@gmail.com)