

# Otimização de Performance em Bancos de Dados

Eficiência e Análise em Ambientes Relacionais

Pedro Henrique Honorio Saito

Laboratório de Métodos de Suporte à Tomada de Decisão

Universidade Federal do Rio de Janeiro

2025-07-28



# Sumário

1. Introdução .....	2
1.1 Contexto .....	3
1.2 Banco de Dados .....	7
2. Problemas .....	8
2.1 Banco de Dados .....	9
2.2 Consulta .....	10
3. Solução .....	12
3.1 Remodelamento do Banco de Dados .....	13
3.2 Aplicação Final .....	14

# 1. Introdução

---

*Resumo* : O LAMDEC, em parceria com a Procuradoria Geral do Município do Rio de Janeiro (PGM), desenvolveu um projeto para **analisar** e **classificar** devedores, otimizando a gestão das Certidões de Dívida Ativa (CDAs).

As **Certidões de Dívida Ativa** formalizam débitos públicos, viabilizando cobranças administrativas ou judiciais.

**IPTU**

**ITBI**

**TAXAS**

**TAP**

**ISS**

**IVVC**

**MULTAS**

**TIS**

O ciclo de vida das CDAs é extenso e não linear, geralmente abrange as seguintes **situações**:

INSCRITA

IRREGULAR

CANCELADA

COBRANÇA

PAGA

EXTINTA

NEGOCIADA

QUITADA

ARREMATACÃO

PARCELADA

SUSPENSA

LEILOADA

O processo, por sua vez, de cobrança passa por diferentes **fases**:

AMIGÁVEL

PROTESTO

JUDICIAL

EXECUÇÃO FISCAL

Os devedores são distribuídos entre **Pessoa Física** ou **Pessoa Jurídica**, identificados, respectivamente, pelo **CPF** e **CNPJ**.

*Problema* : Uma pequena parcela das CDAs possuem devedores atrelados.

O banco de dados da procuradoria é relacional, proprietário e orientado ao **Processamento de Transações Online (OLTP)**, sistemas transacionais destinados ao processamento cotidiano da empresa.



Banco de dados orientado ao **Processamento Analítico Online (OLAP)**. Sistemas OLAP são projetados para análises complexas em grandes volumes de dados, focando em consultas agregadas e tomadas de decisões.



## 2. Problemas

---

- Normalização Irregular das Tabelas
- Relacionamentos N-N entre identidades, CDAs e nomes dos devedores.
- Armazenamento de identificadores (CPF/CNPJ) como BIGINT.
- Desconhecimento de técnicas de **indexação**.
- HistoricoCDA : Tabela com mais de 500 milhões de tuplas.

BARRA DA TNECTIOIJUCA

1905

LEOPOLDINA ESTEVES

```
SELECT
    HistoricoCDA.numCDA,
    MAX( CASE WHEN codSituacaoCDA IN (6, 95) THEN 1 ELSE 0 END ) AS fg_historico_leilao,
    MAX( CASE WHEN codSituacaoCDA IN (7, 96) THEN 1 ELSE 0 END ) AS fg_historico_parcelada,
    MAX( CASE WHEN codSituacaoCDA IN (10, 94) THEN 1 ELSE 0 END ) AS fg_historico_suspensa,
    MAX( CASE WHEN codSituacaoCDA IN (16) THEN 1 ELSE 0 END ) AS fg_historico_arrematacao,
    MAX( CASE WHEN codSituacaoCDA IN (30) THEN 1 ELSE 0 END ) AS fg_historico_negociada,
    MAX( CASE WHEN codSituacaoCDA IN (45) THEN 1 ELSE 0 END ) AS
fg_historico_parcelamento_irregular
INTO lamdec.t_historico_cda_ipitu
FROM HistoricoCDA
LEFT JOIN lamdec.t_tableaux3_ipitu tt3
    on tt3.numCDA = HistoricoCDA.numCDA
WHERE tt3.numCDA IS NULL
GROUP BY HistoricoCDA.numCDA;
```

sql

Vamos analisar a consulta mais de perto..

Vamos analisar a consulta mais de perto..

- Variáveis acessadas: numCDA e codSituacaoCDA.
- Uso em excesso de JOINS.
- Função de agregação MAX( ) e condicional CASE WHEN.
- “Anti-Join” com t\_tableaux3\_ipu.

Vamos analisar a consulta mais de perto..

- Variáveis acessadas: numCDA e codSituacaoCDA.
- Uso em excesso de JOINS.
- Função de agregação MAX( ) e condicional CASE WHEN.
- “Anti-Join” com t\_tableaux3\_ipu.

### Soluções

- Filtragem prévia dos dados.
- Índice clusterizado na coluna numCDA

# 3. Solução

---

- Tratamento de Nulos e datas inconsistentes.
- Criação de um esquema analítico.
- Eleição do campo numCDA para indexação das consultas.
- Técnica de **memoização**.
- Geração de tabelas intermediárias.
- Agendador de processos.
- Informações da **Serpro**.



## 3.2 Aplicação Final

## 3. Solução

