

# ARMAZENAMENTO E PESQUISA DE ARQUIVOS DE MÍDIA EM BANCO DE DADOS RELACIONAIS E NÃO RELACIONAIS

Tiago Cenci

## RESUMO

Este artigo tem o objetivo de responder como funciona o armazenamento de arquivos de mídia em um banco de dados não relacional, e como esse tipo de arquivo pode se beneficiar em uma pesquisa dentro de um banco de dados relacional.

**Palavras-chave:** NoSQL; Banco de Dados Relacional; Armazenamento; Pesquisa; Arquivos de Mídia;

## ABSTRACT

This article aims to answer how media files storage works in a non-relational database and how these types of files can benefit from a search within a relational database.

**Keywords:** NoSQL; Relational Database; Storage; Search; Media Files;

## INTRODUÇÃO

Os bancos de dados NoSQL(“Not Only SQL”) são base de dados que não possuem uma estrutura tabular e armazenam informações de maneira distinta das

tabelas relacionais. Eles apresentam diversos modelos de dados, sendo os principais deles o de documentos, chave-valor, colunas amplas e grafos. Elas proporcionam esquemas flexíveis e se adaptam facilmente a grandes quantidades de informações e alta demanda de usuários, entretanto não existe o conceito de relação entre registros como em um banco de dados relacional. [1]

Já um banco de dados relacional, armazena e fornece acesso a pontos de dados relacionados entre si. Eles são baseados no modelo relacional, uma maneira intuitiva e direta de representar dados em tabelas. Em um banco de dados relacional, cada linha da tabela é um registro com um ID exclusivo chamado chave. As colunas da tabela guardam os atributos dos dados, e cada registro geralmente tem um valor para cada atributo, facilitando o estabelecimento das relações entre os pontos de dados. E neste modelo, existem as relações entre estruturas e registros. [2]

## **O ARMAZENAMENTO DE ARQUIVOS DE MÍDIA EM BANCOS DE DADOS NÃO RELACIONAIS**

Em um banco de dados relacional é possível armazenar arquivos de mídia de duas formas: como um BLOB (Binary Large Object) em um campo de uma tabela ou armazenando um “ponteiro”, para a localização do arquivo no disco. Cada um desses métodos oferece vantagens e desvantagens. Utilizar a técnica de armazenar imagens como um BLOB pode causar um aumento significativo no tamanho do banco de dados, mesmo que as imagens estejam armazenadas em um local centralizado e único. Em contrapartida, optar por armazenar imagens como um ponteiro de arquivo pode reduzir consideravelmente o tamanho do banco de dados, mas isso pode resultar em problemas posteriores se houver arquivos inválidos ou ausentes. [3]

Já em bancos de dados não relacionais, existem formas diferentes de realizar este armazenamento de acordo com o tipo do banco. No MongoDB, por exemplo, podemos usar a especificação “GridFS”, que permite armazenar e recuperar arquivos que excedam o limite de tamanho de documento “BSON” de 16 MB, que seria o caso dos arquivos de mídia na maioria das vezes. O “GridFS” divide o arquivo em “chunks” e os armazena em duas coleções: a coleção de arquivos e a coleção de “chunks”. A coleção de arquivos armazena os metadados do arquivo,

como o nome do arquivo e o tipo "MIME", enquanto a coleção de "chunks" armazena os fragmentos reais do arquivo. Cada "chunk" é armazenado como um documento BSON na coleção de "chunks" e é identificado por um "ObjectID" único. [4]

Existe também o armazenamento de objetos e não por documentos como citado anteriormente, que são ideais para armazenar e recuperar grandes objetos binários, como imagens, arquivos de texto, fluxos de áudio e vídeo, grandes documentos e objetos de dados de aplicativos e imagens de disco de máquina virtual. Eles são projetados para lidar com arquivos muito grandes e fornecer grandes quantidades de armazenamento total para gerenciar todos os arquivos. Os objetos consistem em dados armazenados, metadados e uma ID exclusiva para acessar o objeto. Alguns armazenamentos replicam os dados em vários nós de servidor para permitir leituras paralelas e consultas expansíveis. O compartilhamento de arquivos de rede é um caso especial de armazenamento de objetos que permite acesso a dados altamente escaláveis. [5]

## **A PESQUISA DE ARQUIVOS DE MÍDIAS EM BANCOS DE DADOS RELACIONAIS**

Em um banco de dados relacional, existem algumas possibilidades fascinantes para consulta de dados de mídia, como a possibilidade de consultar esses dados para encontrar informações específicas, como combinações de palavras em trechos de áudio ou pessoas executando certas operações em vídeos. No entanto, essa análise de dados abstratos requer muitos recursos computacionais, o que torna a busca multimídia e a recuperação de dados dependente de anotações do autor para evitar análises em tempo real. Bancos de dados multimídia tornam possíveis pesquisas desafiadoras, mas requerem mecanismos que gerem índices e interfaces para otimizar as consultas. O processo de reaquisição de informações consiste em localizar os registros baseando-se em palavras-chave ou exemplos (similaridade). Por exemplo, um livro, teria as palavras "título", "autor", "sumário", "capa" e etc.

A linguagem que realiza esse tipo de pesquisa, respeita uma estrutura de dados hierárquica definida pelo autor. Então, a consulta ideal depende do autor da mesma, estruturá-la da melhor forma, que se adapte às aplicações e aos dados desejados.

Portanto, para realizar a pesquisa de dados de mídia em um banco de dados relacional, é necessário que haja interpretação de conteúdo. Isso gera esquemas de indexação e análise de algoritmos de imagens, vídeos ou áudios, extremamente complexos. [6]

## CONCLUSÃO

Sobre como funciona o armazenamento de arquivos de mídia em um banco de dados não relacional, foi compreendido que existem diversas formas de se realizar este processo. Tanto no modelo de documentos, como utilizado pelo “MongoDB”, como no modelo de objetos.

Já quanto a como os arquivos de mídia podem se beneficiar em uma pesquisa dentro de um banco de dados relacional, pode-se atestar que a forma de realizar estas pesquisas, são complexas e exigem bastante processamento computacional, mas podem trazer consultas complexas em áudios e vídeos, conforme palavras-chave e sinônimos.

## REFERÊNCIAS

- [1] SCHAEFER, Lauren. **NoSQL Explained**. MongoDB, 2023. Disponível em: <https://www.mongodb.com/nosql-explained>. Acesso em: 12 de abril de 2023.
- [2] Oracle. **What is a Relational Database (RDBMS)?**. Oracle, 2023. Disponível em: <https://www.oracle.com/database/what-is-a-relational-database/>. Acesso em: 12 de abril de 2023.
- [3] SCHULTES, Stan. **Como Armazenar Imagens no Banco de Dados**. Linha de Código, 2023. Disponível em: <http://www.linhadecodigo.com.br/artigo/47/como-armazenar-imagens-no-banco-de-dados.aspx>. Acesso em: 13 de abril de 2023.
- [4] MongoDB. **GridFS**. MongoDB, 2023. Disponível em: <https://www.mongodb.com/docs/manual/core/gridfs/#std-label-gridfs-files-collection>. Acesso em: 13 de abril de 2023.
- [5] TEJADA, Zoiner. **Dados não relacionais e NoSQL**. Microsoft, 2023. Disponível em: <https://learn.microsoft.com/pt-br/azure/architecture/data-guide/big-data/non-relational-data>. Acesso em 14 de abril de 2023.

[6] CODATO, Ana. **PROCESSAMENTO DE CONSULTAS EM BANCOS DE DADOS DE DIFERENTES TECNOLOGIAS.** Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Santa Catarina. Florianópolis, p. 151. 2001.