Faculdade de Tecnologia de Mococa – Mário Robertson de Sylos Filho



**Comparação entre sistemas MongoDB e PostgreSQL**

**COMPARAÇÃO ENTRE SISTEMAS MONGODB E POSTGRESQL**

**Felipe Trepador Quillici; Leonardo Dassan Pereira Schiavon;**

Faculdade de Tecnologia de Mococa

Discentes do curso de Gestão da Tecnologia da Informação

**Professor Tarcísio Jorge Bezerra**

Docente do curso de Gestão da Tecnologia da Informação

**RESUMO**

O presente trabalho tem como principal objetivo, apresentar uma comparação entre dois tipos de sistema gerenciador de banco de dados (SGBD), são eles, PostgreSQL e o MongoDB, possibilitando uma análise mais detalhada dos dois sistemas, e ajudando pessoas que estejam em dúvida sobre qual sistema utilizar, para atingir seus objetivos.

Partindo da ideia de que os sistemas não relacionais (noSQL) como o MongoDB, vem tendo uma crescente demanda nos últimos anos, devido à grande popularização da internet e do uso smartphones, todos os dias são gerados inúmeras quantidades de dados, que precisam ser armazenados e gerenciados. Um bom exemplo que demonstrar com clareza o quanto consumimos dados ultimamente é a popularização de redes sociais, como, Twitter, Facebook, Instagram, etc.

Também será apresentado uma breve comparação entre os dois SGBD’s, analisando seus pontos negativos e positivos, como tipo de estrutura, facilidade de manuseio, escalonamento, entre outros fatores que serão analisados pelo leitor deste trabalho, para que se possa chegar a uma conclusão de qual sistema gerenciador de banco de dados é mais vantajoso para uma determinada finalidade.

No decorrer do trabalho, não ficaremos do lado de nenhum dos dois sistemas, o principal objetivo é informar e deixar com que o leitor tire suas conclusões.

**Palavras-chave**: SGBD, PostgreSQL, MongoDB, Dados, Sistema.

**INTRODUÇÃO**

O crescente avanço da tecnologia apresenta um grande impacto no crescimento de computadores sendo possível a afirmação de que na área em que eles são utilizados é comum que as pessoas tenham conhecimento do termo banco de dados.

O Banco de dados é um repositório de dados relacionais, tais dados são armazenados, possuindo a capacidade de serem gravados, editados, excluídos (caso necessário), podendo ser gerados e mantidos manualmente ou de forma automatizada.

O sistema gerenciador de banco de dados (SGBD) é um conjunto de programas que permite aos usuários criar e manter um banco de dados, facilitando assim os processos de definição, construção, manipulação e compartilhamento de bancos de dados entre vários usuários e aplicações.

O presente trabalho tem como principal objetivo a comparação entre dois sistemas gerenciadores de banco de dados, um relacional (SQL) e outro não relacional (noSQL), a comparação será feita entre MongoDB e PostgreSQL.

Nos dias atuais, com o avanço tecnológico e a popularização da internet a quantidade de dados armazenados é muito maior do que a de alguns anos atrás. Principalmente com o crescente uso da internet e o surgimento das redes sociais, que são acessadas por milhões de pessoas todos os dias, surge a curiosidade em saber como essa quantidade de dados gerados serão armazenados e gerenciados.

 Observa-se o avanço dos modelos de bancos noSQL, que veio para coordenar essa parte do mercado que estava em déficit. Assim se popularizando rapidamente em bancos de dados que necessitavam de grandes quantidades de informação.

**METODOLOGIA**

A pesquisa tem por finalidade gerar um conteúdo acadêmico baseado em obras de outros autores, para que pessoas interessadas no assunto tenham mais conteúdo para aprimorar seus conhecimentos, e servir como base para futuras novas pesquisas.

Inicialmente o presente trabalho apresentará a metodologia que será utilizada, seguido da discussão e resultados obtidos entre os dois sistemas gerenciados de banco de dados, por fim a conclusão colocará os pontos que mais se destacam no trabalho e todo o seu referencial teórico.

Para obter os resultados e respostas acerca da problematização apresentada no trabalho, será feita a análise de dois bancos de dados que seguem caminhos diferentes para obter o mesmo resultado. Na comparação entre os bancos de dados foi analisada a estrutura e as funções usadas. O método de pesquisa utilizado teve como base referências bibliográficas, livros, artigos e sites relacionado ao tema.

**DISCUSSÃO**

**Sistema de Gerenciamento PostgreSQL**

Deu início em 1986, o projeto PostgreSQL na Universidade Berkely na Califórnia (EUA), O professor Michael Stonebraker foi escolhido para criar um sistema de armazenamento de dados, com apoio de alguns órgãos, Amy Research office (ARO) E National Science Foundation (NSF). (MILANI,2008).

O projeto levou um ano para ser concluído, em 1987, sai à primeira versão, entretanto essa não se encontrava totalmente instável, já que, estavam ocorrendo alguns bugs. Em 1991 a empresa Illustra Information Technologies adquiriu o projeto, e logo após incorporou com a empresa Infomix e em 2001 foi comprada pela IBM.

Durante 1994 a ferramenta ficou bastante conhecida dando origem a um programa Postgres95, que tinha como base a linguagem PostgreSQL que possuía pouca compatibilidade na época. Para que ficasse mais acessível, os desenvolvedores Andrew yu e jolly chen incorporaram a linguagem SQL, assim o programa ficou harmonizado com sistemas para as plataformas. (MILANI,2008).

Em 1996 surgiram novas melhorias, onde houve grandes transformações na ferramenta, hoje em dia Postregre95 caiu em desuso o nome e se tornou PostgreSQL, que se encontra na versão 11.0.(MILANI,2008).

O Sistema de Gerenciamento de Banco de Dados relacional tem como função armazenar informações para serem utilizadas posteriormente, essas informações estão presentes na maioria dos setores da nossa sociedade. Algumas delas são bancos, caixas de supermercados, smarthphones entre outros. Os armazenamentos são feitos através de arquivo de texto ou em tabelas.

Os arquivos podem ser acessados por inteiro, alterados, mas há regras previamente definidas. Seguindo recomendação do PostgreSQL. (MILANI,2008).

O PostgreSQL usa procedimentos armazenados e funções em várias linguagens de programação. Além das linguagens predefinidas de C, SQL e PL / pgSQL, podendo ser habilitada facilmente com o suporte para linguagens adicionais, como PL / Perl, PL / Python, PL / V8 (também conhecido como PL / JavaScript), PL / Ruby e PL / R. Este suporte com uma ampla variedade de linguagens permite que você escolha a linguagem com construções que possam resolver melhor o problema em questão. ([O. Obe](https://www.pdfdrive.com/search?q=Regina+O.+Obe);  [S. Hsu](https://www.pdfdrive.com/search?q=Leo+S.+Hsu), 2017 )

De acordo com a pesquisa para criarmos uma tabela no banco de dados, deve-se abrir o terminal de comandos instalado na instalação do PostgreSQL para a criação de objetos em nosso banco. O create table é o comando usado para criar tabelas, então, seguirá este seguinte exemplo, CREATE TABLE Jogador ( id int not null primary key, nome varchar(20), aposentadoria varchar(20), data\_nascimento date);

Pronto, a sua tabela já foi registrada.

PostgreSQL é um software que não precisa de licença para sua utilização, diferente dos outros softwares isso ocorreu quando na universidade de Berkeley na Califórnia o BSD (Berkeley Software Distribution) ele teve sua origem, e é utilizado até hoje, sendo revisado periodicamente. (MILANI,2008), no entanto o PostgreSQL tem compatibilidade com vários sistemas operacionais Linux, Unix, Max Os X Server, Windows. (MILANI,2008). Com isso ele segue um conceito conhecido como ACID (acrônimo de Atomicidade, Consistência, Isolamento e Durabilidade). Conjunto propriedades que vem para garantir as transações no banco de dados, e mesmo que ocorra erros durante a transição ele garantirá que o dado não se perca. (MILANI,2008). O PostgreSQL tem a vantagem de replicações entre servidores, pois sua licença é gratuita, e assim possui fácil acesso a aplicações comerciais, diferente de outros softwares. (MILANI,2008). É possível também aumentar a sua capacidade de processamento de um servidor decorrente as limitações. (MILANI,2008).

Já está embutido no suporte nativo Secure Socket Layer (SSL), o aumento da possibilidade de criação de conexões com mais segurança á partir desses canais e ainda fornece estabilidade para utilização da criptografia SHA1 e o MD5. (MILANI,2008).

Sobre a capacidade e o tamanho do banco de dados as linhas por tabela possuem índice Ilimitado, a Tabela tem 32tb, a linha 1,6tb e o campos possui 1gb. (Carvalho,2017).

**Sistema de Gerenciamento MongoDB**

O MongoDB é um banco de dados não relacional (noSQL) que surgiu em 2009, criado por Eliot Horowitz, ex-fundador do DoubleClick e CTO Dwight Merriman, foi baseado nas experiências de construção em grande escala, alta disponibilidade e robustez. Esse tipo de modelo é completamente distinto do modelo relacional, possuindo muitos diferenciais, como melhor desempenho, custos reduzidos, dados aprimorados, velocidade para o mercado, flexibilidade, além de ter seu código aberto e ser gratuito. (MongoDB, 2020)

O MongoDB é um banco de dados não relacional (noSQL) que surgiu em 2009, criado por Eliot Horowitz, ex-fundador do DoubleClick e CTO Dwight Merriman, foi baseado nas experiências em construção em grande escala, alta disponibilidade e robustez. Esse tipo de modelo é completamente distinto do modelo relacional, tendo muitos diferenciais, como melhor performance, custos reduzidos, dados aprimorados, velocidade para o mercado, flexibilidade, além de ter seu código aberto. (Silva, 2013).

Com o avanço tecnológico, a popularização da internet, o surgimento das redes sociais e o grande aumento na venda de dispositivos moveis, os bancos de dados foram ficando cada vez mais complexos, bancos de dados não relacionais (noSQL) surgem para tornar esse processo mais simples, rápido e robusto. Um grande exemplo de aplicações que necessitam de um grande gerenciamento de dados não estruturados são as redes sociais, como exemplo podemos citar o facebook, onde a quantidade de informações geradas pelos usuários aumenta cada dia mais, o modelo relacional não consegue comportar, não sendo eficaz nesse requisito. Por isso o tipo de banco de dados noSQL que vem sendo cada vez mais usado em empresas de grande porte como, Amazon, Ebay, Sega, Google, Adobe, entre outras. É importante lembrar que os bancos de dados não relacionais não vem para substituir os relacionais como um todo, mas vem como uma solução, onde é necessária uma maior flexibilidade da estruturação do banco. (Silva, 2013).

No exemplo abaixo de armazenamento no MongoDB.

Jogador = {"nome":"pele","aposentadoria":1977,"nascimento":"1940-23-10"}

Esse tipo de banco de dados usa o formato JSON (BSON), que pode ser estendido, por isso é possível acrescentar mais informações. Como no exemplo acima é possível adicionar mais informações ao objeto jogador, como um array de dados por exemplo. (Silva, 2013).

Dado que os documentos têm sua estrutura feita em coleções, não há a exigência de que os documentos armazenados nela tenham todos o mesmo campo, como por exemplo um jogador e um clube que podem ser armazenados em uma coleção, mesmo que não faça sentido nenhum. (Silva, 2013).

O MongoDB atualmente trabalha na versão 4.4, podemos citar como características principais ele ser um sistema open source, de código aberto, além de ser um banco de dados destinados a documentos no estilo JSON (BSON) e preparado para trabalhar na nuvem, tendo suporte de inúmeras plataformas, como por exemplo, Linux, Unix, Windows, Mac OS X. (Silva, 2013).

Atualmente o MongoDB oferece três tipos de Clusters para armazenamento na nuvem, o clusters compartilhados, como o nome já diz, compartilha sua ram e vCPUs, esse cluster é gratuito e oferece um armazenamento de até 512 MB, até 2 GB tendo um custo de 9 dólares mensais, e o plano com armazenamento de até 5 GB tem um custo de 25 dólares mensais. Os dois planos pagos oferecem uma copia de segurança (backup) dos dados armazenados, os dois outros clusters são para armazenamento em servidores locais, e serviços de desenvolvimento de aplicativos, os dois variam muito seus preços e tamanhos de armazenamento, sendo muito customizáveis, subindo o preço conforme a demanda dos clientes.

**RESULTADOS**

O banco de dados MongoDB é destinado a documentos, construídos em uma arquitetura escalável, sendo muito utilizado para armazenamento na nuvem, gerenciando a entrega de dados a aplicativos que o usem. Lidando com cargas de trabalho transacionais, operacionais e analíticas em escala, esse banco de dados é uma ótima escolha para as pessoas que buscam uma maior produtividade, mais agilidade em metodologias, produtividade para o desenvolvedor e suporte a DevOps (Desenvolvedor operacional).

Os modelos não relacionais possuem algumas características muito importantes, que fazem esse modelo ser diferente do relacional. Como por exemplo podemos citar a escalabilidade horizontal e a ausência de esquema. A escalabilidade horizontal torna esse SGBD muito adequado para gerenciar grandes volumes de dados que crescem cada dia mais. Uma maneira de alcançar a escalabilidade horizontal é utilizando “Sharding”, que tem a função de dividir os dados em diversas tabelas que serão armazenadas nos diversos nós de uma rede, a utilização de “Sharding” tem o problema de romper a lógica de relacionamentos, então as aplicações têm que gerenciar a complexidade causada pela fragmentação de informações.

Já a ausência de esquemas é uma característica forte desse SGBD, pois a falta de esquema é o que garante e permite a maneira fácil da escalabilidade e o aumento de disponibilidade, porém não é garantido a integridade dos dados, algo que é bem diferente dos relacionais que mantêm uma estrutura rígida.

O noSQL tem um suporte nativo a replicação, sendo uma maneira de escalar, com essa replicação o tempo gasto para a recuperação de informações se tornando muito mais rápida. Duas das principais maneiras de replicação são:

* Master-Slave: sendo a mais aconselhada quando se tem um grande volume de dados. A cada escrita no banco, nós escravos são criados, e a escrita é feita no nó mestre, e o mestre em cada nó escravo, tornando a leitura muito mais rápida.
* Multi-Master: Cria vários nós mestres, que diminui o gargalo gerado pela abordagem Master-Slave, mas possui o problema de conflito de dados.

Como esse SGBD não está focado nos dados armazenados, mas sim em recuperá-los de forma simples e eficiente, as APIs simples são desenvolvidas para facilitar o acesso, permitindo que qualquer aplicação possa utilizá-las de forma eficiente.

Já o PostgreSQL, como seu nome já diz é um banco de dados relacional (SQL) de nível empresarial, que foi implementado a mais de 30 anos no mercado e vem crescendo e expandindo seus recursos cada dia mais. Contando com uma arquitetura escalável o PostgreSQL tem tudo que uma empresa precisa no requesito banco de dados relacional, possuindo uma ótima compatibilidade, servindo milhares de consultas em centenas de tabelas, levando o SQL ao limite. Ambos os bancos de dados fazem um ótimo trabalho para a finalidade que foram criados, basta você decidir qual deles se encaixa melhor para suas necessidades.

Na tabela abaixo foi realizada a comparação referente Banco de dados PostgreSQL e MongoDB. Assim analisando o desempenho dos dois bancos de dados. O teste descrito mostra o escalonamento, consistência e disponibilidade assim chegando ao resultado da comparação dos dois bancos de dados.

**Tabela 1: comparação entre PostgreSQL e MongoDB.**

|  |  |  |
| --- | --- | --- |
|  | **Banco de Dados Relacional(PostgreSQL)** | **Banco de Dados Não Relacional** **(MongoDB)** |
| **Escalonamento** | É possível o escalonamento em um modelo relacional, porem é complexo de ser realizado, possui natureza estruturada, a inserção dinâmica e transparente de novos nós na tabela não é realizada naturalmente. | Nesse tipo de SGBD é mais comum o escalonamento horizontal, tornando o MongoDB uma solução de gerenciamento de grande volume de dados que cresce exponencialmente . |
| **Consistência** | Há regras de consistência são extremamente rigorosas na consistência das informações. |  No noSQL a consistência nem sempre é mantida entre os pontos de distribuição de dados, levando em conta o teorema de CAP, que diz que não é possível obter três característica em simultâneo, e que os noSQL focam em disponibilidade e tolerância a falhas. |
| **Disponibilidade** | Modelo relacional não suporta uma demanda alta de informações. | Modelo não relacional, suporta uma demanda alta de informações. |

Fonte: Elaborada pelos autores.

**CONCLUSÃO**

Com o desenvolvimento do presente trabalho, foram identificados muitos pontos positivos e negativos em relação aos Sistemas gerenciadores de banco de dados PostgreSQL e MongoDB.

Chegou-se a conclusão, da mesa forma de que cada sistema tem suas características próprias e atende a um propósito diferente, empresas que tenham um volume de dados mais reduzido e precisam de maior confiabilidade, o PostgreSQL é uma ótima opção, porém se a empresa que o utiliza, trabalha com grandes quantidades de dados e precisa de rapidez na hora do acesso, sem a preocupação da ocorrência mais frequente de bugs esse SGBD pode não atender a essas necessidades. Do outro lado temos o MongoDB, um SGBD não relacional orientado a documentos, que é uma boa pedida para pessoas e empresas que trabalham com grandes volumes de dados, e precisam o acessar de forma mais rápida e fácil, principalmente quando esse armazenamento é feito na nuvem, o MongoDB PE se apresenta como uma boa opção.

Ao longo do processo de produção deste trabalho, chegou-se a conclusão por fim, de que não existe um SGBD melhor ou pior, quem necessita dos serviços oferecidos pelos sistemas, deve fazer um estudo antes de contratar o serviço, analisando quais são suas necessidades no momento, e chegando a uma conclusão em relação a qual Sistema gerenciado de Banco de Dados se encaixa melhor no tipo, e quantidade de dados a serem trabalhados.

Apresentar de maneira lógica e clara as principais conclusões do trabalho de acordo com os resultados discutidos no item anterior.

**REFERÊNCIAS BIBLIOGRÁFICAS**

Carvalho, Vinicius. **PostgreSQL Banco de dados para aplicações web modernas.** São Paulo, Editora Casa do Código,2017.

Elmasri,Ramez; B. Navathe, Shamkant. **Sistemas de Banco de dados.** São Paulo, Editora Afiliada,2006.

Hows, David.; Membrey, Peter.; Plugge, Eelco. **Introdução ao MongoDB**. São Paulo, Editora Novatec Editora, 2019.

Milani, Andre. **PostgreSQL guia do programador.** Rio Janeiro, Editora Novatec, 2008.

[O. Obe](https://www.pdfdrive.com/search?q=Regina+O.+Obe), Regina;  [S. Hsu](https://www.pdfdrive.com/search?q=Leo+S.+Hsu), Leo. **PostgreSQL: Up and Running: A Practical Guide to the Advanced Open Source Database.** Gravenstein Highway North, Sebastopol, Editora Melanie Yarbrough United States of America, 2017.

Secco, Ricardo Ribeiro; Silva, Francisco Assis da; Maracci, Francisco Virginio; Pazoti, Mário Augusto; **Análise comparativa entre o banco de dados Cassandra (modelo noSQL) e o PostgreSQL (modelo Relacional) em duas diferentes organizações Empresariais.** Presidente Prudente, 2016.

Cuer, Eder dos Santos; **Comparação de desempenho de bancos de dados SQL e NOSQL.** Marilia, 2014. Disponível em: <<https://aberto.univem.edu.br/bitstream/handle/11077/999/Eder%20dos%20Santos%20Cuer.pdf?sequence=1&isAllowed=y>>. Acesso em 28/10/2020.

Silva, Fabio; **como surgiu o MongoDB.** Dezembro, 2013. Disponível em: <<https://fabiosilva.com.br/2014/08/07/como-surgiu-o-mongodb/>>. Acesso em: 18/10/2020.

Hows, David; Membrey, Peter; Eelco Plugge. **Introdução ao MongoDB.** São Paulo. Editora Novatec, 2019. Disponível em: <https://books.google.com.br/books?id=MvPCDwAAQBAJ&pg=PT50&dq=mongodb&hl=ptBR&sa=X&ved=2ahUKEwirvcmp1sPsAhWtIrkGHchVDPoQ6AEwAnoECAYQAg#v=onepage&q=mongodb&f=false>. Acesso em: 12/11/2020.