

jeferson carvalho augusto

Segurança em banco de dados nosql

são Gonçalo

2019

jeferson carvalho augusto

segurança em banco de dados nosql

Projeto apresentado ao Curso de bacharel em Ciência da Computação da Instituição de ensino Anhanguera.

Orientador: Giomar Oliver Sequeiros Olivera

São Gonçalo

2019

**SUMÁRIO**

[1 **INTRODUÇÃO** 1](#_Toc503358475)

2 [**O PROBLEMA** 2](#_Toc503358476)

3 [**OBJETIVOS** 3](#_Toc503358477)

[3.1 OBJETIVO GERAL 3](#_Toc503358478)

[3.2 OBJETIVOS ESPECÍFICOS 3](#_Toc503358479)

[4 **JUSTIFICATIVA** 4](#_Toc503358480)

[5 **FUNDAMENTAÇÃO TEÓRICA** 5](#_Toc503358481)

[6 **METODOLOGIA** 8](#_Toc503358482)

[7 **CRONOGRAMA DE DESENVOLVIMENTO** 9](#_Toc503358483)

[**REFERÊNCIAS** 10](#_Toc503358484)

1. introdução

Para entender a importância de um banco de dados a melhor definição pode-se ver no livro de Elmasri et al (2010, p. 2), onde cita um banco de dados como ferramenta essencial para a sociedade nos dias atuais. São estruturas lógicas que permite armazenar, consultar e distribuir dados, seu armazenamento é feito de forma estruturada a fim de facilitar o acesso autorizado. Pode-se dizer que a informação é o bem mais precioso de qualquer organização, sejam cartões de crédito, vendas, fotos, vídeos, etc. Todos são armazenados em algum tipo de banco de dados e com toda essa importância precisamos garantir que ninguém sem autorização tenha acesso aos dados.

Com ênfase nos bancos de dados não relacionais (que serão chamados de NoSQL) devido ao seu grande crescimento nos últimos anos, será desenvolvido uma pesquisa qualitativa com a finalidade descritiva a fim de entender se é possível operar essa nova tecnologia de banco de dados NoSql para manter níveis aceitáveis de segurança. Será abordado também o funcionamento dos bancos NoSQL e seus pontos fracos com relação a segurança

Para responder essa questão serão utilizados como fonte de pesquisa livros dos principais autores da área de banco de dados e artigos científicos. Todo o assunto abordado será compreendido a um período anterior a 10 anos considerando a data de entrega desse trabalho.

É esperado que essa pesquisa contribua com o meio acadêmico esclarecendo sobre a importância dos bancos de dados NoQSL e pontos que podem ser observados e até desenvolvidos.

1. PROBLEMA

É possível operar essa nova tecnologia de banco de dados NoSql para manter níveis aceitáveis de segurança?

1. OBJETIVOS

3.1 OBJETIVO GERAL

# Demonstrar brechas de segurança em sistemas NoSQL.

3.2 OBJETIVOS ESPECÍFICOS

* Analisar o funcionamento dos bancos de dados NoSQL mais usados no mercado.
* Ressaltar brechas de segurança que possam existir nos sistemas analisados.
* Demonstrar soluções para compensar falhas de segurança que possam existir nos sistemas discutidos.

1. JUSTIFICATIVA

A proposta desse projeto é para esclarecer ao leitor sobre os riscos que corremos quando optamos por usar um banco de dados NoSQL. A tecnologia evoluiu rapidamente e fez-se necessário criar esse novo modelo de banco de dados para suprir a deficiência no modelo anterior (banco de dados relacional), mas como é um conceito relativamente novo, há algumas falhas que precisam ser corrigidas ou contornadas.

A informação é um bem que possui um valor inestimável e precisa ser muito bem protegida em algumas organizações podemos observa-la como a principal ferramenta para tomada de decisões. Em qualquer sistema quando algum indivíduo explora uma e falha e posteriormente há vazamento de informações, normalmente há um grande impacto para o negócio em questão, sendo assim é de suma importância que seja do conhecimento de todos qualquer tipo de brecha nos sistemas para que os dados sejam protegidos da melhor forma possível.

Como o conceito do banco de dados não relacional está com expectativa de grande crescimento no mercado, além de ser muito recente se comparado aos bancos de dados relacionais, não existem muitas informações disponíveis para os profissionais dessa área com orientações sobre a segurança do serviço.

No meio acadêmico para profissionais em formação que têm o objetivo de aprender a utilizar um banco de dados o conteúdo é ministrado focando o ensino em todos os processos de funcionamento da ferramenta, porem existem brechas que são transparentes nesse primeiro contato e podem ou não serem notadas futuramente com o uso intenso, onde haverá melhor aprendizado da ferramenta. O conteúdo desse artigo é fundamental para todos os alunos de áreas correlatas a computação para complementar o conhecimento sobre banco de dados não relacionais

1. **Fundamentação teórica**

Nesse tópico será introduzida a base teórica sobre a história dos bancos de dados relacionais e NoSQL os avanços tecnológicos em relação a forma de armazenamento de dados, as principais falhas relacionadas a segurança e as soluções encontradas no ambiente NoSQL.

Sempre houve a necessidade de armazenamento de dados, voltando algumas décadas e analisando os sistemas de armazenamento de dados, inicialmente, por volta da década de 1960 o armazenamento de dados era feito de forma hierárquica utilizando uma estrutura de arvore, mas com o avanço tecnológico foram surgindo outros métodos de armazenar dados. Analisando os estudos de Morgado (2016), sabemos que a partir de 1970 surgiu o banco de dados relacional. Esse novo conceito de banco de dados possui um sistema de armazenamento em tabelas, colunas e linhas de forma que cada linha possua um relacionamento com a linha de outra tabela. Esse mecanismo de armazenamento mostrou-se muito eficiente e durou muito anos, até hoje ainda é o modelo mais utilizado no mercado. Mais a frente em 1990 surgiu o banco de dados NoSQL, porem não teve uma aceitação logo no seu começo e os modelos relacionais continuaram a dominar o mercado.

A partir de ano 2000 devido à deficiência do modelo relacional em administrar uma massiva quantidade de dados e a sua incompatibilidade de impedância, ou seja, a diferença entre o modelo relacional e as estruturas de dados na memória, começou a aumentou o interesse pelo modelo de banco de dados NoSQL.

“Esse aumento ocorria em muitas dimensões. Websites começaram a registrar atividades e estruturas de uma forma muito detalhada. Surgiram grandes conjuntos de dados: links, redes sociais, atividades em logs, dados de mapeamento. Devido ao crescimento na quantidade de dados, houve um aumento também no número de usuários.” (SADALAGE; FOWLER, 2013, p. 32).

Ainda segundo Sadalage, et al (2013), a história do banco de dados NoSQL foi focada nos grandes volumes de dados para processamento em clusters. Essa era a necessidade na época então assim foi feito. Segundo Morgado, et al (2016), como não há um mecanismo de segurança eficiente para os bancos de dados NoSQL, essa responsabilidade fica a cargo dos clientes defender a base de dados.

Os bancos de dados NoSQL possuem diferentes modelos de armazenar dados, são eles: grafos, documentos, colunas e chave-valor. Cada modelo possui sua peculiaridade para receber um conjunto de dados. No mercado existem diversas empresas que desenvolveram servidores de banco de dados NoSQL e cada um possui uma característica diferente, mas no geral esse paradigma tem como principais características facilitar a escalabilidade horizontal do sistema, permitir facil implementação, ausência de esquema ou esquema flexível, suporte nativo a replicação (*Master-Slave* ou *Multi-Master*) e fácil acesso aos dados armazenados.

Apesar das suas vantagens, pelo fato de ser uma tecnologia recente se comparado ao já consolidado modelo relacional, existem muitas dúvidas sobre a eficácia da segurança. A crescente utilização de bancos de dados não relacional deve ser acompanhada de atualizações para que o serviço seja eficiente.

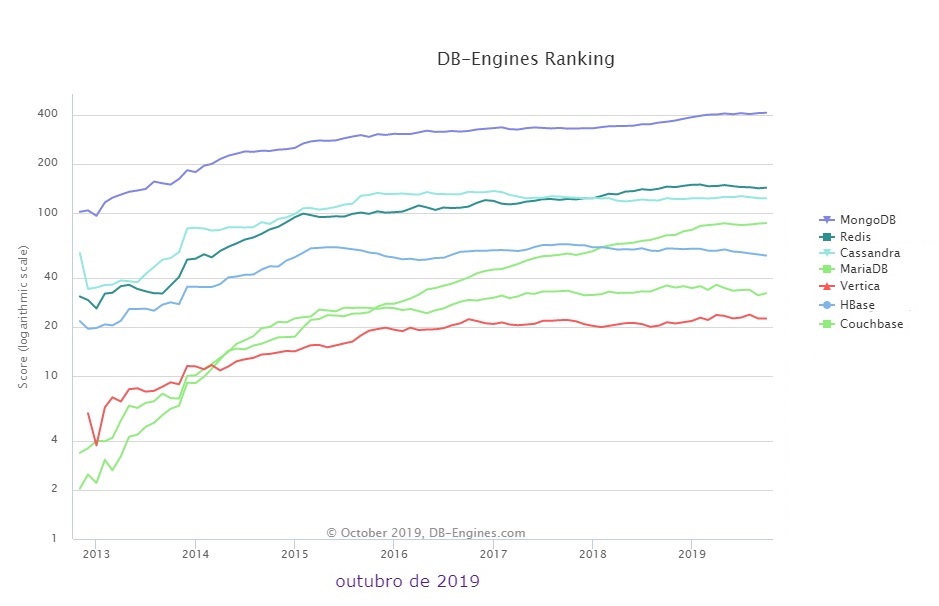


Figura - Estatística de utilização dos Bancos NoSQL

Fonte <https://db-engines.com/en/ranking_trend> (acessado em 18/10/19)

Na Figura 1 temos um gráfico de utilização dos bancos de dados NoSQL em função dos anos. Foi analisado um período de 2013 até 2019, onde podemos notar o crescente avanço da utilização dos bancos de dados NoSQL especialmente o serviço MongoDB teve um grande crescimento nos últimos anos. Segundo Cardoso 2014, o paradigma NoSQL vem ganhando ênfase como a nova opção que está em crescimento focando no problema e não na padronização.

Em relação a segurança em um ambiente NoSQL já foi analisado por Okman et al (2011), foi observado que os principais problemas eram a falta de criptografia dos dados ao serem inseridos no banco de dados, fraca autenticação tanto por parte do cliente, quanto por parte do.

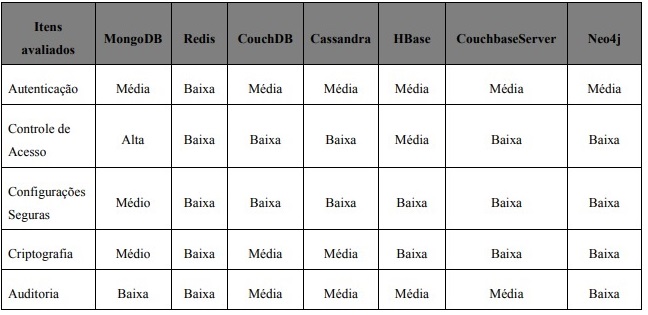


Figura - Níveis de segurança

Fonte: ZAHID et al (2014).

Na Figura 2 temos o comparativos feito por ZAHID et al (2014), sobre alguns dos principais bancos NoSQL e suas características quanto a segurança. Podemos confirmar que é predominante o nível de classificações média e baixa em todos os bancos de dados listados

1. **Metodologia**

Propõe-se nesse projeto uma pesquisa qualitativa com a finalidade descritiva acerca da segurança nos bancos de dados NoSql mais utilizados no mercado. Para desenvolver esse conteúdo será utilizado como fonte bibliográfica artigos científicos (coletados em bibliotecas virtuais como a “IEEE Xplore”), páginas da internet em geral, mas será melhor aproveitado o conteúdo do “scholar.google.com” e livros publicados por autores como Ramez Elmasri e Shamkant Navathe.

A partir das fontes informadas será explorada a história da motivação para que surgisse esse paradigma NoSQL, em seguida será abordado as falhas conhecidas e as melhores praticas para superar essas falhas.

Para que o leitor possa ter um melhor aproveitamento do tema, apesar do trabalho abordar toda a história do banco de dados NoSql, os artigos utilizados para essa pesquisa foram publicados na ultima década. É um grande período, mas irá garantir o melhor entendimento sobre o assunto analisando mais conteúdo.

Para fazer a busca desse trabalho temos os seguintes descritores: NoSql, Big Data, banco de dados, modelo relacional, not only SQL, brechas, base de dados, segurança, SQL e falhas.

1. **Cronograma**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ATIVIDADES** | **2019** | | | | | | **2019** | | | | | | **2020** | | | | | |
| Jan | Fev | Mar | Abr | Mai | Jun | Jul | Ago | Set | Out | Nov | Dez | Jan | Fev | Mar | Abr | Mai | Jun |
| Escolha do tema. Definição do problema de pesquisa |  |  |  |  |  | **x** |  |  |  |  |  |  |  |  |  |  |  |  |
| Definição dos objetivos, justificativa. |  |  |  |  |  |  |  |  | **x** |  |  |  |  |  |  |  |  |  |
| Definição da metodologia. |  |  |  |  |  |  |  |  | **x** |  |  |  |  |  |  |  |  |  |
| Pesquisa bibliográfica e elaboração da fundamentação teórica. |  |  |  |  |  |  |  |  |  | **x** |  |  |  |  |  |  |  |  |
| Entrega da primeira versão do projeto. |  |  |  |  |  |  |  |  | **x** |  |  |  |  |  |  |  |  |  |
| Entrega da versão final do projeto. |  |  |  |  |  |  |  |  |  |  | **x** |  |  |  |  |  |  |  |
| Revisão das referências para elaboração do TCC. |  |  |  |  |  |  |  |  |  |  |  |  | **x** |  |  |  |  |  |
| Elaboração do Capítulo 1. |  |  |  |  |  |  |  |  |  |  |  |  | **x** |  |  |  |  |  |
| Revisão e reestruturação do Capítulo 1 e elaboração do Capítulo 2. |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** |  |  |  |
| Revisão e reestruturação dos Capítulos 1 e 2. Elaboração do Capítulo 3. |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** |  |  |  |
| Elaboração das considerações finais. Revisão da Introdução. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** |  |  |
| Reestruturação e revisão de todo o texto. Verificação das referências utilizadas. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** |  |  |
| Elaboração de todos os elementos pré e pós-textuais. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** |  |
| Entrega da monografia. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** |
| Defesa da monografia. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **x** |

Referências utilizadas

SADALAGE,Pramod J.; FOWLER, Martin- **No SQL Essencial** - Um guia conciso para o mundo Emergente da Persistência Poliglota-São Paulo: Novatec, 2013. 216p.

ELMASRI, Ramez; NAVATHE, Shamkant B., Sistemas de banco de dados 6ª edição-São Paulo: Pearson 2010. 770 p

MORGADO, Cláudia. MODELO DE SEGURANÇA PARA BANCO DE DADOS ORIENTADO A GRAFOS. **Tese (Mestrado em Tecnologia)** – Universidade Estadual de Campinas Faculdade de Tecnologia, Limeira 2016. 96 p Disponível em:<<http://repositorio.unicamp.br/bitstream/REPOSIP/320823/1/Morgado_Claudia_M.pdf>>. Acesso em 18/10/19.

CARDOSO, Rodrigo Aniceto; FREIRE, Renê Xavier-Um Estudo Sobre a Utilização do Banco de Dados NoSQL Cassandra em Dados Biológicos. 2014. **Monografia (Bacharelado em Ciência da Computação)**- Universidade de Brasília, Brasília 2014.50 p Disponível em:<<http://bdm.unb.br/bitstream/10483/7927/1/2014_RodrigoCardosoAniceto_ReneFreireXavier.pdf> >. Acessado em 19/11/19.

OKMAN, Lior; GAL-OZ, Nurit; GONEN, Yaron; GUDES, Ehud; ABRAMOV, Jenny - *Security Issues in NoSQL Databases.* 2011 7 f. Artigo científico- **Digital Library** **IEEE Xplore**, Changsha, China 2012. Disponivel em: < <https://ieeexplore.ieee.org/abstract/document/6120863>>. Acessado em 15/11/19.

ZAHID, A.; MASOOD, R. SHIBLI, M.A.; “Security of Sharded NoSQL Database: A Comparative Analysis” . Conference on Information Assurance and Cyber Security. **Digital Library IEEE Xplorer**. 2014. Disponivel em: < <https://ieeexplore.ieee.org/abstract/document/6861323> >. Acessado em 21/10/19.