IV SIMTEC
Simpósio de Tecnologia da Fatec Taquaritinga

Uma Comparação entre MongoDB E CouchDB

A Comparison between MongoDB and CouchDB

Eder Carlos Salazar Sotto – eder.sotto@fatectq.edu.br
Faculdade de Tecnologia de Taquaritinga – Taquaritinga – São Paulo – Brasil
Kevin Lucas Germano – k.germano55@gmail.com
Faculdade de Tecnologia de Taquaritinga – Taquaritinga – São Paulo – Brasil

RESUMO

Este artigo apresenta uma comparação de performance entre os SGBDs *NoSQL* MongoDB e CouchDB. O estudo de caso realizado apresenta os tempos de execução de conjuntos de operação de inserção, busca, atualização e exclusão (ações estas denominadas CRUD na linguagem dos bancos de dados), dos dois sistemas de gerenciamento de banco de dados, permitindo obter estas métricas de maneira quantitativa, e apresentá-las em forma de tabelas comparativas para cada operação, comparando a performance de ambos os *SGDBs* executando o mesmo conjunto de operações. Pelos resultados obtidos no ambiente de teste criado, é evidente a diferença entre o tempo de inserção, busca, atualização e exclusão entre MongoDB e CouchDB, permitindo assim concluir que o MongoDB é o mais performático para ser utilizado em conjunto com a linguagem PHP no desenvolvimento de aplicações Web.

Palavras-chave: MongoDB. CouchDB. Tempo. PHP.

ABSTRACT

This article shows an introduction to NoSQL, MongoDB and CouchDB, terms related to the non-relational database environment. Its main point is to show the time of each insertion, search, update and delete operation (actions called CRUD in the database language) of two database management systems, named MongoDB and CouchDB. Keeping in mind that database management systems have different times for their operations, it has been discovered how much time it takes to enter, find, change, and delete, thereby creating a test environment needed to obtain the data in a quantitative way and presents them in the form of tables for each operation to make a comparison between both SGDBs. By the results obtained in the created test environment, it is clear the difference between the insertion, search, update and exclusion time between MongoDB and CouchDB, thus making MongoDB the recommended one to be used in along the PHP language for web application development.

Keywords: MongoDB. CouchDB. Time. PHP.

1 INTRODUÇÃO

Muitas empresas coletam uma grade quantidade de dados de seus clientes, como forma de pesquisa, para suas ações futuras através de mineração de dados (Joel De Bortoli,2012).

Assim se faz necessário um sistema de gerenciamento de banco de dados com um tempo de inserção, busca, alteram e exclusão rápido e eficiente. Neste sentido, este artigo visa verificar o tempo para que essas operações entre os dois *SGDBs* (sistema de gerenciamento de banco de dados) sendo eles MongoDB e CouchDB, apresentando-as de maneira quantitativa em formas de tabelas e gráficos com seus devidos resultados.

1.1 OBJETIVO GERAL

Demonstrar uma comparação de desempenho para medir e analisar o tempo de execução de inserir, atualizar, remover e consultar dados armazenados em dois diferentes sistemas de gerenciamento de banco de dados não-relacionais: MongoDB e CouchDB, utilizado a linguagem PHP.

1.2 OBJETIVOS ESPECÍFICOS

- Propor cenários de testes utilizando a linguagem PHP para avaliar o desempenho de ambos os sistemas de gerenciamento de banco de dados em relação ao tempo de execução.
- Analisar o desempenho dos bancos de dados escolhidos
- Apresentar os resultados obtidos em cada cenário.

1.3 ESTRUTURA DO TRABALHO

Este trabalho foi Organizado da seguinte forma.

- Na seção 1 foram apresentados a motivação, uma breve introdução do tema abordado e o objetivo do trabalho.
- Na seção 2 são abordadas as principais características do *NoSQL*, MongoDB e CouchDB.

- Na seção 3, desenvolve-se o estudo de caso, onde são definidos cenários de teste e realizadas as execuções dos testes.
- Na seção 4 é feita avaliação dos resultados obtidos, bem como, a comparação dos entre cada sistema de gerenciamento de banco de dados.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentadas as principais características de como um banco de dados pode ser definido, o que é *NoSQL* e uma introdução à SGDBs o MongoDB e CouchDB

2.1 BANCO DE DADOS

Um banco de dados pede ser definido como um local onde são armazenados dados necessários para a manutenção das atividades de organização, empresas, sites, possuindo vários níveis de interação com o mundo real.

Segundo Date (2004, p. 6), "Um banco de dados é uma coleção de dados persistentes, usada pelos sistemas de aplicação de uma determinada empresa", sem então, um local que possibilita o acesso à informais, para os gerenciamentos dos dados de uma empresa.

Para Elmasri e Navathe (2011, p. 3), na expressão Banco de Dados, subentende-se que:

> Um banco de dados representa algum aspecto do mundo real, às vezes chamado de minimundo ou de universo de discurso (UoD – Universe of Discourse). As mudanças no minimundo são refletidas no Banco de Dados. Um banco de dados é uma coleção logicamente coerente de dados com algum significado inerente. Uma variedade aleatória de dados não pode ser corretamente chamada de banco de dados. Um banco de dados é projetado, construído e populado com dados para uma finalidade específica. Ele possui um grupo definido de usuários e algumas aplicações previamente concebidas nas quais esses usuários estão interessados.

Para tanto, o *Software* Gerenciador de Banco de Dados disponibiliza recursos para definir, construir, manipular, compartilhar, proteger e manter bancos de dados (ELMASRI; NAVATHE, 2011).

2.2 O NOSQL

A origem do termo *NoSQL*, e a ideia de um SGBD não-relacional começou em 1998, sem também um nome de um banco de dados. Seu criador Carlo Strozzi, criou seu sistema de gerenciamento de banco de dados que usa um paradigma de operador de fluxo para acessar dados em vez de sintaxe SQL (TAMANE, 2016).

O autor do termo *NoSQL,* Carlo Strozzi afirma que "é completamente distinto do modelo relacional e, portanto, deveria ser mais apropriadamente chamado "NoREL" ou algo que produzisse o mesmo efeito" (STROZZI, 2007).

"*NoSQL* é um movimento que promove soluções de armazenamento de dados não relacionais." (PORCELLI, 2011).

Os bancos de dados NoSQL nasceram das necessidades de atender os requisitos, que os bancos de dados relacionais não podem atender. Em sua maioria, essas bases representam algumas características que vale a pena ressaltar, como: alta performance, escalabilidade, replicação, suporte a dados estruturados, grafos e sub-colunas (NASCIMENTO, 2010).

A necessidade de uma performance superior e de alta escalabilidade vem do gasto com os atuais bancos de dados relacionais que são muito restritos sendo necessária a distribuição vertical de servidores, sendo assim, elevando a um custo maior conforme a necessidade do uso mais dados, de mais memória e demais disco um servidor (NASCIMENTO,2010).

Nesta área o NoSQL tem uma grande facilidade na distribuição horizontal, ou seja, não necessita de servidores de alta performance.

Alguns dos bancos mais utilizados nos dias de hoje, são NoSQL como o MongoDB.

2.2 MONGODB

MongoDB é um SGDB orientado a documentos de alta performance. Como citado no capítulo anterior, o MongoDB é classificado como um banco de dados NOSQL, tendo como uma de sua característica de ser um software de código aberto (DOCS.MONGODB, 2017).

Os bancos de dados orientados a documentos são bastante diferentes dos tradicionais bancos de dados relacionais. Em vez de armazenar dados em estruturas rígidas, como tabelas, eles os armazenam em documentos vagamente definidos (CUNHA, 2011).

O MogoDB foi desenvolvido na linguagem de programação C++ e tem drivers disponíveis para as seguintes linguagens de programação C, C++, C#, Java, Perl, PHP, Python, Ruby e Scala (DOCS.MONGODB, 2017).

Inicialmente, o MongoDB foi desenvolvido como um componente de serviço pela empresa 10gen em outubro de 2007, somente no ano de 2009 passou a ser um software *open-source* (SOLAGNA; LAZZARETTI, 2016).

2.3 COUCHDB

Para Joe Lennon (2009) CouchDB é:

> CouchDB é um sistema de software livre de gerenciamento de banco de dados orientado a documentos que pode ser acessado através da API de *JavaScript Object Notation* (JSON) RESTful. O termo "Couch" é um acrônimo para "Cluster Of Unreliable Commodity Hardware", que reflete a meta do CouchDB de ser extremamente escalável, oferecendo alta disponibilidade e confiabilidade, mesmo ao executar em hardware que está geralmente sujeito à falhar. O CouchDB foi gravado originalmente em C++, mas em abril de 2008, o projeto foi movido para a plataforma Erlang OTP devido à sua ênfase em tolerância a falhas.

Atualmente o CouchDB conta com *drivers* para várias linguagens de programação, sendo elas: C, C++, C#, Java, Perl, PHP, Python, Ruby e Scala(DOCS.COUCHDB.ORG. 2017).

2.4 PHP

PHP é uma linguagem de script embutida no HTML. Muito da sua sintaxe é emprestada de C, Java e Perl com algumas características específicas do PHP juntas. O objetivo da linguagem é permitir que desenvolvedores web escrevam páginas geradas dinamicamente de forma rápida (SECURE.PHP.NET,2017).

3 DESENVOLVIMENTO

Foi desenvolvido um pequeno software para realização dos testes de performance, onde foi possível a criação de um cenário fictício que serviu como base para a realização dos testes.

A partir desse cenário foram feitas as modelagens dos dados para cada *SGBD,* utilizando a mesma estrutura em ambos os sistemas.

O objetivo dos testes foi comparar o desempenho de cada SGBD através de inserção, alteração, remoção e alteração, visando saber qual dos *SGDBs* MongoDB e CouchDB tem o menor tempo de resposta para execução para as ações citadas.

A Ilustração 1, a presenta a estrutura da *collection* do MongoDB utilizada para os testes.

Ilustração 1 – Estrutura Mongo utilizada nos testes.

```
db.telefones."Pode ser utilizado insert, find,Update, remove"( {
                    nome: "JJ500",
                    Telefone: "3333-3332"
            }
  )
```

Fonte: O Autor (2017)

Na Ilustração 2, a presenta a estrutura utilizada no CouchDB.

Ilustração 2 – Estrutura CouchDB

```
{
  "_id" : "5b5b367ff3a6b6d6fc6a8e9a95002852" ,
  "Nome" : "JJ500" ,
  "Telefone" : "3333-3332"
}
```

Fonte: Autor (2017)

3.1 AMBIENTES DE APLICAÇÃO DOS TESTES

Na Tabela 1 é descrito o equipamento utilizado na aplicação dos testes.

Tabela 1 – Equipamento utilizado nos testes

Equipamento 1	Descrição
Processador	Intel Core i5 2.30 GHz
Memória RAM	8G
HD	1T
Sistema operacional	Windows 10

Fonte: O Autor (2017)

A Tabela 2 apresenta os softwares utilizados.

Tabela 2 – Software utilizados nos testes e suas respectivas versões

Software	Versão
MongoDB	3.4
CouchDB	2.1.0
PHP	7.1
Xampp	3.2.2
Composer	1.5.2
DLL Para xampp MongoDB	1.3.0
API PHP para MongoDB	mongodb/mongo-php-library
API PHP pata CouchDB	PHP-on-Couch/PHP-on-Couch

Fonte: O Autor (2017)

3.2 OS TESTES

Os testes foram divididos em 4 categorias sendo: inserção, consulta, atualização e exclusão.

Foram executados em loops de 1, 100, 1.000 e 10.000 repetições para buscas, bem como, as inserções. As atualizações e exclusões foram executadas sem a utilização de loops, sendo assim, feita em uma única execução.

Todos os testes foram executados três vezes cada, criado assim uma média, a qual permitiu a comparação dos resultados obtidos de cada SGDB.

4 RESULTADOS E DISCUSSÃO

São descritos nesta seção os resultados obtidos através da realização dos testes, onde os SGDBs, através das operações de inserção, seleção, atualização e exclusão, retornaram o tempo de execução de cada ação.

Os testes medem a diferença entre os tempos de execução de todas essas operações nos SGDBs MongoDB e CouchDB, com a finalidade de evidenciar que a diferencia de tempo para cada execução das ações citadas anteriormente.

Para a obtenção das medias foi criado um pequeno script php com a utilização da função microtime.

4.1 REULTADOS MONGODB

A seguir são mostrados os resultados obtidos com o SGBD MongoDB.

Os resultados obtidos na inserção são mostrados na Tabela 3.

Tabela 3 – Resultados da Inserção com MongoDB em segundos

N. Inserção	1	10	100	1.000	10.000
1ª Execução em segundos	0,0028970	0,0084180	0,0626828	0,5367460	4,2534258
2ª Execução em segundos	0,0129501	0,0096790	0,0656988	0,8804371	5,5574191
3ª Execução em segundos	0,0061190	0,0137341	0,0644569	0,5831778	2,6881552
Media em segundos	0,0073220	0,0106103	0,0642795	0,6667869	4,1663333

Fonte: O Autor (2017)

Na Tabela 4 são mostrados os resultados da busca. Em uma *collection* com 1.000 registros

Tabela 4 – Resultados da Buscas com MongoDB em segundos

N. Busca	1	10	100	1.000	10.000
1ª Execução em segundos	0,0064208	0,0185160	0,0376162	1,2514202	12,043915
2ª Execução	0,0092270	0,0205109	0,1303579	1,2024118	14,740020

em segundos					
3ª Execução em segundos	0,0079121	0,0196480	0,1239819	1,4203879	5,3154330
Media em segundos	0,0078533	0,0195583	0,0973186	1,2914066	10,699789

Fonte: O Autor (2017)

Na Tabela 5 são mostrados resultados os dados de alteração.

Tabela 5 – Resultados da Alteração com MongoDB em segundos

N. Alteração	1
1ª Execução em segundos	0,0069289
2ª Execução em segundos	0,0066189
3ª Execução em segundos	0,0067579
Media em segundos	0,0067685

Fonte: O Autor (2017)

Na Tabela 6 são mostrados os dados de exclusão.

Tabela 6 – Resultados da Exclusão com MongoDB em segundos

N. Exclusão	1
1ª Execução em segundos	0,2543458
2ª Execução em segundos	0,0046021
3ª Execução em segundos	0,0046010
Media em segundos	0,0878496

Fonte: O Autor (2017)

4.2 RESULTADOS COUCHDB

A seguir são mostrados os resultados obtidos com o SGDB CouchDB.

Os resultados obtidos na inserção do primeiro equipamento são mostrados na Tabela 7.

Tabela 7 - Resultados da Inserção com CouchDB em segundos

N. Inserção	1 vez	10 vezes	100 vezes	1.000 vezes	10.000 vezes

1ª Execução em segundos	0,2760151	2,9091660	28,740644	281,18008	2884,3909
2ª Execução em segundos	0,3160181	3,0911769	29,103664	281,18508	2884,4109
3ª Execução em segundos	0,3480200	2,8911650	29,226675	285,29631	2886,3813
Média em segundos	0,31335106	2,9638359	29,023661	282,55382	2885,0610

Fonte: O Autor (2017)

Na Tabela 8 são mostrados os dados da busca dos dados. Em 1000 registros.

Tabela 8 – Resultados da Buscas com CouchDB em segundos

N. Busca	1 vez	10 vezes	100 vezes	1.000 vezes	10.000 vezes
1ª Execução em segundos	0,2610149	2,6701543	26,404513	263,09703	2638,4139
2ª Execução em segundos	0,2810158	2,5821475	26,188492	263,95609	2626,9112
3ª Execução em segundos	0,2660150	2,5931478	26,311504	265,22917	2629,6352
Media em segundos	0,2693485	2,6151498	26,301503	264,09409	2631,6534

Fonte: O Autor (2017)

Na Tabela 9 são mostrados os dados de alteração.

Tabela 9 – Resultados da Alteração com CouchDB em segundos

N. Alteração	1 vez
1ª Execução em segundos	0,2670553
2ª Execução em segundos	0,2533456
3ª Execução em segundos	0,2655790
Media em segundos	0.2619933

Fonte: O Autor (2017)

Na Tabela 10 são mostrados os dados de exclusão.

Tabela 10 – Resultados da Exclusão com CouchDB em segundos

N. Exclusão	1 vez
1ª Execução em segundos	0,2302433
2ª Execução em segundos	0,2452014

3ª Execução em segundos	0,2298359
Media em segundos	0,2350935

Fonte: O Autor (2017)

4.3 **COMPARAÇÃO DOS DADOS**

É apresentada na Tabela 11 a comparação entres as médias obtidas do MongoDB e CouchDB.

Tabela 11 – Comparação das médias do MongoDB e CouchDB

Nº da operação	SGDBs	Médias Inserções em segundos	Médias Buscas em segundos
1 vez	MongoDB	0,0073220	0,0078533
	CouchDB	0,31335106	0,2693485
10 vezes	MongoDB	0,0195583	0,0195583
	CouchDB	2,9638359	2,6151498
100 vezes	MongoDB	0,0642795	0,0973186
	CouchDB	29,023661	26,301503
1.000 vezes	MongoDB	0,6667869	1,2914066
	CouchDB	282,55382	264,09409
10.000 vezes	MongoDB	4,1663333	10,699789
	CouchDB	2885,0610	2631,6534

Fonte: O Autor 2017

Na Tabela 11, com os dados obtidos nas médias de inserções em segundos, o MongoDB apresenta média de execução muito superior à obtida com o CouchDB. Em quanto o MongoDB leva 0,0073220 segundos para inserir um registro no banco de dados, o CouchDB leva 0,3122516 segundos em média para fazer a mesma operação. O MongoDB continua superior ao CouchDB inserindo 10, 100, 1.000, 10.000 registros.

Ainda na Tabela 11, os dados obtidos nas médias de busca em segundos, o MongoDB se mantém superior ao CouchDB com os seguintes resultados. Busca em um banco de dados

com mil registros uma vez: 0,0078533, 10 vezes: 0,0195583, 100 vezes: 0,0973186, 1.000 vezes: 1,2914066 e 10.000 vezes: 10,699789.

Tabela 12 – Comparação das médias do MongoDB e CouchDB alteração e exclusão

Nº da operação	SGDBs	Médias Alteração em segundos	Médias Exclusão em segundos
1 vez	MongoDB	0.2619933	0.2619933
	CouchDB	0.2619933	0,2350935

Fonte: O Autor (2017)

Na Tabela 12, na média de alteração, mais uma vez o MongoDB supera o CouchDB e no último tópico da Tabela 12, na média de exclusão, o MongoDB supera por definitivo o CouchDB.

5 CONCLUSÃO

A comparação entre os dois *SGDBs* não-relacionais MongoDB e CouchDB permitiu verificar quanto tempo cada banco de dados leva para inserir, buscar, alterar e excluir os dados do banco utilizando uma aplicação Web em PHP. Seja qual for a operação de manipulação de dados, o MongoDB destacou-se com o menor tempo necessário para realizar estas operações. Pode ser possível que os *drivers* de conexão interfiram no desempenho da execução, portanto, seria de grande relevância para trabalhos futuros uma comparação de desempenho com outros drivers disponíveis em outras linguagens de programação, como Java e C#. No entanto, com os resultados já obtidos é possível concluir que o MongoDB possui desempenho consideravelmente superior se comparado ao CouchDB para realização de qualquer operação de manipulação de dados, quando utilizado em uma aplicação Web desenvolvida em PHP.

REFERÊNCIAS

CUNHA, T. M. de A. Escalabilidade de Sistemas com Banco de Dados NoSQL: um Estudo de Caso Comparativo com MongoDB e MySQL. 2011. 85 f. Centro Universitário da Bahia – Estácio, Salvador.

COUCHDB. 2017. Disponível em: http://docs.couchdb.org/en/2.1.0/ - Acessado em 4 outubro. 2017.

COMPOSER. 2017. Disponível em: https://getcomposer.org/doc/ - Acessado em 4 outubro.

DATE, J.C. Introdução a Sistemas de Banco de Dados. 2004. 7ª ed. Education. Pearson.

ELMASRI, RAMEZ; NAVATHE, SHAMKANT B. 2011 Sistemas de Bancos de Dados. Pearson, 6a. edição em português.

Jean Nascimento. 2010. NoSQL – você realmente sabe do que estamos falando?. Disponível em: https://imasters.com.br/artigo/17043/banco-de-dados/nosql-voce-realmente-sabe-do-que-estamos-falando?trace=1519021197&source=single - Acessado em 9 outubro.

JOE LENNON. 2009. Disponível em: https://www.ibm.com/developerworks/br/library/os-couchdb/index.html - Acessado em 7 outubro.

Joel De Bortoli.2012. Data Mining. Disponível em: http://www.joeldebortoli.com/2012/08/data-mining-mineracao-de-dados.html - Acessado em 7 outubro.

MONGODB. 2017. Disponível em: https://docs.mongodb.com/ - Acessado em 2 outubro. 2017.

MONGODB/MONGO-PHP-LIBRARY. 2017. Disponível em: https://github.com/mongodb/mongo-php-library. Acessado em 4 outubro. 2017

PHP-ON-COUCH/PHP-ON-COUCH**.** 2017. Disponível em**:** https://github.com/PHP-on-Couch/PHP-on-Couch**.** Acessado em 8 outubro. 2017.

PHP. 2017. Disponível em**:** http://php.net/manual/pt_BR/set.mongodb.php - Acessado em 4 outubro. 2017.

PORCELLI, ALEXANDRE. O que é NoSQL? . Disponível em: http://www.devmedia.com.br/o-que-e-nosql-java-magazine-86/18777Java Magazine, 86, jan. 2011. Acessado em: 10 de maio. 2017.

STROZZI.it. 2 de outubro de 2007. Disponível em**:** http://www.strozzi.it/cgi-bin/CSA/tw7/I/en_US/nosql/Home%20Page - Acessado em 9 outubro. 2017.

SECURE.2017. Informações Gerais. Disponível em: https://secure.php.net/manual/pt_BR/faq.general.php - Acessado em 9 outubro. 2017.

SOLAGNA E. A; LAZZARETTI A. T. 2016. UM ESTUDO COMPARATIVO ENTRE O MONGODB E O POSTGRESQL. Disponível em: http://painel.passofundo.ifsul.edu.br/uploads/arq/20160711180550101591419 8.pdf - Acessado em 4 outubro.

TAMANE, SHARVARI, Non-Relational Databases in Big Data. In: Italian Conference on Theoretical Computer Science, 2016, Udaipur, India.