

Comparação de Desempenho entre *Drivers* do MongoDB em Aplicações Node.js

OMITIDO

¹OMITIDO

OMITIDO

Abstract. *In recent years, the number of NoSQL database has grown significantly. Such databases do not have a common language for defining and manipulating data, so this role is assigned to APIs or drivers. There are cases where, for the same database, there are several solutions, which makes it difficult to understand the impacts that the choice of each option has. In this paper, we present a comparative study between two MongoDB drivers in Node.js applications, in which different execution scenarios of CRUD operations were defined for performance analysis based on some metrics. The results demonstrate that, under quantative analysis, MongoClient performs better in the evaluated scenarios.*

Resumo. *Nos últimos anos, o número de banco de dados NoSQL tem crescido significativamente. Tais bases não possuem uma linguagem comum para definição e manipulação dos dados, de modo que esse papel é atribuído a API's ou drivers. Há casos em que, para o mesmo banco de dados, existem diversas soluções, o que dificulta o entendimento dos impactos que a escolha de cada opção tem. Neste artigo, é apresentado um estudo comparativo entre dois drivers para MongoDB em aplicações Node.js, em que foram definidos diferentes cenários de execução de operações CRUD para análise de desempenho com base em algumas métricas. Os resultados demonstram que, sob análise quantitativa, o MongoClient tem melhor desempenho nos cenários avaliados.*

1. Introdução

Nos últimos anos, o crescimento no volume de dados mudou a utilização desses por empresas e organizações. Inicialmente, os dados eram considerados agentes passivos, relacionados às regras de negócio empresarial; tornaram-se potenciais oportunidades de lucro e obtenção de conhecimento através de processos de análise de informações.

A crescente quantidade de dados, fenomeno conhecido como *Big Data*, não somente requer maior espaço de armazenamento, mas uma mudança em sua organização com base em cada contexto, considerando características como volume, variedade, velocidade e valores [Ward and Barker 2013]. A arquitetura dos tradicionais Bancos de Dados Relacionais são baseadas nas propriedades ACID (*atomicity*, *consistency*, *isolation* e *durability*), contudo, em ambientes onde o *Big Data* se faz presente, a alta consistência afeta diretamente os aspectos de disponibilidade e eficiência, que são primordiais, devido ao alto volume, variedade e velocidade presente nesse fenomeno [González-Aparicio et al. 2016].

Nesse cenário, surgem os Banco de Dados *NoSQL* (*Not only SQL*) que em contrapartida promovem maior flexibilidade estrutural, escalabilidade, suporte a replicação e consistência eventual, seguindo o critério BASE (*basic, availability, soft-state e eventual consistency*) [Han et al. 2011]. Nos últimos anos, a popularidade dos Bancos de Dados *NoSQL* tem crescido [Cooper et al. 2010, Edlich 2015] proporcionando diversas soluções conforme características dos dados e do ambiente de cada aplicação.

Diferente do modelo relacional, os Bancos de Dados *NoSQL* não possuem uma linguagem comum entre eles que permita a realização de operações, como o SQL (*Structured Query Language*). Desse modo, cada banco provê uma interface nativa para manipulação dos dados. Contudo, uso de chamadas de sistema para execução de comandos dessas interfaces não são reconhecidas como uma boa prática, que pode incorrer em diversos problemas.

De modo a contornar essa situação, para os diferentes ambientes de desenvolvimento e linguagens de programação, *Drivers* tem sido desenvolvidos de modo a viabilizar a execução dos comandos no banco de dados. Em muitas situações, a decisão de qual combinação entre Banco de Dados *NoSQL* e *driver* a ser empregada pode ser um problema, devido a variedade de possibilidades, assim como o desconhecimento dos pontos positivos e negativos de cada solução.

Neste artigo, por meio de um estudo comparativo de desempenho, busca-se avaliar as duas principais soluções de *drivers* para o MongoDB ¹, respectivamente MongoClient ² e Mongoose ³, em ambientes de aplicação Node.js. O principal fator considerado para a realização dessa análise consiste na definição prévia de esquema para a manipulação dos dados em operações de CRUD (*create-read-update-delete*). Conceitualmente, os Bancos de Dados *NoSQL* não requerem essa predefinição, proporcionando flexibilidade. Contudo não existe nada que impeça sua utilização, principalmente quanto a respeito do desempenho; possibilitando algum impacto relevante.

A escolha do Banco de Dados MongoDB ocorreu devido a sua enorme popularidade recente, empregado em diversas aplicações e linhas de pesquisa; o qual consiste na principal opção dentre os Bancos de Dados que adotam a estratégia de armazenamento orientada a documentos. A respeito do ambiente Node.js, apesar de ser uma tecnologia recente, alguns trabalhos apontam a sua viabilidade no desenvolvimento de aplicações [Chaniotis et al. 2015]. Além disso, com essa escolha, tanto aplicação quanto banco de dados utilizam JavaScript, permitindo a elaboração de um sistema uniforme, em termos de linguagem de programação.

Na análise conduzida neste trabalho, o desempenho do Banco de Dados MongoDB atrelado a cada um dos *Drivers* investigados é analisado sobre perspectivas: de tempo de execução; tempo de uso de processamento; e consumo de memória. A análise é aplicada sob um conjunto de dados genérico, propiciando a identificação de algumas características.

O restante desse trabalho está organizado do seguinte modo: na Seção ?? é realizada uma revisão conceitual sobre banco de dados não-relacional, assim como sobre o

¹<https://www.mongodb.com/>

²<https://mongodb.github.io/node-mongodb-native/>

³<https://mongoosejs.com/>

MongoDB e seus *Drivers*. Na Seção ?? é apresentado o ambiente de execução Node.js assim como detalhes sobre uso de memória que são relevantes no estudo comparativo. A Seção ?? é apresentado o estudo comparativo realizado neste trabalho. Seção 4 são apresentados os resultados quantitativos obtidos, cuja discussão é elaborada na Seção 5. Em seguida, na Seção 6 e, por fim, na Seção 8 as considerações finais desse trabalho.

2. Fundamentação Teórica

Nesta seção, são discutidos alguns conceitos fundamentais para o trabalho: Bancos de Dados NoSQL, as características do MongoDB e dos *Drivers* MongoClient e Mongoose, organização de aplicações Node.js.

2.1. Banco de Dados NoSQL

Os bancos de dados NoSQL, foram desenvolvidos visando armazenar e processar grandes volumes de dados. Em linhas gerais os bancos de dados NoSQL são livres de esquematizações e mais propícios a lidar com dados não estruturados como e-mails, documentos e mídias sociais de maneira eficiente [Mohamed et al. 2014, Ramesh et al. 2016].

O termo NoSQL é comumente utilizado para se referir a uma ampla variedade de armazenamentos de dados nos quais as restrições de transação ACID foram suavizadas visando permitir melhor dimensionamento e desempenho horizontal [Rafique et al. 2018]. Os recursos gerais presentes nos bancos de dados NoSQL são sumarizados em: esquemas menos estruturados, suporte a operações de junção, alta escalabilidade, modelagem de dados simples com linguagem de consulta simples [Ramesh et al. 2016]. Os bancos de dados NoSQL foram categorizados em: armazenamento de documentos, famílias de colunas, chave/valor, gráficos e multimodais [González-Aparicio et al. 2016].

Este trabalho tem como foco a categoria orientada a documentos, a qual possui modelagem de dados estreitamente relacionados a programação orientada a objetos, deste modo, cada documento é visto como um objeto. O conceito de esquema nos Bancos de Dados orientado a documentos é dinâmico, uma vez que, cada documento pode conter campos distintos um dos outros, sendo útil na modelagem de dados não estruturados e polimórficos. Essa categoria permite consultas robustas, em que qualquer combinação de campos no documento pode ser realizada visando consultar dados [Patil et al. 2017]. Em termos de estruturação, seus dados são organizados em coleções de documentos, as quais utilizam uma estrutura semelhante a JSON (*JavaScript Object Notation*) ou XML (*Extensible Markup Language*).

2.2. MongoDB

O MongoDB [Membrey et al. 2011] é um Banco de Dados orientado a documentos que possui código-aberto. Embora não seja considerado um Banco de Dados Relacional, esse apresenta muitas funcionalidades providas por essa categoria, como ordenação, indexação secundária e consultas de intervalo.

A estruturação dos dados não acontece com base em tabelas com suas respectivas colunas e tuplas, mas sim em coleções de documentos. O banco de dados não impõe um esquema previo, entretanto, normalmente todos os documentos em uma coleção são de propósito semelhante ou relacionado [Kanade et al. 2014, Lutu 2015]. Há duas abordagens para modelagem de documentos:

- **Modelo de dados incorporado:** Os dados relacionados são incorporados em uma única estrutura ou documento. Esses esquemas são geralmente conhecidos como modelos sem normalização.
- **Modelo de dados normalizado:** Os dados possuem referências de documentos para registrar relacionamentos entre esses, mas a combinação de documentos deve ser feita diretamente no código-fonte da aplicação.

O armazenamento dos dados ocorre através da serialização de objetos JavaScript, também conhecidos JSON, cuja implementação interna utiliza uma codificação binária chamada BSON [BSONSpecification 2019]. O banco de dados MongoDB disponibiliza diversos *drivers* para linguagens de programação como Java, C++, C#, PHP e Python [Lutu 2015], assim como para aplicações baseadas em Node.js.

Nesse contexto, dentre os *Drivers* existentes, tem-se como destaque o **Mongo-Client** ⁴, consiste na solução oficial e nativa provida organização, provendo um conjunto de funcionalidades que permite a manipulação dos dados e uso de recursos avançados do sistema. Essa solução é caracterizada pela modelagem documentos-objeto (*ODM – Object-Document Modeler*) de modo implícito ao banco de dados.

Assim como o anterior, o **Mongoose** consiste também em um driver para MongoDB, porém esse provê a modelagem de dados utilizando um mecanismo semelhante ao mapeamento de dados em tabelas (*ORM – Object Relational Mapping*) utilizado em banco de dados relacionais [Mardan 2014], executando diversas tarefas de verificação e validação dos dados, como nulidade ou tipagem, previamente definidos por meio da elaboração de um esquema.

2.3. Node.js

Node.js é uma plataforma construída sob o ambiente de execução para JavaScript do navegador Google Chrome, para criação facilitada de aplicações de internet rápidas e escaláveis [NodeJS 2019]. Esse ambiente se baseia em mecanismo orientado a eventos de entrada e saída não-bloqueante, o que viabiliza a interação do usuário enquanto demais tarefas executam em segundo-plano, resultando em aplicações leves e eficientes.

Atualmente, pode ser observado um crescente número de projetos utilizando essa tecnologia, o que é comprovado pela popularidade da plataforma desde as suas versões iniciais, atrelada a crescente utilização da linguagem Javascript ⁵.

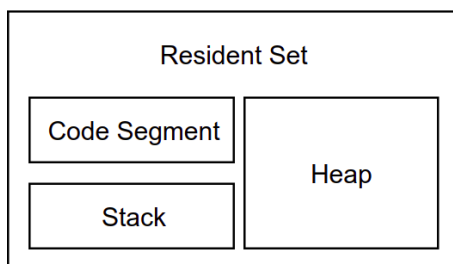


Figura 1. Organização de memória de um processo Node.js – Adaptado de [Gagliardi 2018].

⁴<https://mongodb.github.io/node-mongodb-native/index.html>

⁵<https://github.com/search?l=JavaScript&o=desc&q=stars%3A%3E1&s=stars&type=Repositories>

Toda aplicação Node.js em execução, assim como qualquer processo em geral, requer que uma região da memória seja reservada. A Figura 1 apresenta a organização da memória em processos Node.js.

A primeira região, que engloba as demais, é chamada *Resident Set*, essa corresponde a toda memória utilizada no processo. Em seguida tem-se a região *Code Segment*, a qual armazenada todas as instruções definidas para o programa. A próxima região *Stack* armazena todas as variáveis e estruturas de dados utilizadas durante o tempo de vida dessas. Por fim, tem-se a região *Heap*, a qual armazena dados específicos como objetos, strings e closures; em geral, cada processo aloca esse região com um tamanho predefinido, contudo essa pode ser utilizada apenas parcialmente ou em sua totalidade [Gagliardi 2018].

3. Composição do Experimento

Nesta seção é apresentado a composição do experimento realizado comparando os *drivers* MongoClient e Mongoose, conforme apresentado na Figura 2.

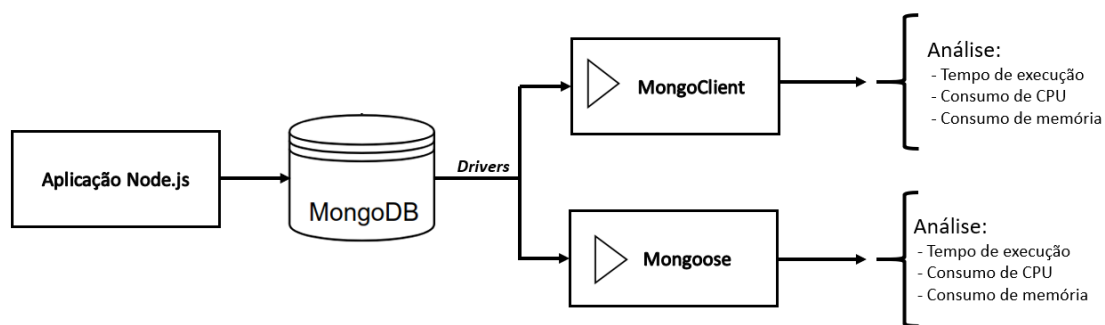


Figura 2. Experimento: Comparação entre os *Drivers* do MongoDB.

A ideia principal consiste em analisar ambos sob o desempenho médio, em termos de algumas métricas, na execução das operações que compõem o CRUD, de modo avaliar o real impacto sob um mesmo cenário em termos de banco de dados e aplicação utilizada.

Como principal ponto de diferença entre os *drivers* consiste na modelagem dos dados, neste estudo, foram conduzidas análises visando identificar se o tamanho médio de cada registro pode influenciar e qual a proporção desse impacto.

As métricas escolhidas para análises de cada um dos cenários conduzidos foram as seguintes:

- Tempo de Execução Médio: consiste tempo médio total de execução de cada operação específica.
- Tempo de Processamento Médio: consiste no tempo médio de uso do processador durante a execução da operação específica.
- Variação Média de Uso de Memória: consiste na variação média de uso de memória RAM durante a execução da operação específica, sendo expressa em kilobytes (KB).

Por fim, os resultados obtidos em cada um dos cenários analisados são apresentados na Seção 4 e discutidos na Seção 5, com base nas seguintes questões de pesquisa:

Q1 – A escolha do driver impacta no tempo de execução de cada uma das operações de CRUD?

Q2 – A escolha do driver pode influenciar no tempo de processamento (CPU) nas operações de CRUD?

Q3 – A escolha do driver impacta de modo relevante quanto ao uso de memória nas operações de CRUD?

3.1. Características dos Conjuntos de Dados

Para possibilitar a execução do experimento foi utilizado um conjunto de dados com cerca 18 mil instâncias, proveniente do seguinte dataset ⁶. Originalmente, todos os registros presentes são compostos por 89 atributos, predominantemente textuais, obtendo um tamanho médio de 1,37KB. A partir do conjunto original, foi construído um conjunto reduzido em número de atributos (6 atributos), com o mesmo total de instância, porém com tamanho médio de 0,13KB.

Deste modo, dois conjuntos de dados foram utilizados para a experimentação, observe a Tabela 1. O principal motivo para a criação do segundo conjunto de dados, deve-se ao objetivo de confrontar o comportamento dos dois *Drivers* do MongoDB ao manipularem dados que possuem uma grande quantidade de atributos ou uma pequena quantidade de atributos.

Tabela 1. Características dos Conjuntos de Dados utilizado no Experimento

Conjunto de Dados 1	Conjunto de Dados 2
18 mil instâncias	18 mil instâncias
89 atributos	6 atributos
1,37 KB	0,13 KB

3.2. Ambiente de Execução

O ambiente de execução para os testes de desempenho consistiu em um computador com Sistema Operacional Ubuntu 18.04.2, processador Intel i3 3217U e memória RAM de 4GB DDR3. Durante a execução dos testes, o ambiente de execução da aplicação Node.js foi definido o uso do *heap* de memória com limite máximo de 3GB, o que restringiu no limite superior do número de operações executadas em cada cenário de teste.

Em cada cenário de execução, foram extraídos dados relativos ao tempo de execução, tempo de uso de CPU e uso de memória RAM. Foram analisados cenários com diferentes quantidades de operações CRUD realizadas, as quais variaram de 1000, 10000, 100000 e 200000; cada qual foi repetido 10 vezes e armazenado a média aritmética de cada execução. A obtenção das métricas de desempenho para tempo de execução, tempo de uso de CPU e memória RAM, foi realizada através da biblioteca JSMeter ⁷.

4. Resultados do Experimento

Todos os resultados a seguir são apresentados sob a perspectiva que cada uma das operações (*create-read-update-delete*) respectivamente, em que 100% dos registros são

⁶<https://www.kaggle.com/karangadiya/fifa19>

⁷<https://github.com/wahengchang/js-meter>

atingidos em cada operação. Cada resultado apresenta o tempo de execução, em uma operação específica, da combinação de um *Driver* com o conjunto de dados grande (original, com todos atributos) ou conjunto de dados pequeno (com número reduzido de atributos).

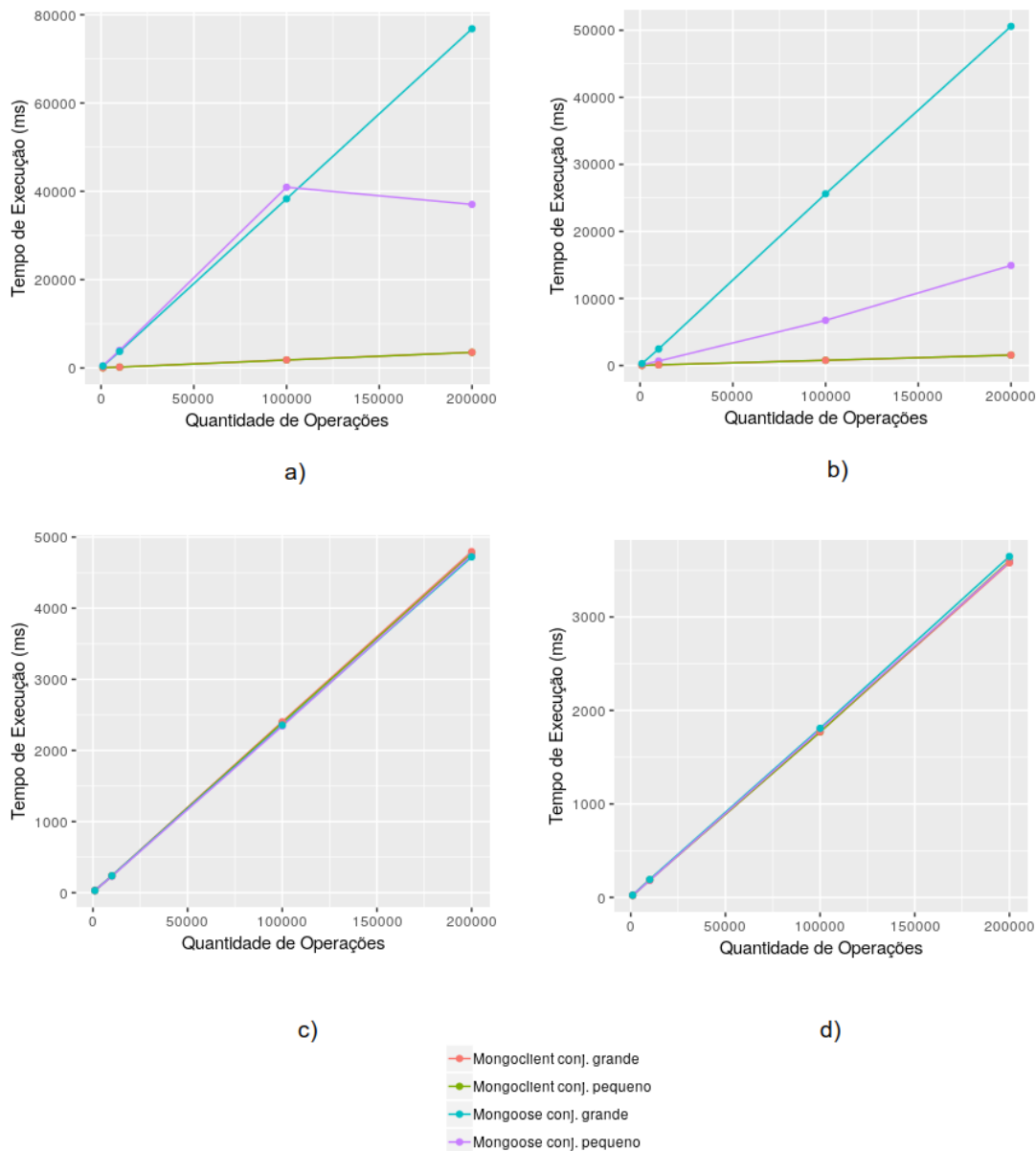


Figura 3. Comparativo de operações em relação ao tempo de execução.

A primeira análise, na Figura 3, apresenta os resultados relativos ao tempo de execução. A Figura 3a apresenta, para operação de inserção, o tempo de execução do *driver* Mongoose maior, para os dois conjuntos, de modo relevante, enquanto para o MongoClient, aparenta não haver diferenças significativas entre os conjuntos. Como exceção, pode-se observar a execução do conjunto pequeno para o Mongoose, em que o tempo para 200 000 operações não mantém a proporcionalidade das demais execuções. Um possível fator que justifique esse comportamento consiste na ocorrência de divisão de conjuntos na

operação de inserção quando a quantidade excede 100 000 itens, contudo, isso não ocorre para o conjunto de registros grandes.

A Figura 3b também apresenta, para operação de busca, o tempo de execução inferior para ambos os conjuntos do MongoClient, em que ambos atuam de modo estritamente similar. Quanto ao Mongoose, as execuções para conjunto reduzido e grande apresentam comportamento crescente e proporcional em detrimento a diferença de tamanho dos registros.

As Figuras 3c e 3d, para as operações de atualização e deleção de registros, ambos os *drivers* para os conjuntos obtiveram tempos de execução estritamente similares e proporcionando a quantidade de operações, não apresentando diferenças significativas, nem mesmo quanto ao tamanho médio dos registros.

A próxima análise, na Figura 4, é referente aos resultados relacionados ao tempo de uso de processamento em cada operação. As Figuras 4a e 4b, assim como análise anterior, para as operações de inserção e busca, apresenta o MongoClient com tempo de execução médio significativamente inferior, em ambos os conjuntos, em detrimento do alto tempo apresentado pelo Mongoose. Para a inserção, também apresenta o caso de exceção para 200 000 operações, cuja possível justificativa é semelhante a análise anterior.

A Figura 4c, representa a operação de atualização, em que se pode-se observar somente o *driver* Mongoose com conjunto de registros maiores apresentou maior tempo de uso de processamento, em detrimento dos demais, que foram semelhantes e com tempo menor relevantemente, mesmo com oscilações. É importante salientar que o tempo de processamento de todas as execuções foram inferiores a 250 ms. A Figura 4d, representa a operação de deleção, cada execução apresentou comportamento relativamente instável, tendo as execuções com *driver* com tempo um pouco maior, contudo não há diferença significativa, porque todas as execuções obtiveram tempo de processamento inferior a 10 ms.

A última análise, apresentada na Figura 5, refere-se a variação média do uso de memória RAM em cada operação. Essa análise apresentou os resultados mais variantes, quanto ao cenários aplicados e quantidade de operações executadas.

As Figuras 5a e 5b apresentam o uso de memória para as operações de inserção e busca, no qual não pode se identificar um padrão de uso de memória, contudo pode-se identificar que predominantemente o *driver* Mongoose tem um consumo maior de memória na operação realizada, em ambos os conjuntos, em detrimento ao casos do *driver* MongoClient. Para ambas operações, há um uso adicional de memória em torno de 30 a 40MB.

Por fim, as Figuras 5c e 5d, respectivamente operações de atualização e deleção, também não apresentam padronização no consumo de memória para ambos os *drivers* e conjuntos. Apesar de apresentar alguns pontos de instabilidade, é possível notar que tais operações consomem pouca memória adicional, aproximadamente 1MB ou menos, mesmo considerando os picos de oscilação.

5. Discussão

Nessa seção são apresentadas discussões a respeito dos resultados obtidos, os quais são analisados sob as perspectivas das questões de pesquisa nas Seções 5.1, 5.2 e 5.3. Adici-

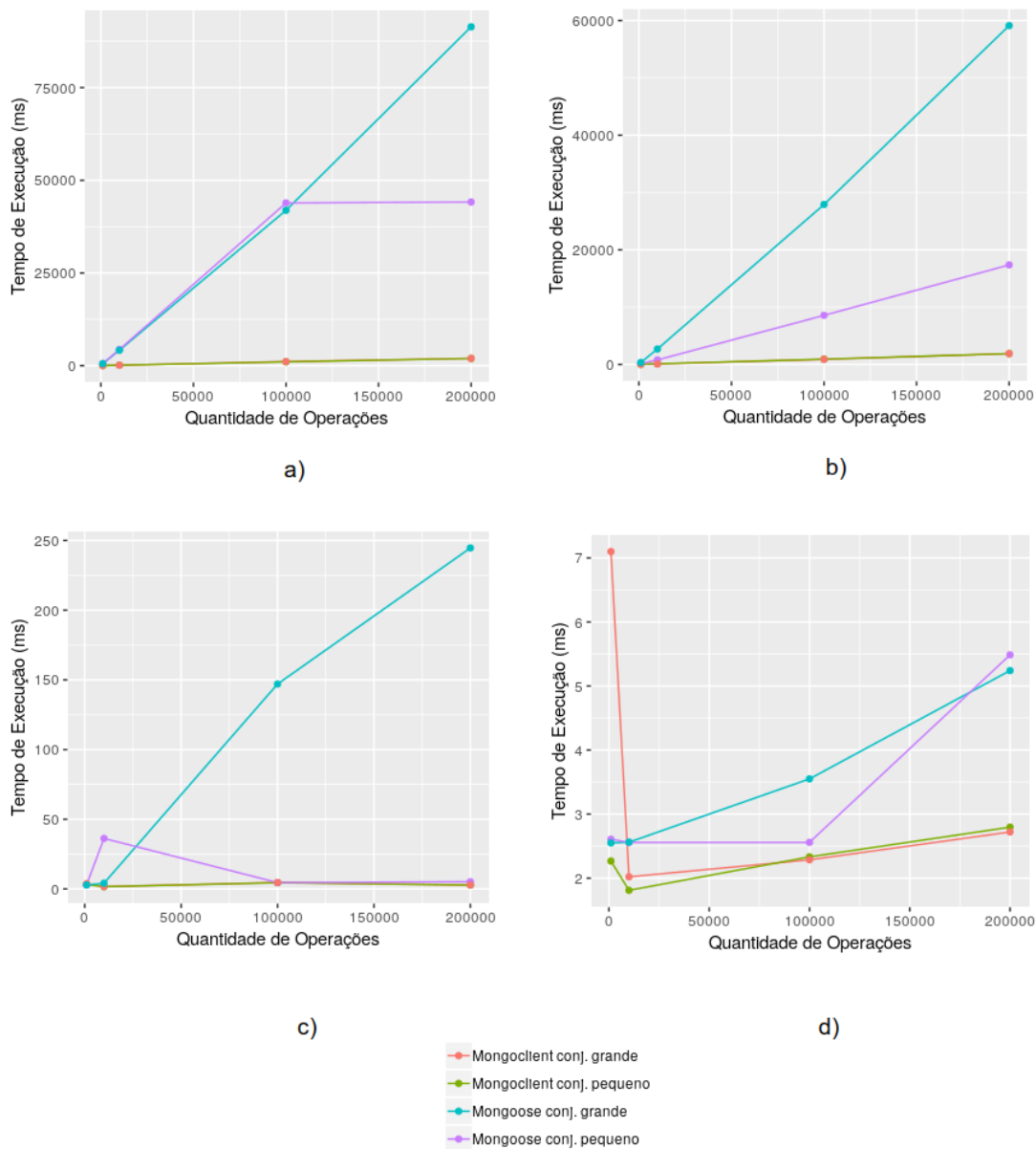


Figura 4. Comparativo de operações em relação ao tempo de uso do processador.

onalmente, na Seção 5.4 é apresentada uma discussão geral do estudo, de modo a cobrir as perspectivas não analisadas pelas questões de pesquisa.

5.1. Discussão da Q1

Para a primeira questão de pesquisa, em linhas gerais, os conjuntos executados utilizando o *driver* Mongoose apresentaram tempo superior ao MongoClient em duas operações, enquanto nas demais operações obteve resultado semelhante, sem diferenças significativas.

Desse modo, sob a perspectiva de tempo médio de execução de cada operação, temos que, a escolha do *driver* pode impactar no desempenho, tendo o MongoClient como melhor opção sob a perspectiva analisada.

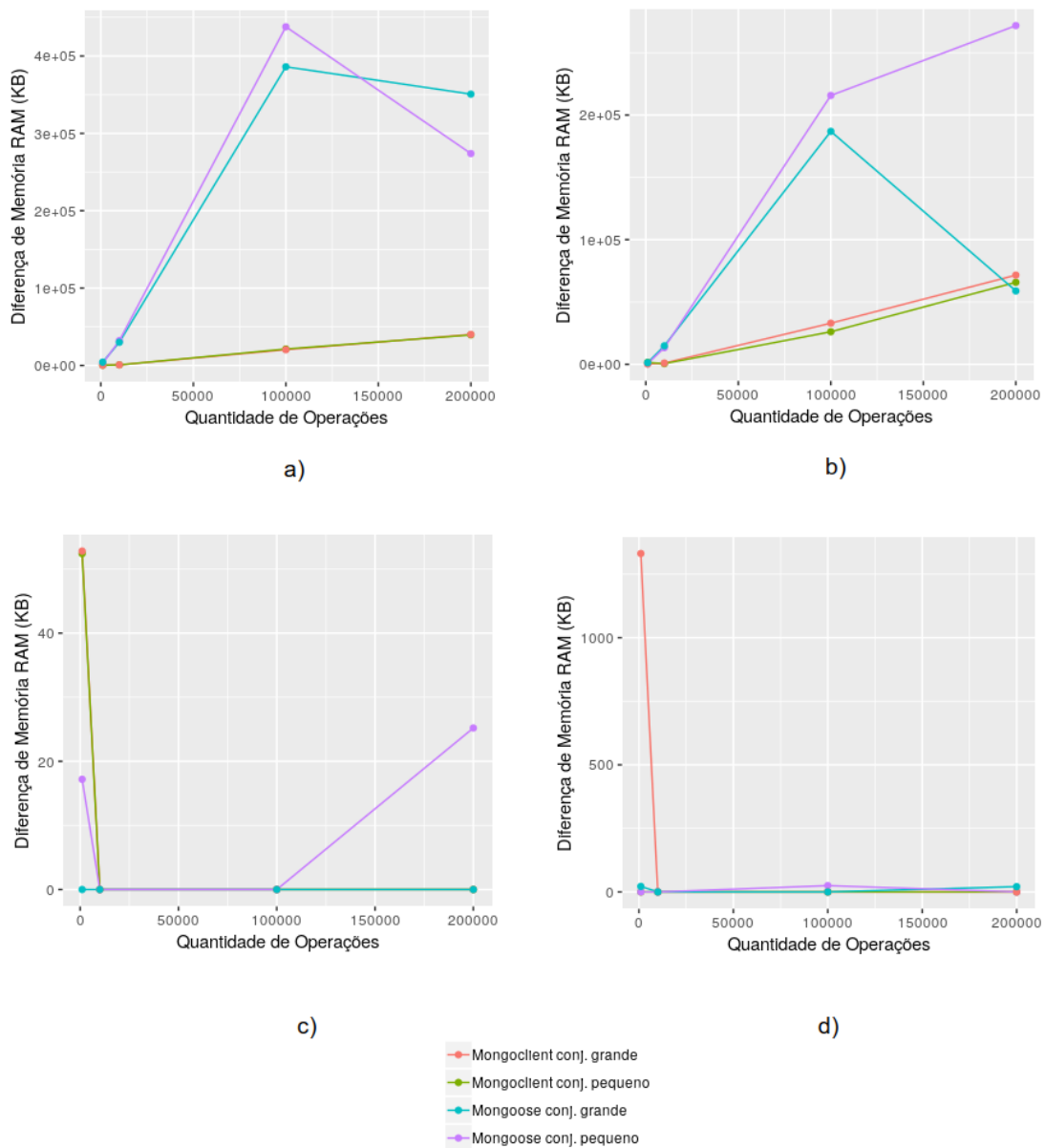


Figura 5. Comparativo de operações em relação ao uso de memória.

5.2. Discussão da Q2

De modo predominante, assim como na questão anterior, os conjuntos executados utilizando o MongoClient obtiveram melhor tempo de processamento em relação ao Mongoose, em ambos os conjuntos, principalmente para as operações de inserção e busca, enquanto nas demais operações (atualização e deleção), também foi registrado melhor desempenho, contudo em proporção menor.

Em suma, em termos de tempo de processamento, temos que, a escolha do *driver* pode influenciar no desempenho, também apresentando MongoClient como melhor opção.

5.3. Discussão da Q3

A respeito a última questão de pesquisa, tem-se que para as operações de inserção e busca, para a maioria dos casos de execução, o *driver* Mongoose apresenta maior consumo de memória, enquanto para as operações de atualização e deleção não há diferenças significantes. Contudo cabe ressaltar que em nenhuma das comparações houve comportamento estável e proporcional de memória.

Desse modo, em termos de consumo de memória, temos que, a escolha do *driver* não impacta de modo relevante para todas as operações, apesar do consumo inferior por parte do *driver* MongoClient.

5.4. Discussão Geral

De modo geral, os dados obtidos indicam que as operações de inserção e busca são as mais custosas quanto ao tempo de execução, para os piores casos, aproximadamente 70 000 a 80 000 ms, enquanto as demais são inferiores a 5 000 ms. Sob a perspectiva de tempo de uso de CPU também se vale a mesma análise, inclusive as operações indicam aproximada proporcionalidade em comparação ao tempo de execução total. Por fim, quanto ao uso de memória, ambas operações também apresentam maior custo, mesmo que não-linear, em níveis próximos de 30 a 40MB, em detrimento das demais operações com uso próximo ou inferior a 1MB.

Quanto a diferença média de tamanho dos registros do conjunto de dados adotado, o *driver* MongoClient apresentou-se de modo indiferente, não apresentando oscilações significativas de desempenho, enquanto o Mongoose apresenta o desempenho diretamente proporcional ao tamanho do registro.

Em termos de comparação entre os *drivers*, o MongoClient apresentou desempenho mais estável e de menor custo sob a perspectiva de todas as métricas adotadas, em detrimento ao Mongoose.

Portanto, conclui-se que, sob o critério exclusivo de desempenho, o *MongoClient* apresenta melhor desempenho em relação ao concorrente. Cabe ressaltar que se quaisquer recursos adicionais providos por uma das opções, como verificação de dados ou facilidade de implementação, consistirem em fatores relevantes, deve-se reavaliar a escolha, contudo, esse estudo tem caráter quantitativo e não visa mensurar a utilização de recursos adicionais que podem variar de contexto e aplicação utilizados.

6. Ameaças à Validação

Esta seção apresenta as possíveis ameaças que podem comprometer os resultados desse estudo comparativo.

Como ameaça à validação interna desse estudo tem a etapa da obtenção dos dados quantitativos quanto às análises conduzidas, desse modo, durante a execução dos testes, todos esses foram conduzidos repetidamente e de modo subsequente para que nenhuma interferência pudesse ser adicionada.

Como ameaça à validação externa, tem-se o conjunto de dados selecionado, o qual está diretamente atrelado aos resultados numéricos obtidos, porém análise é baseado na comparação relativa dos resultados, além a aplicação de teste foi projetada de modo aceitar um conjunto de dados genérico.

7. Trabalhos Relacionados

[Kanade et al. 2014] did a study for NoSQL databases with both normalized and denormalized forms using a similar dataset, and have found that the embedded MongoDB data model provides a much better efficiency as compared to a normalized model.

8. Considerações Finais

Este artigo apresenta um estudo comparativo entre *drivers* para banco de dados MongoDB. Na avaliação foram conduzidas análises quantitativas quanto a tempo de execução, tempo de processamento e uso de memória para as operações de CRUD; além de comparar o impacto da variação do tamanho médio dos registros.

De modo geral, através dos resultados quantitativos, foi obtido que o *driver* MongoClient obtém melhor desempenho médio quanto a todas as operações, principalmente sob as métricas de tempo de execução e processamento, e, de modo menos significativo, quanto ao consumo de memória.

Adicionalmente, como contribuição adicional deste trabalho, tem-se a implementação da ferramenta de testes ⁸, de modo a viabilizar a execução de futuras análises em ambientes Node.js.

Referências

BSONSpecification (2019). Binary json.

Chaniotis, I. K., Kyriakou, K.-I. D., and Tselikas, N. D. (2015). Is node.js a viable option for building modern web applications? a performance evaluation study. *Computing*, 97(10):1023–1044.

Cooper, B. F., Silberstein, A., Tam, E., Ramakrishnan, R., and Sears, R. (2010). Benchmarking cloud serving systems with ycsb. In *Proceedings of the 1st ACM symposium on Cloud computing*, pages 143–154. ACM.

Edlich, S. (2015). Nosql-your ultimate guideto the non-relational universe! nosql.

Gagliardi, V. (2018). How to inspect the memory usage of a process in node.js.

González-Aparicio, M. T., Younas, M., Tuya, J., and Casado, R. (2016). A new model for testing crud operations in a nosql database. In *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)*, pages 79–86.

Han, J., Haihong, E., Le, G., and Du, J. (2011). Survey on nosql database. In *2011 6th international conference on pervasive computing and applications*, pages 363–366. IEEE.

Kanade, A., Gopal, A., and Kanade, S. (2014). A study of normalization and embedding in mongodb. In *2014 IEEE International Advance Computing Conference (IACC)*, pages 416–421. IEEE.

Lutu, P. (2015). Big data and nosql databases: new opportunities for database systems curricula. *Proceedings of the 44th Annual Southern African Computer Lecturers' Association, SACLA*, pages 204–209.

⁸<https://github.com/leandroungari/database-driver>

- Mardan, A. (2014). Boosting your node.js data with the mongoose orm library. In *Practical Node.js*, pages 149–172. Springer.
- Membrey, P., Plugge, E., and Hawkins, D. (2011). *The definitive guide to MongoDB: the noSQL database for cloud and desktop computing*. Apress.
- Mohamed, M., G. Altrafi, O., and O. Ismail, M. (2014). Relational vs. nosql databases: A survey. *International Journal of Computer and Information Technology (IJCIT)*, 03:598.
- NodeJS (2019). Node.js.
- Patil, M. M., Hanni, A., Tejeshwar, C. H., and Patil, P. (2017). A qualitative analysis of the performance of mongodb vs mysql database based on insertion and retrieval operations using a web/android application to explore load balancing — sharding in mongodb and its advantages. In *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pages 325–330.
- Rafique, A., Van Landuyt, D., Lagaisse, B., and Joosen, W. (2018). On the performance impact of data access middleware for nosql data stores a study of the trade-off between performance and migration cost. *IEEE Transactions on Cloud Computing*, 6(3):843–856.
- Ramesh, D., Khosla, E., and Bhukya, S. N. (2016). Inclusion of e-commerce workflow with nosql dbms: Mongodb document store. In *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICIC)*, pages 1–5.
- Ward, J. S. and Barker, A. (2013). Undefined by data: a survey of big data definitions. *arXiv preprint arXiv:1309.5821*.