

IJCIT

(Indonesian Journal on Computer and Information Technology)

Journal Homepage: <http://ejournal.bsi.ac.id/ejurnal/index.php/ijcit>

Perbandingan Performa SQL dan NoSQL Dengan PHP Pada 5 Juta Data

Setiawan Budiman¹, Faisal Fadhila², Vian Ardiyansyah Saputro³
Ema Utami⁴, Khusnawi⁵

^{1,2,3,4,5}Program Studi Magister Teknik Informatika, Universitas Amikom Yogyakarta
Yogyakarta, Indonesia

e-mail: setiawan.1267@students.amikom.ac.id¹, faisal.fadhilla87@gmail.com²,
vian.1258@students.amikom.ac.id³, ema.u@amikom.ac.id⁴, khusnawi@amikom.ac.id⁵

ABSTRAK

Dunia digital sekarang sudah merambah pada seluruh aspek kehidupan. Salah satu yang berperan penting dalam digitalisasi adalah database. Seluruh perusahaan digital pasti akan mempunyai database untuk kepentingan usaha. Semakin lama database yang dikelola akan semakin banyak, oleh karena itu kita membutuhkan aplikasi database yang ideal untuk digunakan. Pada penelitian ini, kami akan melakukan perbandingan kecepatan proses data antara SQL (MySQL) dengan NoSQL (MongoDB) dengan menggunakan VPS webserver Apache dan bahasa program PHP yang umum digunakan oleh banyak perusahaan. Metodologi yang kami gunakan adalah menguji coba kecepatan SELECT, INSERT, DELETE dan UPDATE pada jumlah data mulai 1.000, 10.000, 100.000, 1.000.000, 2.000.000 dan 5.000.000. Khusus untuk SELECT dan INSERT, kami melakukan proses looping data seperti yang umum dilakukan pada program PHP (do while). Hasil dari penelitian ini kami dapatkan bahwa proses ujicoba dengan jutaan data akan terasa lebih cepat saat menggunakan MongoDB dibanding MySQL untuk proses SELECT INSERT dan UPDATE. Namun berbeda ketika menggunakan proses DELETE, MySQL memiliki waktu respon yang lebih baik dibandingkan dengan MongoDB. Hal ini disebabkan karena MongoDB menggunakan program PHP sebagai aplikasi yang menjalankan proses query.

Kata Kunci: *database, nosql, php, sql*

ABSTRACTS

Digital world has been impact almost every aspect of life. Database is the most important thing at digitization. All digital company will have databases. After several period of usage, the database size will improve, therefore we need to use the ideal database system. At this paper, we will compare the speed of processing data between SQL (MySQL) vs NoSQL (MongoDB) which run on VPS Apache webserver using famous PHP script. The methodology that we use for speed measurement will use SELECT, INSERT, DELETE and UPDATE at big data start from 1.000, 10.000, 100.000, 1.000.000, 2.000.000 and 5.000.000. Especially for SELECT and INSERT, we do the looping procedure that usually use at PHP script (do while). The result from this paper declare that processing million of data will be faster on MongoDB if compare with MySQL when we use SELECT, INSERT and UPDATE. However, it is different when using the DELETE process, MySQL has a better response time compared to MongoDB. This is because MongoDB uses a PHP program as an application that runs the query process.

Keywords: *database, nosql, php, sql*



1. PENDAHULUAN

Pertumbuhan pesat di dalam teknologi informasi telah menghasilkan milyaran data setiap harinya, di antara penghasil data tersebut adalah situs jejaring sosial Twitter. Situs jejaring sosial ini memungkinkan pengguna untuk mengirimkan dan membaca pesan singkat sebanyak 280 karakter secara real time, pesan singkat ini biasa kita kenal dengan sebutan kicauan (tweet).

Twitter menyediakan layanan API yang mengijinkan untuk dapat di integrasikan dengan aplikasi lainnya, melalui layanan API yang telah disediakan Twitter yaitu Twitter Streaming API, peneliti dapat melakukan pengambilan data dari Twitter, data yang telah didapatkan kemudian disimpan ke dalam database NoSQL ataupun database SQL. Penelitian sebelumnya yang telah

dilakukan mengenai perbandingan performa antara database NoSQL dan database SQL (Chang & Chua, 2019), Pada paper tersebut menggunakan data Twitter yang didapatkan dengan cara melakukan Twitter Streaming API. Data yang digunakan adalah sebanyak 1.000.000 dan membandingkan langsung dengan bahasa program python. Hasil yang didapatkan pada penelitian tersebut adalah:

- Pada proses penulisan data, disebutkan bahwa MongoDB lebih cepat dibandingkan MySQL.
- Pada proses pengambilan data, MongoDB lebih cepat dibandingkan MySQL.
- Pada proses modifikasi data, MongoDB lebih cepat dibandingkan MySQL.
- Pada proses penghapusan data, MySQL lebih cepat dibandingkan MongoDB.

Tabel 1. Perbandingan Penelitian

| SQL dan NoSQL Database Comparison [1] | | This Paper |
|---------------------------------------|-------------------|-------------------|
| SQL Databasel | MySQL | MySQL |
| NoSQL Database | MongoDB | MongoDB |
| Data Model | Semi Stuctured | Semi Stuctured |
| Data Source | Social Media Data | Social Media Data |
| Total Record | 1.000.000 | 5.000.000 |
| INSERT Operation | Done | Done |
| SELECT Operation | Done | Done |
| UPDATE Operation | Done | Done |
| DELETE Operation | Done | Done |
| Programming Language | Python | PHP |

Selanjutnya pada penelitian evaluasi kinerja antara MySQL dan MongoDB menggunakan aplikasi hypermarket (Damodaran B et al., 2016), pada penelitian tersebut memberikan hasil bahwa MongoDB memerlukan waktu yang lebih sedikit ketika adanya penambahan data, berbeda halnya dengan MySQL yang memerlukan waktu lebih lama dimana jumlah data yang digunakan adalah 25.000.

Kemudian untuk penelitian lainnya yang berkaitan dengan evaluasi kinerja antara MySQL dan MongoDB menggunakan data e-commerce (Aboutorabi et al., 2015), pada penelitian tersebut menggunakan proses INSERT, DELETE dan UPDATE dengan jumlah data sebanyak 1.000.000, dimana hasil dari penelitian tersebut menunjukkan bahwa MongoDB memberikan hasil kinerja lebih baik dengan data yang besar dan kompleks.

Berdasarkan hasil penelitian sebelumnya, maka penulis mencoba membuat sebuah

perbandingan antara MySQL dan MongoDB untuk empat proses yaitu *select*, *insert*, *delete* dan *update* serta menggunakan jumlah data yang lebih banyak yaitu 5.000.000 data. Di dalam penelitian ini menggunakan bahasa program yang berbeda, yaitu dengan Apache webserver dan PHP script.

2. METODE PENELITIAN

Untuk mendukung penelitian ini, kami akan menggunakan data sebanyak 5.000.000 data Twitter. Dalam hal ini kami mendapatkan data awal dari data *Sentiment Analysis With Tweets* milik KazAnova yang terdapat website di Kaggle.com. Data ini terdiri dari 140 karakter *datasheet* yang totalnya adalah 1.600.000 juta data yang di dapatkan dari Twitter API ini memiliki 6 field yaitu: target, ids, date, flag, user, text.

Awalnya data tersimpan pada format txt yang dipisahkan dengan tanda koma (,). Lalu

dengan program MySQL broser yaitu HeidiSQL, kami berhasil mengimport sebanyak 1.600.000 data kedalam server MySQL. Untuk mendukung kebutuhan data makalah ini, kami melakukan duplikasi dari data awal 1.600.000 sampai dengan mencapai total 5.000.000, tetapi agar data tidak terduplikasi sepenuhnya, kami memberikan tambahan text unik di field text data Twitter tersebut, sehingga kami akhirnya memiliki total 5 juta data yang tidak sama untuk bagian field text nya.

Penelitian ini, memberikan perhatian pada cara melakukan proses looping data secara satuan lewat script PHP. Hal ini berbeda dengan research yang dilakukan sebelumnya, karena beberapa literatur review yang kami baca tidak melakukan proses looping dengan script PHP. Mengapa kami memilih menggunakan script PHP? Karena program PHP adalah salah satu bahasa program yang paling banyak digunakan untuk membuat program web khususnya untuk perusahaan UMKM dan beberapa usaha lainnya. Proses yang kami analisa adalah proses SELECT, INSERT, DELETE, UPDATE. Masing-masing proses dibuatkan program php yang berbeda dengan parameter data yang berbeda juga untuk melakukan proses ujicobanya, yaitu menggunakan data sebanyak 1.000, 10.000, 100.000, 1.000.000, 2.000.000 dan 5.000.000 untuk membandingkan masing-masing kinerja kecepatan hasil proses ujicoba masing-masing perintah diatas.

Untuk proses perhitungan waktu, kami menggunakan fungsi microtime() dari PHP yang dihitung sebelum proses berjalan sampai dengan akhir proses selesai. Jadi yang kami tulis diawal program adalah \$waktumulai = microtime(true) dan diakhir program adalah \$waktuakhir = microtime(true). Lalu kami hitung dengan rumus \$waktutotal = \$waktuawal - \$waktuakhir.

a. Proses SELECT

Kami membuat koneksi PDO MySQL untuk melakukan koneksi dan query proses SELECT untuk masing-masing data, mulai dari 1.000, 10.000, 100.000, 1.000.000, 2.000.000 dan 5.000.000 dengan memberikan batasan WHERE text LIKE "%a%". Sehingga bisa

dikatakan kami menguji coba menghitung kecepatan.

b. Proses INSERT

Ini adalah metode pengujian yang kami pilih untuk melakukan perhitungan dan membandingkan antara Mysql dan MongoDB. Untuk MySQL kami menggunakan PDO MySQL untuk looping data dan untuk MongoDB kami menggunakan \$bulkWrite->insert() dengan data looping mulai dari 1.000, 10.000, 100.000, 1.000.000, 2.000.000 dan 5.000.000.

c. Proses DELETE

Pada proses DELETE, kami menggunakan MySQL PDO dengan perintah hapus semua. Pada MongoDB kami menggunakan perintah \$bulk->delete() untuk melakukan penghapusan semua data yang masing-masing dianalisa pada MySQL table dan MongoDB collection mulai dari 1.000, 10.000, 100.000, 1.000.000, 2.000.000 dan 5.000.000.

d. Proses UPDATE

Pada proses UPDATE, kami menggunakan MySQL PDO dengan perintah UPDATE semua data yang mengandung kata kunci NO_QUERY di field flag dan yang diupdate adalah field user menjadi BARU UPDATE. Sehingga semua data terupdate. Pada MongoDB kami melakukan metode update dengan parameter yang sama dengan perintah \$bulk->update() yang dilakukan dan dianalisa pada MySQL table dan MongoDB collection mulai dari 1.000, 10.000, 100.000, 1.000.000, 2.000.000 dan 5.000.000.

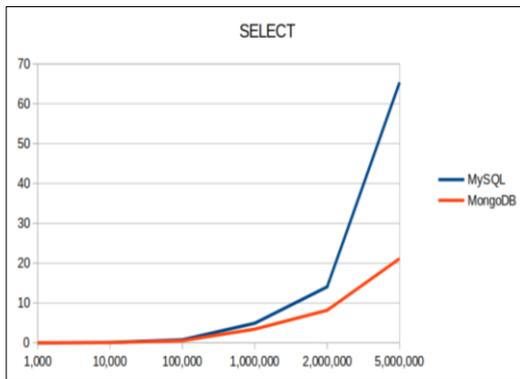
3. HASIL DAN PEMBAHASAN

Dari hasil analisa yang telah kami lakukan menggunakan perintah SELECT, INSERT, DELETE, dan UPDATE yang dijalankan pada server menggunakan 2 CPU dan 2 GB RAM dengan script PHP 7, maka dapat dilihat bahwa pada NoSQL ada penambahan kecepatan saat data melakukan SELECT dan INSERT, sedangkan pada DELETE dan UPDATE tidak terjadi perbedaan yang signifikan antara MySQL dibanding MongoDB.

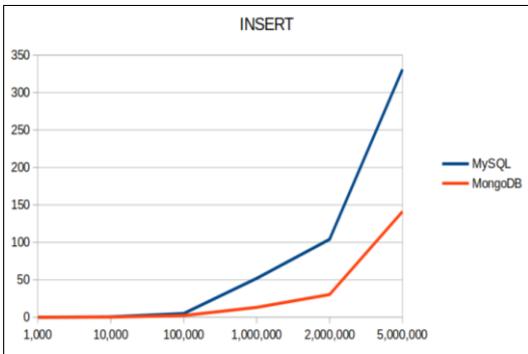
Tabel 2. Hasil Penelitian

| Jumlah Data | Select | | Insert | | Delete | | Update | |
|-------------|--------|---------|--------|---------|--------|---------|--------|---------|
| | MySQL | MongoDB | MySQL | MongoDB | MySQL | MongoDB | MySQL | MongoDB |
| 1.000 | 0.011 | 0.014 | 0.06 | 0.03 | 0.011 | 0.015 | 0.102 | 0.031 |
| 10.000 | 0.096 | 0.068 | 0.54 | 0.21 | 0.03 | 0.12 | 0.25 | 0.3 |
| 100.000 | 0.78 | 0.56 | 5.05 | 2.35 | 0.35 | 1.17 | 2 | 2 |
| 1.000.000 | 4.94 | 3.45 | 51.91 | 13.17 | 3.71 | 12.57 | 19.25 | 12.53 |
| 2.000.000 | 14.07 | 8.19 | 104.09 | 30.35 | 8.33 | 25.02 | 33.72 | 27.52 |
| 5.000.000 | 65.37 | 21.14 | 330.9 | 141.13 | 21.33 | 63.47 | 75.45 | 70.85 |

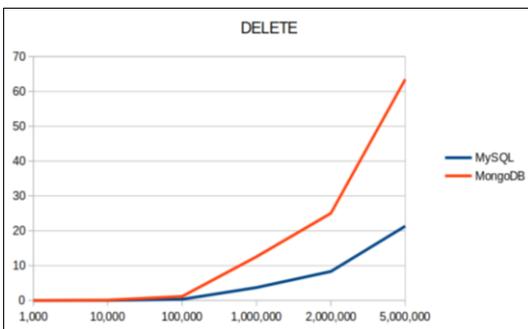
dari hasil pengujian yang tampak pada tabel 2, dapat dilihat pada grafik 1 sampai grafik 4.



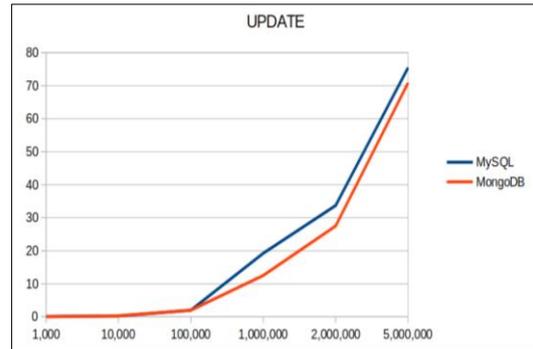
Grafik 1. Perbandingan MySQL dan MongoDB pada proses SELECT



Grafik 2. Perbandingan MySQL dan MongoDB pada proses INSERT



Grafik 3. Perbandingan MySQL dan MongoDB pada proses DELETE



Grafik 4. Perbandingan MySQL dan MongoDB pada proses UPDATE

Penelitian ini menggunakan server cloud VPS dengan CPU 2 CORE dan 2 GB RAM. Kami memilih server ini sebagai spesifikasi yang umum digunakan pada sebuah website dengan kunjungan mencapai 1 juta view per bulan. Kami tidak melakukan pengujian pada server yang memiliki spesifikasi lebih rendah atau lebih tinggi karena keterbatasan waktu untuk proses pembuatan tulisan ini.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa database MySQL dan MongoDB memiliki perbedaan kecepatan ketika melakukan pemrosesan data. Hal ini dapat terlihat ketika menggunakan proses SELECT, INSERT dan UPDATE, MongoDB memiliki waktu respon yang lebih cepat, namun berbeda ketika menggunakan proses DELETE, MySQL memiliki waktu respon yang lebih baik dibandingkan dengan MongoDB. Hal ini disebabkan karena MongoDB menggunakan program PHP sebagai aplikasi yang menjalankan proses query.

Untuk penelitian selanjutnya mengenai perbandingan performa SQL dan NoSQL ini dapat dilakukan dengan penggunaan spesifikasi server

yang lebih baik, metode yang digunakan dan penggunaan sistem operasi yang berbeda.

5. REFERENSI

- Aboutorabi, S. H., Rezapour, M., Moradi, M., & Ghadiri, N. (2015). Performance evaluation of SQL and MongoDB databases for big e-commerce data. *CSSE 2015 - 20th International Symposium on Computer Science and Software Engineering*.
- Chang, M. L. E., & Chua, H. N. (2019). SQL and NoSQL database comparison: from performance perspective in supporting semi-structured data. *Advances in Intelligent Systems and Computing*.
- Damodaran B, D., Salim, S., & Vargese, S. M. (2016). Performance Evaluation of MySQL and MongoDB Databases. *International Journal on Cybernetics & Informatics*.
- Gupta, S., Narsimha. (2015). Performance evaluation of NoSQL – cassandra over relational data store – MYSQL for bigdata. *Int. J. Technol.*
- Parker, Zachary Poe, Scott Vrbsky, Susan V. (2013). Comparing NoSQL MongoDB to an SQL DB. *Proceedings of the Annual Southeast Conference*.
- Abramova, Veronika Bernardino, Jorge Furtado, Pedro. SQL or NoSQL? Performance and scalability evaluation. (2015). *International Journal of Business Process Integration and Management*