

ANALISIS KINERJA SISTEM MONITORING ENERGI LISTRIK DENGAN PENYIMPANAN CLUSTER SQL BERBASIS KUBERNETES MENGGUNAKAN REGRESI POLINOMIAL



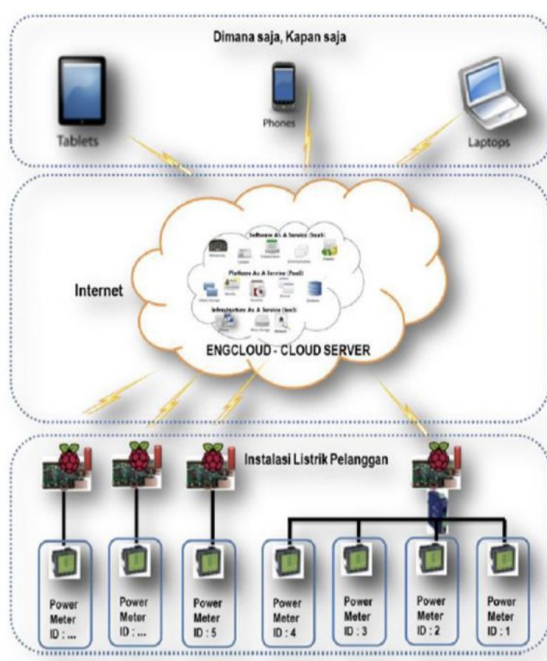
Bagus Dewangga*, Edi Leksono¹, Nugraha¹

¹Magister Teknik Fisika, Fakultas Teknologi Industri, Institut Teknologi Bandung, Bandung, Indonesia.

*Penulis Koresponden

PENDAHULUAN

Penggunaan internet dan IoT terus mengalami peningkatan yang pesat. Hal ini juga terjadi di Laboratorium Manajemen Energi Teknik Fisika ITB, Sistem Monitoring Energi Listrik (SiELis) yang dikembangkan sudah diimplementasikan di banyak gedung dan akan terus mengalami peningkatan. Meskipun begitu, SiELis saat ini memiliki beberapa keterbatasan. Pertama, ketika terjadi kegagalan di database maka seluruh sensor tidak dapat mengirim data ke server. Kedua, ketika terjadi kenaikan laju transaksi data oleh penggunaan user dapat mengakibatkan kegagalan database. Oleh karena itu, reliabilitas dari SiELis perlu ditingkatkan dengan meningkatkan reliabilitas dari database. Clustering database adalah cara yang dapat digunakan agar database SiELis yang berbasis SQL dapat memiliki reliabilitas tinggi. Sementara itu, untuk meningkatkan kemampuan adaptasi terhadap perubahan laju transaksi data adalah dengan konsep containerisasi menggunakan docker dan kubernetes. Metode clustering database ada dua yaitu replikasi dan sharding, metode replikasi memiliki keunggulan reliabilitas, kemudahan implementasi dan dapat digunakan pada berbagai tipe dan struktur data dibandingkan dengan metode sharding. Namun, untuk setiap transaksi input data cluster SQL dengan metode replikasi harus menyimpan data ke setiap server di dalam cluster yang berpotensi meningkatkan latensi sehingga mengurangi kinerja SiELis. Oleh karena itu, diperlukanlah model untuk analisis dan prediksi kinerja SiELis berbasis latensi ketika prediksi latensi database pada penelitian - penelitian sebelumnya adalah model dengan pendekatan regresi polinomial.



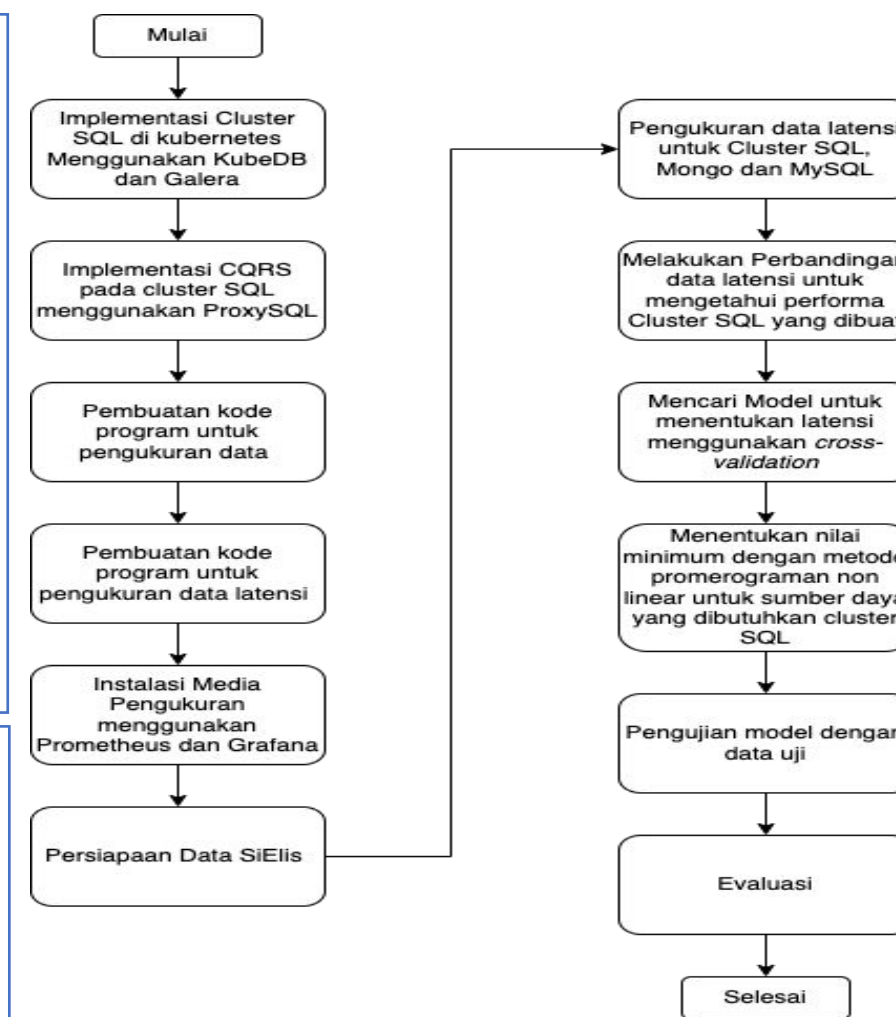
SiELis memiliki beberapa komponen yang pertama adalah komponen pengambilan data menggunakan power meter. Komponen transmisi data berbasis serial. Komponen pengolahan data lokal menggunakan raspberry pi. Komponen database menggunakan MySQL yang diinstall pada sistem awan. Tujuan utama dari sistem ini adalah mengumpulkan informasi penggunaan listrik dan dapat diakses melalui internet.

TUJUAN PENELITIAN

- Merancang dan membangun cluster SQL berbasis kubernetes untuk meningkatkan reliabilitas SiELis.
- Memvalidasi perancangan dan implementasi SiELis Cluster SQL menggunakan parameter latensi dengan model regresi polinomial.

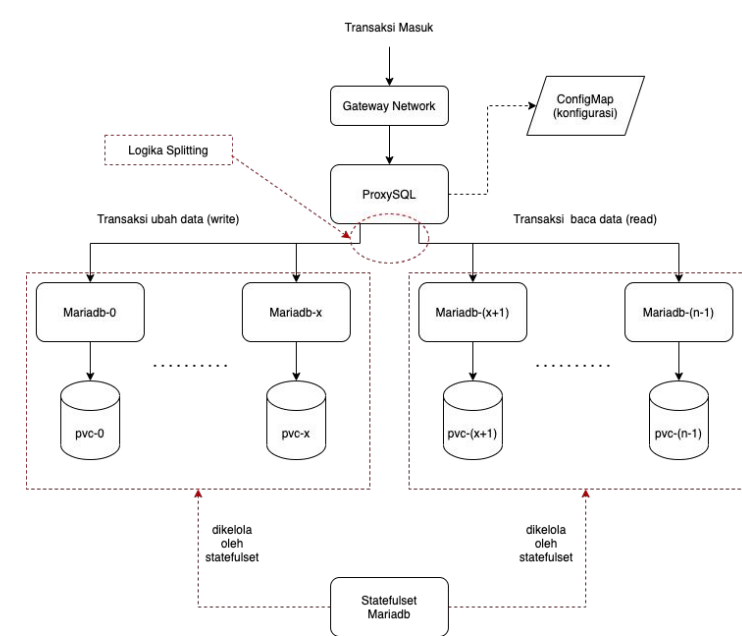
METODOLOGI PENELITIAN

Pelaksanaan penelitian ini dilakukan dalam beberapa tahap. Tahap pertama ada perancangan arsitektur cluster SQL menggunakan kubernetes dan implementasi arsitektur tersebut di Google Clouds. Tahap kedua adalah mempersiapkan data pengujian menggunakan data pengukuran penggunaan energi listrik oleh laboratorium Manajemen Energi. Tahap ketiga adalah melakukan pengujian kinerja dengan parameter latensi dan membandingkannya dengan MySQL tunggal dan MongoDB. Tahap keempat adalah pembuatan model prediksi latensi cluster SQL menggunakan model regresi polinomial. Tahap kelima adalah validasi model, kemudian melihat apakah perlu ada perbaikan jika model yang dihasilkan tidak memberikan pendekatan yang akurat. Tahap akhir adalah penulisan laporan hasil penelitian dan menjelaskan kontribusi keilmuan yang diberikan oleh penelitian ini.



Perancangan Cluster SQL MariaDB menggunakan Kubernetes:

Pada penelitian ini kubernetes yang digunakan adalah kubernetes yang disediakan oleh Google Cloud yaitu GKE (Google kubernetes engine). Setiap entitas MariaDB akan dipasang pada sebuah pod. Siklus hidup pod MariaDB dikelola menggunakan StatefulSets. Ketika satu pod MariaDB berhenti maka media penyimpanannya tidak akan digunakan oleh pod MariaDB yang lain tapi tetap digunakan oleh pod MariaDB tersebut. Transaksi yang masuk ke cluster akan melalui gerbang jaringan (Gateway Network). Setelah melalui gerbang jaringan permintaan transaksi data diteruskan ke ProxySQL. ProxySQL akan mencoba mengidentifikasi tipe transaksi data yang masuk. Ketika transaksi yang masuk adalah transaksi input data maka transaksi data akan diarahkan ke salah satu pod dalam kelompok MariaDB yaitu dari MariaDB-0 sampai MariaDB-x. Kemudian jika tipe transaksi yang masuk adalah pembacaan data yaitu dari MariaDB-(x+1) sampai MariaDB-(n-1).



Perancangan Pengujian Cluster SQL MariaDB:

Pengujian dilakukan dengan membuat sistem pengujian untuk melakukan permintaan transaksi data ke cluster SQL dengan laju transaksi data dalam kelipatan 100 transaksi per detik. Sistem pengujian dibuat dengan menggunakan bahasa pemrograman golang dan dijalankan menggunakan kubernetes Job. Kubernetes Job akan mengatur jumlah pod yang dibutuhkan untuk mencapai jumlah transaksi per detik yang diharapkan. Misalnya jumlah transaksi data yang ingin dicoba adalah 500 transaksi per detik, tiap pod mengelola 100 transaksi per detik maka jumlah pod pengujian yang dibutuhkan adalah 5. Selain itu kubernetes job akan mengatur siklus hidup pod pengujian, ketika pod pengujian berhenti karena terjadi kesalahan maka kubernetes job akan menggantinya dengan pod pengujian baru.

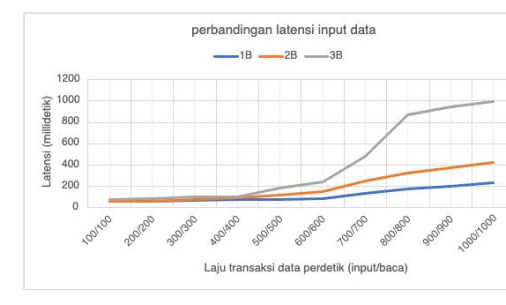
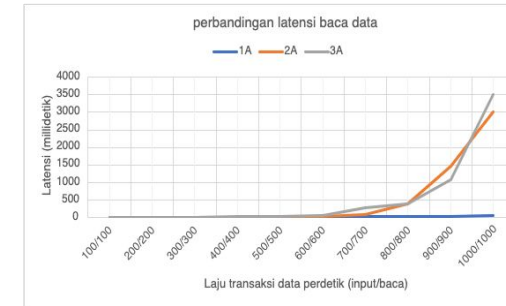
Pemodelan Sistem Menggunakan Regresi Polinomial:

Pemodelan menggunakan bahasa pemrograman python dengan menggunakan modul scikit-learn untuk menentukan hubungan antara fitur dengan latensi yang dihasilkan. Pada penelitian ini digunakan pendekatan polinomial dari order 1 sampai dengan order 5 untuk menentukan model. Selanjutnya dilakukan iterasi untuk setiap order polinomial untuk menentukan hyper-parameter yang paling sesuai dengan model yang cari dengan menggunakan 5-fold cross-validation.

Hasil Pengolahan Data Cluster SQL SiELis:

Untuk perbandingan latensi input data untuk variasi jumlah pod input data lebih banyak yaitu 1B, 2B dan 3B dan untuk perbandingan latensi baca data untuk variasi jumlah pod baca data lebih banyak yaitu 1A, 2A dan 3A terlihat bahwa dengan menurunkan nilai memori dan CPU kemudian menaikkan jumlah pod tidak menaikkan kinerja transaksi data SiELis

Variasi	Jumlah pod MariaDB untuk transaksi penambahan data	Jumlah pod MariaDB untuk transaksi pembacaan data	Memori pod MariaDB (milli Core)	CPU pod MariaDB (MiB)
Variasi 1A	1	2	1234	834
Variasi 1B	2	1	1234	834
Variasi 2A	2	3	740	500
Variasi 2B	3	2	740	500
Variasi 3A	3	4	529	358
Variasi 3B	4	3	529	358



Perbandingan dengan database Mongo MySQL:

SiELis dengan cluster SQL tidak lebih baik dibandingkan dengan MongoDB tetapi lebih baik dari MySQL ketika transaksi data lebih besar dari 900 transaksi per detik untuk input data dan 900 transaksi data untuk baca data.

Pemodelan:

Regresi polinomial latensi input data menunjukkan nilai rata-rata R2 terbaik adalah 0.9202 dengan orde polinomial 3, tetapi dengan data pengukuran menunjukkan nilai kesalahan 19.32%. Sementara itu, untuk regresi polinomial latensi baca data menunjukkan nilai rata-rata R2 terbaik adalah 0.7641 dengan polinomial orde 5, tetapi pengujian data pengukuran menunjukkan nilai kesalahan 66.37%. Nilai kesalahan yang jauh berbeda dengan nilai R2 dari kedua model menunjukkan bahwa terjadi overfitting untuk kedua model polinomial tersebut. Dari overfitting tersebut dapat disimpulkan bahwa pendekatan polinomial untuk keseluruhan data tidak dapat diimplementasikan untuk mengestimasi nilai latensi SiELis dengan cluster SQL, namun diperlukan tahap lebih lanjut agar kesalahan yang didapatkan semakin kecil. pemisahan titik ukur menjadi dua rentang dilakukan untuk mengurangi dampak kesalahan estimasi latensi input data dan baca data. Sehingga kesalahan estimasi latensi input data dapat dikurangi menjadi 9.11% dengan pendekatan polinomial orde 2 dan kesalahan estimasi latensi baca data menjadi 14.34% dengan pendekatan polinomial orde 3

KESIMPULAN

- Menaikan jumlah replikasi cluster SQL untuk SiELis dengan menurunkan nilai CPU dan memori tiap replika tidak memberikan nilai latensi yang lebih baik.
- SiELis menggunakan MongoDB memiliki latensi yang lebih baik dibandingkan menggunakan cluster SQL.
- Untuk perbandingan replikasi cluster SQL 1:2 SiELis menggunakan MySQL memiliki latensi baca data lebih baik dengan laju transaksi data dibawah 900/900 transaksi per detik, tetapi untuk laju transaksi diatas 900/900 SiELis dengan cluster SQL memiliki nilai latensi baca data yang lebih baik dibandingkan menggunakan MySQL.
- Untuk perbandingan replikasi cluster SQL 2:1 SiELis menggunakan MySQL memiliki latensi input data lebih baik dengan laju transaksi data dibawah 900/900 transaksi per detik, tetapi untuk laju transaksi diatas 900/900 SiELis dengan cluster SQL memiliki nilai latensi input data yang lebih baik dibandingkan menggunakan MySQL.