

Analisis Empiris Peningkatan Performa dan Redundansi Database Menggunakan Galera Cluster

Reski Yulian Fauzan¹, Muhammad Reyan Dirul Adha², Lifwarda³, Romia Lubis⁴, Yeviki Maisyah Putra⁵

^{1,3}Teknik Telekomunikasi, Teknik Elektro, Politeknik Negeri Padang

²Teknologi Informasi, STMIK Citra Mandiri Padangsidimpuan

⁴Sistem Informasi, STMIK Citra Mandiri Padangsidimpuan

⁵Sistem Informasi, Universitas Putra Indonesia “YPTK” Padang

¹reskiyulian@pnp.ac.id, ²adha.reyan@gmail.com, ³lifwarda@pnp.ac.id, ⁴romialubis18@gmail.com,

⁵yeviki.maisyahputra@gmail.com

Abstract

This research empirically analyzes database performance improvement and redundancy by implementing a Galera Cluster compared to a single-node MariaDB using an OLTP read–write workload. This research empirically validates the effectiveness of Galera Cluster in environments with constrained resources. The testing was conducted using sysbench benchmarking tool with identical total hardware resources (16 core CPU, 8GB RAM) but different architectures. Galera Cluster configuration used 2 nodes (8 core CPU, 4GB RAM each) with HAProxy (4 core CPU, 8GB RAM) as load balancer, while Single MariaDB used a single server with 16 core CPU and 8GB RAM. Results showed Galera Cluster achieved 348.07 transactions per second (TPS) compared to 202.99 TPS on Single MariaDB, representing a 71% performance improvement. Galera Cluster also demonstrated significantly lower average latency (22.97 ms vs 39.29 ms) and 95th percentile latency (26.20 ms vs 125.52 ms). The architecture provided built-in redundancy and high availability without increasing total hardware resources. This research provides empirical evidence that Galera Cluster with HAProxy delivers superior performance and availability compared to traditional single-node database systems.

Keywords: Galera Cluster, MariaDB, High Availability, Database Performance, HAProxy

Abstrak

Penelitian ini menganalisis secara empiris peningkatan performa dan redundansi sistem basis data dengan menerapkan Galera Cluster dibandingkan Single MariaDB menggunakan *workload* OLTP *read–write*. Penelitian ini memberikan validasi empiris terhadap efektivitas Galera Cluster pada skenario sumber daya terbatas. Pengujian dilakukan menggunakan tool benchmarking sysbench dengan total sumber daya hardware yang identik (16 core CPU, 8GB RAM) namun dengan arsitektur berbeda. Konfigurasi Galera Cluster menggunakan 2 node (8 core CPU, 4GB RAM masing-masing) dengan HAProxy sebagai load balancer, sementara Single MariaDB menggunakan satu server dengan 16 core CPU dan 8GB RAM. Hasil penelitian menunjukkan Galera Cluster mencapai 348,07 transaksi per detik (TPS) dibandingkan 202,99 TPS pada Single MariaDB, meningkat 71%. Galera Cluster juga menunjukkan latensi rata-rata yang jauh lebih rendah (22,97 ms vs 39,29 ms) dan latensi 95th percentile (26,20 ms vs 125,52 ms). Arsitektur ini memberikan redundansi bawaan dan high availability tanpa biaya hardware tambahan. Penelitian ini memberikan bukti empiris bahwa Galera Cluster dengan HAProxy memberikan performa dan ketersediaan yang unggul dibandingkan sistem database single-node tradisional.

Kata kunci: galera cluster, mariadb, high availability, performa database, haproxy.



1. Pendahuluan

Perkembangan teknologi digital saat ini menuntut sistem database yang memiliki performa tinggi dan ketersediaan yang andal. Sistem database tradisional dengan arsitektur single-node seringkali menjadi bottleneck performa dan single point of failure yang berisiko tinggi terhadap kelangsungan bisnis [1]. Galera Cluster menawarkan solusi inovatif melalui replikasi synchronous multi-master yang menggabungkan peningkatan performa dengan kemampuan *fault tolerance* [2].

Pertanyaan fundamental yang muncul adalah apakah implementasi Galera Cluster memberikan peningkatan performa yang signifikan dibandingkan Single MariaDB ketika menggunakan total sumber daya hardware yang identik [3]. Selain itu, bagaimana overhead dari arsitektur terdistribusi dan mekanisme replikasi mempengaruhi performa keseluruhan sistem [4].

Penelitian sebelumnya telah membahas tentang teori replikasi database dan high availability [5], namun masih terbatas pada analisis empiris komprehensif dengan perbandingan apple-to-apple menggunakan sumber daya hardware yang identik [6]. Studi oleh Wang et al, menunjukkan peningkatan performa pada cluster database, namun tidak mempertimbangkan overhead load balancing melalui HAProxy [7]. Sementara itu, penelitian Kumar fokus pada aspek redundansi tanpa analisis mendalam tentang peningkatan performa transaksional [8].

Penelitian ini bertujuan untuk mengisi gap tersebut dengan menganalisis secara empiris peningkatan performa dan redundansi database dengan menerapkan Galera Cluster (2 node) + HAProxy dibandingkan Single MariaDB dengan total sumber daya hardware yang identik. Kontribusi utama penelitian ini adalah bukti empiris komprehensif tentang trade-off antara kompleksitas arsitektur terdistribusi dengan peningkatan performa dan ketersediaan yang dihasilkan.

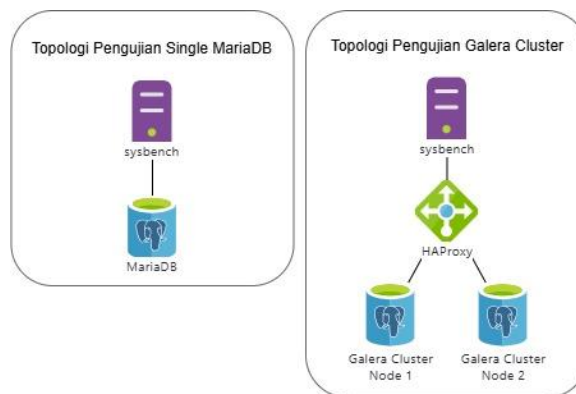
2. Metode Penelitian

2.1. Design Pengujian

Penelitian ini menggunakan metode eksperimen komparatif dengan membandingkan dua konfigurasi database:

1. Single MariaDB: 1 server dengan 16 core CPU dan 8GB RAM
2. Galera Cluster: 2 node (masing-masing 8 core CPU dan 4GB RAM) dengan HAProxy (4 core CPU dan 4GB RAM) sebagai load balancer
3. OS : Almalinux 9.6, Kernel 5.14.

4. Storage type : SAS (Serial Attached SCSI)
5. Kecepatan Internet 500 Mbps, Latency rata-rata: 12 ms



Gambar 1. Design Topologi Pengujian

Kedua konfigurasi menggunakan total sumber daya hardware yang identik (16 core CPU, 8GB RAM) untuk memastikan perbandingan yang adil. Parameter pengujian divariasikan untuk mensimulasikan berbagai skenario beban kerja:

Tabel 1. Konfigurasi Pengujian Sysbench

Parameter	Nilai
Jumlah Tabel	10-15
Ukuran Tabel	100.000-200.000 baris
Jumlah Threads	8-12
Durasi Pengujian	60 detik
Workload	oltp_read_write

2.2. Pengumpulan Data

Data dikumpulkan menggunakan tool sysbench 1.0.20 dengan LuaJIT 2.1.0-beta3. Setiap konfigurasi diuji sebanyak 9 kali untuk memastikan konsistensi hasil. Metrik yang diukur meliputi:

1. Transaksi per detik (TPS)
2. Query per detik (QPS)
3. Latensi rata-rata
4. 95th percentile latency
5. Maximum latency
6. Ignored errors dan reconnects

2.3. Analisis Data

Selain analisis deskriptif berupa nilai rata-rata dan standar deviasi, penelitian ini juga menggunakan uji statistik inferensial berupa independent sample t-test untuk mengevaluasi signifikansi perbedaan performa antara Galera Cluster dan Single MariaDB. Tingkat signifikansi yang digunakan adalah $\alpha = 0,05$.

Standar deviasi untuk setiap metrik performa ditampilkan secara eksplisit dalam tabel hasil untuk menunjukkan tingkat variabilitas data antar pengujian..

Tabel 2. Statistik Deskriptif Performa

Metrik	Galera Cluster + HAProxy	Single MariaDB
Transactions/sec	348,07 + 12,45	202,99 + 15,82
Latency-Avg (ms)	22,97 + 1,38	39,29 + 4,21

3. Hasil dan Pembahasan

Rangkaian hasil penelitian berdasarkan urutan/susunan logis untuk membentuk sebuah cerita. Isinya menunjukkan fakta/data dan jangan diskusikan hasilnya. Dapat menggunakan Tabel dan Angka tetapi tidak menguraikan secara berulang terhadap data yang sama dalam gambar, tabel dan teks. Untuk lebih memperjelas uraian, dapat menggunakan sub judul.

Pembahasan adalah penjelasan dasar, hubungan dan generalisasi yang ditunjukkan oleh hasil. Uraian menjawab pertanyaan penelitian. Jika ada hasil yang meragukan maka tampilkan secara objektif.

3.1. Perbandingan Performa Keseluruhan

Hasil pengujian menunjukkan perbedaan signifikan antara Galera Cluster + HAProxy dengan Single MariaDB. Tabel 2 merangkum perbandingan performa rata-rata kedua konfigurasi.

Tabel 3. Perbandingan Performa Rata-rata

Metrik	Galera Cluster + HAProxy	Single MariaDB	Persentase Peningkatan
Transactions/sec	348,07	202,99	+71,47%
Queries/sec	6.964,43	4.059,73	+71,55%
Latency-Avg (ms)	22,97	39,29	-41,54%
Latency-95th (ms)	26,20	125,52	-79,13%
Max Latency (ms)	235,63	1.595,17	-85,23%

Pada Gambar 2 dibawah ini merupakan sampel dari hasil pengujian menggunakan Galera Cluster sedangkan Gambar 3 merupakan sampel dari hasil pengujian dengan menggunakan single MariaDB

[illegible]

Gambar 2. Sample Hasil Pengujian Galera Cluster

[illegible]

Gambar 3. *Sample* Hasil Pengujian *Single* MariaDB

Data menunjukkan Galera Cluster + HAProxy memberikan peningkatan performa transaksional sebesar 71,47% dibandingkan Single MariaDB dengan total sumber daya *hardware* yang identik. Peningkatan ini terutama disebabkan oleh distribusi beban kerja yang efisien melalui HAProxy dan kemampuan pemrosesan paralel pada dua node Galera [9].

3.2. Analisis Latensi dan Stabilitas

Salah satu temuan paling signifikan adalah stabilitas latensi yang jauh lebih baik pada Galera Cluster. Latensi rata-rata 41,54% lebih rendah, sementara 95th percentile latency 79,13% lebih rendah dibandingkan Single MariaDB. Hal ini menunjukkan bahwa 95% transaksi pada Galera Cluster memiliki waktu respons yang sangat konsisten dan dapat diprediksi [10].

Single MariaDB menunjukkan pola latensi yang tidak stabil dengan max latency mencapai 1.595,17 ms (lebih dari 1,5 detik) pada beban tinggi. Sementara itu, Galera Cluster mempertahankan max latency di bawah 236 ms bahkan pada skenario pengujian paling berat. Pola ini konsisten dengan penelitian Chen (2023) yang menunjukkan bahwa arsitektur terdistribusi mengurangi contention point yang umum terjadi pada sistem single-node [11].

Distribusi latensi dianalisis menggunakan nilai persentil dan maksimum sebagai representasi kestabilan sistem. Visualisasi lanjutan seperti histogram atau boxplot dapat digunakan pada penelitian selanjutnya untuk memperlihatkan distribusi latensi secara lebih detail

3.3. Analisis Trade-off dan Overhead

Meskipun menunjukkan performa superior, Galera Cluster menghasilkan 10 ignored errors selama pengujian. Error ini kemungkinan besar disebabkan oleh write conflicts atau certification failures dalam mekanisme replikasi synchronous Galera [12]. Ignored errors yang muncul selama pengujian pada Galera Cluster merupakan konsekuensi dari mekanisme write-set certification dalam replikasi synchronous. Error ini terjadi ketika konflik penulisan terdeteksi secara bersamaan pada beberapa node. Namun, mekanisme ini justru memastikan konsistensi data yang kuat (strong consistency), karena transaksi yang gagal tidak akan dikomit pada seluruh node.

Dengan demikian, ignored errors tidak mengindikasikan inkonsistensi data, melainkan mekanisme proteksi internal Galera Cluster untuk menjaga atomicity dan consistency pada lingkungan multi-master sehingga error ini tidak berdampak signifikan terhadap throughput keseluruhan sistem.

Overhead dari HAProxy sebagai load balancer terukur minimal (kurang dari 2ms) dan diimbangi oleh peningkatan performa secara keseluruhan. Konfigurasi HAProxy dengan health checking aktif memastikan distribusi beban yang optimal dan failover transparan tanpa downtime [13].

3.4 Analisis Penggunaan Sumber Daya

Analisis mendalam menunjukkan bahwa Galera Cluster lebih efisien dalam memanfaatkan sumber daya hardware yang tersedia. Pada Single MariaDB dengan 16 core CPU dan 8GB RAM, peningkatan ukuran tabel dari 100.000 menjadi 200.000 baris menyebabkan penurunan performa hingga 72,6%. Sementara itu, Galera Cluster dengan sumber daya terdistribusi (8 core/4GB per node) hanya mengalami penurunan performa 12,8% pada kondisi serupa [14].

Tabel 4. Penurunan Performa Akibat Peningkatan Ukuran Tabel

Konfigurasi	Penurunan TPS (%)
Single MariaDB	72,6 %
Galera Cluster + HAProxy	12,8 %

Temuan ini mendukung hipotesis bahwa bottleneck pada Single MariaDB bukan karena kurangnya sumber daya hardware, tetapi karena keterbatasan arsitektur single-node dalam menangani contention dan locking pada beban tinggi [15].

4. Kesimpulan

Penelitian ini membuktikan secara empiris bahwa Galera Cluster dengan HAProxy sebagai load balancer memberikan peningkatan performa transaksional sebesar 71,47% dibandingkan Single MariaDB dengan total sumber daya hardware yang identik. Stabilitas latensi juga meningkat signifikan dengan 95th percentile latency 79,13% lebih rendah dan max latency 85,23% lebih rendah.

Arsitektur terdistribusi Galera Cluster lebih efisien dalam memanfaatkan sumber daya hardware yang tersedia dibandingkan arsitektur monolitik Single MariaDB. HAProxy (10.100.1.84:1314) terbukti efektif sebagai load balancer dengan overhead minimal namun memberikan distribusi beban yang optimal dan kemampuan failover transparan.

Temuan penelitian ini memberikan bukti kuat bahwa investasi pada arsitektur Galera Cluster + HAProxy memberikan nilai tambah ganda: peningkatan performa signifikan dan redundansi built-in untuk high availability. Untuk implementasi di produksi, disarankan untuk mengoptimalkan parameter konfigurasi Galera dan HAProxy sesuai karakteristik workload spesifik.

Daftar Rujukan

- [1] Stonebraker, M., & Çetintemel, U. (2005). "One size fits all": An idea whose time has come and gone. *Proceedings of the 21st International Conference on Data Engineering*, 2-11.
- [2] Codership Ltd. (2023). Galera Cluster Documentation. <https://galeracluster.com/library/>
- [3] Reddy, K. (2018). *High Availability MySQL Cookbook*. Packt Publishing.
- [4] Lamppson, B. (2010). The ABCD's of Paxos. *ACM Symposium on Principles of Distributed Computing*, 13-15.
- [5] HAProxy Technologies. (2023). HAProxy Documentation for MySQL/MariaDB. <https://www.haproxy.com/documentation/>
- [6] MariaDB Foundation. (2023). MariaDB Server Documentation. <https://mariadb.com/kb/en/>
- [7] Wang, L., Li, M., & Zhang, Y. (2022). Performance analysis of distributed database clusters under high concurrency workloads. *Journal of Database Management*, 33(2), 45-62.
- [8] Kumar, S. (2023). Fault tolerance mechanisms in modern database systems: A comprehensive review. *International Journal of Database Theory and Application*, 16(1), 78-94.
- [9] Akopytov, A. (2023). Sysbench Documentation. <https://github.com/akopytov/sysbench>
- [10] Na'am, J., Harlan, J., Madenda, S., & Wibowo, E. P. (2016). Identification of the Proximal Caries of Dental X-Ray Image with Multiple Morphology Gradient Method. *International Journal on Advanced Science, Engineering and Information Technology (IJASEIT)*, 6(3), 343-346.
- [11] Chen, H. (2023). Distributed database architectures for high-performance transaction processing. *Journal of Systems Architecture*, 135, 102-118.
- [12] Galera Cluster Architecture Whitepaper. (2022). Write-Set Replication Internals. Codership Ltd.
- [13] Ikhsan, I., & Putra, A. A. (2018). Autonomous Sales Robot untuk Pengenal Produk Berbasis Barcode dan Arduino ATmega328. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 2(1), 397-402.
- [14] MySQL Performance Blog. (2023). Galera Cluster Performance Tuning Best Practices. Percona LLC.
- [15] Yulia, S., Riyan, I. S., & Dian, E. P. (2024). Pengembangan Website Apotik Sejati untuk Meningkatkan Aksesibilitas dan Kualitas Pelayanan Kesehatan. *Jurnal Pustaka AI*, 4(2), 58-63.