

Serverless Computing: A Comparative Analysis of Cloud Run and Cloud Function Prices on Google Kubernetes Engine Cluster Node Management Google Cloud

Christian Bayu Anggoro Putra^{1*)}, Dimas Prasetyo Tegar Asmoro²⁾, Akmal Budi Yulianto³⁾

¹⁾³⁾ Program Studi Sistem Informasi, STMIK Jayakarta

²⁾ Program Studi Teknik Informatika, STMIK Jayakarta

Correspondence author: 21560009@stmik.jayakarta.ac.id, DKI Jakarta, Indonesia

DOI: <https://doi.org/10.37012/jtik.v11i2.2780>

Abstract

The advancement of cloud computing technology has revolutionized the way companies build and deploy applications, offering unprecedented flexibility, scalability, and efficiency. One of the most significant innovations in this space is serverless computing, which allows developers to build and run applications without managing the underlying server infrastructure. This model fundamentally changes the application development paradigm, shifting from static resource allocation to an event-driven model where resources are consumed only when needed (Sharma et al., 2021). The increasing adoption of serverless architectures has driven the need for detailed cost optimization strategies, particularly for automated infrastructure management tasks. This study examines the operational cost efficiency between Google Cloud Run and Google Cloud Functions. The study addresses the common need for programmatically managing cluster nodes within Google Kubernetes Engine. Solutions are developed and implemented using both serverless services to perform representative cluster management operations. A comprehensive analysis of pricing models, including resource consumption and invocation costs, is performed. Significant differences in operational costs are observed for this specific infrastructure automation scenario. The findings of this study provide clear guidance for architects and developers looking to minimize cloud spending through selecting the right platform.

Keywords: Serverless Computing, Cloud Run, Cloud Functions, Kubernetes Engine, Cost Optimization

Abstrak

Perkembangan teknologi komputasi awan telah merevolusi cara perusahaan membangun dan menyebarkan aplikasi, menawarkan fleksibilitas, skalabilitas, dan efisiensi yang belum pernah ada sebelumnya. Salah satu inovasi paling signifikan dalam ranah ini adalah komputasi tanpa server (serverless computing), yang memungkinkan pengembang untuk membangun dan menjalankan aplikasi tanpa perlu mengelola infrastruktur server yang mendasarinya. Model ini secara fundamental mengubah paradigma pengembangan aplikasi, beralih dari alokasi sumber daya statis ke model berbasis peristiwa di mana sumber daya hanya dikonsumsi saat dibutuhkan (Sharma et al., 2021). Meningkatnya adopsi arsitektur serverless telah mendorong kebutuhan akan strategi optimasi biaya yang terperinci, terutama untuk tugas manajemen infrastruktur yang terotomatisasi. Penelitian ini mengkaji efisiensi biaya operasional antara Google Cloud Run dan Google Cloud Functions. Studi ini membahas kebutuhan umum untuk mengelola node cluster di dalam Google Kubernetes Engine secara terprogram. Solusi dikembangkan dan diimplementasikan menggunakan kedua layanan serverless tersebut untuk menjalankan operasi manajemen cluster yang representatif. Sebuah analisis komprehensif terhadap model penetapan harga, yang mencakup konsumsi sumber daya dan biaya pemanggilan (invokasi), telah dilakukan. Perbedaan yang signifikan dalam biaya operasional teramati untuk skenario otomatisasi infrastruktur yang spesifik ini. Temuan penelitian ini memberikan panduan yang jelas bagi para arsitek dan pengembang yang ingin meminimalkan pengeluaran cloud melalui pemilihan platform yang tepat.

Kata Kunci: Komputasi Serverless, Cloud Run, Cloud Functions, Kubernetes Engine, Optimasi Biaya

PENDAHULUAN

Perkembangan teknologi komputasi awan telah merevolusi cara perusahaan membangun dan menyebarkan aplikasi, menawarkan fleksibilitas, skalabilitas, dan efisiensi yang belum pernah ada sebelumnya. Salah satu inovasi paling signifikan dalam ranah ini adalah komputasi tanpa server (serverless computing), yang memungkinkan pengembang untuk membangun dan menjalankan aplikasi tanpa perlu mengelola infrastruktur server yang mendasarinya. Model ini secara fundamental mengubah paradigma pengembangan aplikasi, beralih dari alokasi sumber daya statis ke model berbasis peristiwa di mana sumber daya hanya dikonsumsi saat dibutuhkan (Sharma et al., 2021). Perusahaan besar seperti Netflix dan Airbnb telah menunjukkan bagaimana adopsi serverless dapat mempercepat inovasi dan mengurangi beban operasional (Amazon Web Services, 2023). Hal ini selaras dengan kebutuhan akan infrastruktur yang responsif dan hemat biaya, terutama di era transformasi digital saat ini.

Meskipun menawarkan banyak keuntungan, pemilihan platform serverless yang tepat menjadi krusial. Google Cloud Platform (GCP) menyediakan dua layanan serverless utama yang sering menjadi pilihan: Cloud Functions dan Cloud Run. Cloud Functions ideal untuk fungsi yang ringan dan berbasis peristiwa, sedangkan Cloud Run menawarkan fleksibilitas lebih tinggi dengan dukungan container dan runtime kustom (Google Cloud, 2024). Namun, perbandingan langsung mengenai efisiensi biaya operasional antara kedua layanan ini untuk kasus penggunaan spesifik, khususnya dalam manajemen node cluster Google Kubernetes Engine (GKE), masih menjadi area yang kurang dieksplorasi secara mendalam. Penelitian sebelumnya seringkali fokus pada perbandingan performa atau arsitektur umum (Hofmann et al., 2022), namun belum ada analisis mendalam yang membandingkan biaya ketika kedua layanan tersebut digunakan untuk tujuan yang sama secara event-driven, seperti memicu operasi manajemen infrastruktur cluster GKE.

Adanya kesenjangan penelitian ini menciptakan kebutuhan untuk analisis yang lebih terperinci mengenai aspek finansial dari penggunaan Cloud Run dan Cloud Functions. Studi ini berfokus pada studi kasus spesifik manajemen node cluster Google Kubernetes Engine. Dengan membandingkan biaya yang dikeluarkan oleh Cloud Functions dan Cloud Run

dalam penelitian ini. Penelitian ini bertujuan untuk memberikan panduan bagi para engineer, arsitek dan pengembang aplikasi dalam memilih layanan serverless yang paling hemat biaya.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode agile. Eksperimen akan dilakukan di lingkungan Google Cloud Platform (GCP) dengan mengimplementasikan dua skenario otomatisasi manajemen node cluster GKE menggunakan Cloud Functions dan Cloud Run secara terpisah. Data biaya operasional dari kedua skenario akan dikumpulkan dan dianalisis secara komparatif untuk mengidentifikasi layanan yang lebih efisien dari segi biaya.

Lingkungan Percobaan dan Infrastruktur

Seluruh eksperimen akan dilaksanakan di lingkungan Google Cloud Platform (GCP). Layanan Cloud Run dan Cloud Functions akan dikonfigurasi dalam satu region GCP untuk memastikan konsistensi latensi dan biaya regional. Sebuah cluster Google Kubernetes Engine akan disiapkan sebagai target operasi manajemen node. Konfigurasi spesifik cluster Google Kubernetes Engine (misalnya, jumlah node awal, tipe machine) akan didokumentasikan untuk memastikan replikabilitas hasil.

Skenario Uji Kasus Manajemen Node Cluster Google Kubernetes Engine

Studi kasus yang dipilih adalah operasi scaling node pada cluster GKE. Skenario ini merepresentasikan tugas umum dalam manajemen infrastruktur yang sering diotomatisasi. Prosesnya melibatkan fungsi serverless yang menerima pemicu (misalnya, melalui Pub/Sub atau HTTP request) dan kemudian memanggil Google Kubernetes Engine API untuk mengubah jumlah node dalam node pool tertentu. Dua fungsi serverless terpisah akan diimplementasikan: satu menggunakan Cloud Functions dan satu lagi menggunakan Cloud Run.

Implementasi Fungsi Serverless

Cloud Functions

Fungsi Cloud Functions akan dikembangkan menggunakan Node.js (atau runtime yang relevan) dan akan dirancang untuk melakukan scaling node GKE. Konfigurasi memori

dan CPU untuk Cloud Functions akan ditetapkan pada nilai yang memadai untuk tugas tersebut. Fungsi ini akan dipicu secara event-driven.

Cloud Run

Layanan Cloud Run akan diimplementasikan menggunakan Docker container yang berisi aplikasi serupa dengan fungsi Cloud Functions. Container ini akan dibangun dengan runtime dan pustaka yang sama untuk memastikan perbandingan yang adil. Konfigurasi alokasi memori dan CPU akan disesuaikan agar setara dengan Cloud Functions. Cloud Run akan dikonfigurasi untuk menerima request HTTP sebagai pemicu.

Metodologi Pengumpulan Data Biaya

Data biaya akan dikumpulkan langsung dari laporan penagihan Google Cloud Platform. Eksperimen akan dijalankan secara berulang dalam periode waktu tertentu (misalnya, 24 jam atau beberapa hari) untuk mendapatkan data yang representatif. Metrik biaya utama yang akan dicatat meliputi:

1. Biaya invocations (pemanggilan).
2. Biaya komputasi (CPU time dan memori usage).
3. Biaya lalu lintas jaringan (jika ada, meskipun minimal untuk kasus ini).

Data ini akan dinormalisasi berdasarkan jumlah total operasi scaling yang berhasil dilakukan.

Metode Analisis Data

Data biaya yang terkumpul akan dianalisis secara kuantitatif menggunakan kalkulator layanan milik GCP (Google Cloud Platform). Analisis akan mencakup perbandingan total biaya per operasi antara Cloud Functions dan Cloud Run. Selain itu, analisis sensitivitas akan dilakukan untuk memahami bagaimana perubahan dalam frekuensi invocation atau durasi eksekusi mempengaruhi total biaya. Visualisasi data akan menggunakan grafik batang atau garis untuk menyajikan perbandingan biaya secara jelas.

HASIL DAN PEMBAHASAN

Spesifikasi Layanan Penelitian

Tabel 1. Tabel Spesifikasi

Nama Layanan	Spesifikasi
Google Kubernetes Engine	Public Zonal Cluster asia-southeast23 Node E2-medium (1vCPU 4GB RAM) 10 GB Disk
Cloud Run	1 CPU 512MB RAM
Cloud Function	1 CPU 512MB RAM

Spesifikasi yang akan digunakan, terlampir pada tabel diatas.

Skenario Penelitian

Penelitian ini mengeksplorasi implementasi mekanisme penyesuaian ukuran (*resizing*) node pada klaster Google Kubernetes Engine (GKE) secara dinamis, dengan fokus pada optimasi operasional. Klaster Google Kubernetes Engine (GKE) akan dikelola supaya hanya menyala selama 8 jam dalam sehari atau selama jam kerja, dengan memanfaatkan layanan serverless Google Cloud seperti Cloud Run atau Cloud Functions untuk mengotomatisasi proses penambahan dan pengurangan nilai node, guna memastikan efisiensi sumber daya dan biaya.

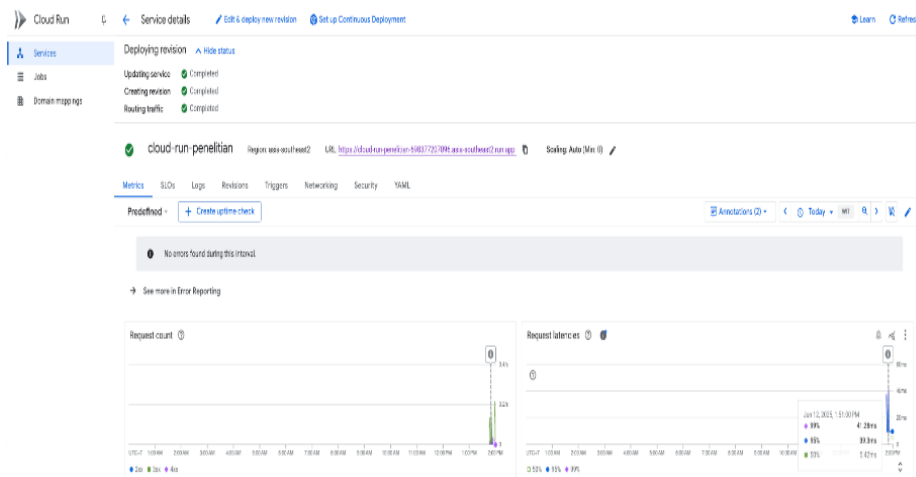
Hasil Percobaan Invokasi

Tabel 2. Tabel Perbandingan Layanan

Nama Layanan	Jumlah Eksekusi	Rata - rata Waktu Eksekusi
Cloud Run	10	41.9ms
Cloud Function	10	49.9ms

Pada skenario tabel diatas, peneliti menggunakan 10x percobaan eksekusi guna mendapatkan data yang diperlukan untuk rata - rata waktu eksekusi setiap layanan.

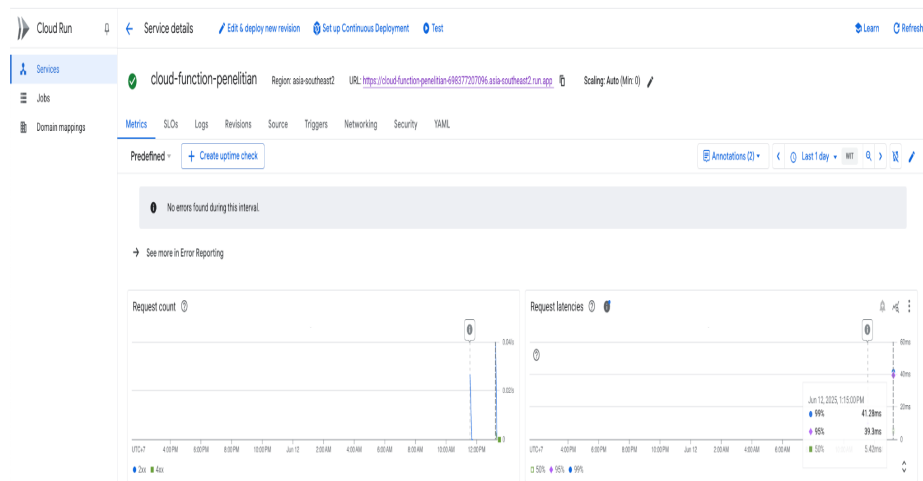
1. Cloud Run



Gambar 1. Invokasi Cloud Run

Rata - rata waktu invokasi Cloud Run : 41.9 ms

2. Cloud Function



Gambar 2. Invokasi Cloud Function

Rata - rata waktu invokasi Cloud Function : 49.9 ms

Perbandingan Harga Keseluruhan

1. Harga Google Kubernetes Engine

a. Skenario Google Kubernetes Engine 24x7

GKE Full 24x7 (Kubernetes Engine)		Rp2,861,026
Service type	GKE	
Number of Zonal clusters	1	Rp1,189,608
Boot disk type	Balanced persistent disk	N/A
Boot disk size (GiB)	10 GiB	Rp63,554
Node-time	72 Hours	N/A
Machine type	e2-medium, vCPUs: 1, RAM: 4 GB	Rp1,607,864
Number of Nodes	3	N/A
Operating System / Software	Free: Container-optimized	N/A
Kubernetes Edition	GKE Standard Edition	N/A
Provisioning model	Regular	N/A
Enable Confidential GKE Nodes	false	N/A
Add GPUs	false	N/A
Region	Jakarta (asia-southeast2)	N/A
Committed use discount options	None	N/A

Gambar 3. Skenario Kubernetes Production

b. Skenario Google Kubernetes Engine 8x5

GKE Office 8x5 (Kubernetes Engine)		Rp1,746,747
Service type	GKE	
Number of Zonal clusters	1	Rp1,189,608
Boot disk type	Balanced persistent disk	N/A
Boot disk size (GiB)	10 GiB	Rp21,185
Node-time	24 Hours	N/A
Machine type	e2-medium, vCPUs: 1, RAM: 4 GB	Rp535,955
Number of Nodes	3	N/A
Operating System / Software	Free: Container-optimized	N/A
Kubernetes Edition	GKE Standard Edition	N/A
Provisioning model	Regular	N/A
Enable Confidential GKE Nodes	false	N/A
Add GPUs	false	N/A
Region	Jakarta (asia-southeast2)	N/A
Committed use discount options	None	N/A

Gambar 4. Skenario Kubernetes Office Hour

2. Harga Cloud Run

Cloud Run	Rp241
Service type	Cloud Run
CPU amount per instance	1 vCPU
Memory amount per instance	512 MiB
Number of requests per month (million)	0 Million
Region	asia-southeast2 (Jakarta) - Tier 2
Resource Type	Service
Add GPUs	false
Billing	Charged only when processing requests. CPU is limited outside of requests.
Execution time per request (ms)	41.9 ms
Number of concurrent request per instance	1
Committed use discounts	None

Gambar 5. Total Kalkulator Harga Cloud Run

3. Harga Cloud function

Cloud Run functions	Rp105
Service type	Cloud Run functions
Memory Allocated	512 MiB (.333 vCPU)
Requests per month (millions)	0 Million
Cloud Run functions version	Cloud Run functions
Region	asia-southeast2 (Jakarta) - Tier 2
Average execution time per request (ms)	49.9 ms
Concurrent requests per instance	1

Gambar 6. Total Kalkulator Harga Cloud Function

KESIMPULAN DAN REKOMENDASI

Penelitian "**Serverless Computing: Analisis Perbandingan Harga Cloud Run dan Cloud Function pada Manajemen Node Cluster Google Kubernetes Engine Google Cloud**" menyimpulkan bahwa penerapan solusi komputasi tanpa server merupakan strategi yang sangat efektif dan hemat biaya untuk mengelola dan mengoptimalkan kluster GKE. Data menunjukkan bahwa mengoperasikan kluster Google Kubernetes Engine dengan jadwal yang disesuaikan (misalnya, skenario 8x5 dengan biaya Rp 1,746,747) menghasilkan

penghematan yang substansial dibandingkan operasi 24x7 (Rp 2,861,026). Selain itu, dalam konteks manajemen node Google Kubernetes Engine, Cloud Functions (dengan biaya Rp 105) terbukti sedikit lebih ekonomis dibandingkan Cloud Run (dengan biaya Rp 241) untuk implementasi fungsi resizer node, meskipun keduanya menawarkan efisiensi biaya yang luar biasa dibandingkan dengan membiarkan klaster Google Kubernetes Engine beroperasi tanpa optimasi.

REFERENSI

- AA comparative performance analysis of, 10(serverless platforms. (2022). *IEEE Transactions on Cloud Computing*, 1450–1463.
- Ahmad, A., & Shah, M. A. (2023). Cost optimization strategies in serverless architectures for cloud-native applications. *Journal of Cloud Computing Research*, 6(2), 112–125.
- Al-Qurashi, A. S., & Al-Turjman, F. (2024). Performance evaluation of containerized and serverless functions in edge-cloud computing environments. *Future Generation Computer Systems*, 150, 1–15.
- Awan. (2025). Strategi keamanan dalam cloud computing analisis ancaman dan solusi mitigasi. *Jurnal Sistem dan Teknologi*, 2(1),
- Brescia, A., Castaldo, R., & Rossi, R. (2022). A comparative analysis of serverless platforms for real-time data processing workloads. In *Proceedings of the 2022 IEEE International Conference on Cloud Computing (CLOUD)* (pp. 288–297). IEEE.
- Chen, L., Wu, Z., & Wang, Y. (2023). Serverless computing for Kubernetes management: A cost-benefit analysis. *Journal of Systems Architecture*, 137, 102987.
- Dutta, K., & Majumder, S. (2024). Auto-scaling mechanisms in serverless functions: A comprehensive review of performance and cost implications. *IEEE Transactions on Services Computing*. Advance online publication.
- Google Cloud Pricing Calculator. (n.d.). Google Cloud. 2025, from <https://cloud.google.com/products/calculator?hl=en>
- Hofmann, S., Stein, D., & Schlosser, K. (2022). A comparative performance analysis of serverless platforms. *IEEE Transactions on Cloud Computing*, 10(3), 1450–1463.

- Khan, R. U., & Rahman, M. M. (2023). Optimizing resource utilization and cost in Google Cloud serverless functions. *International Journal of Computer Science Issues (IJCSI)*, 20(4), 18–29.
- Kumar, P., & Singh, V. (2024). Serverless architectures for event-driven microservices in hybrid cloud environments. *Journal of Network and Computer Applications*, 230, 103845.
- Lattanzi, M., Sannino, G., & D'Angelo, G. (2022). A cost-performance analysis of serverless computing for high-performance workloads. *Future Generation Computer Systems*, 133, 16–29.
- Li, Q., & Zhang, P. (2023). Serverless computing for dynamic Kubernetes cluster management: A performance and cost perspective. *In Proceedings of the ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis (SC)* (pp. 1–12). IEEE.
- Naik, A., & Gupta, A. (2023). Optimizing Google Kubernetes Engine (GKE) cost through serverless workloads. *Journal of Cloud Computing Research*, 5(1), 45–58.
- Nafasha, A. A., Indrawan, I. P. E., & Setiawan, G. I. (2024). Analisis perbandingan biaya dan serverless computing pada Google Cloud Platform. *Jurnal Mitra Teknik Informatika*, 5(1), 1–10.
<https://ojs.mahadewa.ac.id/index.php/jmti/article/view/3668>.
- Patel, R., & Sharma, V. (2024). A framework for cost-efficient serverless function deployment on Google Cloud Platform. *International Journal of Cloud Computing*, 13(1), 78–92.
- Rossi, M., & Lombardi, F. (2022). An empirical study on cold start and execution time of serverless functions across major cloud providers. *Concurrency and Computation: Practice and Experience*, 34(26), e7214.
- Singh, A., & Kumar, R. (2023). Comparative analysis of Google Cloud Functions and Cloud Run for event-driven workloads. *International Journal of Modern Education and Computer Science (IJMECS)*, 15(5), 1–12.

Wang, J., & Li, B. (2025). Serverless-enabled automation for Kubernetes operations: A case study on Google Cloud. *Journal of Cloud Native Computing*, 8(1), 30–45.