

COMPARATIVE ANALYSIS OF CPU AND REQUEST PER-SECOND (RPS) METRICS IN KUBERNETES HORIZONTAL POD AUTOSCALER (HPA)

SKRIPSI



**PROGRAM STUDI ILMU KOMPUTER (S1)
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KEJURUAN
UNIVERSITAS PENDIDIKAN GANESHA**



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



**Balai Besar
Sertifikasi
Elektronik**



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



**Balai Besar
Sertifikasi
Elektronik**

SKRIPSI

DIAJUKAN UNTUK MELENGKAPI TUGAS DAN MEMENUHI SYARAT-SYARAT UNTUK MENCAPAI GELAR SARJANA KOMPUTER

Menyetujui

Pembimbing I	Kadek Yota Ernanda Aryanto, S.Kom., M.T., Ph.D. NIP.197803242005011001
Pembimbing II	I Ketut Purnamawan, S.Kom., M.Kom. NIP.197905112006041004



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



**Balai Besar
Sertifikasi
Elektronik**

Skripsi oleh I Gusti Putu Ranga Adithia Putra ini
telah dipertahankan di depan dewan penguji
Pada tanggal 24 Juni 2026

Dewan Penguji

Ketua	Ir. I Ketut Resika Arthana, S.T., M.Kom. NIP.198412012012121002
Anggota	Ir. Ketut Agus Seputra, S.ST.,M.T. NIP.199008152019031018
Anggota	Kadek Yota Ernanda Aryanto, S.Kom., M.T., Ph.D. NIP.197803242005011001
Anggota	I Ketut Purnamawan, S.Kom., M.Kom. NIP.197905112006041004



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



**Balai Besar
Sertifikasi
Elektronik**

Diterima oleh Panitia Ujian Fakultas Teknik dan Kejuruan
Universitas Pendidikan Ganesha
guna memenuhi syarat-syarat untuk mencapai gelar Sarjana Komputer

Menyetujui

Ketua Ujian	Ir. Made Windu Antara Kesiman, S.T., M.Sc., Ph.D. NIP.198211112008121001
Sekretaris Ujian	I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs. NIP.198910262019031004



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



**Balai Besar
Sertifikasi
Elektronik**

STATEMENT OF AUTHENTICITY

I hereby declare that this academic work entitled "Comparative Analysis of CPU Based and Request Per-Second (RPS) Metrics in Kubernetes Horizontal Pod Autoscaler (HPA)" and all its contents are truly my own work, and I have not committed plagiarism or cited in ways inconsistent with the ethics prevailing in the academic community. I am prepared to bear any risk/sanction imposed on me if violations of academic ethics are found in this work, or if there are claims against its authenticity.

Singaraja, 13th July 2026

Statement made by,



I Gusti Putu Ranga Adithia Putra

NIM. 2215101001

ACKNOWLEDGMENT

All the praise and gratitude the author extends to the Almighty God, for His bless that this undergraduate thesis, entitled "**Comparative Analysis of CPU Based and Request Per-Second (RPS) Metrics in Kubernetes Horizontal Pod Autoscaler (HPA)**", could be completed. This undergraduate thesis is submitted as a requirement to obtain a bachelor's degree in computer science at Ganesha University of Education. In the process of completing this undergraduate thesis, the author has received significant support. Therefore, the author would like to express sincere gratitude to:

1. Prof. Dr. Kadek Rihendra Dantes, S.T., M.T., Dean of the Faculty of Engineering and Vocational, for his encouragement and the facilities provided, which enabled the author to complete the study as planned.
2. Made Windu Antara Kesiman, S.T., M.Sc., Ph.D., Vice Dean I of the Faculty of Engineering and Vocational Studies, for the motivation and facilities provided, which enabled the author to complete this undergraduate thesis.
3. Prof. Dr. Ketut Agustini, S.Si., M.Si., Vice Dean III of the Faculty of Engineering and Vocational Studies, for the motivation provided.
4. Dr. Putu Hendra Suputra, S.Kom., M.Cs., Head of the Informatics Engineering Department, for the motivation given throughout the completion of this undergraduate thesis.
5. Ir. I Ketut Resika Arthana, S.T., M.Kom., Head of UPA TIK UNDIKSHA and also Examiner 1, for the feedback and evaluation provided on the author's undergraduate thesis.
6. Ir. Ketut Agus Seputra, S.ST., M.T., Coordinator of the Bachelor's Program in Computer Science and also Examiner 2, for the feedback and evaluation provided on the author's undergraduate thesis.
7. Kadek Yota Ernanda Aryanto, S.Kom., M.T., Ph.D., as the first academic advisor, for her guidance, direction, advice, and motivation throughout the undergraduate thesis writing process.
8. I Ketut Purnamawan, S.Kom., M.Kom., as the second academic advisor, for his guidance, direction, advice, and motivation during the completion of this work.

9. The staff members of the Faculty of Engineering and Vocational Studies who assisted in the administrative processes.
10. I Gusti Nyoman Sutresna and Ni Ketut Suliasih, the author's parents, along with the author's siblings Krisnanda and Ayu Trisna, for their unwavering support material, motivational, and spiritual throughout the author's academic journey.
11. The author's close friends, Satrya Pudja, Fieter Brain, Restu Cahyana, for their support during both joyful and challenging moments throughout the completion of this undergraduate thesis.
12. My fellow comrades from the IKI Class, and friends whose names cannot be mentioned individually, for their contributions in supporting and encouraging the author during the preparation of this undergraduate thesis

The author acknowledges that this thesis is not without its shortcomings and sincerely welcomes constructive feedback. It is hoped that this work may offer meaningful contributions to technology development.

Singaraja, 13th July 2026

Author

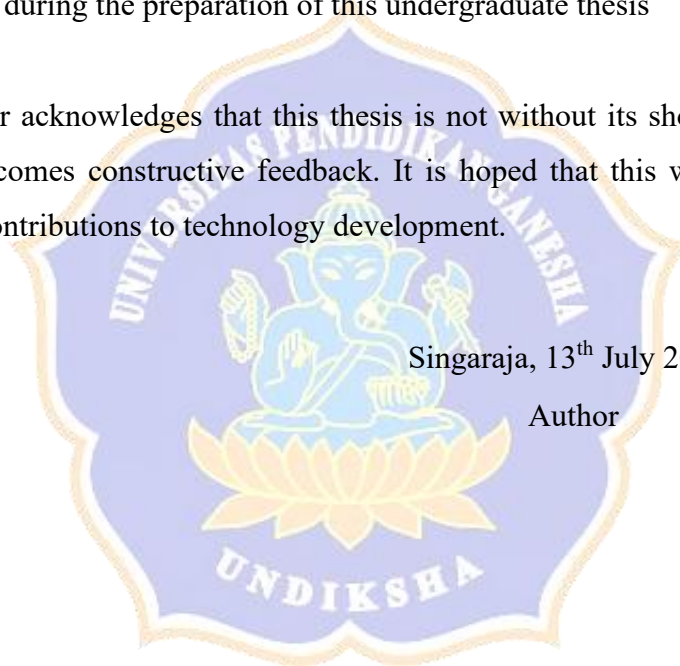
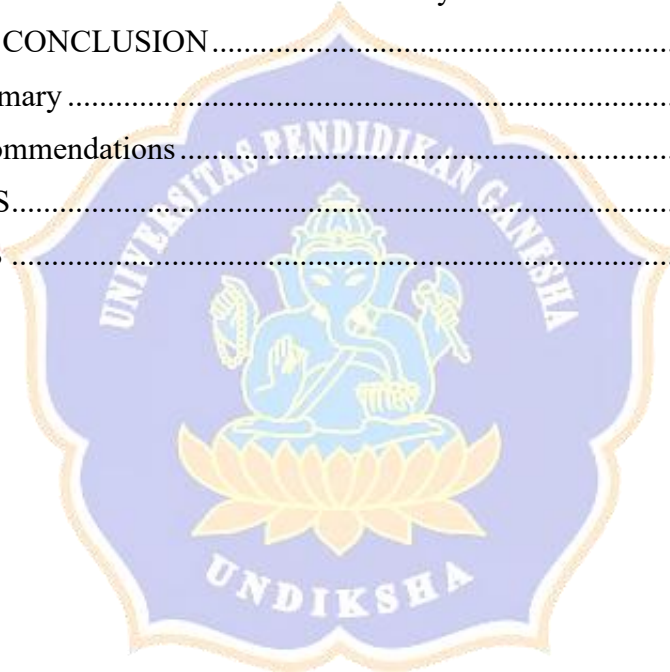


TABLE OF CONTENT

ACKNOWLEDGMENT	i
TABLE OF CONTENT	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF APPENDIXES	viii
CHAPTER I INTRODUCTION	1
1.1 Research Background.....	1
1.2 Research Identification.....	4
1.3 Research Questions	5
1.4 Problem Limitations.....	6
1.5 Research Objective.....	7
1.6 Research Significances.....	8
1.6.1 Theoretical Benefits	8
1.6.2 Practical Benefits	8
CHAPTER II THEORETICAL FRAMEWORK	9
2.1 Theoretical Framework	9
2.1.1 Cloud Computing.....	9
2.1.2 Containerisation	10
2.1.3 Kubernetes.....	11
2.1.3.1 Kubernetes Architecture and Components.....	11
2.1.3.2 Kubernetes Autoscaling	12
2.1.3.2.1 Horizontal Pod Autoscaling (HPA)	12
2.1.3.2.2 Custom Metrics Autoscaling.....	14
2.1.4 Microservice Architecture.....	15
2.1.5 Load Testing and Traffic Patterns.....	15
2.2 Literature Review	16
2.3 Framework of Thinking	22
2.4 Research Hypothesis	23
CHAPTER III RESEARCH METHODOLOGY	24
3.1 System Design.....	24
3.2 Testing Environment.....	25

3.3	Baseline Testing	26
3.4	Horizontal Pod Autoscaler (HPA) Configuration	27
3.5	Load Testing	28
3.5.1	Ramp-up Load	29
3.5.2	Sustained High-Load Scenario	29
3.5.3	Cyclic Load Scenario	30
3.5.4	Gradual Decreasing Load Scenario	31
3.6	Evaluation Metrics	31
3.7	Experimental Replication and Statistical Analysis	33
CHAPTER IV FINDINGS AND DISCUSSION		38
4.1	Research Environment and Implementation Setup	38
4.1.1	Hardware Specifications	38
4.1.2	Operating System	39
4.1.3	Development and Experimentation Tools	39
4.2	Experimental Results and Evaluation	42
4.2.1	Baseline Testing Results	43
4.2.2	HPA CPU Based Results under Ramp-up Load	45
4.2.3	HPA RPS Based Results under Ramp-up Load	46
4.2.4	HPA CPU Based Results under Sustained High Load	48
4.2.5	HPA RPS Based Results under Sustained High Load	50
4.2.6	HPA CPU Based Results under Cyclic Load	52
4.2.7	HPA RPS Based Results under Cyclic Load	53
4.2.8	HPA CPU Based Results under Gradual Decreasing Load	55
4.2.9	HPA RPS Based Results under Gradual Decreasing Load	58
4.3	Error and Bottleneck Analysis	62
4.3.1	Indications of Connection-Level Bottleneck in Nginx Reverse Proxy 62	
4.3.2	Latency Patterns Consistent with Fast-Fail Behavior and TCP Backlog Congestion	66
4.3.3	Limitations of CPU Metrics in Detecting I/O-Bound Bottlenecks	71
4.3.4	Metric Misalignment in CPU-Based Autoscaling for I/O-Bound Workloads	72
4.3.5	Proactive Responsiveness of RPS-Based Autoscaling	74
4.3.6	Statistical Hypothesis Testing	76

4.4	Post-Comparison Optimization of the RPS-Based HPA Configuration	77
4.4.1	Optimization Configuration	79
4.4.2	Optimized Scale-Up Behaviour Configuration.....	80
4.4.3	Optimized RPS-Based HPA Results under Sustained High Load	82
4.4.4	Optimized RPS-Based HPA Results under Cyclic Load	87
4.4.5	Summary of Optimization Impact.....	90
4.5	Comparison with Previous Research	92
4.6	Discussion	94
4.6.1	Key Insights on Autoscaling and System Bottlenecks.....	94
4.6.2	Strengths of the Experimental Design	95
4.6.3	Limitations and Threats to Validity	96
CHAPTER V CONCLUSION		100
5.1	Summary	100
5.2	Recommendations	104
REFERENCES.....		107
APPENDICES		110



LIST OF TABLES

Table 2. 1: The review of relevant research	17
Table 3. 1: HPA Configuration Details	28
Table 3. 2: Evaluation Metrics	32
Table 4. 1: Baseline Testing Result.....	44
Table 4. 2: HPA CPU Based Results under Ramp-up Load	46
Table 4. 3: HPA RPS Based Results under Ramp-up Load.....	48
Table 4. 4: HPA CPU Based Results under Sustained High Load	49
Table 4. 5: HPA RPS Based Results under Sustained High Load	51
Table 4. 6: HPA CPU Based Results under Cyclic Load	53
Table 4. 7: HPA RPS Based Results under Cyclic Load	55
Table 4. 8: HPA CPU Based Results under Gradual Decreasing Load	56
Table 4. 9: HPA RPS Based Results under Gradual Decreasing Load.....	59
Table 4. 10: Client–Backend Request Discrepancy and Average Active Connection under Sustained High Load.....	65
Table 4. 11: Summary of Paired T-Test Results (CPU vs RPS) at Peak Load	77
Table 4. 12: RPS-Based HPA Optimization Strategy	79
Table 4. 13: Optimized RPS-Based HPA Results under Sustained High Load...	83
Table 4. 14: Summary Comparison Before and After Optimization under Sustained High Load	84
Table 4. 15: Optimized RPS-Based HPA Results under Cyclic Load	88
Table 4. 16: Summary Comparison Before and After Optimization under Cyclic Load.....	89

LIST OF FIGURES

Figure 2. 1: Basic architecture of a Kubernetes cluster	13
Figure 2. 2: Framework of Thinking.....	23
Figure 3. 1: High Level System Design.....	24
Figure 3. 2: Load Testing Schema Ramp-up Load	29
Figure 3. 3: Load Testing Schema Sustained High-Load	30
Figure 4. 1: CPU Usage & Error Rate on Baseline Testing.....	63
Figure 4. 2: Latency CPU-based HPA under Ramp-up Load.....	68
Figure 4. 3: Latency RPS-based HPA under Ramp-up Load	69
Figure 4. 4: Latency CPU-based HPA under Sustained High-Load.....	70
Figure 4. 5: Latency RPS-based HPA under Sustained High-Load	71
Figure 4. 6: CPU Usage & Error Rate on High Sustain Load CPU-Based HPA	74
Figure 4. 7: Number of Pods on High Sustain Load CPU-Based HPA	74
Figure 4. 8: Error Rate Comparison Before and After Optimization under Sustained High Load.	85
Figure 4. 9: Pod Replica Comparison Before and After Optimization under Sustained High Load.....	85

LIST OF APPENDIXES

Appendix 1: Curriculum Vitae.....	110
Appendix 2: Frontend Deployment Config File	110
Appendix 3: HPA CPU Config File.....	112
Appendix 4: HPA RPS Config File	113
Appendix 5: Query Prometheus for getting Pod Replica Frontend	113
Appendix 6: Query Prometheus for getting CPU Usage Frontend.....	113
Appendix 7: Load testing script for Baseline & Ramp-up Load	114
Appendix 8: Load testing script for Cyclic Load.....	114
Appendix 9: Load testing script for Sustained High Load.....	115
Appendix 10: Load testing script for Gradual Decrease Load.....	116
Appendix 11: Load Testing Log	117

