Network Resources Optimization through Regional Computing forVehicular Big Data

Sajid Ullah Khan
sk.khan@psau.edu.sa

Prince Sattam Bin Abdulaziz University

Afzal Badshah
Hamdard University

Meshal Alharbi
Prince Sattam Bin Abdulaziz University

Tahir Saleem
Khushal Khan Khattak University Karak

Ghani Ur Reman
Khushal Khan Khattak University Karak

Faheem Khan
Gachon University

Article

Keywords: Vehicular big data, regional computing, network optimization, ITS

Posted Date: April 30th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4306263/v1

Additional Declarations: No competing interests reported.

Network Resources Optimization through Regional Computing for Vehicular Big Data

Afzal Badshah [1], Tahir Saleem [2], Ghani Ur Reman [3], Sajid Ullah Khan [4,*], Faheem Khan [5], and Meshal Alharbi [6]

[1]Department of Computing, Hamdard University, Islamabad Campus, Pakistan

[2,3]Department of Computer Science & Bioinformatics, Khushal Khan Khattak University, Karak, Pakistan

[4]Department of Information systems, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Alkharj 11942, Saudi Arabia

[5] Department of Computer Engineering, Gachon University, South Korea

[6]Department Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Alkharj 11942, Saudi Arabia

Abstract: The transportation system is shifting toward an Intelligent Transportation System (ITS). Research shows that a single Autonomous Vehicle (AV) can generate up to 1 Terabyte (TB) data daily. Data processing occurs in real-time at the vehicle edge to facilitate navigation. However, some of this data is sent to the cloud for real-time navigation support for other AVs. Even slight delays can have significant consequences for these vehicles. All AV data is also necessary in the cloud for navigation and training purposes. Unfortunately, the massive volume of Vehicular Big Data (VBD) congests public networks, making real-time communication unfeasible. This study proposes Regional Computing (RC) as a solution, enabling real-time updates and navigation for AVs. Compared to edge servers, RC offers sufficient resources and considerably reduced delays and costs compared to Cloud Computing (CC). RC substantially reduces communication delay, cost, and network congestion.

Keywords: Vehicular big data, regional computing, network optimization, ITS

1 Introduction

Recently, the automotive industry has witnessed unprecedented growth in data generation due to the proliferation of advanced sensors, onboard computers, and connected vehicle technologies [1]. This influx of data, as shown in Figure 1, termed 'Vehicular Big Data' (VBD) in this paper, presents a unique opportunity to extract valuable insights and enhance various aspects of the automotive domain. Statistics reveal that approximately 1.3 million individuals worldwide die in road accidents yearly. A staggering 95% of these accidents can be attributed to human errors [2], Figure 2 shows the forecast of Autonomous Vehicle (AV). However, the emergence of AVs offers a promising solution by mitigating human mistakes and enhancing road safety. These vehicles leverage advanced technology to analyze road conditions and surrounding environments, reducing costs and environmentally friendly transportation [3]. Key players in this field, including Tesla [4], Waymo [5], Motional [6], and Ford [7], are actively working on autopilot vehicle development. The number of AVs on the road is anticipated to exceed predictions by reaching a significant milestone by 2025 [8].

AVs heavily rely on a multitude of sensors and high-resolution cameras for detecting and avoiding obstacles, resulting in the generation of massive amounts of big data [9]. According to statistics, an average AV produces around 1Terabyte (TB) data daily [10]. While the vehicle's onboard powerful computer processes a significant portion of this data, specific information must be communicated with external servers . These vehicles require access to high-resolution

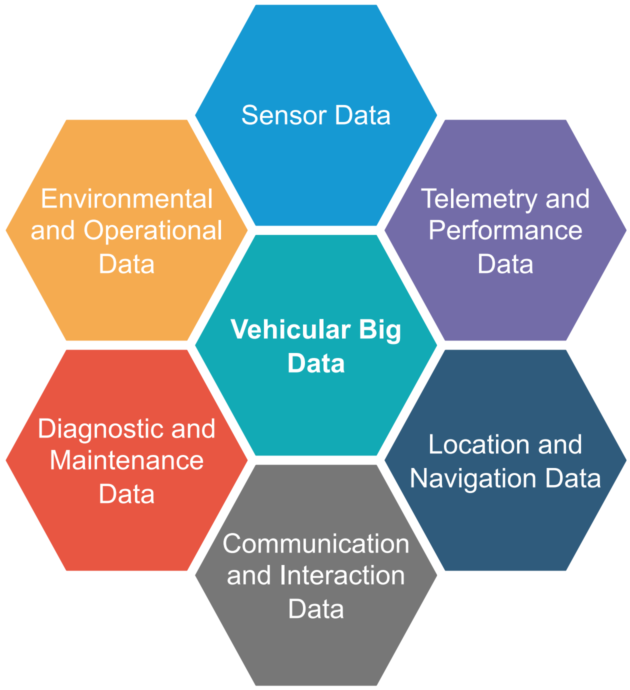
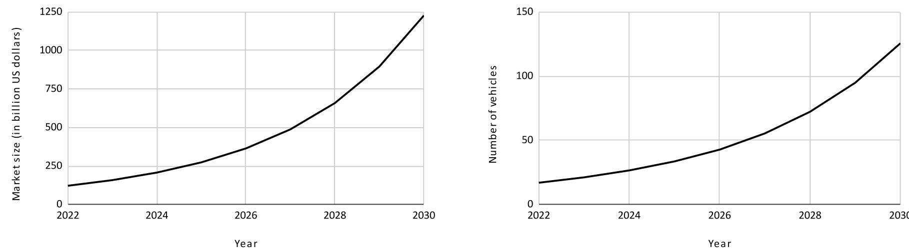


Figure 1: Vehicular Big Data

maps of the area, road conditions, infrastructure details, and data from other vehicles regarding road and location-specific information [11]. Additionally, navigation data and local government road policies and regulations are necessary for optimal functioning. However, the communication of this data with distant cloud servers can introduce potential delays in services or hinder the ability of vehicles to upload their sensor-generated big data to remote servers [12].

AVs and Advanced Driving Assistance System (ADAS)s rely on real-time communication with the cloud for navigation. According to recent reports, these systems utilize various sensors, such as high-resolution cameras, radar, lidar, ultrasonic sensors, and GPS, resulting in significant data generation, up to 4 GB per second [13]. This vast amount of data poses a barrier to the widespread adoption of autonomous cars. The high-resolution map, metadata,



(a) Autonomous vehicles market size (b) Autonomous vehicle forecast

Figure 2: Autonomous vehicle statistics

and other vehicle reports must be downloaded to the car's system. At the same time, the vehicle-generated data must be uploaded to cloud servers for navigation and training of other vehicles [14]. Table 1 show the acronyms used in the article.

Table 1: Acronym Employed in the Article

Symbol	**Representation**	**Symbol**	**Representation**
CC	Cloud Computing	EC	Edge Computing
RC	Regional Computing		
ITS	Intelligent Transport System	AV	Autonomous Vehicle
VBD	Vehicular Big Data		
ADAS	Advanced Driving Assistant System	GPS	Global Positioning System
BS	Base Station	HD	High Definition
RSU	Road Side Units	IoV	Internet of Vehicles
VANETs	Vehicular Ad Hoc Networks	DC	Data Center

Therefore, the major challenges for the future of autonomous vehicles are:

- The current network infrastructure poses constraints on the efficient migration, storage, and processing of the vast volumes of data generated by autonomous vehicles.

- The anticipated proliferation of autonomous vehicles and their accompanying data streams raise concerns regarding network congestion, particularly during peak hours. The transmission of identical datasets to the cloud via public networks may exacerbate this issue.

- Autonomous vehicles necessitate minimal latency to mitigate the risk of accidents. However, the integration of cloud-based services and Vehicle-to-Cloud communication introduces additional latency, potentially impeding the real-time responsiveness crucial for safe autonomous driving.

AVs and ADASs play a central role in shaping the future generation, so they receive extensive investigation in academia and the market. Despite significant progress, further exploration is necessary, particularly in handling Vehicular Big Data (VBD). Western markets are expected to witness the sale of approximately 31 million vehicles equipped with forward-facing cameras by 2025, resulting in a daily data generation of up to 10 Exabytes that will require processing by off-board resources [15].

Therefore, the main objectives of this article remain as follows:

- To proficiently handle the voluminous vehicular big data generated by vehicles on regional servers.

- To transfer the vehicular big data to the cloud during off-peak hours, facilitating navigation and training for other vehicles without overburdening the public network.

- To maintain minimal response times for external queries about roads and relevant information, notwithstanding the significant data output by vehicles.

AVs and ADASs pose an incredible level of complexity, seamlessly connecting with cloud platforms to create high-definition (HD) maps and store data. They

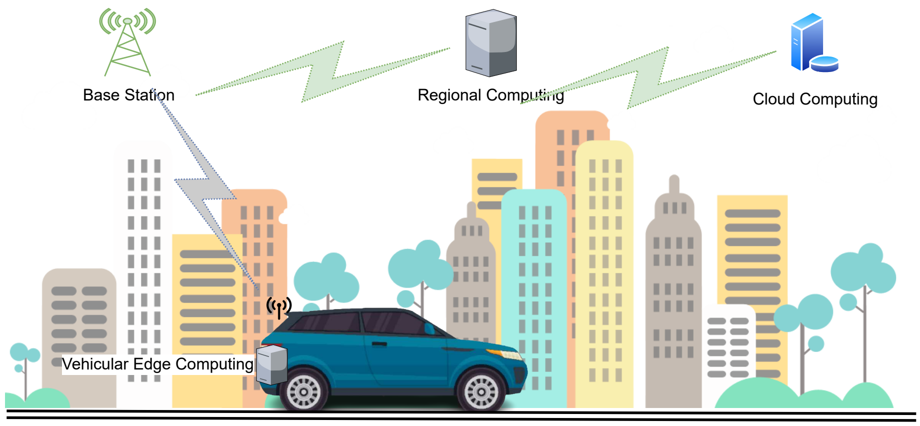


Figure 3: Regional computing for Vehicular Big Data (VBD)

integrate various technologies, including sensing, localization, perception, and decision-making [16].

This article presents the following proposals:

- Implementing regional computing, as shown in Figure 3, to minimize delays between regional servers and vehicles. These servers, strategically placed within a specific region, cover the entire area [17]. Regionally located computing servers will process data within their region, eliminating the need to transmit high-resolution graphics data to the cloud.

- Storing raw data generated by vehicle sensors initially on regional servers, which can then be transferred to the vehicular cloud during off-peak hours for AVs and ADASs navigation and training.

- Enforcing traffic regulations and rules within the Vehicles Regional Computing framework. This server would house collective data and analysis, ensuring adherence to traffic norms and facilitating efficient operations.

The remaining sections cover the following:

Section 2 provides a comprehensive review of the relevant literature and projects.

Section 3 introduces an innovative solution: regional computing, specifically tailored to tackle the aforementioned challenges. Section 4 evaluates the effectiveness of the proposed solution. Section 5 discusses the implementation challenges associated with the model. Finally, section 6 concludes the work, summarizing the key findings and implications.

2 Background

Previous studies have predominantly focused on big data analytics, with fewer discussions on big data transmission and processing. Particularly, processing and storing of this data for vehicle navigation purposes have received limited attention. The literature shows that the vehicle produces a massive amount of data, about 4 TB of data per day [15]. This data is processed in real-time in the vehicle computers and communicated with Roaside Units (RSU)s and the cloud.

Exploring the intricate relationship between the Internet of Vehicles (IoV) and Big Data within vehicular environments, recent research investigates how IoV facilitates the management of vast data volumes generated by connected vehicles. Recent research, as highlighted in a study by [18], delves into the intricate relationship between the Internet of Vehicles (IoV) and Big Data within vehicular environments. This investigation not only explores how IoV enables the management of vast data volumes generated by interconnected vehicles but also delves into the benefits of leveraging Big Data for IoV characterization, performance assessment, and communication protocol design. Furthermore, efforts by [19] to classify vehicular big data and prioritize its handling, as evidenced in [20], underscore the critical importance of effectively managing and prioritizing such data to optimize IoV functionality and performance.

Autonomous vehicle data communication predominantly relies on Vehicular

Ad Hoc Networks (VANETs). However, VANETs encounter challenges in managing large-scale data. To address this, machine learning techniques are utilized to analyze measurement data from VANETs, aiming to identify unfavorable communication scenarios [21]. Alternatively, researchers in [22] propose a novel data movement approach that leverages vehicles themselves for data transfer, rather than solely relying on the underlying infrastructure. Their findings indicate that this method significantly reduces carbon emissions.

Likewise, in [23], authors propose a big data analytical architecture tailored for vehicular data. Their approach advocates for a distributed architecture accommodating data processing, storage, and analysis to effectively handle the immense data volumes. Similarly, another study by authors in [24] challenges the conventional vehicular communication paradigm involving Roadside Units (RSUs) and cloud infrastructure. Instead, they advocate for the deployment of edge computing along roads for communication purposes. This strategy harnesses 5G internet connectivity to expedite communication and circumvent delays associated with cloud-based solutions. Similarly, authors in [25] exploit parked vehicle resources for data processing and transmission, contributing to the optimization of vehicular data management.

Addressing the challenges of vehicular data management, another investigation highlighted in [26] reveals that edge servers may not be the most suitable solution for vehicles, particularly considering their high-speed travel over long distances. Consequently, there is a growing consensus that the data required for autonomous vehicles should be distributed regionally. Similarly, [27] proposes an innovative approach that leverages parked vehicles as roadside units, facilitating collaborative data processing, transformation, and storage between vehicles and road infrastructure.

In the study conducted by [28], the emphasis was placed on the big data generated by vehicles and their communication with the cloud. Employing THz band technology, the researchers tackled this substantial data volume. Their findings indicated that this approach enhances adaptability in the communication environment, offering supplementary traffic information to nearby autonomous vehicles. Additionally, using millimetre waves, as investigated by [29], facilitates the transportation of larger data quantities compared to traditional waves.

The evolution of big data within the expansive Internet of Vehicles (IoV) framework has ushered in unprecedented opportunities for unified transport management and the development of intelligent transport systems. Addressing this paradigm shift, the authors of [30] underscore the associated challenges and propose security requirements alongside a basic system model for the secure collection of big data within the Internet of Vehicles. Moreover, the transformation of this data's security presents an ongoing challenge. In response, authors in [31] put forward a secure system wherein only vehicles registered to the cloud can transmit or receive data, safeguarding vehicular data from unauthorized access.

Similarly, authors in [32] introduce a software architecture grounded in the observer mode, with the goal of establishing regional cloud computing data centers. This architecture prioritizes registering computing resources to a central registry and employing the registry as the system coordinator. By adopting this approach, all vehicles gain consistent access to optimal computing resources from the registry, facilitating efficient data transfer and processing while upholding real-time performance and stability.

3 Proposed Model

The Intelligent Transportation System (ITS) becomes smarter if we add every vehicle learning to the joint system. Uploading every vehicle's data to the cloud

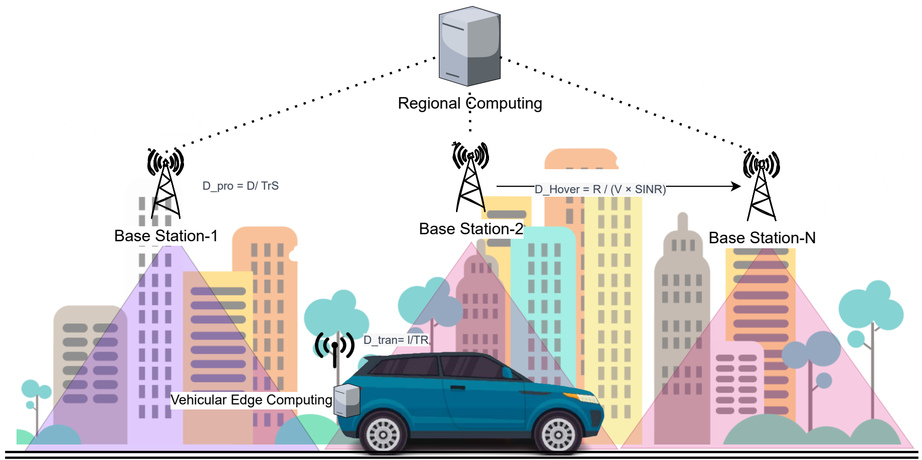


Figure 4: Delay faced by the data communication, Where D_tran is Transmission Delay, L is the data and TR is the Transmission Rate; D_pr is propagation delay, D is the distance and TrS is the transmission speed. D_Hover is hand over delay of base station, R is acceptable signal quality level for the handover decision, V denotes the vehicle velocity. $SINR$ refers to the Signal-to-Interference-plus-Noise Ratio

servers is necessary. However, every vehicle may generate one TB of data per day, which is impossible for the existing system to transport, process, and store [33]. As shown in Figure 4, the proposed system aims to process/store the VBD near the road infrastructure to minimize the delay and cost and laterally transfer this data to the cloud in off-peak hours. This will cause a minimum load on the network, as we usually see that the network gets congested in peak hours and underutilized in off-peak hours.

The proposed method works on three layers: The vehicular Edge Computing Layer, the Regional Computing Layer, and the Cloud Computing Layer.

The edge computing layer is the vehicle's server processing the sensor data in real-time. Regional computing layer process and store the VBD before transferring it to the cloud. The cloud computing layer store this big data and coordinate among all vehicles.

3.1 Vehicular Edge Layer

The vehicle Layer includes the vehicle's sensors, processing, actuators, and storage system. It also includes the local edge servers and vehicle communication to surrounding vehicles, RSUs, pedestrians, etc. During driving and obstacle detection and avoidance, this internal system is used. The sensors, e.g., Lidar, Radar, Sonar, cameras, and navigation system, give input to the system. Along with the internal system, the surrounding vehicles and RSUs also help the vehicle's navigation. The internal computers process the data with powerful algorithms using the already-built data and Artificial Intelligence and direct the actuators that keep the vehicle on the road.

AVs need 5G internet connections to upload their internal data to the cloud or regional servers or get the navigation, road status, and other particular spots and congestion data from the cloud servers. The vehicle is connected to the base station, communicating its data to the regional layer.

$$D_t = D_{tran} + D_{prop} + D_{proc} + D_{que} + D_{hovr} \tag{1}$$

Where D_{tran} represents the transmission delay, D_{prop} stands for the propagation delay, D_{proc} signifies the data processing delay, D_{que} denotes the queuing delay, and D_{hovr} corresponds to the handover delay.

The transmission delay (D_{tran}) accounts for the time taken to transmit data to the transmission medium. It is influenced by factors such as workload (W), channel bandwidth capacity (B), signal-to-noise ratio (SNR), modulation efficiency (ME), and error rate (ER). The calculation is given by:

$$D_{\text{tran}} = \frac{W \times \text{SNR} \times \text{ME}}{B \times \text{ER}} \tag{2}$$

The propagation delay (D_{prop}) represents the time for data to travel on the transmission medium from source to destination. It depends on distance (Dis) and transmission speed (tr_s), and is calculated as:

$$D_{\text{prop}} = \frac{\text{Dis}}{\text{tr}_s} \tag{3}$$

Data processing delay (D_{proc}) reflects the time taken for the system to process the data. It is determined by the size of the data (S) and the processing rate of the machine (P_r). The formula is:

$$D_{\text{proc}} = \frac{S}{P_r} \tag{4}$$

Queuing delay (D_{que}) indicates the time data spends waiting in a queue for processing. This delay depends on packet length (L), arrival rate of packets (a), and packet processing rate (R). The equation is given by:

$$D_{\text{que}} = \frac{L \times a}{R} \tag{5}$$

Handover delay (D_{hovr}) represents the delay in connecting to the next base station due to signal strength. It is influenced by acceptable signal quality level (R), vehicle velocity (V), and Signal-to-Interference-plus-Noise Ratio ($SINR$). The calculation is expressed as:

$$D_{\text{hovr}} = \frac{R}{V \times \text{SINR}} \tag{6}$$

These individual components collectively contribute to the overall delay (D_t) experienced by the data during its journey to the edge server.

Similarly, the data communication cost also increases as the distance the size of the data increases; we know that.

$$C_t = C_{\text{tran}} + C_{\text{prop}} \tag{7}$$

The equation represents the total cost (C_t) experienced by data to reach the edge server, which is the sum of transmission cost (C_{tran}) and propagation cost (C_{prop}).

3.2 Regional Layer

Regional servers within a specific region process and store vehicle data. They act as local hubs for collecting information from vehicles in their vicinity. Each vehicle within the region sends its data to the regional servers. This data may include parameters such as GPS location, speed, acceleration, sensor readings, video, and other relevant information about the vehicle's status and surroundings.

The regional servers process the received vehicle data to extract useful information. This can involve analyzing traffic patterns, road conditions, and congestion levels and identifying potential road issues or hazards. The regional service can provide guidance and information to other regional vehicles based on the processed data. This can include real-time updates about road conditions, traffic congestion, accidents, or any relevant information to help drivers make informed decisions.

The VBD is temporarily stored on this server in peak hours and sent to the cloud in off-peak hours to minimize the congestion on a public network. Furthermore, as we can see from equation 1, the delay (D_t) depends on the distance (D) and workload (L). So, if this big data is processed and stored locally in peak hours, this minimizes the total delay, and the vehicle will get a real-time response. Similarly, the public network is not overburdened.

The regional vehicular big data management algorithm, referenced as Algo-

Algorithm 1 Regional Vehicular Big Data Management Algorithm

```
Input: Vehicular big data from the region's vehicles
Output: Decision on data processing and migration
if Network condition is normal AND Network is highly utilized then
    Process and store the data at regional servers
end
else
    if Data is received during off-peak hours then
        if Local processing at regional servers is sufficient then
            Process and store the data locally at regional servers
        end
        else
            Transfer the required data to the cloud for processing and analysis
            during off-peak hours
        end
    end
    else
        Transfer the required data to the cloud for processing and analysis
    end
end
if Data is needed for training in another region then
    Transfer the data to the cloud for further migration
end
```

rithm 1, is designed for autonomous vehicles and accepts inputs such as video, RADAR, LIDAR, SONAR, GPS, and other sensor data. This algorithm makes decisions on whether to store and process the data at the regional servers or transfer it to the cloud for further processing and analysis.

If the network is operating under normal conditions and is highly utilized, the algorithm proceeds to store and process the data locally at the regional servers. However, in scenarios where the network conditions deviate from normal or experience high utilization, the algorithm assesses the timing of data reception. During off-peak hours, the data is processed and stored at the regional servers, minimizing delay and optimizing real-time response for vehicles. If local processing is insufficient during off-peak hours, the required data is then transferred to the cloud for comprehensive processing and analysis.

Additionally, the algorithm accounts for situations where the data is needed for training in another region. In such cases, the algorithm ensures the seamless migration of the required data to the cloud, facilitating training processes in a different geographical context.

Algorithm 2 Vehicular Cloud Processing

Input: Vehicular big data
Output: Real-time insights
if *Urgency is high* **then**
    Prioritize immediate processing and delivery
end
else
    Preprocess data for quality **Apply** analytics for insights **Perform** traffic analysis, road monitoring, and anomaly detection **Optimize** data storage and retrieval **Collaborate** with regional servers and edge computing **Provide** real-time updates to vehicles and traffic systems **Facilitate** data sharing for broader applications **Monitor** and **Manage** system performance **Implement** mechanisms for data retention and compliance
end

3.3 Cloud Layer

The regional layer plays an active role in autonomous driving and real-time operations, while the cloud layer operates passively. The regional layer depends on the cloud for computing resources and transfers its data to the cloud during off-peak hours to leverage the cloud's capabilities for processing and analysis.

Compared to the regional layer, the cloud possesses a vast amount of VBD from AV worldwide. It serves as a repository for processing and storing this extensive dataset.

$$T_{\text{cloud}} = \frac{W}{R} \tag{8}$$

Equation 8 represents the processing time (T_{cloud}) for collectively processing all vehicular workload at cloud servers. The variables are defined as follows: T_{cloud} is the processing time, W is the total amount of vehicular workload to be processed, and R is the processing rate of the cloud servers, indicating the amount of data processed per unit of time.

The total vehicular workload (W) is calculated as the summation of the amount of data produced by each vehicle (W_i) from 1 to n, as shown in Equation 9:

$$W = \sum_{i=1}^{n} W_i \tag{9}$$

Examining equation number 1 reveals that the roundtrip to the cloud will incur more delay than anticipated. The increased delay is attributed to the cloud having to process a substantial volume of vehicular big data (W). Additionally, the propagation delay and queuing delay experience escalation in tandem with the growing workload and distance.

From the above working, we can see that cloud analyzes the data to gain insights into traffic patterns, congestion, and road conditions. This information can be used to optimize traffic flow and provide real-time updates to autonomous vehicles regarding alternative routes or potential hazards. The cloud employs machine learning and artificial intelligence algorithms to extract valuable information from the data. This enables development and improvement of autonomous driving algorithms, predictive maintenance models, and other intelligent systems.

The cloud leverages the vehicular data to enhance safety and security measures. It can identify and mitigate potential risks, detect anomalies or malicious activities, and provide early warnings to vehicles and authorities. By analyzing the data, the cloud can identify areas for improvement in ADAS, such as fuel efficiency, route planning, and vehicle performance. This optimization can lead to cost savings and enhanced overall performance. The cloud's vast dataset is a valuable resource for researchers, engineers, and developers to study and innovate in autonomous driving. It enables the exploration of new algorithms, technologies, and applications to advance the capabilities of autonomous vehicles. This data may also be used to train the autonomous driving styles.

Energy Calculation

Similar to delay, energy consumption is directly proportional to distance, impacting operational costs for data transfer. The energy consumption in the autonomous vehicles' communication environment is computed as follows:

$$E_{\text{tran}} = \sum_{i=1}^{n} E_{\text{tran}}(i, i+1) \tag{10}$$

Here, $E_{\text{tran}}(i, i+1)$ represents the power consumption between consecutive stages, with power consumption increasing as the stages progress.

Where

$$E_{\text{tran}}(i, i+1) = \frac{D_{i,i+1} \cdot P_i}{T_{i,i+1}} \tag{11}$$

Where $D_{i,i+1}$ is the distance, P_i is the power and $T_{i,i+1}$ is the time. The above equations show that as the distance increases the energy consumption increases, which is directly proportional to cost.

$$E_{\text{other}} = E_{\text{pro}} + E_{\text{stor}} + E_{\text{col}} + k \tag{12}$$

Similarly, E_{other} represents energy consumption in other activities, including processing (E_{pro}), storing (E_{stor}), and cooling the data centres (E_{col}).

Therefore, the total energy consumption of communication is calculated as follows:

$$E_{\text{total}} = E_{\text{tran}} + E_{\text{other}} \tag{13}$$

Here, E_{total} indicates the total power usage, while E_{tran} denotes the total power consumption on data transfer, encompassing wires, switches, routers, and other devices.

It is also acknowledged that,

$$Cost_{\text{oper}} \propto E \tag{14}$$

The power consumption (E) is directly proportional to the operational cost ($Cost_{\text{oper}}$); hence, operational costs increase with rising power consumption.

4 Evaluation

Cloud-Analytic is a widely utilized simulation tool for simulating cloud networks, enabling calculations related to delay and cost assessments. In our preliminary research, we employed this tool to evaluate the impact of VBD on network performance while considering both cloud and regional computing.

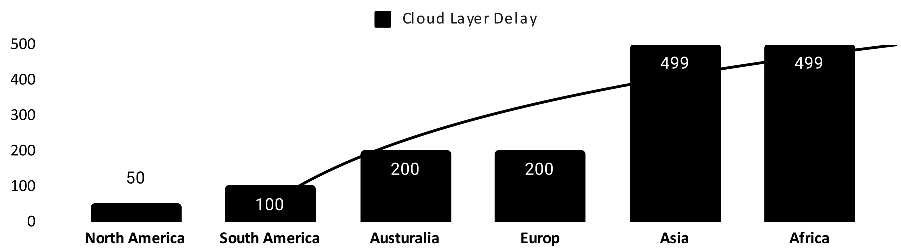


Figure 5: Cloud Layers Delay for Vehicular Big Data

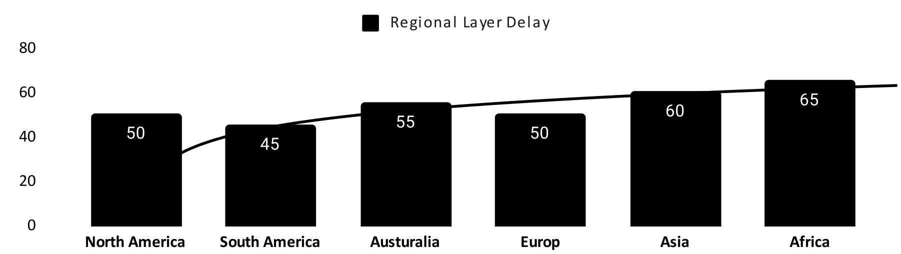


Figure 6: Regional Layers Delay for Vehicular Big Data

4.1 Experimental Setup

The proposed structure of this work consists of two scenarios:

1. In the first, we calculate the delay and cost associated with transferring, storing, and processing VBD on cloud computing servers.

2. In the second scenario, we examine the delay and cost of transferring, storing, and processing VBD on regional computing servers.

In the first scenario, a Data Center (DC) was established in North America (Region 1), while AVs were located in various regions worldwide. A workload of 100 MB was generated to simulate the scenario, and the upload process from the AVs to the cloud server was initiated.

Each Data Center (DC) was equipped with 204,800 MB of RAM, 100,000,000 MB of storage, 1,000,000 MB of bandwidth, 4 CPUs, and a processing speed of 10,000 Million Instructions Per Second (MIPS). Within the DC were five virtual machines (VMs) utilizing the available resources. For the simulation, the peak hours were defined as 01:00 AM to 09:00 PM, while the off-peak hours were set from 01:00 AM to 09:00 AM. During peak hours, 1,000 AVs sent their requests to the server concurrently, whereas during off-peak hours, 100 AVs requested resources simultaneously.

We conducted ten iterations of this experiment, each lasting 60 minutes, and calculated the average communication delay and cost. The results indicate that the delay also increases as the distance between autonomous vehicles (AVs) and the cloud servers increases, as illustrated in Figure 5. Specifically, the delay varies across different geographical regions, with the highest delay observed in Africa (499 ms) and Asia (499 ms), followed by Europe (200 ms) and Australia (200 ms). North America experiences the lowest delay at 50 ms, while South America encounters a delay of 100 ms. Similarly, the cost escalates with greater distances, as depicted in Figure 7. The cost values for each region are as follows: North America ($0.065), South America ($0.192), Australia ($0.228), Europe ($0.192), Asia ($0.196), and Africa ($0.196). These findings underscore the impact of geographical distance on communication delay and cost in cloud-based autonomous vehicle systems.

In the second plot, regional cloud servers were established in each corresponding region, and the regional workload was executed on these servers. Each

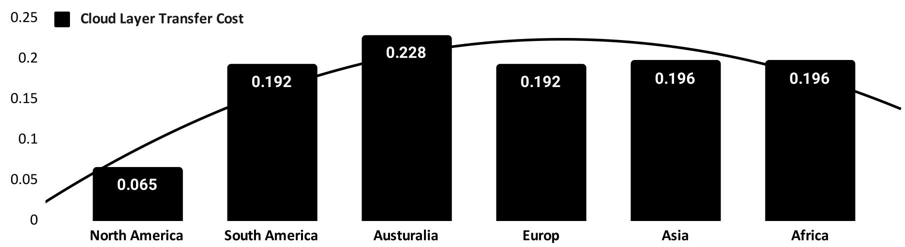


Figure 7: Cloud Layers Cost for Vehicular Big Data

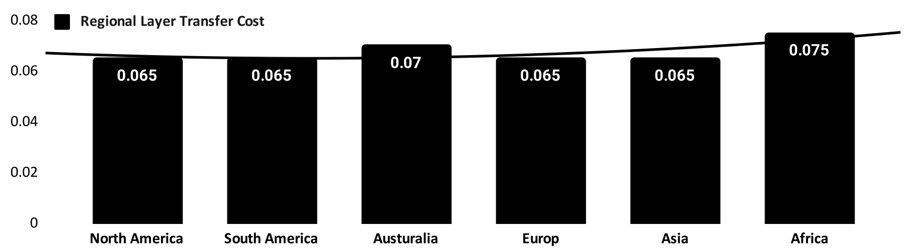


Figure 8: Cloud Layers Cost for Vehicular Big Data

Data Center (DC) was equipped with 204,800 MB of RAM, 100,000,000 MB of storage, 1,000,000 MB of bandwidth, 4 CPUs, and a processing speed of 10,000 Million Instructions Per Second (MIPS). Five virtual machines (VMs) utilized the available resources within each DC. As in the first scenario, the peak hours for this simulation were extended from 01:00 AM to 09:00 PM, while the off-peak hours were defined as 01:00 AM to 09:00 AM. During the peak hours, 1,000 AVs simultaneously sent their requests to the server, while during the off-peak hours, 100 AVs concurrently requested resources.

To ascertain the average communication delay and cost, we employed a 100 MB workload and executed ten iterations of each regional cloud experiment, each lasting 60 minutes. The outcomes reveal that having servers located within the same region significantly reduces both the delay and costs compared to a centralized cloud server. This reduction in delay and costs can be attributed to decreased data transmission distances. Specifically, when the data is communicated from North America, the delay varies across different regions, with South America experiencing the lowest delay at 45 ms, followed closely by North America itself at 50 ms. Australia encounters a delay of 55 ms, while Europe experiences a delay of 50 ms. Asia and Africa both encounter higher delays, at 60 ms and 65 ms respectively. Similarly, the cost varies across regions, with South America, North America, and Europe having the lowest cost at $0.065, and Australia having a slightly higher cost of $0.07. Asia and Africa experience the highest costs, at $0.065 and $0.075 respectively. These findings are illustrated in Figures 8 and 6, demonstrating the results of the second scenario.

5 Discussion

The findings from our study highlight the significant impact of geographical distance on cloud computing performance for autonomous vehicles (AVs). As demonstrated by the results, cloud computing from regions such as North America, South America, Europe, Asia, Australia, and Africa incurs varying degrees of delay and cost. For instance, our analysis reveals that regions closer to the AVs, such as North America and South America, exhibit lower delays and costs compared to regions farther away, such as Asia and Africa. This aligns with the existing literature, which suggests that proximity to regional servers reduces delay and costs significantly [17].

Overall, our results underscore the importance of considering regional computing as a viable alternative to centralized cloud computing for AVs. By leveraging regional computing resources, AVs can experience improved performance in terms of delay and cost, thereby enhancing their reliability and attractiveness as a computing option.

However, specific challenges still need to be addressed within the proposed framework instead of positively impacting delay and cost. The leading chal-

Table 2: Comparison of Vehicular Edge Computing, Regional Computing, and Cloud Computing

Parameters	**Vehicle Edge Computing**	**Regional Computing**	**Cloud Computing**
Vehicles	Limited to a single vehicle	Regional coverage	Vehicles around the world
Area	Limited to the immediate surroundings of the vehicle	Covers a specific region or locality	Global coverage
Delay	Very low latency	Medium-level latency	High-level latency
Data	Data generated by individual vehicle	Aggregated data from regional vehicles	Data from vehicles worldwide
Computation Power	Limited computational capabilities	Moderate computational capabilities	High computational capabilities
Storage	Limited storage capacity	Moderate storage capacity	High storage capacity
Server Mobility	Movable (Deployed on the vehicle)	Immovable (fixed location)	Immovable (fixed location)
Resource Scalability	Low scalability	Medium scalability	High scalability

lenges are;

- Firstly, the ownership cost presents a challenge as the Vehicular Industry must invest significant capital in deploying and operating regional computing servers [34].

- Secondly, regional servers' data processing at the terminal level raises security and privacy concerns that must be addressed [35].

- A third challenge is the variation of cyber rules across different regions, requiring the implementation of region-specific management strategies for these issues [36].

Based on the experimentation above and the literature, we comprehensively compared regional computing and cloud computing for VBD across various parameters such as delay, area, vehicles, data, servers' mobility, computational power, and capacity. The detailed comparison is presented in Table 2.

6 Conclusion

In this article, we introduce the concept of Regional Computing (RC) as a viable solution for managing Vehicular Big Data (VBD) by shifting the focus from cloud computing servers to regional infrastructure, particularly during peak hours. This novel approach not only relieves pressure on the public network but also enhances real-time response capabilities and optimizes the utilization of VBD for vehicle training and communication purposes. Our findings underscore the substantial reduction in workload on the primary network and the consequent improvement in vehicular communication performance, highlighting the potential of RC in addressing the challenges posed by VBD management.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data may be available on a reasonable request to corresponding author.

Authors Contribution

This work was carried out in collaboration between all authors. Detailed methodology: A.B and G.U.R., Formal analysis and Investigations: S.U.K and M.N, Writing, review & editing: T.S. Writing original draft: G.U.R. All authors have read and agreed to the published version of the manuscript.

Acknowledgement

This study is supported via funding from Prince sattam bin Abdulaziz University project number (PSAU/2024/R/1445).

References

[1] Dohoon Lee, David Camacho, and Jason J Jung. Smart mobility with big data: Approaches, applications, and challenges. *Applied Sciences*, 13(12):7244, 2023.

[2] Sirwan K Ahmed, Mona G Mohammed, Salar O Abdulqadir, Rabab G Abd El-Kader, Nahed A El-Shall, Deepak Chandran, Mohammad E Ur Rehman, and Kuldeep Dhama. Road traffic accidental injuries and deaths: A neglected global health issue. *Health science reports*, 6(5):e1240, 2023.

[3] Yuerong Zhang and Maria Kamargianni. A review on the factors influencing the adoption of new mobility technologies and services: autonomous vehicle, drone, micromobility and mobility as a service. *Transport reviews*, 43(3):407–429, 2023.

[4] Tesla. Tesla. `https:www.tesla.com/`, 2024.

[5] Waymo. Waymo. `https://waymo.com/`, 2024.

[6] Motional. Motional. `https://motional.com/`, 2024.

[7] Ford. Ford. `https://ford.com/`, 2024.

[8] Statista. Penetration rate of light autonomous vehicles (l4) worldwide in 2021, with a forecast through 2030. https://www.statista.com/statistics/875080/av-market-penetration-worldwide-forecast/, 2024.

[9] Budi Padmaja, CH VKNSN Moorthy, N Venkateswarulu, and Myneni Madhu Bala. Exploration of issues, challenges and latest developments in autonomous cars. *Journal of Big Data*, 10(1):61, 2023.

[10] Mordorintelligence. Autonomous vehicle data storage. https://premioinc.com/pages/autonomous-vehicle-data-storage, 2024.

[11] Afzal Badshah, Anwar Ghani, Azeem Irshad, Husnain Naqvi, and Saru Kumari. Smart workload migration on external cloud service providers to minimize delay, running time, and transfer cost. *International Journal of Communication Systems*, 34(3):e4686, 2021.

[12] Afzal Badshah, Anwar Ghani, Ali Daud, Anthony Theodore Chronopoulos, and Ateeqa Jalal. Revenue maximization approaches in iaas clouds: Research challenges and opportunities. *Transactions on Emerging Telecommunications Technologies*, 33(7):e4492, 2022.

[13] Jaycil Z Varghese, Randy G Boone, et al. Overview of autonomous vehicle sensors and systems. In *International Conference on Operations Excellence and Service Engineering*, volume 2015. sn, 2015.

[14] SV Aswin Kumer, LSP Sairam Nadipalli, P Kanakaraja, K Sarat Kumar, and K Ch Sri Kavya. Controlling the autonomous vehicle using computer vision and cloud server. *Materials Today: Proceedings*, 37:2982–2985, 2021.

[15] Dohoon Lee, David Camacho, and Jason J Jung. Smart mobility with

big data: Approaches, applications, and challenges. *Applied Sciences*, 13(12):7244, 2023.

[16] Shaoshan Liu, Liyun Li, Jie Tang, Shuang Wu, and Jean-Luc Gaudiot. *Creating autonomous vehicle systems*. Springer, 2018.

[17] Afzal Badshah, Celestine Iwendi, Ateeqa Jalal, Syed Shabih Ul Hasan, Ghawar Said, Shahab S Band, and Arthur Chang. Use of regional computing to minimize the social big data effects. *Computers & Industrial Engineering*, 171:108433, 2022.

[18] Wenchao Xu, Haibo Zhou, Nan Cheng, Feng Lyu, Weisen Shi, Jiayin Chen, and Xuemin Shen. Internet of vehicles in big data era. *IEEE/CAA Journal of Automatica Sinica*, 5(1):19–35, 2018.

[19] Big autonomous vehicular data classifications: Towards procuring intelligence in its. *Vehicular Communications*, 9:306–312, 2017.

[20] Günter Knieps. Internet of things, big data and the economics of networked vehicles. *Telecommunications Policy*, 43(2):171–181, 2019. The data economy and data-driven ecosystems: regulation, frameworks, and case studies.

[21] Nan Cheng, Feng Lyu, Jiayin Chen, Wenchao Xu, Haibo Zhou, Shan Zhang, and Xuemin Shen. Big data driven vehicular networks. *IEEE Network*, 32(6):160–167, 2018.

[22] Murk, Asad Waqar Malik, Imran Mahmood, Nadeem Ahmed, and Zahid Anwar. Big data in motion: A vehicle-assisted urban computing framework for smart cities. *IEEE Access*, 7:55951–55965, 2019.

[23] Theodoros Alexakis, Nikolaos Peppes, Konstantinos Demestichas, and Evgenia Adamopoulou. A distributed big data analytics architecture for vehicle sensor data. *Sensors*, 23(1):357, 2022.

[24] Sahil Garg, Amritpal Singh, Kuljeet Kaur, Gagangeet Singh Aujla, Shalini Batra, Neeraj Kumar, and M. S. Obaidat. Edge computing-based security framework for big data analytics in vanets. *IEEE Network*, 33(2):72–81, 2019.

[25] Nitin Singh Rajput, Upendra Singh, Amit Dua, Neeraj Kumar, Joel J. P. C. Rodrigues, Sheetal Sisodia, Mohammed Elhoseny, and Yahya Lakys. Amalgamating vehicular networks with vehicular clouds, ai, and big data for next-generation its services. *IEEE Transactions on Intelligent Transportation Systems*, 25(1):869–883, 2024.

[26] Christian Prehofer and Shafqat Mehmood. Big data architectures for vehicle data analysis. In *2020 IEEE International Conference on Big Data (Big Data)*, pages 3404–3412. IEEE, 2020.

[27] Nitin Singh Rajput, Upendra Singh, Amit Dua, Neeraj Kumar, Joel J. P. C. Rodrigues, Sheetal Sisodia, Mohammed Elhoseny, and Yahya Lakys. Amalgamating vehicular networks with vehicular clouds, ai, and big data for next-generation its services. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–15, 2023.

[28] Chaofeng Zhang, Kaoru Ota, Juncheng Jia, and Mianxiong Dong. Breaking the blockage for big data transmission: Gigabit road communication in autonomous vehicles. *IEEE Communications Magazine*, 56(6):152–157, 2018.

[29] Sen Du, Junjie Hou, Shijin Song, Yuefeng Song, and Yongxin Zhu. A geographical hierarchy greedy routing strategy for vehicular big data communications over millimeter wave. *Physical Communication*, 40:101065, 2020.

[30] Adnan Mahmood, Hushairi Zen, and Shadi Hilles. Big data and privacy issues for connected vehicles in intelligent transportation systems. *arXiv preprint arXiv:1806.02944*, 2018.

[31] Longhua Guo, Mianxiong Dong, Kaoru Ota, Qiang Li, Tianpeng Ye, Jun Wu, and Jianhua Li. A secure mechanism for big data collection in large scale internet of vehicle. *IEEE Internet of Things Journal*, 4(2):601–610, 2017.

[32] Ji Nan, Ding Qiang, Wang Liangliang, Guo Shaojie, and Hu Shilan. Cloud computing design of intelligent connected vehicles based on regional load. In *Sixth International Conference on Intelligent Computing, Communication, and Devices (ICCD 2023)*, volume 12703, pages 517–523. SPIE, 2023.

[33] Afzal Badshah, Anwar Ghani, Isma Farah Siddiqui, Ali Daud, Muhammad Zubair, and Zafar Mehmood. Orchestrating model to improve utilization of iaas environment for sustainable revenue. *Sustainable Energy Technologies and Assessments*, 57:103228, 2023.

[34] Benedikt Martens, Marc Walterbusch, and Frank Teuteberg. Costing of cloud computing services: A total cost of ownership approach. In *2012 45th Hawaii International Conference on System Sciences*, pages 1563–1572. IEEE, 2012.

[35] Aleksandr Ometov, Oliver Liombe Molua, Mikhail Komarov, and Jari Nurmi. A survey of security in cloud, edge, and fog computing. *Sensors*, 22(3):927, 2022.

[36] Dayanand Lal N, Riyaz Mahomad, Shaik Vaseem Akram, Anuradha Averineni, G. Ravi Babu, and Ch. Naga Sai Satyadeva. Innovative cyber security techniques based on blockchain technology for use in industrial 5.0 applica-

tions. In *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 1316–1320, 2023.