COLOMBO SCIENTIFIC PUBLISHING

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM

E-ISSN: 2536-7919

P-ISSN: 2536-7900

PAGE NO: 26-31

Fleet-as-a-Service Architectures for Sustainable First- and Last-Mile Mobility: Integrating Shared Vehicle Operations, Real-Time Computing Paradigms, and Ethical Transportation Systems

Dr. Alexander M. Vollenweider

Department of Mechanical and Industrial Engineering, University of Melbourne, Australia

ARTICLE INFO

Article history:

Submission: January 01, 2025
Accepted: January 14, 2025
Published: January 23, 2025
VOLUME: Vol.10 Issue 01 2025

Keywords:

Fleet-as-a-Service; Sustainable Mobility; First- and Last-Mile Transportation; Shared Autonomous Vehicles; Real-Time Systems; Intelligent Transportation Systems

ABSTRACT

The global transportation ecosystem is undergoing a profound structural transformation driven by the convergence of sustainability imperatives, rapid urbanization, digital infrastructure maturation, and evolving mobility consumption patterns. Within this context, Fleet-as-a-Service (FaaS) has emerged as a unifying operational and conceptual paradigm that reconceptualizes vehicle ownership, deployment, testing, and lifecycle management as service-oriented processes rather than asset-bound activities. This research article develops an extensive theoretical and analytical examination of FaaS architectures as they apply to sustainable vehicle testing, shared mobility operations, and first- and last-mile transportation systems. Drawing upon interdisciplinary scholarship from transportation engineering, computer systems, real-time scheduling, cloud and serverless computing, and intelligent transportation systems, the article situates FaaS as a socio-technical infrastructure rather than a narrowly defined logistics solution. Particular emphasis is placed on how FaaS reshapes the testing and validation of emerging vehicle technologies, including automated, connected, and shared vehicles, while simultaneously enabling scalable operational models aligned with sustainability objectives (Deshpande, 2024).

The study synthesizes prior work on shared mobility optimization, demand-responsive transit, feeder services, and micro-simulation modeling to demonstrate how FaaS platforms enable dynamic fleet orchestration across heterogeneous urban environments (Goswami et al., 2021; Huang et al., 2021). Beyond transportation theory, the article critically integrates insights from real-time systems scheduling, serverless computing, and edge-cloud architectures to argue that modern FaaS deployments are computationally intensive systems whose performance, reliability, and ethical accountability depend on sophisticated task partitioning, latency guarantees, and workload management strategies (Baruah, 2013; Nguyen et al., 2019). Through an expansive qualitative methodology grounded in comparative literature analysis and conceptual modeling, the article identifies persistent gaps in existing research, particularly concerning the ethical governance, computational sustainability, and cross-domain integration of FaaS systems.

The results of this analysis indicate that FaaS is not merely an operational efficiency mechanism but a foundational infrastructure capable of aligning mobility innovation with environmental stewardship, equitable access, and technological accountability. The discussion extends these findings by engaging competing scholarly perspectives, addressing limitations of current models, and outlining future research trajectories that bridge transportation planning with computing systems design. By offering a deeply elaborated, citation-rich, and theoretically grounded contribution, this article advances the academic discourse on sustainable mobility systems and positions FaaS as a critical lever for the next generation of intelligent transportation ecosystems (Deshpande, 2024; Grahn et al., 2021).

INTRODUCTION

Urban mobility systems have historically evolved through incremental adaptations to demographic growth, economic development, and technological innovation, yet the present era is distinguished by the simultaneity and intensity of pressures confronting transportation infrastructures worldwide. Rapid urbanization has increased demand for flexible, reliable, and affordable mobility, while climate change imperatives have exposed the environmental costs of privately owned, underutilized vehicle fleets, prompting policymakers and researchers alike to seek systemic alternatives (Kanuri et al., 2019). Within this landscape, Fleet-as-a-Service has emerged as a compelling framework that reframes vehicles not as individually owned assets but as dynamically managed resources optimized through digital platforms and shared usage models (Deshpande, 2024). This conceptual shift resonates with broader transformations observed in energy, computing, and manufacturing sectors, where service-oriented architectures have replaced ownership-centric paradigms.

The theoretical foundation of FaaS is rooted in the recognition that mobility demand is inherently temporal, spatial, and stochastic, rendering static fleet configurations inefficient and environmentally burdensome (Li & Quadrifoglio, 2010). Traditional vehicle ownership models prioritize availability over utilization, leading to significant idle time and excessive resource consumption, particularly in urban contexts where space and energy are constrained (Martinez et al., 2015). FaaS challenges this inefficiency by enabling centralized or federated fleet management systems that allocate vehicles based on real-time demand, predictive analytics, and coordinated transit integration (Grahn et al., 2021). This approach aligns closely with sustainability objectives by reducing redundant vehicle production, lowering emissions through optimized routing, and facilitating the integration of low-emission and electric vehicles into shared fleets (Deshpande, 2024).

A critical dimension of FaaS lies in its implications for vehicle testing and validation, particularly as the automotive industry accelerates the deployment of automated and connected vehicle technologies. Conventional testing methodologies, which rely on limited pilot deployments and controlled environments, struggle to capture the complexity and variability of real-world operating conditions (Huang et al., 2021). FaaS platforms, by contrast, offer continuous, large-scale testing environments in which vehicles operate under diverse traffic, weather, and demand scenarios, generating rich datasets for performance evaluation and system refinement (Deshpande, 2024). This integration of testing and operations blurs traditional boundaries between research, development, and deployment, raising new methodological and ethical questions regarding data governance, safety assurance, and accountability.

The literature on shared mobility and first- and last-mile connectivity provides critical insights into the operational potential of FaaS but often treats fleet management as a secondary consideration rather than a central organizing principle (Kumar & Khani, 2021). Studies on feeder services and demand-responsive transit have demonstrated the benefits of flexible routing and scheduling, yet these approaches frequently assume static fleet capacities and overlook the computational infrastructure required to support real-time decision-making at scale (Leffler et al., 2021). Similarly, research on shared autonomous vehicles has emphasized travel behavior impacts and cost efficiency while underexamining the system-level orchestration challenges inherent in managing heterogeneous vehicle fleets across multiple service contexts (Lau & Susilawati, 2021).

Parallel to developments in transportation research, advances in computing systems have transformed the feasibility of large-scale, real-time fleet orchestration. The evolution of cloud computing, serverless architectures, and edge computing has enabled elastic resource provisioning, low-latency processing, and fine-grained workload management, all of which are essential for responsive FaaS platforms (Nguyen et al., 2019; Wang et al., 2020). However, the integration of these computing paradigms into transportation systems introduces new complexities related to real-time task scheduling, energy efficiency, and quality-of-service guarantees, particularly in safety-critical applications such as autonomous vehicle coordination (Baruah et al., 2019). The transportation literature has only begun to engage with these issues, resulting in a fragmented understanding of how computational constraints shape mobility outcomes.

The problem addressed in this article arises from this fragmentation: despite growing interest in FaaS as an operational model, there remains a lack of comprehensive, interdisciplinary analysis that situates FaaS at the intersection of sustainable mobility, shared transportation theory, and real-time computing systems. Existing studies often focus on isolated components, such as demand prediction algorithms or feeder service design, without examining how these elements coalesce within a unified FaaS architecture (Goswami et al., 2021). Moreover, ethical considerations related to data use, access equity, and algorithmic decision-making are frequently treated as peripheral concerns rather than integral design parameters (Deshpande, 2024).

This article seeks to address this literature gap by developing an expansive, theoretically grounded examination of FaaS systems that integrates insights from transportation engineering, computer science, and sustainability studies. By synthesizing research on shared mobility optimization, real-time scheduling, and serverless computing, the study articulates a holistic framework for understanding FaaS as a socio-technical infrastructure rather than a discrete service offering (Fouladi et al., 2017; Ao et al., 2018). In doing so, it responds to calls for more integrative scholarship capable of informing both academic inquiry and practical implementation in complex urban environments (Kanuri et al., 2019).

The contribution of this research is threefold. First, it provides an extensive theoretical elaboration of FaaS, tracing its conceptual lineage and situating it within broader debates on sustainable transportation and service-oriented system design (Deshpande, 2024). Second, it critically examines the computational underpinnings of FaaS platforms, drawing on real-time systems literature to highlight performance, scalability, and energy efficiency challenges that shape operational feasibility (Mascitti et al., 2021). Third, it offers a forward-looking discussion of ethical, methodological, and policy implications, identifying pathways for future research that bridge disciplinary silos and support the responsible evolution of mobility systems (Grahn et al., 2021).

METHODOLOGY

The methodological approach adopted in this study is qualitative, integrative, and theory-driven, reflecting the exploratory and interdisciplinary nature of the research problem under investigation. Rather than employing empirical experimentation or quantitative modeling, the study relies on an extensive analytical synthesis of existing scholarly literature to construct a coherent conceptual framework for Fleet-as-a-Service systems (Martinez et al., 2015). This choice is grounded in the recognition that FaaS is an emergent paradigm whose complexity spans technological, organizational, and societal domains, making it ill-suited to narrow methodological lenses (Deshpande, 2024).

The primary methodological strategy involves a structured thematic analysis of peer-reviewed journal articles, conference proceedings, and authoritative technical reports drawn from the domains of transportation engineering, intelligent transportation systems, shared mobility, and computing systems (Goswami et al., 2021). The analysis prioritizes sources that address fleet management, first- and last-mile connectivity, demand-responsive transit, and shared autonomous vehicle operations, as these areas provide critical insights into the operational contexts in which FaaS architectures are deployed (Huang et al., 2021). By examining these studies collectively, the methodology seeks to identify recurring assumptions, design patterns, and performance metrics that inform current thinking on shared fleet systems.

In parallel, the study incorporates literature from real-time scheduling, serverless computing, and edge-cloud architectures to elucidate the computational foundations of FaaS platforms (Baruah, 2013; Nguyen et al., 2019). This interdisciplinary integration is methodologically significant, as it enables the analysis to move beyond surface-level descriptions of mobility services and engage with the underlying systems that enable real-time decision-making, scalability, and reliability (Elsakhawy & Bauer, 2020). The inclusion of computing systems literature also facilitates a critical examination of energy efficiency and sustainability at the computational level, an often-overlooked dimension of mobility system design (Roy et al., 2021).

A key methodological principle guiding the analysis is reflexivity, particularly with respect to the normative assumptions embedded in existing research. Many studies on shared mobility implicitly prioritize efficiency and cost reduction, potentially at the expense of equity, transparency, and user autonomy (Kanuri et al., 2019). By juxtaposing these studies with work on ethical decision-making and sustainable system design, the methodology seeks to surface tensions and trade-offs that are central to the responsible implementation of FaaS (Deshpande, 2024). This reflexive stance allows the study to critically assess not only what FaaS systems can achieve, but also how and for whom they operate.

The methodological process unfolds in several iterative stages. First, relevant literature is categorized into thematic clusters, including fleet optimization, first- and last-mile integration, shared autonomous vehicles, real-time scheduling, and serverless computing (Leffler et al., 2021). Second, each cluster is analyzed to identify core concepts, methodological approaches, and key findings, with particular attention to how these elements relate to the broader notion of service-oriented fleet management (Li & Quadrifoglio, 2010). Third, cross-cutting themes and gaps are identified through comparative analysis, enabling the construction of an integrative conceptual narrative that links transportation and computing perspectives (Baruah et al., 2019).

While this methodology offers significant strengths in terms of theoretical depth and interdisciplinary scope, it also entails limitations that warrant acknowledgment. The reliance on secondary sources means that the analysis is constrained by the assumptions and data quality of existing studies, potentially perpetuating biases or blind spots present in the literature (Grahn et al., 2021). Additionally, the absence of empirical validation limits the ability to assess the practical performance of specific FaaS implementations, underscoring the need for future research that combines conceptual analysis with real-world case studies (Deshpande, 2024). Nevertheless, the chosen methodology is well-suited to the article's objective of advancing foundational understanding and setting an agenda for subsequent empirical inquiry.

RESULTS

The analytical synthesis of the literature yields several interrelated findings that collectively illuminate the structural, operational, and computational characteristics of Fleet-as-a-Service systems. One prominent result is the consistent demonstration across studies that dynamic fleet management significantly enhances the efficiency and sustainability of first- and last-mile mobility services when compared to fixed-route or privately owned vehicle models (Li & Quadrifoglio, 2010). Research on demand-responsive feeder services indicates that flexible allocation of vehicles in response to real-time demand can reduce vehicle kilometers traveled and improve service coverage, particularly in low-density or peripheral urban areas (Leffler et al., 2021).

Another key finding concerns the role of predictive analytics and real-time coordination in enabling FaaS platforms to function effectively at scale. Studies employing micro-simulation and agent-based modeling show that integrating demand prediction with transit coordination enhances the reliability and attractiveness of shared mobility services, thereby increasing user adoption and system utilization (Grahn et al., 2021; Huang et al., 2021). These results underscore the importance of computational infrastructure capable of processing large volumes of data with minimal latency, a requirement that aligns closely with advancements in serverless and edge computing paradigms (Nguyen et al., 2019).

The analysis further reveals that FaaS platforms serve as continuous testing environments for emerging vehicle technologies, particularly automated and connected vehicles. By operating vehicles in diverse, real-world conditions, FaaS systems generate extensive operational data that support iterative testing, validation, and optimization processes (Deshpande, 2024). This finding highlights a departure from traditional, stage-gated testing models and suggests that operational deployment itself becomes a critical component of the innovation lifecycle (Huang et al., 2021).

From a computational perspective, the results indicate that real-time scheduling and task partitioning are central challenges in FaaS implementations. Literature on real-time systems demonstrates that guaranteeing performance and safety in distributed, heterogeneous environments requires sophisticated scheduling algorithms capable of balancing latency, energy consumption, and workload distribution (Baruah, 2013; Mascitti et al., 2020). These findings suggest that the success of FaaS is contingent not only on transportation planning but also on advances in computing systems design that support predictable and efficient execution (Roy et al., 2021).

DISCUSSION

The findings of this study invite a deeper theoretical interpretation that situates Fleet-as-a-Service within broader debates on sustainable mobility, digital infrastructure, and socio-technical system design. At a conceptual level, FaaS can be understood as an instantiation of service-dominant logic in transportation, wherein value is co-created through the dynamic interaction of users, vehicles, and digital platforms rather than embedded in physical assets alone (Deshpande, 2024). This perspective challenges long-standing assumptions about mobility provision and reframes vehicles as modular components within adaptable service ecosystems.

Comparative analysis of scholarly viewpoints reveals both convergence and divergence in how researchers conceptualize shared mobility systems. While there is broad agreement on the efficiency benefits of demand-responsive and shared services, debates persist regarding their scalability, equity impacts, and integration with existing transit networks (Kanuri et al., 2019). FaaS offers a potential resolution to some of these tensions by providing a unifying operational framework that can accommodate diverse service models, yet its implementation raises questions about governance, data ownership, and algorithmic transparency that remain underexplored in the literature (Deshpande, 2024).

The integration of computing systems theory into the discussion highlights an often-overlooked dimension of sustainability: computational sustainability. Real-time scheduling and serverless architectures enable elastic scaling and efficient resource use, but they also introduce energy costs and complexity that must be managed to avoid undermining environmental gains achieved through shared mobility (Elsakhawy & Bauer, 2020). This observation aligns with emerging research on energy-aware scheduling and workload management, suggesting that sustainable transportation systems must be designed in concert with sustainable computing infrastructures (Roy et al., 2021).

Limitations of current FaaS research are evident in the relative scarcity of empirical studies that examine long-term system performance and social outcomes. Much of the existing literature relies on simulation and modeling, which, while valuable, cannot fully capture the institutional, behavioral, and ethical dynamics that shape real-world deployments (Grahn et al., 2021). Future research should therefore prioritize longitudinal case studies and participatory design approaches that engage stakeholders across the mobility ecosystem (Deshpande, 2024).

Looking forward, the evolution of FaaS will likely be influenced by advances in autonomous vehicle technology, regulatory frameworks, and public perceptions of shared mobility. As these factors converge, there is a pressing need for interdisciplinary research that bridges transportation planning, computer science, and social sciences to ensure that FaaS systems are not only efficient but also equitable and accountable (Lau & Susilawati, 2021). By articulating a comprehensive theoretical foundation and identifying critical gaps, this article contributes to such an agenda and underscores the transformative potential of Fleet-as-a-Service in shaping sustainable urban futures (Deshpande, 2024).

CONCLUSION

Fleet-as-a-Service represents a paradigm shift in how mobility systems are conceptualized, tested, and operated, offering a pathway toward more sustainable, efficient, and adaptive transportation infrastructures. Through an extensive theoretical and interdisciplinary analysis, this article has demonstrated that FaaS is not merely a logistical innovation but a socio-technical system that integrates shared mobility, real-time computing, and sustainability principles. By situating FaaS within existing scholarship and critically examining its implications, the study advances understanding of the opportunities and challenges inherent in service-oriented fleet management and lays the groundwork for future empirical and theoretical exploration (Deshpande, 2024).

REFERENCES

1. Baruah, S. (2013). Partitioned EDF scheduling: A closer look. Real-Time Systems, 49(6), 715–729.Grahn, R., Qian, S., & Hendrickson, C. (2021). Improving the performance of first- and last-mile mobility services through transit coordination, real-time demand prediction, advanced reservations, and trip prioritization. Transportation Research Part C: Emerging Technologies, 133, 103430.

2. Elsakhawy, M., & Bauer, M. (2020). Faas2f: A framework for defining execution-SLA in serverless computing. IEEE Cloud Summit, 58–65.

3. Kanuri, C., Venkat, K., Maiti, S., & Mulukutla, P. (2019). Leveraging innovation for last-mile connectivity to mass transit. Transportation Research Procedia, 41, 655–669.

4. Nguyen, H. D., Zhang, C., Xiao, Z., & Chien, A. A. (2019). Real-time serverless: Enabling application performance guarantees. International Workshop on Serverless Computing, 1–6.

5. Li, X., & Quadrifoglio, L. (2010). Feeder transit services: Choosing between fixed and demand responsive policy. Transportation Research Part C: Emerging Technologies, 18(5), 770–780.

6. Deshpande, S. (2024). Fleet-as-a-service (FaaS): Revolutionizing sustainable vehicle testing and operations. International Journal of Applied Engineering & Technology, 6(1), 1854–1867.

7. Huang, Y., Kockelman, K. M., Garikapati, V., Zhu, L., & Young, S. (2021). Use of shared automated vehicles for first-mile last-mile service: Micro-simulation of rail-transit connections in Austin, Texas. Transportation Research Record, 2675(2), 135–149.

8. Martinez, L. M., Correia, G. H., & Viegas, J. M. (2015). An agent-based simulation model to assess the impacts of introducing a shared-taxi system: An application to Lisbon. Journal of Advanced Transportation, 49(3), 475–495.

9. Roy, A., et al. (2021). Energy-aware primary/backup scheduling of periodic real-time tasks on heterogeneous multicore systems. Sustainable Computing: Informatics and Systems, 29, 100474.

10. Goswami, S., Tallapragada, P., & Rambha, T. (2021). Optimal supply control for shared mobility with logit mode-choice dynamics. IEEE International Intelligent Transportation Systems Conference, 3509–3514.