Edge Computing as a Modern Trend in Information Technology: A Comprehensive Review.

Awiti Gideon Appiah (Student, University of Energy and Natural Resource),
Dr. Lazarus Kwao (Lecturer, Sunyani Technical University),
Benjamin Opoku Atuahene (Student, University of Energy and Natural Resource)

How to Cite this Article:

Appiah, A. G. & Atuahene, B. O. (2026). Edge Computing as a Modern Trend in Information Technology: A Comprehensive Review.. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(7), 1-9. https://doi.org/10.55041/ijcope.v2i7.218

https://doi.org/10.55041/ijcope.v2i7.218

Abstract

The rapid growth of Internet of Things (IoT) devices, together with the increasing demand for real-time data processing, has exposed the limitations of traditional centralized cloud computing architectures. Edge computing has emerged as a transformative paradigm that brings computation, storage, and networking closer to the data source, thereby reducing latency, conserving bandwidth, and enhancing privacy. This paper presents a comprehensive review of edge computing as a modern trend in information technology. We systematically examine the architectural foundations, enabling technologies, and deployment models that underpin edge computing ecosystems. Furthermore, we analyze prominent application domains including autonomous vehicles, smart cities, industrial IoT (IIoT), healthcare, and augmented or virtual reality where edge computing delivers measurable performance improvements. A critical assessment of open challenges such as security vulnerabilities, resource constraints, interoperability, and orchestration complexity is also provided. Finally, we outline future research directions, including the convergence of edge computing with artificial intelligence (Edge AI), 6G networks, digital twins, and serverless edge architectures. This review aims to serve as a foundational reference for researchers and practitioners seeking to understand the current state and future direction of edge computing within the broader information technology landscape.

Keywords: Edge Computing, Internet of Things, Fog Computing, Cloud-Edge Continuum, Latency Reduction, Edge AI, Distributed Systems, Real-Time Processing.

1. Introduction

The digital transformation of the 21st century has been characterized by an unprecedented production of connected devices, data-intensive applications, and user expectations for instantaneous service delivery. According to recent industry projections, the global count of IoT-connected devices is expected to exceed 75 billion by 2030, generating data volumes measured in zettabytes annually [1]. Traditional cloud computing architectures, which rely on centralized data centers located at considerable geographic distances from end users and data sources, are increasingly unable to meet the stringent latency, bandwidth, and privacy requirements of next-generation applications [2].

Edge computing addresses these limitations by decentralizing computational resources and positioning them at the "edge" of the network that is in close proximity to the data-generating devices and end users [3]. This paradigm shift represents one of the most significant architectural evolutions in modern information technology, fundamentally altering how data is collected, processed, stored, and acted upon. The core proposition of edge computing is simple yet profound, rather than transmitting raw data to distant cloud servers for processing, computation is performed locally or at intermediate network nodes, enabling real-time decision-making and significantly reducing network congestion [4].

The relevance of edge computing extends across numerous domains. In autonomous driving, for instance, vehicles must process sensor data with sub-millisecond latency to make life-critical navigation decisions with which the centralized cloud architectures cannot reliably satisfy [5]. In industrial manufacturing, predictive maintenance systems powered by edge computing can detect equipment anomalies in real time, preventing costly downtime [6]. Similarly, in healthcare, edge-enabled wearable devices can perform on-device analysis of patient vitals, triggering immediate alerts without depending on continuous cloud connectivity [7].

Despite its promise, edge computing presents significant challenges. The distributed and heterogeneous nature of edge infrastructure complicates resource management, security enforcement, and software deployment. Edge devices often operate under severe constraints in terms of computational power, memory, and energy supply, necessitating novel optimization techniques [8]. Furthermore, the absence of universally adopted standards for edge architectures hinders interoperability and large-scale deployment [9].

This paper makes the following contributions:

1. **A comprehensive taxonomy** of edge computing architectures, distinguishing between edge computing, fog computing, mobile edge computing (MEC), and the cloud-edge continuum.
2. **A detailed analysis** of enabling technologies, including containerization, microservices, lightweight virtualization, and edge-native orchestration frameworks.
3. **An in-depth review** of application domains where edge computing demonstrates transformative potential.
4. **A critical discussion** of open challenges and research gaps that must be addressed for the maturation of edge computing ecosystems.
5. **A forward-looking perspective** on emerging trends, including Edge AI, integration with 6G networks, and sustainable edge computing.

The remainder of this paper is organized as follows: Section 2 provides background and traces the evolution of edge computing. Section 3 details the architectural foundations. Section 4 examines enabling technologies. Section 5 reviews application domains. Section 6 discusses challenges and open issues. Section 7 outlines future research directions. Section 8 concludes the paper.

2. Background and Evolution

2.1 From Cloud Computing to Edge Computing

Cloud computing, formalized in the late 2000s, revolutionized information technology by offering on-demand access to scalable computational resources over the Internet [10]. The centralized model of cloud computing provided economies of scale, simplified infrastructure management, and enabled the rapid deployment of global services. However, the physical distance between end devices and cloud data centers inherently introduces latency, typically ranging from tens to hundreds of milliseconds, depending on network conditions and geographic proximity [11].

The emergence of latency-sensitive and bandwidth-intensive applications such as real-time video analytics, augmented reality (AR), and autonomous systems,necessitated a re-evaluation of the centralized paradigm. Researchers and industry stakeholders recognized that processing data closer to its source could dramatically improve application responsiveness, reduce backhaul traffic, and enhance data sovereignty [12].

2.2 Historical Milestones

The conceptual foundations of edge computing can be traced to Content Delivery Networks (CDNs) deployed in the late 1990s, which cached web content at geographically distributed edge servers to reduce load times [13]. The formalization of edge computing as a distinct research domain gained momentum in the 2010s, driven by several key developments:

- **2012–2014:** The concept of "fog computing" was introduced by Cisco Systems to describe a decentralized computing infrastructure positioned between IoT devices and cloud data centers [14].
- **2014–2016:** The European Telecommunications Standards Institute (ETSI) established the Mobile Edge Computing (MEC) initiative, later renamed Multi-access Edge Computing, to leverage telecommunications infrastructure for edge processing [15].
- **2017–2019:** Open-source projects such as EdgeX Foundry, OpenEdge, and KubeEdge matured, providing standardized frameworks for edge application development and deployment [16].
- **2020–2023:** The integration of edge computing with 5G networks accelerated, with network operators deploying Multi-access Edge Computing (MEC) platforms at base stations to support ultra-reliable low-latency communication (URLLC) use cases [17].
- **2024–2026:** The emergence of generative AI and large language models has catalyzed interest in "Edge AI," driving research into model compression, federated learning, and AI inference at the edge [18].

2.3 Defining the Edge

The term "edge" is inherently contextual. In a network topology, the edge refers to the periphery that is the points where end devices connect to the network infrastructure. However, the precise location of the edge varies depending on the application and deployment context. In an IoT scenario, the edge may be a gateway device in a smart building. In a telecommunications context, the edge may be a base station or central office. In a vehicular context, the edge may be an on-board computing unit or Roadside System Unit (RSU) [19].

This contextual variability has given rise to a layered conceptualization of the edge, often described as a "cloud-edge-device continuum," wherein computational resources are distributed across multiple tiers with varying capabilities, latencies, and proximity to data sources [20].

3. Architectural Foundations

3.1 Three-Tier Architecture

The most widely adopted conceptual model for edge computing is the three-tier architecture:

- **Device Tier:** Comprises end-user devices and IoT sensors that generate data. These devices may possess limited computational capabilities (e.g., microcontrollers) or moderate capabilities (e.g., smartphones, industrial controllers).
- **Edge Tier:** Consists of edge nodes such as gateways, routers, base stations, and micro data centers that provide computational resources proximate to devices. The edge tier performs data filtering, aggregation, analytics, and short-term storage.
- **Cloud Tier:** Encompasses centralized data centers that handle resource-intensive tasks such as large-scale model training, long-term data storage, and global analytics [21].

Figure 1: *Edge Computing as three tier architecture.*

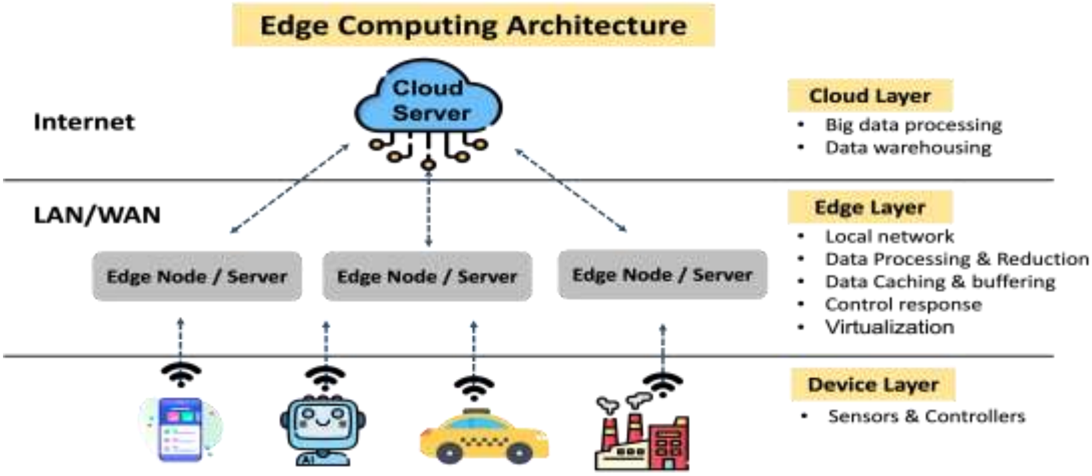

Figure : Edge computing architecture overview
Source : The research team

3.2 Fog Computing vs. Edge Computing

While often used interchangeably, fog computing and edge computing exhibit subtle distinctions. Fog computing, as originally defined, emphasizes a hierarchical, multi-layer architecture that extends cloud services to the network edge through intermediary fog nodes [14]. Edge computing, in contrast, more broadly refers to any paradigm that places computation at or near the data source, without necessarily implying a hierarchical structure [22]. In practice, the distinction has blurred, and both terms are often subsumed under the broader umbrella of "edge computing."

3.3 Multi-Access Edge Computing (MEC)

MEC framework integrates edge computing capabilities directly into the radio access network (RAN) of mobile telecommunications systems. By collocating application servers with cellular base stations, MEC enables ultra-low latency services for mobile users and provides application developers with access to real-time network information, such as user location and radio conditions [15]. MEC is a cornerstone of 5G-enabled edge computing and is expected to play a critical role in 6G architectures.

3.4 Cloud-Edge Continuum

Contemporary research increasingly views cloud and edge not as binary alternatives but as endpoints of a computational continuum. The cloud-edge continuum model posits that workloads can be dynamically placed across a spectrum of resources that is from on-device processors to regional edge nodes to centralized cloud servers which is based on real-time assessments of latency requirements, resource availability, energy constraints, and data sensitivity [23]. This model requires sophisticated orchestration mechanisms capable of managing workload migration, state synchronization, and quality-of-service (QoS) enforcement across heterogeneous infrastructure.

4. Enabling Technologies

4.1 Containerization and Lightweight Virtualization

Containers, particularly Docker containers, have become the de facto unit of software deployment in edge computing due to their lightweight nature, portability, and rapid startup times [24]. Unlike virtual machines (VMs), containers share the host operating system kernel, resulting in significantly lower overhead that is a critical consideration for resource-constrained edge devices. For devices with even more limited resources, lightweight alternatives such as WebAssembly (Wasm) modules and unikernels are being explored [25].

4.2 Edge Orchestration Frameworks

The management of distributed edge workloads requires orchestration frameworks that can schedule, deploy, monitor, and scale applications across heterogeneous edge nodes. Prominent frameworks include:

- **Kubernetes with Edge Extensions:** Projects such as KubeEdge, K3s, and MicroK8s extend Kubernetes to edge environments, enabling container orchestration with reduced resource footprints [26].
- **OpenStack StarlingX:** An open-source platform for edge cloud infrastructure management [27].
- **Azure IoT Edge and AWS Greengrass:** Proprietary platforms that extend cloud services to edge devices with local execution capabilities.

4.3 Edge AI and Machine Learning

The deployment of AI and machine learning (ML) models at the edge which is termed Edge AI, enables real-time inference without reliance on cloud connectivity.

Key techniques include:

- **Model Compression:** Techniques such as quantization, pruning, and knowledge distillation reduce the size and computational requirements of deep learning models, enabling deployment on resource-constrained devices [28].
- **Federated Learning:** A distributed ML paradigm where models are trained collaboratively across edge devices without centralizing raw data, thereby preserving privacy while leveraging distributed data [29].
- **TinyML:** The application of machine learning to ultra-low-power microcontrollers, enabling intelligence in devices with minimal computational resources [30].

4.4 Software-Defined Networking (SDN) and Network Function Virtualization (NFV)

SDN and NFV provide the networking foundations for flexible, programmable edge infrastructure. SDN decouples the control plane from the data plane, enabling dynamic traffic management and routing optimization. NFV virtualizes network functions, allowing them to be deployed as software instances on commodity hardware at the edge [31].

4.5 Blockchain for Edge Security

Blockchain technology has been proposed as a mechanism to enhance security and trust in edge computing environments. Decentralized consensus mechanisms can enable secure device authentication, tamper-proof logging of edge transactions, and decentralized access control in scenarios where a central authority is absent or untrusted [32].

5. Application Domains

5.1 Autonomous Vehicles and Intelligent Transportation

Autonomous vehicles generate terabytes of data per hour from LiDAR, cameras, radar, and other sensors. Real-time processing of this data for object detection, path planning, and decision-making demands latencies below 10 milliseconds, a requirement that necessitates on-vehicle and roadside edge computing [5]. Vehicle-to-everything (V2X) communication further benefits from edge infrastructure for cooperative perception and traffic optimization.

5.2 Smart Cities and Urban Analytics

Edge computing underpins smart city applications including intelligent traffic management, environmental monitoring, public safety surveillance, and smart lighting. Edge nodes deployed throughout urban environments process sensor data locally, reducing the bandwidth requirements for transmitting continuous video streams and enabling real-time anomaly detection [33].

5.3 Industrial IoT and Industry 4.0

In manufacturing and industrial settings, edge computing enables predictive maintenance, quality control, robotics coordination, and digital twin simulations. By processing sensor data from industrial equipment at the edge, manufacturers can detect anomalies, optimize production processes, and reduce unplanned downtime without exposing sensitive operational data to external cloud services [6].

5.4 Healthcare and Remote Patient Monitoring

Edge computing facilitates real-time analysis of patient data from wearable devices, implantable sensors, and medical imaging systems. On-device and near-device processing enables immediate clinical alerts, reduces the risk associated with network outages, and enhances patient privacy by minimizing the transmission of sensitive health data [7].

5.5 Augmented Reality, Virtual Reality, and the Metaverse

Immersive experiences in AR and VR demand high-bandwidth, low-latency rendering and interaction. Edge computing enables split-rendering architectures, wherein computationally intensive graphics rendering is offloaded from headsets to proximate edge servers, reducing device weight and power consumption while maintaining high visual fidelity [34].

5.6 Telecommunications and 5G/6G Networks

Telecommunications operators are deploying edge computing infrastructure to support network slicing, service chaining, and URLLC use cases. MEC platforms collocated with 5G base stations enable operators to offer differentiated services with guaranteed latency and throughput parameters [17].

6. Challenges and Open Issues

6.1 Security and Privacy

The distributed nature of edge computing significantly expands the attack surface compared to centralized cloud architectures. Edge nodes are often deployed in physically accessible locations, making them vulnerable to tampering, physical attacks, and unauthorized access [35]. Furthermore, the heterogeneity of edge devices complicates the uniform enforcement of security policies. Key security challenges include:

- Secure authentication and identity management for billions of heterogeneous devices.
- Protection of data in transit and at rest across untrusted edge infrastructure.
- Mitigation of distributed denial-of-service (DDoS) attacks targeting edge nodes.
- Ensuring compliance with data protection regulations (e.g., GDPR, HIPAA) in distributed processing environments.

6.2 Resource Constraints and Heterogeneity

Edge devices and nodes exhibit extreme heterogeneity in terms of computational capacity, memory, storage, power supply, and network connectivity. Designing software and algorithms that can operate effectively across this diverse spectrum remains a fundamental challenge [8]. Resource management techniques must account for dynamic variations in workload demand, device availability, and network conditions.

6.3 Interoperability and Standardization

The absence of universally adopted standards for edge computing architectures, APIs, and communication protocols creates fragmentation and vendor lock-in. While organizations such as ETSI, the Open Computing Foundation (OCF), the Linux Foundation Edge (LF Edge), and the Industrial Internet Consortium (IIC) are developing standards, convergence remains incomplete [9].

6.4 Orchestration and Management Complexity

Managing workloads across thousands or millions of geographically distributed, heterogeneous edge nodes requires orchestration systems of unprecedented scale and sophistication. Challenges include:

- Dynamic workload placement and migration across the cloud-edge continuum.
- Service discovery and load balancing in highly dynamic environments.
- Consistent state management and data synchronization across distributed nodes.
- Automated fault detection, recovery, and self-healing [36].

6.5 Energy Efficiency and Sustainability

The proliferation of edge computing infrastructure raises concerns about energy consumption and environmental sustainability. Edge data centers, while smaller than hyperscale cloud facilities, collectively represent a significant and growing energy demand. Research into energy-efficient hardware, green computing practices, and renewable energy integration at the edge is increasingly important [37].

6.6 Reliability and Fault Tolerance

Edge environments are inherently less reliable than controlled data center environments due to variable network connectivity, power instability, and physical exposure. Designing fault-tolerant systems that can maintain service continuity in the face of node failures, network partitions, and intermittent connectivity is a critical research challenge [38].

7. Future Research Directions

7.1 Edge AI and Generative Models at the Edge

The deployment of large language models (LLMs) and generative AI at the edge represents a frontier of active research. Techniques for running compressed, distilled, or quantized versions of large models on edge hardware—while preserving acceptable output quality—are under intensive investigation [18]. This direction has profound implications for privacy-preserving AI, offline-capable intelligent assistants, and real-time content generation.

7.2 Integration with 6G Networks

As 6G networks are developed (expected commercial deployment circa 2030), the integration of edge computing with native AI, terahertz communication, and integrated sensing and communication (ISAC) will create new architectural possibilities. 6G is anticipated to feature "network-native AI," wherein intelligence is embedded at every layer of the network, including the edge [39].

7.3 Digital Twins and Edge Computing

The convergence of edge computing with digital twin technology enables the creation of real-time, high-fidelity virtual representations of physical systems. Edge computing provides the low-latency data processing required to keep digital twins synchronized with their physical counterparts, enabling applications in smart manufacturing, urban planning, and healthcare simulation [40].

7.4 Serverless Edge Computing

Serverless computing models, which abstract infrastructure management from developers, are being adapted for edge environments. Serverless edge computing promises to simplify application development and deployment by enabling function-as-a-service (FaaS) execution at the edge, with automatic scaling and resource management [41].

7.5 Sustainable and Green Edge Computing

Future research must address the environmental impact of large-scale edge computing deployment. This includes the development of energy-aware scheduling algorithms, the use of renewable energy sources for edge data centers, carbon-aware workload placement, and hardware recycling strategies [37].

7.6 Quantum-Edge Computing Convergence

Although in its nascent stages, the potential integration of quantum computing resources with edge infrastructure is an area of speculative but potentially transformative research. Quantum processors deployed at strategic edge locations could accelerate optimization, cryptography, and simulation tasks that are intractable for classical edge hardware [42].

8. Conclusion

Edge computing has evolved from a supplementary architectural concept to a foundational pillar of modern information technology. By decentralizing computation and placing it in proximity to data sources and end users, edge computing addresses the critical limitations of centralized cloud architectures that is latency, bandwidth, privacy, and reliability. This comprehensive review has examined the architectural foundations, enabling technologies, application domains, and challenges that define the current state of edge computing.

The trajectory of edge computing is deeply intertwined with broader technological trends, including the proliferation of IoT, the advancement of AI, the deployment of 5G/6G networks, and the growing imperative for data sovereignty and sustainability. As these trends converge, edge computing will increasingly serve as the computational substrate upon which intelligent, responsive, and resilient digital systems are built.

However, significant challenges remain. The security, interoperability, orchestration, and sustainability of large-scale edge computing ecosystems require sustained research investment and cross-industry collaboration. The establishment of robust standards, the development of energy-efficient hardware and algorithms, and the creation of developer-friendly tools and platforms are essential for realizing the full potential of edge computing.

We anticipate that the coming decade will witness the maturation of edge computing from a promising paradigm into a ubiquitous infrastructure layer—pervasive, intelligent, and indispensable to the fabric of digital society.

References

[1] Statista, "Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2030," *Statista Research Department*, 2025.

[2] M. Satyanarayanan, "The emergence of edge computing," *Computer*, vol. 50, no. 1, pp. 30–39, 2017.

[3] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.

[4] P. Garcia Lopez et al., "Edge-centric computing: Vision and challenges," *ACM SIGCOMM Computer Communication Review*, vol. 45, no. 5, pp. 37–42, 2015.

[5] K. Abboud, H. A. Omar, W. Zhuang, and W. Li, "Stochastic modeling of single-hop cluster-based vehicular safety applications," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5739–5753, 2015.

[6] J. Wan et al., "Context-aware cloud robotics for material handling in cognitive manufacturing internet of things," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2272–2281, 2018.

[7] M. M. Dhanvijay and S. C. Patil, "Internet of Things: A survey of enabling technologies in healthcare and its applications," *Computer Networks*, vol. 153, pp. 113–131, 2019.

[8] Y. C. Hu, M. Patel, D. Sabella, N. Sprecher, and V. Young, "Mobile edge computing—A key technology towards 5G," *ETSI White Paper*, no. 11, pp. 1–16, 2015.

[9] Linux Foundation Edge (LF Edge), "State of the Edge Report," 2024. [Online]. Available: https://stateoftheedge.com

[10] M. Armbrust et al., "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[11] B. Varghese, N. Wang, S. Barbhuiya, P. Kilpatrick, and D. S. Nikolopoulos, "Challenges and opportunities in edge computing," in *Proc. IEEE International Conference on Smart Cloud (SmartCloud)*, 2016, pp. 20–26.

[12] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proc. MCC Workshop on Mobile Cloud Computing*, 2012, pp. 13–16.

[13] E. Nygren, R. K. Sitaraman, and J. Sun, "The Akamai network: A platform for high-performance internet applications," *ACM SIGOPS Operating Systems Review*, vol. 44, no. 3, pp. 2–19, 2010.

[14] F. Bonomi, R. Milito, P. Natarajan, and J. Zhu, "Fog computing: A platform for internet of things and analytics," in *Big Data and Internet of Things: A Roadmap for Smart Environments*, Springer, 2014, pp. 169–186.

[15] ETSI, "Mobile Edge Computing (MEC): Framework and Reference Architecture," *ETSI GS MEC 003 V1.1.1*, 2016.

[16] EdgeX Foundry, "EdgeX Foundry: A Vendor-Neutral Open Source Framework for IoT Edge Computing," Linux Foundation, 2023.

[17] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1657–1681, 2017.

[18] J. Shao and J. Zhang, "Communication-computation trade-off in resource-constrained edge inference," *IEEE Communications Magazine*, vol. 58, no. 12, pp. 20–26, 2020.

[19] H. Elazhary, "Internet of Things (IoT), mobile cloud, cloudlet, mobile IoT, IoT cloud, fog, mobile edge, and edge emerging computing paradigms: Disambiguation and research directions," *Journal of Network and Computer Applications*, vol. 128, pp. 105–140, 2019.

[20] O. Skarlat, M. Nardelli, S. Schulte, M. Borkowski, and P. Leitner, "Optimized IoT service placement in the fog," *Service Oriented Computing and Applications*, vol. 11, no. 4, pp. 427–443, 2017.

[21] A. Yousefpour, C. Fung, T. Nguyen, K. Kadiyala, F. Jalali, A. Niakanlahiji, J. Kong, and J. P. Jue, "All one needs to know about fog computing and related edge computing paradigms: A complete survey," *Journal of Systems Architecture*, vol. 98, pp. 289–330, 2019.

[22] M. Aazam, I. Khan, A. A. Alsaffar, and E. Huh, "Cloud of Things: Integrating Internet of Things and cloud computing and the issues involved," in *Proc. International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, 2014, pp. 414–419.

[23] D. Sabella, A. Vaillant, P. Kuure, U. Rauschenbach, and F. Giust, "Mobile-edge computing architecture: The role of MEC in the Internet of Things," *IEEE Consumer Electronics Magazine*, vol. 5, no. 4, pp. 84–91, 2016.

[24] D. Merkel, "Docker: Lightweight Linux containers for consistent development and deployment," *Linux Journal*, vol. 2014, no. 239, p. 2, 2014.

[25] A. Haas et al., "Bringing the web up to speed with WebAssembly," in *Proc. ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI)*, 2017, pp. 185–200.

[26] J. Jiang, L. Hu, C. Xue, and L. Zhu, "An edge computing-enabled NOMA transmission system for 5G wireless networks," in *Proc. IEEE International Conference on Communications (ICC)*, 2019, pp. 1–6.

[27] OpenStack StarlingX, "StarlingX: Complete Cloud Infrastructure Software Stack for Edge Computing," Open Infrastructure Foundation, 2024.

[28] Y. Cheng, D. Wang, P. Zhou, and T. Zhang, "Model compression and acceleration for deep neural networks," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 126–136, 2018.

[29] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017, pp. 1273–1282.

[30] V. J. Reddi et al., "TinyML: Machine learning with TensorFlow Lite on Arduino and ultra-low-power microcontrollers," *O'Reilly Media*, 2020.

[31] B. Han, V. Gopalakrishnan, L. Ji, and S. Lee, "Network function virtualization: Challenges and opportunities for innovations," *IEEE Communications Magazine*, vol. 53, no. 2, pp. 90–97, 2015.

[32] M. A. Khan and K. Salah, "IoT security: Review, blockchain solutions, and open challenges," *Future Generation Computer Systems*, vol. 82, pp. 395–411, 2018.

[33] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22–32, 2014.

[34] M. Chen, Y. Hao, Y. Li, C. F. Lai, and D. Wu, "On the computation offloading at ad hoc cloudlet: Architecture and service modes," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 18–24, 2015.

[35] R. Roman, J. Zhou, and J. Lopez, "On the features and challenges of security and privacy in distributed internet of things," *Computer Networks*, vol. 57, no. 10, pp. 2266–2279, 2013.

[36] C. Pahl and P. Jamshidi, "Microservices: A systematic mapping study," in *Proc. International Conference on Web Services (ICWS)*, 2016, pp. 137–146.

[37] A. Andrae and T. Edler, "On global electricity usage of communication technology: Trends to 2030," *Challenges*, vol. 6, no. 1, pp. 117–157, 2015.

[38] Y. Xiao, Y. Jia, C. Liu, X. Cheng, J. Yu, and W. Lv, "Edge computing security: State of the art and challenges," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1608–1631, 2019.

[39] M. Latva-aho and K. Leppänen, "6G research visions: Key drivers and research directions," *University of Oulu White Paper*, 2019.

[40] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2405–2415, 2019.

[41] E. Jonas et al., "Cloud programming simplified: A Berkeley view on serverless computing," *UC Berkeley Technical Report*, 2019.

[42] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 201

