Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 11
November 2024

Towards Scalable IoT Architectures: Merging Cloud and Edge Computing for Optimized Performance

Tamasa Priyadarsini [1],
[1]Asst. Professor, Dept. of CSE, Gandhi Institute of Excellent Technocrats, Ghangapatna
[1]Tamasa.priyadarsini@gmail.com,

Abstract

The Internet of Things (IoT) is a rapidly advancing technology, revolutionizing the way everyday objects connect and interact. By embedding hardware into various devices, IoT creates a network of internet-connected objects that can transform our lives and work. IoT devices enable consumers to manage energy use, enhance convenience, and interact with their environment in novel ways. For instance, mobile devices can now control home appliances remotely using smart infrared wireless technology. Through the IoT, users can connect with services provided by component suppliers, connectivity providers, and data analysis experts. As IoT grows, businesses can develop more sophisticated products that improve energy efficiency, lower costs, and make cutting-edge technology more accessible. By 2030, the number of IoT devices globally is expected to exceed 29 billion, reflecting their increasing role across consumer markets and industry sectors. IoT architecture consists of three primary layers: the physical layer for data collection, the information layer for data processing and storage, and the application layer for delivering services to users. However, with the rise of IoT comes a need for robust data management solutions that address processing power, latency, and security limitations, particularly in cloud and edge computing environments. As interconnected devices grow, optimizing IoT systems through the convergence of cloud and edge computing becomes essential to ensure scalability, real-time decision-making, and secure data management.

Introduction:

One area of technology that is advancing quickly is the Internet of Things (IoT). Through connecting various devices in an embedded hardware platform, IoT offers an environment of Internet-connected objects surrounding us. Internet-enabled items can reinvent life and benefit people enormously [2]. Customers can save energy by putting a mobile device with a smart infrared wireless IR remote control on their cell phones and using it to control gadgets remotely. Customers can access the functional components of the Internet of Things (IoT), including component suppliers, connectivity providers, and data analysis providers, among others, to establish connections with the IoT world. With the use of IoT platforms, businesses will be able to develop more advanced technological iterations of their current products, lowering their energy and power costs and making them more reasonably priced. Over 29 billion Internet of Things (IoT) devices are expected to exist globally by 2030, almost doubling from 15.1 billion in 2020 [1]. By 2030, India will have more IoT device users than any other nation, with about 8 billion of them living there.

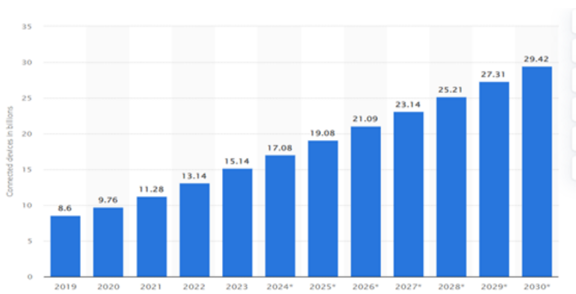

Fig. 1: Number of Internet of Things (IoT)

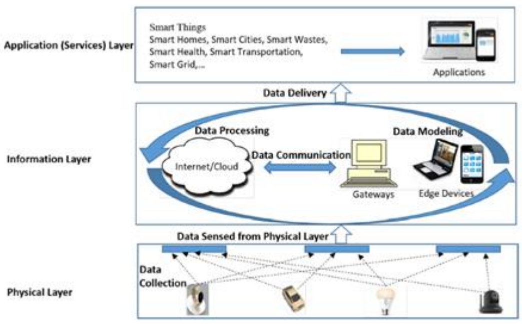

Fig.-2: A generic architecture of IoT systems

IoT devices are employed throughout several consumer markets and industry verticals; by 2020, about 60% of all IoT-connected devices globally, spanning from 2019 to 2023, will be associated with consumers. This fraction should stay at this level over the next ten years. Major industrial verticals such as waste management, retail and wholesale, gas, steam and electricity, transportation and storage, and government are among the more than 100 million IoT devices linked to these verticals. By 2030, there will be more than eight billion IoT devices globally, spanning all industry sectors. Consumer internet and media devices, including smartphones, constitute the most important application for Internet of Things devices in the consumer sector. By 2030, there will be over one billion IoT devices in use cases such as IT infrastructure, smart grid, asset tracking and monitoring, and linked (autonomous) cars.

IoT Architecture

The three main tiers that comprise the generic architecture of Internet of Things systems are the physical layer (data collection), the information layer (sharing, analytics, and storage), and the application layer (data utilisation), as seen in Fig. 2. The network technologies and sensors that comprise an IoT ecosystem provide data to the Physical Layer. Sensors can talk to cloud-based servers or other sensors in IoT services. Due to their low processing power and low radio frequency operation mode, IoT devices cannot send big messages. This layer is where physical devices like RFID (radio frequency identification) sensors and others are used for data collection. Just as crucial as protecting data transported to or from the device is safeguarding the hardware. The IoT systems' inadequate hardware and application security brought the most frequent risks in the physical layer. Devices at this level must be safeguarded to safeguard owners' privacy and the data gathered.

After processing the data that has been gathered, the information layer acts in accordance with the demands and specifications of the application. After that, the data is distributed or stored to the application layer. Because it makes service creation and integration easier, this layer has drawn the interest of researchers recently [3]. To protect user privacy and guarantee the authenticity of the data being collected, security and privacy components are necessary. This layer allows for the processing of data and decision-making, guaranteeing the safety and preservation of the privacy of the moving data, data owners, and data consumers, among other uses for various algorithms and methodologies. Based on the final service implementation, the Application layer concentrates on the feature definition for service provisioning. The planned features of the system are made available to the user via this layer. Additionally, it receives all of the Information Layer's features [3].

Combination of cloud and IoT

With its establishment and growth, cloud computing now offers a platform that makes it easier for apps to store and analyse data. In order to store and process data, IoT sensors typically rely on cloud computing [4]. Cloud offers new business models and opportunities to showcase created solutions for enhancing existing information systems [5]. However, two concerns are made possible by cloud computing technology: security and privacy [6]. The goal of integrating cloud computing and the Internet of objects is to provide consumers with dependable and adaptable decision-making tools while transforming common resources like machines, work processes, and sensors into smart objects [7]. The IoT and cloud computing represent a synergistic alliance that may revolutionize several industries and enhance the capabilities of connected things. Cloud computing serves as a powerful backbone for IoT ecosystems by providing scalable and on-demand access to computing resources, storage, and services. This partnership addresses the inherent challenges of IoT devices, such as restricted memory and processing power, by offloading intensive computations and data storage to the cloud. Through this collaboration, IoT devices can leverage the cloud's vast computational capabilities, allowing for real-time data analysis, complex analytics, and machine learning applications. Moreover, cloud services facilitate seamless connectivity and communication among diverse IoT devices, enabling them to work together efficiently. The integration of cloud and IoT not only enhances the performance of individual devices but also fosters the development of intelligent, interconnected systems that can generate valuable insights and support informed policymaking. However, addressing data security, privacy, and latency concerns is essential to fully unlock this combination's potential and confirm its widespread adoption across various sectors.

Combining cloud and IoT opens up new possibilities for cost-effective solutions and business models. Cloud services enable a more flexible and scalable infrastructure, allowing organizations to deploy and manage IoT applications with greater efficiency. This flexibility is particularly advantageous for businesses that experience fluctuating workloads or are looking to scale their IoT deployments rapidly. Pay-as-you-go cloud computing eliminates the need for large upfront expenditures in hardware and infrastructure, making IoT adoption more accessible to a broader range of enterprises, including smaller businesses and startups.

Furthermore, the collaboration between cloud computing and IoT supports the progress of edge computing abilities. Edge computing processes data closer to the source, which reduces latency and enhances real-time responsiveness. By combining cloud resources with edge computing, organizations can strike a balance between centralized processing for complex tasks and localized processing for time-sensitive operations. This not only improves the overall performance of IoT systems but also minimizes the reliance on constant high-bandwidth connectivity.

Combination of Cloud and Edge

Edge Computing Combined with the Cloud Now that we are familiar with the concept of the cloud let me explain. The cloud provides services like platform, storage, and other resources via the internet from remote data centres, allowing us to use our resources anywhere at any time. We gain from efficiency in that we can save money on material resources. Among the drawbacks of the cloud are its high latency, constrained capacity, data security measures, and Internet connectivity [8]. These days, there is a rise in internet usage as well as the proliferation of IoT applications, including those for smartphones, computers, sensors, Google apps, and various social media platforms. We utilise these gadgets and Internet of Things apps on a daily basis. Additionally, in order for data streams to transfer to the cloud and end user, all data sources must be kept in a single, centralised cloud. Because of this, real-time data analysis in cloud computing is more important than high-speed internet connectivity. This has led to security concerns as well as issues with service latency in the current cloud. Therefore, edge computing—the processing of real-time data—needs intelligence and specialised service architecture.

Edge computing-based cloud-based Internet of Things apps

Edge computing and cloud services are increasingly integrated to support IoT applications, including devices such as smartphones, smart homes, smart grids, smart healthcare, smart cities, etc. The main goal is for cloud servers to deliver services efficiently at the edge of multiple servers. Edge computing is needed to improve the quality of service (QoS) of these devices by reducing latency, optimizing bandwidth utilization, reducing power consumption, and improving security[4]. Despite its many advantages over cloud computing, edge computing can’t completely replace it. Many edge devices still rely on centralized application servers in large data centres because cloud providers hold most of the computing, storage and networking resources needed for advanced services. This close collaboration between edge and cloud infrastructures works well. It also ensures a scalable IoT ecosystem.

IoT-Based Cloud with Edge Computing Application

Cloud computing and edge computing are two pivotal paradigms shaping the landscape of modern information technology, and their convergence finds a dynamic application in the realm of the IoT. In the case of IoT applications, this connection between cloud and edge computing is essential for maximising data processing, storage, and transmission. Cloud computing provides a powerful centralised infrastructure, enabling scalable storage and computing resourcesNevertheless, edge computing aims to improve real-time processing capabilities, decrease latency, and move computation closer to the data source. This integration offers a comprehensive solution in the Internet of Things, where numerous devices generate enormous volumes of data.

In IoT applications, the coexistence of cloud and edge computing enables effective data processing and essential decision-making at the edge while utilising the cloud's scalability and storage capabilities. This is particularly advantageous in scenarios where low-latency responses are essential, such as automatic automobiles, smart cities, and manufacturing computerization. Additionally, it optimises resources by offloading certain tasks to the cloud, ensuring that edge devices operate optimally [8]. As the IoT ecosystem expands, the collaborative utilization of cloud and edge computing will likely become increasingly indispensable for unlocking the full potential of connected devices and enabling innovative applications across diverse domains.In real-world applications, combining cloud and edge computing in IoT transforms industries and enhances user experiences. One prominent example is healthcare, where wearable sensors can nonstop monitor vital signs. Edge devices process this real-time data locally to identify critical health events while transmitting aggregated information to the cloud for in-depth analysis and long-term trend monitoring. This architecture ensures prompt emergency response and facilitates personalized healthcare insights through cloud-based analytics.

Another compelling application lies in the area of smart cities. Edge computing empowers connected equipment like cameras and sensors to process data on-site, enabling swift responses to traffic congestion, accidents, or environmental changes. Simultaneously, cloud platforms aggregate and analyze data from various edge devices to derive actionable insights for urban planning and resource optimization. This dynamic combination enhances the overall efficiency and responsiveness of smart city initiatives.These real-world instances highlight the cloud's adaptability and edge computing's influence on Internet of Things applications. The collaborative synergy between local processing and centralized analytics not only addresses the challenges of data volume and latency but also opens up new options intended for innovation across diverse sectors, ultimately reshaping how we interact with technology in our daily lives.

The advantages of mobile edge computing

Mobile Edge Computing (MEC) presents a paradigm shift in how data is processed, and services are delivered in mobile networks. One of the primary merits of MEC lies in its ability to reduce latency significantly. By moving data storage and processing closer to the network's edge, MEC minimizes the round-trip time for data between the device and a centralized cloud server. This low-latency advantage is critical for real-time responsiveness applications like augmented reality, virtual reality, and autonomous vehicles. MEC not only enhances the user experience by ensuring faster response times but also enables the development of innovative and immersive applications previously constrained by latency limitations [9]. Furthermore, Mobile Edge Computing contributes to efficient network resource utilization. By offloading computational tasks to edge servers, MEC helps alleviate the burden on the core network infrastructure. This enhances the overall network efficiency and reduces the load on centralized data centres, resulting in financial and energy savings. The distributed nature of MEC facilitates load balancing, ensuring that resources are utilized optimally across the network. This merit is particularly significant in the context of the burgeoning demand for data-intensive applications, as it allows mobile networks to cope with the increasing data traffic more effectively, ultimately improving the scalability and sustainability of the entire network ecosystem.

EDGE COMPUTING IN MACHINE LEARNING

Machine Learning (ML) coupled with Edge Computing heralds a transformative synergy It overcomes the drawbacks of conventional cloud-centric methods. Edge Computing involves processing data nearer to the generation source, minimizing latency, and then reducing the burden on centralized servers. Integrating ML algorithms at the edge empowers devices to make intelligent, real-time decisions without the need for continuous reliance on cloud connectivity. By deploying ML models at the edge, devices can analyse and respond to data locally, optimizing efficiency, and preserving bandwidth[10]. This not only enhances the overall performance of systems but also ensures privacy by limiting the transmission of sensitive information to the cloud.

The convergence of Machine Learning and Edge Computing introduces new opportunities and challenges for researchers and practitioners. Efficient utilization of limited resources at the edge, developing lightweight ML models, and exploring federated learning to preserve privacy are pivotal research areas. Furthermore, understanding the trade-offs between computation, communication, and model accuracy is essential for optimizing the performance of edge-based ML systems. As this interdisciplinary field continues to evolve, collaboration between computer scientists, engineers, and domain experts becomes paramount to harnessing the full potential of machine learning with edge computing in diverse applications ranging from healthcare and smart cities to agriculture and beyond.

Edge and Cloud Characteristics

Table 1 shows the distinctions between edge computing and cloud computing.

Table 1 Features that set Edge separate from the cloud

Characteristics	Cloud	Edge
Processing Speed	Slower	Faster
Bandwidth Usage	High	Low
Real time data processing	Millisecond	Microsecond
Storage	Centralized	Decentralized
Deployment cost	High	Low
Multi Environment	Support	Support
IoT device identification	more time required	Less timer required

Research challenges and future perspectives for IoT Cloud Edge

A thorough examination of the numerous research issues arising from the convergence of IoT, Cloud Computing, and Edge Computing is necessary. One significant challenge lies in efficiently managing massive sizes of information produced by IoT devices. The sheer scale and diversity of data types present obstacles in terms of storing, processing, then transmission. Researchers are tasked with developing innovative solutions that optimize data handling to ensure minimal latency and resource consumption

Another pressing challenge in IoT Cloud Edge research is the need for robust security and privacy measures. With the proliferation of interconnected devices, the attack surface for malicious actors widens, necessitating sophisticated security protocols. Privacy concerns also come to the forefront, as sensitive data traverses the network between IoT devices, the Edge, and the Cloud[6]. Addressing these challenges requires advanced cryptographic techniques and a comprehensive understanding of the legal and ethical implications surrounding data privacy. Furthermore, the dynamic nature of Edge environments introduces challenges related to resource constraints and varying network conditions. Optimizing resource utilization, balancing workloads, and adapting to changing network environments are critical aspects that demand attention in IoT Cloud Edge research [11]. Developing adaptive algorithms and frameworks that can dynamically adjust to the fluctuating conditions of the Edge will be essential for ensuring the reliability and scalability of IoT applications. Overall, navigating these multifaceted challenges is imperative in order to realize the maximum potential of IoT Cloud Edge systems and foster a resilient and secure connected ecosystem.

Conclusion

In summary, this research paper has explored the evolving landscape of IoT, cloud computing, and edge computing, highlighting the transformative potential of their integration across various industries. The combination of these technologies enhances data processing capabilities, reduces latency, and optimizes resource utilization, making real-time decision-making more efficient. Through an in-depth analysis of the interactions between IoT, cloud, and edge computing, this study emphasizes the importance of a collaborative framework that balances centralized cloud resources with localized edge processing. As the digital ecosystem continues to evolve, addressing the challenges related to data management, security, and scalability in IoT systems is crucial. This research underscores the need for ongoing innovation, standardization, and cooperative efforts to fully unlock the transformative power of IoT, cloud, and edge computing. By addressing these challenges, the future of interconnected systems will be more resilient, efficient, and responsive to the growing demands of an increasingly connected world.

References:

[1]https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/

[2] Ju, J., Kim, M., Ahn, J.: 'Prototyping business models for IoT service', Procedia Comput. Sci., 2016, 91, pp. 882–890, doi.org/10.1016/j.procs.2016.07.106

[3] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," Comput. Netw., vol. 54, no. 15, pp. 2787–2805, 2010.

[4] Bittencourt, L., Immich, R., Sakellariou, R., et al.: 'The internet of things, fog and cloud Continuum: integration and challenges', Internet Things, 2018, 3-4, pp. 134–155

[5] Wang, X.V., Wang, L.: 'A cloud-based production system for information and service integration: an internet of things case study on waste electronics', Enterp. Inf. Syst., 2017, 11, (7), pp. 952–968

[6] Chahid, I., Marzouk, A.: 'A secure Iot data integration in cloud storage systems using ABAC access control policy', Int. J. Adv. Eng. Res. Sci., 2017, (8), pp. 34–37

[7] Parto Dezfouli, M.: 'A secure MTconnect compatible iot platform for machine monitoring through integration of fog computing, cloud computing, and communication protocols', Georgia Institute of Technology, 2017

[8] M. Aydos, Y. Vural and A. Tekerek, "Assessing Risks and Threats with Layered Approach to Internet of Things Security," Measurement and Control, Vol. 52 no.5-6, pp. 338–353, 2019

[9] L. Cui, S. Chen, Y. Pan, Z. Ming and M. Xu, "A Decentralized and Trusted Edge Computing Platform for Internet of Things," IEEE Internet of Things Journal, vol. 7, no.5, May 2020.

[10] J. K. Verma, S. Kumar, O. Kaiwartya, Y. Cao, J. Lloret, C. P. Katti, and R. Kharel, "Enabling green computing in cloud environments: Network virtualization approach toward 5g support," Transactions on Emerging Telecommunications Technologies, vol. 29, no. 11, p. e3434, 2018.

[11] L.P. Qian, B. Ji, Y. Wu, L. Huang, D.H. Tsang, Hybrid IoT: Integration of hierarchical multiple access and computation offloading for IoT-based smart cities, IEEE Netw. 33 (2) (2019) 6–13.