Advancing IoT with Edge Computing: A Comparative Analysis of Cloud, Fog, and Edge Paradigms

Mohammad A Shahzad[1]
[1]University of Gloucestershire, United Kingdom
mohammad.shahzad@hotmail.co.uk[1]

Abstract

Recent advancements in Internet technology have significantly transformed the interaction between humans and electronic devices, leading to the rise of the Internet of Things (IoT). This paper explores the evolution and impact of edge computing, a technology that processes data near its source rather than relying solely on cloud computing. By comparing cloud, fog, and edge computing, we highlight the advantages of edge computing in reducing latency, enhancing security, and improving efficiency. The paper also addresses the challenges and opportunities in edge computing, emphasizing its potential applications in smart homes, cities, and collaborative environments. Edge computing represents a pivotal shift towards more responsive and efficient data processing, crucial for the future of IoT.

Keywords: Edge Computing, Internet of Things, Cloud Computing, Fog Computing, Smart Cities

Article Number: IOARP-JCN04-006
IOARP Journal of Communication and Networks
Volume 4, Issue 2, December 2024
Appeared on Digital Library: 15 January 2025
ISSN 3049-9356
www.ioarp.org

1. Introduction

Recent developments in Internet technology have transformed the way humans and electronic devices interact on a regular basis. Nowadays, various electronic devices are used for different purposes, such as security monitoring systems in homes or navigational systems in the wild. This activity at the edge of the network drives the extensive implementation of network-enabled devices and the interconnection of physical world activities with computer-based platforms, known as the Internet of Things (IoT).

A lot of IT departments in different organizations in the world are looking to install computing technologies specifically intended for the IoT. For this reason, edge data computing and analytics are required at the edge and in the area where huge amount of information is captured. Just as the cloud was the first of its kind off-premises network, edge can also be considered another off-premises network. Since its inception around 2005, cloud computing has enormously revolutionized human beings' lives in perspective of the way they live, work and study [1]. Software as a service (SaaS), for example Apps like Twitter, Google, Flicker and Facebook etc., are being highly utilized in our daily routine. In addition to the development of SaaS, more scalable and efficient infrastructures, along with processing engines, have been introduced to support cloud services. These advancements significantly enhance business operations, with examples including MapReduce, Apache Spark, Apache Hadoop, and Google File System [2].

The term IoT is around for a while now in the networking discipline, however, initially, it was introduced in 1999. The first-ever IoT was used for supply chain management and after that was highly utilize in other areas for example home, healthcare, transport and environment etc. Recently with IoT, we have entered in the post-cloud era, where it is expected that there would be huge quantity of information that will be generated by the devices that are immersed in our everyday life routine. A large number of Apps should be developed at the edge to consume this information wisely with a little effort required.

Edge computing represents an initial step in HPC innovation, shifting deep processing power and high-performance data analytics closer to where IoT equipment produces, retrieves, and shares data at high speeds. According to a report by International Data Corporation (IDC) in 2019, 45 percent of IoT generated information would be stored, processed, analyzed and worked upon close to or at the edge of the network [3]. It is expected that around 20 billion devices are connected to the Internet today. Some IoT Apps may need quick response time, some may deal with private information, while some may generate a huge amount of information that could generate heavy traffic for the underlying networks to handle. So, it is noticeable that cloud computing is not capable enough to support these Apps. Performing real-time activities at the edge requires high-end computing technologies that can streamline access to relevant information, reduce expenses and risks associated with data transport across the network, and ultimately enable more cost-effective business decisions.

2. Edge Computing

Edge computing is a new and useful technology that enables data processing at the edge of a network, as an alternative to relying solely on cloud or central data warehouses. It optimizes cloud computing systems by performing data processing activities close to the data originators. Here, "Edge" refers to all computing and network devices or resources along the link between cloud data centers and sources. For instance, a smartphone acts as the edge between the user and the cloud, a gateway in a smart home serves as the edge between home devices and the cloud, and a micro data center or cloudlet functions as the edge between a mobile device and the cloud. The motivation behind edge computing is to bring computing closer to the data originators. In our opinion, edge computing is similar to fog computing, but edge computing focuses more on the devices, while fog computing emphasizes the infrastructure.

Edge computing is a totally new technology which decreases the usage of communication bandwidth required between central data center and sensors by carrying analytics and knowledge generation at or close to the originator of the information [4]. As this new technology enables computation power at the edge of the network for downstream data on behalf of cloud services and upstream data on behalf of IoT services, it requires leveraging resources that might not always be connected to a network, such as smartphones, sensors, laptops, and tablets. Edge computing technology encompasses various fields, including mobile data acquisition, wireless sensor networks (WSNs), mobile signature analysis, and cooperative distributed P2P ad-hoc networking. Additionally, processing power can be divided into cloud/fog computing, grid/mesh computing, mobile edge computing, dew computing, distributed data storage, cloudlets, and retrieval. It also involves autonomic self-healing networks, remote cloud services, augmented reality, and more. It is important to remember that at the edge, the things could not only request services and contents from the cloud, but also carry out computing tasks through the cloud. Edge computing can handle tasks such as computing offloading, caching, processing, and data storage, as well as distributing requests and delivering services from the cloud to users. For these applications to be effective, it is crucial that the edge is carefully designed to meet users' requirements, particularly in terms of privacy protection, reliability, and security.

3. Comparison of Cloud, Fog, and Edge Computing

3.1. Cloud Computing

Cloud computing, often referred to as "the cloud," involves the act of storing, managing, and processing data online. It relies on sharing computing resources rather than using local personal devices or servers to support various programs and applications. The term "cloud" is a metaphor for "the Internet," so cloud computing can be described as "Internet-based computing," where servers, storage, and applications are delivered to organizations' PCs and other electronic devices via the Internet [5]. In simple terms, cloud computing provides services outside an organization's firewall on shared systems. All applications and services are accessed through the web instead of a hard drive. These services are paid for by cloud customers or businesses and are typically managed by the cloud provider rather than by individual cloud customers.

3.2. Fog Computing

Also called fog networking or fogging, it is a decentralized computing infrastructure. Data storage, computing, and applications are distributed in a logical and efficient manner between the data source and the cloud. Fog computing essentially extends cloud computing and services closer to the edge of the network, bringing the benefits and processing power of the cloud nearer to the data source.

The aim of fogging is to enhance competency and decrease the amount of traffic to the cloud for processing, analysis and storage. This is mostly performed to enhance competency, and it can also be utilized for security reasons. The most famous fog computing applications are smart home, smart city, smart grid, vehicle networks and software defined networks. Edge devices and sensors are used to produce and collect data. They do not have the computing and storage capabilities to carryout advance analytic and machine learning tasks [6]. Although servers in cloud have the capabilities to perform these tasks but usually, they are not closer to source to process the data and respond quickly.

The term fog computing and edge computing are sometimes used interchangeably because both entail bringing intelligence and processing power close to the data source. On the other hand, the main difference between these lies in where the intelligence and processing power is placed in the network. In fog computing, intelligence is placed at the Local Area Network (LAN) and information is transmitted from end user to a gateway where it is then transmitted to sources for processing and reply. In edge computing on the other hand, intelligence and power of the edge gateway lie in the devices such as programmable automation controllers.

3.3. Edge Computing

Edge computing is known for data processing power at the edge of the network rather than holding processing power in a cloud or at central data warehouse. Edge computing offers numerous benefits compared to other types of computing. For example, in industrial IoT applications such as power production, smart traffic lights, or manufacturing, edge devices capture streaming data that can be used to prevent part failures, reroute traffic, optimize production, and avoid product faults. Edge devices can include ATM machines at banks, retail stores using beacons to push in-store incentives to mobile applications, smartphones, and gateway devices that accumulate information/data from other end users before forwarding it to the central cloud. When data analysis is carried out at the edge of the network, it is referred to as “edge analytics” [7].

4. Importance of Edge Computing

Edge computing shows the future of computing. In 2011, more smartphones were sold in the world compared to personal computers for the first time in the history. It is expected that smartphones will shortly become the leading computing platform in the world. Recently, mobile phones have achieved computational power and other resources almost on par with laptops, combining storage capacity, outstanding battery life, networking, and a multitude of internal and external sensors into a single platform. Here are some reasons why we need edge computing [4].

4.1. Push from Cloud Services

Shifting computing to the cloud has advantages, as it is a well-organized method for data processing due to the superior computing power of the cloud compared to edge devices. However, when the quantity of data produced at the edge increases, the speed of data exchange can decrease or even jam the cloud-based computing model. For example, a Boeing 787 produces approximately 5GB of data per second, but the available bandwidth between the airplane and either satellites or ground stations is insufficient for managing data transmission. Similarly, an autonomous vehicle generates 1GB of data per second and requires real-time processing to make appropriate decisions. If all this data needs to be sent to the cloud for processing, the response time may be too long. Therefore, it is crucial to process information at the edge to achieve shorter response times, more efficient processing, and reduced network pressure.

4.2. Pull From IoT

Very soon many other electronic devices would become part of IoT and they would act as data producers and consumers. These devices could be anything from an LED bar, air quality sensors, a streetlight or even an internet connected microwave oven. It is expected that close to 2 billion things will be connected to the Internet within the next few years. Hence, design and implementation at the edge of the network will become vital. Raw information generated by these things would be massive and conventional cloud computing approach will not be efficient enough to handle all this information load. It simply means that most information generated by IoT would never be transmitted to the cloud, but it would be consumed at the edge of the network.

Figure 1 demonstrates the cloud computing infrastructure. Raw information is produced by the data producers and sent to the cloud. Data consumer forwards a request for consuming data from cloud. A consumer request is denoted by red dotted line, and the green dotted line shows the result of that request from the cloud. However, this model is not adequate for IoT because the large quantity of data at the edge of the network results in unnecessary bandwidth usage and other computing resource consumption. Another reason this model is inadequate for IoT is that privacy protection requirements pose an obstacle for cloud computing. Finally, many IoT devices have limited energy, and wireless communication consumes more energy compared to wired connections. Therefore, offloading or handing over some computing tasks to the edge of the network can result in more energy-efficient utilization.

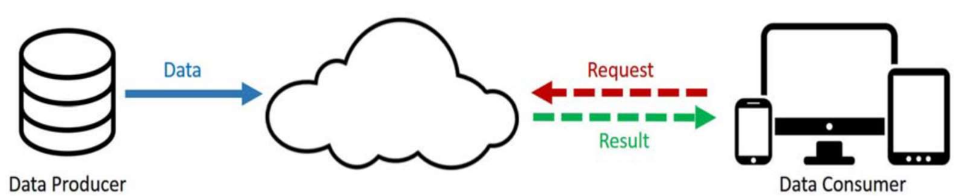

Figure 1: Cloud computing paradigm.

4.3. Change from Data Consumer to Producer

In the cloud computing model, the things at the edge of the network are typically playing the role of data consumers e.g., watching a YouTube video on a mobile phone. The shift from being solely data consumers to becoming data producers and consumers requires many functions to be present at the edge of the network. For example, it is now common for users to take pictures or make videos for various reasons and then share this information through cloud services like Facebook, YouTube, Instagram, or Twitter. Currently, YouTube sees 72 hours of new video content uploaded every minute, Facebook users share about 2.5 million pieces of information per minute, Twitter users tweet approximately 300,000 times per minute, and Instagram users post around 220,000 new photos per minute. However, these images or videos can be quite large, resulting in significant bandwidth usage for uploading. To address this, video clips can be compressed and adjusted to an appropriate resolution at the edge before being uploaded to the cloud, thereby minimizing bandwidth usage. Another example is wearable health devices. The physical information gathered by these devices at the edge is typically private. Processing this data at the edge of the network can maintain user privacy, making it preferable to uploading raw data to the cloud.

Edge Computing Advantages

Edge computing enforces a few restrictions on the selection of particular technology platforms, applications or services, and it is important for every one of these platforms, applications or services to be designed or configured properly for efficiency of edge computing. The following are some of the advantages of edge computing [4].

- Edge computing application services notably reduce the amount of information needed to be moved, the subsequent traffic, and the distance the information needs to travel. Therefore, it reduces network transmissions, their cost, reduce latency, and improve quality of service (QoS).
- This computing approach reduces the reliance on the core computing environment, thereby preventing or eliminating significant traffic congestion and potential points of failure.
- Security is enhanced as encrypted data travels towards the network core. When the information is about to reach the enterprise, it is verified because it passes through guarded firewalls and other security checkpoints, where compromised or fake information, viruses, and malicious hackers can be detected early on.
- The capability of virtualization, which involves logically pooling Central Processing Unit (CPU) resources as needed in real-time, enhances scalability. The edge computing model typically operates on a “charge for network services” basis. It can be argued that the primary consumers of edge computing are organizations seeking to scale business application performance linearly with the growth of their subscriber base.

5. Challenges and Opportunities

The following are some of the challenges, solutions and opportunities which are important for the detailed understanding of edge computing [4].

5.1. Programmability

In cloud computing, users’ programs and their code reside on the cloud. The cloud provider has full control over where the computing is performed within the cloud, and users are generally unaware or only partially aware of how the applications run. This is a significant advantage of cloud computing, as the infrastructure is transparent to all users. Typically, a program is written in a programming language and compiled for a target platform because it runs in the cloud.

In contrast, edge computing offloads computation from the cloud to edge devices, which are often heterogeneous platforms. In edge computing, the runtime of each device or node can vary, making it challenging for programmers to write applications that can be deployed in this model.

To address the programmability of edge computing, the concept of a computing stream has been proposed. This can be described as a series of functions or computations applied to data along its propagation path. These functions or computations can encompass all or some functionalities of an application and can take place anywhere along the link, as long as the application specifies where the computing needs to be performed. This computing stream represents an application-defined computing flow, allowing information to be processed in a distributed and efficient manner across information-producing nodes, edge devices, and the cloud environment.

5.2. Naming

In edge computing, the number of devices is significantly higher. At the top of the edge nodes, many applications are running, each with its own structure regarding how the service is offered. Edge computing requires a similar node naming scheme compared to traditional computer networks, making it vital for addressing, programming, node identification, and information exchange between nodes. Unfortunately, as of today, no efficient naming schemes exist for edge computing. It is crucial for edge computing research to fully understand different communication and network protocols to communicate with the heterogeneous nodes in the model. The naming mechanism in edge computing must handle the mobility of nodes/devices, extremely dynamic network topology, privacy and security protection, and scalability to accommodate the vast number of unreliable devices.

Traditional naming schemes like DNS and Uniform Resource Identifiers work well for current networks but are not flexible enough to support dynamic edge networks. Many edge devices can be extremely mobile and have resource limitations. Due to these limitations, IP-based naming methods can be too complex and resource-intensive for edge computing.

New naming schemes like Named Data Networking (NDN) and MobilityFirst can also be implemented in edge computing. NDN provides a hierarchical naming structure for content/data-centric networks, which is user-friendly for service management and offers excellent scalability for edge computing. However, it requires additional proxies to integrate with other communication protocols like ZigBee or Bluetooth. Another issue with these naming schemes is security, as it is challenging to separate hardware information from service providers.

MobilityFirst can isolate names from network addresses to achieve superior mobility support, making it suitable for edge services where devices are highly mobile. However, MobilityFirst requires the use of Global Unique Identifiers (GUIDs) for naming, which is unnecessary for fixed information aggregation services at the edge, such as in home environments. Another limitation of MobilityFirst for edge computing is the complexity in service management, as GUIDs are not user-friendly.

5.3. Data Abstraction

Data abstraction is a big problem in edge computing as compared to cloud computing. For example, in a smart home environment within the IoT, large amounts of data are generated, and nearly all devices report data to the edge operating system (edgeOS). In edge computing, many devices periodically report sensed data to the gateway. For instance, a thermometer might report the temperature every minute, but this data is only used by the end user a few times per day. Similarly, a home CCTV system used for security reasons forwards information to the gateway, but this information is only stored in the database temporarily before being overwritten by new recordings.

Based on this observation, human involvement in edge computing is minimized, with edge nodes consuming and processing all the data and interacting with users proactively. In this scenario, information needs to be preprocessed at the gateway level, including tasks such as noise and low-quality data removal, event detection, and privacy protection. Processed information is then forwarded to the upper layer for future service provision. There are many challenges in this process.

First, the information arriving from many nodes comes in different formats, as shown in Figure 2. For privacy and security concerns, programs running on the gateway need to be shielded from raw information. Furthermore, they should extract the information they are interested in from an integrated data table. The table can be easily described with time, name, ID, and data (e.g., {1111, 11:35:57 PM

02/03/2017, kitchen.oven4.temperature5, 89}), allowing every edge node's information to be fitted in. However, the details of sensed information may be obscured, which might affect its utilization.

Secondly, it is challenging to decide the degree of data abstraction. If a large amount of raw information is filtered out, some services or programs may not gain sufficient knowledge. On the other hand, retaining a huge amount of raw information poses a challenge for information storage.

Finally, information reported by nodes at the edge can sometimes be unreliable due to low-accuracy sensors, hazardous environments, and unreliable wireless connections. In this scenario, abstracting valuable data from unreliable data sources remains a challenge for IoT application and system developers.

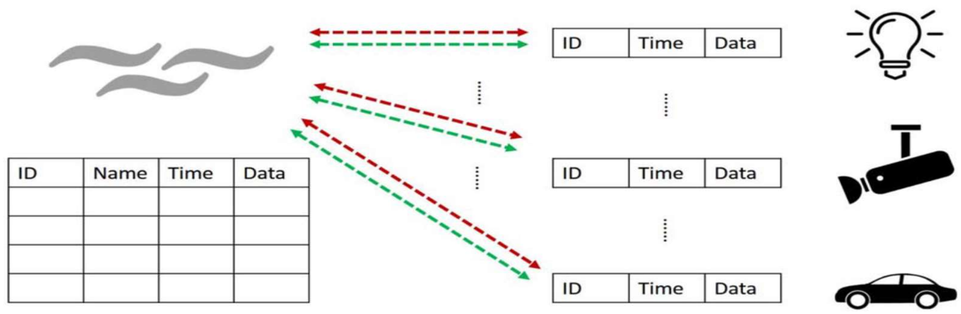

Figure 2: Data abstraction issue for edge computing

5.4. Service Management

The following four fundamental features could be maintained to guarantee a reliable service management at the edge of the network.

5.4.1. Differentiation

The expansion of IoT consumption is increasing rapidly, and it is expected that multiple services will be implemented at the edge of the network, such as in smart homes. All these services will have varying priorities. For example, critical services like device diagnostics and failure alarms should be processed as soon as possible, ahead of regular services. Health-related services, such as fall detection or heart failure detection, must also have a higher priority compared to routine services like entertainment.

5.4.2. Extensibility

Extensibility is a significant issue at the edge of the network. Unlike a mobile system, IoT devices can be exceptionally dynamic. When an owner buys a new device, can it be seamlessly integrated into the existing service? Or when a device is replaced due to wear and tear, can the old service easily accommodate the new node? These issues can be resolved with a flexible and extensible design of the service management layer in the edge operating system.

5.4.3. Isolation

Isolation is also a big issue at the edge of the network and in the mobile OS. If an application stops responding or collapses, then the complete system would normally collapse and reboot. In distributed systems, shared resources can be managed using various synchronization techniques, such as locking or token ring. However, this problem may be more complicated in a smart edgeOS. There can be many applications that share common data resources, such as the control of lights. If one application is not responding, a user should still be able to manage their lights without causing the entire edgeOS to collapse. Similarly, if a user deletes the application that controls the lights, the lights should continue to function rather than experiencing a lost connection to the edgeOS. This issue can potentially be solved through a deployment/undeployment framework. If the OS can detect the issue before an application is installed, it can warn the user and help avoid potential access problems.

5.4.4. Reliability

Reliability is also vital issue at the edge of the network. It can be discussed from various perspective. In the context of service, it is usually difficult to determine the cause of a service failure. For example, if an air conditioner is not working, the potential reasons could be a cut power cord or a compressor failure. In the context of the system, it is crucial for the edgeOS to maintain the network topology of the entire system, ensuring that every component can forward status and diagnostic information to the edgeOS. Finally, in the context of data, reliability issues often arise from the information sensing and communication processes.

5.4.5. Privacy and Security

Privacy and security of information is the key service at the edge of the network. In a home equipped with IoT devices, many private details can be inferred from sensed usage information. For example, gas or electricity usage readings could reveal whether the home is occupied. Providing services without compromising privacy is a challenge. Some private information can be removed from data before processing, such as masking entire faces in a video. Keeping computing at the edge, which is beneficial in the home, can be an effective approach to protect privacy and data security.

6. Edge Computing Applications

The following are five potential applications of the edge computing [8].

6.1. Cloud Offloading

In cloud computing environment, data and the requests are processed in the central cloud which means that most of the computation occurs in the cloud. This computation at the central cloud might suffer from longer latency. In the edge computing, however, things at the edge contain computation capabilities and offers an opportunity to offload part of the workload from central cloud.

6.2. Video Analytics

The widespread use of smartphones and network cameras makes video analytics a promising technology. Cloud computing is not suitable for applications requiring video analytics due to information transmission latency and privacy issues. In contrast, video analytics in edge computing is more suitable. For example, in the case of searching for a missing child using the edge computing model, the request for the missing child could be generated in the cloud and pushed to all devices in a target region. Each device, such as a smartphone, could execute the request, search its local camera information, and only report the results back to the central cloud.

6.3. Smart Home

IoT can greatly enhance the home environment. Many smart devices, such as smart lights, smart TVs, and robot vacuums, are already available on the market. However, simply adding a Wi-Fi unit to existing electronic devices and connecting them to the cloud is not sufficient for a smart home. In addition to connected devices, less expensive wireless sensors and controllers can be installed in rooms, pipes, walls, and floors. These devices will generate a significant amount of information, and to manage the data transmission load and protect privacy, this information would typically be processed within the home. This characteristic makes the cloud computing model unsuitable for a smart home. However, the edge computing paradigm is a perfect choice. With an edge gateway running a specific edgeOS in the home, devices can be connected and managed effortlessly. Information can be processed locally to reduce the load on Internet bandwidth, and services can be installed on the edge operating system for efficient management and delivery.

6.4. Smart City

Edge computing model is flexible and can be easily extended from a home to a city. Edge computing makes it possible to process some data locally close to the data source. With this model, a request can be created from the top of the computing model and could be processed at the edge. Hence, edge computing can be a good choice for smart city management.

6.5. Collaborative Edge

The information created by stakeholders are hardly ever shared with other devices because of privacy issues and the frightening cost of information exchange in some cases. Therefore, the choice of collaboration among many stakeholders is restricted. An edge is a physically small data center that connects to the cloud and end users, with information processing capabilities as part of the logical concept. The idea of a collaborative edge, which connects the edges of many stakeholders that are geographically dispersed regardless of their physical location and network configuration, is proposed.

7. Conclusion

This paper has explored the transformative potential of edge computing within the context of the Internet of Things. By comparing cloud, fog, and edge computing, we have highlighted the unique advantages of edge computing, such as reduced latency, enhanced security, and improved efficiency. The paper has also addressed the challenges and opportunities associated with edge computing, including programmability, naming, data abstraction, and service management. Through detailed analysis, we have demonstrated how edge computing can effectively support various applications, from smart homes and cities to collaborative environments. Ultimately, this paper underscores the critical role of edge computing in advancing IoT technologies and shaping the future of data processing.

References

[1] M. Armbrust et al., “A view of cloud computing,” Commun. ACM, vol. 53, no. 4, pp. 50–58, 2010.

[2] S. Ghemawat, H. Gobioff, and S.-T. Leung, “The Google file system,” ACM SIGOPS Oper. Syst. Rev., vol. 37, no. 5, pp. 29–43, 2003.

[3] D. Evans, “The Internet of Things: How the next evolution of the Internet is changing everything,” CISCO White Paper, vol. 1, pp. 1–11, 2011.

[4] Y. Li and L. Xu, “Edge Computing: Vision and Challenges”, IEEE INTERNET OF THINGS JOURNAL, VOL. 3, NO. 5, OCTOBER 2016

[5] V. Beal, “Cloud Computing”, http://www.webopedia.com/TERM/C/cloud_computing.html

[6] M. Rouse, “fog computing (fog networking, fogging)”, http://internetofthingsagenda.techtarget.com/definition/fog-computing-fogging

[7] C. Raphael, “What's the difference between edge computing and cloud computing?”, https://www.quora.com/Whats-the-difference-between-edge-computing-and-cloud-computing,

[8] S. Yi, Z. Hao, Z. Qin and Q. Li, “Fog Computing: Platform and Applications”, Third IEEE Workshop on Hot Topics in Web Systems and Technologies, 2015