Introduction to edge computing

Chapter · September 2020

DOI: 10.1049/PBPC033E_ch1

4 authors:

Auday Al-Dulaimy
Mälardalen University

Yogesh Sharma
Thompson Rivers University

Michel Gokan Khan
Karlstads Universitet

Javid Taheri
Karlstads Universitet

Chapter 1
Introduction to edge computing

Auday Al-Dulaimy[1], Yogesh Sharma[1], Michel Gokan Khan[1] and Javid Taheri[1]

Edge computing is the model that extends cloud computing services to the edge of the network. This model aims to move decision-making operations as close as possible to data sources since it acts as an intermediate layer connecting cloud data centres to edge devices/sensors. Transferring all the data from the network edge to the cloud data centres for processing may create a latency problem and outstrip the network's bandwidth capacity. To resolve this issue, it might be best to process data closer to the devices/sensors. This chapter will take a deep dive into edge computing, its applications, and the existing challenges related to this model.

1.1 Evolution of computing systems

Computing has come a long way from where it started. At the very beginning, a computer could only perform one task at a time. Several distributed computers had to run in parallel when performing multiple tasks, and distributed systems were formed by connecting those computers, which usually communicated and coordinated their actions through message exchange. Then, personal computers and multitasking operating systems emerged and made it possible to run multiple tasks on the same computer. This enabled the systems' developers to build and run an entire system within one or more connected computers. As the price of computing power and storage fell, organisations all over the world started using distributed systems. The breakthrough time of distributed systems came when Internet-based companies became so large that they needed to build distributed systems that spanned across the World as data centres. The standardisation of the concept of distributed computing led to the development of other models, including cluster computing, autonomic computing, utility computing, and grid computing. Engineers and developers then started to think about a way to create multiple virtual computers within the same machine. This led to the concept of virtualisation by which the same computer could act as multiple computers all running at the same time. On top of these predecessor models, the

[1]Department of Mathematics and Computer Science, Karlstad University, Karlstad, Sweden

construction of the cloud computing model was a natural step forward from grid computing [1]. Cloud computing leveraged the existing models and used the virtualisation concept to provide users with new services including new features and characteristics [2].

That was a good idea, but it was not the best option when it came to the utilisation of the resource of the host computer. Running multiple operating systems required using additional resources that were not needed when running one operating system. This led to the invention of containers in which multiple programs required separate runtime when using the same host operating system kernel. Containers act as virtual machines without the need for a separate overhead operating system. Due to its innovative characteristics and services, the cloud computing model has captured significant attention among individual users, academia, industries, and even governments. Cloud computing services are offered by providing access to a wide range of infrastructures hosted on cloud data centres [3]. Despite the increasing usage of cloud computing, its inherent issues remain unresolved. These include unreliable latency, lack of mobility support, location-awareness, as well as privacy and security issues. Other models have emerged in an attempt to solve these problems, but edge computing is still the most promising state-of-the-art computing system that continues to evolve.

1.2 An overview of edge computing

The idea of having computational resources near the data sources might not be something new [4]. The term 'edge computing' first appeared in 2002, to state that from a business perspective, applications should be served and moved from the cloud data centre to the network edge [5]. The term was then used in 2004 to describe a system that distributed program methods and the corresponding data to the edge of the network in order to enhance the system's performance [6]. Edge computing continued to advance until it was able to resolve many of the problems associated with cloud computing, as it provided elastic resources to end users at the edge of the network, at a time when cloud computing could only provide resources distributed and hosted on cloud data centres in the core network.

This section defines edge computing, discusses its concept, introduces representative application scenarios, and identifies various aspects of issues that may arise when designing and implementing edge computing systems. It also highlights some opportunities and challenges and serves as a guidance for potential future work in related techniques.

1.2.1 Motivation

As stated in Section 1.1, computing systems evolved from using one computer to perform one or more tasks to providing centralised and consolidated services and applications in the cloud data centres that emerged in the last decade. Recent technological developments such as powerful domestic connection boxes, high-capacity mobile end-user devices, wireless networks, and the increasing user concerns about reliability, privacy, and autonomy all call for controlling computing applications, data, and services from the logical extreme (the 'edge') of the Internet rather than from the central nodes (the 'core') [7]. The centralised nature of remote computing resources hinders their ability to handle huge data traffic originating from geographically dispersed edge devices. Thus, it is necessary to push the servers to the edge of the network, and it is actually an unavoidable trend as shifting towards (the 'edge') becomes indispensable. Several factors make edge computing an essential necessity, and such factors must be examined independently from the end users' and operators' perspective. The following requirements propel computing to move to the edge and they explain why edge computing is an essential model [8,9].

Push from cloud services: Placing all computing tasks on the cloud has proved to be an effective method for data processing, because the computing power on the cloud surpasses the capacity of the things on the edge. However, the network suffers a deadlock compared to the speed of the rapidly developing data processing. As the amount of data generated at the edge continues to increase, the speed of data transportation impedes the traffic flow in the cloud-based computing model. Two examples can be given here: The first example, in [10], states that a Boeing 787 generates about 5 gigabytes of data every second, but the bandwidth between the aircraft and either the satellite or the base station on the ground is insufficient for data transmission. The second example in [11] states that a vehicle generates 1 gigabyte of data every second, and that the vehicle needs real-time processing in order to make correct decisions. Sending all data to the cloud for processing would ineffectively prolong the response time. It would also pose a challenge on the reliability of the current network bandwidth and its ability to support a large number of vehicles in one area. Thus, processing data at the edge ensures less response time and network pressure as well as better processing.

Pull from Internet of things (IoT): In the near future, most electric devices will be included in the IoT, and they will play the roles of both data producers and consumers alike, such as air quality sensors, light-emitting diode (LED) bars, streetlights, and even an Internet-connected microwave oven [8]. We can conclude that, in a few years, the number of things on the edge will exceed billions. Thus, they will produce massive raw data that the standard cloud computing model will not be able to handle. This leads to the fact that most IoT data will not make its way to the cloud, and will rather be consumed at the network edge. The classical structure of the cloud computing model is inadequate for IoT for many reasons: First, the amount of data at the edge is too large and it utilises a lot of unnecessary bandwidth and computing resources. Second, the requirement for privacy protection hinders cloud computing in IoT. Third, energy constraints restrict most of the IoT end nodes, and the wireless communication module usually consumes a lot of energy, so it might be more energy efficient to discharge some computing tasks to the edge.

Change from data consumer to producer: The front-end devices at the edge usually act as a data consumer in the cloud computing model, such as watching a YouTube video on your smart phone, but nowadays people are also generating data via their mobile devices. Switching from data consumer to data producer/consumer requires more peripheral placement. For instance, people nowadays normally take pictures or record videos and then share the data via a cloud service such as Twitter, Snapchat, and Instagram. Twitter users sent around 473,400 tweets per minute in 2018; Snapchat users shared 2,038,333 snaps; and Instagram users posted 49,380 new photos in that year [12]. However, uploading a large image or a video clip would occupy a lot of bandwidth, and this requires adjusting the video clip to an appropriate resolution before uploading it on the cloud. Wearable health devices can also be an example. As physical data collected by the things at the edge of the network are often privately owned, users privacy could be better protected by processing data at the edge rather than uploading them on the cloud.

Decentralised cloud and low-latency computing: Centralised cloud computing is not always the ideal strategy for geographically distributed applications. To enhance the service provided, the computing must be done closer to the data source. For any web-based application, this benefit can be generalised [13]. Recently, users have become more interested in using location-aware applications, and this entices application providers to improve their service by computing on edge nodes that are closer to users. Edge devices generate many data streams, and performing the analytics on a distant cloud obstructs real-time decision-making. Using current cloud infrastructures in real-time applications causes serious latency issues between an edge device and the cloud, and any potential delay would not be in line with the requirements of latency-sensitive applications. Video streaming generates most of the mobile traffic, and it is still difficult to maintain users' satisfaction when the majority of the network traffic flows to the same source. In the same way, gamers face similar latency obstacles when using multimedia applications [9]. In this case, employing edge nodes closer to users can minimise network latency to enhance the computations performed on the cloud.

Surmounting resource limitations of front-end devices: User devices have somewhat restrained hardware resources when compared to compute and/or storage resources of the servers hosted within a cloud data centre. Front-end devices can be mainly divided into two types [14]: devices carried by people and others placed in the environment. The main role of the two front-end device types is to obtain real-time data by capturing sensory input in the form of text, audio, video, touch, or motion, and then cloud services process the transferred data. These devices, however, are unable to perform complex analytics due to their hardware limitations. Therefore, data often needs to be transmitted to the cloud to meet computational processing requirements and to send back useful information to the front-end device. However, the service does not have to use all the data it receives from the front-end device in order to build analytical cloud workloads. Data can possibly be filtered or even analysed at edge nodes, which may have spare computer resources for managing data [9].

Sustainable energy consumption: Numerous energy-consuming data centres have been recently established worldwide. However, the financial and environmental impact of the high energy consumption at these data centres is a major issue that needs to be tackled, and enhancing the energy efficiency of such data centres is a main challenge in the cloud computing model. As the number of applications moving on to the cloud is increasing, the growing energy demand may become unsustainable.

For this purpose, it is highly important to develop energy-efficiency strategies and to propose new approaches to enhance energy efficiency in cloud data centre [3,15]. To some extent, integration of practical power management strategies could contribute to a reduction in energy consumption. Some analytical tasks could be done on edge nodes, such as base stations or routers, that are closer to the data source, rather than overloading data centres with minor tasks that could be carried out at edge notes with minimal energy consumption [9].

Dealing with data explosion and network traffic: Cloud providers expand their infrastructure to include IoT devices/sensors at the edge of the network, so as to meet the rapidly growing demands for cloud services, and the number of these devices/sensors has significantly increased. Cisco Internet Business Solutions Group predicts that 50 billion devices will be connected to the Internet by 2020 [16], and according to the International Data Corporation (IDC), approximately 80 billion devices will be connected to the internet by 2025 [17]. Data generating from those IoT devices/sensors continues to increase rapidly. The work in [9] states that 43 trillion gigabytes of data are expected to be produced by 2020. The architecture of the cloud computing model is not designed to work with the volume, variety, and speed of data generated by edge devices/sensors. Thus, data centres have to expand to carry out monitoring tasks and analytical workloads. However, this raises further concerns about sustainable energy consumption at data centres, because resource limitations on the edge devices make it difficult to perform analytics on the edge device. Collective analytics (of multi-edge devices) cannot be performed on edge devices either. The volume of network traffic to and from clouds data centres is another concern with increasing data generation, as it causes delays edge devices and depletes the network bandwidth capacity. In this case, using very close nodes in the network to complement the device computations or those of the data centre would solve the issue of handling data growth and the distribution of traffic in a network.

Smart computation techniques: In order to reduce latency and minimise the impact of energy consumption of data generated by end users to and from cloud data centres, there is a need to adopt a collaborative approach in the manipulation of network resources and the distribution of computation among these sources. The work in [9] identifies two cases of computation distribution. In the first case, the generated data are filtered on the front-end device by the application pipeline, and then edge nodes carry out the workload analysis, while the more complex tasks are done at the cloud nodes. The second case allows data centres to offload computations requiring limited resources on to edge nodes, or it allows the edge nodes to use volunteer devices to enhance computational capacities. Edge nodes have the ability to facilitate computations closer to the source, and they can also incorporate strategies to remotely enhance front-end device capabilities.

1.2.2 Definition

Many definitions for edge computing have been proposed over the past years. However, no standard definition exists. The most widely adopted definitions are as follows.

As defined by the authors in [8] and [18], '*Edge Computing refers to the enabling technologies allowing computation to be performed at the edge of the network, on downstream data on behalf of cloud services and upstream data on behalf of IoT services*'. In [19], '*Edge Computing is the model that pushes localised processing in the advanced manner, i.e., closer to the data source*'. As defined in [20], '*Edge Computing is an umbrella term covering the latest trend of bringing the computational resources to the proximity of the end devices*'. The author in [21] stated that '*Edge Computing is moving processing close to where data is being generated*'. Another work [22] defined edge computing as follows: '*Edge Computing is a new paradigm in which substantial computing and storage resources (variously referred to as cloudlets, micro data centers, or fog nodes) are placed at the Internet's edge in close proximity to mobile devices or sensors.*' In [23], '*Edge Computing refers to applications, services, and processing performed outside of a central data center and closer to end users. The definition of "closer" falls along a spectrum and depends highly on networking technologies used, the application characteristics, and the desired end user experience*'.

Edge computing can, therefore, be defined as the model that optimises cloud computing systems by processing data close to its source at the edge of the network. It enables technologies to place computing/storage resources at a close proximity to the data source, mainly to the edge of the network.

1.2.3 Architecture

The philosophy of edge computing is primarily based on performing computation at the edge that is in the proximity of a data source. Edge computing resources can be a network or a computing resource run among end users on the one side, and fog nodes and cloud data centres on the other. The interconnected sensors/devices at the IoT layer generate and transfer data among each other using a modern communication network infrastructure. Generated data are processed at other layers determined by specific application requirements. Edge computing architecture can be explained in four layers: IoT layer, edge layer, fog layer, and cloud layer. Figure 1.1 illustrates the architecture of edge computing.

In this architecture, IoT layer includes millions of devices/sensors which constantly produce data, exchange important information amongst themselves through a modern communication network infrastructure, and monitor and control critical smart-world infrastructures. All these IoT devices/sensors are end users for edge computing.

IoT and edge computing are rapidly evolving in an independent manner. However, the edge computing platform can assist IoT in resolving a number of key issues and improving performance. In general, IoT devices/sensors can benefit from the high computational capacity and large storage of the edge, fog, and cloud computing layers. However, edge computing has further advantages over fog and cloud computing for IoT, even though its occupational capacity and storage are more limited. IoT specifically requires fast response time rather than high computational capacity and large storage. Edge computing offers a tolerable computational capacity, enough

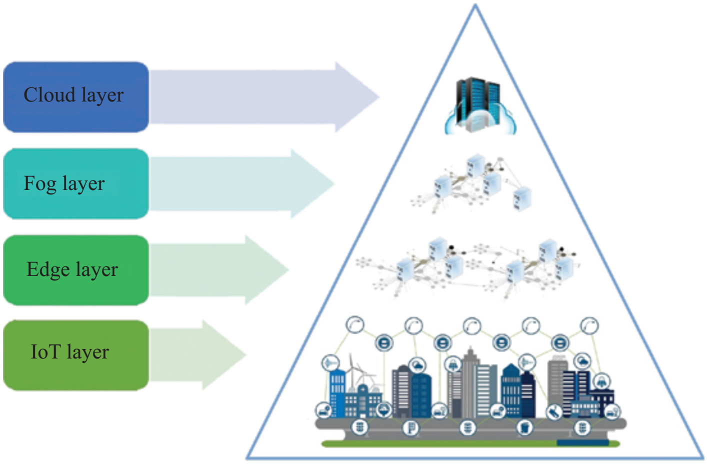

Figure 1.1 Edge computing architecture

storage space, and fast response time to satisfy IoT application requirements. Edge computing can benefit from IoT by extending the edge computing structure so that it can handle the distributed edge computing nodes. IoT devices can be used as edge nodes to provide services. As the number of IoT devices continues to grow, IoT and edge computing are likely to become inseparable. Edge computing can also benefit from the fog computing layer, as the resources capabilities on the fog layer are relatively stronger than those on the edge layer.

In the literature, edge computing and fog computing are sometimes used interchangeably, or considered as one layer. However, edge computing focuses more on the side of the things, while fog computing focuses on the side of infrastructure. The cloud computing layer consists of multiple high-performance servers and storage devices. Among all layers, the cloud computing layer has the highest computation power and storage capacity.

1.2.4 Characteristics

In general, IoT and edge computing have similar characteristics with regard to the IoT layer. On the other hand, edge computing has a lot of common characteristics with fog computing.

The prime objectives of edge computing and fog computing are similar. Both models bring cloud computing capabilities to the edge of the network. They enable the computation and storage capacities located within a close proximity to end users to reduce service latency and to save network bandwidth for delay-sensitive applications [24]. Both edge and fog computing models have distributed, hierarchical, and decentralised architectures which are different from the centralised cloud computing architecture. Their service locations are also at a close proximity to end users. Edge computing is located on edge devices, while fog computing takes place in network

Table 1.1 Characteristics of IoT, edge, fog, and cloud computing

Characteristic	IoT	Edge	Fog	Cloud
Deployment	Distributed	Distributed	Distributed	Centralised
Components	Physical devices	Edge nodes	Fog nodes	Virtual resources
Location awareness	Aware	Aware	Aware	Not aware
Computational	Limited	Limited	Limited	Unlimited
Storage	Very limited	Limited	Limited	Unlimited
Data	Source	Process	Process	Process
Distance to data source	The source	The nearest	Near	Far
Response time	No response time	The fastest	Fast	Slow
Nodes count	The largest	Very large	Large	Small

edge devices that are one or a few network hops away from the edge. When compared to cloud computing resources, those of the edge and fog computing models (e.g. computing, communication, and storage resources) are quite limited. Network edge devices also have a relatively higher resource and service capability than those of edge devices.

In some works, the edge and fog computing models are considered to be one layer. However, there are still some fundamental differences between the two, even if they share the same goals [25]. Edge devices are unable to implement several IoT applications in the edge computing layer, as the limited sources would cause resource contention and would also increase processing latency. In contrast, fog computing can successfully overcome these limitations and prevent resource contention at the edge by smoothly incorporating edge devices with cloud resources. Fog computing coordinates, balances, and improves the utilisation of geographically distributed network edge devices, and it exploits the cloud resources for that purpose. In addition, edge computing is focused on the things, while fog computing is more focused on the infrastructure. Table 1.1 further demonstrates a comparison of the characteristics of all layers.

1.3 Applications

Edge computing has many applications, as described in the following subsections [8,26].

1.3.1 Smart systems

In smart systems, which can be called '***SmartX***', network communication technologies are interlinked with sensors and actuators that are meant to send signals to the system to take action. Smart systems are viewed as an extension of IoT technologies. The integration of sensors offers countless opportunities for data collection, physical system management, as well as allocation and optimisation of resources. The utilisation of smart systems enables edge computing to reduce latency to the lowest level possible and to increase storage capacity in devices lacking computational efficiency. In the same way, edge data analysis can ensure a high degree of resilience in compromised systems. Smart systems have several significant areas, including smart homes, smart grids, smart cities, smart farms, smart transportation, smart healthcare, and others.

Smart home: IoT can be very beneficial in a home environment. Some products, such as smart lights, smart TVs, and robot vacuums, have already been developed and sold in the market. However, smart homes are not created by just adding a Wi-Fi module to the current electrical device and connecting it to the cloud. In addition to the connected device, cheap wireless sensors and controllers must be deployed in the rooms, pipes, floors, and even walls of a smart home. Such devices would report a substantial amount of data, which should primarily be consumed in the smart home if we take into account data transportation pressure and privacy protection issues. Regardless, edge computing is considered to be a perfect model for smart home establishment. Things can be easily connected and controlled in the smart home by using an edge gateway that runs a specialised edge operating system (EdgeOS). The generated data can be locally processed to reduce Internet bandwidth usage, and the EdgeOS also offers a better service in terms of management and delivery [27].

Smart city: The edge computing model can feasibly extend from a single home to an entire community or even a city. In this model, computing should be done as close as possible to the data source. This design allows the generation of a request at the top of the computing paradigm and processing it at the edge. A smart city utilises distributed IoT devices supported by various sensors. These devices improve transportation and traffic management, air quality monitoring (e.g. pollution, temperature, and humidity), smart parking, smart lighting, and smart garden irrigation systems [28]. Edge computing can be a perfect platform for a smart city due to the following characteristics [8]:

- Large data quantity: Due to the traffic overload, it is not possible to establish centralised cloud data centres that can handle vast amounts of data (generated from public safety, health, utility, transportation systems, etc.). In this case, edge computing may become an effective solution, as data would be processed at the edge of the network.
- Low latency: Edge computing also works well with applications that require low, predictable latency, such as those used for healthcare emergency or public safety purposes, as this model can save data transmission time and facilitate the network structure. It is more efficient to make decisions and diagnostics on the edge of the network than to collect information and make decisions at a central cloud.
- Location awareness: Given its location awareness, edge computing works better than cloud computing when used for geographic-based applications, such as transportation and utility management applications. In this model, data can be collected and processed in its geographical location without being transported to the cloud.

Smart healthcare: Healthcare systems would extremely benefit from IoT. In general, cloud computing frees IoT devices/sensors battery-draining computing tasks and provides virtually unlimited sources. Large-scale data sets can be collected from various sensors to carry out the analytical tasks mentioned above. However, such a simple sensor-to-cloud architecture is not feasible in many applications operating within healthcare information systems. Regulations sometimes prevent the storage of patients' data outside the hospital premises. Moreover, some applications cannot entirely rely on remote data centres, as network and data centre failure would jeopardise patient safety. Edge computing is one possible solution to narrow the gap between sensors and analytics in healthcare information systems.

Healthcare has recently undergone some of its most profound transformations, and the introduction of wireless sensor technology was one of the main factors behind this change. In addition to providing further access to biometric parameters, sensors are getting smaller in size and can be worn without interfering with everyday life. This is essential for ongoing data collection. Such technologies allow patients to use many sensors and fitness trackers, and in the future, these sensors will serve as a tool to monitor the health condition of any human being regardless of their health status.

Sensory data can be more useful in healthcare systems when adopting other driving factors, such as big data analysis and machine learning. In addition to providing computer-assisted analysis of medical images, large data analysis can be used to investigate treatment effectiveness, identify patients who are at risk of developing chronic diseases, ensure patients' adherence to treatment plans, optimise processes, and develop personalised care plans. At this point, sensors used to monitor patients must be wireless and wearable. However, this limits the sensors' size and affects their energy quality memory and processing capacity. Moreover, data can only be valued in context, and it must be accumulated from various sensors. Thus, sensors send data to other more capable computing devices for accumulation, storage, and analysis. However, in a broader environment, it is not possible to adopt such vertical approaches. Their specific infrastructure is expensive and difficult to maintain when several patients need to be equipped with many sensors individually. In this case, the IoT offers an alternative approach. Using a common infrastructure allows sensor devices to transmit their data to more comprehensive applications. This added flexibility in computing offers new opportunities to meeting current healthcare challenges. Improving mobility and the integration of patients will ensure ongoing monitoring, as mentioned above, and will enable the utilisation of new applications. In many countries, healthcare systems face great challenges because chronic diseases increase with age.

There is also a growing shortage of nursing staff in many countries. At the same time, there is a high demand for cost reduction while maintaining high-quality patient care. As a result, the healthcare industry is promoting a model that focuses on information delivery. A part of this model that delivers medical services enables remote monitoring of patients, as this increases accessibility to the patients and improves healthcare quality, efficiency, and sustainability. It also reduces the overall healthcare cost. Today, the manual measuring of biometric parameters and transfer of data between systems in hospitals is extremely time-consuming. Remote monitoring would be time saving for caregivers, and automated monitoring would replace manual monitoring. Processes taking place within the hospital would also improve, as remote monitoring would enable effective utilisation of resources. Many processes are planned manually, and therefore done in a sequence, rather than using resources effectively. Moreover, sensors would make it easier to obtain correct information about current health conditions, equipment location, caregivers, and patients. Sensors can also provide a more accurate image of patients by continuously capturing data and providing insights into a wide range of biometric parameters. Medical treatment and diagnostics will be revolutionised. Data would not be processed in isolated silos; it would be combined with other sources and viewed within context. The healthcare industry is moving towards preventive medicine. Reactive healthcare involves treating patients only when they face health issues, while preventive care monitors individuals to keep them healthy and out of hospitals. Moreover, patients can be discharged from hospitals earlier and monitored at home. It primarily means removing boundaries between hospitals, homes, and any other healthcare facilities, so healthcare becomes an ongoing process [29].

Smart grid: When monitoring devices are deployed, the smart grid collects huge amounts of energy-related data at an unprecedented speed. The smart grid becomes data-driven, and this requires extracting important data from a large data set. The traditional data extraction approach enhances computing efficiency in a temporal dimension, but it only works for one task in the smart grid. Furthermore, current solutions do not take into account the geographical distribution of computing capacity within a large-scale smart grid [30].

In the field of grid technology and implementation, the smart grid is considered to be the next generation. To reap the benefits of the smart grid (e.g. safety, security, and self-healing), many smart meters, sensors, and actuators must be used to collect and share measured data in the smart grid. Thus, edge computing can potentially meet the requirements of the smart grid deployment.

Smart transportation: The deployment of a cloud-based vehicle control system is necessary to ensure safety and efficiency in self-driving vehicles because it can collect data from sensors via a vehicle-to-vehicle network. This system can control and coordinate a large number of vehicles. A real-time vehicle management system certainly entails strict requirements, such as low latency, which can be provided by edge computing [26].

Smart surveillance: Smart surveillance systems monitor, screen, and track activities to provide security in public places like airports. Several aspects, such as the screening of objects and people, maintaining a database of potential threats, biometric identification, and video surveillance, work in tandem to monitor activities. The scope of the surveillance applications' work includes the control of access to specific areas, human identification, and event detection. This is added to monitoring traffic, hospitalised patients, and activities in public places, including shopping malls, industries, banks, and government institutions. Any surveillance system should be able to compress large video recordings effectively and thus facilitate subsequent processing [31]. Smart surveillance applications that are based on cloud computing face major challenges. These applications require real time for object detection and tracking by processing video streams collected from broadly distributed data sources, such as networked cameras and smart mobile devices. However, the transfer of large amounts of raw data to cloud centres has an uncertain time frame and it overloads communication networks. Moreover, the remote transfer of data can jeopardise data protection and privacy, as it would be more exposed to attackers. Surveillance video streams are often viewed as a tool for forensic analysis, rather than as a proactive tool to help thwart suspicious activities before the damage is actually done. Therefore, new technologies are needed to handle critical and security-sensitive tasks locally. IoT technology is paving the way for a new post-cloud era. Thousands of the smart 'things' we use in our daily life generate and process a lot of data on the edge of the network [32]. However, edge computing can provide the real-time management that surveillance systems require.

Smart farm: To meet future demands of food production, the agricultural sector has to integrate IoT in various production, management, and analysis processes. Smart farms can use edge computing to operate autonomous vehicles (tractors) and to perform remote monitoring and real-time analytics. IoT devices/sensors can provide data on crop production, rainfall, pest infection, and soil nutrition. Such information is helpful for production and can also help improve farming techniques in the long run [33].

1.3.2 Video analytics

The rapid spread of mobile phones and network cameras is giving rise to video analytics technology. Cloud computing does not work for applications requiring video analytics, as it entails data transmission latency and privacy issues. An example of this is finding a lost child in a city. Today, various types of cameras are deployed in urban areas and vehicles. A camera somewhere might capture an image of the child. However, the camera data might not be uploaded on the cloud for privacy issues or traffic cost. This makes it extremely hard to leverage the data of cameras in a vast area. The data might already be available on the cloud, but it would take too long to upload and analyse large amounts of data to find the missing child. In this case, the edge computing model can be used to generate a search request for the child from the cloud and transmit it to all the things in the target area. For example, a smartphone can receive the request, search its local camera data, and only report the results to the cloud. This model makes it possible to leverage data and computing power to get faster results than those obtained through the cloud computing model [8].

1.3.3 Collaborative edge

Cloud computing has certainly become a platform for processing large amounts of data in academics and industries. In cloud computing, data should be held or transmitted to the cloud for processing. However, stakeholders rarely share data due to privacy issues and the very high cost of data transfer, and thus, collaboration among several stakeholders can be limited. The edge may be part of the logical concept, as it is a small physical data centre with data processing capabilities that connects end users to the cloud. It proposes a collaborative edge that connects the edges of many geographically distributed stakeholders regardless of their location and network structure [34]. These instantly connected edges give stakeholders the chance to share and coordinate data.

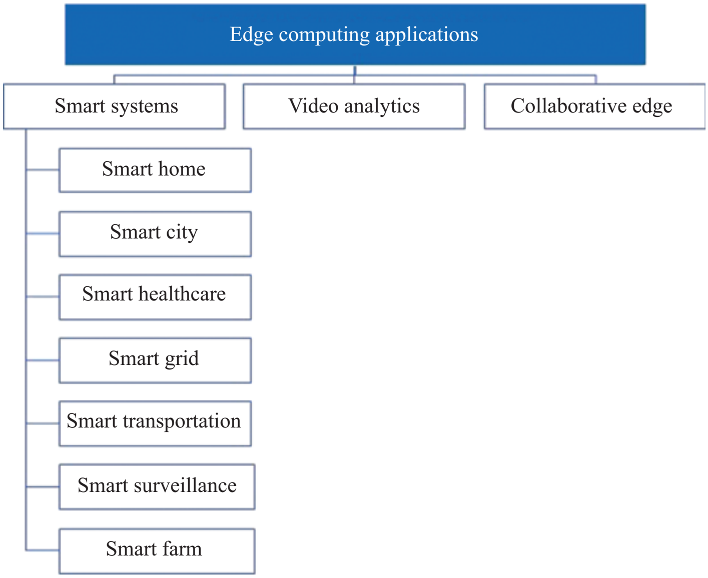

Figure 1.2 Edge computing applications

Applications that process geographically distributed data (e.g. healthcare data) require companies in several domains to collaborate and share data. Edge computing can meet this challenge through merging geographically distributed data and creating virtually shared data views. The virtual data are then shared with end users via a predefined service interface terminal, and an application will use this public interface to deliver complex end-user services. Collaborative edge participants deliver these services, and the computation only occurs at the participants' data centre to ensure data privacy and integrity [8].

Edge computing applications can be summarised in Figure 1.2.

1.4 Open challenges

Edge computing open challenges can be summarised in the following subsections [8,9,26,35].

1.4.1 Naming

It is safe to assume that there are a lot of things in edge computing. Many applications run at the top of the edge, and each has its own structure and service method. Like any other computer system, the naming scheme in edge computing is very significant for programming, identification of things, and data communication. However, there is no standardised naming mechanism for the edge computing model yet. Edge practitioners often need to learn various communication and network protocols so they can interact with a variety of things in the system. The naming scheme for edge computing has to tackle the portable nature of things, the highly dynamic network geographical distribution, privacy and safety, and the scalability of an enormous number of unreliable things [8].

1.4.2 System integration

Supporting different types of devices and meeting various service demands are a huge challenge in the edge computing environment. Edge computing integrates the combination of different platforms, network typologies, and servers. Therefore, the heterogeneous nature of this system makes it difficult to program and manage data and resources for many applications that run on various platforms in different locations. From a programming perspective, all applications and user programs are deployed and operated on cloud servers in cloud computing. Cloud providers, such as Google and Amazon, place these applications and programs in the appropriate locations and hardware to ensure their smooth operation. Most users do not know how these applications function or where the data and resources are located. Cloud computing offers manageable, centralised cloud service as one of its advantages. Furthermore, developers need to use just one programming language to develop applications for a specific platform because the cloud application is deployed on just one particular cloud service provider. Edge computing is, by contrast, very different from cloud computing. While there are many advantages to distributed topology, edge nodes are usually divergent platforms. Thus, developers would face serious challenges when developing an application that might be deployed and run on an edge computing platform. Certain schemes have been designed to address the challenges of programming in edge computing, but none of these schemes is designed for specific IoT purposes. The first step in IoT is finding the edge nodes. Before that, IoT devices cannot identify the type of platforms deployed in their vicinity. Moreover, many programs on the server side need to be deployed on the edge nodes, and this poses another challenge for edge node providers who need to deploy and manage those programs. As for data management, different storage servers are run by various operating systems. This represents another challenge for file naming, resource allocation, file reliability management, etc. As many IoT devices are simultaneously generating and uploading data, naming data resources becomes a difficult task [26].

1.4.3 Programmability

Users in the cloud computing model program and deploy their code on the cloud. The cloud provider then decides where cloud computing tasks are executed. Users do not usually know how the application is served. However, cloud computing is known for its transparent infrastructure. The program is usually written in a certain programming language and compiled in the cloud for a specific target platform, because the program only runs on the cloud. In contrast, computation is done from the cloud in edge computing, and the edge nodes are probably heterogeneous platforms. Thus, these nodes have different runtimes, and it is hard for the programmer to write an application that can be deployed in the edge computing model [8].

1.4.4 General-purpose computing on edge nodes

Hypothetically speaking, edge computing can run on several nodes located between the edge device and the cloud, including access points, base stations, gateways, traffic aggregation points, routers, switches, etc. For instance, base stations include Digital Signal Processors (DSPs) that are tailored to handle their workload. However, in practice, base stations may not be able to handle analytical workloads because DSPs are not designed to carry out general-purpose computing tasks. Furthermore, whether these nodes can perform additional computing tasks remains unknown. It might be useful to use such base stations at an off-rush time to leverage the computational capabilities of several available computing cores. Many business suppliers have made the first step towards software solutions that use edge computing. For instance, Nokia developed a software solution for mobile edge computing (MEC) that is meant to establish base station sites for the edge computing model. Likewise, Cisco's IOx16 provides an implementation environment for its incorporated service routers. As these solutions are hardware specific, they might not work in a heterogeneous environment. It would be challenging to develop portable software solutions that can function in different environments. Currently, research is underway to upgrade edge node resources that would support general-purpose computing. General-purpose CPUs can replace specialised DSPs as an alternative solution, but this would be a large investment [9].

1.4.5 Resource management

The integration of IoT and edge computing requires a thorough, comprehensive understanding and reinforcement of resource management. Network traffic and latency have a major impact on IoT devices and often result in computation deficiency and resource shortage because it requires utilising more power to re-transfer data in jammed settings. Edge computing can help reduce latency in devices, and decentralised resources will be instrumental in motivating and sharing these assets. These resources can be managed by various means, provided that it involves inexpensive computation. However, the disparity between service providers, devices, and applications causes a lot of complexities, and such interactions should be taken into consideration. In particular, the motivating factors of Edge/IoT resource management are synchronised with those of smart systems. The direct service of a system – which has multiple resource providers and a huge variety of applications and user needs – can be allocated, shared, and priced by optimisation global welfare or some other metric through competitive bidding or any other strategy [26].

1.4.6 Service management

To ensure the system's reliability, service management at the network edge requires supporting the following four essential features: differentiation, extensibility, isolation, and reliability [8].

Differentiation: The rapid development of IoT deployment will lead to the deployment of many services at the edge of the network, such as smart homes.

The priorities of these services will vary. For instance, critical services, such as diagnostics of things and failure alarm should be delivered faster than other services. Another example is health-related services, such as fall detection or heart failure detection, which should have a higher priority over other services like entertainment ones.

Extensibility: Extensibility presents a major challenge at the edge of the network. The things in the IoT are very dynamic. A versatile service management layer can resolve some of the current issues. When a user purchases a new thing, they need to add the current service with ease. Moreover, when a thing wears out and gets replaced, the previous service should easily adopt a new node.

Isolation: Isolation is another issue on the network's edge. When an application fails or crashes in a mobile OS, the whole system usually crashes or reboots. In a distributed system, various synchronisation mechanisms, such as a lock or a token ring, can be used to manage the shared resource. This problem, however, is more complicated in a smart EdgeOS. Many applications could share the same data resource, light control, for example. If one application crashes or does not respond, the user can still control their lights as this would not make the entire EdgeOS crash. Another example is when a user eliminates the only application that controls lights from the system, and the lights still work without losing connection to the EdgeOS. The introduction of a deployment framework may resolve this issue. If the OS detects the issue before the application is deployed, it will alert the user to avoid any access issues. Another aspect of this challenge is isolating a user's private data from third-party applications. For instance, your activity tracking application should not have access to your electricity usage data. To resolve this issue, the EdgeOS service management layer must be equipped with a reliable control access mechanism.

Reliability: Reliability is also another major challenge at the edge of the network, and it is viewed from various service, system, and data aspects. From a service aspect, it is sometimes difficult to identify the cause of a service failure in an accurate manner. For example, an air conditioning system stops working because the cable is cut, the compressor is broken, or even the battery of a temperature controller is dead. A sensor node might have lost connection to the system as a result of a battery outage, bad connection, component failure, etc. Thus, when some nodes lose connection at the network's edge, it is not enough to just maintain the current service; it would be more useful to provide the action after the node fails. From a system aspect, maintaining the network topology of the entire system is truly essential for the EdgeOS. All system components can send status/diagnosis data to the edge. This feature allows simple system deployment of services, such as failure detection, replacement of things, and data quality detection. From a data aspect, data sensing and communication are mainly responsible for reliability issues. Things at the network's edge fail for various reasons, and they report inaccurate data under uncertain conditions, such as low battery level. There have been various proposals for new communication protocols for data collection. Such protocols can support a large number of sensor nodes and the highly dynamic network condition. However, the connection is not as reliable as the one in Bluetooth or Wi-Fi. It remains a challenge for a system to provide a reliable service by leveraging many reference data sources and data history records, even when the sensing data and communication are not so reliable.

1.4.7 Discovering edge nodes

The literature thoroughly examines the resources and services provided in a distributed computing setting. Various techniques are used in monitoring tools and service brokerages to facilitate such examination in both closely integrated environments as well as separate ones. To enhance performance, using a technique like benchmarking supports the mapping of tasks by setting up decision-making onto the most suitable resources. However, using the network's edge requires exploring mechanisms to help identify suitable nodes for a decentralised cloud configuration. Due to the large number of devices at this layer, such mechanisms cannot be manual. Furthermore, the mechanisms have to work well with heterogeneous devices from several generations as well as modern ones. The benchmarking techniques need to determine the resources availability and efficiency in a very rapid manner. These mechanisms should facilitate smooth nodes' integration and elimination at various hierarchical levels in the computational workflow without increasing latency or compromising the user's experience. Users appreciate the system's ability to tackle faults and to enable automatic recovery on the node in a reliable and proactive manner. In this context, current cloud approaches do not practically discover edge nodes [9].

1.4.8 Partitioning and offloading tasks

The evolution of computing environments has brought forth many techniques that facilitate the distribution of tasks over multiple geographic locations. Workloads are distributed and executed across various locations, and the distribution of tasks is usually carried out via a language or management tool. However, if we use edge nodes to offload computations, it will be challenging to partition computations in an effective manner. This should be done automatically, regardless of the edge nodes' storage and computing capacities, by developing schedulers that deploy tasks onto edge nodes [9].

1.4.9 Uncompromising QoS and QoE

Edge nodes ensure that quality of service (QoS) and quality of experience (QoE) are effectively delivered to users. One of the edge computing principles is to avoid overloading nodes with intensive computing workloads. In this case, it is difficult to ensure the nodes' high performance and reliability when delivering the intended workloads as well as workloads received from a data centre or edge devices. The user of an edge device or a data centre still expects some service even when the edge node is fully exploited. For instance, overloading a base station might affect the service provided to the connected edge devices. This requires realisation of rush hours of edge nodes' usage in order to partition and schedule tasks in a flexible manner. A management framework can be very useful in this sense, but it causes issues in terms of monitoring, scheduling, and rescheduling at infrastructure and platform levels [9].

1.4.10 Smart system support

Smart systems essentially interlink network communication technologies with sensors and actuators to achieve system awareness and actuators that are meant to send signals to the system to take action. Smart systems are viewed as an extension of IoT technologies. The integration of sensing devices offers countless opportunities for data collection, system management, and resource allocation and optimisation. Smart systems are mainly utilised in areas such as smart grids, smart cities, smart transportation, smart healthcare, and others. The increasing utilisation of smart systems enables edge computing to reduce latency to the lowest level possible and to increase storage capacity in devices that are computationally deficient. In the same way, edge data analysis can also ensure a high degree of resilience in compromised systems [26].

1.4.11 Advanced communication

Edge computing has currently changed remote computation and storage as it has removed barriers to offer fast, low-latency, and high-computation applications. In the same way, future 5G cellular network technologies, including ultra-dense networks (UDNs), massive MIMO (multiple-input and multiple-output), and millimetre-wave, are improving as they move towards reducing latency, increasing throughput, and providing support for interconnected groups in dense networks. This will inevitably develop communication technology. 5G is considered to be the next generation in communication technology. It aims to provide users with pervasive network connectivity and access to data. In this regard, 5G, IoT, and edge computing can be integrated to achieve flexible and efficient communication. 5G technology can also help make many IoT applications more efficient [26].

1.4.12 Privacy and security

Privacy and security are important issues that require careful consideration. Cloud and edge computing paradigms both have numerous security issues and challenges. Edge computing specifically involves connecting various multiple technologies (e.g. peer-to-peer systems, wireless networks, and virtualisation), and this requires a comprehensive integrated model to safeguard and manage each technology platform and the system as a whole. However, the advancement of edge computing may result in some unpredicted issues. Certain situations that remain unexamined, such as the interplay of various edge nodes and the local and global migration of services, all create the potential for malicious behaviour. Moreover, the built-in features of edge computing further determine the viability of privacy and security measures.

Although the distributed framework offers many advantages to IoT, it is difficult to maintain privacy and security in such distributed structures. In terms of privacy, edge computing can provide an efficient platform to future IoT. Processing data at the edge enables the edge computing model to utilise privacy-sensitive data belonging to end users. Sensing data from IoT systems is stored at edge nodes, which are more susceptible than cloud servers. Thus, edge computing should utilise efficient data protection mechanisms to protect data privacy. As for security issues, authenticating gateways on different levels is a main problem in edge computing.

For example, smart meters used at residential homes have their own IP address [26]. Offloading an application from the end user's device to the edge server requires transferring data from the mobile device to the edge server. It allows intruders to access data by committing security violations. End users' devices are battery powered, and therefore require applying lightweight security and safety mechanisms. Security and privacy solutions should also be flexible so they can promote the edge computing goal of reducing execution time. However, it is difficult to find a reliable lightweight solution for a complex security problem in a diverse environment [35]. We can conclude here that security and safety are the main issues preventing advances in edge computing.

1.4.13 Using edge nodes publicly and securely

Hardware resources at data centres, supercomputing centres, and private organisations can be converted into computing services using virtualisation. As providers identify associated risks for users, they provide computing on a pay-as-you-go basis. The computing marketplace has, therefore, become highly competitive as it offers numerous options that meet the standards of the service-level agreements (SLAs). However, using alternative devices, such as switches, routers, or base stations, as publicly available edge nodes would create some issues. First, there is a need to identify the risk associated with the devices owned by public and private organisations as well as the entities that employ them. Second, the device's intended purpose cannot be compromised when it serves as an edge computing node, such as when using a router to manage Internet traffic. Third, multi-tenancy on edge nodes requires using a security-oriented technology. For instance, containers represent a lightweight technology used on edge nodes, but they need to demonstrate potent security features. Fourth, the edge node user must be guaranteed a minimum level of service. In addition, to develop appropriate pricing models for accessible edge nodes, the workload, computation, data location, data transmission, and cost of maintenance and energy bills all need to be addressed [9].

1.4.14 Monitoring, accounting, and billing

It is important to manage edge computing resources, accounting, and billing to ensure QoS and to set the appropriate charge for the offered edge computing service, and this requires service providers to apply a sustainable business model. However, it is difficult to design such a model given the customer's mobile nature and the limited services range. A user usually uses an edge node service for a limited time, such as students who use the service at a university cafeteria during lunch hours. This short-time use makes it difficult to maintain a business model. In addition, when a user moves and execution is transferred from one edge platform to another, it becomes difficult to divide charges between the involved service providers and to maintain a business model for monitoring, accounting, and billing. This requires a business model that includes multiple levels of granular charging [35].

1.4.15 Social collaboration

It is difficult for various service providers to collaborate and accomplish a common goal. Standardisation and competition are the main obstacles that hinder such social

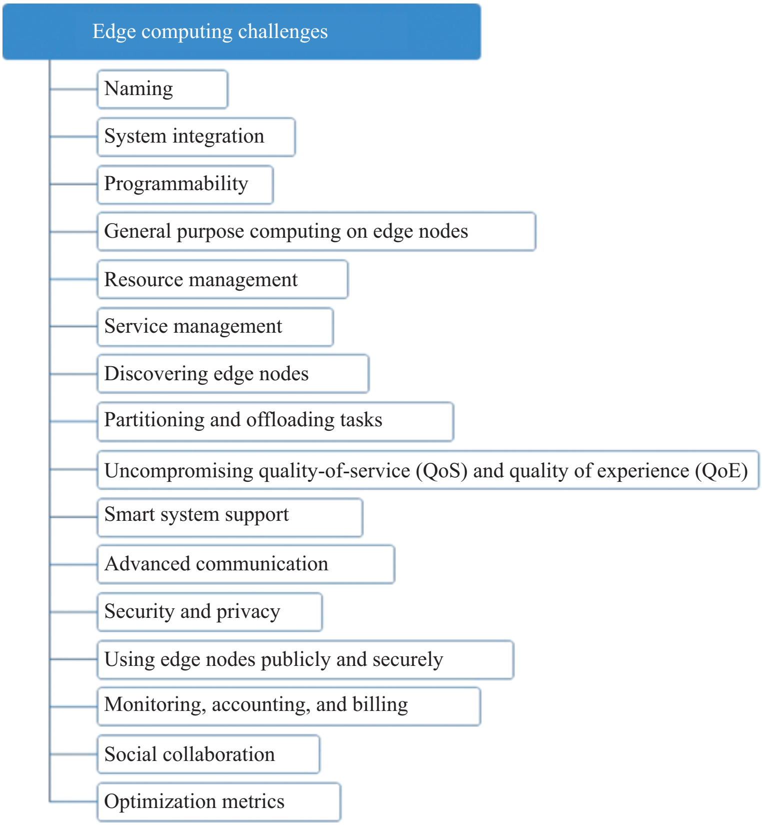

Figure 1.3 Edge computing challenges

collaboration. Various companies offer different edge services. However, it is difficult to achieve social collaboration among them given the competitive marketplace. The heterogeneous nature of the devices provided by different companies also presents an obstacle for achieving such collaboration. Overcoming the collaboration challenge would enhance data analytics efficiency [35].

1.4.16 Optimisation metrics

Edge computing involves several layers with different computing capacities. The workload distribution can be problematic because we need to choose the layer that would handle the workload or the number of tasks to assign for each part. Many allocation strategies can be applied to complete a workload. For example, the workload can be evenly distributed on each layer or completed at each layer as much as possible. In extreme cases, the workload would be fully processed on the cloud. When choosing an optimal allocation strategy, some measures must be taken into consideration, such as latency, bandwidth, energy, and cost [8].

Edge computing challenges can be summarised in Figure 1.3.

1.5 Conclusions

Edge computing extends cloud computing by bringing the services closer to the end user at the network edge. This chapter aims to explore the edge computing model. Although this model offers numerous benefits, it presents many new challenges. The chapter highlights the significance of edge computing as it provides an insight into its definition, architecture, and distinguishing characteristics as compared to fog and cloud computing characteristics. It also demonstrates the applications the edge computing enables. There is a need to address the challenges of adopting the edge computing model, as this model will become the core component of the future computing landscape. The chapter serves as a guideline for related future directions.

References

[1] Shawish A, and Salama M. Cloud computing: paradigms and technologies. In: Xhafa F, and Bessis N, editors. Inter-cooperative Collective Intelligence: Techniques and Applications. Studies in Computational Intelligence, Vol. 495. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014.

[2] Al-Dulaimy A, Zantout R, Itani W, and Zekri A. Job submission in the cloud: energy aware approaches. The 24th World Congress on Engineering and Computer Science (WCECS 2016), Vol. 1, International Association of Engineers (IAENG), San Francisco, USA, 2016.

[3] Al-Dulaimy A, Itani W, Zantout R, and Zekri A. Type-aware virtual machine management for energy efficient cloud data centers. Journal of Sustainable Computing: Informatics and Systems (SUSCOM), Elsevier, Vol. 19, pp: 185-203, 2018.

[4] Chang C, Srirama S, and Buyya R. Internet of things (IoT) and new computing paradigms In: Buyya R, Srirama S, editors. Fog and Edge Computing: Principles and Paradigms. Wiley. 2019. pp. 1–23.

[5] Margulius D. Apps on the edge, InfoWorld. Last accessed July 2019. https://www.infoworld.com/article/2677229/apps-on-the-edge.html.

[6] Pang HH, and Tan K. Authenticating query results in edge computing. In: Proceedings. 20th International Conference on Data Engineering; 2004. pp. 560–571.

[7] Garcia Lopez P, Montresor A, Epema D, *et al.* Edge-centric computing: vision and challenges. ACMSIGCOMM Comput Commun Rev. 2015;45(5):37–42.

[8] Shi W, Cao J, Zhang Q, Youhuizi L, and Lanyu X. Edge computing: vision and challenges. IEEE Internet of Things Journal. 2016;3(5):637–646.

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

[10] Finnegan M. Boeing 787s to create half a terabyte of data per flight, says Virgin Atlantic, Computer World, UK. Last accessed July 2019. https://www.computerworlduk.com/data/boeing-787s-create-half-terabyte-of-data-per-flight-says-virgin-atlantic%2D3433595/.

[11] van Rijmenam M. Self-driving cars will create 2 petabytes of data: what are the big data opportunities for the car industry? DataFloq. Last accessed July 2019. https://datafloq.com/read/self-driving-cars-create-2-petabytes-data-annually/172.

[12] Data Never Sleeps 6.0, DOMO. Last accessed July 2019. https://www.domo.com/learn/data-never-sleeps-6.

[13] Zhu J, Chan DS, Prabhu MS, Natarajan P, Hu H, and Bonomi F. Improving web sites performance using edge servers in fog computing architecture. In: 2013 IEEE Seventh International Symposium on Service-Oriented System Engineering; 2013. pp. 320–323.

[14] Sawasaki N, Ishihara T, Mouri M, Murase Y, Masui S, and Nakamoto H. Front-end device technology for human centric IoT. Fujitsu Scientific & Technical Journal. 2016 10;52:61–67.

[15] Al-Dulaimy A, Itani W, Zekri A, and Zantout R. Power management in virtualized data centers: state of the art. Journal of Cloud Computing. Advances, Systems and Applications (JoCCASA), Springer. 2016;5(1):6.

[16] Evans D. The Internet of things: how the next evolution of the Internet is changing everything. Cisco Internet Business Solutions Group (IBSG); 2011. https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf.

[17] International Data Corporation (IDC). In: IDC Directions Conference; 2016.

[18] Cao J, Zhang Q, and Shi W. Introduction. In: Cao J, Zhang Q, Shi W, editors. Edge Computing: A Primer. Springer; 2018. pp. 1–9.

[19] Singh SP, Nayyar A, Kumar R, and Sharma A. Fog computing: from architecture to edge computing and big data processing. The Journal of Supercomputing. 2019;75:2070–2105.

[20] Baktir AC, Ozgovde A, and Ersoy C. How can edge computing benefit from software-defined networking: a survey, use cases, and future directions. IEEE Communications Surveys Tutorials. 2017 Fourthquarter;19(4):2359–2391.

[21] Lea P (editor). Cloud and fog topologies, in Internet of Things for Architects. Packt Publishing; 2018. pp. 338–371.

[22] Satyanarayanan M. The emergence of edge computing. Computer. 2017; 50(1):30–39.

[23] SNIPS. Last accessed 2019. https://snips.ai/content/intro-to-edge-computing/?utm_campaign=2019_Edge\&gclid=Cj0KCQjwjYHpBRC4ARIsAI%2D3GkFHj9hwwA6upDLKCk3oTsg8sFJnGfXKYHux8U6aliFqhyFBG9yh_x4aAiXhEALw_wcB#what-is-edge-computing.

[24] Khan SU. The curious case of distributed systems and continuous computing. IT Professional. 2016;18(2):4–7.

[25] Hu P, Dhelim S, Ning H, and Qiu T. Survey on fog computing: architecture, key technologies, applications and open issues. J Network and Computer Applications. 2017;98:27–42.

[26] Yu W, Liang F, He X, *et al.* A survey on the edge computing for the Internet of things. IEEE Access. 2018;6:6900–6919.

[27] Cao J. SOFIE: Smart operating system for Internet of everything, PhD Thesis, Graduate School of Wayne State University. 2018.

[28] Samie F, Bauer L, and Henkel J. IoT technologies for embedded computing: a survey. In: 2016 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS); 2016. pp. 1–10.

[29] Kraemer F, Brten A, Tamkittikhun N, and Palma D. Fog computing in healthcare: a review and discussion. IEEE Access. 2017;9206–9222.

[30] Hou W, Ning Z, Guo L, and Zhang X. Temporal, functional and spatial big data computing framework for large-scale smart grid. IEEE Transactions on Emerging Topics in Computing. 2019;7(3):369–379.

[31] Babu RV, and Makur A. Object-based surveillance video compression using foreground motion compensation. In: 2006 9th International Conference on Control, Automation, Robotics and Vision; 2006. pp. 1–6.

[32] Xu R, Nikouei SY, Chen Y, *et al.* Real-time human objects tracking for smart surveillance at the edge. In: 2018 IEEE International Conference on Communications (ICC); 2018. pp. 1–6.

[33] Porambage P, Okwuibe J, Liyanage M, Ylianttila M, and Taleb T. Survey on multi-access edge computing for Internet of things realization. IEEE Communications Surveys Tutorials. 2018 Fourthquarter;20(4):2961–2991.

[34] Bonomi F, Milito R, Zhu J, and Addepalli S. Fog computing and its role in the Internet of things. In: Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing. MCC’12. New York, NY, USA: ACM; 2012. pp. 13–16. Available from: http://doi.acm.org/10.1145/2342509.2342513.

[35] Ahmed E, Ahmed A, Yaqoob I, *et al.* Bringing computation closer toward the user network: is edge computing the solution? IEEE Communications Magazine. 2017;55(11):138–144.