Processing IoT Data: From Cloud to Fog-It's Time to Be Down to Earth

Chapter · May 2018

DOI: 10.4018/978-1-5225-4044-1.ch007

Pijush Kanti Dutta Pramanik
National Institute of Technology, Durgapur, India

Prasenjit Choudhury
National Institute of Technology, Durgapur

Chapter 7

Processing IoT Data:
From Cloud to Fog—It's Time to Be Down to Earth

Pijush Kanti Dutta Pramanik
National Institute of Technology Durgapur, India

Saurabh Pal
Bengal Institute of Technology, India

Aditya Brahmachari
National Institute of Technology Durgapur, India

Prasenjit Choudhury
National Institute of Technology Durgapur, India

ABSTRACT

This chapter describes how traditionally, Cloud Computing has been used for processing Internet of Things (IoT) data. This works fine for the analytical and batch processing jobs. But most of the IoT applications demand real-time response which cannot be achieved through Cloud Computing mainly because of inherent latency. Fog Computing solves this problem by offering cloud-like services at the edge of the network. The computationally powerful edge devices have enabled realising this idea. Witnessing the exponential rise of IoT applications, Fog Computing deserves an in-depth exploration. This chapter establishes the need for Fog Computing for processing IoT data. Readers will be able to gain a fair comprehension of the various aspects of Fog Computing. The benefits, challenges and applications of Fog Computing with respect to IoT have been mentioned elaboratively. An architecture for IoT data processing is presented. A thorough comparison between Cloud and Fog has been portrayed. Also, a detailed discussion has been depicted on how the IoT, Fog, and Cloud interact among them.

DOI: 10.4018/978-1-5225-4044-1.ch007

INTRODUCTION

Typically, IoT devices are attributed to very limited computation and storage capacity. To get over this limitation, Cloud Computing has been the most favoured platform for processing IoT data, which provides on-demand and scalable resources for computing and storage. The sensor data are transported to the Cloud data centre, where they are processed, and the outcome is sent to the subscribed applications. Furthermore, the data centres may store the IoT data, if necessary, for analysis to extract further knowledge which helps in business decision making. The Cloud platform has become popular for IoT data processing mainly because of economic reason. By opting Cloud Computing, organisations have freed themselves from the hassle of establishing their own computing setup and maintenance. But as we are heading towards the smart world for a smart living, uses of sensors and wireless networks locally has been on the rise, and the data generated locally is increasingly consumed locally (Chiang, 2015). In other words, instead of at the remote centralised Cloud data centre, the data gravity is shifting more and more towards the neighbouring to the data source or, formally what we call, the edge of the network. For these applications, it is extremely crucial to be facilitated by low and predictable communication latency for real-time interaction, location awareness, and support for mobility and large-scale networks (Milunovich, Passi, & Roy, 2014). The traditional Cloud Computing architecture lacks in these aspects. IoT requires a different computing architecture that enables distributed processing of IoT data with mobility support and quick response whenever and wherever wanted. Fog Computing perfectly befits this scenario. Fog Computing is particularly suited for applications that demand real-time response with predictable and minimal latency (Milunovich, Passi, & Roy, 2014). The edge devices such as set-top-boxes, access points, routers, switches, base stations etc. are becoming ever more powerful in terms of computing, storage and networking. Hence, they are being considered as capable candidates to perform computational jobs. Considering that, Fog Computing can play a big role in processing the huge amount of data generated from billions of distributed IoT sensors. Fog Computing is not to replace the Cloud Computing rather it augments Cloud Computing by extending its services to the edge of the network. Principally, both Cloud and Fog serve the end users by providing data, computing resource, storage, and application services. But Fog is differentiated from the Cloud with respect to its proximity to the source and sink, its distribution irrespective of the geography and last but not the least its support for mobility (Mora, 2014). In the case of Cloud-based IoT data processing, every single bit of data would have to be shipped to the data centre. When the size of data to be processed grows enormously (and that is the exact case of IoT), it becomes very expensive to move them around. Since in Fog Computing data are being processed locally, the burden of transporting these data is lessened. The processed data are sent to the Cloud only if they are to be stored for further analysis and historical purposes. Also, since the data are processed very close to the source, the end-user service becomes very prompt which is very crucial for maintaining QoS in real-time and machine-to-machine (M2M) applications. Handling services in the Fog provide better user experience and more efficient and effective applications of IoT data. In this chapter, we have advocated for employing Fog Computing for IoT applications while discussing and comparing it with Cloud Computing in several aspects.

The rest of the chapter is organized as follows. A brief review of IoT and Cloud Computing is presented in section 2. The section 3 discusses how IoT data is processed in the Cloud, along with the advantages and issues. In section 4, we shall discuss the basics of Fog Computing including its characteristics and architecture. We shall identify the differences between Cloud and Fog in section 5. Section 6 addresses the processing of IoT data in the Fog. An architecture has been laid out for this. This section also discusses how IoT, Fog, and Cloud interact with each other. The prevalent commercial Fog Computing solutions are also discussed briefly. Section 7 and 8 respectively mentions the applications and challenges of Fog Computing. And section 9 concludes the chapter.

A Brief Introduction to IoT and Cloud Computing

Internet of Things (IoT) is the network of physical things, connected together to share data among themselves and other computing devices. The 'things' may be sensors, automobile, kitchen appliances, electronic devices, building, elevator or other devices. These interconnected 'things' collect and exchange data to share their state information. With an application of intelligent middleware, 'things' will be transformed into intelligent entities thereby blending the physical and virtual world together making the machine and human interactions very personalized. According to IBM:

The IoT is expected to make the physical world every bit as easy to search, utilize, and engage with as the virtual world (Milunovich, Passi, & Roy, 2014).

The sole objective of IoT is automation and monitoring, automating every activity which involves digital interventions.

In terms of practical realization, IoT applications in real time produce huge data within a time lag of fraction of second as a constant stream. To process the data, further, it requires high-speed data processing in continuum for data analysis to find the valuable insights. This puts a lot of strain on traditional private data centres owns by individual businesses in terms of network load management, centralized high data storage, processing, scalability etc. To realize this feat the Cloud Computing came into picture which acts as mere enabler or catalyst (NxtraData, 2016). Cloud Computing often referred to as "Cloud", as defined by IBM "is the delivery of on-demand computing resource – everything from applications to data centres – over the Internet". The Cloud provides computing resources like software, platform and infrastructure as pay-and-use services. The key characteristics that have made Cloud Computing popular are (Rouse, 2017; What is Cloud Computing? A beginner's guide, 2017):

- **Self-Service Provision:** This feature enables to choose any type of computing resources as per the need.
- **Elasticity:** Computing resource is scaled up or down based on the need.
- **Pay Per Use:** Resources can be paid only for use.
- **Performance:** Cloud Computing runs over secure data centres which are regularly upgraded with fast and efficient computing hardware. The network latency is very low.
- **Reliability:** The data backup is taken at multiple sites.

The objective of Cloud Computing is to establish a high-performance scalable virtual system with enormous data capacity and virtually capable of serving all type of processing jobs. This environment gives facility to all business enterprises and start-ups in revolutionarily cost reduction for what being spent on putting up computing infrastructure like private data centres.

These two technologies are completely different, where IoT act a platform and Cloud as a service. Though these technologies are conceptually independent still they are complementary to each other. IoT generates huge data, where Cloud provides a way for these data to reach their destination (storage, processing – data analytics).

Even though Cloud-based IoT model is a suitable solution, the deployment of IoT application has many challenges originated from economic consideration, social concern, technical limitation and administrative issues. Cloud may not be suitable for deploying all kinds of IoT-based applications, for example, where the data generated by 'things' are useful in its very locality or other entities in local proximity.

IoT DATA PROCESSING IN THE CLOUD

Typically, IoT devices are attributed with very limited computation and storage capacity and as a reason, Cloud Computing is there as a choice of platform for processing IoT data for quite a long time.

IoT produces huge data and needs huge storage and real-time processing. Comprehending the current contextual situation of 'things' by analysing the present and past data is inherently complex. The current contextual information or the data insight may help in decision making to take current action and future predictions. In this perspective, Cloud Computing is a possible solution, offering IoT-based applications the advantage of huge data storage and computational power to process out the complex computation and other software services in dynamic, scalable and virtualized manner at a very low cost (Alamri et al., 2013).

The mechanism for IoT data processing through Cloud is straightforward. Cloud linked to wireless sensor node (WSN) through gateways (Cloud gateway and sensor gateway) incorporated at both ends of the link. These gateways would allow data collection, aggregation and flow management. The sensor gateway collects the huge data streamed from sensors, compresses it and sends it to Cloud gateway. Whereas the Cloud gateway further decompresses the sensor data and store it in large Cloud storage servers (Alamri et al., 2013).

Advantages

There are several advantages of Cloud-IoT model, which had earned this model reputation in all respects. The advantages are described as follows (Alamri et al., 2013):

- **Increased Data and Processing Power:** Cloud provides enormous storage facility and processing power. Organizations can keep IoT sensor data easily in the Cloud without the hassle of creating its own private data storage. The huge processing power helps to process big complex data for the large-scale application.
- **Scalability:** The large routing architecture of Cloud model allows the IoT-based application to scale up in the Cloud as the need for new computing resources and services arises. This enables existing IoT-based application to scale up to large sizes based on new requirements without having to invest heavily in the new resources added.

- **Collaboration:** Cloud enables huge IoT data stored in storage server to be shared among different IoT application and group of users.
- **Dynamic Provisioning of Services:** Cloud provides varying services which allows processing the relevant information dynamically whenever and wherever they are needed. The API available in Cloud for various services enables the IoT applications to communicate with the data source. Cloud maintains 99.99% uptime, making its services available practically to anywhere and anytime as long as the IoT's have Internet connection (Coles, 2017).
- **Multi-Tenancy:** The multi-tenancy attribute of Cloud allows instances of IoT applications to share the same service infrastructure of Cloud in a varying manner. Further, Cloud allows integrating several services (Infrastructure, platform and software) from different service providers available on Clouds and Internet to meet the tailor-made demand of the user.
- **Flexibility:** Cloud provides flexibility to support IoT applications to scale up based on the business requirement and IoT-based application development and other required services through Cloud's customizable software services. The Cloud flexibility can be realized in terms of scalability, storage option, control choice and security. Scalability allows Cloud to support dynamic workload of IoT. Whereas storage option gives flexibility of choice to store data based on business model into private, public or hybrid storage. Cloud gives the flexibility on how the IoT application get controlled by Cloud (Sauerwalt, 2017).
- **Agility of Services:** By accommodating changing business demands Cloud allows rapid application development, testing and deployment. In this perspective, IoT can gain access to more resources (expensive hardware and software/applications) very rapidly as well as relinquish them to Cloud when the task finishes. The Cloud agility thus efficiently allows IoT applications to adapt to the rapidly changing business need and policies in a very cost-effective manner (Kumari, 2015).
- **Resource Optimization:** IoT Cloud model enables resource optimization by enabling resources (infrastructure, platform and software) sharing among several numbers of application. This reduces the cost of operation and gains in the service quality. The IoT and Cloud model is benefited to all size of organization – small, medium or big by the concept of resource sharing based on requirement and pay for use scheme thus optimizes the resource use.
- **Analysis:** The huge scalable processing power of Cloud and other data processing services available make data analysis job easy. This makes Cloud very attractive for various kinds of data analysis jobs over the accumulated sensor data to get valuable data insight into the future decision making.
- **Visualization:** The Cloud provides tools (visualization API) that help to visualize sensor data pattern in terms of the diagram and allows applying statistics to predict the future data pattern.

Issues

Despite the increasing usage of Cloud, its application to IoT raises many challenges and is proved to be disadvantageous to IoT.

- **Network and Communication:** Cloud Computing is an Internet-based computing model, where Internet act as the backbone of the communication network. The Internet is a non-homogenous and loosely controlled structured, having numerous types and topologies, varying network speed and heterogeneous technologies. Internet communication paths are very dynamic and thus the data communication path often changes. The non-homogenous and loosely controlled nature raises many issues for Cloud-IoT model like network latencies and bandwidth constraints that badly affect Quality of Service (QoS) (Firdhous, Ghazali, & Hassan, 2014).

- **Latency Constraint:** Network latency caused by communication delay and delay jitters is one of the big issues in an Internet-based model. Any real-time applications which allow interaction in actual time are quite badly affected by network latency. IoT application in the industrial control system and other real-time systems like a vehicle to vehicle communication, a vehicle to signal system communication, Virtual Reality System, gaming and drone flight demands end to end device communication latency less than a millisecond. Since Cloud is distantly apart from the source of the event or the user it may cause a huge time delay in information communication. To meet the stringent time requirement of IoT-based application it is desirable that the data processing, analysis and decision making should be taken close to the data source or target user (Chiang & Zhang, 2016).
- **Network Bandwidth Constraint:** The huge number of connected 'things' produces data exponentially at high speed. Sending these enormous amounts of data to Cloud at high speed in the continuum is network constraining and bandwidth intensive which causes resource starvation. Catering the demand for very high bandwidth may not be possible, practically, in all IoT cases, as the devices are connected to Cloud through multi-hop and varying network structures where network devices have variable data transfer capacity (Chiang & Zhang, 2016).
- **Servicing Heterogeneous Devices:** IoT consists of heterogeneous devices of severely limited resources and each communicates with varying protocols. To fulfil their computation need, these devices rely on the Cloud. Interacting directly with all these devices is quite unrealistic and prohibitive for Cloud, as because each requires resource-intensive processing and handling complex or unknown protocols. For example, sending and receiving data from IoT device to and from Cloud requires data encryption and decryption. These processes carried at both ends need a sophisticated algorithm and are a resource and time-consuming process. Furthermore, IoT node ends require regular firmware updates from Cloud to handle sophisticated encryption and decryption techniques, which rather seems to be highly unrealistic (Chiang & Zhang, Fog and IoT: An Overview of Research Opportunities, 2016).
- **Security Challenges:** The IoT data, due to the considerable distance between the source and the Cloud, travel through multiple hops and complex network structures which makes it vulnerable to security issues. As the data travel through network edges, crossing multiple nodes, it becomes vulnerable to attack and corruption. Even though the data are encrypted, the more it goes away from user deep into the Internet, the more time it spends in transit, the higher the risk becomes (Firdhous, Ghazali, & Hassan, 2014).

INTRODUCTION TO FOG COMPUTING

We know that when Cloud descends to the ground, it is named as Fog. Similarly, when the computation has been shifted from remote Cloud to the system that is close to the data source is termed as Fog Computing. Fog Computing is the term used by Cisco for representing edge computing. In edge computing, the processing is done at the devices that reside at the edge of the network. The edge devices include routers, switches, Wi-Fi access points, set-top-boxes, base stations etc. These devices no longer are used