International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 03 | Mar 2019 www.irjet.net p-ISSN: 2395-0072

EDGE COMPUTING

THE NEXT COMPUTATIONAL LEAP

ASWATHAMAN.G

B.E Student,Department of Computer Science & Engg.
Sri Ramakrishhna Engineering College, Coimbatore, India.

Abstract - *The expansion of the Internet of things and highly persuasive, rich cloud services have improved and pushed the horizon of a new computational field called edge computing, which gives the ability to process data at the edge devices. Edge computing has the ability to reduce the response time, battery life constraints, cutting bandwidth cost and more importantly data security and privacy. In this paper, we discuss the definition of Edge computing and also its ability to transform the world we live in with highly fast and secure devices that can make smart homes and smart cities. We also address the challenges and opportunities in this domain to inspire more research and product development.*

Keywords – highly persuasive, edge computing, bandwidth cost, data security, privacy, smart homes and cities.

I. INTRODUCTION

Cloud computing has tremendously changed the way we approach applications since its introduction in 2005. Various cloud services like the icloud, google cloud, dropbox etc has helped us to manage our day to day activities and business activities in a very efficient and influential way. The Internet of things is the concept of "making a computer sense information without the aid of human intervention"[1] is used in various sectors like HealthCare, home automation etc. With IoT, there is a large quantity of data that has been generated by the devices and it is been sent to cloud for processing and also more importantly for training the model using that data. Transferring such a large amount of data through the internet increases the latency time and may cause a security breach.

These problems can be overcome by performing the calculations in the edge devices itself. This decreases the latency time because the data is not transferred anywhere and privacy is also ensured by the same. By the support of both the cloud computing and IoT platform, we can envision a new field of computation called the edge computing which gives the ability to perform on-device inference using platforms like caffe2, tensorflow lite etc. We can start by analyzing the need for edge computing and their ability to make changes to the way we approach the IoT platform and the potential in it to transform the way devices work today.

The remaining sections of this paper are organized in such a way, Section 2 discusses the need for computations in edge devices. Section 3, we discuss its areas of applications and Section 4, talks about the challenges and opportunities in this field. Finally, we conclude in Section 5.

Figure 1. Edge Computing Platform

II. EDGE COMPUTING

The edges of the network are the birthplace of data, therefore, it is more efficient to process the data at the edge itself. Cloud computing is not the most efficient way for processing data when it is produced in the edge devices. Edge computing helps to cut off lots of bandwidth cost, reduces data transfer latency time and make the and inference more secure and unbreachable. Many other concepts like fog computing, micro datacenter, cloudlet etc were introduced because cloud computing did not serve to be the best and the most optimal way of making inference using Deep learning models or in simple terms cloud computing did not help us to make intelligent Application in edge devices by deploying Deep Learning models in the cloud.

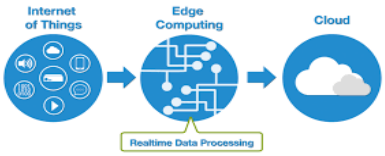

Figure 2. Egde Computing

A. The need for edge computing

Deploying models in the cloud and making the inferences by transferring the data via request and response from the edge devices served well until there were no highly capable processors in edge devices that were capable of performing some highly complex operations and run heavy models like MobileNet and InceptionV3 etc[2]. The newly launched processors like kirin970, Snapdragon 855 and Mediatek helio is released with dedicated GPU's and NPU's (Neural Processing Unit) for making more faster inferences in smartphones[3]. For example, 5GB of data is generated by a Boeing 787 every second[4] and in autonomous vehicles like Google's self-driving car generates 1GB of data every second and it requires real-time processing for making accurate and quick decisions[5].

The data producers produce data and send it to the cloud and the data consumers send a request to the cloud to fetch the required data for consumption. But, the structure is not sufficient for the data processing to happen very efficiently in the edge device because of the enormous amount of data generated by the data producers which cost a huge amount of bandwidth to transfer that data. The latency time taken by the whole process makes the whole project to lack behind real-time. By making inference in the edge devices most of the data is not sent to the cloud itself so that a lot of bandwidth cost is cut off and the response of the inferences made with the model is near real-time. Most importantly the edge devices are mostly energy constraint and the wireless communication modules used in the edge devices are very energy consuming while transferring large amounts of data. So, by deploying the models in the edge devices we can make offload a large amount of workload and consume less energy for transferring data from cloud to edge devices and vice versa.

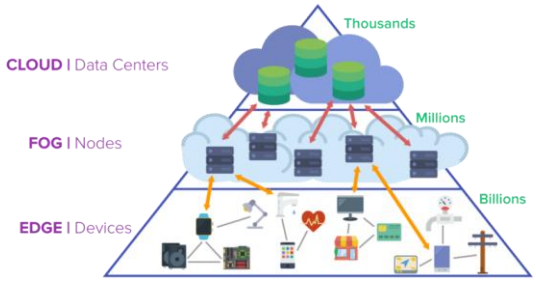

Figure 3. Need for Egde Computing

B. Advantages of Edge Computing

In Edge Computing we process the generated data at the same place where it is produced and this approach has a lot of advantages compared to the traditional way of computations done with the models in the edge devices with the help of various cloud-based services. A face-recognition application app was done as a proof-of-concept by a group of researchers where the output results were astonishing in which they found that the inference time was reduced from 900 to 169ms when

the computations were moved from cloud to edge and also they found that the energy consumption was reduced from 30%-40% after transferring the computations to edge[6].

III. APPLICATIONS OF EDGE COMPUTING

Case 1: Smart City

The edge computing applications can be expanded in a large city scale applications that helps to automate and use lots of intelligent applications that can make a city more secure and more organized. The edge computing platform can serve to be a very optimal solution for making an efficient smart city. By considering these characteristics in mind.

1) Enormous amounts of data: A highly populated city produces hundreds of PB data per data. These data highly contributes to public safety, healthcare, utility, and logistics etc, processing these enormous amounts of data by sending it to the cloud will obviously make the applications to lag behind real-time and makes it unreliable to take immediate and emergency operations like generating SOS , detecting accidents or crime at the very instance of happening etc. By embedding the calculation into the edge devices these kinds of issues can be neglected.

2) Location: Edge computation platforms can serve better when there is a need for making computations that require location awareness and it helps us to find an optimized solution of processing the location data without transporting it to the cloud.

Case 2: Autonomous Vehicles

The autonomous vehicles like Google's self-driving cars make use of the edge platform to the best possible extent because autonomous vehicles have a requirement of making sudden and immediate life-saving decisions in a very limited amount of time where it should decide on avoiding an accident. Completely relying on the cloud platform might not be a very good decision because it may lead to a disaster due to its high latency and also the vehicle may misbehave in case if it is driven in a low network connectivity area. So here the Edge computing platform comes to its rescue and makes it efficient and reliable due to its on-device computational capacity that does on rely on the network connectivity.

IV. OBSTACLES AND OPPORTUNITIES

The very important issues to be addressed as obstacles in edge computing is naming, programmability, security, privacy, and data abstraction.

The first issue we address here is the programmability which might cause a problem when the edge devices are heterogeneous. So, this might cause a lot of inter-platform dependency issues irrespective of the model's efficiency. The computing stream concept which is a serial of functions/computations applied on the data along the data propagation path to make the necessary computations that are platform dependent. Security and privacy is an opportunity for the edge computing field to take advantage of the other cloud-based applications. The naming issues arise because of the large number of heterogeneous edge nodes that are present in the network and they may contain different architectures irrespective of the cloud system's capability and too much sparsity with lots of unconventional naming systems it becomes tough to address all the nodes in a simple way. So, it is important to follow a proper naming schema or system to eliminate the naming problems caused due to the highly heterogeneous edge devices in the network.

There are some platforms that help the developers to make intelligent application by deploying Deep learning models in the edge devices. Platforms like tensorflow mobile, tensorflow lite, ncnn, caffe, caffe2, etc have helped application perform intelligent calculations with the help of Deep Learning models embedded within them. Tensorflow Mobile is a platform introduced by Google to make such intelligent applications. It supports iOS, Android and also the raspberry pi platforms for embedding the deep learning models in the application

and the use of protocol buffers to serialize the objects helps the platform to work better. But, this platform is about to be deprecated and it will be replaced with the tensorflow lite platform which is a successor of tensorflow mobile. The tensorflow lite has a lot of advantages over tensorflow mobile because it makes use of the object serialization protocol called flat buffers. Flat buffers invented by Google for object serialization helps to reduce the computational time, power consumption and model size by eliminating the process of unpacking and parsing the data. There are platforms like caffe developed by Berkeley, ncnn developed by Tencent, caffe2 developed by Facebook are also highly useful platforms that can be used for deploying deep learning and making intelligent applications in mobile and IoT platform.

V. CONCLUSION

Recently there are many services that are being transported from cloud to edge devices for processing data in shorter response time and improved reliability. There are several places where the edge computing can flourish and make devices work better and solve many important problems in an efficient way. The collaborative edge can connect the user to the could both physically and logically make way for many new applications to flow in with a lot of advantages and better options for solving user problem and queries near real time.

REFERENCES

[1] H. Sundmaeker, P. Guillemin, P. Friess, and S. Woelfflé, "Vision and challenges for realizing the Internet of things," vol. 20, no. 10, 2010.

[2] Challenges and Opportunities in Edge Computing Blesson Varghese, Nan Wang, Sakil Barbhuiya, Peter Kilpatrick and Dimitrios S. Nikolopoulos School of Electronics, Electrical Engineering, and Computer Science Queen's University Belfast, UK.

[3] AI Benchmark: Running Deep Neural Networks on Android Smartphones Andrey Ignatov Qualcomm, Inc. andrey@vision.ee.ethz.ch

[4] Boeing 787s to Create Half a Terabyte of Data Per Flight, Says Virgin Atlantic. Accessed on Dec. 7, 2016.

[5] Self-Driving Cars Will Create 2 Petabytes of Data, What are the Big Data Opportunities for the Car Industry? Accessed on Dec. 7, 2016.

[6] B.-G. Chun, S. Ihm, P. Maniatis, M. Naik, and A. Patti, "CloneCloud: Elastic execution between mobile device and cloud," in Proc. 6th Conf.Comput. Syst., Salzburg, Austria, 2011.