RESEARCH ARTICLE OPEN ACCESS

Big Data- fitting in the shoes of Internet of Things

Ravi Gorli
Department of Computer Science and Engineering
GIT, Gitam University
India

ABSTRACT

Big data, concept denotes data which is huge due to increase in volume, variety from different sources. Considering a general example of Facebook and Twitter where data arises from chats, status, pictures, likes and tweets. For every second there is increase in users along with the data. It has increasing from millions towards trillions, If we consider the developing technology i.e., IoT where the world is moving towards smart technologies and inventions in every sector for making the world more comfortable and efficient, So here the Smart technology which comes as smart home, wastage management, vehicles, transport and so on, So here the question arises that what is the role of Big data in IoT? The concept of IoT is sensors and actuators. A sensor is one which is going to retrieve the input using different methods depending on the type of input, there are different sensors which will sense the input by using different characteristics, so here the sensor will sense the input for every fraction of second and forward to processor for processing with an actuator. It should be considered and maintained at the central server, where the volume is increasing and the data retrieved by central server is of different types with different structures. Here, the problem arises with the limit for every storage aspect and retrieving data in a more efficient way. So here comes the concept for Big data in IoT which is going to follow different procedures for storing the data in different locations along with criteria for maintenance of structured and unstructured data, demonstrating the efficient methodology in retrieving only the relevant data from IoT using the core concepts of Big data, There by justifying our title.

Keywords:- Actuator, Big data, IoT, sensor, Server

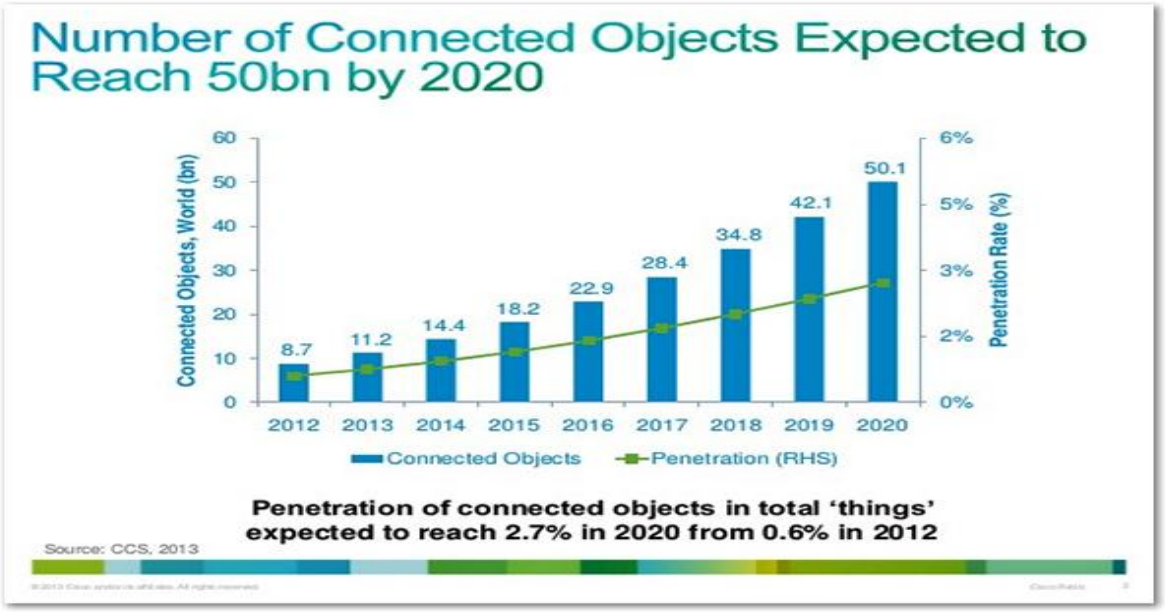
I. INTRODUCTION

Big Data – fitting in the shoes of Internet of Things, Big data is the data where the word reflects data in terms of huge volumes, Big Data is defined as big means huge in terms of volume and data is information represented in bits (Huge Volume of Information in Bits). As data is collected from multiple input devices, it should be continuously, processed and stored in a large memory [4]. Big Data is characterized by four V's i.e. Volume, Variety, Velocity, and Veracity. To make it clean and reliable, big data comes in large amounts (volume), is a mixture of structured and unstructured information (variety), arrives at speed (velocity), and can be accurate (veracity). Big Data is of a hot topic then [5].

II. INTERNET OF THINGS (IOT)

The Internet of Things (IoT) is a system of interrelated computing devices [1], mechanical and digital machines, objects, animals or people that are provided with indenters and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction. A thing, in the Internet of Things, can be a person with a heart monitor implant, a farm animal with a biochip transponder, an automobile that has built-in sensors to alert the driver when tire pressure is low -- or any other natural or man-made object that can be assigned an IP address and provided with the ability to transfer data over a network. IoT has evolved from the convergence of wireless technologies, micro-electromechanical systems (MEMS), micro services and the internet. The convergence has helped tear down the silo walls between operational technologies (OT) and information technology (IT), allowing unstructured machine-generated data to be analysed for insights that will drive improvements.

The below figure shows the increase in the number of objects from 2010 to 2020 in a vast increase of their need [2].



Kevin Ashton [2], confounder and executive director of the Auto-ID Centre at MIT, first mentioned the Internet of Things in a presentation he made to Proctor & Gamble in 1999. Here's how Ashton explains the potential of the Internet of Things:

"Today computers -- and, therefore, the Internet -- are almost wholly dependent on human beings for information. Nearly all of the roughly 50 petabytes (a petabyte is 1,024 terabytes) of data available on the Internet were first captured and created by human beings by typing, pressing a record button, taking a digital picture or scanning a bar code[2].

The problem is, people have limited time, attention and accuracy -- all of which means they are not very good at capturing data about things in the real world. If we had computers that knew everything there was to know about things -- using data they gathered without any help from us -- we would be able to track and count everything and greatly reduce waste, loss and cost. We would know when things needed replacing, repairing or recalling and whether they were fresh or past their best.", Practical applications of IoT technology can be found in many industries today, including precision agriculture, building management, healthcare, energy and transportation. Connectivity options for electronics engineers and application developers working on products and systems for the Internet of Things

III. DATA STORAGE

When we talk about IoT, [5] one of the first things that comes to mind is huge, continuous stream of data hitting companies' data storage. Data centres must be equipped to handle this additional load of heterogeneous data. In response to this direct impact on big data storage infrastructure, many organizations are moving toward the Platform as a Service (PaaS) model instead of keeping their own storage infrastructure, which would require continuous expansion to handle the load of big data. PaaS is a cloud-based, managed solution that provides scalability, flexibility, compliance, and a sophisticated architecture to store valuable IoT data. Cloud storage options include private, public, and hybrid models. If companies have sensitive data or data that is subject to regulatory compliance requirements that require heightened security, a private cloud model might be the best fit. Otherwise, a public or hybrid model can be chosen as storage for IoT data.

IV. BIG DATA TECHNOLOGIES

When selecting the technology stack for big data processing, the tremendous influx of data that the IoT will deliver must be kept in mind. Organizations will have to adapt technologies to map with IoT data. Network, disk, and compute power all will be impacted and should be planned to take care of this new type of data. From a technology perspective, the most important thing is to receive events from IoT-connected devices. The devices can be connected to the network using Wi-Fi, Bluetooth, or another technology, but must be able to send messages to a broker using some well-defined protocol. One of the most popular and widely used protocols is Message Queue Telemetry Transport (MQTT). Mosquitto is a popular open-source MQTT broker.

Once the data is received, the next consideration is the technology platform to store the IoT data. Many companies use Hadoop and Hive to store big data. But for IoT data, No SQL document databases like Apache Couch DB are more suitable because they offer high throughput and very low latency. These types of databases are schema-less, which supports the flexibility to add new event types easily. Other popular IoT tools are Apache Kafka for intermediate message

brokering and Apache Storm for real-time stream processing.

V. BIG DATA ANALYTICS

IoT and big data basically are two sides of the same coin. Managing and extracting value from IoT data is the biggest challenge that companies face. Organizations should set up a proper analytics platform/infrastructure to analyse the IoT data. And they should remember that not all IoT data is important [4].

A proper analytics platform should be based on three parameters: performance, right-size infrastructure, and future growth. For performance, a bare-metal server, a single-tenant physical server dedicated to a single customer, is the best fit. For infrastructure and future growth, hybrid is the best approach. Hybrid deployments, which consist of cloud, managed hosting, colocation, and dedicated hosting, combine the best features from multiple platforms into a single optimal environment. Managed Service Providers (MSPs) are also working on their platforms to handle IoT data. MSP vendors are typically working on the infrastructure, performance, and tools side to cover the entire IoT domain [3].

An IoT device generates continuous streams of data in a scalable way, and companies must handle the high volume of stream data and perform actions on that data. The actions can be event correlation, metric calculation, statistics preparation, and analytics. In a normal big data scenario, the data is not always stream data, and the actions are different. Building an analytics solution to manage the scale of IoT data should be done with these differences in mind.

The growth of the IoT heralds a new age of technology, and organizations that wish to participate in this new era will have to change the way they do things to accommodate new data types and data sources. And these changes likely are just the beginning. As the IoT grows and businesses grow with IoT, they will have many more challenges to solve.

So, are IoT and Big Data are different things or two sides same coin?[6]

Let's take two instances and try to identify which statement defines IoT and which one is about big data.

1. Over 204 million emails are sent every minute. This generates 1.8 million likes on Facebook and around 278 thousand tweets.
2. With the rapid growth in amount data generated in recent years' industry is expected to reach $54.3 billion by 2017.

Both can mean IoT and/or Big Data. It is difficult to judge. These two concepts almost merge with each other. The end result is, the digital advancements that we see around us. Big data is what IoT means.

VI. IOT + BIG DATA ANALYTICS

Reports are in galleons on how smartphone use has increased[10] manifold times in recent years. Every person has a smartphone and an internet connection. Apps, these days, are all about letting you do things on-the-go.

We have to understand one basic fact. IoT has generated huge amounts of data. They are of high quality and extremely sensitive data. So, what next?

Not all data are required. How will you do the filtering? How will you analyse the data? How will you dig in deeper to understand how this data can benefit?

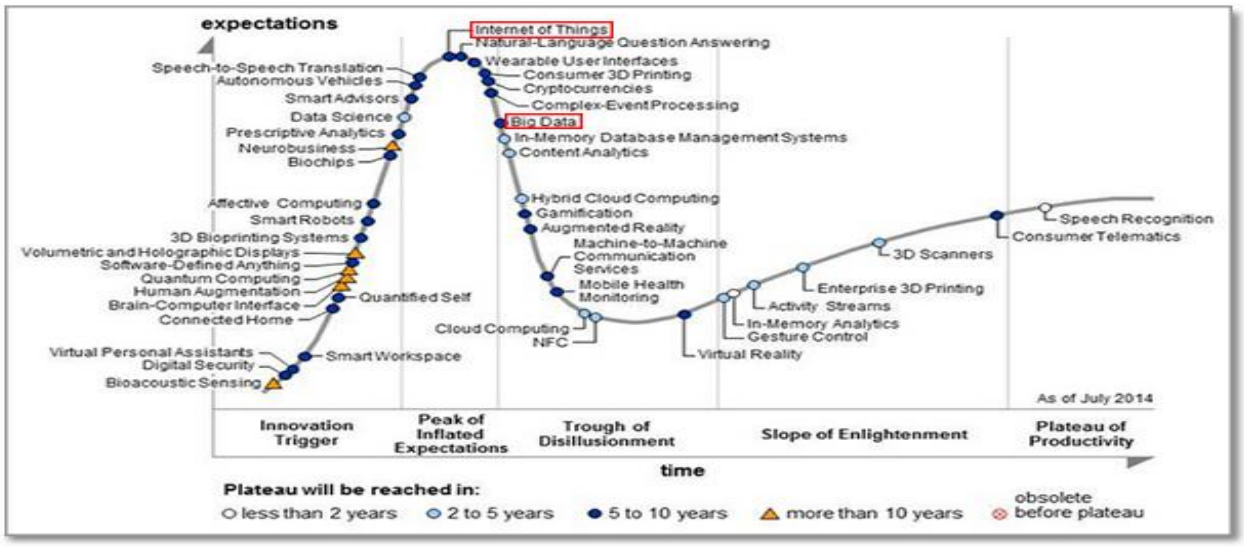
Here comes the role of analytics. Big data analytics is the answer to all these questions. Only when you are able to filter out data and process it as per your needs, you get to reap the benefits of having data at your disposal.

"The critical challenge is using this data when it is still in motion – and extracting valuable information from it". [7]

1. Every minute, we send 204 million emails, generate 1.8 million Facebook likes, send 278 thousand tweets, and upload 200 thousand photos to Facebook. *Is this statement about big data or the Internet of Things?*
2. 12 million RFID tags (used to capture data and track movement of objects in the physical world) were sold in 2011. By 2021, it's estimated this number will increase to 209 billion as [*big data or the Internet of Things?*] takes off.
3. The boom of [*big data or the Internet of Things?*] will mean that the amount of devices that connect to the internet will rise from about 13 billion today to 50 billion by 2020.

The [*big data or the Internet of Things?*] industry is expected to grow from US$10.2 billion in 2013 to about US$54.3 billion by 2017. [8]

"To make the Internet of Things useful, we need an Analytics of Things. This will mean new data management and integration approaches, and new ways to analyse streaming data continuously".[4]



The above figure shows the evaluation of technologies day by day here the red block shows the place of IoT and big data in this generation [7].

The IoT will massively increase the amount of data available for analysis by all manner of organisations. However, there are significant barriers to overcome before the potential benefits are fully realised. The problem is, people have limited time, attention and accuracy -- all of which means they are not very good at capturing data about things in the real world. If we had computers

That knew everything there was to know about things -- using data they gathered without any help from us -- we would be able to track and count everything and greatly reduce waste, loss and cost. We would know when things needed replacing, repairing or recalling and whether they were fresh or past their best.

Data was getting seriously big even before IoT[11] entered the picture. If big data is “Heart” of advancement [3]. IoT is the “Soul”. IoT represents a uniform platform for human interaction with the physical world, and it will turn Internet into a rare medium.. So, basically the future will be a connected life. Here, when sensors are attached to variety of things in IoT, these sensors can take multitude of possible measurements; this is when Big Data enters into the zone. These are also the ones to be needed in IoT in order to handle the huge trillions of data. Imagine one sensor senses and sends every second it may be small to handle, if it goes to minutes, hours, days and years and also increase in the sensors from different connectivity’s, then think how about the cycle repeats for multiple sensors at a time, so it ultimately needs a huge storage and maintenance. This is why Big Data and IoT handshake each other.

VII. CONCLUSION

In this paper we have given the introduction of IoT and Big data and their roles in different technologies with their wide spread of use in this generation, Coming to the title of Big data fitting in the shoes of IoT, Here this is explained with different criteria taking examples in real time, By the end of this paper we have given the need for IoT and Big data for handshaking their hands with their need to each other. The view is given with consideration information from different websites. In next to this the actual implementation of their roles with a generalized application is going to be explained in the following paper next to this.

REFERENCES

[1]. Ravi Gorli , "Interlinking OF IoT, Big data, Smart Mobile app with Smart Garbage Monitoring", International Journal of Computer Sciences and Engineering, Volume-05, Issue-01, Page No (0-0), Jan -2017, E-ISSN: 2347-2693

[2]. J. P. Conti, “The Internet of Things”, Communications Engineer, 2006, pp.20-25.

[3]. https://www.edgefx.in/different-types-of-sensors-with-applications/

[4]. http://searchbusinessanalytics.techtarget.com/essentialguide/IoT-analytics-guide-Understanding-Internet-of-Things-data

[5]. http://www.datamation.com/applications/why-big-data-and-the-internet-of-things-are-a-perfect-match.html

[6]. http://www.zdnet.com/article/the-internet-of-things-and-big-data-unlocking-the-power

[7]. https://www.analytixlabs.co.in/blog/2016/10/08/iot-and-big-data-they-work-in-tandem/
[8]. http://data-informed.com/the-impact-of-internet-of-things-on-big-data/
[9]. https://www.linkedin.com/pulse/iot-big-data-scientists-bill-mccabe
[10].http://www.sas.com/en_us/insights/articles/big-data/big-data-and-iot-two-sides-of-the-same-coin.html
[11].http://www.zdnet.com/article/the-internet-of-things-and-big-data-unlocking-the-power/