

# Overview of Big Data with IOT

Win Myat Thuzar<sup>1</sup>, Mie Mie Aung<sup>2</sup>, Thida San<sup>3</sup>, Thae Thae Han<sup>4</sup>  
<sup>1</sup>Information Technology Supporting and Maintenance Department, University of Computer Studies (Meiktila), Myanmar  
<sup>2,3,4</sup>Information Technology Supporting and Maintenance Department, University of Computer Studies (Meiktila), Myanmar  
[winmyatthuzar8@gmail.com](mailto:winmyatthuzar8@gmail.com)

## Abstract:

*Today, Internet of Things (IoT) is involved in our daily life, such as health monitoring, industry, urban planning and home automation and so on. IoT devices generate continuous streams of data in a scalable manner. As the data is turned and crossed over terabytes and leading to petabytes, there should be a solution to manage the overwhelming increase in data. Big data is the solution for the data problem. Big data is one of the future's data dream. Unlimited amount of data can be stored in a secured manner by using big data. So, the demand for Big Data is increasing more. As IoT and Big Data are two trends in the present era, combining those will really create a technical revolution for the future generations. In this paper, we are going to present various scenarios of using big data with IOT.*

## Keywords

*IOT, Big Data, Internet of Things, Big Data Analytics, Machine to machine leaning (M2M)*

## 1. Introduction

Internet of Things (IoT) can be considered as the future of internet applications which involves machine to machine learning (M2M). The applications of IoT spread across various domains like health care, retail management and disaster management. In present era IoT is the added essence in development of technology in smart way. IoT is generally documented with sensors and actuators to operate things without any physical contact. Imagine a world where everything is interconnected through IOT, where things can automatically get operated and dropped without any human intervention like door lock, tube lights, vehicles etc.

The data produced by the IOT is growing rapidly because of the large scale development of various applications. The actionable intelligence can be derived through fusion of Big Data and real time analytics with IoT. Big Data and IoT can be viewed as two sides of a coin. The connection between Big

Data and the objects on internet benefits of IoT can be easily reaped.

## 2. Internet of Things (IoT)

The Internet of Things (IoT) refers the ever-growing network of physical objects that feature an IP address for internet connectivity, and the communication that occurs between these objects and other internet-enable devices and systems. In simple word, Internet of Things (IoT) is an ecosystem of connected physical object that are accessible through the internet. It is also refer to as Machine-to-Machine (M2M), Skynet or Internet of Everything.

The real value of Internet of Things is about creating smarter products, delivering intelligent insights and providing new business outcomes. As millions of devices get connected, internet of things will trigger a massive inflow of Big Data. The key challenge is visualizing and uncovering insights from various types of data (structured, unstructured, images, contextual, dark data, real-time) and in context of applications.

## 3. Big Data

Big Data is massive set of unstructured or semi-structured data from web traffic, social media, sensor etc. It has petabytes, exabytes of data. Big data can get information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices etc. Big data often characterized by the 3V's, Volume, Velocity and Variety. Volume refers to the amount of data, velocity refers to the speed of data processing and variety refers to the number of types of data. Huge amount of data are generated from multiple different kinds of IoT devices. Big data are operated to produce output data by using multiple platforms such as business intelligence and analytics application, data visualization, machine learning, secret sause algorithm, pan-enterprise search and so on.

### 3.1. Role of Big Data in IoT

The big data technologies and initiatives are rising to analyse this data for gaining insights that can help in making strategic decisions. The concept evolved at the beginning of 21<sup>st</sup> century, and every technology giant is now making use of big data technologies. Big data refers to vast and voluminous data sets that may be structured or unstructured. This massive amount of data is produced every day by businesses and users. Big data analytics is the process of examining the large data sets to underline insights and patterns. The Data analytics field in itself is vast.

IoT is acting as a major source for data that are held for analysis purpose. So, the role of big data in IoT comes into the important point. The internet of things is going to trigger a massive set of big data because sensors spread across almost every industry. For the future of big data analytics, IoT will have the biggest impact. Big data analytics is analyzing IoT generated data to improve decision making. The role of big data in IoT is to process a large amount of data in real time basis and storing them using different storage technologies.

Although big data and IoT almost sounds similar, they are different actually. Big data is basically data and a lot of it. IoT is connectivity. Both are useful and should work hand-in-hand. Big data is basically all the data that is available everywhere and that is extracted, compiled and stored in a huge storage device. The IoT is actually like the sources that help in accumulating all that data through objects and sensors. To bring in the technical point of view, IoT is more like collecting all the data, analyzing their patterns, and taking into consideration and new insights or points of view that might come up.

### 3.2. Big Data Processing Technology

In general, big data technology domain can be categorized into the following four parts of layer; data source layer, data storing layer, data processing/analysis layer, and data output/ presentation layer [2].

Data Source Layer incorporates structured, unstructured and/or semi-structure data captured from transitions, interactions and observations system such as Facebook, twitter, IoT. This is wide variety of data, coming in huge volume with high velocity has to be seamlessly merged and consolidated so that the analytics engines, as well as the visualization tools, can operate on it as one single big data set.

Big data storage layer are needed all kinds of distributed computing technologies such as Hadoop Distributed File System (HDFS) and NoSQL. HDFS can store the collected huge amount data into a

distributed storage system. HDFS can monitor and control all distributed data on its hierarchy using HTTP protocol. NoSQL is highly adopted for manipulating unstructured data in a different manner from the conventional relation based database architecture. And Big Table can be used for the big data storing context.

Data analysis layer includes the technologies to analyze unstructured and informal text mining. Such analysis technologies are based on NLP, Natural Language Processing. They can find meaningful keywords, relations, or patterns from the flat and complicated texts. In parallel computing resources, a software framework for processing big data is MapReduce. It provides an effective text analysis results. For terabytes level data, Big Query, one of online OLAP systems, can provide real-time analyzing facility.

Big data presentation layer include data manipulating tools for unprofessional analysts and data visualization technique. Info-Graphic tools in an intuitive and impressive way for the analyzed results. Representative analysis tool, R provides statistical computation, modelling, cutting edge data mining methods. This classification and their flow are illustrated in Figure 1.

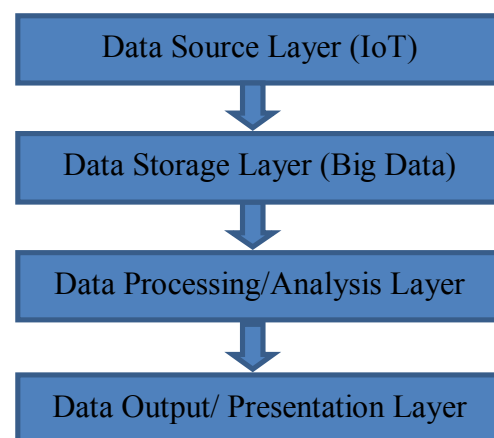
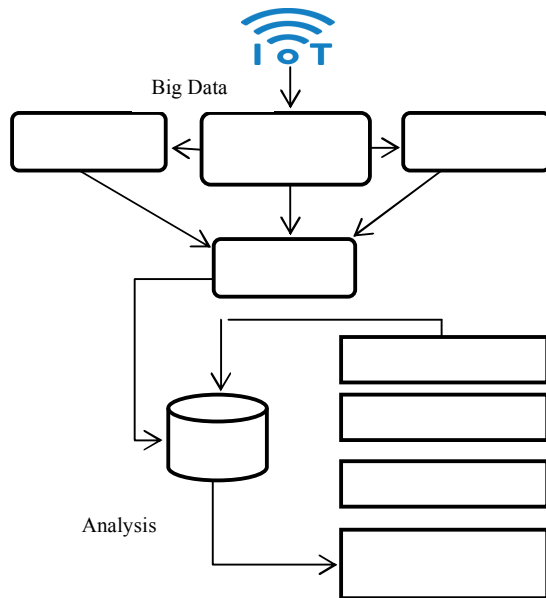


Figure 1. Framework of Big Data

IoT big data processing follows four sequential step; [3]

1. A large amount of unstructured data is generated by IoT devices which are collected in big data system. This IoT generated big data largely depend on their 3V factor.
2. In the big data system which is basically a shared distributed database, the huge amount of data is stored in big data files.

3. Analyzing the stored IoT big data using analytic tool like Hadoop MapReduce or Spark.
  4. Generating the reports of analyzed data.
- Processing steps are as shown in Figure 2.



**Figure 2. IOT Big Data Processing**

### 3.3. Data Storage Technology of IoT Data

Combination of data storage technology is needed because IOT data storage is complex. Most of these data are semi-structured or unstructured. Hence, looking for the flexibility to best fit the data in database is needed. Relational databases have the shortcoming as it is a poor fit for semi-structured or unstructured data and also can't easily incorporate new data types. NoSQL fills these gaps as its data model maps better with the needs.

NoSQL databases are a popular choice for IoT data. NoSQL is used for analytics because it supports high throughput and low latency. Moreover, their schema-less approach is flexible, allowing new types of data to be added dynamically. NoSQL databases used for IoT data include Couchbase, Apache Cassandra, Apache CouchDB, MongoDB and Apache HBase. Big data can be considered as two perspectives, operational data and analytical data. Operational data are stored in operational databases and mostly deals with online live data. Analytical data is a large amount of data to collect insights from it. Hence, NoSQL database can management in a better way to storage big data with an operational database.

### 4. Relationship between Big Data and IoT

In order to function, the IoT needs to collect and store a large amount of data. But what is the use of all these collected data? It is important to process and collect large volumes of real time data (eg from sensors, traffic lights, and any connected IoT device) to make the IoT work properly, and to improve safety and productivity, both in any field and for any type of connected object. When big data analyst is hired, nosql database and IoT data that can be done for integrated systems between big data.

It's not just that there is the only interdependent relation between big data and IoT. As they help each other, in addition to that they hugely impact each other. Fact is the more the IoT grows it will place more demand on businesses regarding big data capabilities. As the IoT generated data is increasing at a huge rate, conventional data storage technology is already being pushed to its limits. As a result, it demands more advanced and innovative storage solutions to handle these growing workloads resulting in updating the infrastructure of an organization's big data storage. Similarly, the IoT big data combined applications accelerate the scope of research in both the fields. So, both IoT and big data technology carry inter-dependency and need further development.

### 5. Literature Survey

In[1], author presents the main foundations of Big Data deployment in smart homes. The proposed architecture can be applied to different smart cities applications. Author describes multiple objectives of smart cities based on big data analysis. This approach is to analyze potential of the applicability of IoT techniques to provide profitable services of smart cities, such as the management of the energy consumption in smart buildings, and the detection of travel profiles in smart transport. Author [2] said that Millions of things are connected through IoT, and these contain enormous of data. The Data processing and transmission is a hard task. This paper mainly aims at role of Big Data in IoT and discuss about protocols and structures. In [5], author proposed an IoT and big data based system for restaurants and described its benefits, challenges and opportunities. It also explains how the system can be integrated to the operations of restaurant. In [4] paper, author researched about R association rules with market basket dataset. Author found the relationship between things depending on what customers buy. By using R, it can be analyze the product easily in many ways.

## 6. Conclusion

As per the study, by the year 2020, around 4.4 trillion GB of data will be generated through IoT and more than ten billion of sensors and devices will be connected to the internet. Furthermore, all of these devices will gather, analyze, share and transmit data in real time. Hence, IoT device would not hold the functionalities and capabilities without the data. Role of big data analytics become challenges associated with the successful deployment of IoT. So, combination of IoT and Big Data is needed to be reliable and strong smatter lives in future.

## 7. References

- [1] A.Rajasekar, J.Samyuktha and S. Meharaj, "Applicability of Big Data Techniques to Smart Home", IJESC Volume 8 Issue No.3, 2018
- [2] Do-Eun Cho, Si Jung Kim, and Sang-Soo Yeo, "Double Privacy Layer Architecture for Big Data Framework", International Journal of Software Engineering and Its Applications Vol. 10, No. 2 (2016), pp. 271-278
- [3] <https://learning.naukri.com/articles/big-data-and-iot/>
- [4] Jihyun Song, Kyeongjoo Kim and Minsoo Lee, "A Big Data Analysis and Mining Approach for IoT Big Data", International Journal of Advances in Computer Science and Technology, Volume 7 No.1, January 2018
- [5] Md. Muminur Rahman, "IoT & Big Data Based Applications for Restaurant Questions of Opportunities, Challenges, Benefits & Operations", ResearchGate, March 2017
- [6] Umar Ahsan, Abdul Bias "A Review on Big Data Analysis and internet of Things" IEEE journals.