# 402C-INTERNET OF THINGS

**(MCA-S.V.UNIVERSITY,TIRUPATI)**

**IV-SEMESTER**

**STUDY MATERIAL**

**PREPARED BY: MISS.C.YAMINI M.C.A,**

**Department of Computer Science,**

**KMM INSTITUTE OF POST GRADUATION STUDIES.**


**KMM INSTITUTE OF POST GRADUATION STUDIES**

**Ramireddipalle, TIRUPATI-517102**

**UNIT-1**

**1.Internet of Things (IoT)** is the networking of physical objects that contain electronics embedded within their architecture in order to communicate and sense interactions amongst each other or with respect to the external environment. In the upcoming years, IoT-based technology will offer advanced levels of services and practically change the way people lead their daily lives. Advancements in medicine, power, gene therapies, agriculture, smart cities, and smart homes are just a very few of the categorical examples where IoT is strongly established.

***IoT is network of interconnected computing devices which are embedded in everyday objects, enabling them to send and receive data.***

Over 9 billion 'Things' (physical objects) are currently connected to the Internet, as of now. In the near future, this number is expected to rise to a whopping 20 billion.

**Main components used in IoT:**

- **Low-power embedded systems:** Less battery consumption, high performance are the inverse factors that play a significant role during the design of electronic systems.
- **Sensors :** Sensors are the major part of any IoT applications. It is a physical device that measures and detect certain physical quantity and convert it into signal which can be provide as an input to processing or control unit for analysis purpose.

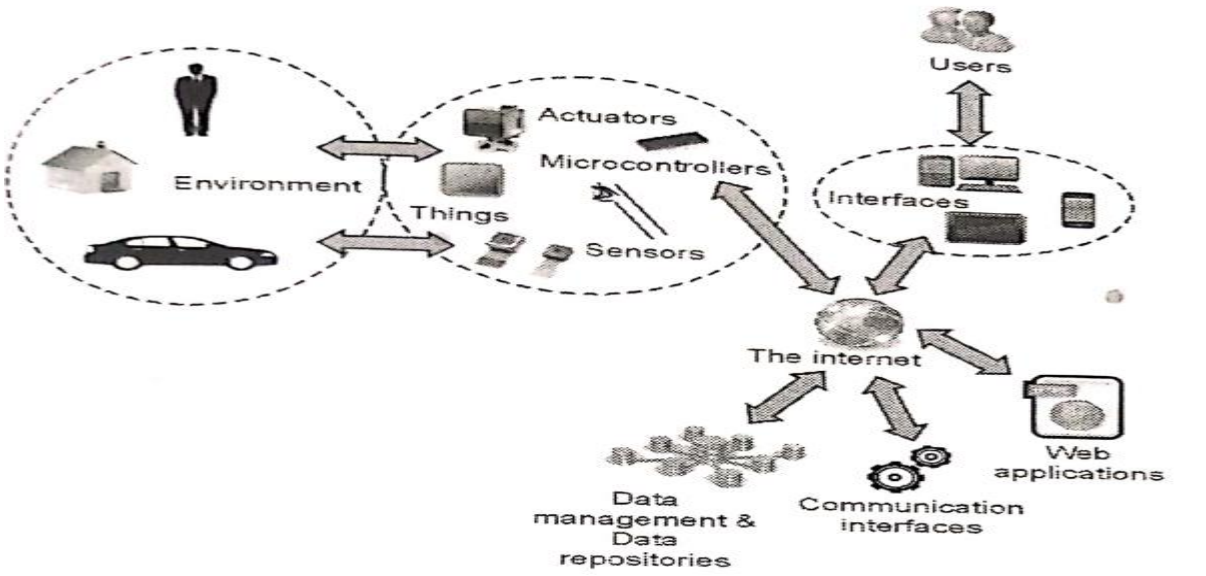
1. Different types of Sensors :
2. Temperature Sensors
3. Image Sensors
4. Gyro Sensors
5. Obstacle Sensors
6. RF Sensor
7. IR Sensor
8. MQ-02/05 Gas Sensor
9. LDR Sensor
10. Ultrasonic Distance Sensor

- **Control Units :** It is a unit of small computer on a single integrated circuit containing microprocessor or processing core, memory and programmable input/output devices/peripherals. It is responsible for major processing work of IoT devices and all logical operations are carried out here.
- **Cloud computing:** Data collected through IoT devices is massive and this data has to be stored on a reliable storage server. This is where cloud computing comes into play. The data is processed and learned, giving more room for us to discover where things like electrical faults/errors are within the system.
- **Availability of big data:** We know that IoT relies heavily on sensors, especially in real-time. As these electronic devices spread throughout every field, their usage is going to trigger a massive flux of big data.
- **Networking connection:** In order to communicate, internet connectivity is a must where each physical object is represented by an IP address. However, there are only a limited number of addresses available according to the IP naming. Due to the growing number of devices, this naming system will not be feasible anymore. Therefore, researchers are looking for another alternative naming system to represent each physical object.

**Working with IoT Devices :**

- *Collect and Transmit Data :* For this purpose sensors are widely used they are used as per requirements in different application areas.
- *Actuate device based on triggers produced by sensors or processing devices :* If certain condition is satisfied or according to user's requirements if certain trigger is activated then which action to performed that is shown by Actuator devices.

- *Receive Information :* From network devices user or device can take certain information also for their analysis and processing purposes.
- *Communication Assistance :* Communication assistance is the phenomena of communication between 2 network or communication between 2 or more IoT devices of same or different Networks. This can be achieved by different communication protocols like : MQTT , Constrained Application Protocol, ZigBee, FTP, HTTP etc.




**2.Characteristics of IoT:**

- Massively scalable and efficient
- IP-based addressing will no longer be suitable in the upcoming future.
- An abundance of physical objects is present that do not use IP, so IoT is made possible.
- Devices typically consume less power. When not in use, they should be automatically programmed to sleep.
- A device that is connected to another device right now may not be connected in another instant of time.
- Intermittent connectivity – IoT devices aren't always connected. In order to save bandwidth and battery consumption, devices will be powered off periodically when not in use. Otherwise, connections might turn unreliable and thus prove to be inefficient.
- ***Desired Quality of any IoT Application :***
- ***Interconnectivity***

*It is the basic first requirement in any IoT infrastructure. Connectivity should be guaranteed from any devices on any network then only devices in a network can communicate with each other.*

- ***Heterogeneity***

*There can be diversity in IoT enabled devices like different hardware and software configuration or different network topologies or connections but they should connect and interact with each other despite of so much heterogeneity.*

- ***Dynamic in nature***

*IoT devices should dynamically adapt themselves to the changing surroundings like different situation and different prefaces.*

- ***Self adapting and self configuring technology***

*For example surveillance camera. It should be flexible to work in different weather conditions and different light situations (morning, afternoon, or night).*

- ***Intelligence***

*Just data collection is not enough in IoT, extraction of knowledge from the generated data is very important. For example, sensors generate data, but that data will only be useful if it is interpreted properly. So intelligence is one of the key characteristics in IoT. Because data interpretation is the major part in any IoT application because without data processing we can't make any insights from data . Hence big data is also one of the most enabling technology in IoT field.*

- ***Scalability***

*The number of elements (devices) connected to IoT zone is increasing day by day. Therefore, an IoT setup should be capable of handling the expansion. It can be either expand capability in terms of processing power, Storage, etc. as vertical scaling or horizontal scaling by multiplying with easy cloning*

- ***Identity***

*Each IoT device has a unique identity (e.g., an IP address). This identity is helpful in communication, tracking and to know status of the things. If there is no identification then it will directly effect security and safety of any system because without discrimination we can't identify with whom one network is connected or with whom we have to communicate. So there should be clear and appropriate discrimination technology available between IoT networks and devices.*

- ***Safety***

*Sensitive personal details of a user might be compromised when the devices are connected to the Internet. So data security is a major challenge. This could cause a loss to the user. Equipment in the huge IoT network may also be at risk. Therefore, equipment safety is also critical.*

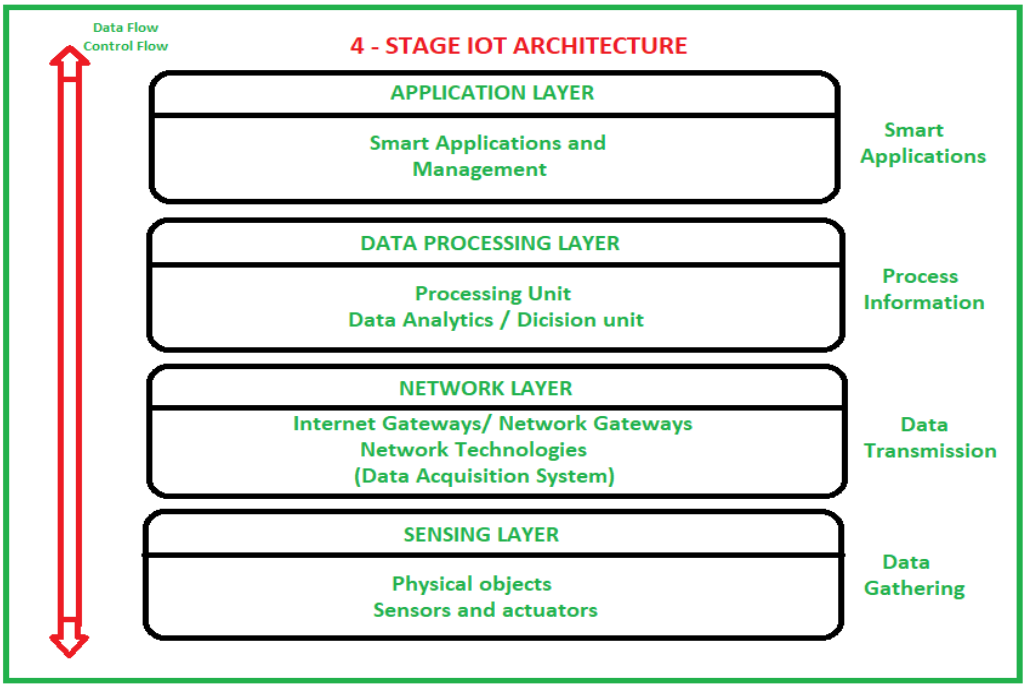
- ***Architecture***

*It should be hybrid, supporting different manufacturer's products to function in the IoT network.*

**3.Architecture of Internet of Things (IoT)**

Internet of Things (IoT) technology has a wide variety of applications and use of Internet of Things is growing so faster. Depending upon different application areas of Internet of Things, it works accordingly as per it has been designed/developed. But it has not a standard defined architecture of working which is strictly followed universally. The architecture of IoT depends upon its functionality and implementation in different sectors. Still, there is a basic process flow based on which IoT is built.

So. here in this article we will discuss basic fundamental architecture of IoT i.e., 4 Stage IoT architecture.




So, from the above image it is clear that there is 4 layers are present that can be divided as follows: Sensing Layer, Network Layer, Data processing Layer, and Application Layer. These are explained as following below.

1. **Sensing Layer –**
   Sensors, actuators, devices are present in this Sensing layer. These Sensors or Actuators accepts data(physical/environmental parameters), processes data and emits data over network.
2. **Network Layer –**
   Internet/Network gateways, Data Acquisition System (DAS) are present in this layer. DAS performs data aggregation and conversion function (Collecting data and aggregating data then converting analog data of sensors to digital data etc). Advanced gateways which mainly opens up connection between Sensor networks and Internet also performs many basic gateway functionalities like malware protection, and filtering also some times decision making based on inputted data and data management services, etc.
3. **Data processing Layer –**
   This is processing unit of IoT ecosystem. Here data is analyzed and pre-processed before sending it to data center from where data is accessed by software applications often termed as business applications where data is monitored and managed and further actions are also prepared. So here Edge IT or edge analytics comes into picture.
4. **Application Layer –**
   This is last layer of 4 stages of IoT architecture. Data centers or cloud is management stage of data where data is managed and is used by end-user applications like agriculture, health care, aerospace, farming, defense, etc.

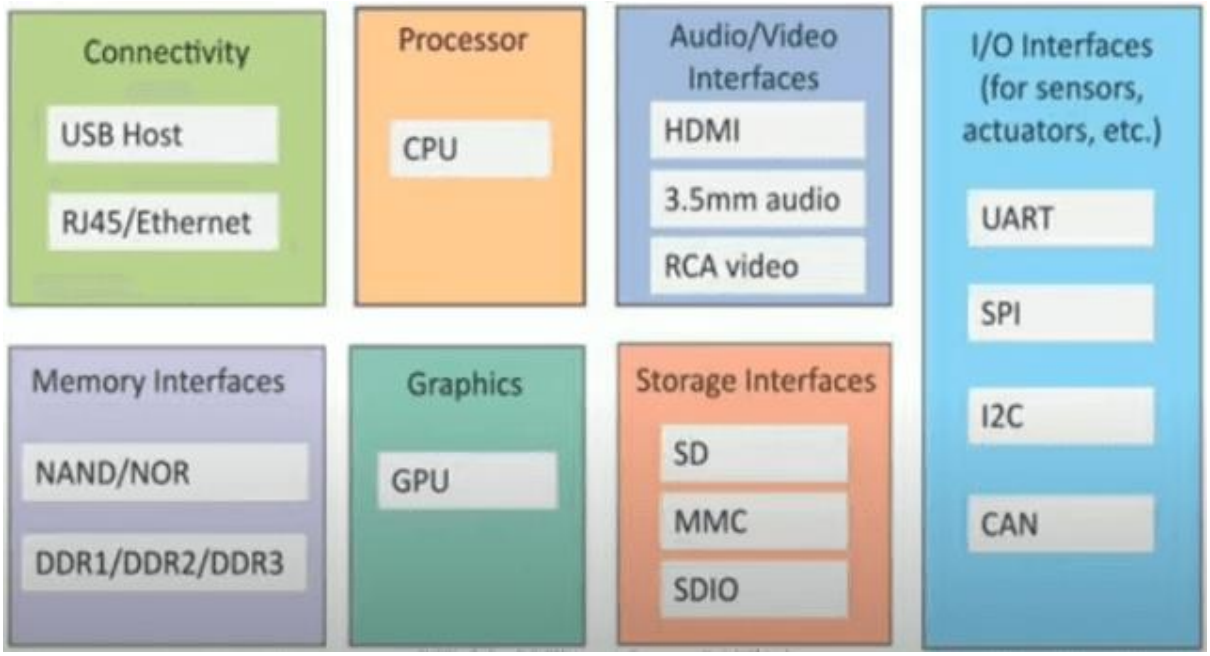
**4.Physical Design of IoT**

The **physical design** of an IoT system is referred to as the **Things/Devices** and protocols that are used to build an IoT system. all these things/Devices are called Node Devices and every device has a unique identity that performs remote sensing, actuating and monitoring work. and the protocols that are used to establish communication between the Node devices and servers over the internet.

### Things/Devices

Things/Devices are used to build a connection, process data, provide interfaces, provide storage, and provide graphics interfaces in an IoT system. all these generate data in a form that can be analyzed by an analytical system and program to perform operations and used to improve the system.

for example temperature sensor that is used to analyze the temperature generates the data from a location and is then determined by algorithms.




### Connectivity

Devices like USB hosts and ETHERNET are used for connectivity between the devices and the server.

### Processor

A processor like a CPU and other units are used to process the data. these data are further used to improve the decision quality of an IoT system.

### Audio/Video Interfaces

An interface like HDMI and RCA devices is used to record audio and videos in a system.

### Input/Output interface

To give input and output signals to sensors, and actuators we use things like UART, SPI, CAN, etc.

### Storage Interfaces

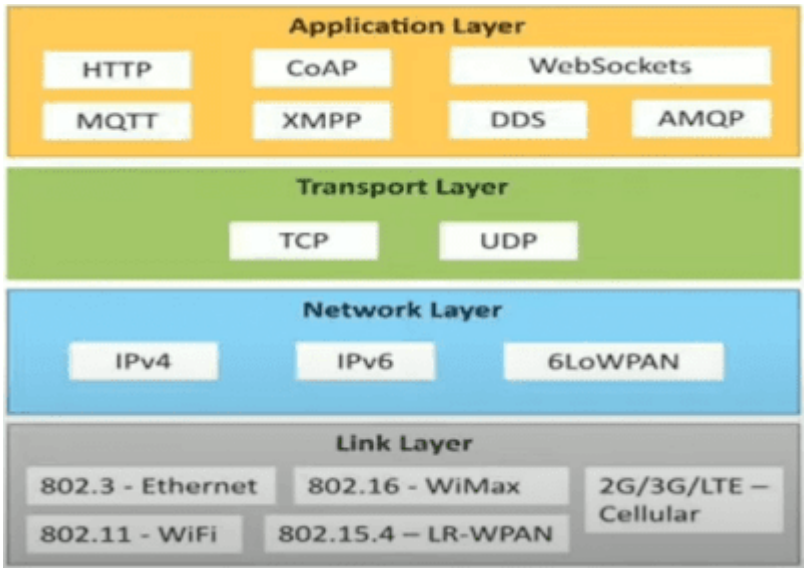
Things like SD, MMC, and SDIO are used to store the data generated from an IoT device.

Other things like DDR and GPU are used to control the activity of an IoT system.

### IoT Protocols

These protocols are used to establish communication between a node device and a server over the internet. it helps to send commands to an IoT device and receive data from

an IoT device over the internet. we use different types of protocols that are present on both the server and client-side and these protocols are managed by network layers like application, transport, network, and link layer.




### Application Layer protocol

In this layer, protocols define how the data can be sent over the network with the lower layer protocols using the application interface. these protocols include HTTP, WebSocket, XMPP, MQTT, DDS, and AMQP protocols.

#### *HTTP*

Hypertext transfer protocol is a protocol that presents in an application layer for transmitting media documents. it is used to communicate between web browsers and servers. it makes a request to a server and then waits till it receives a response and in between the request server does not keep any data between two requests.

#### *WebSocket*

This protocol enables two-way communication between a client and a host that can be run on an untrusted code in a controlled environment. this protocol is commonly used by web browsers.

#### *MQTT*

It is a machine-to-machine connectivity protocol that was designed as a publish/subscribe messaging transport. and it is used for remote locations where a small code footprint is required.

### Transport Layer

This layer is used to control the flow of data segments and handle the error control. also, these layer protocols provide end-to-end message transfer capability independent of the underlying network.

### *TCP*

The transmission control protocol is a protocol that defines how to establish and maintain a network that can exchange data in a proper manner using the internet protocol.

### *UDP*

a user datagram protocol is a part of an internet protocol called the connectionless protocol. this protocol is not required to establish the connection to transfer data.

### Network Layer

This layer is used to send datagrams from the source network to the destination network. we use IPv4 and IPv6 protocols as host identification that transfers data in packets.

### *IPv4*

This is a protocol address that is a unique and numerical label assigned to each device connected to the network. an IP address performs two main functions host and location addressing. IPv4 is an IP address that is 32-bit long.

### *IPv6*

It is a successor of IPv4 that uses 128 bits for an IP address. it is developed by the IETF task force to deal with long-anticipated problems.

### Link Layer

Link-layer protocols are used to send data over the network's physical layer. it also determines how the packets are coded and signaled by the devices.

### *Ethernet*

It is a set of technologies and protocols that are used primarily in LANs. it defines the physical layer and the medium access control for wired ethernet networks.

### *WiFi*

It is a set of LAN protocols and specifies the set of media access control and physical layer protocols for implementing wireless local area networks.

## 5.Logical Design of IoT

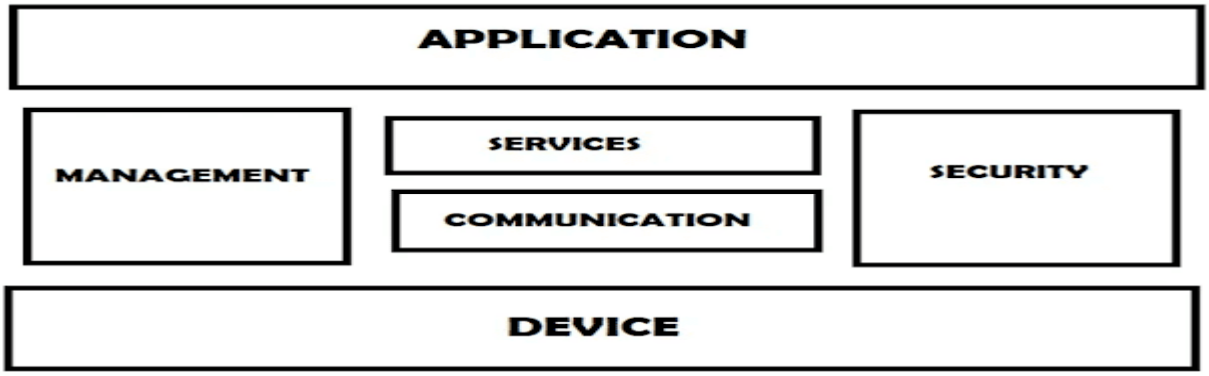
**The logical design** of an IoT system refers to an abstract representation of entities and processes without going into the low-level specifies of implementation. it uses **Functional Blocks**, **Communication Models**, and **Communication APIs** to implement a system.

### Logical Design of Internet of Things(IoT)

1. IoT Functional Blocks
2. IoT Communication Models
3. IoT Communication APIs

## IoT Functional blocks

An IoT system consists of a number of functional blocks like Devices, services, communication, security, and application that provide the capability for sensing, actuation, identification, communication, and management.




These functional blocks consist of devices that provide monitoring control functions, handle communication between host and server, manage the transfer of data, secure the system using authentication and other functions, and interface to control and monitor various terms.

### Application

It is an interface that provides a control system that use by users to view the status and analyze of system.

### Management

This functional block provides various functions that are used to manage an IoT system.

### Services

This functional block provides some services like monitoring and controlling a device and publishing and deleting the data and restoring the system.

### Communication

This block handles the communication between the client and the cloud-based server and sends/receives the data using protocols.

### Security

This block is used to secure an IoT system using some functions like authorization, data security, authentication, 2-step verification, etc.

### Device

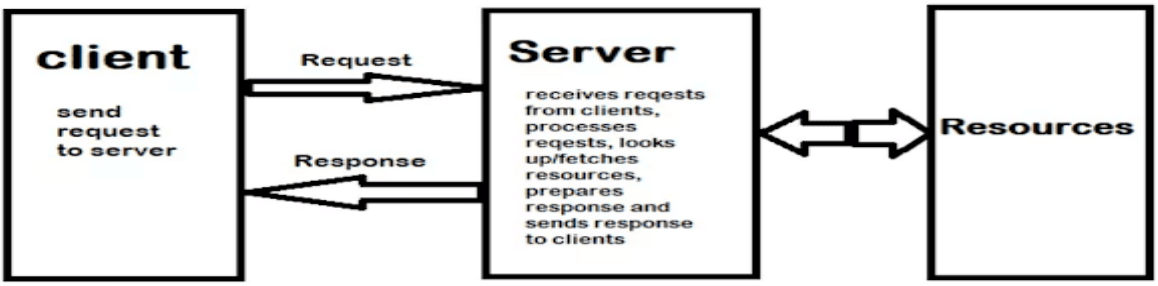
These devices are used to provide sensing and monitoring control functions that collect data from the outer environment.

## IoT Communication Models

There are several different types of models available in an IoT system that is used to communicate between the system and server like the request-response model, publish-subscribe model, push-pull model, exclusive pair model, etc.

## Request-Response Communication Model

This model is a communication model in which a client sends the request for data to the server and the server responds according to the request. when a server receives a request it fetches the data, retrieves the resources and prepares the response, and then sends the data back to the client.




**Request-Response Communication Model**

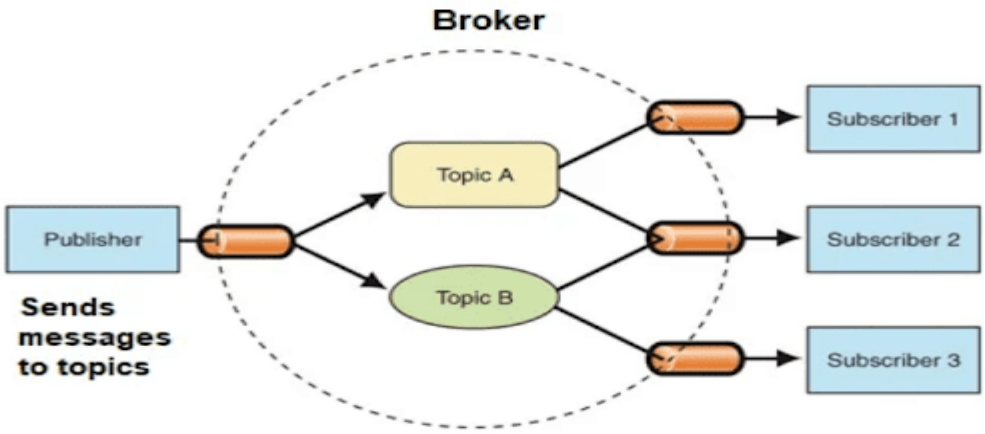
In simple terms, we can say that in the request-response model server send the response of equivalent to the request of the client. in this model, HTTP works as a request-response protocol between a client and server.

### *Example*

When we search a query on a browser then the browser submits an HTTP request to the server and then the server returns a response to the browser(client).

## Publish-Subscribe Communication Model

In this communication model, we have a broker between publisher and consumer. here publishers are the source of data but they are not aware of consumers. they send the data managed by the brokers and when a consumer subscribes to a topic that is managed by the broker and when the broker receives data from the publisher it sends the data to all the subscribed consumers.

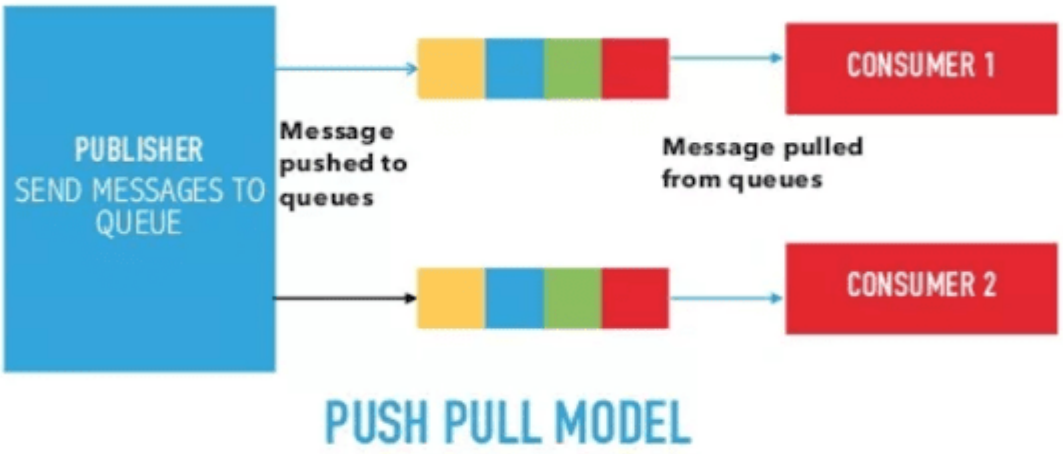



***Example***

On the website many times we subscribed to their newsletters using our email address. these email addresses are managed by some third-party services and when a new article is published on the website it is directly sent to the broker and then the broker sends these new data or posts to all the subscribers.

**Push-Pull Communication Model**

It is a communication model in which the data push by the producers in a queue and the consumers pull the data from the queues. here also producers are not aware of the consumers.




***Example***

When we visit a website we saw a number of posts that are published in a queue and according to our requirements, we click on a post and start reading it.

### Exclusive Pair Communication Model

It is a bidirectional fully duplex communication model that uses a persistent connection between the client and server. here first set up a connection between the client and the server and remain open until the client sends a close connection request to the server.


### IoT communication APIs

These APIs like REST and WebSocket are used to communicate between the server and system in IoT.

### REST-based communication APIs

Representational state transfer(REST) API uses a set of architectural principles that used to design web services. these APIs focus on the systems' resources that how resource states are transferred using the request-response communication model. this API uses some architectural constraints.

#### *Client-server*

Here the client is not aware of the storage of data because it is concerned about the server and similarly the server should not be concerned about the user interface because it is a concern of the client. and this separation is needed for independent development and updating of server and client. no matter how the client is using the response of the server and no matter how the server is using the request of the client.

#### *Stateless*

It means each request from the client to the server must contain all the necessary information to understand by the server. because if the server can't understand the request of the client then it can't fetch the request data in a proper manner.

#### *Cacheable*

In response, if the cache constraints are given then a client can reuse that response in a later request. it improves the efficiency and scalability of the system without loading the extra data.

### WebSocket based communication API

This type of API allows bi-directional full-duplex communication between server and client using the exclusive pair communication model. this API uses full-duplex communication so it does not require a new connection setup every time when it requests new data. WebSocket API begins with a connection setup between the server and client and if the WebSocket is supported by the server then it responds back to the client with the successful response after the setup of a connection server and the client can send data to each other in full-duplex mode.

## 6.IoT Enabled Technologies

**IoT-enabled technologies consist** of **big data**, **digital twins**, **cloud computing**, **sensors**, **communications protocols**, **analytics software**, **edge devices**, etc.

### Big Data

Smart objects are connected to the IoT, and more data is collected from them in order to perform analytics to regulate trends and associations that lead to insights. For example, a jetliner with 6,000 sensors generates 2.5 terabytes of data per day, In such a way, "big data" refers to these large data sets that need to be collected, stored, queried, analyzed, and generally managed in order to deliver on the promise of the IoT - insight. The technical challenge of big data is that the IoT system must deal with not only the data collected from smart objects but also additional data that is needed to execute such analytics for an eg- public and private data sets related to weather, GIS, financial, seismic, map, GPS, crime, etc. Furthermore, today smart objects come online, so IoT operators used three metrics to explain the big data: volume (ie., the amount of data they collect from their IoT sensors measured in gigabytes, terabytes, and petabytes); velocity (i.e., the speed at which data is collected from the sensors); and variety (types of structured and unstructured data collected, especially when compared to video and picture files.

Big Data analytics can help organizations to better understand the information contained within the data and will also help identify the data that is most important to the business and future business decisions, Analysts working with Big Data typically want the knowledge that comes from analyzing the data.

**Some examples of big data generated by IoT systems are described as follows:**

1. Sensor data is generated by IoT systems such as weather monitoring stations.
2. Machine sensor data is collected from sensors embedded in industrial and energy systems for monitoring their health and detecting Failures.
3. Health and fitness data generated by IoT devices such as wearable fitness bands
4. Data generated by IoT systems for location and tracking of vehicles
5. Data generated by retail inventory monitoring systems IoT systems such as weather monitoring stations.

### Characteristics of Big Data

Big data can be described by the following characteristics:

- **Volume** - The quantity of generated and stored data. The size of the data determines the value and potential insight, and whether it can be considered big data or not.
- **Variety** - The type and nature of the data. This helps people who analyze it to effectively use the resulting insight. Big data draws from text, images, audio, and video; plus it completes missing pieces through data fusion.
- **Velocity** - In this context, the speed at which the data is generated and processed to meet the demands and challenges that lie in the path of growth and development. Big data is often available in real time. Compared to small data, big data are produced more continually. Two kinds of velocity related to Big Data are the frequency of generation and the frequency of handling, and recording. and publishing.
- **Veracity** – It is the extended definition for big data, which refers to the data quality and the data value. The data quality of captured data can vary greatly, affecting the accurate analysis.

### Digital Twin

John Vickers, manager of NASA's National Center for Advanced Manufacturing. introduced the concept of Digital TWIN in 2003. This concept determines the digital copy of a physical asset that grows in a virtual environment over the physical asset's lifetime. We know that sensors within the object collect real-time data, and a set of models forming the digital twin is updated with all of the same information. Hence, an inspection of the digital twin would reveal the same information as a physical inspection of the smart object itself, remotely. The digital twin of the smart object is used to not only optimize operations of the smart object through reduced maintenance costs and downtime but to improve the next generation of its design.

### Cloud Computing

Cloud computing is like a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Cloud computing provides 3 service models. They are very essential for the IoT because it allows any user with a browser and an internet connection to transform smart object data into actionable intelligence.

Fig. **Cloud computing**

So cloud computing provides "the virtual infrastructure for utility computing integrating applications, monitoring devices, storage devices, analytics tools, visualization platforms, and client delivery to enable businesses and users to access IoT-enabled Applications on demand anytime, anyplace and anywhere,

Cloud computing resources can be provisioned on demand by the users, without requiring interactions with the cloud service provider. The process of provisioning resources is automated. Cloud computing resources can be accessed over The network using standard

access mechanisms that provide platform-independent access through the use of heterogeneous client platforms such as workstations, laptops, tablets, and smartphones.

### Sensors

Sensors are proficient in detecting events or changes in a specific quantity (e.g pressure) communicating the event or change data to the cloud (directly or via a gateway) and, in some circumstances, receiving data back from the cloud (e.g., a control command) or communicating with other smart objects.

### Communication Protocols

Wired and wireless communication technologies have also improved and nearly every type of electronic device should be connected to sensors embedded in smart objects to send and receive data over the cloud for collection, storage, and eventual analysis. The protocols for allowing IoT sensors to relay data include wireless technologies such as RFID, NFC, Wi-Fi, Bluetooth, Bluetooth Low Energy (BLE), XBee, ZigBee, Z-Wave, Wireless M-Bus, SIGFOX, and NuelNET, as well as satellite connections and mobile networks using GSM, GPRS, 3G LTE, or WiMAX.

Communication protocols form the backbone of IoT systems and enable network connectivity and coupling to applications. Communication protocols allow devices to exchange data over the network. Multiple protocols often describe different aspects of a single communication. A group of protocols designed to work together is known as a protocol suite; when implemented in software they are a protocol stack.

Internet communication protocols are published by the Internet Engineering Task Force (IETF). The IEEE handles wired and wireless networking, and the International Organization for Standardization (ISO) handles other types. The ITU-T handles telecommunication protocols and formats for the public switched telephone network (PSTN). As the PSTN and Internet converge, the standards are also being driven towards convergence.

### Analytics Software

Within the IoT ecosystem, Application Service Providers (ASPS) - which may or may not differ from the companies who sell and service the smart objects - provide software to companies that can transform "raw" machine (big) data collected from smart objects into actionable intelligence. The software performs data mining and employs mathematical models and statistical techniques to provide insight to users. That is, events, trends, and patterns are extracted from big data sets in order to present the software's end-users with insight in the form of portfolio analysis, predictions, risk analysis, automation, and corrective, maintenance and optimization recommendations.

### Edge Devices

Smart objects embedded with sensors connect via the Internet to various service provider systems. Because of these edge devices. It can be any device like a router, routing switch, Integrated Access Device (IAD), Multiplexer, or Metropolitan Area Network (MAN) and Wide Area Network (WAN) access device which provides an entry point from the global, public Internet into an ASP's or other enterprise's private network. For example, edge devices may translate between different network protocols, and provide first-hop security, initial Quality of Service (QoS), and access/ distribution policy functionality.

**7.A brief history of the Internet of Things**

The term Internet of Things is 16 years old. But the actual idea of connected devices had been around longer, at least since the 70s. Back then, the idea was often called "embedded internet" or "pervasive computing". But the actual term "Internet of Things" was coined by Kevin Ashton in 1999 during his work at Procter&Gamble. Ashton who was working in supply chain optimization, wanted to attract senior management's attention to a new exciting technology called RFID. Because the internet was the hottest new trend in 1999 and because it somehow made sense, he called his presentation "Internet of Things".

Even though Kevin grabbed the interest of some P&G executives, the term Internet of Things did not get widespread attention for the next 10 years.

**8.Things in iot**

he Internet of Things (IoT) describes the network of physical objects—"things"—that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the internet. These devices range from ordinary household objects to sophisticated industrial tools. With more than 7 billion connected IoT devices today, experts are expecting this number to grow to 10 billion by 2020 and 22 billion by 2025. Oracle has a network of device partners.

**What technologies have made IoT possible?**

While the idea of IoT has been in existence for a long time, a collection of recent advances in a number of different technologies has made it practical.

- **Access to low-cost, low-power sensor technology.** Affordable and reliable sensors are making IoT technology possible for more manufacturers.
- **Connectivity.** A host of network protocols for the internet has made it easy to connect sensors to the cloud and to other "things" for efficient data transfer.
- **Cloud computing platforms.** The increase in the availability of cloud platforms enables both businesses and consumers to access the infrastructure they need to scale up without actually having to manage it all.
- **Machine learning and analytics.** With advances in machine learning and analytics, along with access to varied and vast amounts of data stored in the cloud, businesses can gather insights faster and more easily. The emergence of these allied technologies continues to push the boundaries of IoT and the data produced by IoT also feeds these technologies.
- **Conversational artificial intelligence (AI).** Advances in neural networks have brought natural-language processing (NLP) to IoT devices (such as digital personal assistants Alexa, Cortana, and Siri) and made them appealing, affordable, and viable for home use.

**9.IoT Framework?**

Normally, when large data is being generated and transmitted across a number of devices, there has to be a specific point where everything is collected and combined.

This specific point is very essential in a network, as it combines all data, making it possible to understand the data being generated.

However, the smooth transmission and generation of data don't just happen. Rather, it is usually made possible by the Internet of Things Framework, (IoT framework). So, just what is IoT framework?

The Internet of Things (IoT) Framework can be described as being an ecosystem, comprising of several connected devices that communicate with each other, over the Internet. These

connected devices usually work to transfer and sense data over the Internet, while requiring very little human intervention.

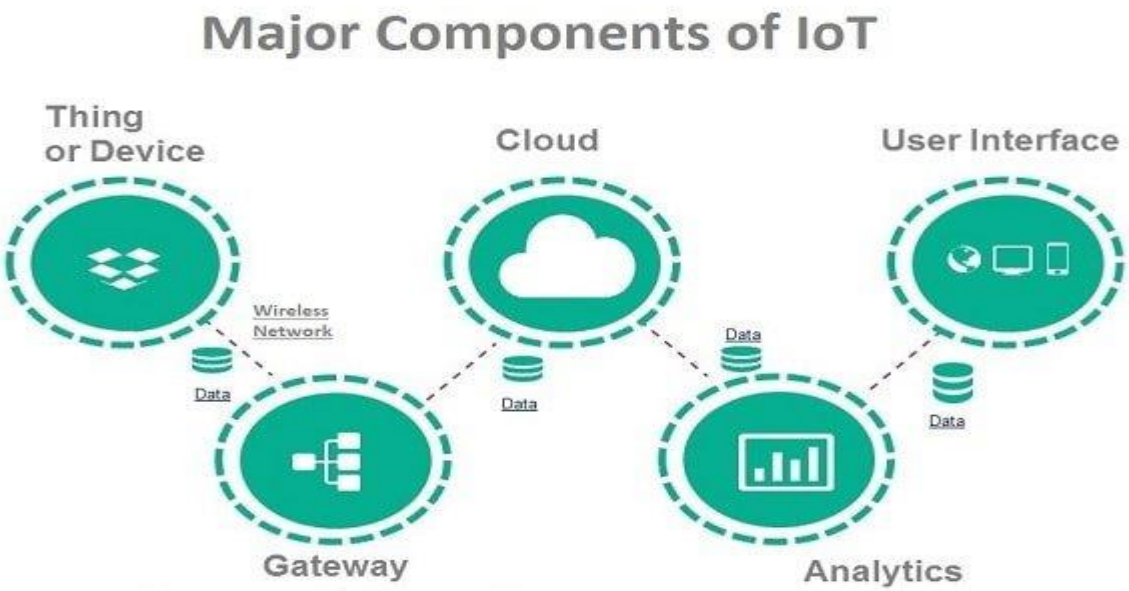
The IoT framework is what makes it possible for the connected devices to have smooth communication over the Internet. It is no wonder, then, that it is referred to as the 'Internet of Things' framework, or in other words, the framework that facilitates the interaction of 'Things' (devices) over the Internet.


**10.**IoT Framework Overview

The IoT framework is a very important element of technology in the modern world, finding application in almost every sector. For instance, one of the major applications of the IoT is in the designing of smart homes.

The IoT framework concept is also applied in the designing of different physical objects, such as thermostats, electrical devices, security and alarm systems, as well as vending machines, among many other objects.

**Main Components of the Internet of Things Framework**




major components of IoT Framework

The IoT Framework is comprised of four major components, as discussed below;

***1) Device Hardware***

The device hardware component of the IoT framework requires some basic knowledge on architecture. The user is also required to have an idea on the working of the different micro-controllers, as well the sensors.
Examples of hardware devices that form part of this IoT framework component are sensors, micro-controllers and controllers.

***2) Device Software***

In order for the device software of the IoT framework to function properly, the included writing applications are required to configure the controller, then operate them remotely. The user is required to have a basic understanding of how an API works inside the micro-controllers, as well how libraries are usually made for programming.

***3) Communication and Cloud Platform***

The cloud platform is one of the most crucial parts of the IoT framework. It calls for the basic knowledge of all communication, whether wireless or wired. The user is also required to have a good understanding of IoT integration, as well as the working of the cloud technology.
In summary, we can say the communication and Cloud Platform of the IoT Framework is where all communications happen.

***4) Cloud Application***

The cloud application is a type of software program, which mainly consists of components that can be accessed quite easier and faster. These components can be either local or even cloud-based. The cloud application works to improve the system, such that its maximum potential is realized.
In other words, the cloud application can be defined as the written application of an IoT framework, that binds all the local hardware devices, as well as the cloud-based devices.

**Open Source IoT Frameworks**

In order to understand what an IoT framework open source is, consider these three facts;

1) Every consumer wishes that they can use any technology device of their choice, without being limited or forced to use devices from just one particular vendor (for instance, some smart-watches require that they are paired with only smartphones from the same vendor).

2) All dealers of IoT devices wish that the integration of their devices can be made much easier, and even possible with a large number of different technology ecosystems.

3) Those who develop applications wish that they could support many different devices, which do not require any developed vendor-specific codes.

## 11.Comparison between the IoT and M2M

There are various differences between the IoT and M2M. Some of the popular comparisons between the IoT and M2M are discussed in the below tabular form.

| Features | IoT | M2M |
|---|---|---|
| **Abbreviation** | IoT stands for the Internet of Things. | M2M stands for Machine-to-Machine communication. |
| **Intelligence** | Devices include objects that are responsible for decision-making processes. | In M2M, there is a limited amount of intelligence observed. |
| **Communication Protocol Used** | IoT has used internet protocols like FTP, Telnet, and HTTP. | Communication technology and Traditional protocols are uses in M2M technology. |
| **Connection Type Used** | The connection of IoT is through the network and using various types of communication. | M2M uses a point to point connection. |
| **Scope** | It has a wide range of devices, yet the scope is large. | It has a limited Scope for devices. |
| **Business Type Used** | It has Business to Consumer (B2C) and Business to Business (B2B). | It has Business to Business (B2B) communication. |
| **Data Sharing** | In IoT, data sharing depends on the Internet protocol network. | In M2M, devices may be connected through mobile or any other network. |
| **Open API Support** | IoT technology supports Open API integrations. | In M2M technology, there is no Open API support. |
| **Example** | Big Data, Cloud, Smart wearables, and many more. | Data and Information, sensors, and many more. |

## Unit-2

### sensor network

1.A **sensor network** comprises a group of small, powered devices, and a wireless or wired networked infrastructure. They record conditions in any number of environments including industrial facilities, farms, and hospitals. The sensor network connects to the internet or computer networks to transfer data for analysis and use.

Sensor network nodes cooperatively sense and control the environment. They enable interaction between persons or computers and the surrounding environment.

**Wired vs. Wireless Sensor Networks**

Sensor networks can be wired or wireless. **Wired sensor networks** use ethernet cables to connect sensors. **Wireless sensor networks (WSNs)** use technologies such as Bluetooth, cellular, wifi or near field communication (NFC) to connect sensors.

WSNs are easier to deploy and maintain and offer better flexibility of devices. With the rapid development of sensors and wireless technologies, WSNs have become a key technology of the IoT. WSNs don't need the physical network infrastructure to be modified.

**Operation of a Sensor Network**

Sensor networks typically include sensor nodes, actuator nodes, gateways, and clients. Sensor nodes group inside the sensor field and form networks of different topologies. The following process describes how sensor networks operate:

- A sensor node monitors the data collected by the sensor and transmits this to other sensor nodes.
- During the transmission process, data may be handled by multiple nodes as it reaches a gateway node.
- The data is then transferred to the management node.
- The management node is managed by the user and determines the monitoring required and collects the monitored data.

**2.Sensor Nodes**

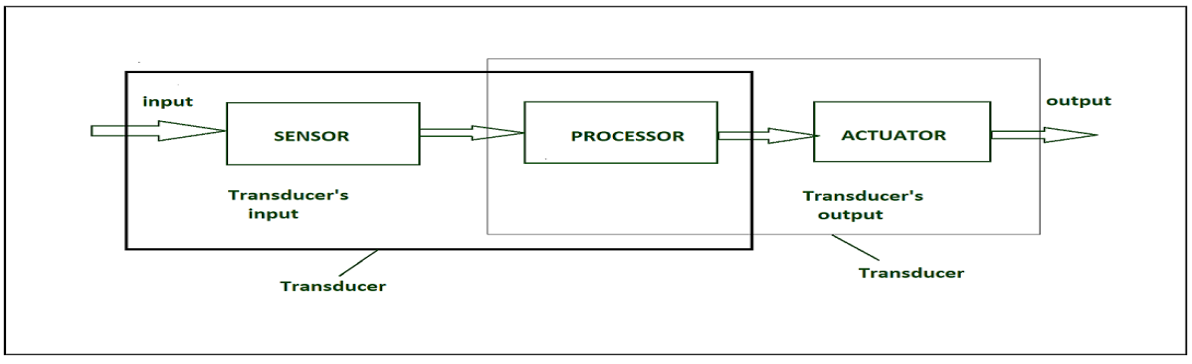
There are many nodes in a sensor network. These nodes are the detection stations. There is a sensor/transducer, microcontroller, transceiver, and power source:

- A sensor senses the physical condition, and if there is any change, it generates electrical signals.
- The signals go to the microcontroller for processing.
- A central processor sends commands to the transceiver and data is transmitted to a computer.

**Sensors**

The sensor is the bond of a sensor network node. Examples of sensors include temperature sensors, accelerometers, infrared detectors, proximity sensors, and motion detectors.

Figure 3, shows the sensors used in a self-driving (autonomous) car.




**Transducer :**

- A transducer converts a signal from one physical structure to another.
- It converts one type of energy into another type.

- It might be used as actuators in various systems.

**Sensors characteristics :**

1. Static
2. Dynamic

**1. Static characteristics :**

It is about how the output of a sensor changes in response to an input change after steady state condition.

- **Accuracy –**
  Accuracy is the capability of measuring instruments to give a result close to the true value of the measured quantity. It measures errors. It is measured by absolute and relative errors. Express the correctness of the output compared to a higher prior system. Absolute error = Measured value – True value
  Relative error = Measured value/True value
- **Range –**
  Gives the highest and the lowest value of the physical quantity within which the sensor can actually sense. Beyond these values, there is no sense or no kind of response. e.g. RTD for measurement of temperature has a range of -200`c to 800`c.
- **Resolution –**
  Resolution is an important specification towards selection of sensors. The higher the resolution, better the precision. When the accretion is zero to, it is called threshold. Provide the smallest changes in the input that a sensor is able to sense.
- **Precision –**
  It is the capacity of a measuring instrument to give the same reading when repetitively measuring the same quantity under the same prescribed conditions.
  It implies agreement between successive readings, NOT closeness to the true value.
  It is related to the variance of a set of measurements.
  It is a necessary but not sufficient condition for accuracy.
- **Sensitivity –**
  Sensitivity indicates the ratio of incremental change in the response of the system with respect to incremental change in input parameters. It can be found from the slope of the output characteristics curve of a sensor. It is the smallest amount of difference in quantity that will change the instrument's reading.
- **Linearity –**
  The deviation of the sensor value curve from a particular straight line. Linearity is determined by the calibration curve. The static calibration curve plots the output amplitude versus the input amplitude under static conditions.
  A curve's slope resemblance to a straight line describes the linearity.
- **Drift –**
  The difference in the measurement of the sensor from a specific reading when kept at that value for a long period of time.
- **Repeatability –**
  The deviation between measurements in a sequence under the same conditions. The measurements have to be made under a short enough time duration so as not to allow significant long-term drift.

**Dynamic Characteristics :**

Properties of the systems

- **Zero-order system –**
  The output shows a response to the input signal with no delay. It does not include energy-storing elements.
  Ex. potentiometer measure, linear and rotary displacements.
- **First-order system –**
  When the output approaches its final value gradually.
  Consists of an energy storage and dissipation element.
- **Second-order system –**
  Complex output response. The output response of the sensor oscillates before steady state.

**Sensor Classification :**

- Passive & Active
- Analog & digital
- Scalar & vector

1. **Passive Sensor –**
   Can not independently sense the input. Ex- Accelerometer, soil moisture, water level and temperature sensors.
2. **Active Sensor –**
   Independently sense the input. Example- Radar, sounder and laser altimeter sensors.
3. **Analog Sensor –**
   The response or output of the sensor is some continuous function of its input parameter. Ex- Temperature sensor, LDR, analog pressure sensor and analog hall effect.
4. **Digital sensor –**
   Response in binary nature. Design to overcome the disadvantages of analog sensors. Along with the analog sensor, it also comprises extra electronics for bit conversion. Example – Passive infrared (PIR) sensor and digital temperature sensor(DS1620).
5. **Scalar sensor –**
   Detects the input parameter only based on its magnitude. The answer for the sensor is a function of magnitude of some input parameter. Not affected by the direction of input parameters.
   Example – temperature, gas, strain, color and smoke sensor.
6. **Vector sensor –**
   The response of the sensor depends on the magnitude of the direction and orientation of input parameter. Example – Accelerometer, gyroscope, magnetic field and motion detector sensors.

**3.Different Types of Sensors**

All these sensors are used for measuring one of the physical properties like Temperature, Resistance, Capacitance, Conduction, Heat Transfer etc.

1. Temperature Sensor
2. Proximity Sensor
3. Accelerometer
4. IR Sensor (Infrared Sensor)
5. Pressure Sensor
6. Light Sensor
7. Ultrasonic Sensor
8. Smoke, Gas and Alcohol Sensor
9. Touch Sensor
10. Color Sensor
11. Humidity Sensor
12. Position Sensor
13. Magnetic Sensor (Hall Effect Sensor)
14. Microphone (Sound Sensor)
15. Tilt Sensor
16. Flow and Level Sensor
17. PIR Sensor
18. Touch Sensor
19. Strain and Weight Sensor

**Temperature Sensor**

One of the most common and most popular sensors is the Temperature Sensor. A Temperature Sensor, as the name suggests, senses the temperature i.e., it measures the changes in the temperature.

There are different types of Temperature Sensors like Temperature Sensor ICs (like LM35, DS18B20), Thermistors, Thermocouples, RTD (Resistive Temperature Devices), etc.

Temperature Sensors can be analog or digital. In an Analog Temperature Sensor, the changes in the Temperature correspond to change in its physical property like resistance or voltage. LM35 is a classic Analog Temperature Sensor.

Coming to the Digital Temperature Sensor, the output is a discrete digital value (usually, some numerical data after converting analog value to digital value). DS18B20 is a simple Digital Temperature Sensor.

Temperature Sensors are used everywhere like computers, mobile phones, automobiles, air conditioning systems, industries etc.

**Proximity Sensors**

A Proximity Sensor is a non-contact type sensor that detects the presence of an object. Proximity Sensors can be implemented using different techniques like Optical (like Infrared or Laser), Sound (Ultrasonic), Magnetic (Hall Effect), Capacitive, etc.

Some of the applications of Proximity Sensors are Mobile Phones, Cars (Parking Sensors), industries (object alignment), Ground Proximity in Aircrafts, etc.

**Infrared Sensor (IR Sensor)**

IR Sensors or Infrared Sensor are light based sensor that are used in various applications like Proximity and Object Detection. IR Sensors are used as proximity sensors in almost all mobile phones.

There are two types of Infrared or IR Sensors: Transmissive Type and Reflective Type. In Transmissive Type IR Sensor, the IR Transmitter (usually an IR LED) and the IR Detector (usually a Photo Diode) are positioned facing each other so that when an object passes between them, the sensor detects the object.

The other type of IR Sensor is a Reflective Type IR Sensor. In this, the transmitter and the detector are positioned adjacent to each other facing the object. When an object comes in front of the sensor, the infrared light from the IR Transmitter is reflected from the object and is detected by the IR Receiver and thus the sensor detects the object.

Different applications where IR Sensor is implemented are Mobile Phones, Robots, Industrial assembly, automobiles etc.

**Ultrasonic Sensor**

An Ultrasonic Sensor is a non-contact type device that can be used to measure distance as well as velocity of an object. An Ultrasonic Sensor works based on the properties of the sound waves with frequency greater than that of the human audible range.

Using the time of flight of the sound wave, an Ultrasonic Sensor can measure the distance of the object (similar to SONAR). The Doppler Shift property of the sound wave is used to measure the velocity of an object.

**Light Sensor**

Sometimes also known as Photo Sensors, Light Sensors are one of the important sensors. A simple Light Sensor available today is the Light Dependent Resistor or LDR. The property of LDR is that its resistance is inversely proportional to the intensity of the ambient light i.e., when the intensity of light increases, its resistance decreases and vise-versa.

By using LDR is a circuit, we can calibrate the changes in its resistance to measure the intensity of Light. There are two other Light Sensors (or Photo Sensors) which are often used in complex electronic system design. They are Photo Diode and Photo Transistor. All these are Analog Sensors.

There are also Digital Light Sensors like BH1750, TSL2561, etc., which can calculate intensity of light and provide a digital equivalent value.

**Smoke and Gas Sensors**

One of the very useful sensors in safety related applications are Smoke and Gas Sensors. Almost all offices and industries are equipped with several smoke detectors, which detect any smoke (due to fire) and sound an alarm.

Gas Sensors are more common in laboratories, large scale kitchens and industries. They can detect different gases like LPG, Propane, Butane, Methane (CH4), etc.

Now-a-days, smoke sensors (which often can detect smoke as well gas) are also installed in most

**Alcohol Sensor**

As the name suggests, an Alcohol Sensor detects alcohol. Usually, alcohol sensors are used in breathalyzer devices, which determine whether a person is drunk or not. Law enforcement personnel uses breathalyzers to catch drunk-and-drive culprits.

**Touch Sensor**

We do not give much importance to touch sensors but they became an integral part of our life. Whether you know or not, all touch screen devices (Mobile Phones, Tablets, Laptops, etc.) have touch sensors in them. Another common application of touch sensor is trackpads in our laptops.
Touch Sensors, as the name suggests, detect touch of a finger or a stylus. Often touch sensors are classified into Resistive and Capacitive type. Almost all modern touch sensors are of Capacitive Types as they are more accurate and have better signal to noise ratio.

**Color Sensor**

A Color Sensor is an useful device in building color sensing applications in the field of image processing, color identification, industrial object tracking etc. The TCS3200 is a simple Color Sensor, which can detect any color and output a square wave proportional to the wavelength of the detected color.

**Humidity Sensor**

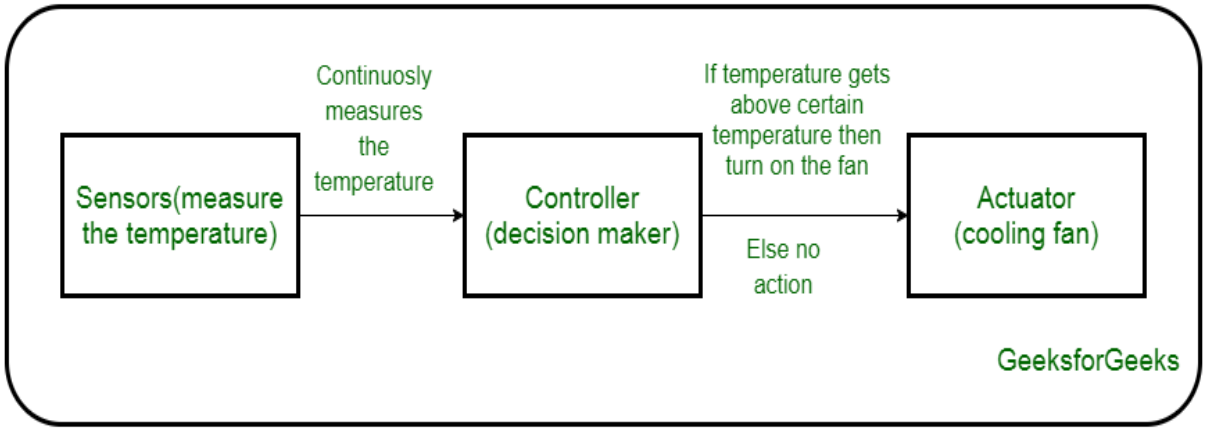
If you see Weather Monitoring Systems, they often provide temperature as well as humidity data. So, measuring humidity is an important task in many applications and Humidity Sensors help us in achieving this.
Often all humidity sensors measure relative humidity (a ratio of water content in air to maximum potential of air to hold water). Since relative humidity is dependent on temperature of air, almost all Humidity Sensors can also measure Temperature.
Humidity Sensors are classified into Capacitive Type, Resistive Type and Thermal Conductive Type. DHT11 and DHT22 are two of the frequently used Humidity Sensors in DIY Community (the former is a resistive type while the latter is capacitive type).

**Tilt Sensor**

Often used to detect inclination or orientation, Tilt Sensors are one of the simplest and inexpensive sensors out there. Previously, tilt sensors are made up of Mercury (and hence they are sometimes called as Mercury Switches) but most modern tilt sensors contain a roller ball.

**4.Actuators in IoT**

An IoT device is made up of a Physical object (“thing”) + Controller (“brain”) + Sensors + Actuators + Networks (Internet). An actuator is a machine component or system that moves or controls the mechanism or the system. Sensors in the device sense the environment, then control signals are generated for the actuators according to the actions needed to perform.




**Types of Actuators :**

**1. Hydraulic Actuators –**

A hydraulic actuator uses hydraulic power to perform a mechanical operation. They are actuated by a cylinder or fluid motor. The mechanical motion is converted to rotary, linear, or oscillatory motion, according to the need of the IoT device. Ex- construction equipment uses hydraulic actuators because hydraulic actuators can generate a large amount of force.

**Advantages :**

- Hydraulic actuators can produce a large magnitude of force and high speed.
- Used in welding, clamping, etc.
- Used for lowering or raising the vehicles in car transport carriers.

**Disadvantages :**

- Hydraulic fluid leaks can cause efficiency loss and issues of cleaning.

- It is expensive.
- It requires noise reduction equipment, heat exchangers, and high maintenance systems.

**2. Pneumatic Actuators –**

A pneumatic actuator uses energy formed by vacuum or compressed air at high pressure to convert into either linear or rotary motion. Example- Used in robotics, use sensors that work like human fingers by using compressed air.

**Advantages :**

- They are a low-cost option and are used at extreme temperatures where using air is a safer option than chemicals.
- They need low maintenance, are durable, and have a long operational life.
- It is very quick in starting and stopping the motion.

**Disadvantages :**

- Loss of pressure can make it less efficient.
- The air compressor should be running continuously.
- Air can be polluted, and it needs maintenance.

**3. Electrical Actuators –**

An electric actuator uses electrical energy, is usually actuated by a motor that converts electrical energy into mechanical torque. An example of an electric actuator is a solenoid based electric bell.

**Advantages :**

- It has many applications in various industries as it can automate industrial valves.
- It produces less noise and is safe to use since there are no fluid leakages.
- It can be re-programmed and it provides the highest control precision positioning.

**Disadvantages :**

- It is expensive.
- It depends a lot on environmental conditions.

Other actuators are –

- **Thermal/Magnetic Actuators –**
  These are actuated by thermal or mechanical energy. Shape Memory Alloys (SMAs) or Magnetic Shape-Memory Alloys (MSMAs) are used by these actuators. An example of a thermal/magnetic actuator can be a piezo motor using SMA.
- **Mechanical Actuators –**
  A mechanical actuator executes movement by converting rotary motion into linear motion. It involves pulleys, chains, gears, rails, and other devices to operate. Example – A crankshaft.
- Soft Actuators
- Shape Memory Polymers
- Light Activated Polymers

- With the expanding world of IoT, sensors and actuators will find more usage in commercial and domestic applications along with the pre-existing use in industry.

## 5.Examples of IoT Devices

### Home Security

The key driver behind smart and secure homes is IoT. A variety of sensors, lights, alarms and cameras (all of which can be controlled from a smartphone) are connected via IoT to provide 24x7 security.

### Activity Trackers

Smart home security cameras provide alerts and peace of mind. Activity trackers are sensor devices that can monitor and transmit key health indicators in real-time. You can track and manage your blood pressure, appetite, physical movement and oxygen levels.

### Industrial Security and Safety

IoT-enabled detection systems, sensors and cameras can be placed in restricted areas to detect trespassers. They can also identify pressure buildups and small leaks of hazardous chemicals and fix them before they become serious problems.

**Augmented Reality Glasses**

Augmented Reality (AR) glasses are wearable computer-enabled glasses that help you get extra information such as 3D animations and videos to the user's real-world scenes. The information is presented within the lenses of the glasses and can help users access Internet applications.

**Motion Detection**

Motion sensors can detect vibrations in buildings, bridges, dams and other large-scale structures. These devices can identify anomalies and disturbances in the structures that could lead to catastrophic failures. They can also be used in areas susceptible to floods, landslides, and earthquakes.

**6.IoT Development Board**

**IoT boards** are useful hardware structures that we use to prototype a new IoT project. As we discussed above, the custom hardware results in expensive to design and manufacture, and **development boards** comes to rescue to avoid that.Let's get straight to the most popular IoT Development Boards:

1. Raspberry Pi
2. Omega 2
3. Particle Photon
4. Beagle bone –
5. Jetson Nano
6. ESP 32
7. Banana Pi
8. Arduino Nano 33 IoT
9. Tessel 2
10. i.MX 8

**1. Raspberry Pi**The raspberry pi Development Board is a small credit **card size computer**. That works on Linux based operating systems and is good for embedded projects. Raspberry boards can be easily plugged in to your monitor, computer or TV. It uses a standard keyboard and mouse. Even amateur users depend on it for configuring their digital media systems and surveillance cameras.

**Features :**

- Processor: 1.2GHz, 64-bit quad-core ARMv8 CPU
- 802.11n Wireless LAN
- Bluetooth 4.1
- Bluetooth Low Energy (BLE)
- 1GB RAM
- 4 USB ports
- 40 GPIO pins
- Full HDMI port
- Combined 3.5mm audio jack and composite video
- Camera interface (CSI)
- Display interface (DSI)
- Micro SD card slot
- videoCore IV 3D graphics core

**2. Omega 2**Omega 2 is one of Onion's Linux-based WiFi *development boards* that allow makers of all skill levels to build connected hardware. This highly integrated **board** comes with a powerful processor and flexible GPIOs. The Platform lets you prototype hardware devices using familiar tools like Git, npm, pip, as well as high-level programming languages like Python, Javascript, and PHP.

**Features** :

- Linux Operating System, powerful processor, and flexible GPIOs.
- Compact size that easily fits into any project design.
- Modular design for a vast range of flexibility.
- Arduino compatible.
- Integrated Wi-Fi;

- Connectivity is expandable with 2G, 3G, Ethernet, Bluetooth®, Bluetooth Low Energy (BLE), GPS.
- U.FL Connector for external Wi-Fi antenna attachment.
- FCC and CE Certified.

**3. Particle Photon**

Particle Photon Board consists of an STM32 **microcontroller**, Wi-Fi, Switches, and LEDs. Simple to use, powerful, and connected to the cloud. Powered by a Cypress Wi-Fi chip alongside a powerful STM32 ARM Cortex M3 microcontroller, it is ideal for prototyping IoT projects.

**Features** :

- Processor: STM32F205 120Mhz ARM Cortex M3
- Real-time operating system (Free RTOS)
- Memory: 1MB flash, 128KB RAM
- Open source design
- On-board Wi-Fi module
- On-board RGB status LED.
- 18 Mixed-signal GPIO and advanced peripherals
- Soft AP setup
- B802.11b/g/n Wi-Fi
- roadcom BCM43362 Wi-Fi chip

**4. Beagle Bone**

The Beagle bone is a low power open-source **single-board computer** produced by Texas instruments. The board can boot Linux in under 10 seconds also you can start developing in less than 5 minutes with just a single USB cable.

It is a computer installed inside of a larger electronics project. The beagle board carries two rows of GPIO (general purpose Input/Output) pins mounted along each side of the board. That allow it to communicate with a wide range of servos, sensors, outputs and other hardware, making it act as the brain of large & complex projects.

Its capabilities can be extended using plug-in boards referred to as “capes”. that are easily available for LCD, motor control, VGA, prototyping, battery power, and other functionalities.

**Features :**

- DDR memory: 512 MB
- Ability to run Ruby, Python, and INO Sketches directly in the Cloud9 IDE,
- Ethernet: On-chip 10/100 Ethernet
- JTAG: Optional
- Memory: 4GB eMMC memory
- Power Options: Via USB or 5V DC input
- Price (USD) Per Unit: $55.00 (Suggested Retail Price)
- Processor: 1GHz AM3359 Sitara ARM Cortex-A8

**5. Jetson Nano**

Jetson Nano is a power-efficient and low-cost development board. Provides total performance to run modern AI workloads in a small form factor. Additionally, It has the ability for heavy workload applications like image classification, object detection, segmentation, and speech processing. It is capable to run multiple **neural network apps** at the same time.

**Features:**

- GPU: 128-core NVIDIA Maxwell™ architecture-based GPU.
- CPU: Quad-core ARM® A57.
- Video: 4K @ 30 fps
- Camera: 1/3” AR0330 CMOS Image sensor with 2.2 µm pixel.
- Memory: 4 GB 64-bit LPDDR4; 25.6 gigabytes/second.
- Connectivity: Gigabit Ethernet.

- OS Support: Linux for Tegra®.

**ESP 32**

ESP32 is a dual core low-footprint system **development board** powered by the latest ESP-WROOM-32 module that can be easily placed into a solderless breadboard. It has a pre-integrated antenna, power amplifier, low-noise amplifiers, filters, and power management module. Because of this, it's easy to build and test circuits as well as making projects related to IoT integrating with the cloud platform.

**7. Banana Pi**

Banana Pi is a line of low-cost credit card-sized single-board computers(SBC). IT is a router-based development board, which efficiently runs on various open-source operating systems including OpenWRT and Android, Lubuntu, Ubuntu, Debian, and Raspbian. Well, the hardware design of banana pi was influenced by the Raspberry Pi and it is compatible with Raspberry Pi boards.

**8. Arduino Nano 33 IoT**

The Arduino **Nano 33 IoT** is a dual-processor device that is perfect for experimentation. It offers a practical and low-cost solution for inventors seeking to add Wi-Fi connectivity to their projects with minimal previous experience in networking. The board is compatible with the Arduino IoT Cloud, where you can create IoT applications in a few simple steps

**6.Arduino IDE and Board Types:**

The Arduino board is invented for the electronics students to use this in their projects. The Arduino boards are provided as open source that helps the user to build their projects and instruments according to their need. This electronic platform contains microcontrollers, connections, LEDs and many more. There are various types of Arduino boards present in the market that includes Arduino UNO, Red Board, LilyPad Arduino, Arduino Mega, Arduino Leonardo. All these Arduino boards are different in specifications, features and uses and are used in different type of electronics project.

**Types of Arduino**

**1. Arduino UNO**

The development of Arduino UNO board is considered as new compared to other Arduino boards. This board comes up with numerous features that helps the user to use this in their project. The Arduino UNO uses the Atmega16U2 microcontroller that helps to increase the transfer rate and contain large memory compared to other boards. No extra devices are needed for the Arduino UNO board like joystick, mouse, keyboard and many more. The Arduino UNO contain SCL and SDA pins and also have two additional pins fit near to RESET pin.

The board contains 14 digital input pins and output pins in which 6 pins are used as PWM, 6 pins as analog inputs, USB connection, reset button and one power jack. The Arduino UNO board can be attached to computer system buy USB port and also get power supply to board from computer system. The Arduino UNO contains flash memory of size 32 KB that is used to the data in it. The other feature of the Arduino UNO is compatibility with other shield and can be combined with other Arduino products.

**2. LilyPad Arduino**

The LilyPad Arduino is considered as other Arduino board type that is designed for integrating with wearable projects and e-textile projects. This board comes in round shape that helps to decrease the snagging and can be easily connected to other devices. This board uses the Atmega328 microcontroller and Arduino bootloader in it. This board uses very less external component in it that makes the design easy and compatible.

The board requires 2 volt to 5 volt power supply and use large size pin holes so that it can be easily connect to other devices. This board is widely used for controlling different device that includes motor, light and switch. The components of this board like sensor board, input board and output board can be washable because this board is used in clothing industries.

**3. Arduino Mega**

This boards is considered as the microcontroller that uses the Atmega2560 in it. There are total 54 input pins and output pins in it in which 14 pins are of PWM output, 4 pins are of hardware port, 16 pins as analog inputs. The board also contain one USB connection, ICSP header, power jack and one REST pin.

There are additional pins that act as crystal oscillator having frequency of 16 MHz. The board also has flash memory of 256KB size that uses to store the data in it. The Arduino Mega board can be

attached to computer system via USB connection and power supply can be provided to board by using battery or AC to DC adapter. As the board has large number of pins fitted in it that make the board suitable for projects that requires more number of pins in it.

**4. Arduino Leonardo**

This board is considered as the microcontroller that uses the Atmega32u4 in it. There are total 20 digital input pins and output pins in it, in that 7 pins are used As PWM and 12 pins used as analog inputs. The board also contain one micro USB connection, power jack, and one RESET button fit in it. There are additional pins which act as crystal oscillator of frequency 16 MHz.

The Arduino Leonardo board can be attached to computer system via USB connection and power supply can be provided to board by using battery or AC to DC adapter. The microcontroller used by the Arduino Leonardo has in-built USB connection that removes the dependency of extra processor in it. As there is no additional USB connection in the board, it helps the board to act as mouse or keyboard for the computer system. The Arduino Leonardo is considered as cheapest Arduino boards compare to other Arduino products.

**5. Arduino Red Board**

The Arduino Red board is another type of Arduino board that uses the mini USB cable for getting programmed and the Arduino IDE is used for this purpose. This board is compatible with Windows 8 operating system and there is no need to change the security settings to make this board working. The Red board uses the FTDI chip and USB chip for the connection to other device. As the design of red board is very simple it can be easily integrate with other projects. The only requirement if to plug the red board and select appropriate option and can upload program in no time. The barrel jack can be used to control the USB cable of the Arduino Red board.

**6. Arduino Shields**

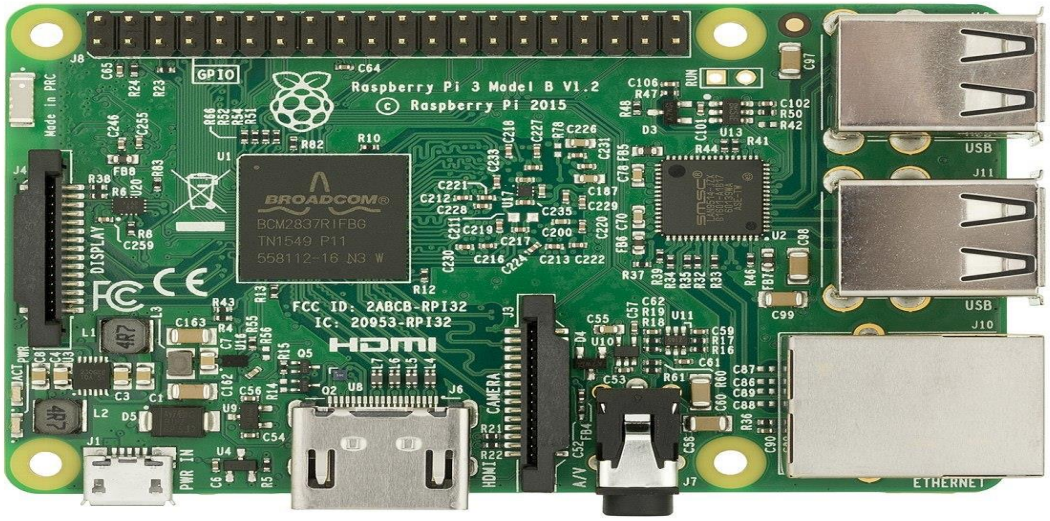
The Arduino shields are considered as pre-build circuit boards that are used to connect other Arduino boards. The Arduino shield are placed on top of Arduino boards and enhance the capability of board to get connected to internet network, controlling of motor, controlling of LCD and also help to establish wireless communication. There are different type of shields available for the use. It includes Wireless Shields, Ethernet Shield, Proto Shield and GSM shield. This helps to increase the compatibility of the Arduino boards.

**7.What is a Raspberry Pi?**

Raspberry pi is the name of the “credit card-sized computer board” developed by the Raspberry pi foundation, based in the U.K. It gets plugged in a TV or monitor and provides a fully functional computer capability.

It is aimed at imparting knowledge about computing to even younger students at the cheapest possible price.

Although it is aimed at teaching computing to kids, but can be used by everyone willing to learn programming, basics of computing, and building different projects by utilizing its versatility.



RASPBERRY PI DEVELOPMENT KIT

This is a nice collection of the basic and advanced accessories required for getting started with the Raspberry Pi.

**Kit Contents**

- Raspberry Pi (Model B+, 512MB RAM)
- 8GB SD Card with pre-installed OS (Raspbian Wheezy)
- HDMI Cable (1.5M)
- HC-SR04 Sensor
- PIR Sensor
- IR Sensor
- Humidity and Temperature Sensor
- 4-port Quantum Hub
- Power Adapter (2AMPS) with Micro USB Cable
- LCD (16*2)
- Cat 5E Ethernet Cable (2M)
- Light Sensor
- Push Buttons (5)
- Buzzer
- 10 Jumper wires
- Card Reader
- Breadboard
- 10 Leds
- Resistor Box
- Potentiometer
- Tactile switch (2)
- IN4007
- RTC-DS1307 IC
- Male headers
- Crystal Oscillator
- Small Dot Board

**8.Components of RFID Technology and Applications**

RFID or Radio Frequency Identification technology has been around for decades. RFID is one of the cost effective solutions which transfer data wirelessly within the proximity. What are the components of RFID technology and what are the applications?

**Components of RFID Technology**

RFID technology consists of four components such as RFID tags, antenna, RFID receiver (transceiver) and software.

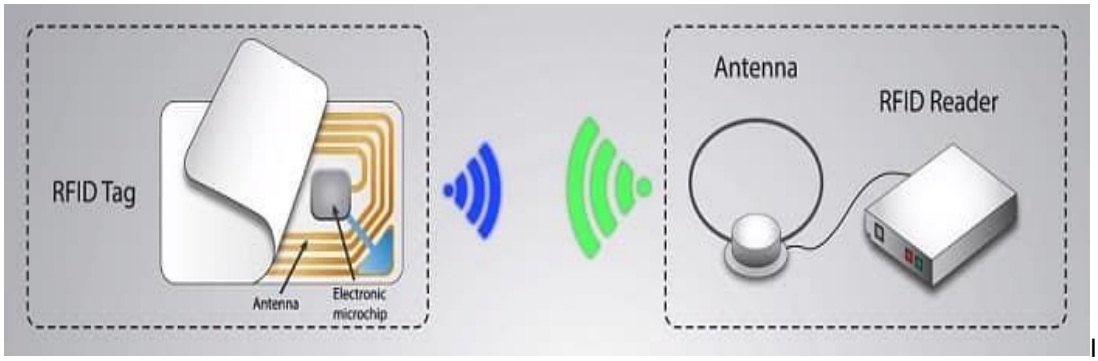


Image:

1. RFID Tag

RFID tags are small devices consists of an electronic microchip embedded inside and an antenna. The microchip has the unique identification number of the RFID tag.

Passive RFID tag does not have a power source; it will receive power from radio signals transmitted from the RFID receiver. These tags will operate when the reader is at the proximity of the tags (line of sight not required).
Antenna coil will act as power source and medium to transfer data to the reader.

***Types of Tags***

Passive Tags: Does not have a power source, uses power from the reader to operate.
Battery Assisted Passive Tags: Logic circuit chip uses battery power. Need RF signals from the reader to activate and function.
Active Tags: Uses a power source like battery, does not require power from source/reader.

**2. Antenna**

RFID antennas are designed to operate at a specific frequency for each applications in which it operates. These antennas are often mounted on the RFID reader and easily accessible for tags to tap on it.


In some handheld devices, antenna is often attached to the device. Size and shape of the antenna depends on the application and the operating frequency of the system.

3. RFID Reader

RFID reader is one of the significant hardware component in the RFID system which read information from the RFID devices/tags and connected to the network to transfer the information to the database.

***Specification of RFID Reader***

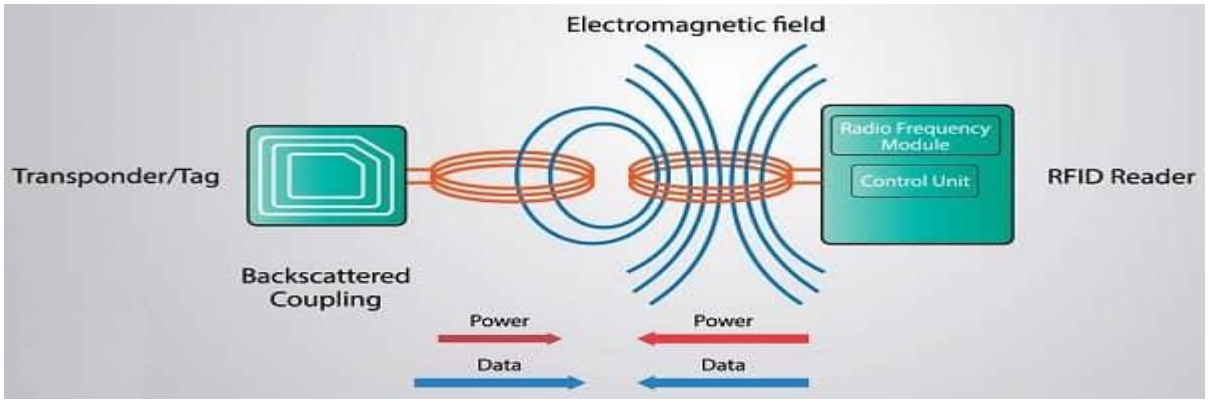
Frequency: Operating frequency is one of the specifications of the RFID reader.

| **Frequency Band** | **Range** | **Data Rate** |
|---|---|---|
| LF: 120–150 kHz | 10 cm | Low |
| HF: 13.56 MHz | 0.1–1 m | Low to moderate |
| UHF: 433 MHz | 1–100 m | Moderate |
| UHF: 865–868 MHz | 1–12 m | Moderate to high |
| 902–928 MHz | | |
| microwave: 2450–5800 MHz | 1–2 m | High |
| microwave: 3.1–10 GHz | up to 200 m | High |

4. Software

RFID technology uses specific software depends on service providers. This software controls the RFID reader, initiate scan and retrieve information from the tags and stores the information to a local computer or send to the cloud storage.
RFID tags can be erased and re-used using control software.

How RFID works?




RFID technology works based on the principle of inductive coupling which include a source antenna and receiver antenna. Each RDID tag will have a microchip which contains a unique identification number, model, manufacturing date, expiry date, access information etc.…

RFID reading is a process of accessing information of a tag using a reader. When the user initiates a scanning, the tags are placed near the RFID reader or vice versa.

**What are the applications of RFID technology?**

1. Retail and Supply Chain

**RFID technology is revolutionizing the operational efficiency across the supply chain. Retail stores can efficiently manage stocks and customers can access information about a particular product or self checkout from the stores.**

**Many fashion retailers have already adapted RFID technology; it helped them to modernize the store and improved customer satisfaction. RFID technology helps faster check outs, reduce number of support staff.**

**Theft control is another advantage of RDIF technology in retail stores. RFID readers will detect any items pass through the exit without payment is completed**

2. Access control

**Access control and security is one of common applications of RFID technology. Employers can assign different level of access to each work group personnel in office environment, manufacturing plants, hospitals, airports and shops etc.… Using modern RFID technology, highly secure access control system can be implemented and monitored remotely.**

**Companies are using personal identification badges with RFID technology for employees.**

**3. Medical and hospital**

RFID technology can be used in medicine and healthcare to track movement of medical equipment, update medicine stock and authorized access for medical professionals. A simple RFID tag on patient's wrist does not require scanning (like barcode scan); it helps medical practitioners to efficiently handle patients for different procedures.
Implementation of RFID technology facilitates hospitals to improve efficiency, avoid errors during medication and better customer satisfaction. An RFID can store more data like personal information, allergy to some medicine, chronic conditions like diabetes etc.
Furthermore, RFID implementation can reduce wait time during busy hours in hospitals and clinics. Healthcare personnel can track the patient's location within hospitals and guide them for other procedures if necessary.
Pharmaceutical companies use RFID technology to prevent counterfeit products reaching the market.

**4. Logistics and shipping**

RFID tags were primarily developed to improve efficiency of logistics and shipping. Manual recording of inventory movements are not feasible for larger quantity of items from warehouse, manufacturing and storage area.
RFID readers can easily read hundreds of tags within seconds with accuracy. Once the items move out a warehouse premises, data will be updated to the database without manual action.
RFID gates are used for inventory management in big warehouse storage areas. RFID readers installed on the gates will record inventory movement in and out of a warehouse location.

**5. Automation of manufacturing**

RFID technology is one of the significant enablers for smart factory concept to help automate different stages during manufacturing. During each stages of manufacturing, RFID readers will record the movement of products and update to the database without additional action from the operator.
RFID technology is suitable for large scale production sites where hundreds or thousands of products are being monitored and status information is recorded real-time. Manually updating this information is time consuming and cause error while entering the data.

**6. Animal tracking**

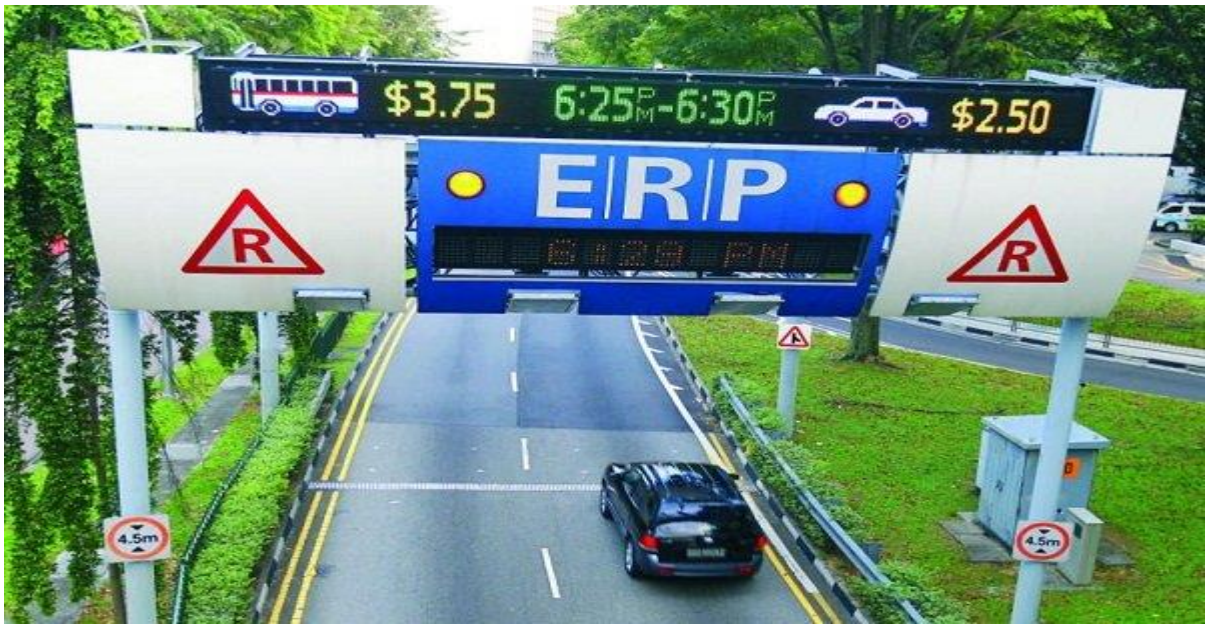
Implementation of RFID tags on livestock helps farmers to update, identify and easier tracking. Manually updating large amount of data is not an easy task especially on a remote location.
Using a handheld reader, information (age, weight, vaccination data etc….) can be easily accessed within seconds. Veterinary doctors will be able to retrieve information about a pet by just scanning the tag (without going through records).

**7. Baggage handling in aviation**

Numbers of airline passengers have been increasing every year and it will continue to grow in coming years. Baggage handling of huge number of passengers is a heavy task for airlines, mishandling of baggage often costs millions of dollars every year.

Efficient usage of RFID technology helps airlines to solve this issue and significantly reduce mishandling of luggage. RFIF tags does not require is line of sight angle to read information, this is an advantage compared to barcode scanning.

**8. Toll gate systems / Electronic Road Pricing**



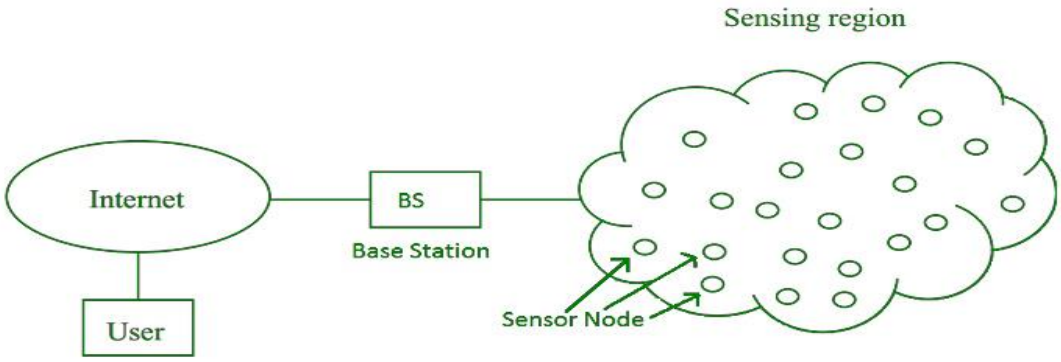

Electronic toll gates and Electronic Road Pricing (ERP) are using RFID technology to allow access and charge for entry to an area. These are mostly active tags with payment option enabled.

RFID readers can read these tags from up to 20 – 30 feet above the vehicles.

# 8.Wireless Sensor Network (WSN)

is an infrastructure-less wireless network that is deployed in a large number of wireless sensors in an ad-hoc manner that is used to monitor the system, physical or environmental conditions.

Sensor nodes are used in WSN with the onboard processor that manages and monitors the environment in a particular area. They are connected to the Base Station which acts as a processing unit in the WSN System.
Base Station in a WSN System is connected through the Internet to share data.




WSN can be used for processing, analysis, storage, and mining of the data.
Applications of WSN:

Internet of Things (IOT)
Surveillance and Monitoring for security, threat detection
Environmental temperature, humidity, and air pressure
Noise Level of the surrounding
Medical applications like patient monitoring

Agriculture
Landslide Detection
Challenges of WSN:

Quality of Service
Security Issue
Energy Efficiency
Network Throughput
Performance
Ability to cope with node failure
Cross layer optimisation
Scalability to large scale of deployment
Components of WSN:
Sensors:
Sensors in WSN are used to capture the environmental variables and which is used for data acquisition. Sensor signals are converted into electrical signals.
Radio Nodes:
It is used to receive the data produced by the Sensors and sends it to the WLAN access point. It consists of a microcontroller, transceiver, external memory, and power source.
WLAN Access Point:
It receives the data which is sent by the Radio nodes wirelessly, generally through the internet.
Evaluation Software:
The data received by the WLAN Access Point is processed by a software called as Evaluation Software for presenting the report to the users for further processing of the data which can be used for processing, analysis, storage, and mining of the data.
What Is Wireless Sensor Network?
WSNs stands for Wireless Sensor Networks can be defined as a self-configured and infrastructure-less wireless network to observe physical or environmental conditions, like temperature, pressure, motion, sound, vibration, or pollutants, and to directly pass their data or information through the network to a sink which is also called the main location where the information is often observed and analyzed.
A base station or sink seems like an interface between the users and the network. It can convert back some required information from the network by injecting some queries and gathering results from the sink. Typically a wireless sensor network contains many thousands of sensor nodes.
The sensory nodes can communicate with each other by using radio signals. The wireless sensor nodes are equipped with sensing and radio transceivers, computing devices, and power components.
A sensor node in a wireless sensor network is inherently resource-constrained, also it has limited processing speed, storage capacity, and communication bandwidth. After the sensor nodes are installed, they're responsible for self-organizing an appropriate network infrastructure often with multi-hop communication with them.
Types of Wireless Sensor Networks
There are five types of Wireless Sensor Networks depending on the environment. Different Types of WSNs are:
**1. Terrestrial Wireless Sensor Networks:** Terrestrial WSNs are used for communicating base stations efficiently, and comprise thousands of wireless sensor nodes deployed either in an unstructured (ad hoc) or structured (Pre-planned) manner.
In an unstructured mode (ad hoc), the sensor nodes are randomly distributed within the target area that's dropped from a set plane.
In WSNs, the battery power is limited, however, the battery is provided with solar cells as a secondary power source. The conservation of energy of the WSNs gets by using low duty cycle operations, optimal routing, minimizing delays, and so on.
**2. Underground Wireless Sensor Networks:** In terms of deployment, maintenance, equipment cost considerations, and careful planning, underground wireless sensor networks are more expensive than terrestrial WSNs.

The Underground Wireless sensor networks UWSNs comprises several sensory nodes that are hidden in the ground to observe underground conditions.

Additional sink nodes are located above the bottom to transfer information from the sensor nodes to the base station, These underground WSNs deployed into the ground are difficult to recharge. The sensor battery nodes equipped with limited battery power are also difficult to recharge. Additionally, the underground environment makes wireless communication a challenge because of the high attenuation and signal loss level.

**3. Underwater Wireless Sensor Networks:** About more than 70% of the earth's planet is occupied with water. These networks contain several sensor nodes and vehicles deployed underwater. Autonomous underwater devices and vehicles are used to collect data from these sensor nodes. A challenge of underwater communication may be a long propagation delay, and bandwidth and sensor failures. Underwater, WSNs are equipped with a limited battery that can't be recharged or replaced.

The difficulty of energy conservation for underwater WSNs involves the development of underwater communication and networking techniques.

**4. Multimedia Wireless Sensor Networks:** Multimedia wireless sensor networks are proposed to enable tracking and monitoring of events in the sort of multimedia, like video, imaging, and audio.

**5. Mobile Wireless Sensor Networks MWSNs:** Mobile WSNs networks comprise a group of sensor nodes that can be moved on their own and can be interacted with the physical environment. The mobile nodes can also compute sense and communicate respectively.

‘

**UNIT-3**

# 1.Wireless Technologies for IoT

The global technological advancement is all about dumb physical things interacting and solving problems for 'us,' an intelligent species. The exciting world of 'network of things' or Internet of Things (IoT) is definitely good at sensing our needs and satisfying them.

Communication through radio frequency (RF), light and sound, and Bluetooth allows physical things to communicate with each other without wires, making the IoT a reality.

Radio frequency

RF communication is all about transmitting data through the various available radio frequencies. There are various technologies like Wi-Fi, Zigbee, SigFox and LoRa that communicate at different frequencies and cover a variety of distances. Wi-Fi is used for both short- and mid-range network communication, making it apt for both smart homes and industrial applications. On the other hand, LoRaWAN is the best suited for long-range communication requirements such as in Smart Cities.

Wi-Fi

This is the most commonly used communication standard, assigned as 802.11 by the Institute of Electrical and Electronics Engineers (IEEE). It operates in the 2.4-5GHz band. With the rollout of 5G network, Wi-Fi will play an even bigger role in wireless transmission.

Wi-Fi technology is promoted by the not-for-profit Wi-Fi Alliance. The alliance certifies products with Wi-Fi interoperability. Wi-Fi Alliance also owns the Wi-Fi trademark. However, it should not be the criteria to choose a product. A lot of products support Wi-Fi technology but aren't certified due to the cost involved. Certification only increases the chances of interoperability.

SigFox

In the IoT, the challenge lies in connecting multiple devices across distributed locations in a cost- and energy-efficient manner. Devices are expected to run on a single coin-cell for years without manual intervention.

Gaurav Sareen, country head – India, SigFox, explains, “Volume of data transmitted in M2M is much less than in a voice or video call. So, technology in cellular data becomes overkill for IoT applications. This is where communication networks like Sigfox come in. The transceiver sleeps most of the time and wakes up only to transmit data. Thus it saves a lot of power.”

SigFox can be used in a multitude of applications. Globally, Smart Cities are readily adopting this technology. SigFox is being greatly used in smart waste management, fire-hydrant control, streetlight control, air-quality monitoring and real-time water-level updates. It has a commercialised global network. For Sigfox subscribers no separate subscription is required when travelling across different countries in the network. A lot of ramifications happened downstream in this way.

Sareen says, “We were able to create standardisation for the IoT. Today, the biggest challenge in IoT ecosystem is a lot of fragmentation as there are many connectivity technologies and protocols being adopted. We are trying to drive the market towards standardisation so that device companies can provide you devices off-the-shelf or application companies can provide you applications off-the-shelf. In a business sense, in comparison to LoRA, we do not charge any licence fee. Today, if a device company wants to roll out devices like smart meters that will collect data and forward it to a central command centre over our network, it is not required to pay us any technology licence fee. This will encourage such companies to develop devices that connect immediately out of the shelf.”

ZigBee

ZigBee is a lightweight and flexible RF-based communication technology. Mostly used for industrial applications, it has been assigned standard 802.15.4 by IEEE. ZigBee works in the same bandwidth as Wi-Fi, that is 2.4GHz, but consumes less power and offers a high level of security through 128-bit encryption. However, it enables limited data exchange.

ZigBee-based sensor networks are widely used in multilevel parking monitoring systems, warehouse measurement and control systems, and various environmental monitoring systems. Sensors are spread over the area to transmit information, and the products required to communicate are tagged in the range for flawless data travel.

IoT communication through light

Various companies and products use light as a medium to transfer data between various IoT objects. Li-Fi, IR and laser are three of the most popular IoT communication technologies in this category.

Li-Fi

Similar to Wi-Fi, Li-Fi uses optical wireless communication technology. It utilises visible, infrared and ultrasonic spectrum to transfer data. Li-Fi finds applications in underwater remotely-operated vehicles (as light is a better medium to communicate under water than radio frequencies) and hospitals to communicate patient data (as light travels faster).

Infrared (IR)

Personal communication is best enjoyed with IR. Electromagnetic invisible light range allows data communication over a short distance. Devices used as the medium are personal computers and digital assistants.

Bluetooth

Wireless headphones connect to your cellphone using Bluetooth technology. Assigned 802.15.1 standard by IEEE, Bluetooth is used in various personal and commercial devices. The fact that it can be used for small products like wireless headphones and smart watches makes it a favourite among tech-savvy individuals. Bluetooth Low Energy, a Bluetooth variant launched in 2011, consumes less power.

**2.WPAN :** Wireless Personal Area Network in short known as WPAN. Actually WPAN is PAN (Personal Area Network) where the interconnected devices are centered around a person's workspace and connected through wireless medium. That's why it is also called as Person's centered short range wireless connectivity. Typically the range is within about 10 meters means very short range. Plugging in is one of the key concept in WPAN as within a close proximity WPAN-equipped devices communicate with each other as like they are connected through cable.
Unlike WLAN (Wireless Local Area Network) where there is a requirement of infrastructure setup, in WPAN connection involves little or no infrastructure. In general, if we will see this WPAN provides power efficient and inexpensive solutions to be implemented for a wide range of devices within a short range distance.
As in WPAN mostly the connection between the devices happens within the building or in a room (short range) So, let's take an example of WiFi connection through which two friends are chatting, sharing documents with each other in room.

Some more examples of WPAN includes Wireless mouse, Wearable devices, USB flash drives, Digital cameras, Bluetooth, Wi-Fi, Thermostats, Security systems, Lighting controls, Motion sensors, Personal server and Leak sensors etc.

As per IEEE, Wireless Personal Area Network is classified into 3 classes i.e

1. **High-rate WPAN (HR-WPAN) :** It is defined in the IEEE 802.15.3 standard. Data throughput is > 20 Mbps.
2. **Medium-rate WPAN (MR-WPAN) :** It is defined in the IEEE 802.15.1 standard. Data throughput is 1 Mbp.
3. **Low-rate WPAN (LR-WPAN) :** It is defined in the IEEE 802.15.4 standard. Data throughput is < 0.25 Mbps.

**Applications of WPAN :**

- Short range connectivity for multimedia applications
- Hands free devices connection
- Industrial sensor applications

**WPAN Topologies :**

As mostly use of WPAN is within short range so it is mainly used for general purpose uses and with little industrial requirement implementations which supports below network connectivity arrangements (network topology)

1. Star Topology
2. Mesh Topology
3. Cluster Tree Topology

**Technologies used in WPAN :**

1. Bluetooth
2. ZigBee
3. Infrared
4. Z-wave
5. Wireless Body Area Network (WBAN)

**Features of WPAN :**

- Low cost, Little or No infrastructure setup
- Short range communication
- Small personal network , use anywhere
- Wide range of devices
- Low power consumption
- No complex connectivity

**Advantages of WPAN :**

- Security
- Portability
- Easy Connectivity
- Stability

**Disadvantages of WPAN :**

- Short range
- Transfer speed

## 3.IEEE 802.15.4 Technology

IEEE 802.15.4 is a low-cost, low-data-rate wireless access technology for devices that are operated or work on batteries. This describes how low-rate wireless personal area networks (LR-WPANs) function.

**Properties:**

**1. Standardization and alliances:** It specifies low-data-rate PHY and MAC layer requirements for wireless personal area networks (WPAN).

IEEE 802.15. Protocol Stacks include:

- **ZigBee:** ZigBee is a Personal Area Network task group with a low rate task group 4. It is a technology of home networking. ZigBee is a technological standard created for controlling and

sensing the network. As we know that ZigBee is the Personal Area network of task group 4 so it is based on IEEE 802.15.4 and is created by Zigbee Alliance.

- **6LoWPAN:** The 6LoWPAN system is used for a variety of applications including wireless sensor networks. This form of wireless sensor network sends data as packets and uses IPv6 – providing the basis for the name – IPv6 over Low power Wireless Personal Area Networks.
- **ZigBee IP:** Zigbee is a standards-based wireless technology that was developed for low-cost and low-power wireless machine-to-machine (M2M) and internet of things (IoT) networks.
- **ISA100.11a:** It is a mesh network that provides secure wireless communication to process control.
- **Wireless HART:** It is also a wireless sensor network technology, that makes use of time-synchronized and self-organizing architecture.
- **Thread:** Thread is an IPv6-based networking protocol for low-power Internet of Things devices in IEEE 802.15. 4-2006 wireless mesh network. Thread is independent.

**2. Physical Layer:** This standard enables a wide range of PHY options in ISM bands, ranging from 2.4 GHz to sub-GHz frequencies. IEEE 802.15.4 enables data transmission speeds of 20 kilobits per second, 40 kilobits per second, 100 kilobits per second, and 250 kilobits per second. The fundamental structure assumes a 10-meter range and a data rate of 250 kilobits per second. To further reduce power usage, even lower data rates are possible. IEEE 802.15.4 regulates the RF transceiver and channel selection, and even some energy and signal management features, at the physical layer. Based on the frequency range and data performance needed, there are now six PHYs specified. Four of them employ frequency hopping techniques known as Direct Sequence Spread Spectrum (DSSS). Both PHY data service and management service share a single packet structure so that they can maintain a common simple interface with MAC.

**3. MAC layer:** The MAC layer provides links to the PHY channel by determining that devices in the same region will share the assigned frequencies. The scheduling and routing of data packets are also managed at this layer. The 802.15.4 MAC layer is responsible for a number of functions like:

- Beaconing for devices that operate as controllers in a network.
- used to associate and dissociate PANs with the help of devices.
- The safety of the device.
- Consistent communication between two MAC devices that are in a peer-to-peer relationship.

Several established frame types are used by the MAC layer to accomplish these functions. In 802.15.4, there are four different types of MAC frames:

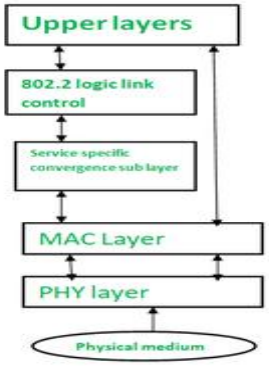
- frame of data
- Frame for a beacon
- Frame of acknowledgement
- Frame for MAC commands

**4. Topology:** Networks based on IEEE 802.15.4 can be developed in a star, peer-to-peer, or mesh topology. Mesh networks connect a large number of nodes. This enables nodes that would otherwise be out of range to interact with each other to use intermediate nodes to relay data.

**5. Security:** For data security, the IEEE 802.15.4 standard employs the Advanced Encryption Standard (AES) with a 128-bit key length as the basic encryption technique. Activating such security measures for 802.15.4 significantly alters the frame format and uses a few of the payloads. The very first phase in activating AES encryption is to use the Security Enabled field in the Frame Control part of the 802.15.4 header. For safety, this field is a single bit which is assigned to 1. When this bit is set, by taking certain bytes from its Payload field, a field known as the Auxiliary Security Header is formed following the Source Address field.

**6. Competitive Technologies:** The IEEE 802.15.4 PHY and MAC layers serve as a basis for a variety of networking profiles that operate in different IoT access scenarios. DASH7 is a competing radio technology with distinct PHY and MAC layers.

**The architecture of LR-WPAN Device:**




*IEEE 802.15.4*

**Advantages of IEEE 802.15.4:**

IEEE 802.15.4 has the following advantages:

- cheap cost
- long battery life,
- Quick installation
- simple
- extensible protocol stack

**Disadvantages of IEEE 802.15.4:**

IEEE 802.15.4's drawbacks include:

- IEEE 802.15.4 causes interference and multipath fading.
- doesn't employ a frequency-hopping approach.
- unbounded latency
- interference susceptibility

**Applications of IEEE 802.15.4:**

IEEE 802.15.4 Applications:

- Wireless sensor networks in the industry
- Building and home automation
- Remote controllers and interacting toys
- Automotive networks

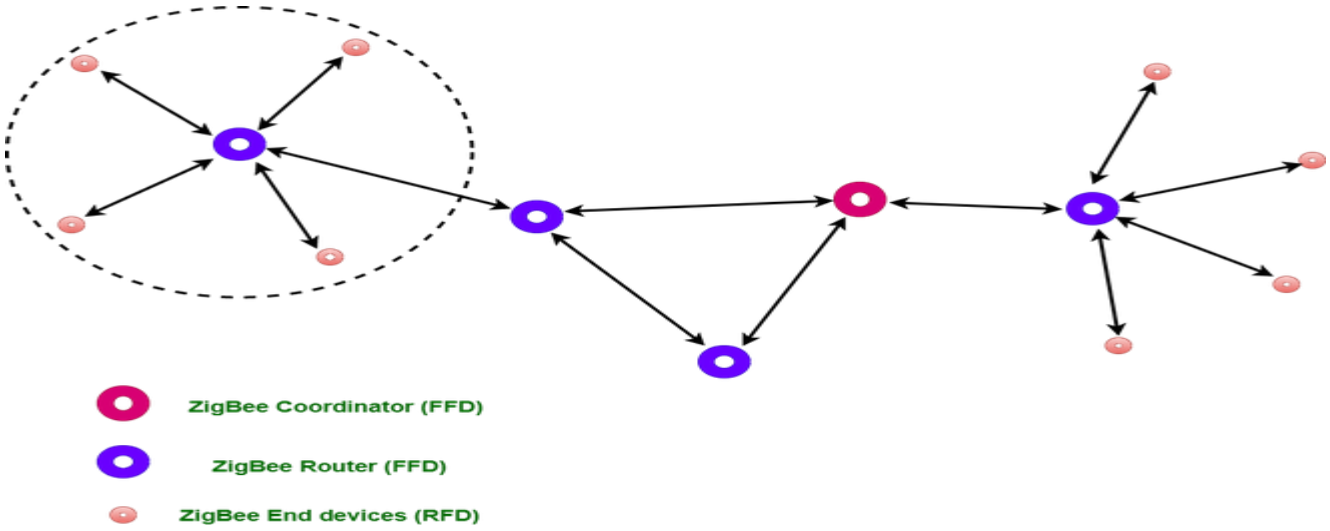
# 4.ZigBee

ZigBee is a Personal Area Network task group with low rate task group 4. It is a technology of home networking. ZigBee is a technological standard created for controlling and sensing the network. As we know that ZigBee is the Personal Area network of task group 4 so it is based on IEEE 802.15.4 and is created by Zigbee Alliance.

ZigBee is a standard that addresses the need for very low-cost implementation of Low power devices with Low data rates for short-range wireless communications.

**Types of ZigBee Devices:**

- **Zigbee Coordinator Device:** It communicates with routers. This device is used for connecting the devices.
- **Zigbee Router:** It is used for passing the data between devices.
- **Zigbee End Device:** It is the device that is going to be controlled.




**General Characteristics of Zigbee Standard:**

- Low Power Consumption
- Low Data Rate (20- 250 kbps)
- Short-Range (75-100 meters)
- Network Join Time (~ 30 msec)
- Support Small and Large Networks (up to 65000 devices (Theory); 240 devices (Practically))
- Low Cost of Products and Cheap Implementation (Open Source Protocol)
- Extremely low duty cycle.
- 3 frequency bands with 27 channels.

**Operating Frequency Bands** (Only one channel will be selected for use in a network):

1. **Channel 0**: 868 MHz (Europe)
2. **Channel 1-10**: 915 MHz (the US and Australia)
3. **Channel 11-26**: 2.4 GHz (Across the World)

**Zigbee Network Topologies:**

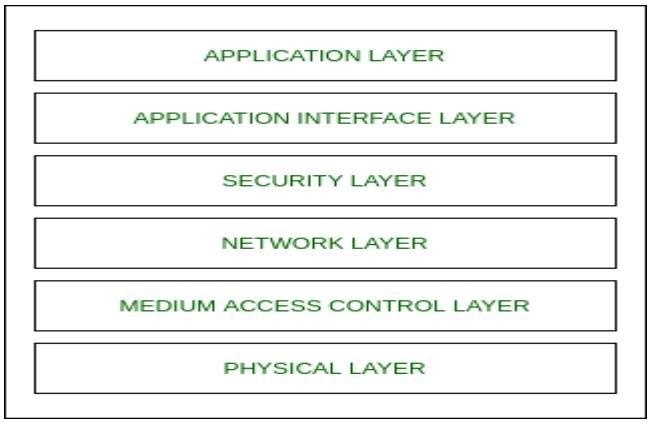
- **Star Topology** (ZigBee Smart Energy): Consists of a coordinator and several end devices, end devices communicate only with the coordinator.
- **Mesh Topology** (Self Healing Process): Mesh topology consists of one coordinator, several routers, and end devices.
- **Tree Topology**: In this topology, the network consists of a central node which is a coordinator, several routers, and end devices. the function of the router is to extend the network coverage.

**Architecture of Zigbee:**

Zigbee architecture is a combination of 6 layers.

1. Application Layer
2. Application Interface Layer
3. Security Layer

4. Network Layer
5. Medium Access Control Layer
6. Physical Layer




- **Physical layer:** The lowest two layers i.e the physical and the MAC (Medium Access Control) Layer are defined by the IEEE 802.15.4 specifications. The Physical layer is closest to the hardware and directly controls and communicates with the Zigbee radio. The physical layer translates the data packets in the over-the-air bits for transmission and vice-versa during the reception.
- **Medium Access Control layer (MAC layer):** The layer is responsible for the interface between the physical and network layer. The MAC layer is also responsible for providing PAN ID and also network discovery through beacon requests.
- **Network layer:** This layer acts as an interface between the MAC layer and the application layer. It is responsible for mesh networking.
- **Application layer:** The application layer in the Zigbee stack is the highest protocol layer and it consists of the application support sub-layer and Zigbee device object. It contains manufacturer-defined applications.

**Channel Access:**

1. **Contention Based Method** (Carrier-Sense Multiple Access With Collision Avoidance Mechanism)
2. **Contention Free Method** (Coordinator dedicates a specific time slot to each device (Guaranteed Time Slot (GTS)))

**Zigbee Applications:**

1. Home Automation
2. Medical Data Collection
3. Industrial Control Systems
4. meter reading system
5. light control system

# 5.Hart & Wireless Hart in IoT

- **Wireless HART** is the latest release of Highway Addressable Remote Transducer (HART) Protocol.
- HART standard was developed for networked smart field devices.
- The wireless protocol makes the implementation of HART cheaper and easier.
- HART encompasses the most number of field devices incorporated in any field network.
- Wireless HART enables device placements more accessible and cheaper - such as the top of a reaction tank, inside a pipe, or at widely separated warehouses.
- Main difference between wired and unwired versions is in the physical, data link and network layers.
- Wired HART lacks a network layer.

**HART Physical Layer**

- Derived from IEEE 802.15.4 protocol.
- It operates only in the 2.4 GHz ISM band.
- Employs and exploits 15 channels of the band to increase reliability.

**HART Data Link Layer**

- Collision free and deterministic communication achieved by means of super-frames and TDMA.
- Super-frames consist of grouped 10ms wide timeslots.
- Super-frames control the timing of transmission to ensure collision free and reliable communication.
- This layer incorporates channel hopping and channel blacklisting to increase reliability and security.
- Channel blacklisting identifies channels consistently affected by interface and remove them from use.

**HART Network & Transport Layers**

- Cooperatively handle various types of traffic, routing, session creation, and security.
- WirelessHART relies on **Mesh networking** for its communication, and each device is primed to forward packets from every other devices.
- Each device is armed with an updated network graph (i.e., updated topology) to handle routing.
- **Network layer (HART) = Network + Transport + Session layers (OSI)**.

**HART Application Layer**

- Handles communication between gateways and devices via a series of **command and response messages**.
- Responsible for **extracting** commands from a message, **executing** it and generating responses.
- This layer is seamless and does not differentiate between wireless and wired versions of HART.

**HART Congestion Control**

- Restricted to 2.4Ghz ISM band with channel 26 removed , due to its restricted usage in certain areas.
- Interference-prone channels avoided by using channel switching post every transmission.
- Transmissions synchronized using 10ms slots.
- During each slot, all available channels can be utilized by the various nodes in the network allowing for the propagation of 15 packets through the network at a time, which also minimizes the risk of collisions.

## 6.NFC:

NFC (Near-Field Communication) is a series of communication protocol that allows two electronic devices to communicate over a distance. Since its release, NFC has been used in more and more different applications, for security measure, convenience, and even transactions!

How does NFC work?

NFC is based on radio-frequency identification (RFID) technology, which allows compatible hardware to use radio waves to both controls and communicate with otherwise unpowered and passive electronic tags

More specifically, NFC transmits via Electromagnetic Induction, which can induce electric currents within passive components as well. This means that passive devices can be powered by the electromagnetic field produced by an active NFC component, and don’t need their own power supply.

Data is transmitted at a frequency of 13.56 megahertz over NFC. You have the option of sending data at 106, 212, or 424 kilobits per second. That's quick enough for a variety of data transfers, from contact information to sharing photos and music.

**NFC Three Modes of Data Exchange**

The NFC standard currently has three distinct modes of operation.

- **Reader/Writer mode** – One-way data transmission where the active device, which may be your smartphone, establishes a link with another device in order to read data from it. This mode is used by NFC advertisement tags.

- **Peer-to-Peer mode** – This enables two NFC-enabled devices to share different types of data. Both devices transform from active to passive when sending and receiving data in this mode. Most common use in smartphones.

- **Card Emulation mode** – The NFC device can be used to make purchases or tap into public transportation networks as a smart or contactless credit card, i.e. Google Pay and Apple Pay

# 7.Z-Wave

**Z-Wave Protocol Stack :**

Z-Wave protocol stack contains five layers physical layer, MAC layer, transport layer, network layer, and application layer.

- **PHY layer:** This layer has many functions but the important one is modulation and coding. In this layer, data is transferred in 8-bit blocks and the most significant bit is sent first.
- **MAC layer:** MAC layer as the name suggests takes care of medium access control among slave nodes based on collision avoidance and backoff algorithms. also, it takes care of network operations based on Home ID, Node ID, and other parameters in the z-wave frame.
- **Transport layer:** Z-Wave transport layer is mainly responsible for retransmission, packet acknowledgment, and packet origin authentication. the z-wave layer consists of four basic frame types:
    - Single cast frame
    - ACK frame
    - Multicast frame
    - Broadcast frame
- **Network layer:** Z-Wave network layer controls the frame routing from one node to another node.
- **Application layer:** This layer is responsible for decoding and execution of commands in the z-wave network.

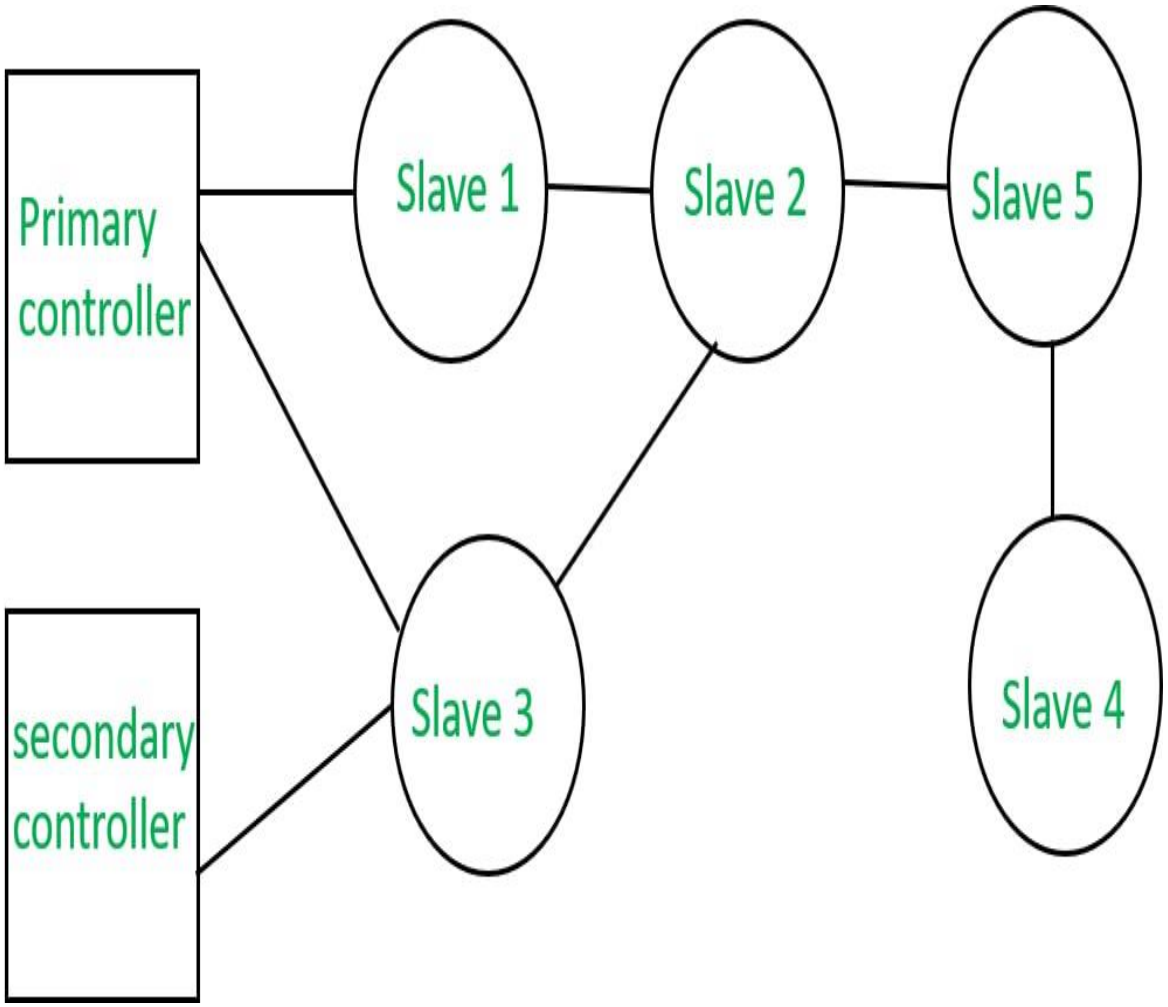
**Z-Wave Components :**

The components of z-wave include controllers, slave nodes, Home ID, Node ID, and routing tables.

- **Controllers:** A controller is a unit that has the ability to compile a routing table of the network and can calculate routes to the different nodes. There are two types of controllers –
    - Primary controller: Primary controller is the device that contains a description of the z-wave network and controls the output. It assigns network ID or Home ID or Node ID to the z-wave during the enrollment process.
    - Secondary controller: It also has a Network ID and it remains constant to maintain routing tables.
- **Slave nodes:** Slave nodes are the nodes that do not contain routing tables but may contain a network map. slave nodes have the ability to receive frames and respond to them if necessary.

- **Home ID:** The ID used by z-Wave for the separation of the network from each other is called Home ID. It is created by the primary controller and is 32-bit in size.
- **Node ID:** The identification number or an address that is given to every device during the process of inclusion is called Node ID.
- **Routing table:** It is used by controllers for calculating routes.

The following diagram shows us z-wave network .




**Characteristics of Z-Wave :**

- Uses RF for signaling and control
- Frequency : 900 MHz (ISM)
- Range : 30 meter
- Data rates : upto 100 kbps
- FSK Modulation

**Applications of Z-Wave :**

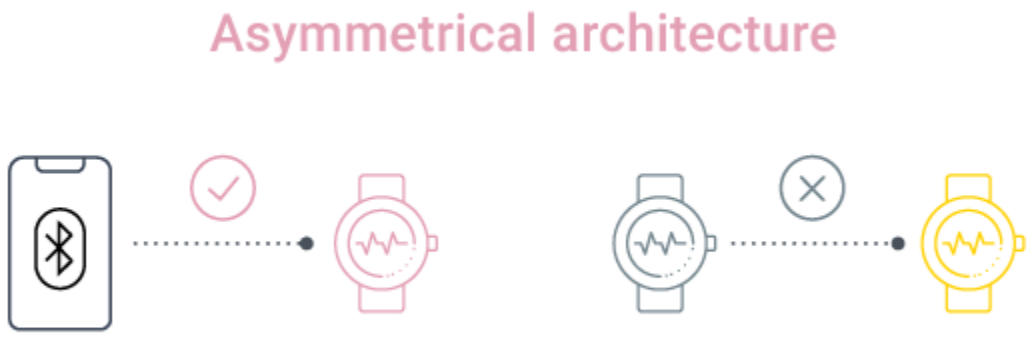
- Home automation
- Water management using flood sensors
- Fingerprint scanner

# 8.Bluetooth Low Energy (BLE)

**What is Bluetooth Low Energy?**

Bluetooth Low Energy is a wireless, low-power personal area network that operates in the 2.4 GHz ISM band. Its goal is to connect devices over a relatively short range. BLE was created with IoT applications in mind, which has particular implications for its design. For example, IoT devices tend to be constrained and require extended battery use, so BLE favors low power consumption over continuous data transfer. In other words: when not in use, it goes into sleep mode to conserve energy.

What is essential to understand when talking about BLE-equipped devices is the architecture behind the technology, particularly its asymmetry. A device may function either in a central or peripheral role. Imagine your smartphone and your smart band: the more "advanced" and complex smartphone is the central device; the smart band, which has limited functionality – is the peripheral. Neither two central nor two peripheral devices can talk to each other. Communication is only possible between a central and a peripheral device. To overcome that limitation, a device may have both a central and a peripheral mode configured, as many smartphones do.




The device mode should not be confused with its Generic Attribute Profile (GATT), where a device can either act as a server or a client. Once two devices establish a connection, the one that sends out the data is the server, and the one that receives it is the client. So if a smart band sends, let's say, heart rate monitor readings to the smartphone, then it acts as a server. But if the smartphone sends a software update to the smart band, then the smartphone is the server.

**Is BLE the same as Bluetooth?**

One more important question that we need to answer when talking about Bluetooth Low Energy is how it differs from Bluetooth Classic – the technology we all know for its famous icon that we click on when we want to pair our devices. Does it differ at all?

The answer is yes. BLE is an independent standard that is incompatible with the "classic" Bluetooth. The latter was first introduced commercially over 20 years ago and is now essentially no longer being developed by the Bluetooth Special Interests Group (SIG). However, "not being developed" doesn't mean "not being used." You will frequently find it in devices that require continuous connection, predominantly audio devices, such as wireless speakers or headphones. Meanwhile, SIG introduced Bluetooth Low Energy in its 2010 Bluetooth 4.0 specification (with later, 2016 Bluetooth 5 specification, being devoted exclusively to BLE). Its main focus was on the growing market of health- and fitness-related devices and smart home and indoor location.

| | Bluetooth Classic | Bluetooth Low Energy |
|---|---|---|
| communication | continuous, bidirectional | short data transfers in one direction |
| range | 100 m | <100 m |
| energy consumption | 1 W | 0.01–0.50 W |
| data rate | 1–3 Mbit/s | 125 kbit/s – 2 Mbit/s |
| latency | 100 ms | 6 ms |
| voice capable | yes | no |

# 9.BACnet and Modbus

Communication protocols often come up in conversations about future-proofing buildings or buildings that are “open.” While the language that a protocol speaks is important, the transmission of the protocol is also critical. A protocol might be in use for the next decade or so, but if the communication media to support that protocol is problematic to install or no longer in use—whether it be through wireless or a physical wire—then nothing is going to help the building owner in the future. This article discusses the importance of both the protocol and the protocol transmission.

BACnet and Modbus are the two open communication protocol standards that building management systems (BMS) often utilize today in applications such as energy monitoring and temperature, lighting, and occupancy controls. Multiple manufacturers and governing authorities developed and implemented these communication protocols, which allowed for the “open standard.” BACnet was developed in 1987 and became an ANSI standard in 1995. Trusty Modbus has been around since 1979.

**BACnet: Widely known and used**

Created and driven by ASHRAE, BACnet (Building Automation Communication network) is the most widely used communication protocol in the industry. The two main types of BACnet implementations are BACnet MS/TP and BACnet/IP.

BACnet MS/TP (master-slave/token passing) is an older implementation where system integrators run twisted pair wiring (RS-485 standard) through the building as a separate network. BACnet MS/TP communicates at slow baud rates and cannot handle defaults or changes in the physical network configuration. It must be installed with specific end-of-line resistors and biasing resistors. Despite its slow speed, as compared to other communication protocols and implementations, and its limited fault tolerance, BACnet MS/TP is probably the most common installation today.

BACnet/IP is an emerging implementation based on Internet Protocol (IP), meaning it can communicate across an Ethernet cable or IP network. Instead of running a standard MS/TP RS-485 twisted pair of wire, system integrators can plug in a BACnet/IP controller or device as they would a computer. The devices have a network jack that accepts an Ethernet cable. Then a system integrator or IT/OT professional sets up a network and it works.

Generally speaking, building automation engineers or system integrators prefer BACnet, regardless of whether it’s MS/TP or /IP. Integrators can enter a building, plug in a computer, conduct a BACnet scan, see the devices, see what data points (such as ambient temperature or occupancy) are in those devices, and then add these points to the BMS or building automation system (BAS) database.

Of course, integrators must deal with network pieces and parts, such as firewall rules, subnetting, and BBMDs (BACnet/IP broadcast management devices). However, BACnet/IP is a more fault-tolerant installation than if on a serial bus or twisted pair installation due to the nature of IP network designs.

***Modbus, the reliable runner-up***

Modbus is widely used in industrial environments, such as electrical switchgears. Factories use Modbus for programmable logic controllers (PLCs), and data centers use it for power distribution units (PDUs). Modbus is also prevalent in Europe and the Asia-Pacific region.

Like BACnet, Modbus offers different implementations. Remote terminal unit (RTU) uses a twisted pair of serial bus wires, while transmission control protocol (TCP) uses the Ethernet compatibility implementation.

Unlike BACnet, Modbus does not offer network discoverability. Integrators need a Modbus Register—essentially a blueprint or roadmap of the communication points in a building—along with the data point address numbers. Integrators manually type in those address numbers and then configure each to read the proper value. Though its setup requires more upfront work than BACnet, Modbus works quite nicely once it is up and running.

BACnet and Modbus are long-standing protocols. However, for BAS and, more specifically, the Internet of Things (IoT), one of the latest protocols on the block is Message Queue Telemetry Transport (MQTT). With minimal bandwidth requirements, this lightweight, open messaging protocol provides an efficient way to communicate. As edge devices such as sensors become smarter, MQTT will likely continue to gain market share as the IoT protocol of choice.

## 10.Internet Protocol version 6 (IPv6)

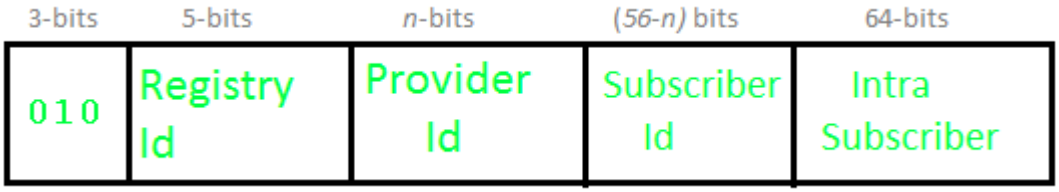
IP v6 was developed by Internet Engineering Task Force (IETF) to deal with the problem of IP v4 exhaustion. IP v6 is a 128-bits address having an address space of 2^128, which is way bigger than IPv4. In IPv6 we use Colon-Hexa representation. There are 8 groups and each group represents 2 Bytes.


In IPv6 representation, we have three addressing methods :

- Unicast
- Multicast
- Anycast

**1. Unicast Address –**
Unicast Address identifies a single network interface. A packet sent to a unicast address is delivered to the interface identified by that address.
**2. Multicast Address –**
Multicast Address is used by multiple hosts, called as Group, acquires a multicast destination address. These hosts need not be geographically together. If any packet is sent to this multicast address, it will be distributed to all interfaces corresponding to that multicast address.

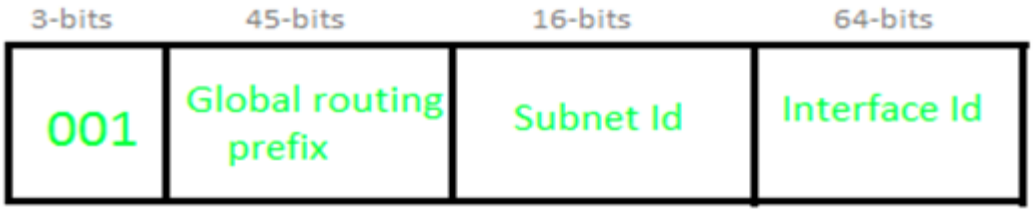
**3. Anycast Address –**
Anycast Address is assigned to a group of interfaces. Any packet sent to an anycast address will be delivered to only one member interface (mostly nearest host possible).
**Note:** Broadcast is not defined in IPv6.
**Types of IPv6 address:**
We have 128 bits in IPv**Provider-based Unicast address :**
These are used for global communication.




The First 3 bits identify it as of this type.
**Registry Id (5-bits):** Registry Id identifies the region to which it belongs. Out of 32 (i.e. 2^5), only 4 registry IDs are being used.

| Registry Id | Registry |
|---|---|
| 10000 | Multi regional (IANA) |
| 01000 | RIPE NCC |
| 11000 | INTER NIC |
| 00100 | APNIC |

**Provider Id:** Depending on the number of service providers that operate under a region, certain bits will be allocated to the Provider Id field. This field need not be fixed. Let's say if Provider Id = 10 bits then Subscriber Id will be 56 – 10 = 46 bits.
**Subscriber Id:** After Provider Id is fixed, the remaining part can be used by ISP as a normal IP address.
**Intra Subscriber:** This part can be modified as per the need of the organization that is using the service.
**Geography based Unicast address :**

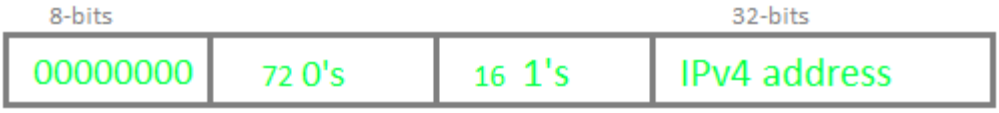



**Global routing prefix:** Global routing prefix contains all the details of Latitude and Longitude. As of now, it is not being used. In Geography-based Unicast address routing will be based on location.
**Interface Id:** In IPv6, instead of using Host Id, we use the term Interface Id.
**Some special addresses:**
**Unspecified –**

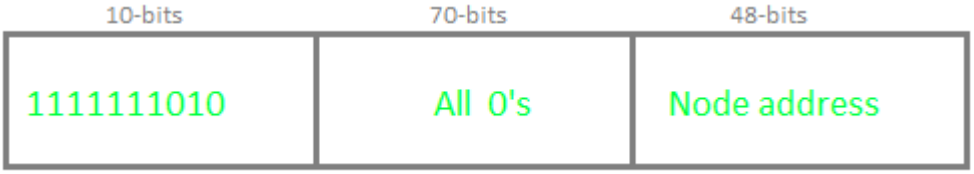
**Loopback –**


**IPv4 Compatible –**


**IPv4 mapped –**




**Local Unicast Addresses :**
There are two types of Local Unicast addresses defined- *Link-local* and *Site-Local*.
**Link-local address:**




A link-local address is used for addressing a single link. It can also be used to communicate with nodes on the same link. The link-local address always begins with 1111111010 (i.e. FE80). The router will not forward any packet with Link-local address.

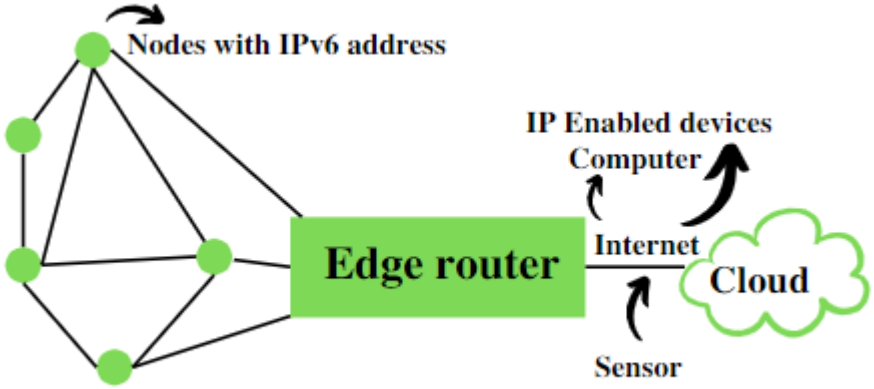
**Site local address:**


6 address but by looking at the first few bits we can identify what type of address it is.

## 11.6LoWPAN

LoWPAN is an IPv6 protocol, and It's extended from is IPv6 over Low Power Personal Area Network. As the name itself explains the meaning of this protocol is that this protocol works on Wireless Personal Area Network i.e., WPAN.
WPAN is a Personal Area Network (PAN) where the interconnected devices are centered around a person's workspace and connected through a wireless medium. You can read more about WPAN at WPAN. 6LoWPAN allows communication using the IPv6 protocol. IPv6 is Internet Protocol Version 6 is a network layer protocol that allows communication to take place over the network. It is faster and more reliable and provides a large number of addresses.
6LoWPAN initially came into existence to overcome the conventional methodologies that were adapted to transmit information. But still, it is not so efficient as it only allows for the smaller

devices with very limited processing ability to establish communication using one of the Internet Protocols, i.e., IPv6. It has very low cost, short-range, low memory usage, and low bit rate.




- It is a technology that makes the individual nodes IP enabled.
- 6LoWPAN can interact with 802.15.4 devices and also other types of devices on an IP Network. For example, Wi-Fi.
- It uses AES 128 link layer security, which AES is a block cipher having key size of 128/192/256 bits and encrypts data in blocks of 128 bits each. This is defined in IEEE 802.15.4 and provides link authentication and encryption.

**Basic Requirements of 6LoWPAN:**

1. The device should be having sleep mode in order to support the battery saving.
2. Minimal memory requirement.
3. Routing overhead should be lowered.

**Features of 6LoWPAN:**

1. It is used with IEEE 802.15,.4 in the 2.4 GHz band.
2. Outdoor range: ~200 m (maximum)
3. Data rate: 200kbps (maximum)
4. Maximum number of nodes: ~100

**Advantages of 6LoWPAN:**

1. 6LoWPAN is a mesh network that is robust, scalable, and can heal on its own.
2. It delivers low-cost and secure communication in IoT devices.
3. It uses IPv6 protocol and so it can be directly routed to cloud platforms.
4. It offers one-to-many and many-to-one routing.
5. In the network, leaf nodes can be in sleep mode for a longer duration of time.

**Disadvantages of 6LoWPAN:**

1. It is comparatively less secure than Zigbee.
2. It has lesser immunity to interference than that Wi-Fi and Bluetooth.
3. Without the mesh topology, it supports a short range.

**Applications of 6LoWPAN:**

1. It is a wireless sensor network.
2. It is used in home-automation,
3. It is used in smart agricultural techniques, and industrial monitoring.

## 12.RPL

The RPL protocol is a distance vector proactive routing protocol that creates a tree-like routing topology called the destination-oriented directed acyclic graph (DODAG), rooted towards one or more nodes called the root node or sink node. The directed acyclic graphs (DAGs) are created based on a user-specified specific objective function (OF). The OF defines the method to find the best-optimized route among the number of sensor devices [14]. The IETF ROLL working group standardized the objective function zero (OF0) [19] and the minimum rank with hysteresis objective

function (MRHOF) as default routing metrics defined in RFC 6719 [20]. The OF0 finds the shortest path to the sink node by selecting the candidate parent node with minimum rank in terms of the distance from the sink (i.e., its position in the routing tree). The MRHOF finds the routes through the sensor nodes that minimize the link cost associated with the routes. It selects the new routing path if the cost associated with it is less than the current path cost by a given threshold value. This is known as 'hysteresis.' As prescribed in the standard, MRHOF utilizes the expected transmission count (ETX) metric which calculates the link quality. Furthermore, ETX considers link-layer congestion but does not reflect node level congestion. Therefore, selecting a routing path based on the smallest hop count and link quality in a heterogeneous traffic environment does not lead to an efficient load-balancing solution. The nodes closer to the sink node suffer from packet loss due to a high relay burden in a dense networking environment. The chosen parent node in RPL can have multiple child nodes; consequently, the overloaded preferred parent becomes prone to failure because its energy drains much faster than other nodes. The inefficient OFs lead to building a routing topology that experiences an excessively unbalanced load and energy distribution, particularly for those nodes that are closest to the sink node.

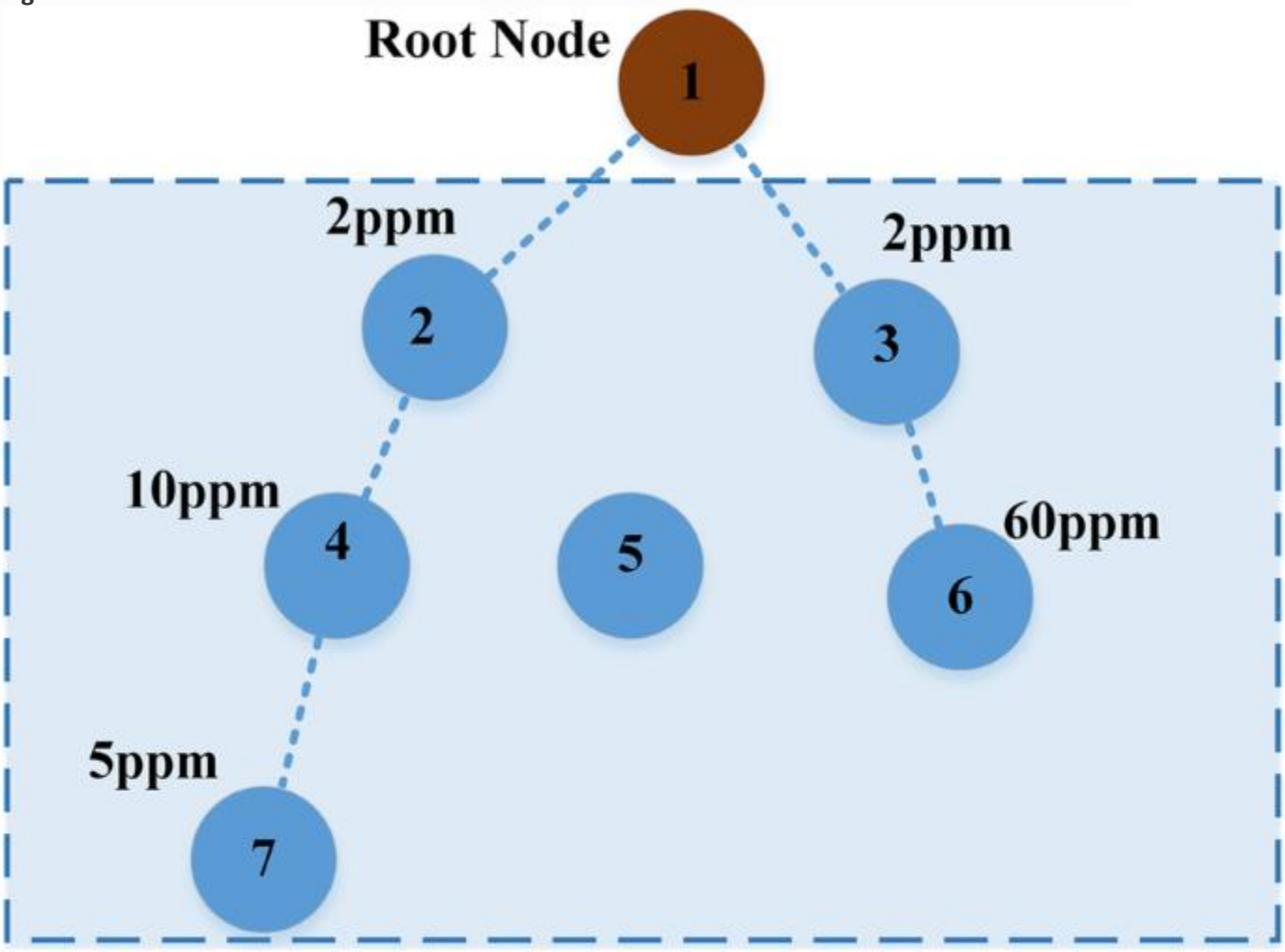
The diverse applications of LLNs include scenarios ranging from basic temperature measurements to high-volume multimedia services that require efficient communication support. The LLN heterogeneous traffic environment suffers from severe congestion and packet loss by not utilizing the full network capacity. A high relay burden, unbalanced load, and limited resources eventually lead to node failure. The ETX-based node broadcasts probe packets at time intervals to assess the link quality. The receiving node rebroadcasts the probe packet which further increases the network congestion. If there is a node failure, RPL initiates two types of repairs: local repair and global repair. In the local repair scenario, a child node routes the packets through its sibling node or the child node switches to the parent node. The global repair is initiated by a gateway or sink node. In both cases, the network incurs overall delay and additional control overhead, which, in turn, becomes a detriment to the overall network performance. When the number of failed nodes increases significantly, the network is split in such a way that communicating with the sink node is not possible through any path, resulting in a non-operational network [14, 20].

The generic definition of heterogeneity is a lack of uniformity. In terms of heterogeneous traffic load, some of the nodes are traffic-intensive while others generate traffic with a low traffic rate. In a normal scenario, the sensor nodes form a tree topology to transfer data towards the root node. The traffic flow is homogeneous with a constant period. In this case, the resulting traffic pattern is predictable and load imbalance is unlikely. In the heterogeneous case, the transmission interval, as well as transmission load, differ in each node; thus, the load fluctuates unpredictably causing a significant load imbalance problem.

In a heterogeneous traffic load scenario, the load imbalance may not be due to the subtree size, as illustrated in Fig. 2, where node 5 has a choice to select the node 2 or node 3 as its parent node. As node 3 contain only one child node, node 5 is likely to select node 3 as its parent node. This scenario is likely to happen in the same traffic load case. However, in this network, the accumulated workload at node 3 is greater than that of node 2 due to the high packet generation rate of node 6, i.e., 60 ppm (packets per minute). Therefore, having more nodes in the routing subtree does not always mean that more traffic flows through that node in a heterogeneous traffic network. In [20], it is observed that parent nodes with large subtree sizes have large queue losses. However, in a heterogeneous traffic pattern, it may not always be true that a parent node with large subtree will have large queue losses. In heterogeneous traffic, the imbalanced loads and queue losses are primarily due to the unbalanced traffic generation as compared with an unbalanced subtree size. The ETX primarily reflects the link-level losses rather than node-level packet losses. On the contrary, the hop count only reflects the number of hops irrespective of traffic load and queue losses. In the same way, utilizing only the queue size as a decision metric does not truly reflect the load condition. The workload of each node is an important parameter in an uneven traffic load pattern. Similarly, the protocol design for such cases must consider latency or average end-to-end (E2E) delay and the amount of overhead as an important performance parameter. Therefore, it is necessary to consider

a new routing metric that solves the load balancing problem in a heterogeneous traffic load scenario.

**Fig. 2**




Illustrating the load imbalance problem due to heterogeneous traffic rate

## 12.REST API

**R**epresentational **S**tate **T**ransfer (REST) is an architectural style that defines a set of constraints to be used for creating web services. **REST API** is a way of accessing web services in a simple and flexible way without having any processing.

REST technology is generally preferred to the more robust Simple Object Access Protocol (SOAP) technology because REST uses less bandwidth, simple and flexible making it more suitable for internet usage. It's used to fetch or give some information from a web service. All communication done via REST API uses only HTTP request.


In **HTTP** there are five methods that are commonly used in a REST-based Architecture i.e., POST, GET, PUT, PATCH, and DELETE. These correspond to create, read, update, and delete (or CRUD) operations respectively. There are other methods which are less frequently used like OPTIONS and HEAD. In **HTTP** there are five methods that are commonly used in a REST-based Architecture i.e., POST, GET, PUT, PATCH, and DELETE. These correspond to create, read, update, and delete (or CRUD) operations respectively. There are other methods which are less frequently used like OPTIONS and HEAD.

- **GET:** The HTTP GET method is used to **read** (or retrieve) a representation of a resource. In the safe path, GET returns a representation in XML or JSON and an HTTP response code of 200 (OK). In an error case, it most often returns a 404 (NOT FOUND) or 400 (BAD REQUEST).

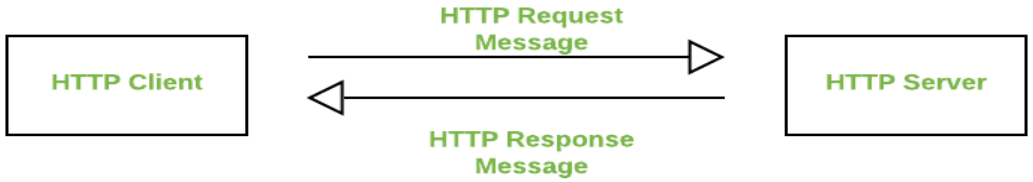
- **POST:** The POST verb is most often utilized to **create** new resources. In particular, it's used to create subordinate resources. That is, subordinate to some other (e.g. parent) resource. On successful creation, return HTTP status 201, returning a Location header with a link to the newly-created resource with the 201 HTTP status.

## 13.AMQP :

AMQP is an acronym used for the Advanced Message Queuing Protocol. It is a protocol that is used for communication between applications. It is a lightweight, protocol which supports the applications for transfer of data. This protocol is used for its scalability and modularity with the technologies.

**HTTP :**

HTTP is an acronym used for Hyper Text Transfer Protocol. It is a protocol that is used for the communication between client and server. This protocol is responsible for the response from the server-side to the client response. It is a base protocol for the communication of web services.




HTTP Model

**14.Constrained Application Protocol (COAP):** The constrained application protocol is a client server-based protocol. With this protocol, the COAP packet can be shared between different client nodes which are commanded by the COAP server. The server is responsible to share the information depending on its logic but has not acknowledged it. This is used with the applications which support the state transfer model. **Message Query Telemetry Transport (MQTT):** The message query telemetry transport protocol is a communication-based protocol that is used for IoT devices. This protocol is based on the publish-subscribe methodology in which clients receive the information through a broker only to the subscribed topic. A broker is a mediator who categorizes messages into labels before being delivered.

**Difference between COAP and MQTT protocols:**

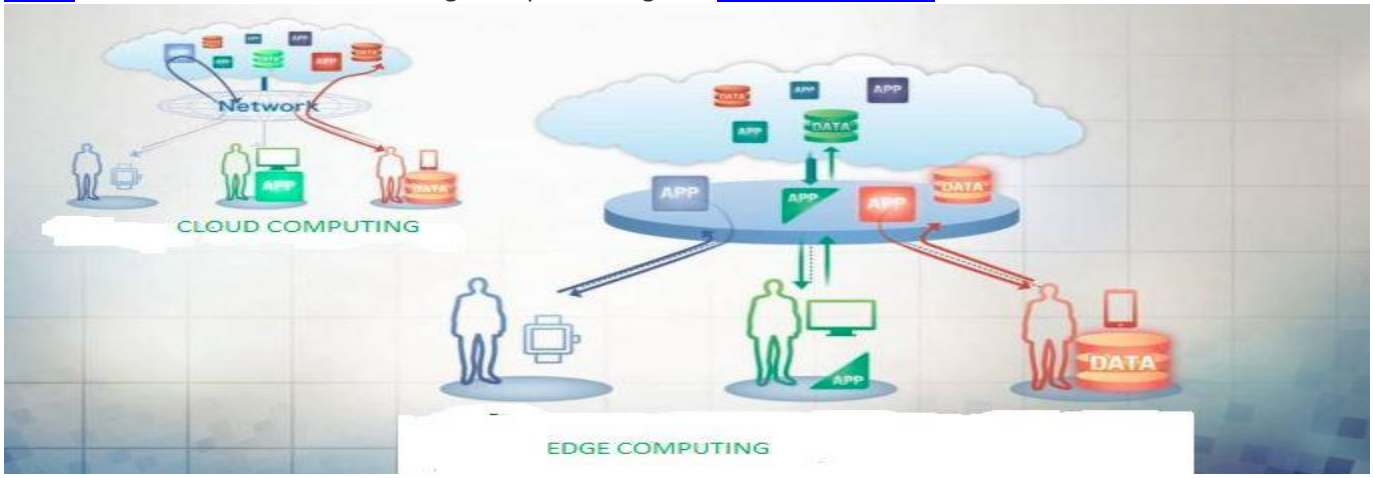
| Basis of | COAP | MQTT |
|---|---|---|
| Abbreviation | Constrained Application Protocol | Message Query Telemetry Transport |
| Communication Type | It uses Request-Response model. | It uses Publish-Subscribe model |
| Messaging Mode | This uses both Asynchronous and | This uses only Asynchronous |

| Basis of | COAP | MQTT |
|---|---|---|
| | Synchronous. | |
| Transport layer protocol | This mainly uses User Datagram protocol(UDP) | This mainly uses Transmission Control protocol(TCP) |
| Header size | It has 4 bytes sized header | It has 2 bytes sized header |
| RESTful based | Yes it uses REST principles | No it does not uses REST principles |
| Persistence support | It does not has such support | It supports and best used for live data communication |
| Message Labelling | It provides by adding labels to the messages. | It has no such feature. |
| Usability/Security | It is used in Utility area networks and has secured mechanism. | It is used in IoT applications and is secure |
| Effectiveness | Effectiveness in LNN is excellent. | Effectiveness in LNN is low. |
| Communication Model | Communication model is one-one. | Communication model is many-many. |

## 15.Edge Technology

Edge Technology aims at making Internet Of Things (IOT) with 100 thousands of sensors in next decade, with the increased usage and manipulation of large data it becomes important to get used to this technology which refers to computing on sensor itself. 2019 is predicted as the year of edge technology and will remain so in the coming years.
In a variety of situations edge computing is deployed. One is when IOT devices is centrally connected to cloud due to poor connectivity of devices. By the year 2020, there will be approximately 1.5 GB worth of data is generated per day. With many devices connected to the internet and generating data, its not possible for cloud alone to handle this huge data all by itself.

**Edge** can relate to data processing as well as local processing of the real time data. The various edge components that can be counted upon are Data processing, Rule Engine, Local Database. Cloud is more concerned with the big data processing and data warehousing.



**Why Edge Computing?**

- This technology increases the efficient usage of bandwidth by analyzing the data at edges itself unlike the cloud which requires transfer of data from the IOT requiring large bandwidth, making it useful to be used in remote location with minimum cost.
- It allows smart applications and devices to respond to data almost at the same time which is important in terms of business ad self driving cars.
- It has the ability to process data without even putting on a public cloud, this ensures full security.
- Data might get corrupt while on an extended network thus affecting the data reliability for the industries to use.
- Edge computation of data provides a limitation to the use of cloud.

**Edge vs Fog Computing:**

Edge is more specific towards computational processes for the edge devices. So, fog includes edge computing, but would also include the network for the processed data to its final destination.

**Real Life Application Of Edge Technology:**

1. **Autonomous Vehicles –**
   GE Digital partner, Intel, estimates that autonomous cars, with hundreds of on-vehicle sensors, will generate 40 TB of data for every eight hours of driving. Therefore, wheels—edge computing plays a dominant role. Sending all the data to cloud is unsafe and impractical. The car immediately response to the events which has valuable data when coupled into digital twin and performance of other cars of its class.
2. **Fleet Management –**
   Let's example considering a trucking company, the main goal is to combine and send data from multiple operational data points like wheels, brakes, battery , etc to the cloud. Health key operational components are analysed by the cloud. Thus,essentially a fleet management solution encourages the vehicle to lower the cost.

**5 Key Benefits Of Edge Computing:**

1. Faster response time.
2. Security and Compliance.
3. Cost-effective Solution.
4. Reliable Operation With Intermittent Connectivity.

**Edge Cloud Computing Services:**

- IOT (Internet Of Things)
- Gaming
- Health Care
- Smart City
- Intelligent Transportation
- Enterprise Security

# Unit-4
# DATA HANDLING AND ANALYTICS

### What is Data?

The quantities, characters, or symbols on which operations are performed by a computer, which may be stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media

### What is Big Data?

**Big Data** is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently. Big data is also a data but with huge size.

### Types Of Big Data

Following are the types of Big Data:

1. **Structured**
2. **Unstructured**
3. **Semi-structured**

### Structured

Any data that can be stored, accessed and processed in the form of fixed format is termed as a 'structured' data. Over the period of time, talent in computer science has achieved greater success in developing techniques for working with such kind of data (where the format is well known in advance) and also deriving value out of it. However, nowadays, we are foreseeing issues when a size of such data grows to a huge extent, typical sizes are being in the rage of multiple zettabytes.

***Do you know? $10^{21}$ bytes*** equal to ***1 zettabyte*** or ***one billion terabytes*** forms ***a zettabyte***.

Looking at these figures one can easily understand why the name Big Data is given and imagine the challenges involved in its storage and processing.

***Do you know?*** Data stored in a relational database management system is one example of a **'structured'** data.

### Examples Of Structured Data

An 'Employee' table in a database is an example of Structured Data

| Employee_ID | Employee_Name | Gender | Department | |
|---|---|---|---|---|
| 2365 | Rajesh Kulkarni | Male | Finance | 650 |
| 3398 | Pratibha Joshi | Female | Admin | 650 |
| 7465 | Shushil Roy | Male | Admin | 500 |
| 7500 | Shubhojit Das | Male | Finance | 500 |
| 7699 | Priya Sane | Female | Finance | 550 |

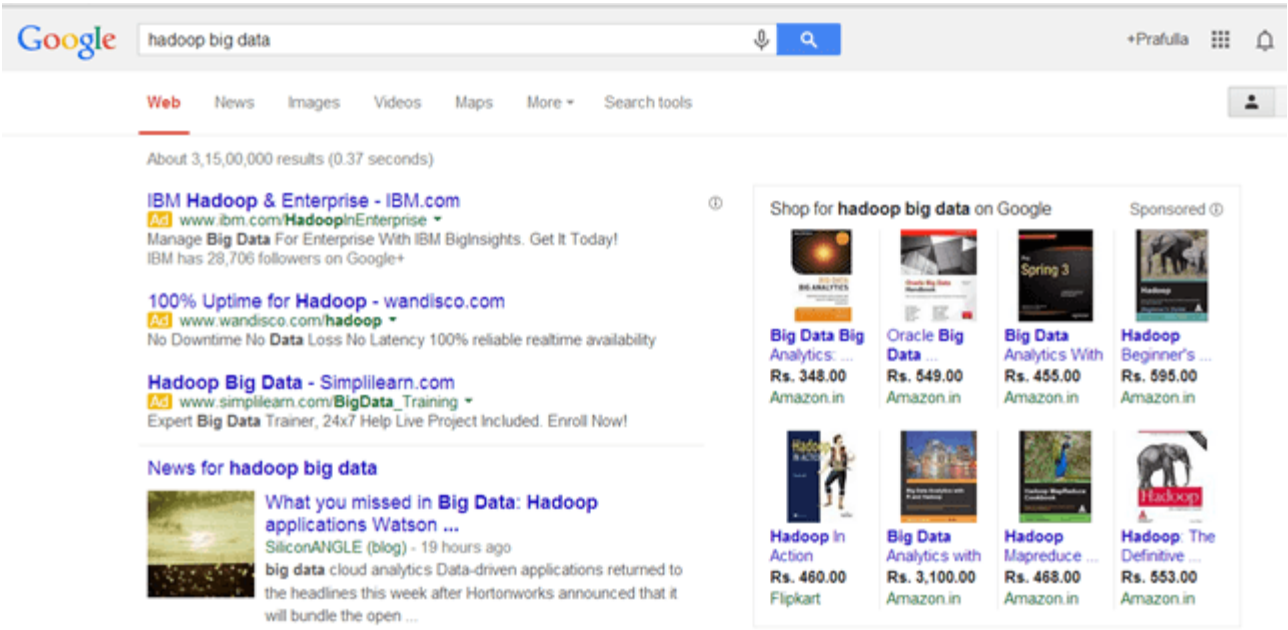
### Unstructured

Any data with unknown form or the structure is classified as unstructured data. In addition to the size being huge, un-structured data poses multiple challenges in terms of its processing for deriving

value out of it. A typical example of unstructured data is a heterogeneous data source containing a combination of simple text files, images, videos etc. Now day organizations have wealth of data available with them but unfortunately, they don't know how to derive value out of it since this data is in its raw form or unstructured format.

**Examples Of Un-structured Data**

The output returned by 'Google Search'




Example Of Un-structured Data

**Semi-structured**

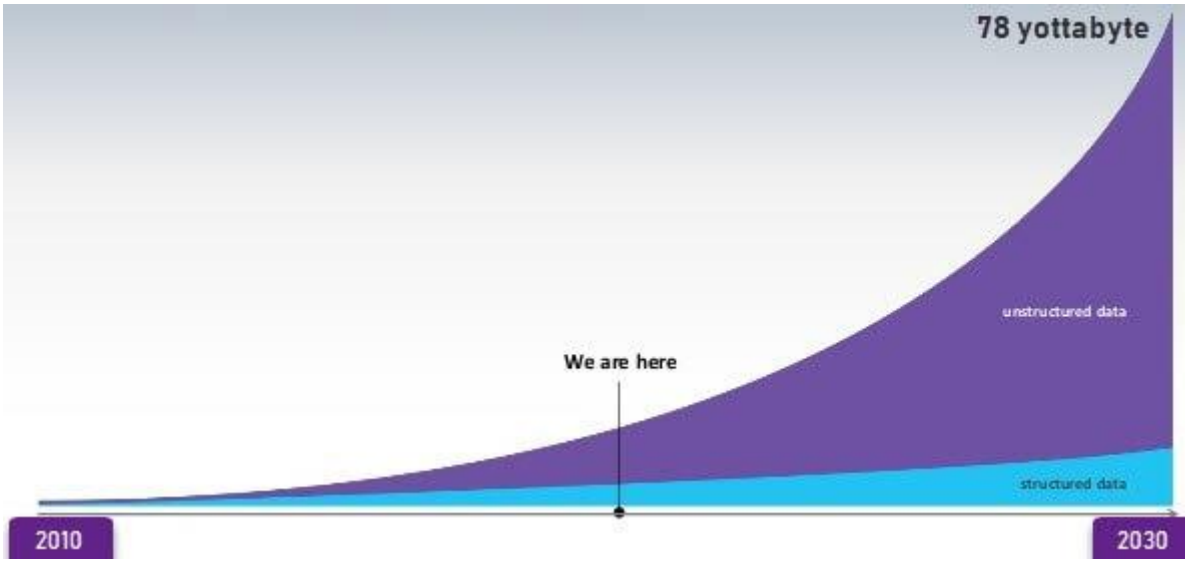
Semi-structured data can contain both the forms of data. We can see semi-structured data as a structured in form but it is actually not defined with e.g. a table definition in relational DBMS. Example of semi-structured data is a data represented in an XML file.

Examples Of Semi-structured Data

Personal data stored in an XML file-

<rec><name>Prashant Rao</name><sex>Male</sex><age>35</age></rec>
<rec><name>Seema R.</name><sex>Female</sex><age>41</age></rec>
<rec><name>Satish Mane</name><sex>Male</sex><age>29</age></rec>
<rec><name>Subrato Roy</name><sex>Male</sex><age>26</age></rec>
<rec><name>Jeremiah J.</name><sex>Male</sex><age>35</age></rec>

**Data Growth over the years**




Data Growth over the years

Please note that web application data, which is unstructured, consists of log files, transaction history files etc. OLTP systems are built to work with structured data wherein data is stored in relations (tables).

**Characteristics Of Big Data**

Big data can be described by the following characteristics:

- Volume
- Variety
- Velocity
- Variability

***(i) Volume –*** The name Big Data itself is related to a size which is enormous. Size of data plays a very crucial role in determining value out of data. Also, whether a particular data can actually be considered as a Big Data or not, is dependent upon the volume of data. Hence, **'Volume'** is one characteristic which needs to be considered while dealing with Big Data solutions.

***(ii) Variety –*** The next aspect of Big Data is its **variety**.

Variety refers to heterogeneous sources and the nature of data, both structured and unstructured. During earlier days, spreadsheets and databases were the only sources of data considered by most of the applications. Nowadays, data in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. are also being considered in the analysis applications. This variety of unstructured data poses certain issues for storage, mining and analyzing data.

***(iii) Velocity –*** The term **'velocity'** refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data.

Big Data Velocity deals with the speed at which data flows in from sources like business processes, application logs, networks, and social media sites, sensors, Mobile devices, etc. The flow of data is massive and continuous.

***(iv) Variability –*** This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the data effectively.

## Big Data Technology

Big data technology is defined as software-utility. This technology is primarily designed to analyze, process and extract information from a large data set and a huge set of extremely complex structures. This is very difficult for traditional data processing software to deal with.

Among the larger concepts of rage in technology, big data technologies are widely associated with many other technologies such as deep learning, machine learning, artificial intelligence (AI), and Internet of Things (IoT) that are massively augmented. In combination with these technologies, big data technologies are focused on analyzing and handling large amounts of real-time data and batch-related data.

## Types of Big Data Technology

Before we start with the list of big data technologies, let us first discuss this technology's board classification. Big Data technology is primarily classified into the following two types:

### Operational Big Data Technologies

This type of big data technology mainly includes the basic day-to-day data that people used to process. Typically, the operational-big data includes daily basis data such as online transactions, social media platforms, and the data from any particular organization or a firm, which is usually needed for analysis using the software based on big data technologies. The data can also be referred to as raw data used as the input for several Analytical Big Data Technologies.

Some specific examples that include the Operational Big Data Technologies can be listed as below:

- Online ticket booking system, e.g., buses, trains, flights, and movies, etc.
- Online trading or shopping from e-commerce websites like Amazon, Flipkart, Walmart, etc.
- Online data on social media sites, such as Facebook, Instagram, Whatsapp, etc.
- The employees' data or executives' particulars in multinational companies.

### Analytical Big Data Technologies

Analytical Big Data is commonly referred to as an improved version of Big Data Technologies. This type of big data technology is a bit complicated when compared with operational-big data. Analytical big data is mainly used when performance criteria are in use, and important real-time business decisions are made based on reports created by analyzing operational-real data. This means that the actual investigation of big data that is important for business decisions falls under this type of big data technology.

Some common examples that involve the Analytical Big Data Technologies can be listed as below:

- Stock marketing data
- Weather forecasting data and the time series analysis
- Medical health records where doctors can personally monitor the health status of an individual
- Carrying out the space mission databases where every information of a mission is very important

## Top Big Data Technologies

We can categorize the leading big data technologies into the following four sections:

- Data Storage
- Data Mining
- Data Analytics
- Data Visualization


## Data Storage

Let us first discuss leading Big Data Technologies that come under Data Storage:

- **Hadoop**: When it comes to handling big data, Hadoop is one of the leading technologies that come into play. This technology is based entirely on map-reduce architecture and is mainly used to process batch information. Also, it is capable enough to process tasks in batches. The Hadoop framework was mainly introduced to store and process data in a distributed data processing environment parallel to commodity hardware and a basic programming execution model.

  Apart from this, Hadoop is also best suited for storing and analyzing the data from various machines with a faster speed and low cost. That is why Hadoop is known as one of the core components of big data technologies. The **Apache Software Foundation** introduced it in Dec 2011. Hadoop is written in Java programming language.
- **MongoDB**: MongoDB is another important component of big data technologies in terms of storage. No relational properties and RDBMS properties apply to MongoDb because it is a NoSQL database. This is not the same as traditional RDBMS databases that use structured query languages. Instead, MongoDB uses schema documents.

The structure of the data storage in MongoDB is also different from traditional RDBMS databases. This enables MongoDB to hold massive amounts of data. It is based on a simple cross-platform document-oriented design. The database in MongoDB uses documents similar to JSON with the schema. This ultimately helps operational data storage options, which can be seen in most financial organizations. As a result, MongoDB is replacing traditional mainframes and offering the flexibility to handle a wide range of high-volume data-types in distributed architectures.
**MongoDB Inc.** introduced MongoDB in Feb 2009. It is written with a combination of C++, Python, JavaScript, and Go language.

- **RainStor**: RainStor is a popular database management system designed to manage and analyze organizations' Big Data requirements. It uses deduplication strategies that help manage storing and handling vast amounts of data for reference.
  RainStor was designed in 2004 by a **RainStor Software Company**. It operates just like SQL. Companies such as Barclays and Credit Suisse are using RainStor for their big data needs.
- **Hunk**: Hunk is mainly helpful when data needs to be accessed in remote Hadoop clusters using virtual indexes. This helps us to use the spunk search processing language to analyze data. Also, Hunk allows us to report and visualize vast amounts of data from Hadoop and NoSQL data sources.
  Hunk was introduced in 2013 by **Splunk Inc**. It is based on the Java programming language.
- **Cassandra**: Cassandra is one of the leading big data technologies among the list of top NoSQL databases. It is open-source, distributed and has extensive column storage options. It is freely available and provides high availability without fail. This ultimately helps in the process of handling data efficiently on large commodity groups. Cassandra's essential features include fault-tolerant mechanisms, scalability, MapReduce support, distributed nature, eventual consistency, query language property, tunable consistency, and multi-datacenter replication, etc.
  Cassandra was developed in 2008 by the **Apache Software Foundation** for the Facebook inbox search feature. It is based on the Java programming language.

### Data Mining

Let us now discuss leading Big Data Technologies that come under Data Mining:

- **Presto**: Presto is an open-source and a distributed SQL query engine developed to run interactive analytical queries against huge-sized data sources. The size of data sources can vary from gigabytes to petabytes. Presto helps in querying the data in Cassandra, Hive, relational databases and proprietary data storage systems.
  Presto is a Java-based query engine that was developed in 2013 by the **Apache Software Foundation**. Companies like Repro, Netflix, Airbnb, Facebook and Checkr are using this big data technology and making good use of it.

- **RapidMiner**: RapidMiner is defined as the data science software that offers us a very robust and powerful graphical user interface to create, deliver, manage, and maintain predictive analytics. Using RapidMiner, we can create advanced workflows and scripting support in a variety of programming languages.
  RapidMiner is a Java-based centralized solution developed in 2001 **by Ralf Klinkenberg, Ingo Mierswa, and Simon Fischer** at the Technical University of Dortmund's AI unit. It was initially named YALE (Yet Another Learning Environment). A few sets of companies that are making good use of the RapidMiner tool are Boston Consulting Group, InFocus, Domino's, Slalom, and Vivint.SmartHome.
- **ElasticSearch**: When it comes to finding information, elasticsearch is known as an essential tool. It typically combines the main components of the ELK stack (i.e., Logstash and Kibana). In simple words, ElasticSearch is a search engine based on the Lucene library and works similarly to Solr. Also, it provides a purely distributed, multi-tenant capable search engine. This search engine is completely text-based and contains schema-free JSON documents with an HTTP web interface.
  ElasticSearch is primarily written in a Java programming language and was developed in 2010 by **Shay Banon**. Now, it has been handled by Elastic NV since 2012. ElasticSearch is used by many top companies, such as LinkedIn, Netflix, Facebook, Google, Accenture, StackOverflow, etc.

## Data Analytics

Now, let us discuss leading Big Data Technologies that come under Data Analytics:

- **Apache Kafka**: Apache Kafka is a popular streaming platform. This streaming platform is primarily known for its three core capabilities: publisher, subscriber and consumer. It is referred to as a distributed streaming platform. It is also defined as a direct messaging, asynchronous messaging broker system that can ingest and perform data processing on real-time streaming data. This platform is almost similar to an enterprise messaging system or messaging queue.
  Besides, Kafka also provides a retention period, and data can be transmitted through a producer-consumer mechanism. Kafka has received many enhancements to date and includes some additional levels or properties, such as schema, Ktables, KSql, registry, etc. It is written in Java language and was developed by the **Apache software community** in 2011. Some top companies using the Apache Kafka platform include Twitter, Spotify, Netflix, Yahoo, LinkedIn etc.
- **Splunk**: Splunk is known as one of the popular software platforms for capturing, correlating, and indexing real-time streaming data in searchable repositories. Splunk can also produce graphs, alerts, summarized reports, data visualizations, and dashboards, etc., using related data. It is mainly beneficial for generating business insights and web analytics. Besides, Splunk is also used for security purposes, compliance, application management and control.

Splunk Inc. introduced **Splunk** in the year 2014. It is written in combination with AJAX, Python, C ++ and XML. Companies such as Trustwave, QRadar, and 1Labs are making good use of Splunk for their analytical and security needs.

- **KNIME**: KNIME is used to draw visual data flows, execute specific steps and analyze the obtained models, results, and interactive views. It also allows us to execute all the analysis steps altogether. It consists of an extension mechanism that can add more plugins, giving additional features and functionalities.
  KNIME is based on Eclipse and written in a Java programming language. It was developed in 2008 by **KNIME Company**. A list of companies that are making use of KNIME includes Harnham, Tyler, and Paloalto.
- **Spark**: Apache Spark is one of the core technologies in the list of big data technologies. It is one of those essential technologies which are widely used by top companies. Spark is known for offering In-memory computing capabilities that help enhance the overall speed of the operational process. It also provides a generalized execution model to support more applications. Besides, it includes top-level APIs (e.g., Java, Scala, and Python) to ease the development process.
  Also, Spark allows users to process and handle real-time streaming data using batching and windowing operations techniques. This ultimately helps to generate datasets and data frames on top of RDDs. As a result, the integral components of Spark Core are produced. Components like Spark MlLib, GraphX, and R help analyze and process machine learning and data science. Spark is written using Java, Scala, Python and R language. The **Apache Software Foundation** developed it in 2009. Companies like Amazon, ORACLE, CISCO, VerizonWireless, and Hortonworks are using this big data technology and making good use of it.
- **R-Language**: R is defined as the programming language, mainly used in statistical computing and graphics. It is a free software environment used by leading data miners, practitioners and statisticians. Language is primarily beneficial in the development of statistical-based software and data analytics.
  R-language was introduced in Feb 2000 by **R-Foundation**. It is written in Fortran. Companies like Barclays, American Express, and Bank of America use R-Language for their data analytics needs.
- **Blockchain**: Blockchain is a technology that can be used in several applications related to different industries, such as finance, supply chain, manufacturing, etc. It is primarily used in processing operations like payments and escrow. This helps in reducing the risks of fraud. Besides, it enhances the transaction's overall processing speed, increases financial privacy, and internationalize the markets. Additionally, it is also used to fulfill the needs of shared ledger, smart contract, privacy, and consensus in any Business Network Environment. Blockchain technology was first introduced in 1991 by two researchers, **Stuart Haber** and **W. Scott Stornetta**. However, blockchain has its first real-world application in Jan 2009 when Bitcoin was launched. It is a specific type of database based on Python, C++, and JavaScript.

ORACLE, Facebook, and MetLife are a few of those top companies using Blockchain technology.

### Data Visualization

Let us discuss leading Big Data Technologies that come under Data Visualization:

- **Tableau**: Tableau is one of the fastest and most powerful data visualization tools used by leading business intelligence industries. It helps in analyzing the data at a very faster speed. Tableau helps in creating the visualizations and insights in the form of dashboards and worksheets.
  Tableau is developed and maintained by a company named **TableAU**. It was introduced in May 2013. It is written using multiple languages, such as Python, C, C++, and Java. Some of the list's top companies are Cognos, QlikQ, and ORACLE Hyperion, using this tool.
- **Plotly**: As the name suggests, Plotly is best suited for plotting or creating graphs and relevant components at a faster speed in an efficient way. It consists of several rich libraries and APIs, such as MATLAB, Python, Julia, REST API, Arduino, R, Node.js, etc. This helps interactive styling graphs with Jupyter notebook and Pycharm.
  Plotly was introduced in 2012 by **Plotly company**. It is based on JavaScript. Paladins and Bitbank are some of those companies that are making good use of Plotly.

**Data Acquisition:**Data acquisition is the processes for bringing data that has been created by a source outside the organization, into the organization, for production use.

**The Data Acquisition Process**

What is exciting about data acquisition to data professionals is the richness of its process. Consider a basic set of tasks that constitute a data acquisition process:

- A need for data is identified, perhaps with use cases
- Prospecting for the required data is carried out
- Data sources are disqualified, leaving a set of qualified sources
- Vendors providing the sources are contacted and legal agreements entered into for evaluation
- Sample data sets are provided for evaluation
- Semantic analysis of the data sets is undertaken, so they are adequately understood
- The data sets are evaluated against originally established use cases
- Legal, privacy and compliance issues are understood, particularly with respect to permitted use of data
- Vendor negotiations occur to purchase the data
- Implementation specifications are drawn up, usually involving Data Operations who will be responsible for production processes
- Source onboarding occurs, such that ingestion is technically accomplished
- Production ingest is undertaken

- Big Data storage methods
- There are currently two well-established big data storage methods:
- **Warehouse Storage** – Similar to a warehouse for storing physical goods, a data warehouse is a large building facility which its primary function is to store and process data on an enterprise level. It is an important tool for big data analytics. These large data warehouses support the various reporting, business intelligence (BI), analytics, data mining, research, cyber monitoring, and other related activities. These warehouses are usually optimised to retain and process large amounts of data at all times while feeding them in and out through online servers where users can access their data without delay.

Data warehouse tools make it possible to manage data more efficiently as it enables being able to find, access, visualise and analyse data to make better business decisions and achieve more desirable business results. Additionally, they are built with the consideration of exponential data growth in mind. There is no risk of the warehouses being cluttered up by the increasing amount of data that is being stored.

The greatest benefit of data warehouses is the ability to translate raw data into information and insight. Data warehouses offer an effective way to support queries, analytics, reporting, as well as providing forecasts and trends based the collected data. Design and data cleansing must be supported by the right storage. Normally, data warehouses depend on large storage capacities that are robust, have lower costs, and perform well.

You might have heard of the term 'Hadoop' being thrown around every once in a while but still don't know what it is, which is fine. Although it is an entire topic on its own, we'll explain it briefly. Hadoop is a software framework meant for distributed storage and processing of big data to handle massive amounts of data and computation. Hadoop revolutionises big data analytics for enterprise storage. However, if you want to read more in-depth on Hadoop and its implications, read our article on it here.

**Cloud Storage** – The other method of storing massive amounts of data is cloud storage, which is something more people are familiar with. If you have ever used iCloud or Google Drive, this means you were using cloud storage to store your documents and files. With cloud storage, data and information are stored electronically online where it can be accessed from anywhere, negating the need for direct attached access to a hard drive or computer. With this approach, you can store virtually boundless amount of data online and access it where.

The cloud provides not only readily-available infrastructure, but also the ability to scale this infrastructure quickly to manage large increases in traffic or usage.

The cloud also provides easy accessibility and usability. When you want to access your data in the cloud, all you need to do is enter your credentials and you will have access. All you need is an internet connection and a device for accessing the cloud such as a mobile phone or laptop computer. Cloud storage has greatly improved productivity and efficiency of businesses as employees are able to instantaneously share, access, and edit files remotely.

**What is Hadoop?**

Hadoop is an open source software programming framework for storing a large amount of data and performing the computation. Its framework is based on Java programming with some native code in C and shell scripts.

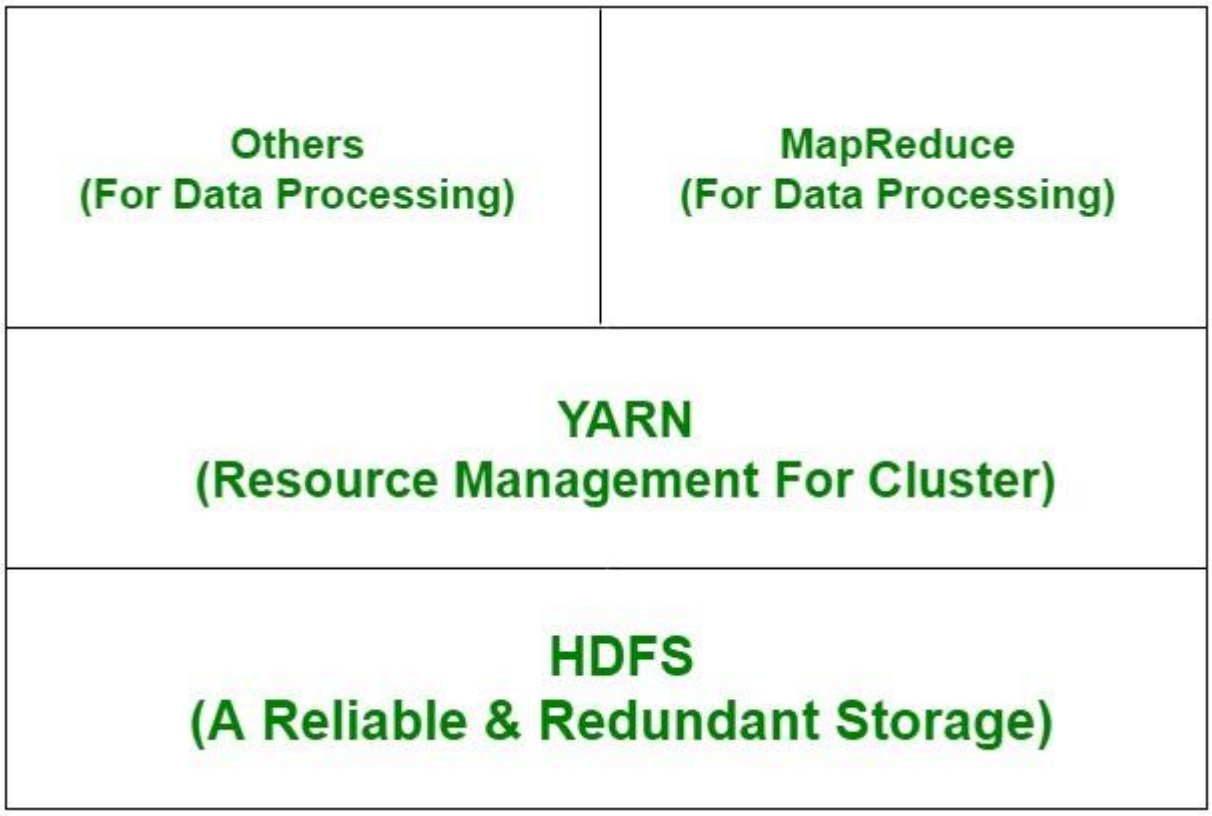
**History of Hadoop**

**Apache Software Foundation** is the developers of Hadoop, and it's co-founders are **Doug Cutting** and **Mike Cafarella**.

It's co-founder Doug Cutting named it on his son's toy elephant. In October 2003 the first paper release was Google File System. In January 2006, MapReduce development started on the Apache Nutch which consisted of around 6000 lines coding for it and around 5000 lines coding for HDFS. In April 2006 Hadoop 0.1.0 was released.

**Hadoop Distributed File System**

It has distributed file system known as HDFS and this HDFS splits files into blocks and sends them across various nodes in form of large clusters. Also in case of a node failure, the system operates and data transfer takes place between the nodes which are facilitated by HDFS.




**Advantages of HDFS:**

It is inexpensive, immutable in nature, stores data reliably, ability to tolerate faults, scalable, block structured, can process a large amount of data simultaneously and many more.

**Disadvantages of HDFS:**

It's the biggest disadvantage is that it is not fit for small quantities of data. Also, it has issues related to potential stability, restrictive and rough in nature.

Hadoop also supports a wide range of software packages such as Apache Flumes, Apache Oozie, Apache HBase, Apache Sqoop, Apache Spark, Apache Storm, Apache Pig, Apache Hive, Apache Phoenix, Cloudera Impala.

**Some common frameworks of Hadoop**

1. Hive- It uses HiveQl for data structuring and for writing complicated MapReduce in HDFS.
2. Drill- It consists of user-defined functions and is used for data exploration.
3. Storm- It allows real-time processing and streaming of data.
4. Spark- It contains a Machine Learning Library(MLlib) for providing enhanced machine learning and is widely used for data processing. It also supports Java, Python, and Scala.
5. Pig- It has Pig Latin, a SQL-Like language and performs data transformation of unstructured data.
6. Tez- It reduces the complexities of Hive and Pig and helps in the running of their codes faster.

Hadoop framework is made up of the following modules:

1. Hadoop MapReduce- a MapReduce programming model for handling and processing large data.
2. Hadoop Distributed File System- distributed files in clusters among nodes.
3. Hadoop YARN- a platform which manages computing resources.
4. Hadoop Common- it contains packages and libraries which are used for other modules.

**Advantages and Disadvantages of Hadoop**

**Advantages:**

- Ability to store a large amount of data.
- High flexibility.

- Cost effective.
- High computational power.
- Tasks are independent.
- Linear scaling.

**Disadvantages:**

- Not very effective for small data.
- Hard cluster management.
- Has stability issues.
- Security concerns.

**Hadoop Distributed File System(HDFS)**

This is where Hadoop comes in. It provides one of the most reliable filesystems. HDFS (Hadoop Distributed File System) is a unique design that provides storage for *extremely large files* with streaming data access pattern and it runs on *commodity hardware*. Let's elaborate the terms:

- ***Extremely large files***: Here we are talking about the data in range of petabytes(1000 TB).
- ***Streaming Data Access Pattern***: HDFS is designed on principle of *write-once and read-many-times*. Once data is written large portions of dataset can be processed any number times.
- ***Commodity hardware:*** Hardware that is inexpensive and easily available in the market. This is one of feature which specially distinguishes HDFS from other file system.

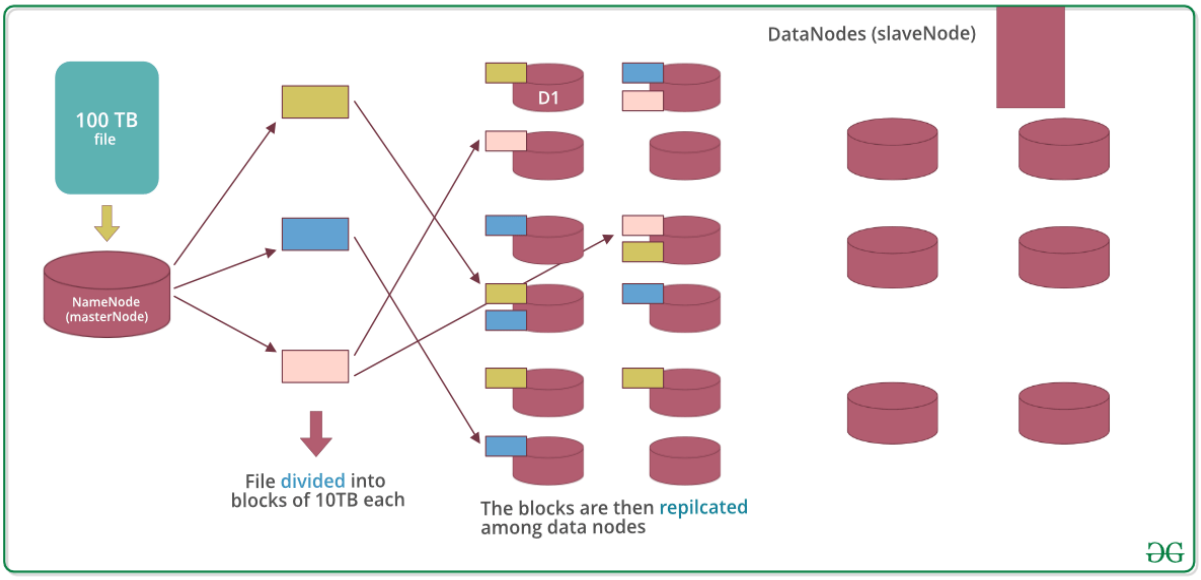
**Nodes:** Master-slave nodes typically forms the HDFS cluster.

1. **NameNode(MasterNode):**
    - Manages all the slave nodes and assign work to them.
    - It executes filesystem namespace operations like opening, closing, renaming files and directories.
    - It should be deployed on reliable hardware which has the high config. not on commodity hardware.
2. **DataNode(SlaveNode):**
    - Actual worker nodes, who do the actual work like reading, writing, processing etc.
    - They also perform creation, deletion, and replication upon instruction from the master.
    - They can be deployed on commodity hardware.

**HDFS daemons:** Daemons are the processes running in background.

- **Namenodes:**
    - Run on the master node.
    - Store metadata (data about data) like file path, the number of blocks, block Ids. etc.
    - Require high amount of RAM.
    - Store meta-data in RAM for fast retrieval i.e to reduce seek time. Though a persistent copy of it is kept on disk.
- **DataNodes:**
    - Run on slave nodes.
    - Require high memory as data is actually stored here.

**Data storage in HDFS:** Now let's see how the data is stored in a distributed manner.




Lets assume that 100TB file is inserted, then masternode(namenode) will first *divide* the file into blocks of 10TB (default size is *128 MB* in Hadoop 2.x and above). Then these blocks are stored across different datanodes(slavenode). Datanodes(slavenode)*replicate* the blocks among themselves and the information of what blocks they contain is sent to the master. Default replication factor is *3* means for each block 3 replicas are created (including itself). In hdfs.site.xml we can increase or decrease the replication factor i.e we can edit its configuration here.

**Map Reduce in Hadoop**

One of the three components of Hadoop is Map Reduce. The first component of Hadoop that is, Hadoop Distributed File System (HDFS) is responsible for storing the file. The second component that is, Map Reduce is responsible for processing the file.

Suppose there is a word file containing some text. Let us name this file as sample.txt. Note that we use Hadoop to deal with huge files but for the sake of easy explanation over here, we are taking a text file as an example. So, let's assume that this sample.txt file contains few lines as text. The content of the file is as follows:

**Job Tracker** (a master service). Job Tracker traps our request and keeps a track of it.

Record Readers. Thus we can also say that as many numbers of input splits are there, those many numbers of record readers are there.

In Hadoop terminology, each line in a text is termed as a 'record'. How record reader converts this text into (key, value) pair depends on the format of the file. In Hadoop, there are four formats of a file. These formats are Predefined Classes in Hadoop.

**Data Analytics?**

Insights about the market and customers are essential for business success. But there have always been challenges in getting those insights. In today's digital era, you need a data analytics solution that integrates the best of analytics and data management capabilities to quickly and easily access the data and analyze the information you need—when and where you need it.

## Four main types of data analytics

### 1. Predictive data analytics

Predictive analytics may be the most commonly used category of data analytics. Businesses use predictive analytics to identify trends, correlations, and causation. The category can be further broken down into **predictive modeling** and **statistical modeling**; however, it's important to know that the two go hand in hand.

For example, an advertising campaign for t-shirts on Facebook could apply predictive analytics to determine how closely conversion rate correlates with a target audience's geographic area, income bracket, and interests. From there, predictive modeling could be used to analyze the statistics for two (or more) target audiences, and provide possible revenue values for each demographic.

### 2. Prescriptive data analytics

Prescriptive analytics is where AI and big data combine to help predict outcomes and identify what actions to take. This category of analytics can be further broken down into **optimization** and **random testing**. Using advancements in ML, prescriptive analytics can help answer questions such as "What if we try this?" and "What is the best action?" You can test the correct variables and even suggest new variables that offer a higher chance of generating a positive outcome.

### 3. Diagnostic data analytics

While not as exciting as predicting the future, analyzing data from the past can serve an important purpose in guiding your business. Diagnostic data analytics is the process of examining data to understand cause and event or why something happened. Techniques such as drill down, data discovery, data mining, and correlations are often employed.

Diagnostic data analytics help answer why something occurred. Like the other categories, it too is broken down into two more specific categories: **discover and alerts** and **query and drill downs**. Query and drill downs are used to get more detail from a report. For example, a sales rep that closed significantly fewer deals one month. A drill down could show fewer workdays, due to a two-week vacation.

Discover and alerts notify of a potential issue before it occurs, for example, an alert about a lower amount of staff hours, which could result in a decrease in closed deals. You could also use diagnostic data analytics to "discover" information such as the most-qualified candidate for a new position at your company.

### 4. Descriptive data analytics

Descriptive analytics are the backbone of reporting—it's impossible to have business intelligence (BI) tools and dashboards without it. It addresses basic questions of "how many, when, where, and what."

Once again, descriptive analytics can be further separated into two categories: **ad hoc reporting** and **canned reports**. A canned report is one that has been designed previously and contains information around a given subject. An example of this is a monthly report sent by your ad agency or ad team that details performance metrics on your latest ad efforts.

Ad hoc reports, on the other hand, are designed by you and usually aren't scheduled. They are generated when there is a need to answer a specific business question. These reports are useful for obtaining more in-depth information about a specific query. An ad hoc report could focus on your corporate social media profile, examining the types of people who've liked your page and other industry pages, as well as other engagement and demographic information. Its hyperspecificity helps give a more complete picture of your social media audience. Chances are you won't need to view this type of report a second time (unless there's a major change to your audience).

### Business-driven insights and dealing with a fast-paced market

In a constantly changing business environment, it may be hard to predict your next move. That's where data analytics comes in. By quickly accessing data across teams and the enterprise, you can drive better decisions by getting deeper insights about:

- Who your customers are and how to reach them
- The market, including competitors
- What has happened in the past
- What's happening now
- What the future holds for your business

## Local Analytics

“Local Analytics has simplified the process of analyzing a new site by building a user-friendly platform with centralized information. This tool is easily saving us countless hours of additional work.”

## Cloud analytics

Cloud analytics is the process of storing and analyzing data in the cloud and using it to extract actionable business insights. Similar to on-premises data analytics, cloud analytics algorithms are applied to large data collections to identify patterns, predict future outcomes and produce other information useful to business decision makers.

However, cloud analytics is generally a more efficient alternative to on-premises analytics — which requires businesses to purchase, house and maintain expensive data centers. While on-premises analytics solutions give companies internal control over data privacy and security, they are difficult and expensive to scale. Cloud analytics, on the other hand, benefits from the scalability, service models and cost savings of cloud computing.

Businesses generate terabytes of data in the course of daily operations. Today, most of this data — sourced from websites, social media, IT devices and financial software, among other things — exists in the cloud. Cloud analytics tools and analytics software are particularly efficient for processing these huge data sets, producing insights in easily digestible formats and creating insights from data in the cloud available on demand, resulting in a better and more streamlined user experience.

Cloud analytics tools and analytics software are particularly efficient for processing these huge data sets, producing insights in easily digestible formats on demand that result in a better and more streamlined user experience.

1. Online Data Storage

Cloud Computing allows storage and access to data like files, images, audio, and videos on the cloud storage. In this age of big data, storing huge volumes of business data locally requires more and more space and escalating costs. This is where cloud storage comes into play, where businesses can store and access data using multiple devices.

The interface provided is easy to use, convenient, and has the benefits of high speed, scalability, and integrated security.

2. Backup and Recovery

Cloud service providers offer safe storage and backup facility for data and resources on the cloud. In a traditional computing system, data backup is a complex problem, and often, in case of a disaster, data can be permanently lost. But with cloud computing, data can be easily recovered with minimal damage in case of a disaster.

3. Big Data Analysis

One of the most important applications of cloud computing is its role in extensive data analysis. The extremely large volume of big data makes it impossible to store using traditional data management systems. Due to the unlimited storage capacity of the cloud, businesses can now store and analyze big data to gain valuable business insights.

4. Testing and Development

Cloud computing applications provide the easiest approach for testing and development of products. In traditional methods, such an environment would be time-consuming, expensive due to the setting up of IT resources and infrastructure, and needed manpower. However, with cloud computing, businesses get scalable and flexible cloud services, which they can use for product development, testing, and deployment.

5. Antivirus Applications

With Cloud Computing comes cloud antivirus software which is stored in the cloud from where they monitor viruses and malware in the organization’s system and fixes them. Earlier, organizations had to install antivirus software within their system and detect security threats.

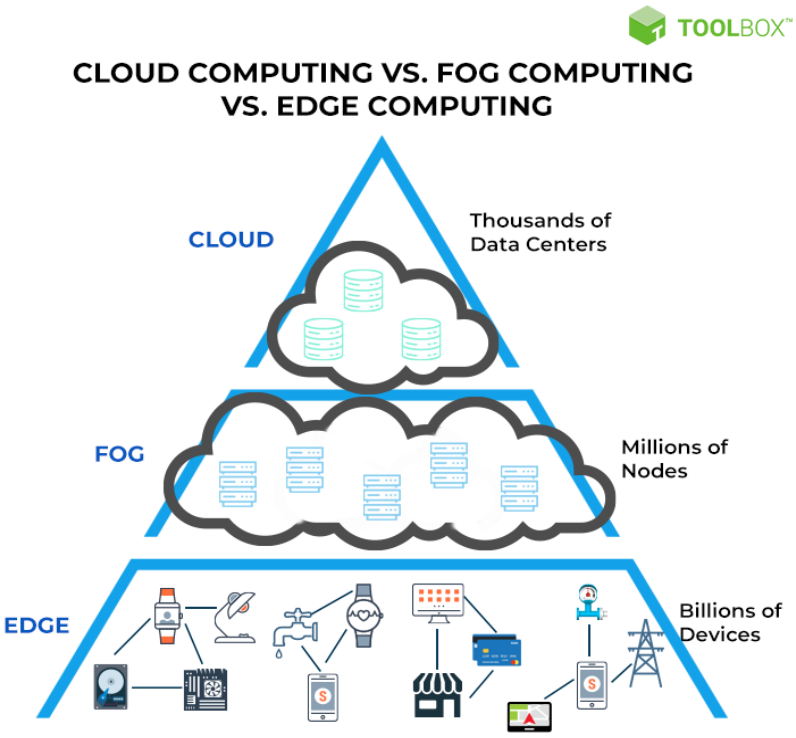
6. E-commerce Application

Ecommerce applications in the cloud enable users and e-businesses to respond quickly to emerging opportunities. It offers a new approach to business leaders to make things done with minimum amount and minimal time. They use cloud environments to manage customer data, product data, and other operational systems.

7. Cloud Computing in Education
E-learning, online distance learning programs, and student information portals are some of the key changes brought about by applications of cloud computing in the education sector. In this new learning environment, there's an attractive environment for learning, teaching, experimenting provided to students, teachers, and researchers so they can connect to the cloud of their establishment and access data and information.

**Edge vs. Fog Computing**

***Edge computing and fog computing can both be defined as technological platforms that bring computing processes closer to where data is generated and collected from. This article explains the two concepts in detail and lists the similarities and differences between them.***




**Cloud Computing vs. Fog Computing vs. Edge Computing**

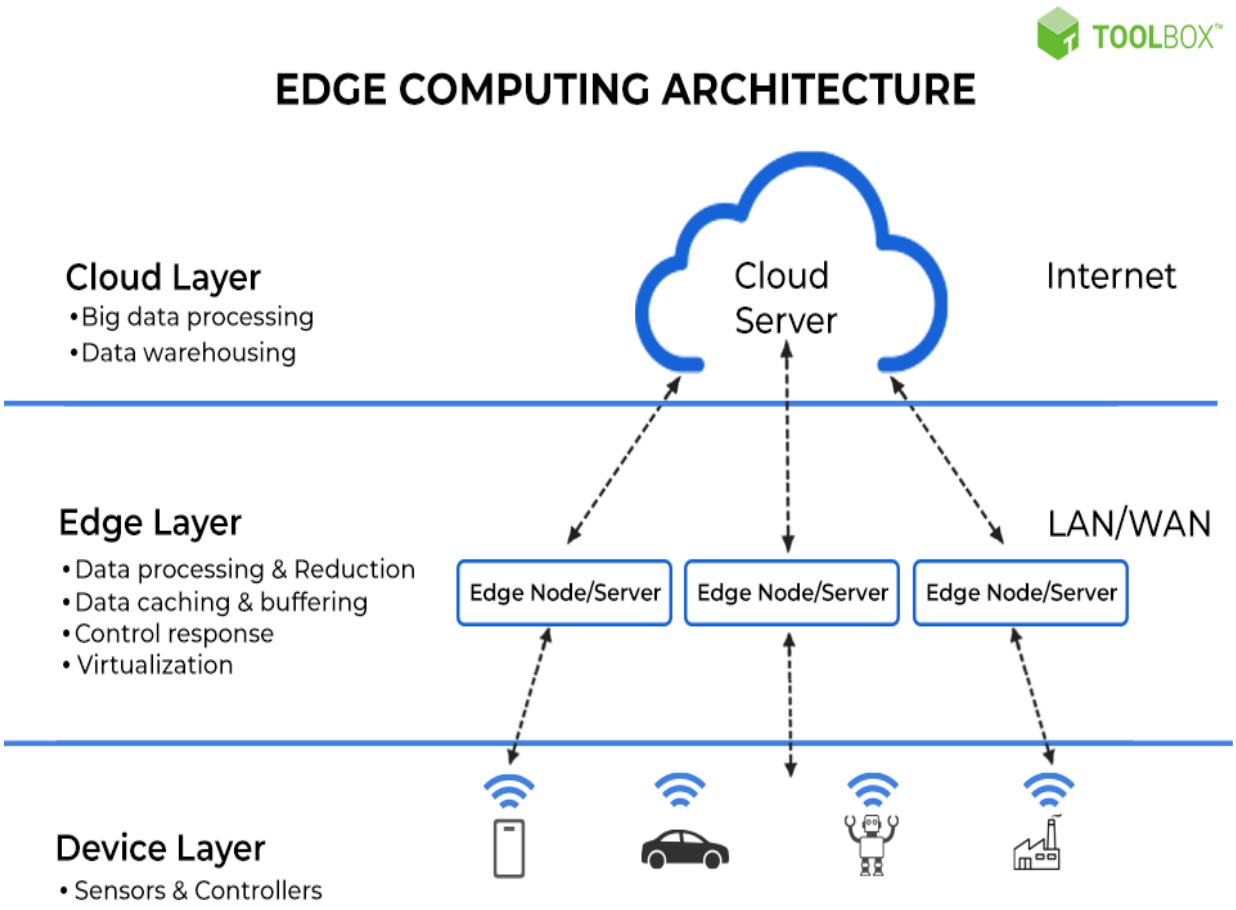
As the name implies, edge computing occurs exactly at 'the edge' of the application network. In terms of topology, this means that an 'edge computer' is right next to or even on top of the endpoints (such as controllers and sensors) connected to the network. The data is then either partially or entirely processed and sent to the cloud for further processing or storage.

However, edge computing can lead to large volumes of data being transferred directly to the cloud. This can affect system capacity, efficiency, and security. Fog computing addresses this problem by inserting a processing layer between the edge and the cloud. This way, the 'fog computer' receives the data gathered at the edge and processes it before it reaches the cloud.

Fog computing also differentiates between relevant and irrelevant data. While relevant data is sent to the cloud for storage, irrelevant data is either deleted or transmitted to the appropriate local platform. As such, edge computing and fog computing work in unison to minimize latency and maximize the efficiency associated with cloud-enabled enterprise systems.

Edge computing defined

Edge computing brings processing and storage systems as close as possible to the application, device, or component that generates and collects data. This helps minimize processing time by removing the need for transferring data to a central processing system and back to the endpoint. As a result, data is processed more efficiently, and the need for internet bandwidth is reduced. This keeps operating costs low and enables the use of applications in remote locations that have unreliable connectivity. Security is also enhanced as the need for interaction with public cloud platforms and networks is minimized. Examples of edge devices are sensors, laptops, and smartphones.




**Edge Computing Architecture**

Edge computing is useful for environments that require real-time data processing and minimal latency. This includes applications such as autonomous vehicles, the internet of things (IoT), software as a service (SaaS), rich web content delivery, voice assistants, predictive maintenance, and traffic management.

Naturally, edge computing is not a replacement for the cloud. In fact, these two technologies work with each other to add value through data. In edge networks, cloud computing is often dedicated to completing tasks that require more computing power, such as large-scale artificial intelligence (AI) and machine learning (ML) operations.

In traditional business applications, endpoints such as employees' computers are used to collect or produce data. The data is then transmitted to an enterprise application using some combination of local area networks (LAN) and wide area networks (WAN) such as the internet. Once the data is processed, the output is transmitted back to the endpoint.

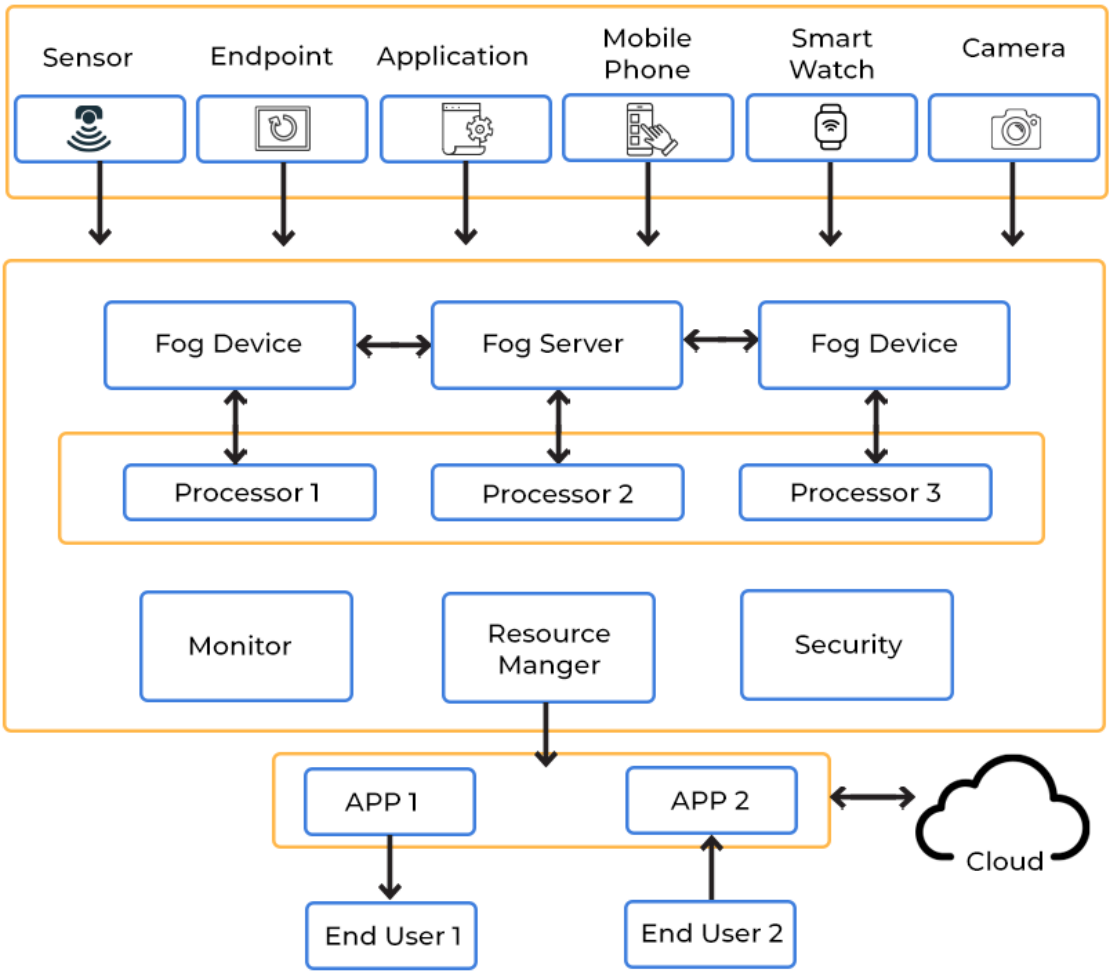
However, the number of devices connected to enterprise networks and the volume of data being generated by them are scaling at a pace that is too rapid for traditional data centers to keep up with. In fact, Gartner projects that 75% of enterprise data will be generated outside of centralized systems by 2025. Such a situation could lead to tremendous strain on both local networks and the internet at large.

## Fog computing defined

Fog computing places a decentralized enterprise computing layer between the source of data and a central cloud platform. Like edge computing, fog computing also brings the processing power closer to where the data is extracted from. While fog computing enhances efficiency, it can also be leveraged for cybersecurity and regulatory compliance. The term 'fog computing' was coined by Cisco — just like fog is formed close to the ground, fog computing takes place close to the network edge.

TOOLBOX™

# FOG COMPUTING ARCHITECTURE




Wearable smart devices such as fitness trackers are an excellent example of fog computing. Such devices rely on linked smartphones to process the data they collect and instantly show the output to the user. This removes the need for these devices to transmit data to a remote cloud platform that the manufacturer would probably need to create and maintain.

Like edge computing, fog computers are not meant to replace cloud computing. Instead, 'fogging' complements the cloud by performing less intensive analytics and processing tasks at the edge. This reduces the pressure on the cloud and allows it to focus on more long-term, resource-intensive tasks. Numerous fog computers process data in real time and create analytical summaries. This metadata is then shared with a central cloud platform, where it is analyzed to generate actionable insights.

**UNIT-5**

**Applications of IOT**

**1.Smart Home and Office:** Smart home applications with the use of smart sensors are becoming popular now. Any smart device can be configured and connected to the internet and control using simple mobile application.

**Smart Door access control system**Smart locks and door access systems are one of the most popular and cost effective solutions of Internet of Things. Smart locks are easy to implement and control using a web interface or Smartphone application.Integration with RDIF tags, smart door accessing systems can be securely implemented. Users can grant access to the doors using mobile app and lock again once the person leaves the premises.

**Smart lighting for home and office**

Smart lighting is one the attractive smart home application using internet of things. In addition to energy saving, it also enables us to manage effectively. Light ambience can be changed using smart hub devices or smart phone app.

Smart lighting can be configured to respond to voice commands and motion detectors / proximity sensors. These sensors will activate the light when someone enters the room or leaving the room. Moreover, it can be configured to turn on when the ambient light is below certain threshold (turn on during sun light is low).

**Automated Gate and garage**

Using smart sensor technology and internet of things, gates and garages can be controlled (operated) conveniently. Once you are about to enter the house or after leaving the premises, you may open or close the gate using mobile devices.

**Smart thermostats and humidity controllers**

Smart thermostats are cost effective and convenient smart home solutions which can be controlled using an internet connection and smart hub device (or using Smartphone app).

**Common sensors for home/office automation:**

- Motion / proximity sensors
- Voice controlled sensor
- Light sensor
- Temperature and humidity sensors
- Smoke/fire sensor
- Precipitation sensor

**Traffic Management**

Analyzing traffic over a period of time gives an insight of possible trends and pattern that could occur during peak hours. It will help to inform commuters to take alternative routes to avoid congestion and delay.

**Smart lighting on streets**

Smart lighting is an effective solution to save energy in the cities. Smart sensors can detect presents of people or vehicles in the proximity and increase light intensity when someone pass by.

Once the person or vehicle is away from that area, smart light will automatically reduce light intensity to save energy. During emergency situations, maximum light intensity will be activated to support recovery activities.

Since the smart lighting systems are connected to control and monitoring network, any faulty light units will be automatically reported and necessary maintenance will be initiated.

**Pollution monitoring and reporting**

Increasing air pollution is one of the challenges we are facing in every growing cities. In order to solve this issue, smart sensors are deployed across the cities to continuously monitor any changes.

Some of the common sensors are temperature, air quality (like CO2 level, haze, and smoke), moisture etc… Interconnected smart sensors collects data, sends these data to the monitoring stations and initiates warning messages during bad air quality detection.

**Smart Parking Solutions**

Smart sensors installed on parking area are collecting information about availability of parking slots and updating it to the database real time. Once the spot is occupied, it will be updated without any delay.

Service providers and customers can plan and manage parking issues with the use of smart parking solutions.

**Water / waste management**

Populations in cities are increasing every year, based on statistics this trend will grow in coming years. Increase in population contributes to increase in wastes as well.

Many cities are adapting recycling of water using water treatment units. With IoT system, the amount of waste water, consumption in a geographical area and trend of waste produced can be analyzed effectively.

IoT and smart sensor technology enables us to manage this issue efficiently. With smart waste management system, authorities will be able to predict the amount of waste produced in a particular location, how to process properly, trigger clearance of waste and analyze data for future planning etc.…

Example: smart sensors implemented on trash bins can send alerts to the waste management system once the bin is full (or reached threshold limit). If the waste quantity in the bin is low, it will not be emptied.

With analytics solutions, an overview of waste generated in every part of the city, how much wastes are generated in duration can be easily assessed. This information will be used to plan during the city expansion and upgrading projects.

Fleets for waste collection and treatment can be managed and any changing trends can be predicted via smart analytics solutions.

**Wearable Devices**

Wearable smart devices introduced as smart watches around a decade ago and many more functions were added since then. Now our smart watches and wearable are capable of reading text messages, showing notifications of other apps, tracking location, monitor workout status, remind schedules and continuously monitoring health conditions.

With Internet of Things, wearable technology can be used beyond these functions. Major smart wearable manufacturers are developing special operating systems and applications dedicated for smart wearable devices.

Wearable smart devices introduced as smart watches around a decade ago and many more functions were added since then. Now our smart watches and wearable are capable of reading text messages, showing notifications of other apps, tracking location, monitor workout status, remind schedules and continuously monitoring health conditions.

With Internet of Things, wearable technology can be used beyond these functions. Major smart wearable manufacturers are developing special operating systems and applications dedicated for smart wearable devices.

**Healthcare:** Healthcare industry has been utilizing the possibilities of Internet of Things for life saving applications. Starting from collecting vital data from bed side devices, real-time diagnosing process, accessing medical records and patient information across multiple departments, the entire system of patient care can be improved with IoT implementation.
IoT will offer convenience for medical practitioners, improve accuracy in the information (helps to reduce error in the data), increase overall efficiency and saves time for each procedures.

Doctors can monitor patient's status remotely and suggest necessary procedures when required.

Data loss and mistakes will be reduced to a lower level with IoT devices. Most of the modern medical devices can be connected to the network and data can be accessed securely (In future, all devices will have the capability to connect the network).

**Autonomous Driving**

Autonomous driving has been evolving with the use of artificial intelligence and smart sensor technology in Internet of Things. Earlier generation of autonomous vehicle (partial automation) will assists drivers to drive safely, avoid collisions and warn about the conditions of the road and vehicle.

**Agriculture and Smart farming**

There are lot of challenges in the agriculture and farming industry to produce more crops and vegetable to feed increasing human population. Internet of Things can assists farmers and researchers in this area to find more optimized and cost effective ways to increase production.

In developed countries, young generation is not attracted to conventional farming and agriculture. Lack of support staff could lead to productivity; authorities have to find alternative ways to overcome this issue.

**Smart Farming**

Internet of Things offers many solutions for convenient tracking of animals with the use of smart RFID tags. Farmers can easily record data of each animal with implementation of IoT and smart tags.

For example: movement (cow, sheep) from a particular location, age and weight of individual and vaccination details can be stored in database and easily accessed by just scanning the smart tag.

**Industrial IoT for manufacturing**

Manufacturing industry is one of the early adopters of Internet of Things which entirely changed several stages of a product development cycle. Industrial IoT will help optimize various stages of product manufacturing such as:

- Monitoring of supply chain and inventory management
- Optimization in product development
- Automate mass production processes
- Quality testing and product improvement
- Improves packaging and management
- Process optimization using data collected from huge number of sensor networks
- Cost effective solution for overall management of factories

**Logistic and fleet management**Smart logistics is a complex task since the goods must be handled with greater care and efficiency. Apart from moving from one location to another location, service providers have to make sure perfect condition is maintained during transportation.

Smart sensors capable of connecting to IoT network continuously monitoring the GPS location, temperature, humidity, shock and tilt angle of the container used for transpiration. Data collected from these sensors are processed and analyzed in a central cloud system.

Logistics team can access this information from anywhere using an internet connection. Movement of fleet can be monitored real-time and updated to customers about the progress of delivery.

Smart grid concept is an enhancement of existing power grids with sensors deployed on the transmission lines and individual customer outlets. Theses sensors helps to notify any failure, abnormality in the line, understand the nature of usage and behavior pattern over time.

These data can be used to find out areas of improvement, lossy nodes during transmission, and peak time usage statistics with the use of smart meters and sensors. Energy companies can use this information to improve existing grids and implement new changes during upgrade and thus reduce carbon emission.