IOT on Google Cloud

Adesh Sadekar, Sai Mohith, Praneet Botke, R Thinakar

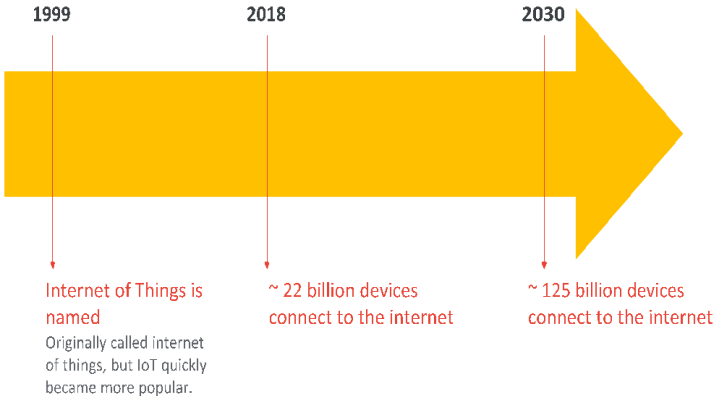
By :- Team STACK

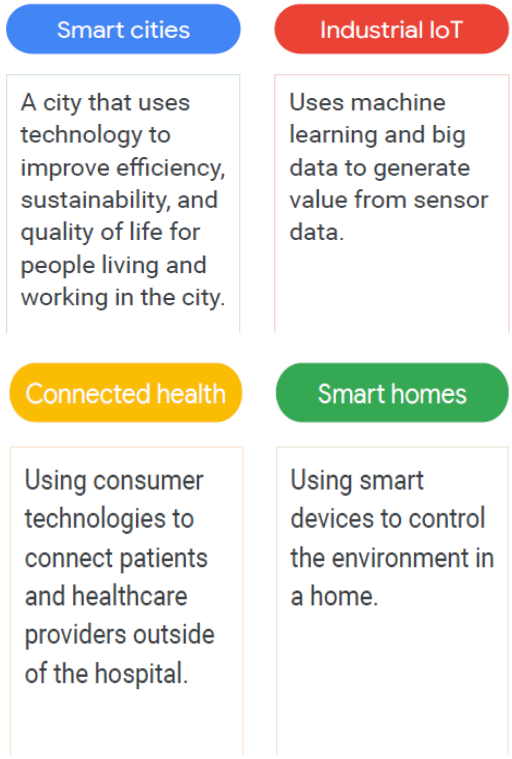
Abstract:-

The prevalence of Internet of Things (IoT) is growing rapidly with increasing influence on our daily life. IoT has received considerable attention in both academia and industry recently and there has been significant studies on security and privacy aspects of IoT and cyber-physical systems. Researchers in academia have developed novel access control models and mechanisms for IoT. On the industry side, companies including cloud services providers like Microsoft, Amazon and Google have deployed Cloud-Enabled IoT Platforms to ensure wide scale adoption. However, there is a lack of consensus between formal IoT access control models proposed in the literature and real-world Cloud-Enabled IoT deployments.

Introduction:-

The amount of attention IoT is getting is growing exponentially. In fact, the term "Internet of Things" wasn't created until 1999. Since then, the field of IoT has grown tremendously. In 2018, the number of installed IoT devices was estimated to be about 22 billion; by 2030, the number is expected to reach 125 billion.



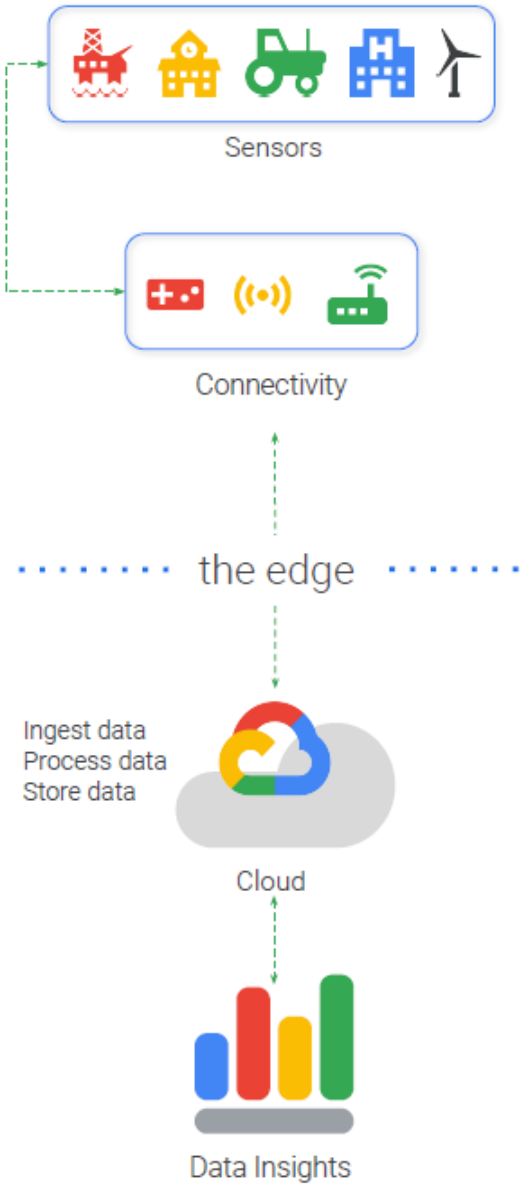


IoT architecture :-

IoT architectures must be capable of scaling connectivity of devices, data ingestion, data processing, and data storage. They must be able to do this quickly while still producing real-time data insights. Sending ever-increasing amounts of data to the cloud slows processor times and requires more bandwidth to transfer and store data.

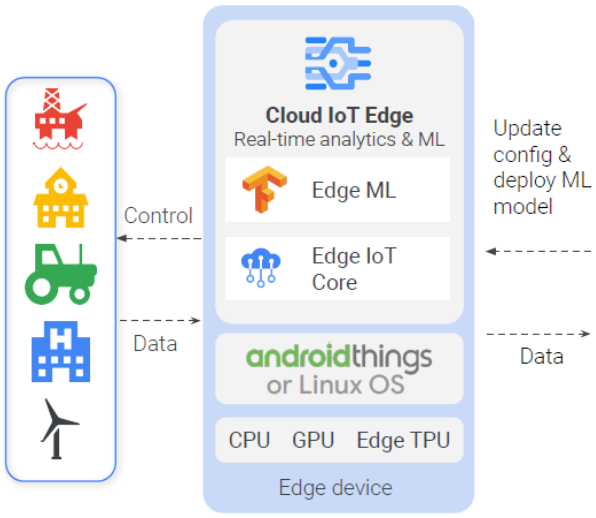
To mitigate this demand, distributed computing known as fog or edge computing is gaining popularity. The edge refers to the geographic distribution of computing nodes in the network as Internet of Things devices, which are at the "edge" of a network. This in turn increases the demand for devices that are capable of cleaning, processing, and analyzing data locally. The result is that only cleaned metadata is sent to the cloud.

Scaling also means that the ability to easily monitor and maintain thousands of devices must also scale. An asynchronous, scalable communication stack is crucial in bidirectional communication with devices. For example, what happens when thousands of devices must be updated or if they transmit data at the same time? A system that allows for asynchronous communication would be less brittle. A communication protocol that separates sending and receiving, such as Message Queuing Telemetry Transport (MQTT), discussed later in this course, is a necessity in IoT architecture. At the same time, there are cases when commands sent to a device must be accomplished immediately, resulting in a need for synchronous (or near synchronous) behavior.



Google Cloud IoT architecture :-

To create an IoT network capable of responding in near-real time, Google's Cloud IoT Architecture must be capable of doing data import, process, storage, and analysis for hundreds of millions of events per hour from devices all over the world.



To accomplish this, Google's IoT architecture can be divided into three stages: data gathering, data ingest and processing, and data analysis.

Data Gathering

Data The first stage, data gathering, occurs at the sensors and devices. Sensors gather data from the environment and send it to the cloud, either directly or through an intermediary device.

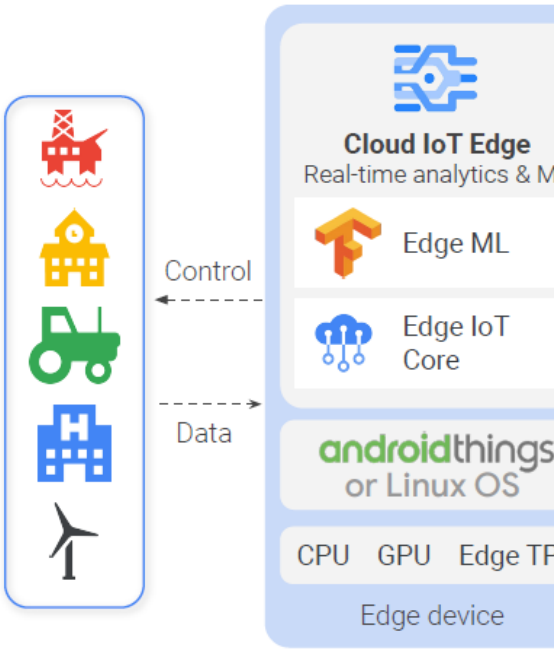
A device will prepare the data for transmission to the cloud. Depending on the network, preparation can include cleaning, preprocessing, analysis, and even machine learning inference.

In Google Cloud IoT architecture the data gathering stage can include Cloud IoT Edge. Cloud IoT Edge is a collection of devices capable of doing real-time analytics and ML. It extends Google Cloud's data processing and ML to billions of edge devices. Edge devices can act on data in real time and predict outcomes locally.

Cloud IoT Edge can run on the Android Things OS or a Linux-based OS. There are two components of Cloud IoT Edge: Edge IoT Core and Edge ML. In this course we will briefly discuss sensors, devices, and device communication. Machine learning will be covered in other courses.

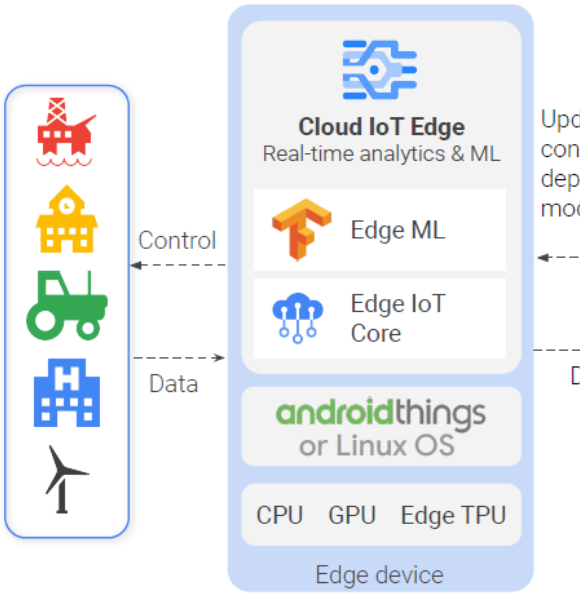
Ingest and process data

Google IoT Cloud processing data encompasses Cloud IoT Core, Cloud Functions, Pub/Sub, and Dataflow. This is a fully managed solution for ingesting and processing IoT data. Using Cloud IoT Core devices are securely connected to the cloud. Pub/Sub receives messages from devices and publishes them for subscribers to read. You use Dataflow to create data pipelines from the device to its destination, which can be BigQuery, Cloud Storage, or Bigtable. For this course you use Google templates to create your pipelines. Use Cloud Functions to create custom pipelines.



Data Analytics and ML

Data analysis and ML can be done on the Edge or on the cloud. Google's Cloud IoT Core Data Analytics and ML are fully integrated with IoT data. Often, the value of IoT analytics comes from combining data from the physical world with data from other sources; for example, online information systems or customer-relationship data. This data often accumulates in various storage systems in Cloud Storage and is accessible to BigQuery and Cloud Bigtable. Combining historical data, metadata, and real time streaming data can lead to deeper and actionable data insights.



Cloud IoT platform stages :-

Google Cloud IoT platform includes the three stages necessary for an IoT pipeline: data ingestion, data processing, and data analysis. For each of the three stages in the pipeline, several smaller tasks need to be completed. Google Cloud IoT platform has created a fully integrated tool set for each stage.

Ingest data :-

Ingest, manage and optimize your IoT device data securely.

Cloud Iot Core is a fully managed service designed to:

- Help connect, manage, and ingest data from globally dispersed devices.
- Easily and securely ingest event streams from anywhere, at any scale, for simple, reliable, real-time stream analytics.
- Seamlessly move IoT data across Google Cloud services.
- Ingest data with Cloud IoT Core and distribute data with Pub/Sub.

Cloud Iot Edge extends Google Cloud's powerful data processing and machine learning to billions of edge devices, such as robotic arms, wind turbines, and oil rigs, so they can act on the data from their sensors in real time and predict outcomes locally.

Process data :-

Process, clean and store data on-demand solutions that scale.

- ➢ You apply data transformations with Dataflow. This is simplified stream or batch data processing with equal reliability and expressiveness.
- ➢ Cloud Functions is a lightweight compute solution for developers to create single- purpose, standalone functions that respond to Cloud events without the need for you to manage a server or runtime environment.
- ➢ Cloud Storage is an online file storage service that is used for storing and accessing data on Google Cloud. It offers secure, scalable, high- performance access to your data.
- ➢ Cloud Bigtable is a sparsely populated table that can scale to billions of rows and thousands of columns, enabling you to store terabytes or even petabytes of data.
- ➢ Bigtable is ideal for storing very large amounts of single-keyed data with very low latency. It supports high read and write throughput at low latency, and it is an ideal data source for MapReduce operations.
- ➢ Cloud Spanner optimizes performance by automatically sharding your data based on request load and size of the data.

Analyze data :-

With Google Cloud IoT platform you can analyze, visualize and predict outcomes, and generate actionable insights.

- BigQuery is a highly scalable enterprise data warehouse that helps you understand your devices' behavior at scale. BigQuery is an enterprise data warehouse that stores and queries large data by enabling super-fast SQL queries.
- Vertex AI is a managed service that enables developers and data scientists to build and bring machine learning models to production. Vertex AI offers training and prediction services that can be used together or individually.
- Google Data Studio turns your data into informative dashboards and reports that are easy to read, easy to share, and fully customizable.
- Focus on generating insights faster, identifying patterns, and sharing trends across your organization and partners using Datalab, Google Data Studio, and Dataprep by Trifacta.
- Datalab is based on the open source Jupyter project. It is an interactive tool for large-scale data exploration, analysis, and visualization. IoT data can be combined with other data in Datalab to develop valuable data insights.
- Google Data Studio turns your data into informative dashboards and reports that are easy to read, easy to share, and fully customizable.

Device management is scalable with Cloud IoT :-

Devices in an IoT network must be securely connected to the network, new devices must be easily added, and all devices must be easily updated when necessary. IoT networks often contain hundreds, thousands, or even millions of devices. Managing a network means managing all the devices in the network. With Cloud IoT you can easily and securely connect, manage, and ingest IoT data from globally dispersed devices at scale. Google Cloud IoT ensures: A device has at least an ID and basic metadata before it can connect to Google Cloud. Credentials and authentication are checked before allowing a device to connect to Google Cloud. Device is authorized to publish or subscribe to a topic on Pub/Sub. A secure network connection with Google Cloud. Devices are registered and tracked when they are used. Details about the device include: Heartbeat, telemetry event received, config set, config acknowledge, and errors. Cloud IoT Core is where users create registries and devices. A Pub/Sub topic is selected when a registry is created. Authorizations and keys are associated with each device as it is added to the registry. Device management on Cloud IoT covers the three main concerns of sensor and device management: adding new devices, monitoring devices, and updating devices.

Adding new devices

Google Cloud IoT ensures:

A device has at least an ID and basic metadata before it can connect to Google Cloud. Credentials and authentication are checked before allowing a device to connect to Google Cloud.

Device is authorized to publish or subscribe to a topic on Pub/Sub

. A secure network connection with Google Cloud.

Devices are registered and tracked when they are used.

Details about the device include: Heartbeat, telemetry event received, config set, config acknowledge, and errors.

Monitoring devices

Cloud IoT Core monitors the daily operations and status of devices with Cloud Logging. The level of Cloud Logging for each device is specified when the device is created. The four levels of Cloud Logging are:

None - no logging of the device is maintained by Cloud Logging

Error - record only error messages associated with the device

Info - log errors, status, and state of the device

Debug - record debug level information for the device

Updating devices

When working with large networks, updating devices individually is not practical, or in some cases, not possible. Cloud IoT gives you the option to push updates over the air (OTA).

Cloud IoT is a complete solution :-

Cloud IoT is able to offer a complete IoT solution because:

The cloud acts as the server. This means the network can be dynamically allocated to satisfy demand.

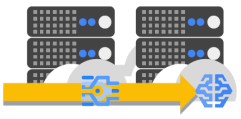
You don't own the hardware which means: you pay for only those resources that you use, you don't have to manage infrastructure (just your sensors and devices), and scaling to suit your needs is fast and easy.

It is serverless by design

Cloud IoT includes all the ML models and AI capacity available with Vertex AI. So you can train your machine learning models at scale, host them in the cloud, and deploy them to make inferences about new data.

You can also use an Edge TPU (Tensor processing unit) board and perform machine learning at the device. Edge TPUs are designed to complement Google's Cloud TPU offering, so you can accelerate ML training in the cloud, then have lightning- fast ML inference at the edge. Your sensors become more than data collectors—they make local, real-time, intelligent decisions. This means data processing and inference is done on the board. Performing ML on the edge means increased privacy, and reduced latency.

It has intelligence built-in with ML and AI capabilities



Cloud IoT Core enables secure authentication with IoT devices. The identity is unique, verifiable, trustable, and be able to validate device firmware.

Hardware root of trust is based on at least two concepts that must be available at the same time:

- the device must be verifiable, trustable, and have a unique identity
- you must be able to validate the device firmware

IoT Core achieves root-of-trust in both hardware and software by using Microchip's CryptoAuthentication device. The authorization keys generated by the device are isolated from software and the board hardware. The possibility of leaking private keys in the certification chain process is greatly reduced.

It is secure with hardware-root-of-trust