Scalability of IoT Solutions

Scalability

- Scalability will be key to handling the explosive growth in the Internet of Things (IoT).
- This means that IoT applications must have the ability to support an increasing number of connected devices, users, application features, and analytics capabilities, without any degradation in the quality of service

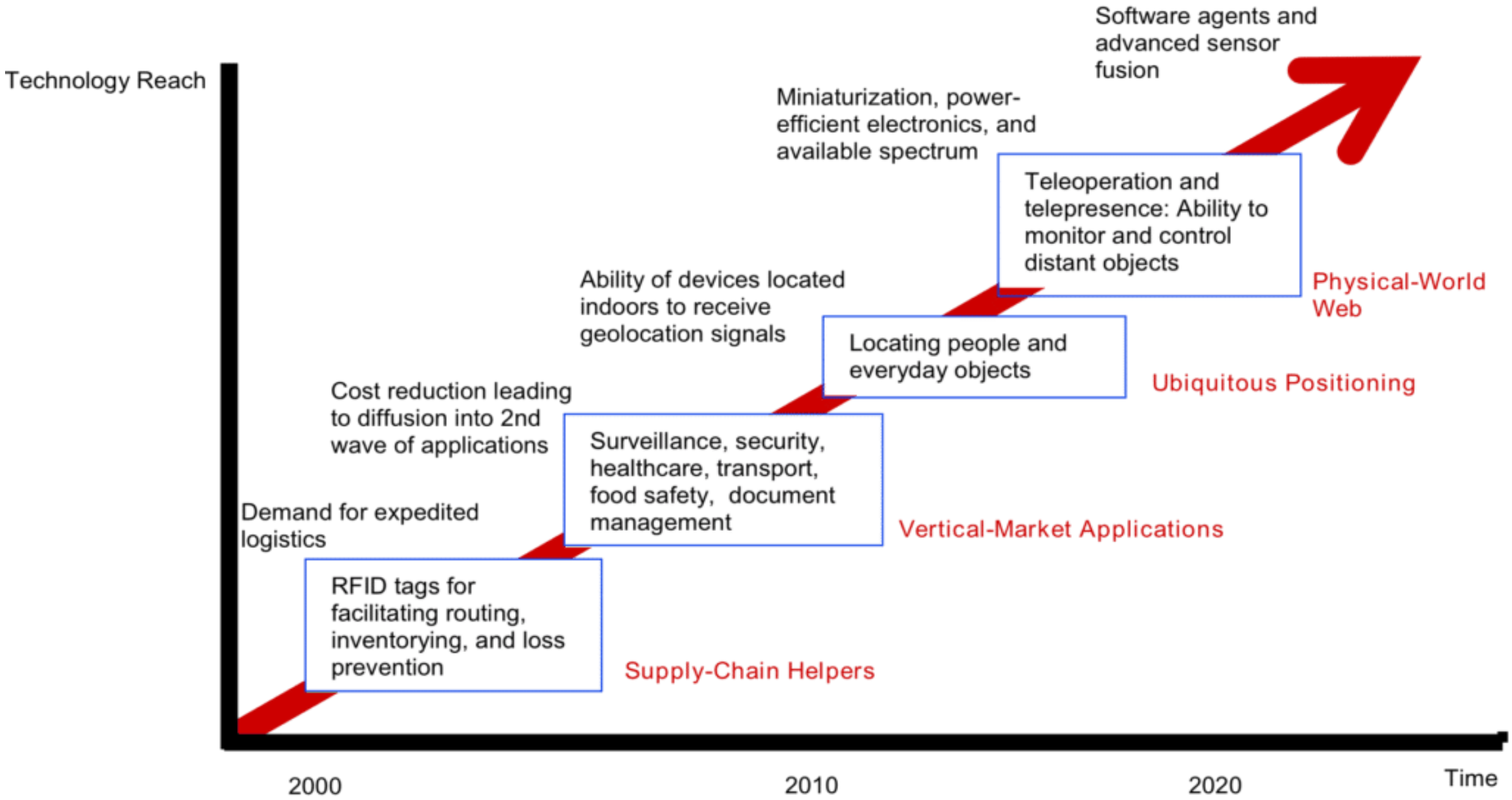
Roadmap for developing complete IoT solutions

- Scalable IoT applications are also essential to
  - monitoring,
  - securing, and
  - managing

  an increasing number of devices through a proportionate increase in the resources.

- Dear Students (All)
- Kindly join by 10 a.m today.
- It will not be ethical to make chief guest to wait for our event.
- To please us, he accepted our request though he is from completely different time zone
- In turn we need to please him by joining at 10.00 a.m
- Thanks for your understanding and support.
- This is your event just For you……

TECHNOLOGY ROADMAP: THE INTERNET OF THINGS



Source: SRI Consulting Business Intelligence

NAST
TRAINING & CONSULTANCY

Road Map for Complete IoT Solution

Infrastructure Layer

Device

Devices consist of Sensors & Sensor Technology. They collect a wide variety of data ranging from temperature, pressure, location, weather/ environment conditions, grid parameters, health essentials of a patient, etc.

Gateway

Gateways as the name suggests, are the gateways to the internet for all the devices or things that need to interact with it. They help to connect the sensor nodes in the internal network with the external Internet, by collecting data from the sensor nodes and further transmitting it to the internet infrastructure.

Application Layer

Cloud

The data transmitted through the gateway is then stored and processed securely within the Cloud in frastructure using the Big Data analytics engine. The data thus processed performs intelligent actions This is what makes the 'Smart Devices' !

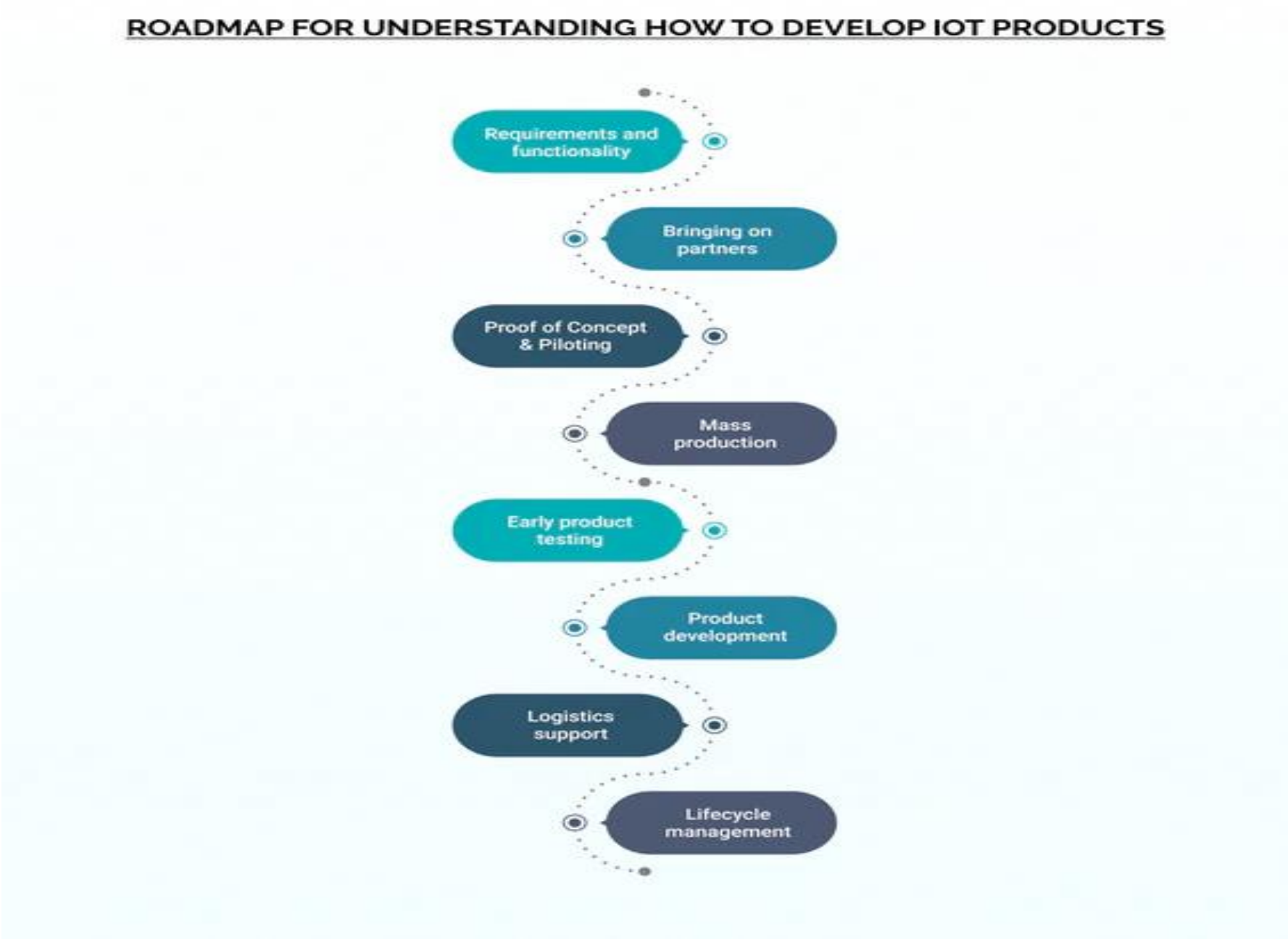
App/ Dashboard

The Applications help the end users to control and monitor their devices from remote locations. They not only send important information on the hand-held devices or PCs but also help to send commands back to the Smart Devices.

Participants :

Validators : NASSCOM

Roadmap for IoT Development



Strategies for implementation,

There are several questions that must be answered in order to ensure the design of scalable IoT applications:

Five Key Strategies for Enhancing the Scalability of IoT Applications

1. **Use automated bootstrapping**

- When the number of devices increases, any manual process with respect to device bootstrapping, registration, software configuration, security configuration, and upgrade is no longer viable.
- Any process that requires a human operator to install, configure, deploy, and manage devices is also not feasible.
- The device software and management service in the platform tier must enable automated bootstrapping, registration, monitoring, and upgrade.
- Devices must come equipped with the necessary bootloaders, security keys, and other features.

2. Control the IoT data pipeline

IoT applications require a data processing pipeline consisting of front-end data collectors and a set of data curation (cleaning, enrichment, transformation) functions applied on streaming data.

Data loggers that send the data to destinations such as databases, files, queues, and applications also play an important role.

The data pipeline needs to be well balanced in its ability to sustain a continuous flow of streaming data.

It must also be able to handle situations such as a temporary surge in the input rate, performance issues, and disruptions in downstream systems.

3. Apply the three-axis scaling approach

- IoT applications can be scaled along all three dimensions – the X, Y, and Z axes.
- Scaling along the X axis involves using more resources to divide the incoming requests between multiple servers, so that any server can handle requests. This is also called horizontal scalability.
- Scaling along the Y axis consists of dividing the workload based on actions.
- Z-axis scaling is a way of dividing responsibilities based on the incoming request data or the response data

4. Develop micro-services architecture

- In order to scale IoT applications, especially in the platform tier, it is useful to break down each application into multiple independent functional units, each of which performs one dedicated function.
- Each of these functional units should be independently deployable and executed.
- The functional units can send messages to each other. This style of architecture is called micro-services architecture

5. Adopt multiple Data Storage technologies

- A 'one size fits all' approach may not work with respect to the choice of technology, especially for IoT data storage.
- Different parts of an IoT application such as the user interfaces, interactive querying, streaming analytics, batch oriented analytics, and machine learning algorithms must be built using the best-suited technology components.
- Data query and retrieval requirements, coupled with the analytics algorithms that run on the selected data, should determine the choice of data storage or the database technology.

IoT App Development Platforms and Tools

Methods, platforms and tools.

Web and Mobile Interfaces

Open Source IoT Tools / Framework

Nodered (Free)

Contiki OS - Cooja (Free)

Zetta (Free)

OpenRemote (Free)

DeviceHive (Free Trial)

ThingSpeak (Free Trial)

Mainflux (Free Trial)

Thinger.io (Free Trial)

Arduino (Free)

IoT App Development Platforms and Tools- Cloud Based

- IBM Watson
- Microsoft Azure IoT
- Google Cloud
- Amazon Web Services (AWS)
- Cisco IoT Cloud Connect
- Oracle IoT

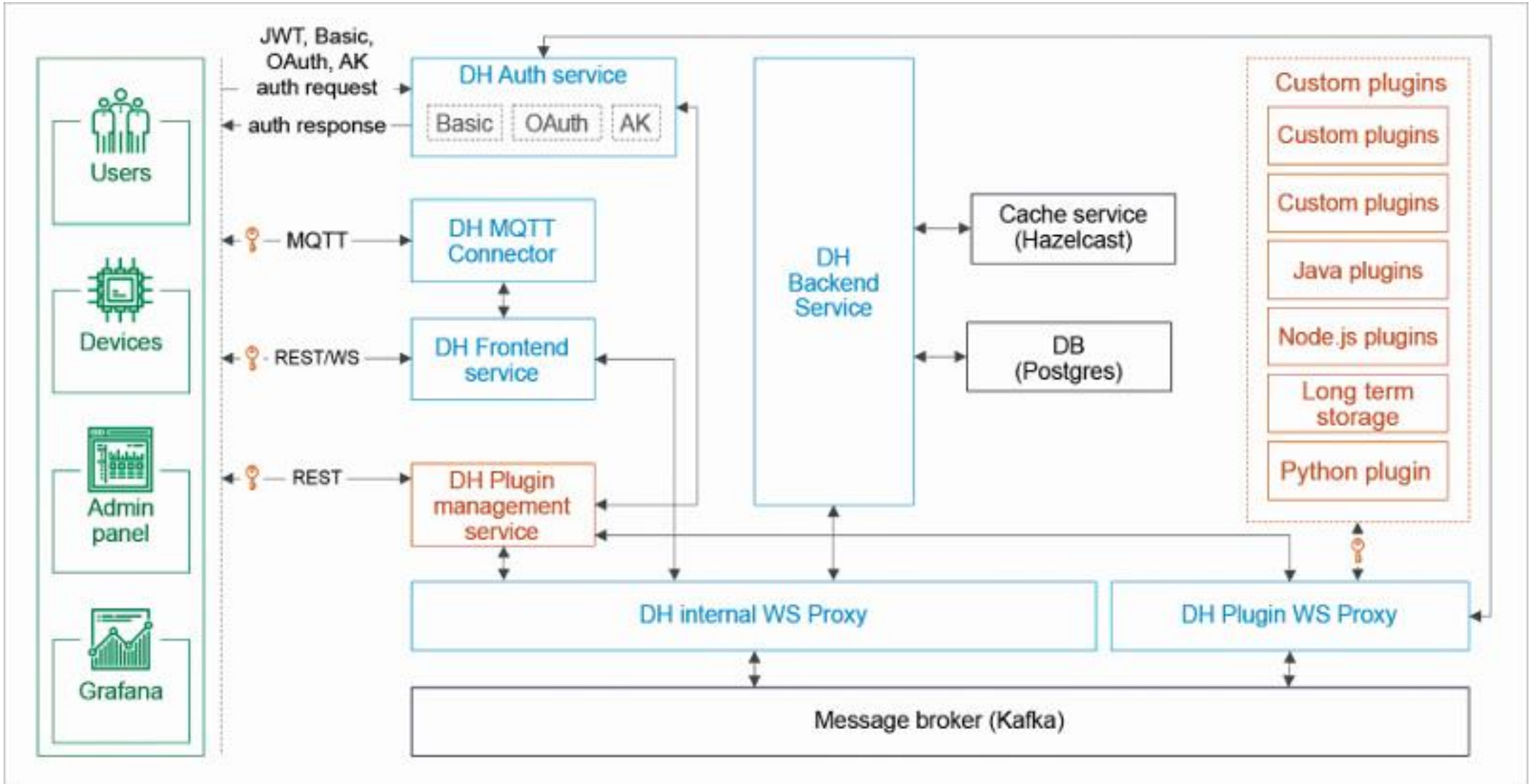
Zetta

- **Zetta is the first API-oriented open source IoT framework that basically serves for non-stop streaming loads of data.**

- https://github.com/zettajs/zetta/wiki
- https://www.zettajs.org/

DeviceHive

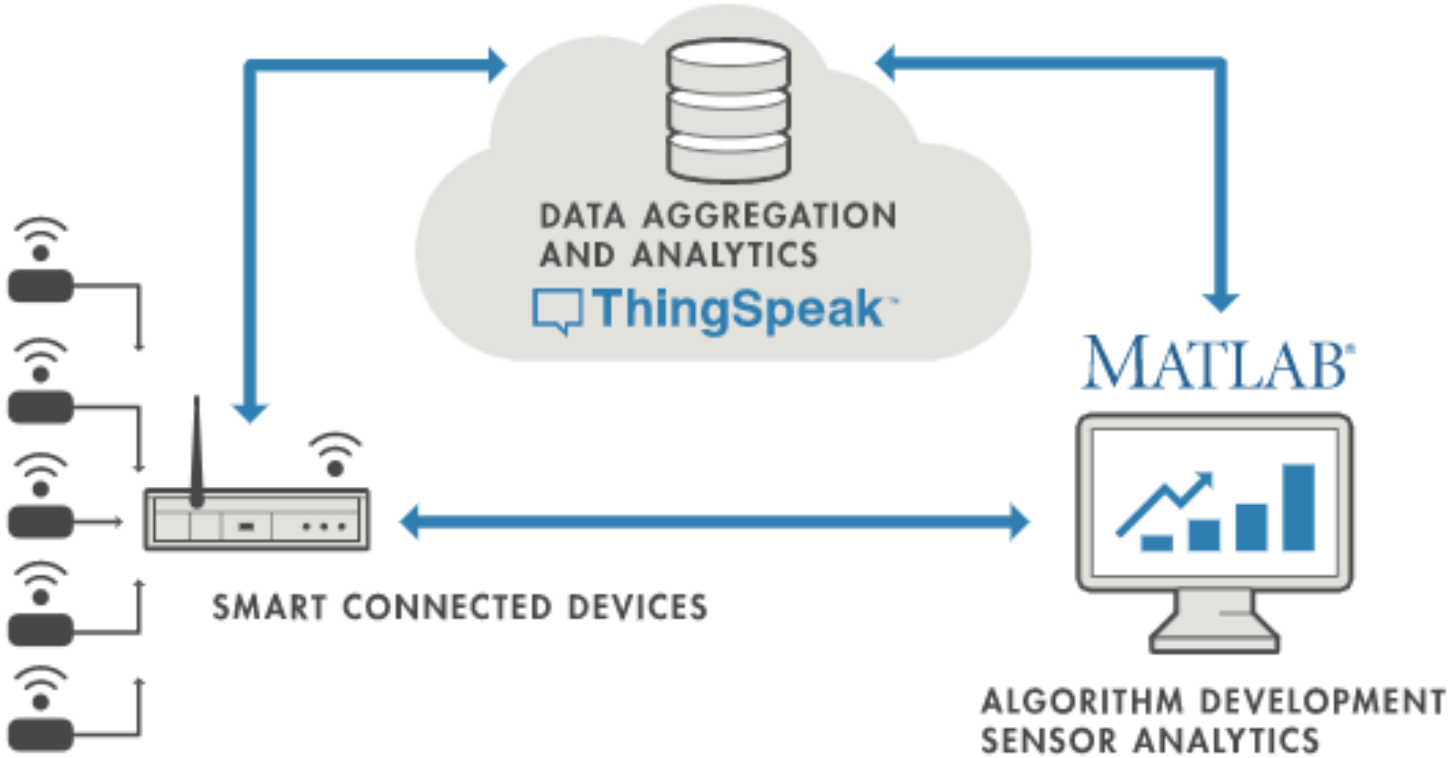
- https://devicehive.com/
- **DeviceHive** is an **open-source IoT cloud service** management platform, licensed under the Apache License Version 2.0, with a particular focus on **big data analytics**.
- DeviceHive offers robust tools to set up communication between **smart IoT devices**.
- It fills the gap between **cloud development, embedded, and mobile app development**.



https://www.byteant.com/blog/5-best-open-source-iot-frameworks/

ThingSpeak

- **ThingSpeak is a IoT platform that tightly collaborates with MathWorks**



https://www.mathworks.com/help/thingspeak/

Mainflux

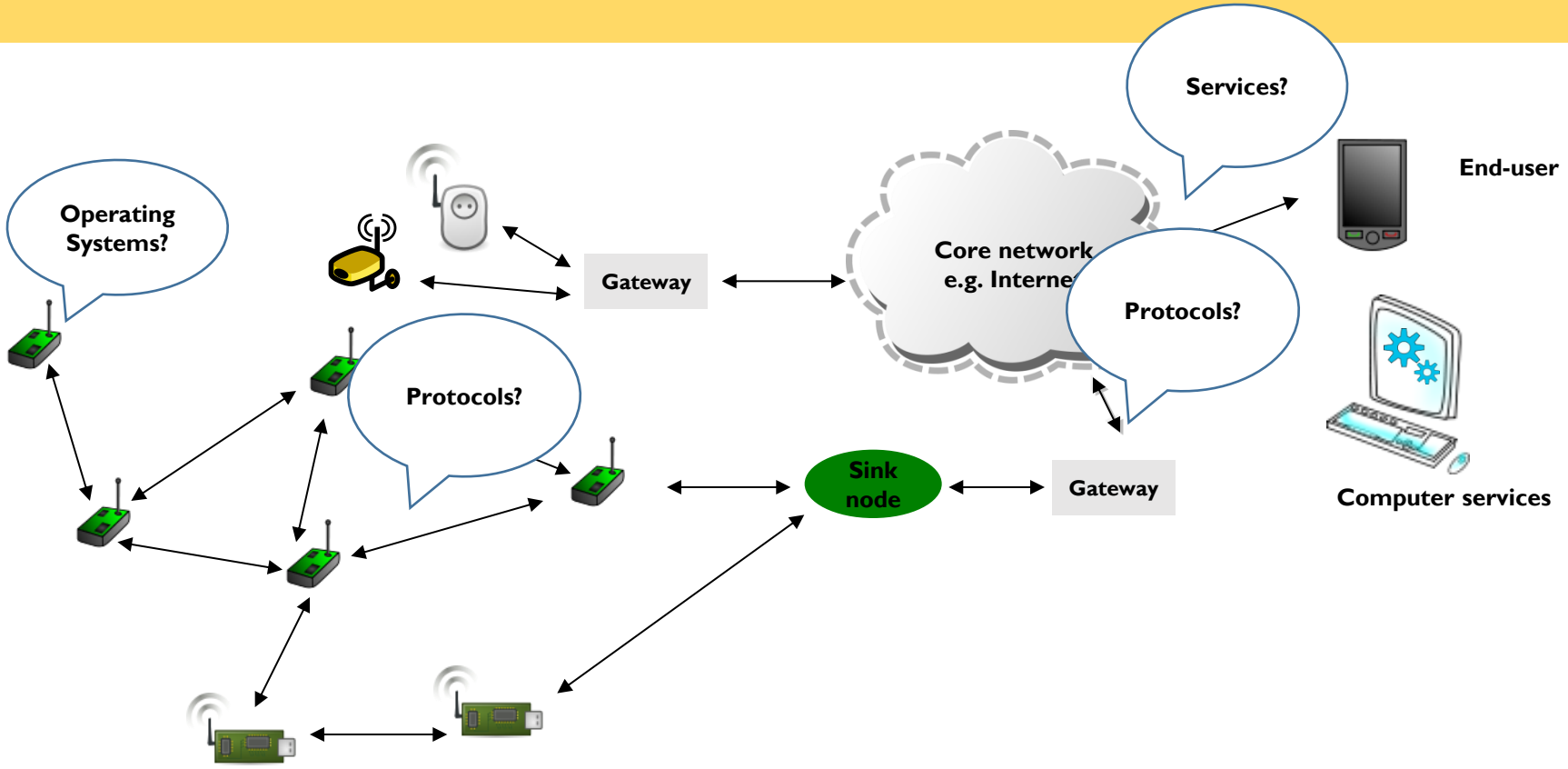
- **Data collection and management, core analytics, and event scheduling.**

https://mainflux.readthedocs.io/en/latest/

Arduino

- Offering an appropriate blend of **IoT hardware and software**, Arduino is a **simple-to-use IoT platform**.
- The software of Arduino comes in the plan of the **Arduino programming language and Integrated Development Environment (IDE).**

Wireless Sensor (and Actuator) Networks



- The networks typically run Low Power Devices
- Consist of one or more sensors, could be different type of sensors (or actuators)

Contiki

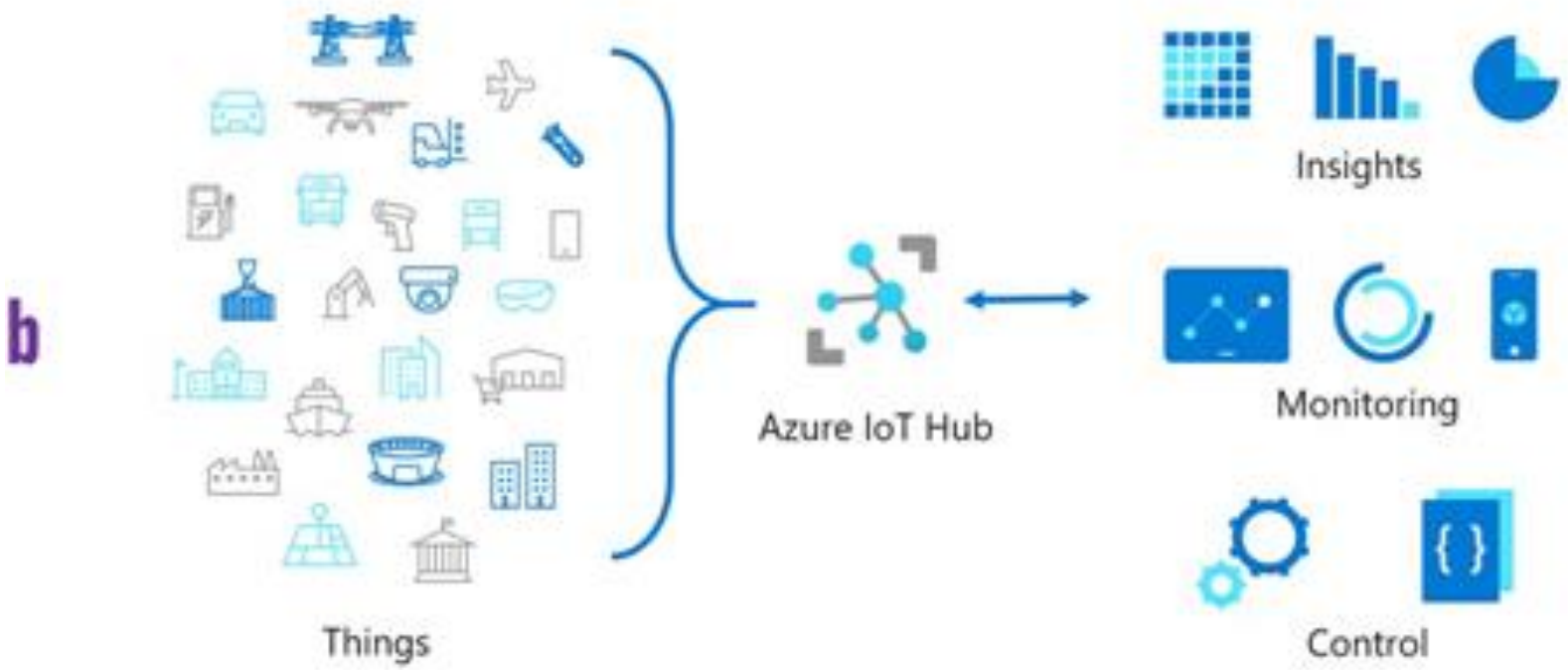
- Contiki is an **open source** operating system for **the Internet of Things**.
  - runs on **networked embedded systems and wireless sensor nodes**.
- It is designed for **microcontrollers** with small amounts of memory.
- A typical Contiki configuration **is 2 kilobytes of RAM and 40 kilobytes of ROM.**
- Contiki provides IP communication, both for IPv4 and IPv6.
  - It has an IPv6 stack
  - Contiki supports 6lowPAN header compression and the CoAP application layer protocol.

Source: http://www.contiki-os.org

Some Azure IoT Services and Methods

- Azure-IoT Hub
- Azure-Event Hub
- Azure Stream Analytics
- Azure- data lake
- Azure-Synapse (Data Ware House)
- Azure-Data Factory

Azure-IoT Hub



IoT Hub is a managed service, hosted in the cloud, that acts as a central message hub for bi-directional communication between your IoT solutions and the devices it manages.

Azure-Event Hub

b

Event Hubs is a fully managed, real-time data ingestion service that's simple, trusted, and scalable. Stream millions of events per second from any source to build dynamic data pipelines and immediately respond to business challenges.

Note - " Additionally, IoT Hub uses Event Hubs for its telemetry flow path, so your IoT solution also benefits from the tremendous power of Event Hubs."

Azure-Event Hub

- Event Hubs provides a distributed stream processing platform with low latency and seamless integration, with data and analytics services inside and outside Azure to build your complete big data pipeline.

Scenarios used by Event Hubs

- Anomaly detection (fraud/outliers)
- Application logging
- Analytics pipelines, such as clickstreams
- Live dashboarding
- Archiving data
- Transaction processing
- User telemetry processing
- Device telemetry streaming

Features

Ingest millions of events per second

Continuously ingress data from hundreds of thousands of sources with low latency and configurable time retention.

Enable real-time and micro-batch processing concurrently

Seamlessly send data to Blob storage or Data Lake Storage for long-term retention or micro-batch processing with Event Hubs capture.

Get a managed service with elastic scale

Easily scale from streaming megabytes of data to terabytes while keeping control over when and how much to scale.

Easily connect with the Apache Kafka ecosystem

Seamlessly connect Event Hubs with your Kafka applications and clients with Azure Event Hubs for Apache Kafka®.

Build a server less streaming solution

Natively connect with Stream Analytics to build an end-to-end server less streaming solution.

Difference between IoT Hub and Event Hub

- Both Azure IoT Hub and Azure Event Hubs are **cloud services** that can ingest large amounts of data and process or store that data for business insights.
- IoT Hub was developed to address the unique requirements of connecting IoT devices to the Azure cloud while Event Hubs was designed for big data streaming**. Microsoft recommends using Azure IoT Hub to connect IoT devices to Azure**.
- Azure Event Hubs is the **big data streaming** service of Azure. It is designed for high throughput data streaming scenarios where customers may send billions of requests per day.

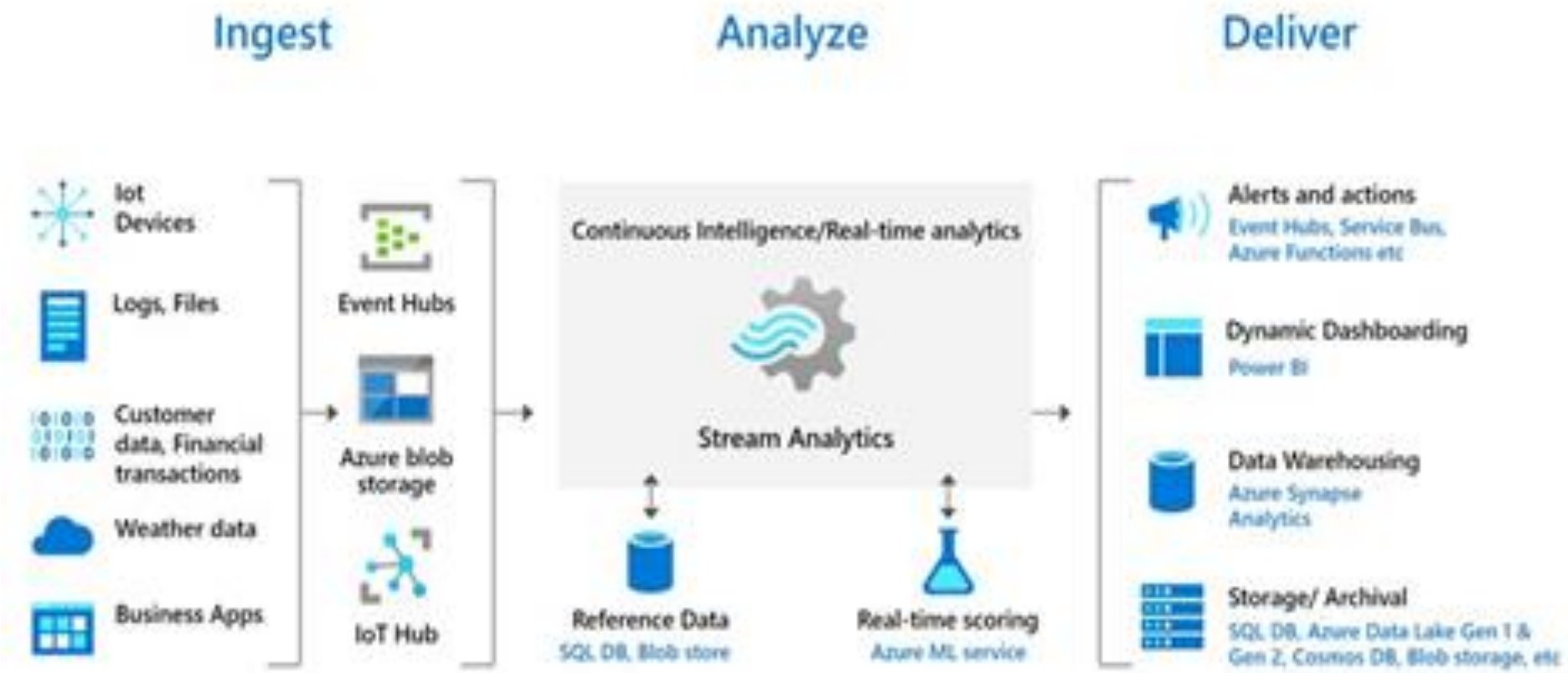
Azure Stream Analytics

- A real-time analytics and **complex event-processing engine** that is designed to analyze and process high volumes of fast streaming data from multiple sources simultaneously.
- Patterns and relationships can be identified in information extracted from a number of input sources including devices, sensors, clickstreams, social media feeds, and applications.
- These patterns can be used to trigger actions and initiate workflows such as **creating alerts, feeding information to a reporting tool, or storing transformed data for later** use
- Stream Analytics is available on Azure IoT Edge runtime, enabling to process data on IoT devices.

Examples on Azure Stream Analytics

- Analyze real-time telemetry streams from IoT devices.
- Web logs/clickstream analytics.
- Geospatial analytics for fleet management and driverless vehicles.
- Remote monitoring and predictive maintenance of high value assets.
- Real-time analytics on Point of Sale data for inventory control and anomaly detection.

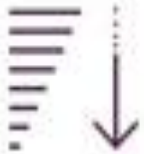
Azure-stream Analytics



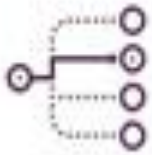
Working principle

- Azure Stream Analytics job consists of an input, query, and an output.
- Stream Analytics ingests data from Azure Event Hubs (including Azure Event Hubs from Apache Kafka), Azure IoT Hub, or Azure Blob Storage.
- Send data to services such as Azure Functions, Service Bus Topics or Queues to trigger communications or custom workflows downstream.
- Send data to a Power BI dashboard for real-time dashboarding.
- Store data in other Azure storage services (e.g. Azure Data Lake, Azure Synapse Analytics, etc.) to train a machine learning model based on historical data or perform batch analytics.

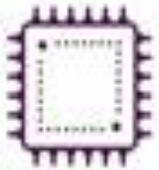
Contd..



End-to-end analytics pipeline that is production-ready in minutes with familiar SQL syntax and extensible with JavaScript and C# custom code.



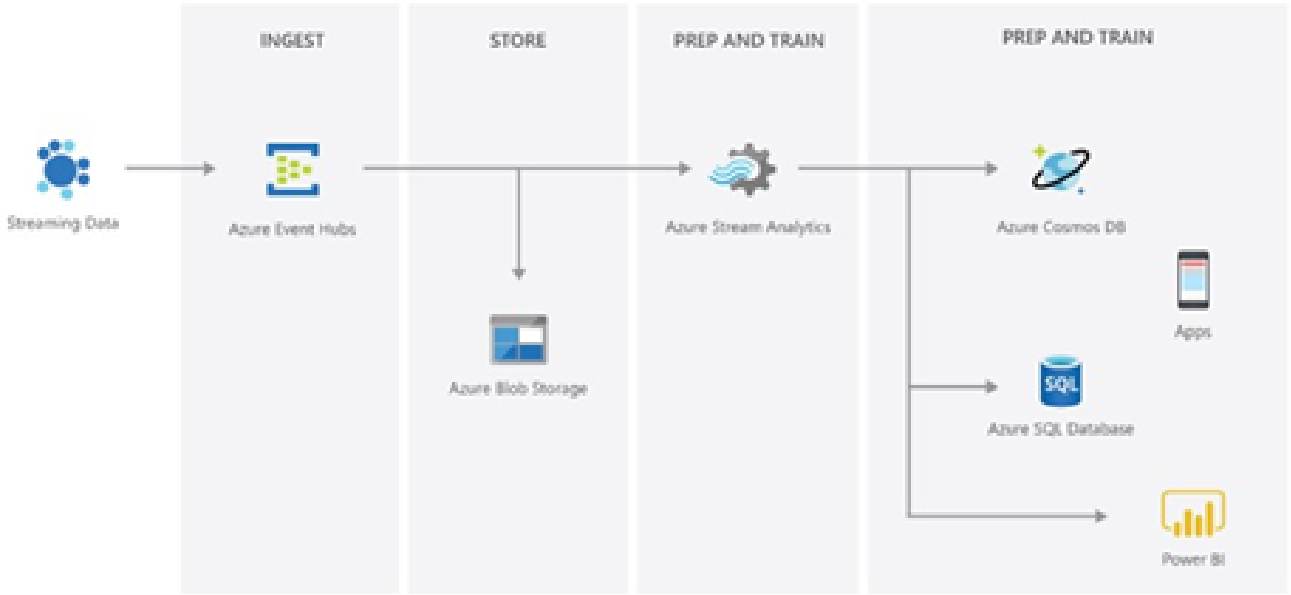
Rapid scalability with elastic capacity to build robust streaming data pipelines and analyze millions of events at sub second latencies.



Hybrid architectures for stream processing with the ability to run the same queries in the cloud and on the edge.

Enterprise-grade reliability with built-in recovery and built-in machine learning capabilities for advanced scenarios.

Azure-Data Lake



Azure Data Lake includes all the capabilities required to make it easy for developers, data scientists, and analysts to store data of any size, shape, and speed, and do all types of processing and analytics across platforms and languages

Contd.,

Massive scalability
Near limitless storage for analytics data.

Cloud object store pricing
Same low-cost data storage model as Azure Blob Storage

Fewer file and folder transactions
Atomic transactions for fewer compute cycles and faster job Execution.

Granular file and folder security
POSIX-compliant, fine-grained access control lists (ACLs)

Simplified ingestion in a single store
Consolidated data storage using the Data Lake Storage Gen2 or Blob Storage REST API.

Full Azure Blob Storage feature set
Data lifecycle policy management; hot, cool, and archive tiers; and high availability/disaster recovery support.

Role-based access and storage account firewalls
Multi-layer security to govern data access so only users from authorized IPs can perform analytics.

Common data model (CDM) support
Ability to exchange data with powerful applications like Microsoft Dynamics 365 (for CRM) and Power BI.

Azure Data Factory

- Azure Data Factory is the platform that solves data oreineted scenarios. It is the *cloud-based ETL and data integration service that allows you to create data-driven workflows for orchestrating data movement and transforming data at scale*.
- Using Azure Data Factory, you can create and schedule data-driven workflows (called pipelines) that can ingest data from disparate data stores.
- You can build complex ETL processes that transform data visually with data flows or by using compute services such as Azure HDInsight Hadoop, Azure Databricks, and Azure SQL Database.
- Azure Data Factory is a managed cloud service that's built for these complex hybrid **extract-transform-load (ETL), extract-load-transform (ELT), and data integration projects.**

Reference

- https://www.tcs.com/content/dam/tcs/pdf/discover-tcs/Research-and-Innovation/Build_a_Scalable_Platform_pdf.pdf