Continuous Computing Plane

From Cloud to Edge

Remo Marconzini, 883256

Niccolò Puccinelli, 881395

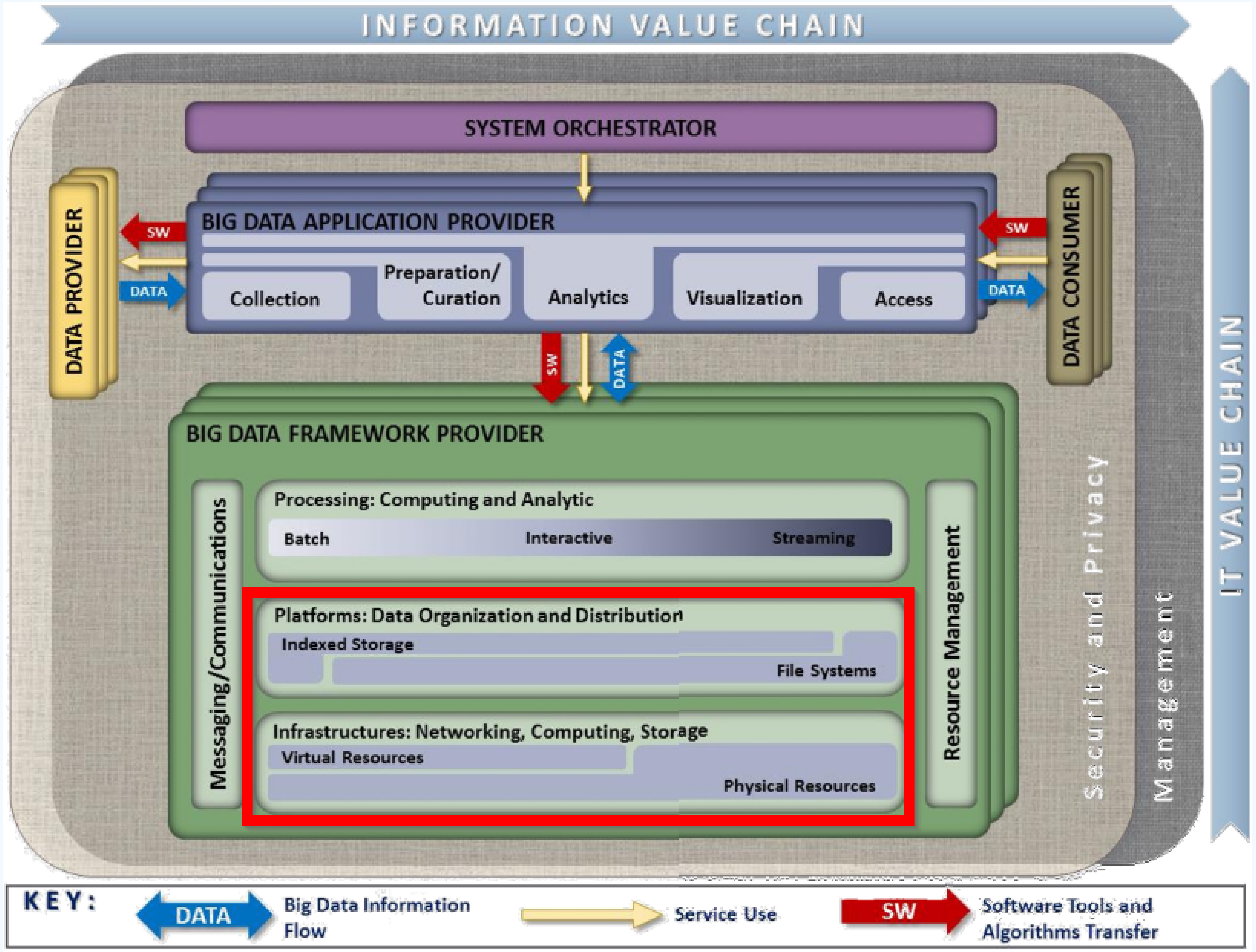
UNIVERSITA' DEGLI STUDI DI MILANO BICOCCA

Outline

- Cloud Computing
- Introduction to Edge Computing
- Key concepts and architecture
- Fog Computing
- Orchestration and the case of Cloudify
- Federated Learning
- About Data Science
- Future prospects

CLOUD COMPUTING

Cloud computing is a model for enabling **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 (NIST, 2011).



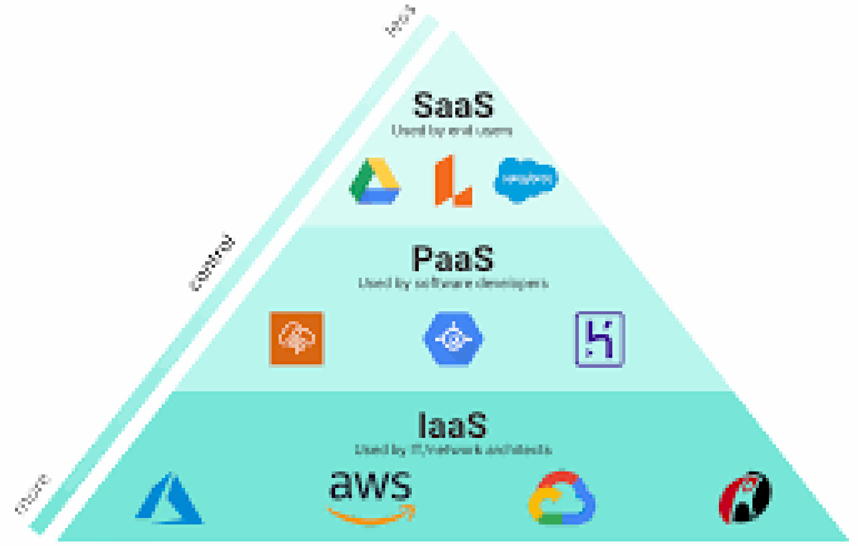
NIST
National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

NIST REFERENCE

This cloud model promotes **availability** and is composed of:

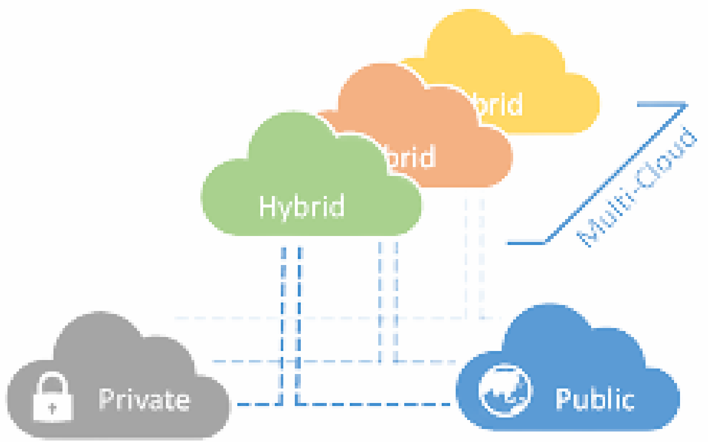
5 essential characteristics:

- High scalability and elasticity
- High availability and reliability
- High manageability and interoperability
- High accessibility and portability
- High performance and optimization



3 Service Models:

- Software as-a-Service (SaaS)
- Platform as-a-Service (PaaS)
- Infrastructure as-a-Service (IaaS)

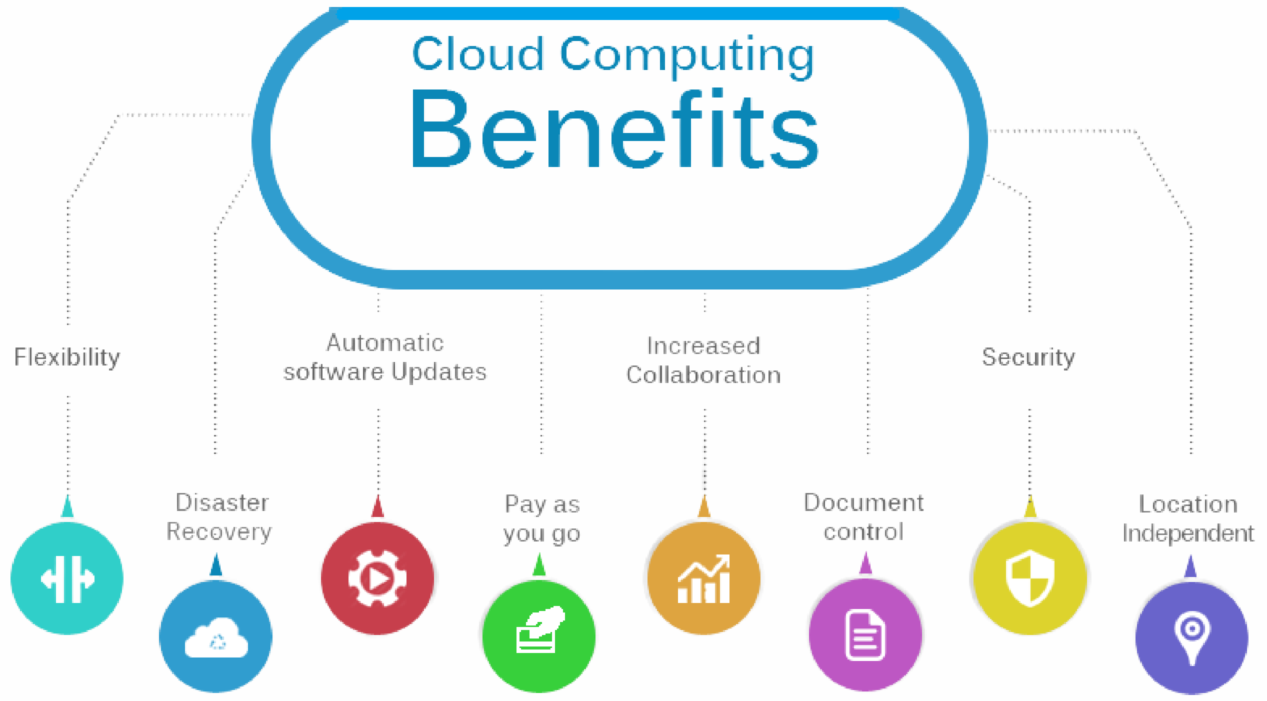


4 Deployment Models:

- Public Cloud
- Private Cloud
- Hybrid Cloud
- Community Cloud

"GOOD AND BAD"
WITH CURRENT CLOUD COMPUTING MODEL

- Cloud is designed to be: **highly realiable**, **highly scalable** and it provides **elasticity** for computing, storage and network resources.
- Cloud provides **business benefits**, such as an on-demand **pay-as-you-go** service to the users, which lowers the owning cost for the companies (from **Capex** to **Opex**).
- Cloud architectures address the difficulties related to the **large-scale data processing**.

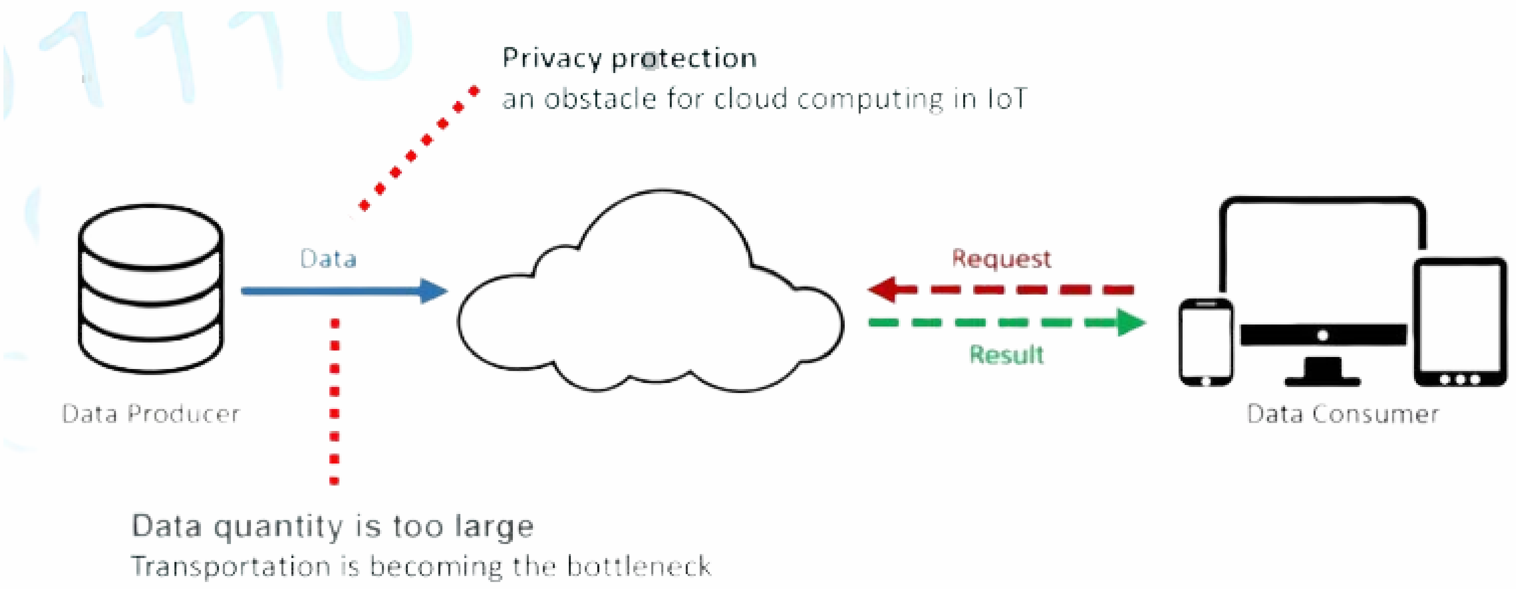


We can simply deliver all the computing services such as networking, database and analytics over the internet...

"GOOD AND BAD"
WITH CURRENT CLOUD COMPUTING MODEL

... however, such a **centralized model** will face significant challenges for the IoT world.
The Internet is **evolving rapidly** towards the future Internet of Things (IoT), which will potentially network billions or even trillions of devices.
Due to this, the current cloud computing model will have to deal with:

- **Volume** and **Velocity** of data accumulation of IoT devices.
- **Security** and **Privacy** concerns.
- **Latency** due to the distance between edge IoT devices and Datacenters.
- **Scalability** challenges.
- The **costs** of storing, processing and analysing data from IoT devices.



WHY IS THE CURRENT MODEL NOT ENOUGH?

- Applications that require **Autonomy**.

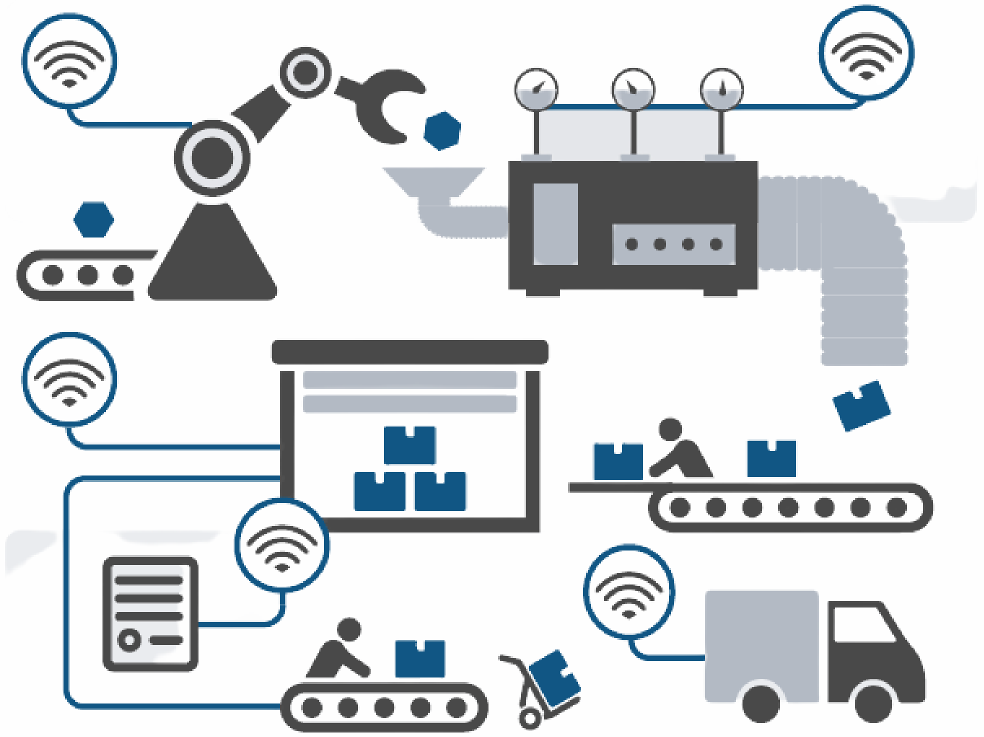
  Completing tasks with little or no human interaction.
  - Industry 4.0.
  - Self-driving car.

- Applications that can't tolerate **Latency**.

  Latency can be reason for failure.
  - Health Care.
  - Financial transactions.

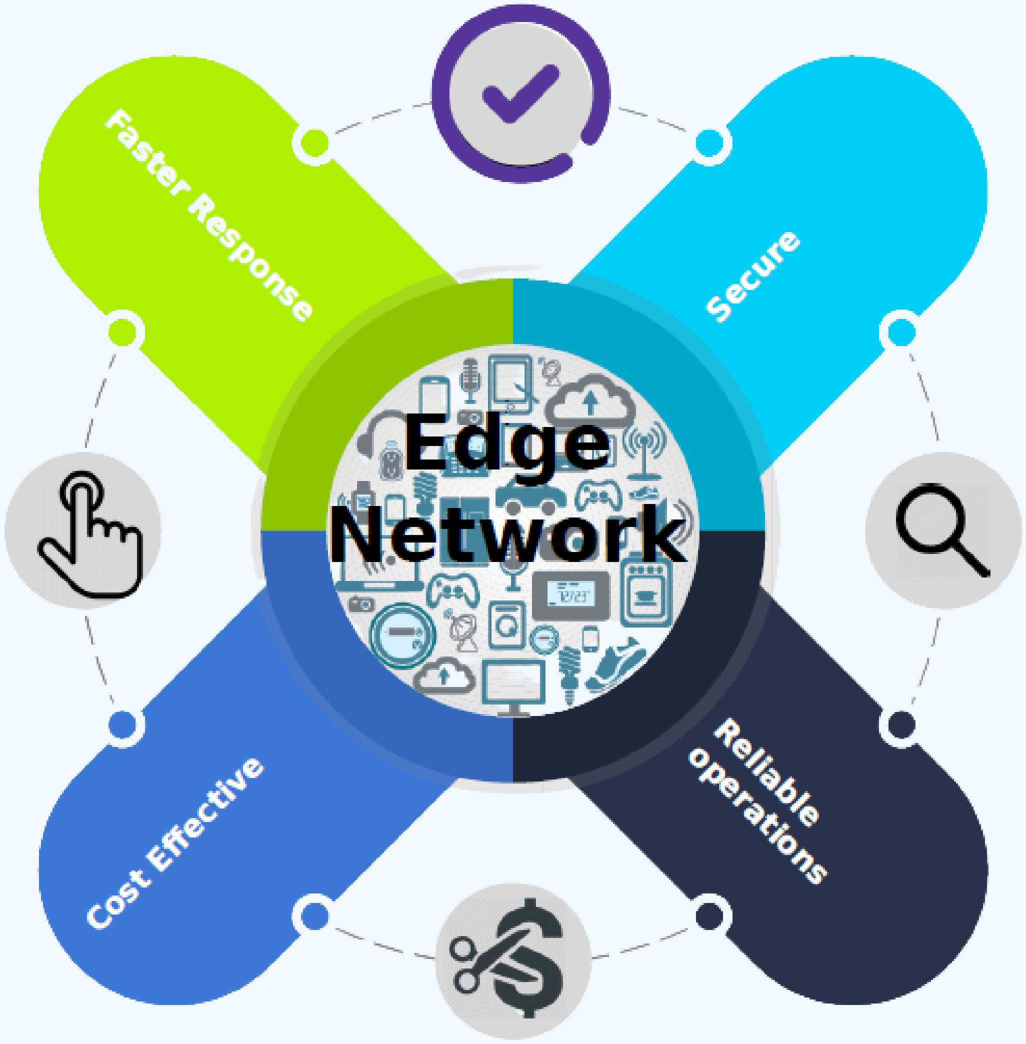
- Applications that require significant **Bandwidth**.

  Generate significant amount of data.
  - Smart surveillance system.
  - Traffic management.



AND EDGE COMPUTING CAN...

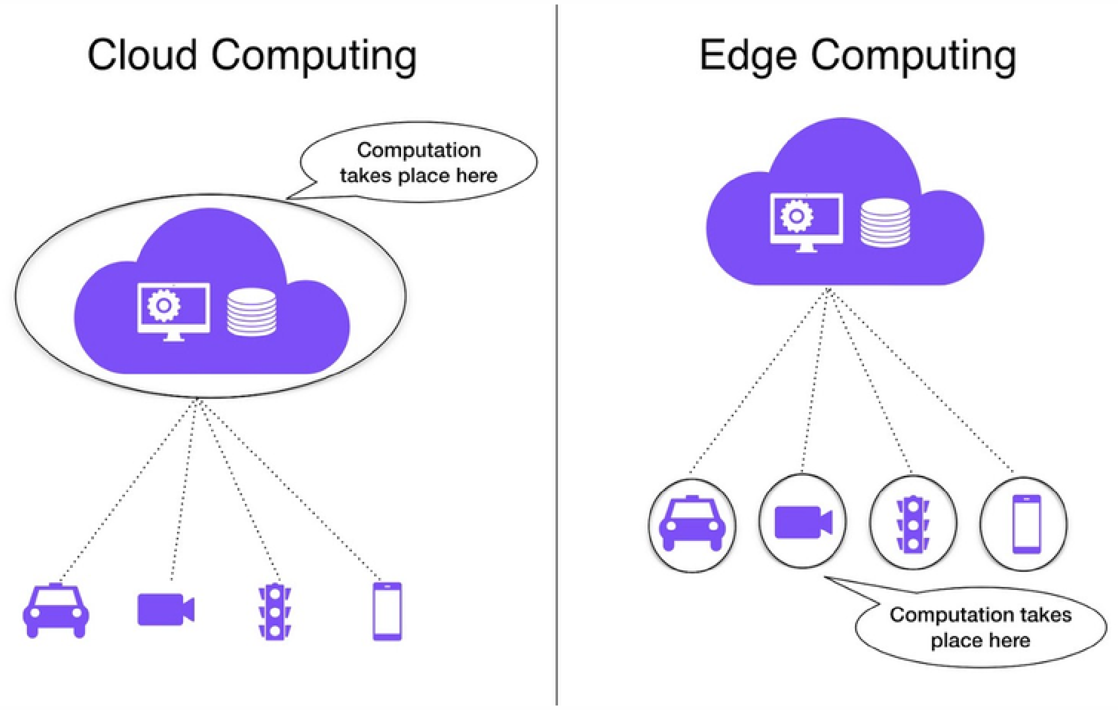
- Reduce **response times** by processing data at source.
- Reduce data transport **costs** to the cloud center.



- Improve **security** by keeping data at source.
- Improve the **reliability** of many operations by being able to work without a connection.

WHAT IS EDGE COMPUTING ?

Edge computing is a **distributed** IT architecture that moves the **compute power** physically closer to where **data is generated**, usually an IoT device or sensor.
Named for the way compute power is brought to the edge of the network or device, edge computing allows for **faster** data processing, increased **bandwidth** and ensured **data sovereignty**, in contrast to Cloud computing where the computation takes place in a central location, increasing the associated costs and response time.



WE NEED MORE <u>SPEED</u>...

Businesses use edge computing to improve the response time of their remote devices. Edge computing makes **real-time computing** possible in locations where it would not normally be feasible.

- In order for a car to drive down any road safely, it must observe the road in **real-time** and stop if a person walks in front of the car. In such a case, **processing visual information** and **decision making** are performed at the edge.

Centralized cloud systems introduce delays in data transmission caused by network latency...

WE CAN'T TOLERATE LATENCY...

The healthcare industry is drowning in data. The **low latency**, **mobility** and **data processing** capabilities of edge computing at a network edge can enable faster and more accurate diagnosis and treatment.

- Edge computing can improve emergency services by allowing paramedics **to make faster and more accurate diagnoses on-site**, and providing hospitals with more detailed information on incoming patients.

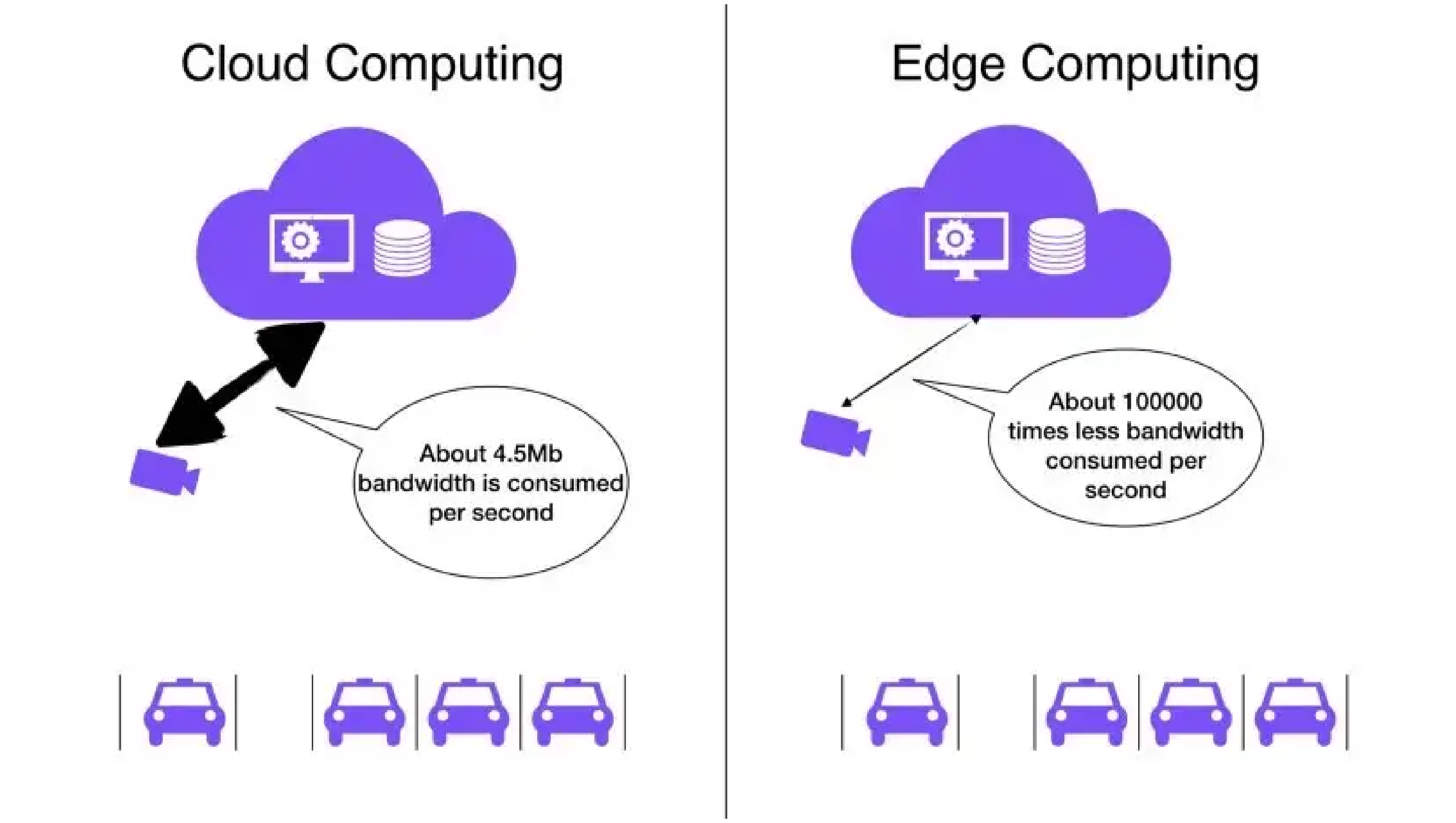
In a centralized cloud system, the execution of these applications is often perceived as insecure due to the sensitivity of patient data...

TOO MUCH BANDWIDTH!

Computing at the Edge means less data transmission as most of the **heavy-lifting** is done by the edge devices. Instead of sending the raw data to the Cloud, most of the processing is done at the edge, and only the result is sent to the Cloud.

In a Smart Park System scenario...

- **Cloud computing**. Would have to send high definition video/images to the cloud to know how many parking spaces are available. **Higher bandwidth**.
- **Edge computing**. Should only send to the central cloud the number of available parking spaces. **Bandwidth reduction**.

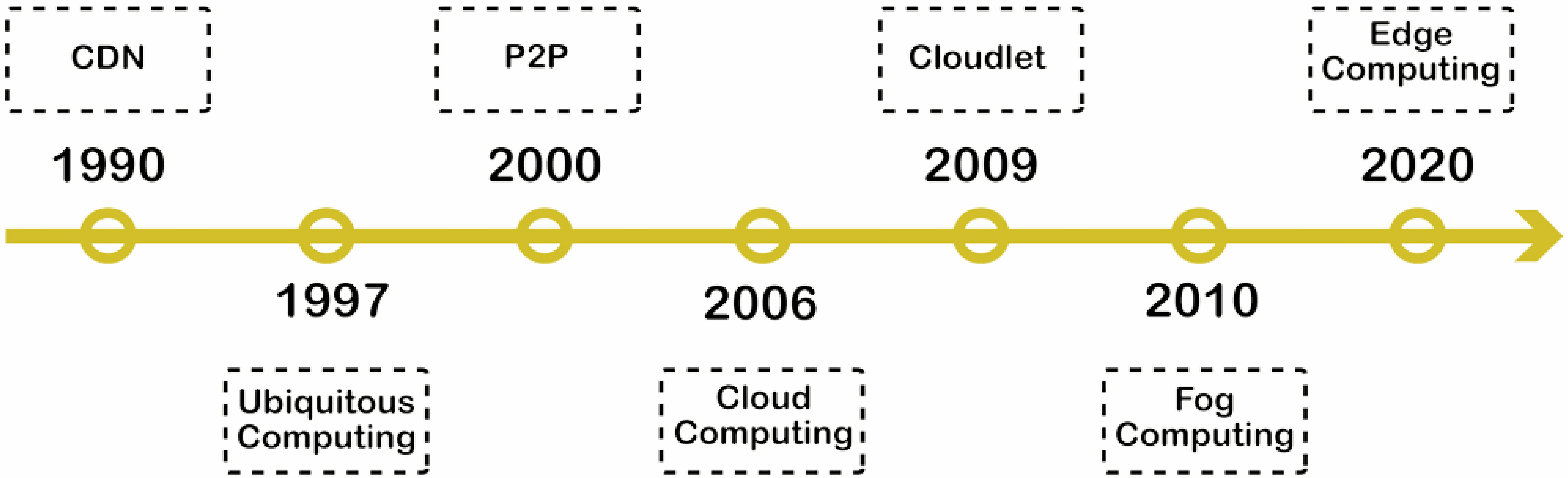


MILESTONES OF EDGE COMPUTING

Akamai, **Content delivery network**

Proximity-routing in overlay P2P networks

Satyanarayanan et al., **The case for VM-based cloudlets in mobile computing**



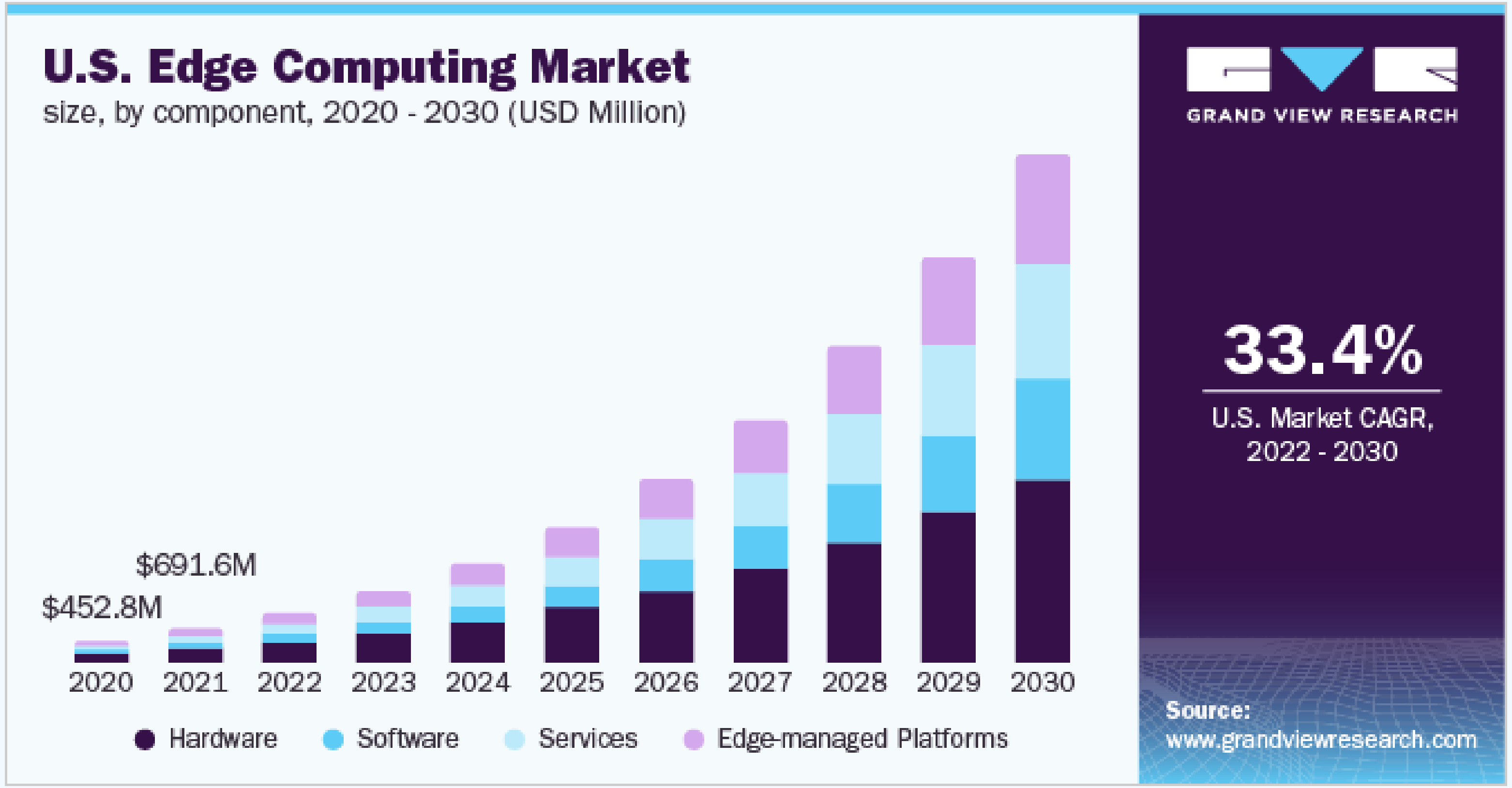
Nobel et al, **Agile application-aware adaptation for mobility**

Amazon, **Elastic Compute Cloud**

Cisco, **Introduced the term "fog computing"**

WHAT DOES MARKET SAY ?

- CAGR of **33.4%** from 2022 to 2030.
- In terms of components, **hardware** holds 40% of the market in 2021.
- In terms of applications, **IoT** holds 28% of the market in 2021.

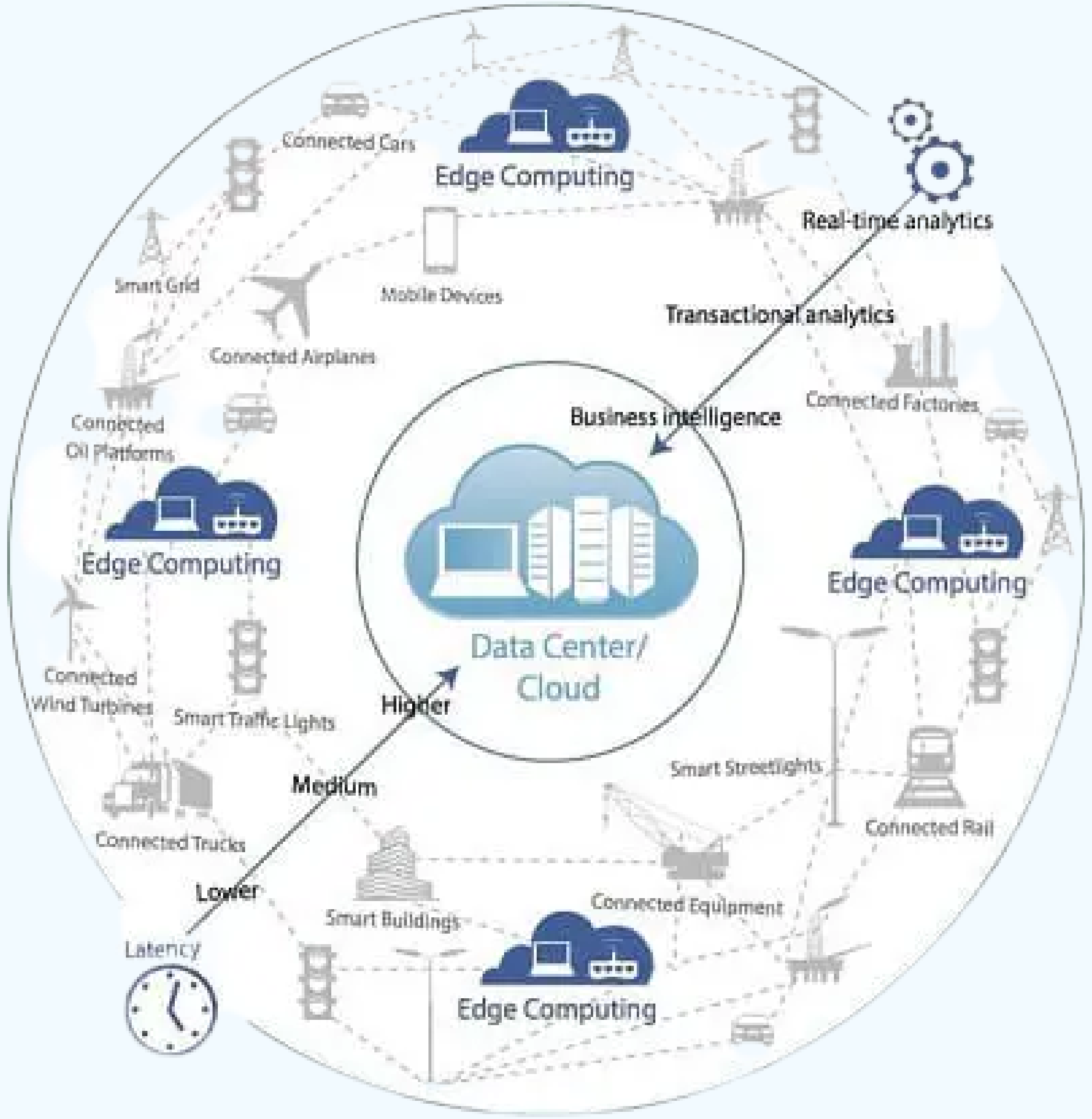


According to **Gartner**, "Enterprises that have deployed edge use cases in production will grow from about 5 percent in 2019 to about 40 percent in 2024."

https://www.grandviewresearch.com/industry-analysis/edge-computing-market https://www.gartner.com/en/documents/3986290

KEY CONCEPTS OF EDGE COMPUTING

- Real-time data processing with low latency.
- Geographically distributed.
- Reliable & resilient.
- Data sovereignty.
- User interaction.
- IoT.



REAL-TIME DATA PROCESSING

- **Latency** is the time needed to send data between two points on a network. Although communication ideally takes place at the speed of light, large physical distances coupled with network congestion or outages can delay data movement across the network.

- By **distributing servers and storage** where the data is generated, edge computing can handle many devices using bandwidth exclusively from the **local devices generating the data**, making latency virtually non-existent.

- Local servers can perform essential **edge analysis** and **data pre-processing** to make **real-time decisions,** before sending the **essential data** to the central cloud.

RELIABLE & RESILIENT

- Edge computing continues to operate even when communication channels are **intermittently available** or **temporarily down** (i.e. edge computing deployments on an oil rig).

- Edge computing doesn't necessarily require a **constant connection** to transmit data to the central cloud; instead, it can choose to move **only the processed information** to the central cloud when the **connection is available**.

- Edge computing also improves **resilience** by reducing the central **point of failure**, improving the **reliability** of the entire connected environment.

DATA SOVEREIGNTY

- Moving huge amounts of data isn't just a technical problem. **Data's journey** across national and regional boundaries can pose additional problems for **data security**, **privacy** and other **legal issues**.

- Edge computing can be used to keep data close to its source and within the bounds of prevailing **data sovereignty laws**, such as the European Union's **GDPR**.

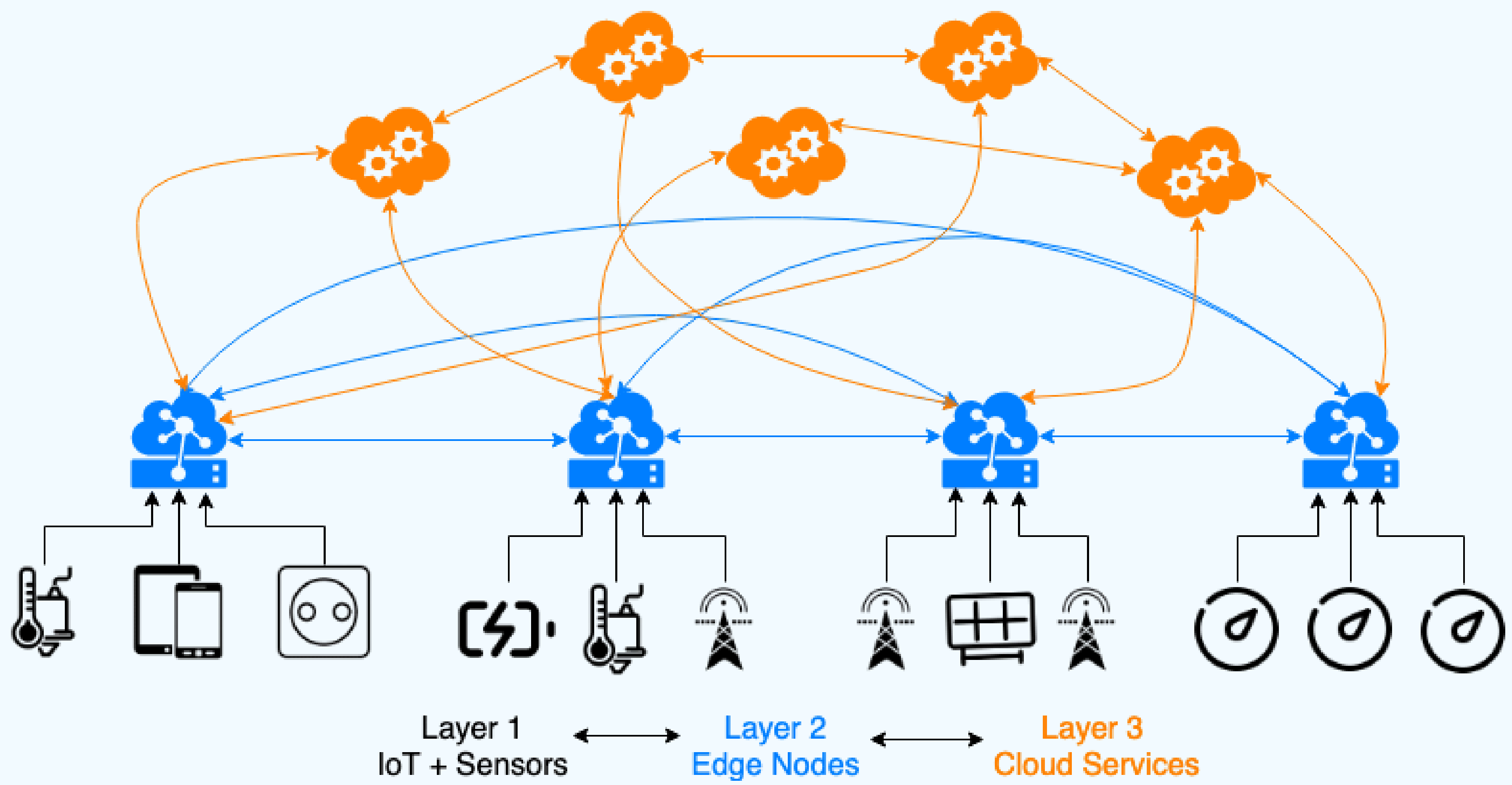
- This can allow raw data to be processed locally, **obscuring** or **securing** any **sensitive data** before sending anything to the cloud or primary data center, which can be in other **jurisdictions**.

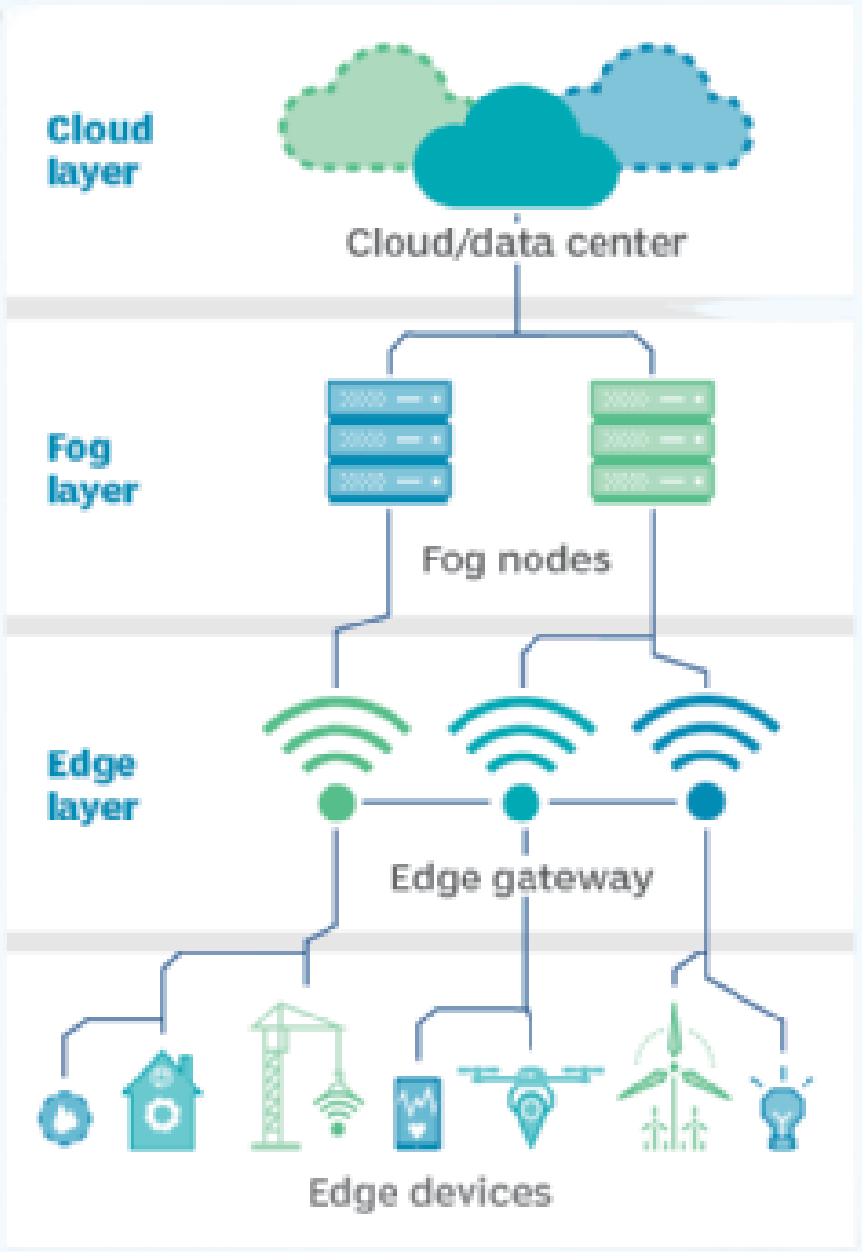
USER INTERACTION & IoT

In terms of interacting with users, edge computing can enable devices such as smartphones and IoT devices to perform **processing tasks** without needing to constantly send data back and forth to a central Cloud. This can lead to a **better user experience**, as well as improved security and privacy.

- For example, an IoT device such as a **smart thermostat** can use edge computing **to process sensor data** locally and make adjustments to the temperature in **near real-time**, without needing to send that data to a central server for processing.

EDGE COMPUTING ARCHITECTURE



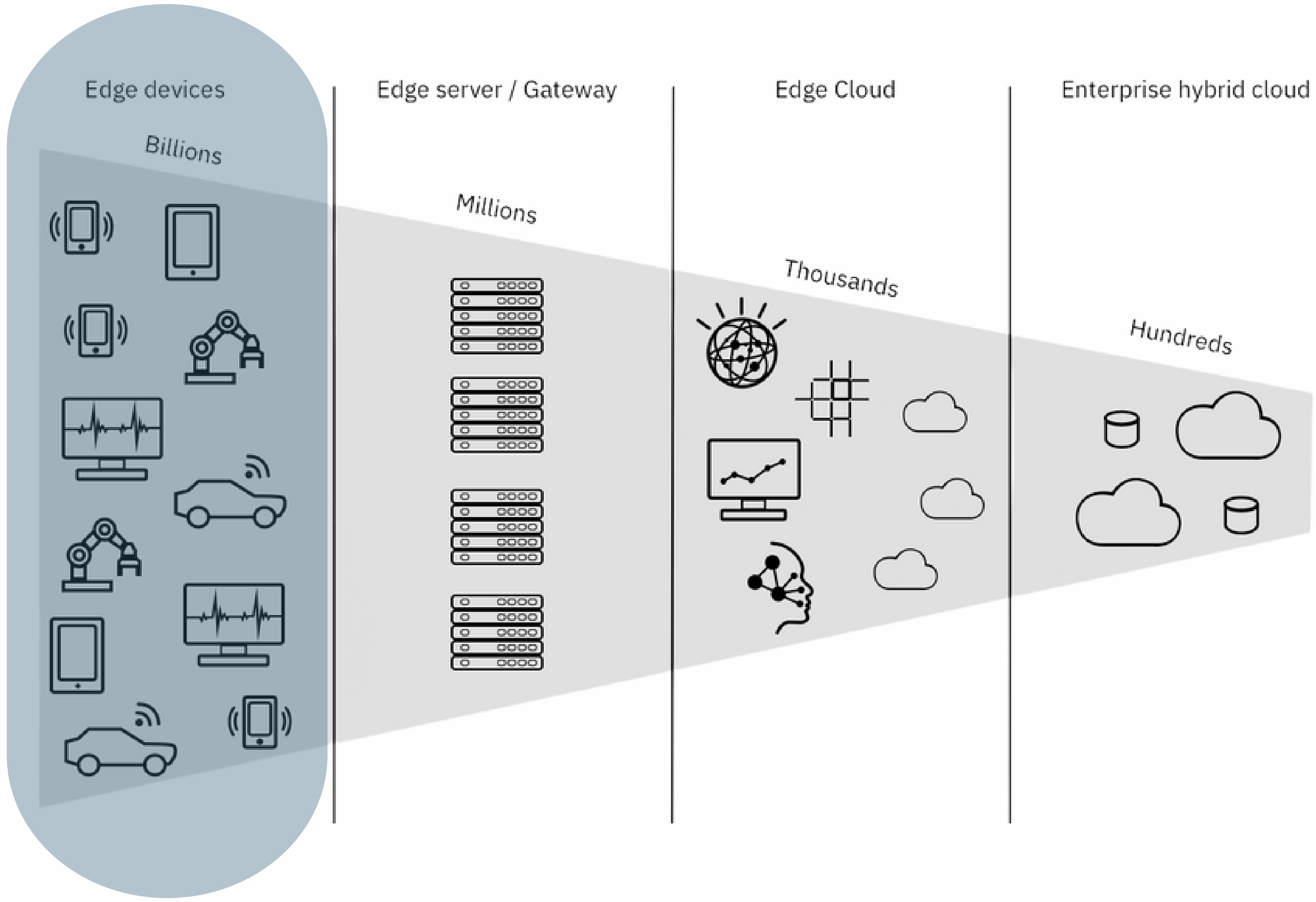


Edge computing **hardware** often consists of several physical components...

EDGE COMPUTING HARDWARE & NETWORKING

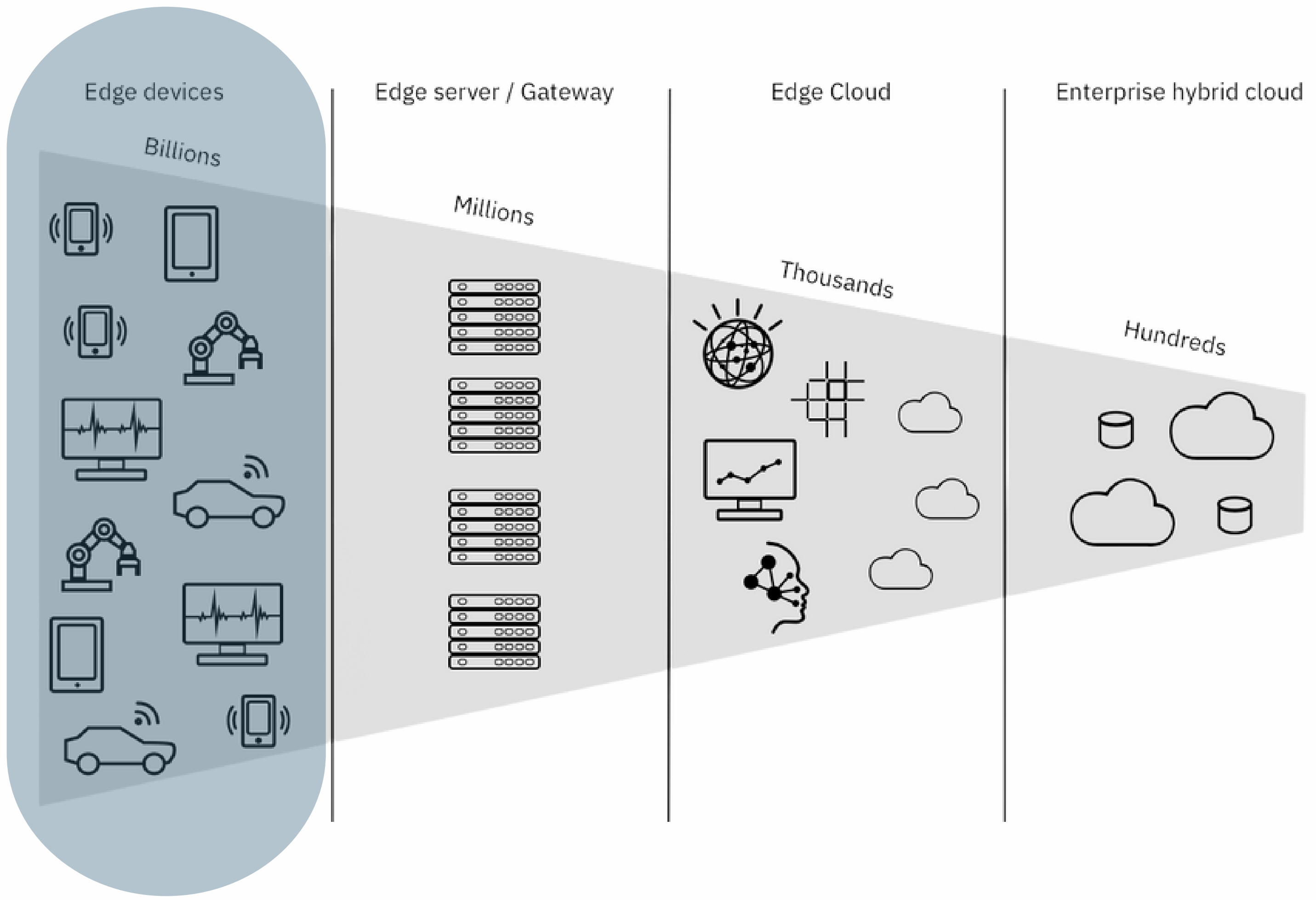
Edge devices

- The edge and IoT devices are equipped to **run analytics**, apply **ML algorithm**, and even **store some data locally** to support operations at the edge.
- They typically have **limited computational resources**



EDGE COMPUTING HARDWARE & NETWORKING
Edge devices

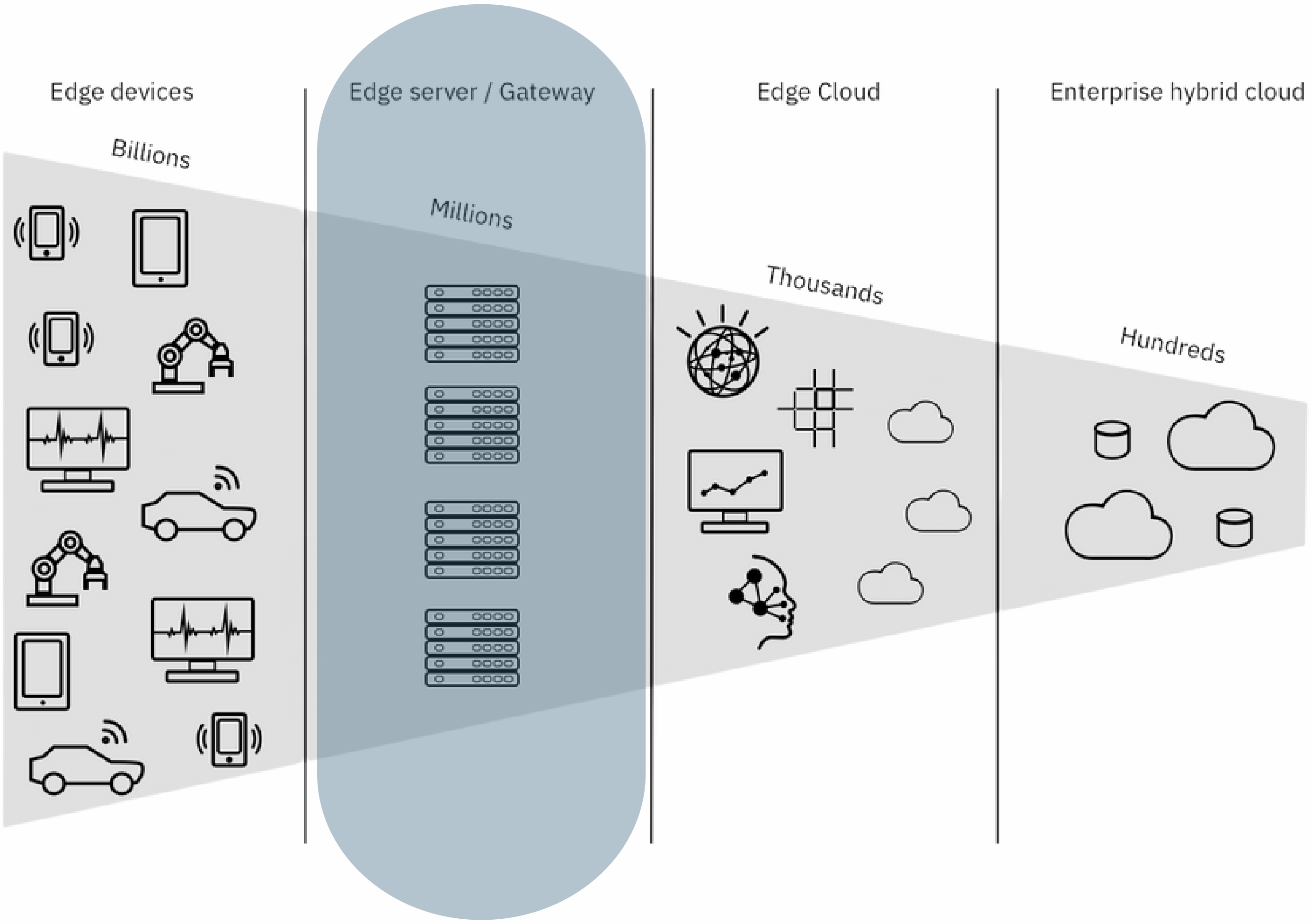
- **Routers** are edge devices that **connect networks**. For example, a router at the edge may be used to connect an enterprise's **LANs** with a **WAN** or the internet.
- **Switches**, which are also referred to as access nodes, **connect several devices** in order to create a network.
- **IoT devices**, like smart cameras, thermometers, robots, drones, vibration sensors ecc...



EDGE COMPUTING HARDWARE & NETWORKING

Gateway

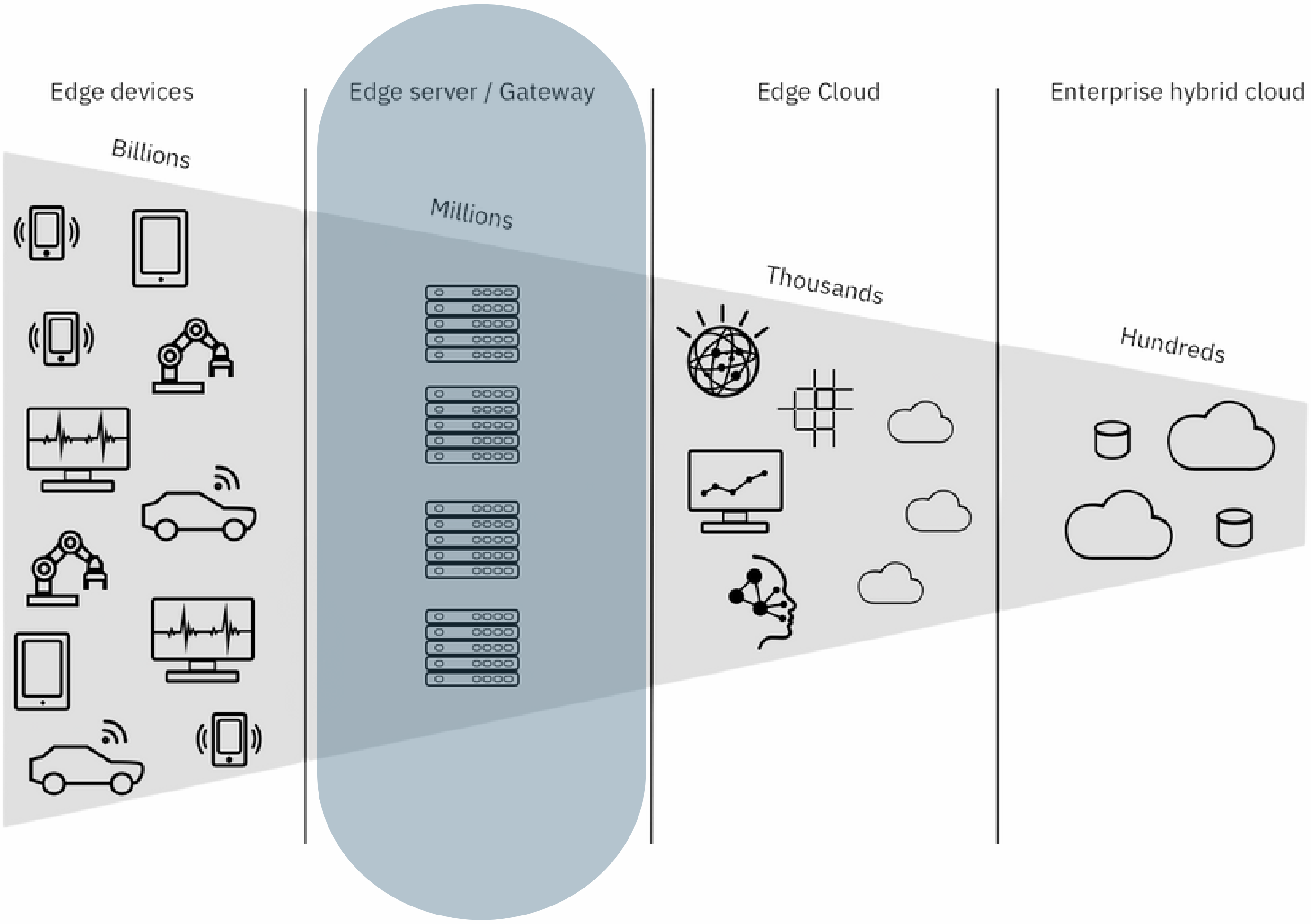
- **Gateways** are a crucial part of edge computing architecture that extends **cloud capabilities** to local edge devices and performs essential network functions like:
  - enabling **wireless connectivity**.
  - providing **firewall protection**.
  - **processing** and **transmitting** edge devices **data**.



EDGE COMPUTING HARDWARE & NETWORKING

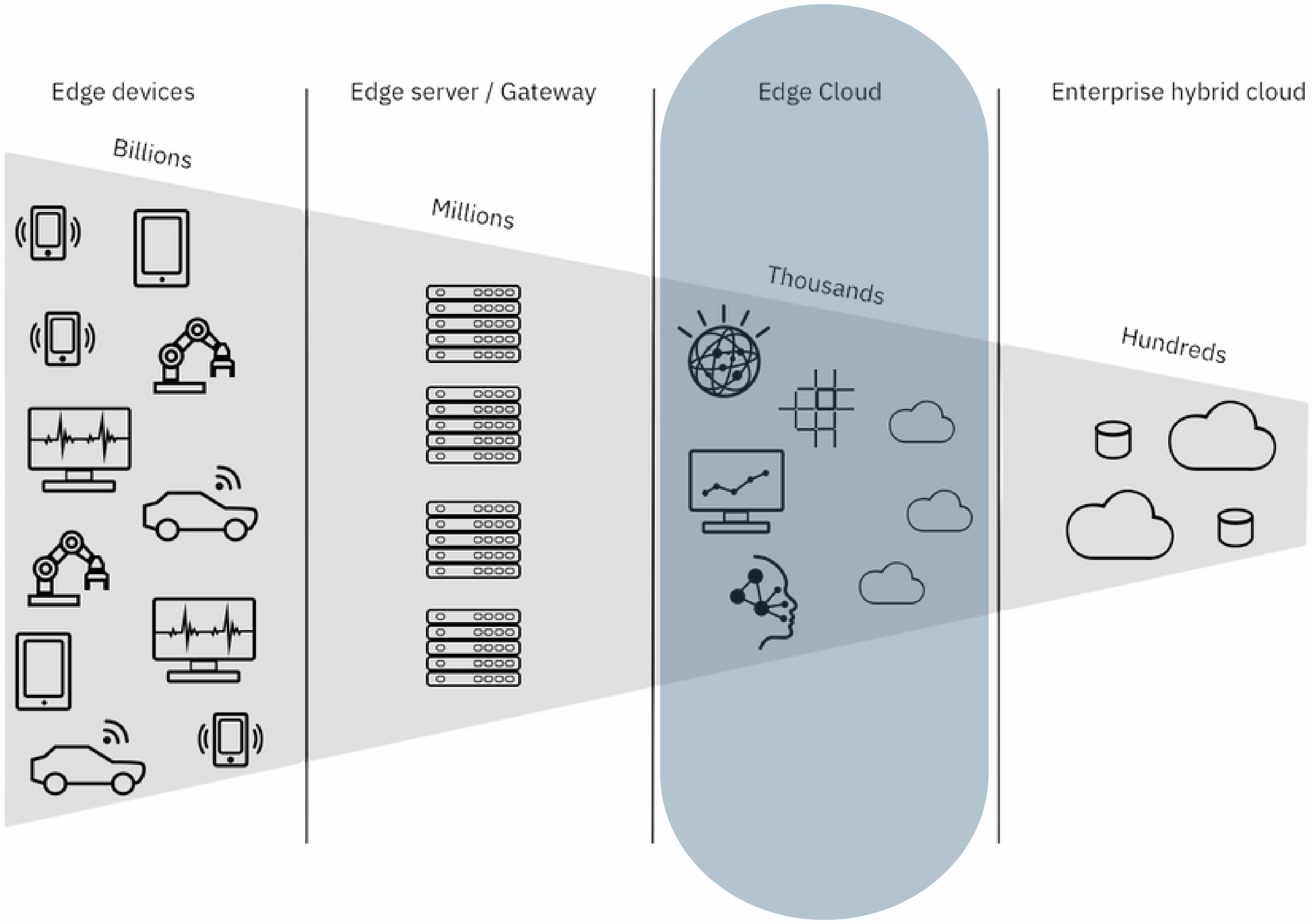
Gateway

- It's like having a **mini cloud** with efficiency, security, **low latency** and **local autonomy**.
- They are in constant communication with the devices by **using agents** that are installed on each of the devices. **Million of gateway** maintain a pulse on the **billions of devices**.
- Some **edge devices** can serve as a limited **gateway** or host network functions, but edge gateways are more often **separate** from edge devices.



EDGE COMPUTING HARDWARE & NETWORKING
Edge Cloud

- New networking technologies have resulted in the **edge network** or edge cloud, which can be viewed as a **local cloud** for devices to communicate with.
- The edge Cloud reduces the **distance** that data must travel from the **devices** and thus decreases latency and bandwidth issues.



EDGE COMPUTING HARDWARE & NETWORKING

Some HW characteristics

- **Temperature resistant**: Edge hardware is often placed outside in freezing, sweltering, and wet climates.

- **Protected from cyberattacks:** Edge devices tend to be more vulnerable to bad actors. So, they must be equipped with security tools like firewalls and network-based intrusion detection systems.

- **Tamper resistant**. Because edge computing devices are often used in far-flung locations where they cannot be consistently monitored.

CONTINUOUS COMPUTING PLANE

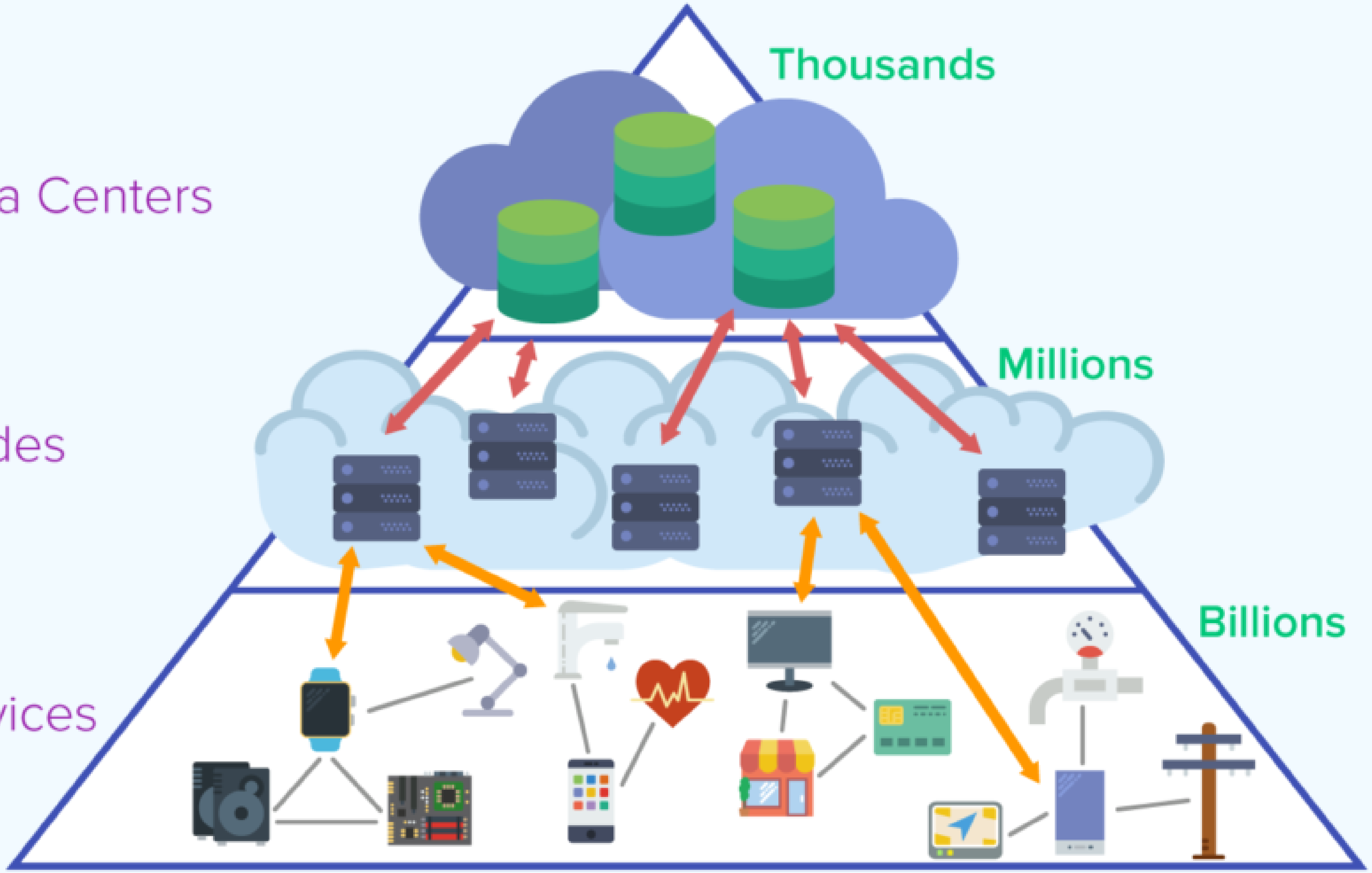
New architectural approach that allows for data and workloads to be processed and stored in various locations ranging **from the cloud to the edge**.

Multiple benefits coming from different environments

CLOUD | Data Centers

FOG | Nodes

EDGE | Devices



How to integrate Cloud and Edge computing?

- Combining strengths of both technologies for a more **efficient** and **effective** computing system.

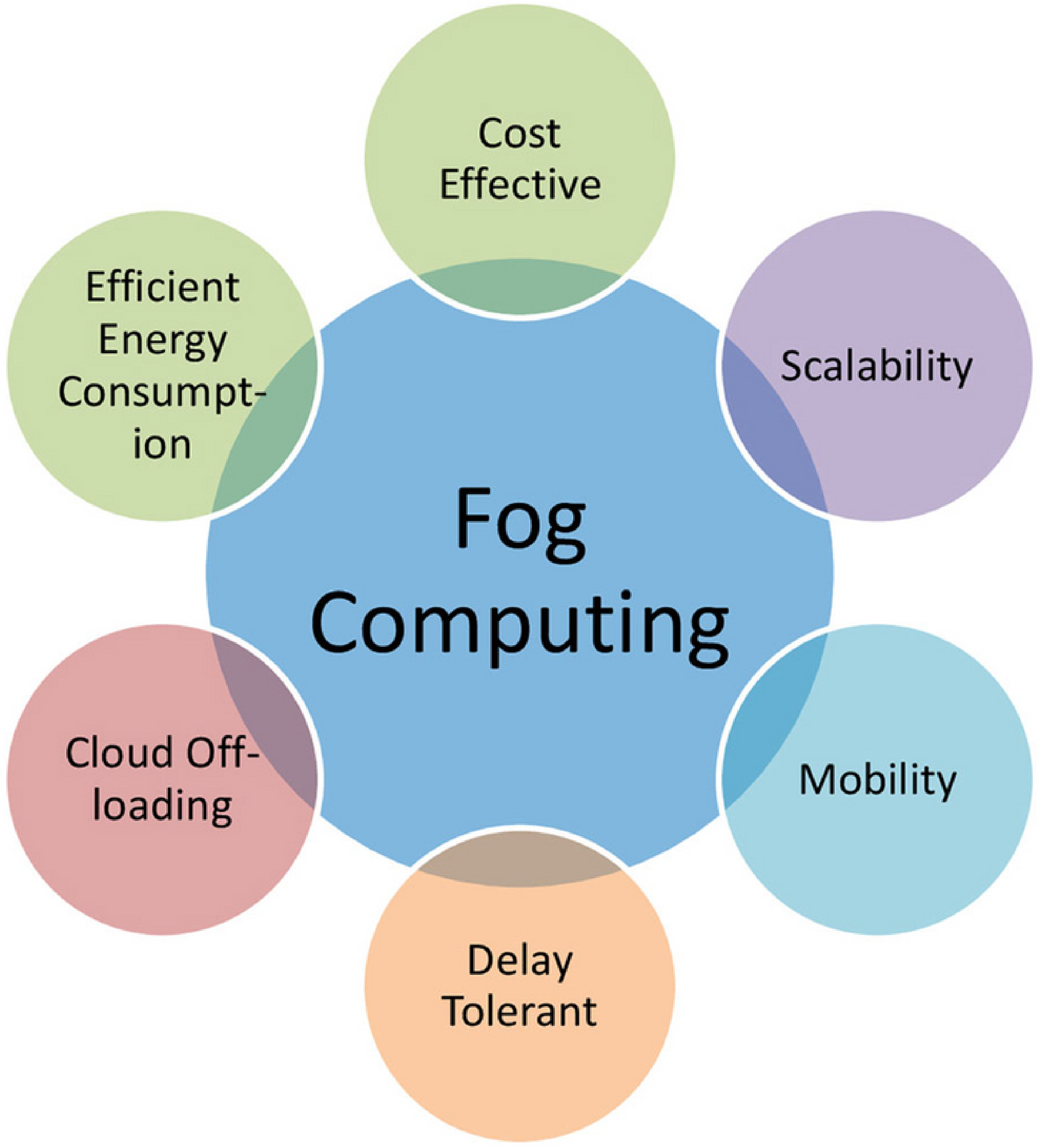
- Common issues:
  - **High-latency communication** between centralized cloud and edge devices must be minimized.
  - Slow **resource management** and **scaling**.
  - Edge devices **low-computational resources** (e.g. sensors, smartband) can't store and process complex real-time data.
  - **Bottleneck** between edge devices and cloud: if all (or most) devices are connected to the same cloud data center, the network needs a very high bandwidth to avoid the bottleneck.
  - **Security** and **privacy** issues in data transfer.

- Need for an **intermediate** layer.

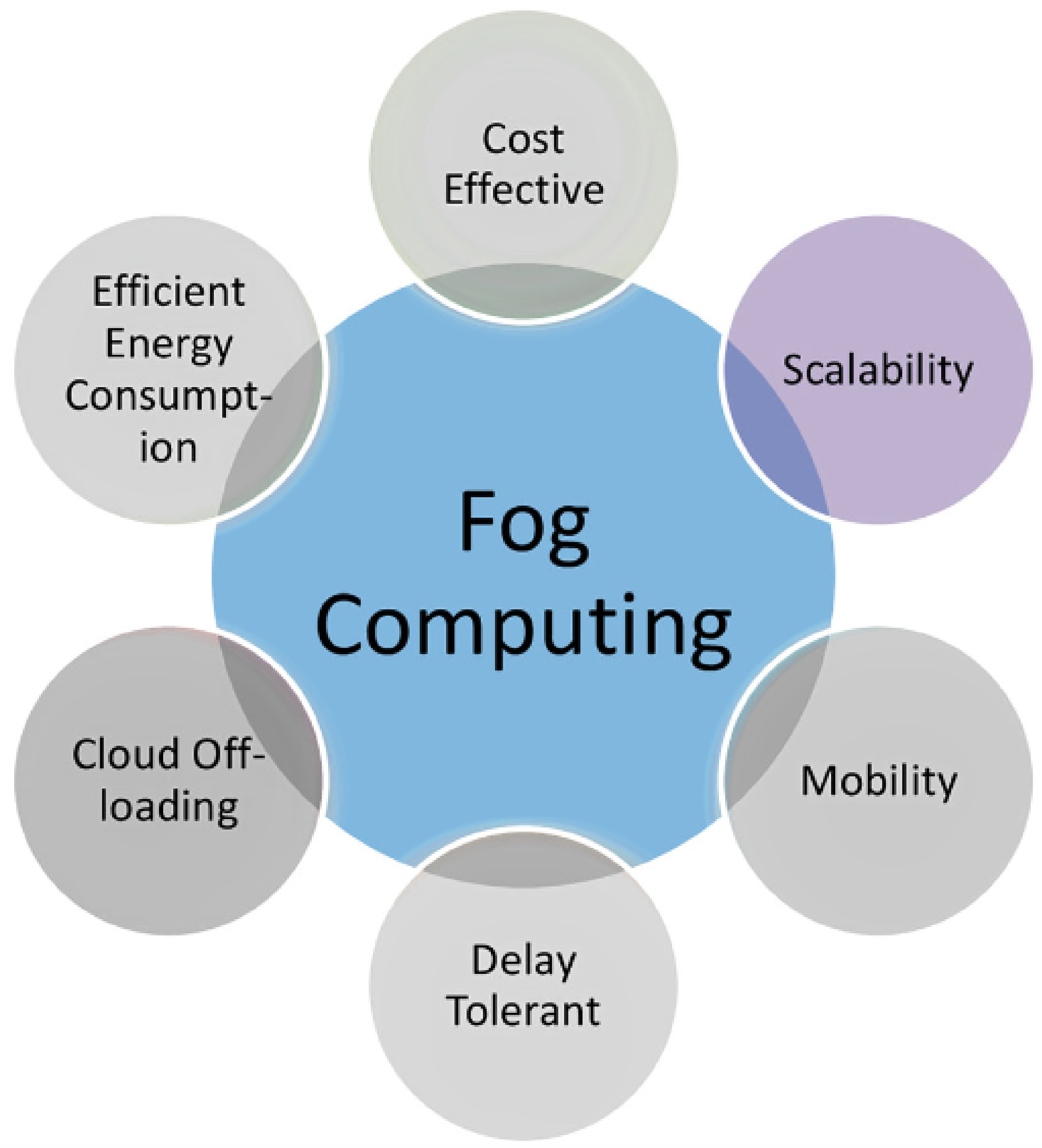
FOG COMPUTING

- **Intermediate layer** between cloud and edge computing.

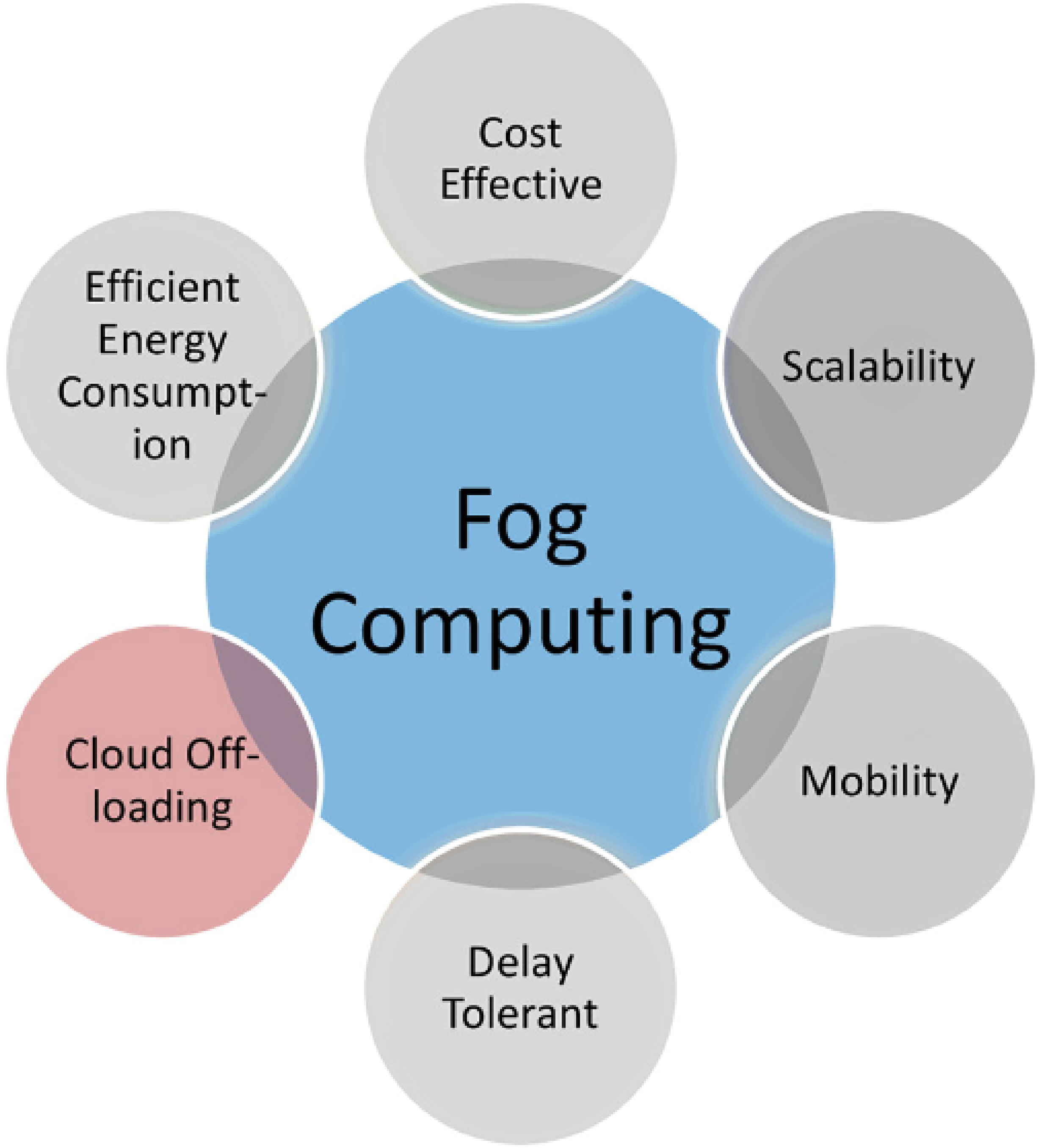
- **Edge clouds (nodes)**
  - **Distributed** infrastructure that places cloud computing resources **closer to the edge** of the network (sort of small data-centers).
  - Provide **low-latency** access to data and services for devices and users that are located near the edge of the network.
  - Equipped with **processing, storage and networking** capabilities.
  - Data **filter** between edge and cloud.



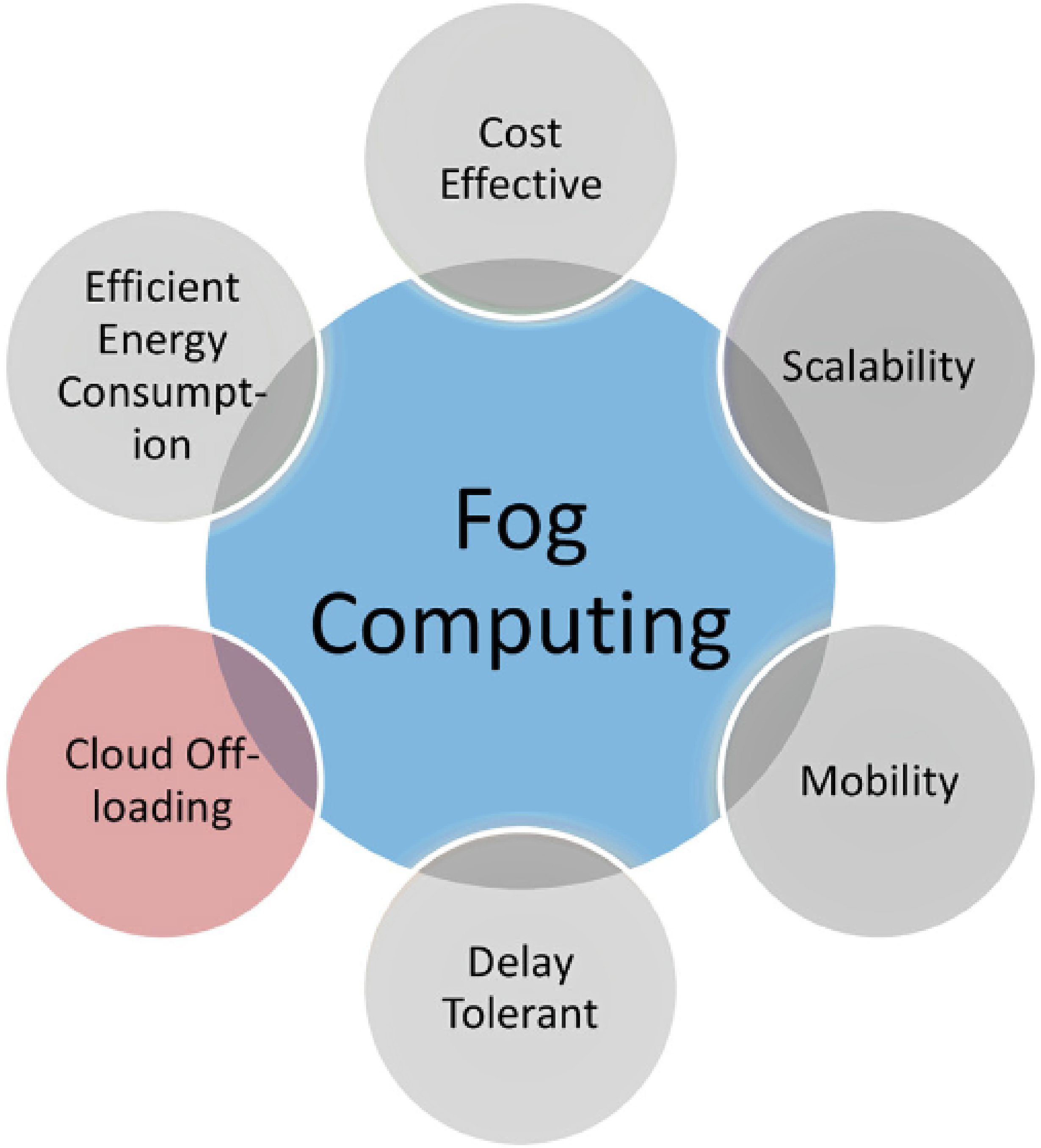
- **Scalability**
  - **Distributed architecture**: additional resources can be added **closer to where they are needed**.
  - **Decentralized management**: resources managed and controlled at the edge, so they can be added or removed without affecting the entire system.
  - **Dynamic resource allocation** based on the needs of users and devices.
  - **Offloading** to the cloud (hybrid architecture).



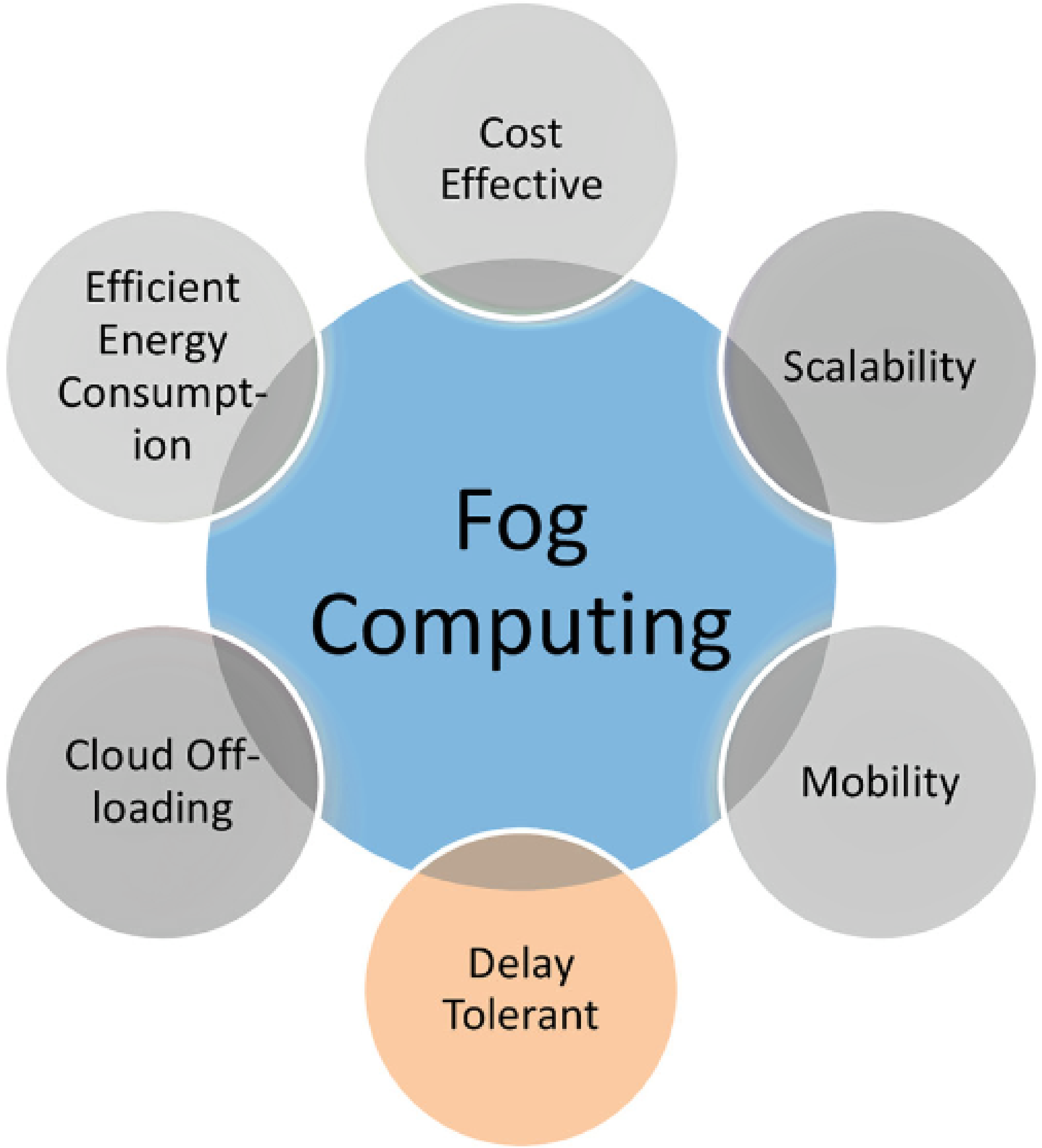
- **Cloud offloading**
  - **Transferring** certain computing tasks or data to a **more powerful resource** for processing.
  - To **nearby fog nodes**:
    - Tasks requiring low latency (e.g. real-time video processing)
    - Data that are too large to be transmitted over long distances.
  - To a **central cloud**:
    - Tasks that require more powerful processing capabilities, such as machine learning or big data analysis.



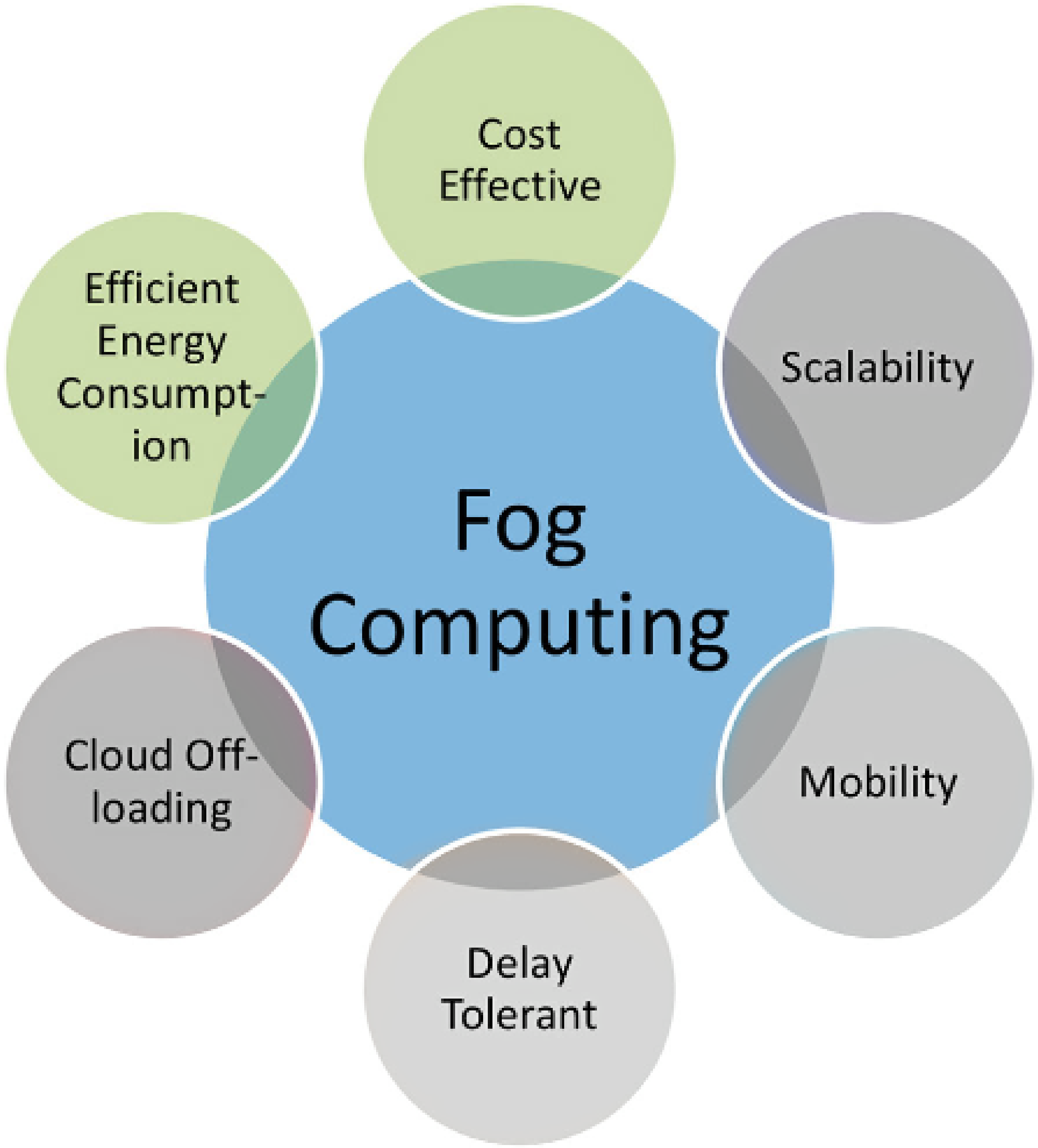
- **Common algorithms**
  - **Static**: predefined rules based on the features of the application and the device resources.
  - **Dynamic**: based on the current state of the device resources, the network conditions and the processing requirements of a specific task.
  - **Hybrid**: combines static and dynamic strategies.
  - **ML-based**: uses machine learning algorithms to analyze data and make offloading decisions.
- Improved **performances** and **scalability**, as it allows for more efficient use of resources.



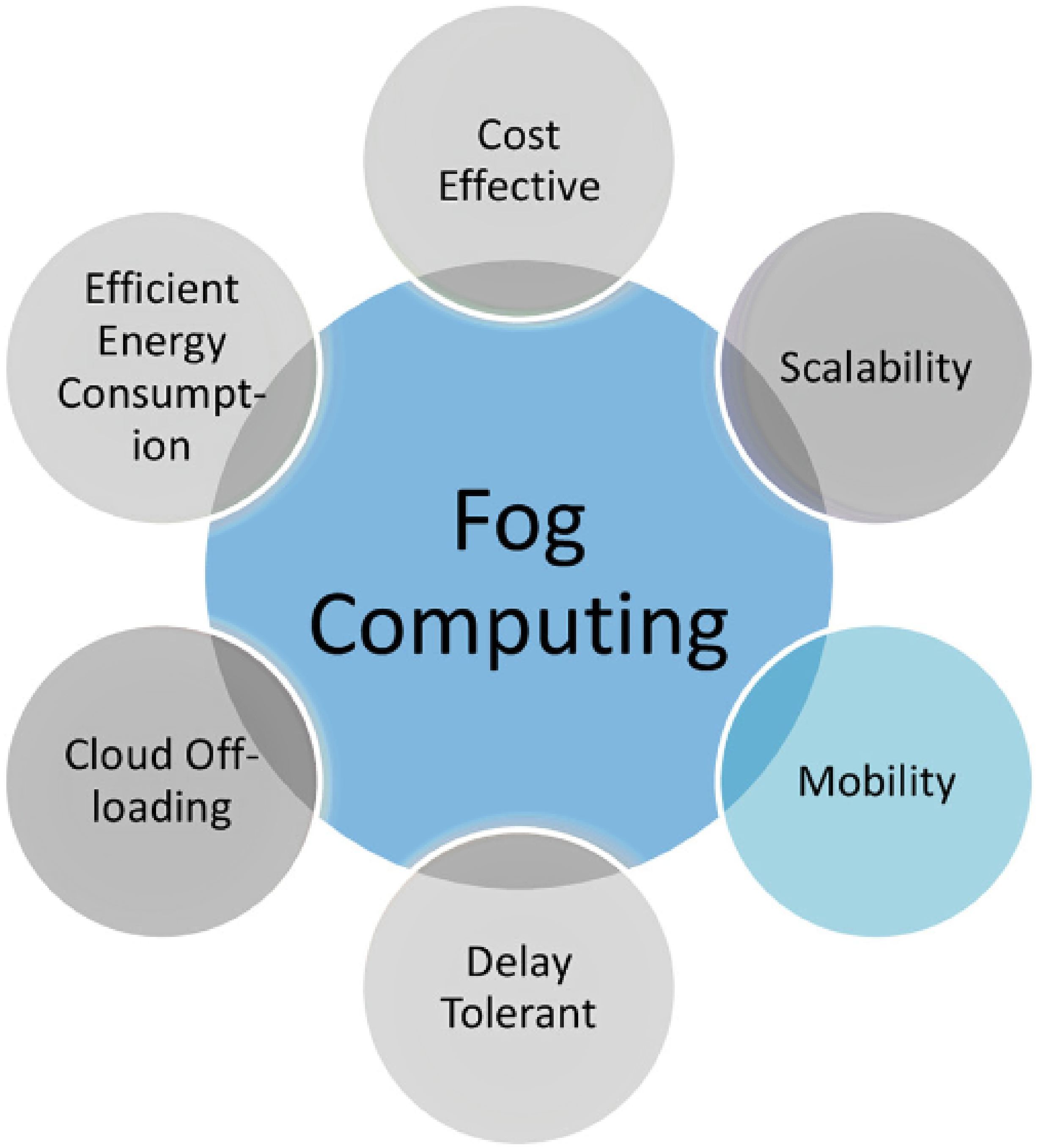
- **Delay tolerant**
  - Ability of the system to continue functioning even in presence of network delays or disruptions.
  - Achieved through:
    - **Local processing**.
    - **Redundancy**: multiple fog nodes and devices.
    - **Offloading**: transfers certain computing tasks or data in case of network failure or delay.
    - **Scheduling** and **prioritization**.



- **Cost effectiveness** and **efficient energy consumption**.
  - Data processed closer to the source.
  - No need for continuous transmission to centralized data-center.
  - Reduced dependence from centralized cloud resources.
  - Greener computing.



- **Mobility**
  - Ability of the system to adapt to changes in the location and movement of devices and users.
  - Achieved through:
    - **Handover management**: continuity of service and low-latency connections as devices move between different fog nodes.
    - **Context awareness**: providing services tailored to the specific context of the user with the support of IoT devices.
    - **Mobile edge computing (MEC)**: MEC servers can provide low-latency, high-bandwidth access to data and services for mobile devices (especially with **5G**).



ORCHESTRATION

- Need for an **orchestration** system.

- Distribution and coordination of **resources**, **workloads** and **applications** across different computing environments.

- Metrics defined by **SLA**: bandwidth, users needs, processing capabilities, ...

- Different approaches.
  - **Centralized**.
  - **Local**.
  - **Distributed**.

Cloud orchestration domain

- Cloud characterized by:
  - **Centralization**. Many hosts in few central locations.
  - **Virtually unlimited pool** of resources. Compute, network, and storage resources viewed as unlimited and capable to serve multiple clients at one time.
  - Fast network **connectivity** within data-centers.

- So we need:
  - Rapid **elasticity**. Steady and predictable performance at the lowest possible cost.
  - **Moveable workloads** to solve healing and scaling issues.
  - Data-center **security**.

Edge orchestration domain

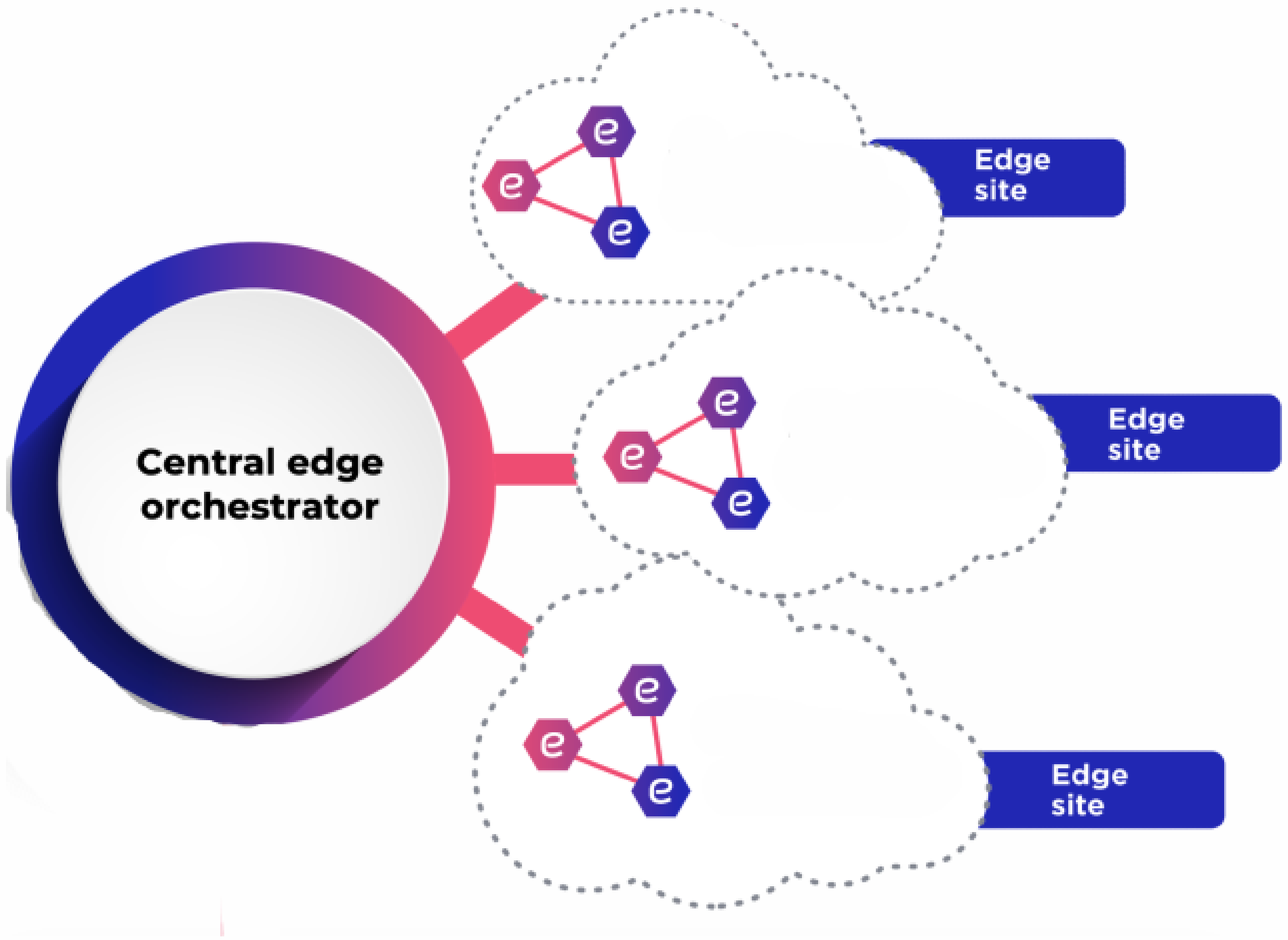
- Edge site characterized by:
  - **Distribution**. Few hosts in many distributed locations.
  - **Limited resources**.
  - **Network limitations**. Frequent network outages.

- So we need:
  - **Autonomy**. Offline operability for a long period of time.
  - **Resilient networking**. Services designed for network outages and slow connections.
  - **Static workloads**. Workloads have to stay where they are deployed, unless offloading.
  - Local **security**.

Centralized orchestration

- **Single centralized controller** to manage and coordinate the deployment and operation of edge devices and services.

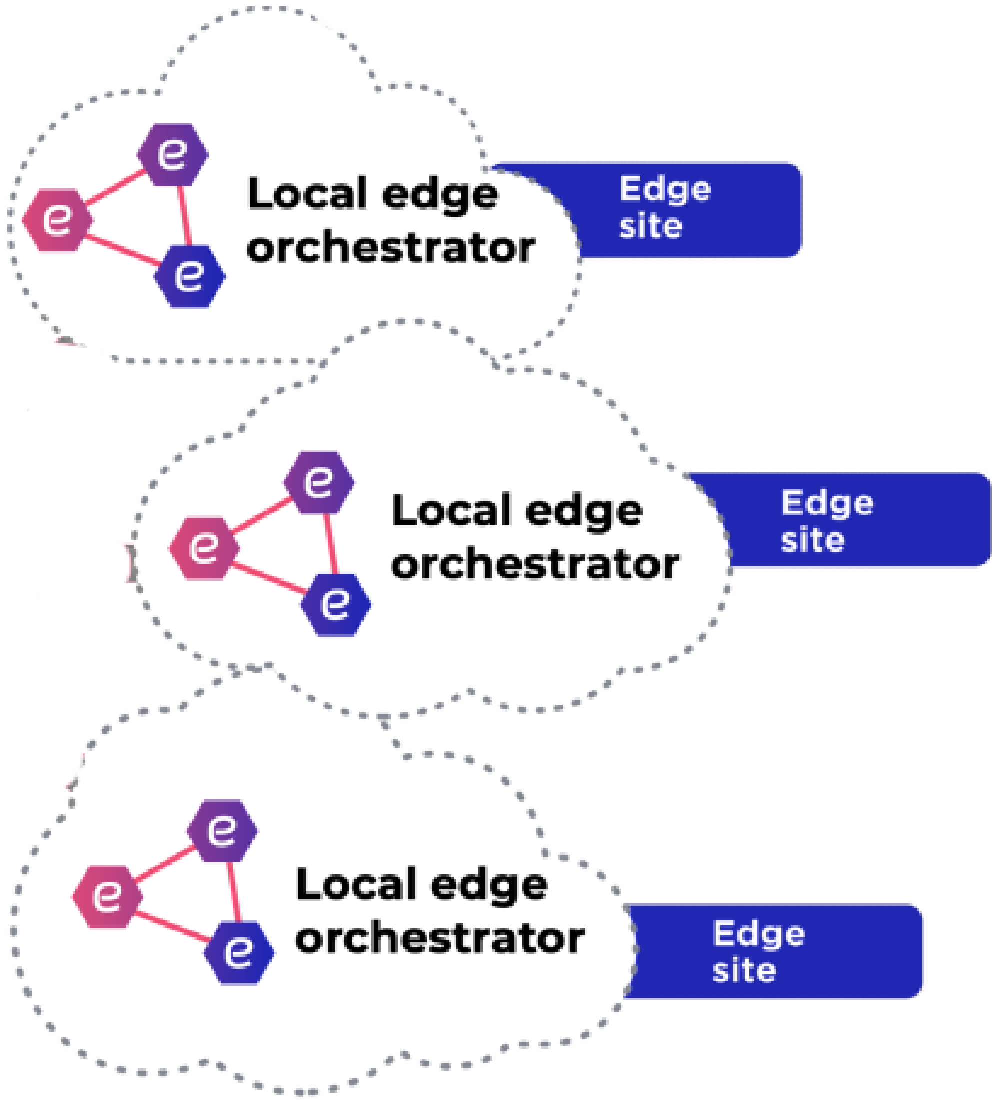
- Pros:
  - **Easier monitoring**.
  - **Security**: access control and policies.
- Cons:
  - **Single point of failure** (also security issue).
  - **Limited scalability**, not suitable for highly distributed systems.
  - **Latency and bandwidth issues** as the number of devices and services increase (bottleneck).



Local edge orchestration

- Edge devices and services to manage and coordinate their own deployment and operation, with **minimal centralized control**.

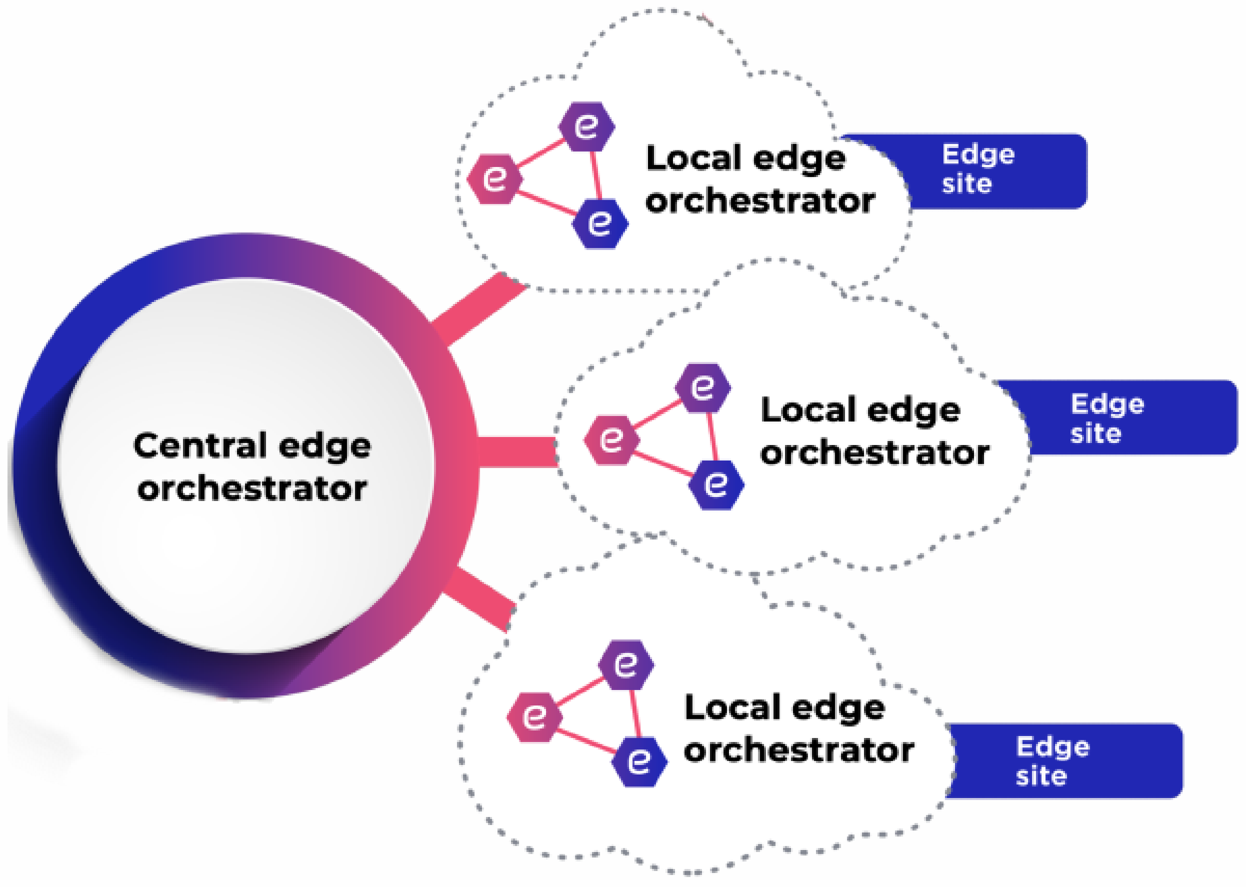
- Pros:
  - **Decentralization and locality**.
    - No single point of failure.
    - Reduced latency.
  - **Autonomy**, but devices need to be equipped with the necessary software and algorithms.
- Cons:
  - **Complex integration and interoperability**.
  - **Limited visibility and control**: it's difficult to monitor and control the behavior of the system.



Distributed orchestration

- **Distributed network of controllers** to manage and coordinate the deployment and operation of edge devices and services.
- Benefits from both **central** and **local** orchestration.

- Pros:
  - **Scalability and flexibility**, suitable for highly distributed systems.
  - **Reliability and redundancy** provided by multiple controllers.
  - **Low latency** for local real-time operations.
  - **Security**: no single point of failure and policies control.
- Cons:
  - **Complex synchronization and interoperability**.

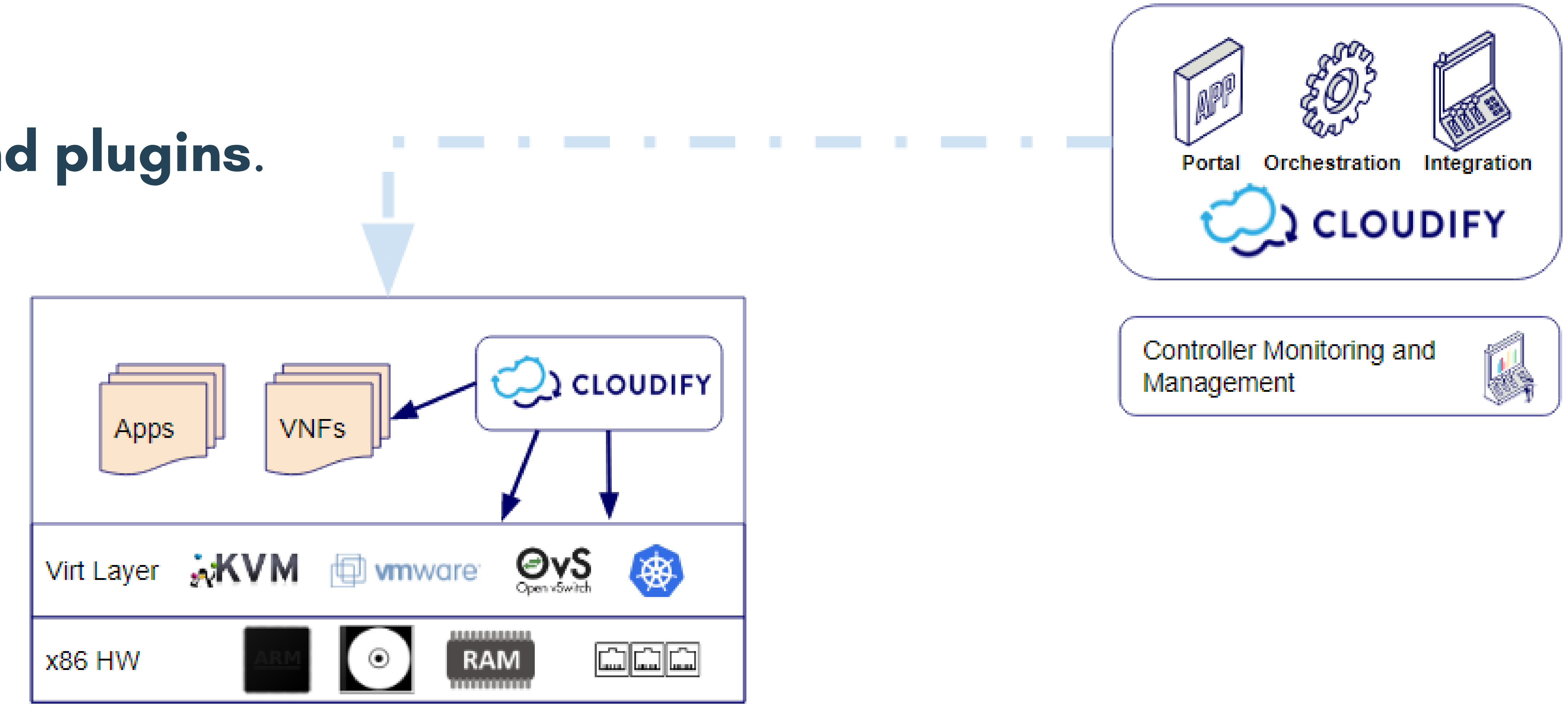


Cloudify

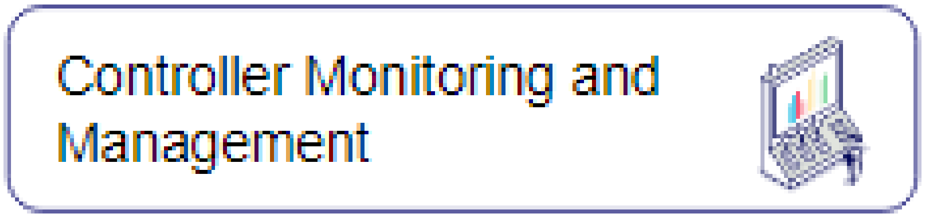
The case of Cloudify

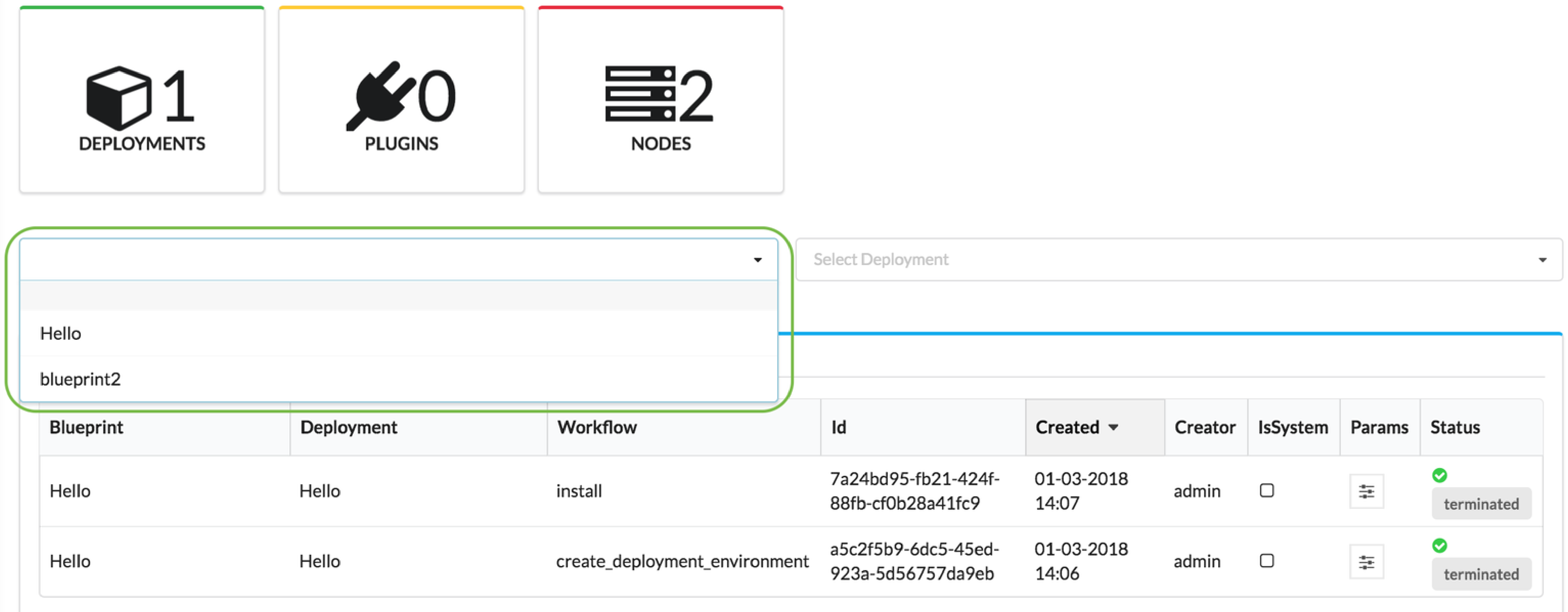
- Open-source unified platform for **management**, **deployment** and **automation** of applications and services **across cloud and edge environments**.

- 3 main components communicating with each other:
  - **Management** console.
  - **Orchestration** engine.
  - Light-weight **local agents and plugins**.



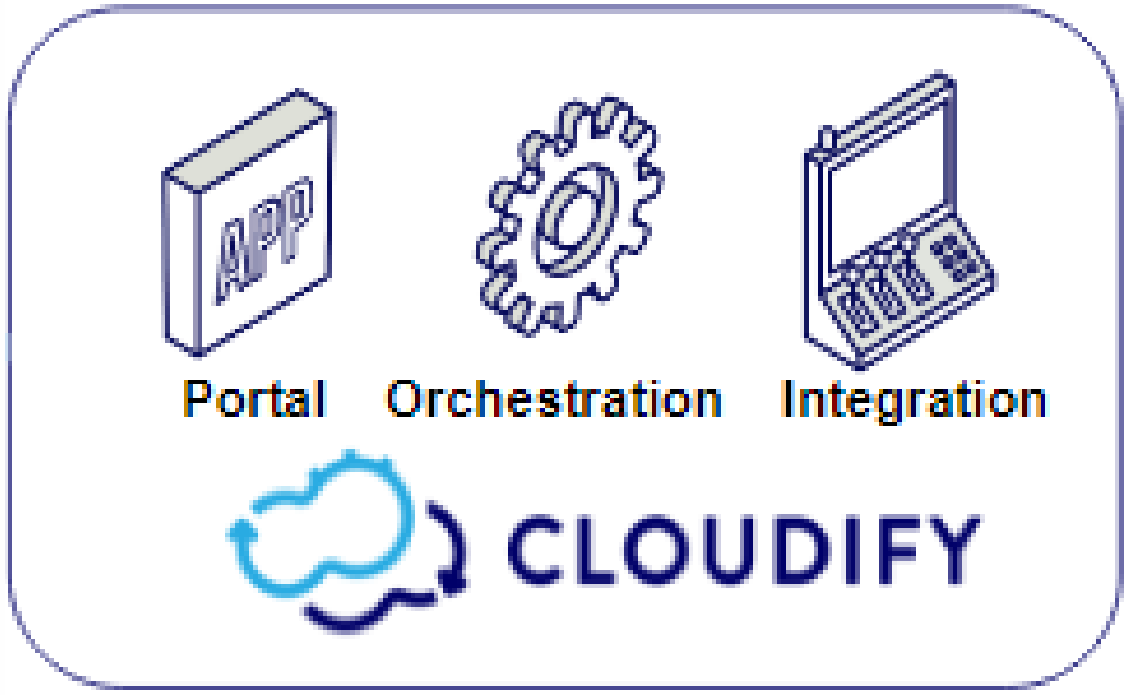
- ***Management console***
  - User web-based **interface**.
  - Managing the entire **application life-cycle**, including provisioning, scaling and monitoring.
  - Provides status and performance metrics of the edge devices.
  - Communicates with the other components.





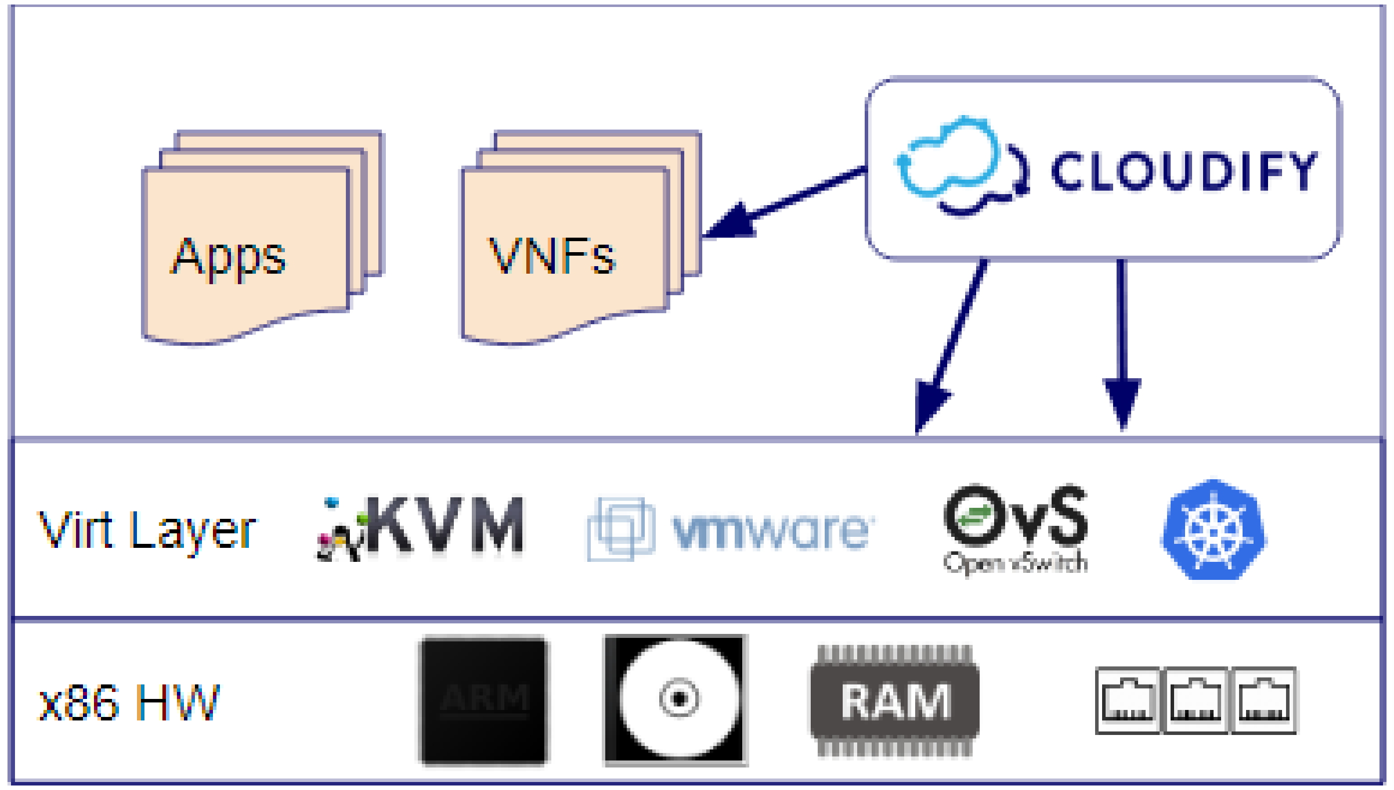
- **Orchestration engine**
  - **Automating deployment and management** of applications and services on the edge devices.
  - 2 ways to specify the desired configuration of edge devices.
    - **Pre-defined templates** with pre-defined policies for scaling and updates.
    - Custom **blueprints**.

  - Cloudify blueprints.
    - **YAML** files defining **components** (such as VMs, load balancers and DBs), **relationships** and **policies** that make up an application or a service.



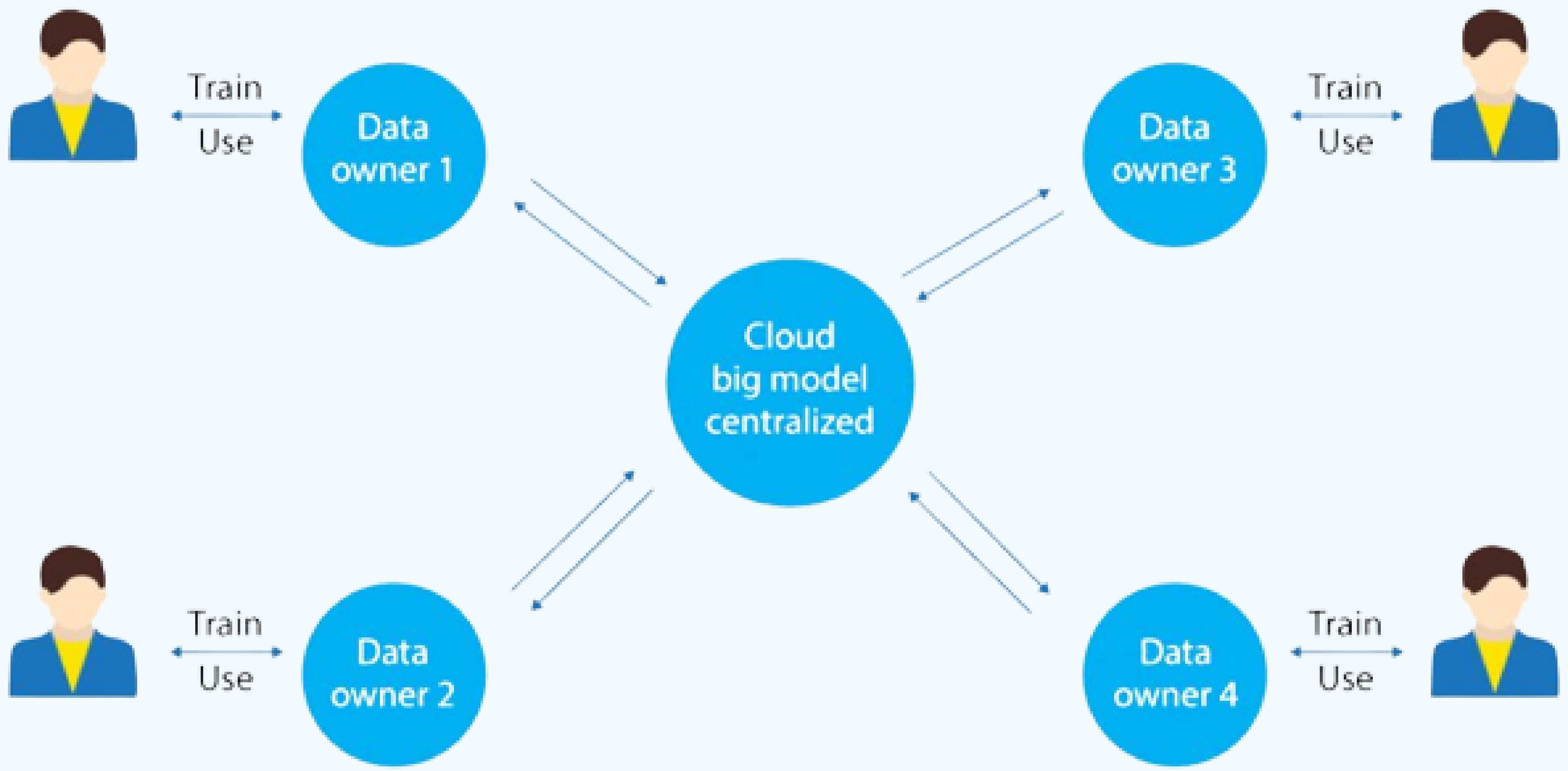
- **Local edge agent**
  - **Autonomous local edge lightweight orchestrator**.
  - Can support a **virtualization layer** or run on a **bare-metal** machine.
  - Can work with a container orchestration system like **Kubernetes** or orchestrate the box **directly**.
  - Communicates with management console and orchestration engine.

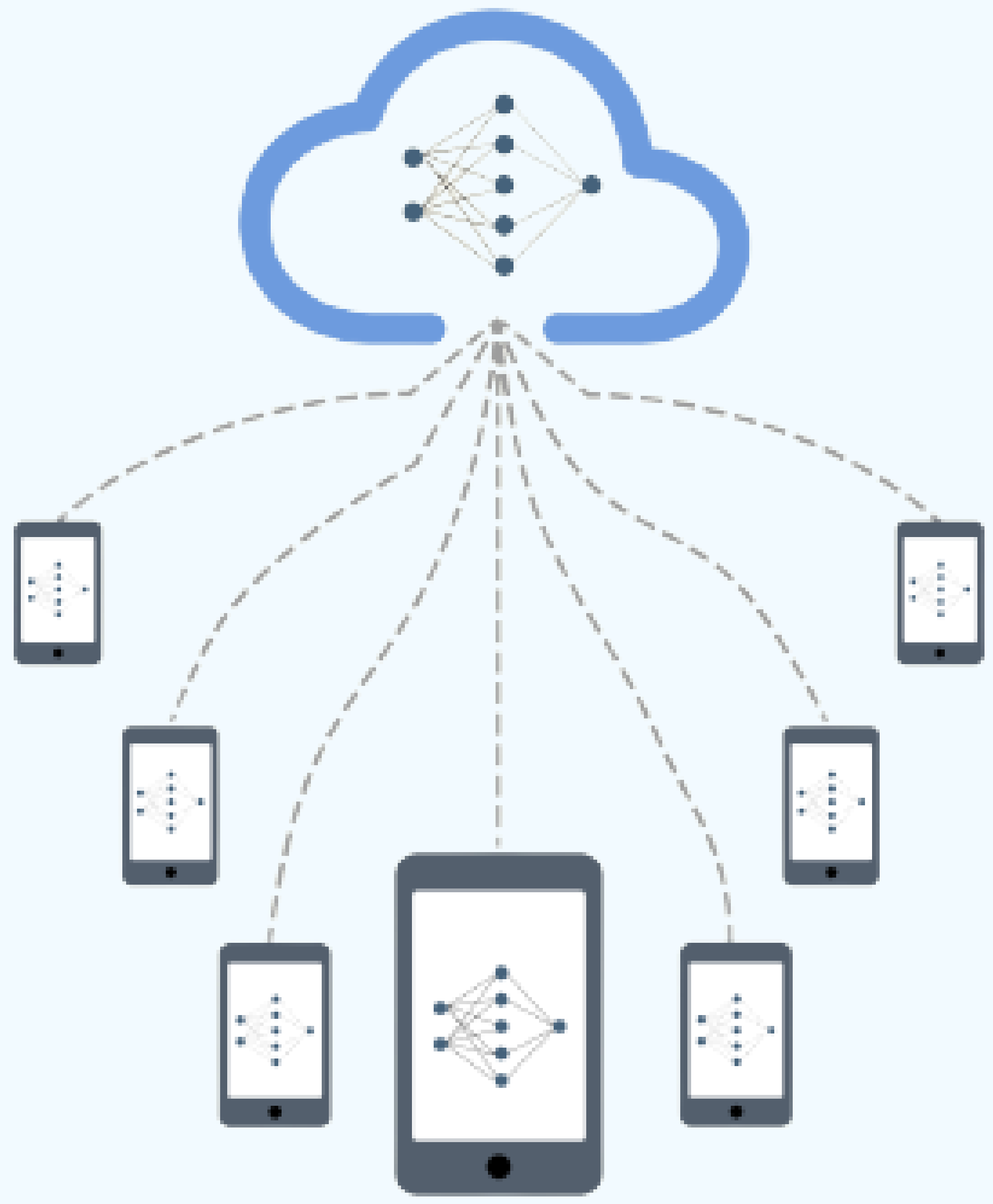
- **Plugins** to extend the functionality of the platform.
  - **Infrastructure** plugins to integrate Cloudify with infrastructure providers for virtual resources (e.g. AWS, Microsoft Azure).
  - **Application** plugins to integrate Cloudify with software systems and tools, such as databases or messaging systems.



FEDERATED LEARNING

Distributed ML approach developed by Google in 2016.



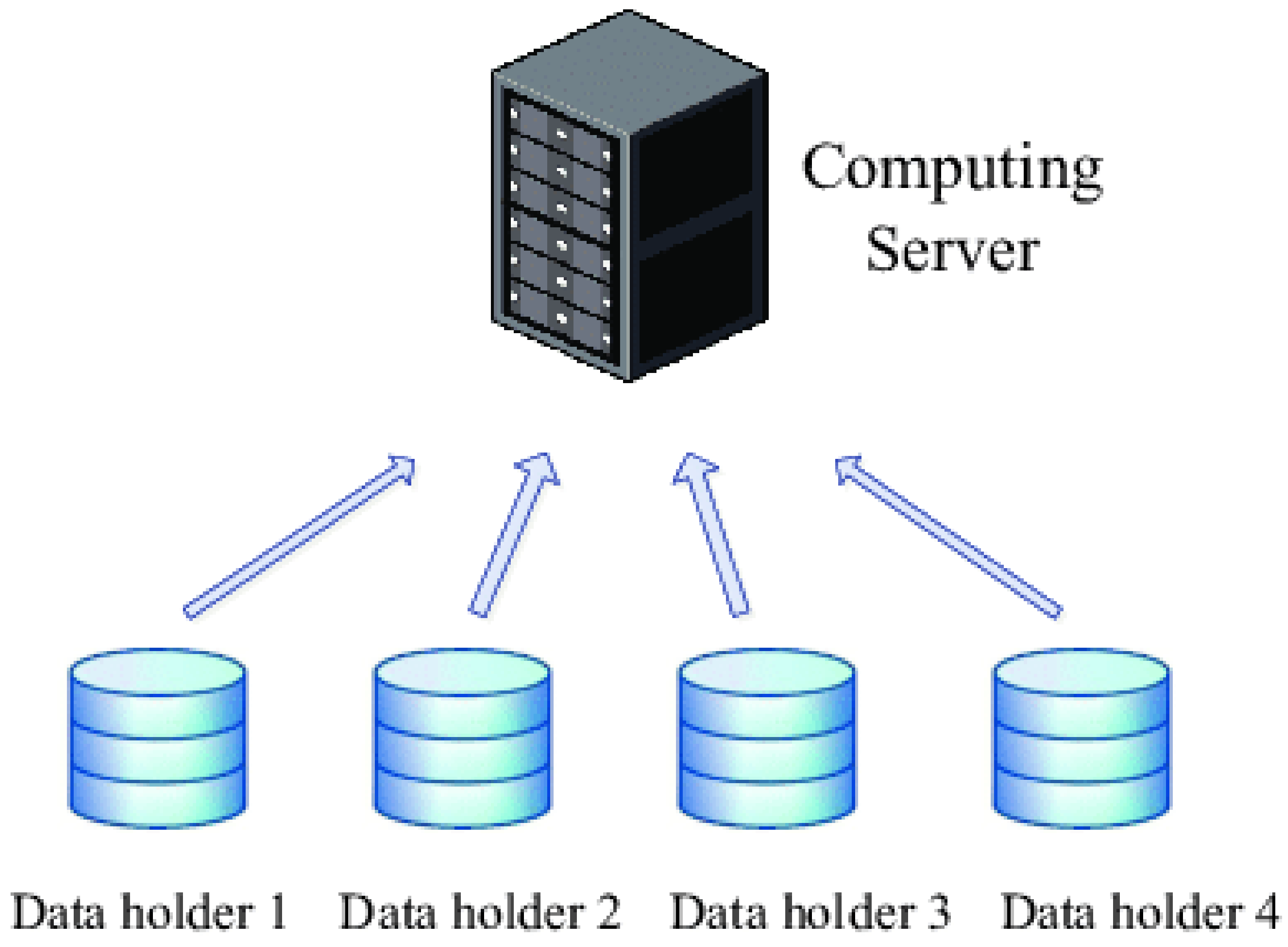


Exploiting edge computing

Centralized approach

- Data storage and model training on **server cloud**.
- Data must be transmitted from the edge to a **central data-center**, in order to train the model.

- Drawbacks:
  - **Communication costs**. We need a very high bandwidth to avoid **bottleneck**.
  - **Reliability**. Performance degradation and failures due to wireless connection.
  - **Security and privacy**. Data could be used inappropriately by companies or could be stolen in an attack to the centralized infrastructure.

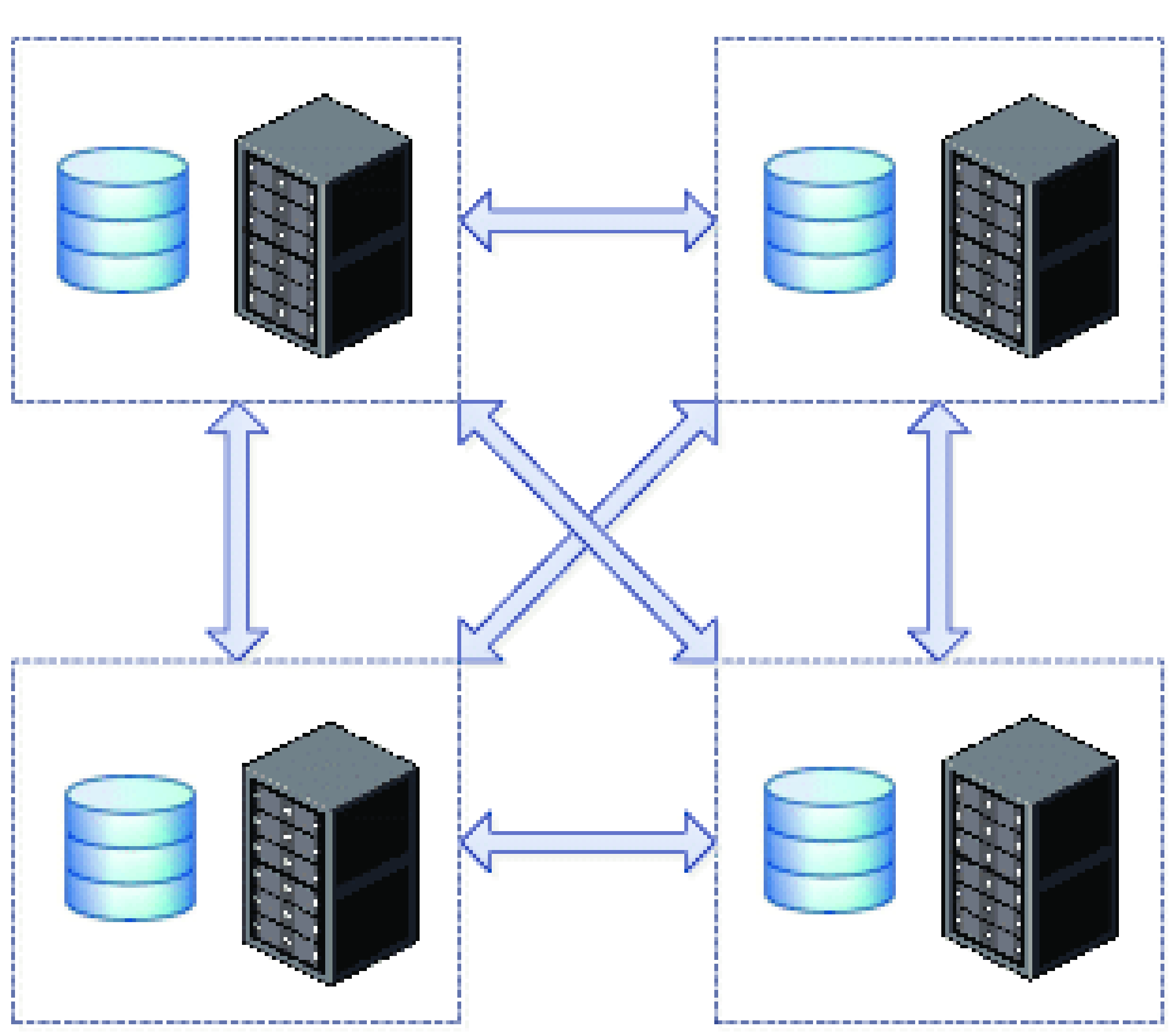


Distributed approach

- **Co-training** between different devices while maintaining **locality**.
- Edge devices do not share their local dataset, **only the model training parameters**.

- Advantages:
  - We use local data, which are not frequently sent to a remote server.
    - **Reduced latency**.
    - **Reduced bandwidth required**.
    - **Improved reliability**.
    - **Improved privacy and security**.

- **Very suitable for edge computing**.



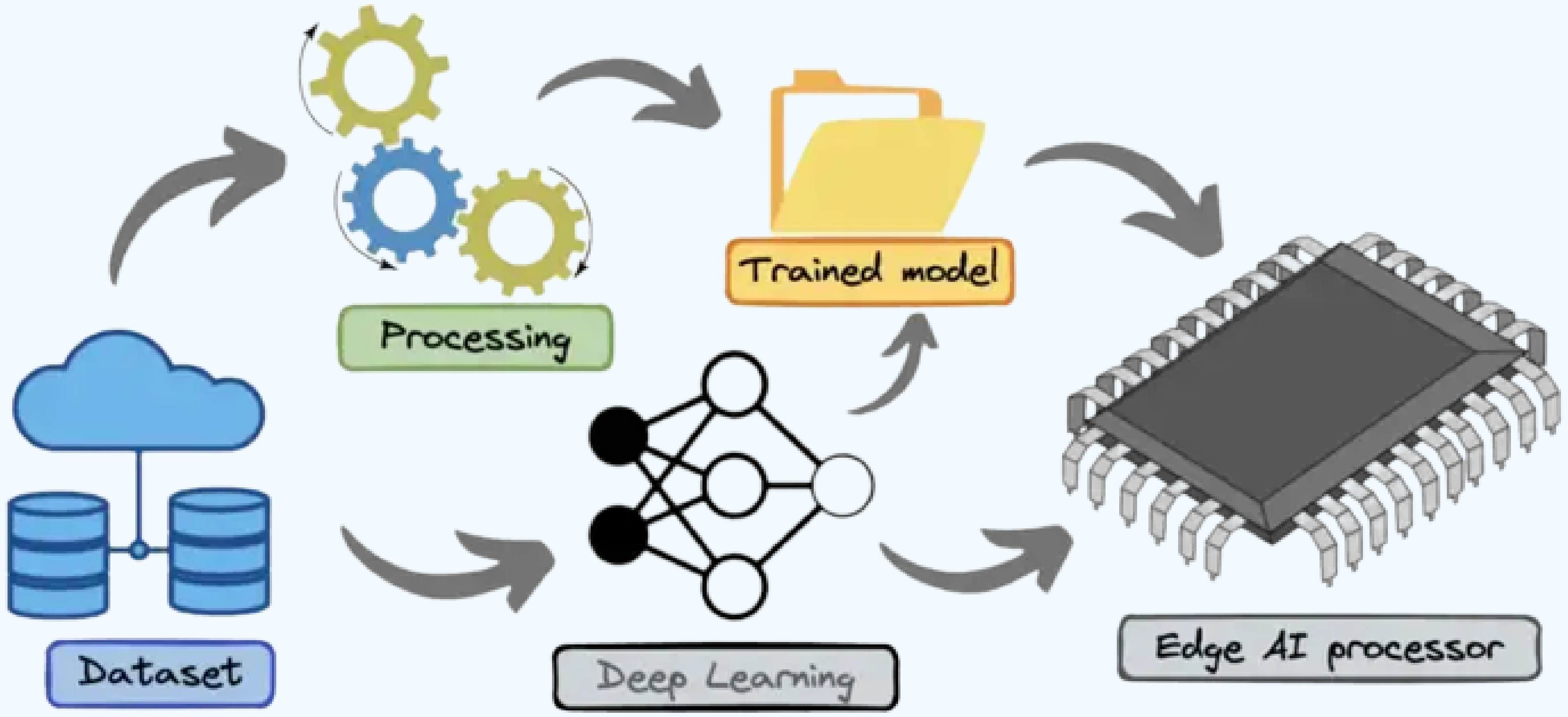
FL steps

1. The central/edge server **determines an ML model** to be trained locally.
2. A **subset of clients** is chosen (randomly or through algorithms such as **active learning**, wherein clients are selected based on the uncertainty of their predictions).
3. The server **multicasts** the model to selected clients, who **download the global parameters and train the model locally**.
4. Each client sends its own **local model updates** to the server.
5. The server receives the updates and **aggregates them** using aggregation algorithms (e.g. **FedAvg**, which averages model updates across all clients) to generate a **new global model**.

The server never accessed client data!

WHAT ABOUT DATA SCIENCE

- **Image classification**.
  - Self-driving cars.
  - Surveillance cameras.
- **Object detection**.
  - Tracking people or vehicles.
- **Natural Language Processing**.
  - Improving text-to-speech application.
  - Improving voice commands processing.



REAL-TIME!

THE FUTURE OF EDGE COMPUTING

- **Increased use of edge computing in IoT applications**, as more and more devices become connected to the internet.
  - **Real-time data analysis and decision making** in several areas (e.g. healthcare).
- **Advancements in edge hardware**.
  - More powerful edge devices and larger batteries.
- **Development of new edge computing platforms**.
  - Easier implementation of edge computing solutions for businesses and organizations.
- **Improved data security**, as edge computing and ML become more popular.
  - Better security features and ML algorithms to detect and prevent security breaches.

Thanks for your attention

REFERENCES

- X. Wang, L. T. Yang, X. Xie, J. Jin and M. J. Deen, "**A Cloud-Edge Computing Framework for Cyber-Physical-Social Services**" in IEEE Communications Magazine, vol. 55, no. 11, pp. 80-85 2017.
- J. Ren, Y. Pan, A. Goscinski and R. A. Beyah, "**Edge Computing for the Internet of Things**" in IEEE Network, vol. 32, no. 1, pp. 6-7, 2018.
- J. Pan and J. McElhannon, "**Future Edge Cloud and Edge Computing for Internet of Things Applications**" in IEEE Internet of Things Journal, vol. 5, no. 1, pp. 439-449, 2018.
- Shabnam Kumari, Surender Singh, Radha , " **Fog Computing: Characteristics and Challenges**" , International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) , Volume 6, Issue 2, March - April 2017 , pp. 113-117.
- Xu, Z., Zhang, Y., Li, H. et al. **Dynamic resource provisioning for cyber-physical systems in cloud-fog-edge computing**. J Cloud Comp 9, 32 (2020).
- Dante D. Sánchez-Gallegos, Diana Carrizales-Espinoza, Hugo G. Reyes-Anastacio, J.L. Gonzalez-Compean, Jesus Carretero, Miguel Morales-Sandoval, Alejandro Galaviz-Mosqueda, **From the edge to the cloud: A continuous delivery and preparation model for processing big IoT data, Simulation Modelling Practice and Theory**, vol. 105, 2020, pp. 102-136.
- Sulieman, N.A.; Ricciardi Celsi, L.; Li, W.; Zomaya, A.; Villari, M. **Edge-Oriented Computing: A Survey on Research and Use Cases**. Energies 2022, 15, 452.
- Abreha, H.G.; Hayajneh, M.; Serhani, M.A. **Federated Learning in Edge Computing: A Systematic Survey**. Sensors 2022, 22, 450.

REFERENCES

- https://blogs.nvidia.com/blog/2022/01/05/difference-between-cloud-and-edge-computing/
- https://towardsdatascience.com/you-need-to-move-from-cloud-computing-to-edge-computing-now-e8759eb9690f
- https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-edge-computing/#:~:text=With%20edge%20computing%2C%20businesses%20can,the%20location%20where%20it's%20collected
- https://www.slideshare.net/chetansk/edge-computing-109820603
- https://www.accenture.com/us-en/insights/cloud/edge-computing-index
- https://www.stackpath.com/edge-academy/what-is-edge-orchestration/
- https://cloudify.co/blog/birth-of-edge-orchestrator-cloudify/
- https://techleadcuriosity.com/edge-computing-and-machine-learning-unlocking-the-potential-of-real-time-data-analysis-b53876530d15
- https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication500-292.pdf

REFERENCES

- https://slideplayer.com/slide/13778757/
- https://www.volico.com/whats-driving-the-adoption-of-edge-computing-in-the-healthcare-sector/
- https://blog.bosch-si.com/bosch-iot-suite/cloud-and-edge-computing-for-iot-a-short-history/#:~:text=The origin of edge computing,such as images and videos.
- https://www.grandviewresearch.com/industry-analysis/edge-computing-market
- https://www.techtarget.com/searchdatacenter/definition/edge-computing
- https://www.techtarget.com/searchcloudcomputing/tip/Explore-edge-computing-- services-in-the-cloud
- https://www.techtarget.com/searchnetworking/answer/What-is-the-network-edge-and-how-is-it-different-from-edge-computing
- https://developer.ibm.com/articles/edge-computing-architecture-and-use-cases/
- https://www.ibm.com/cloud/architecture/architectures/edge-computing/