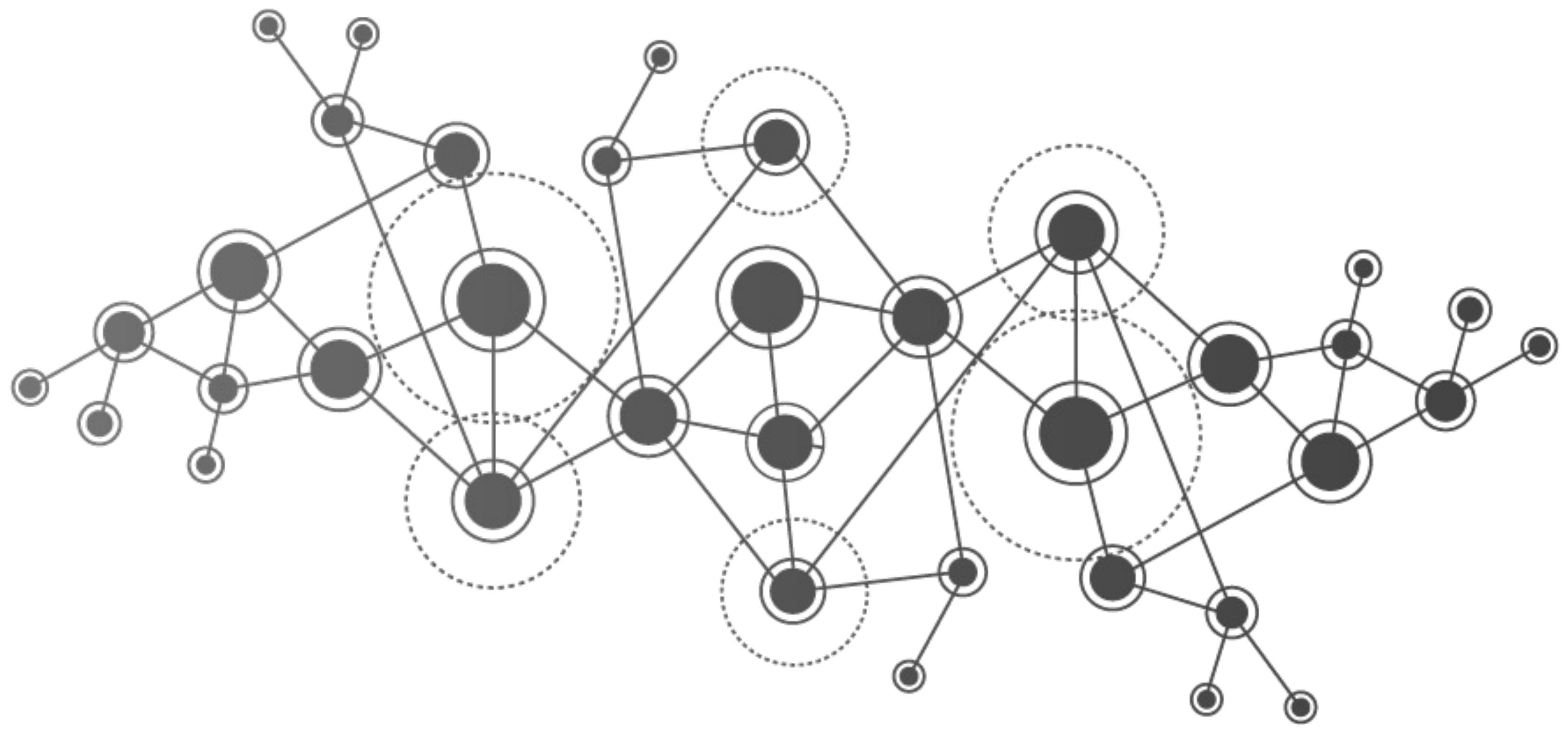
UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA



Intelligent Internet of Things

Edge Computing for the IoT

Prof. Marco Picone

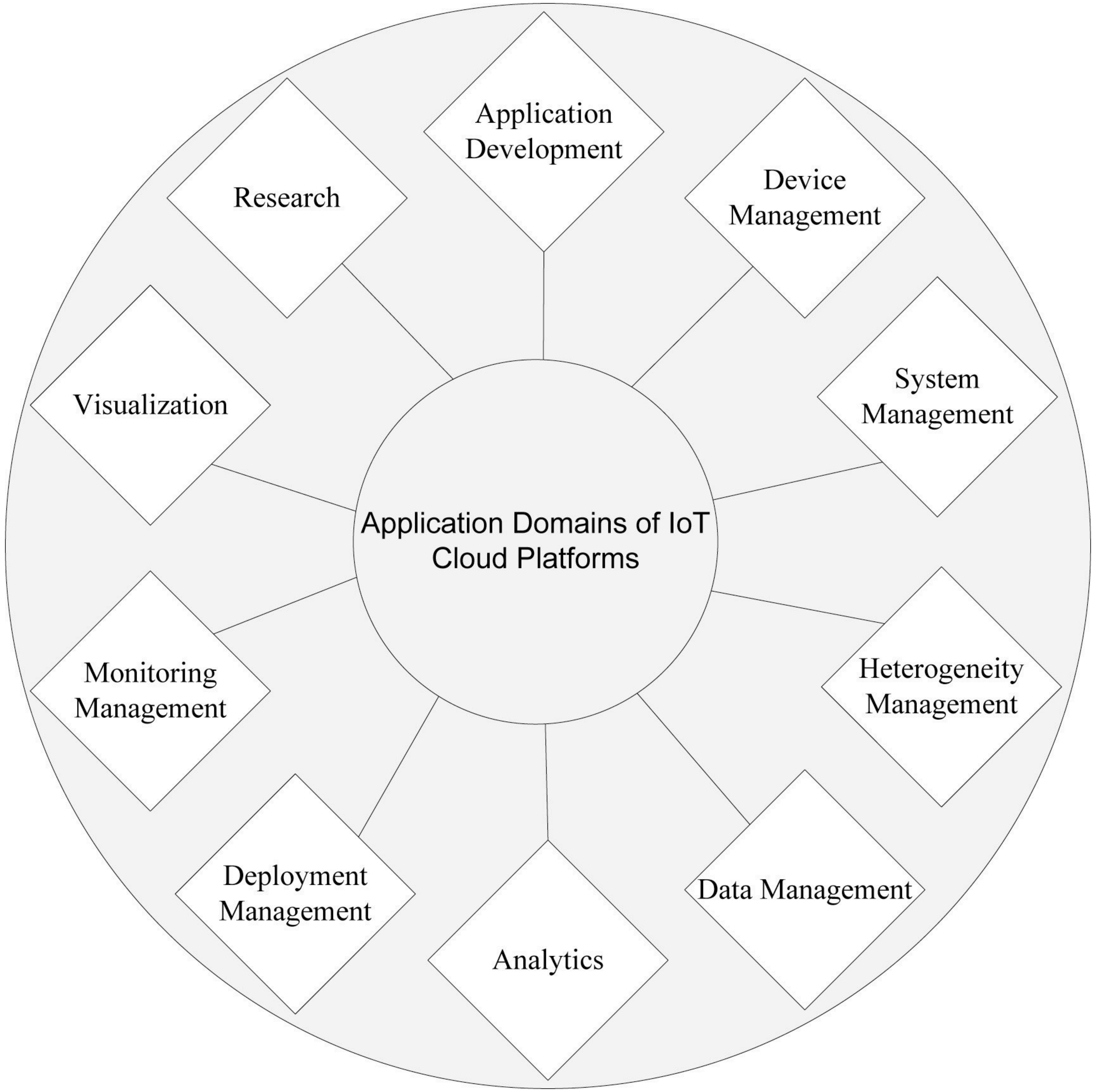
A.A 2023/2024

Edge Computing for the Internet of Things

- Cloud Computing & Internet of Things
- The Cloud “Addiction” Problem
- Edge/Fog Computing Vision
- Distributed IoT Applications
- IoT Hub & Gateways (Edge)
- IoT Hub -> Cloud & Edge Synchronization

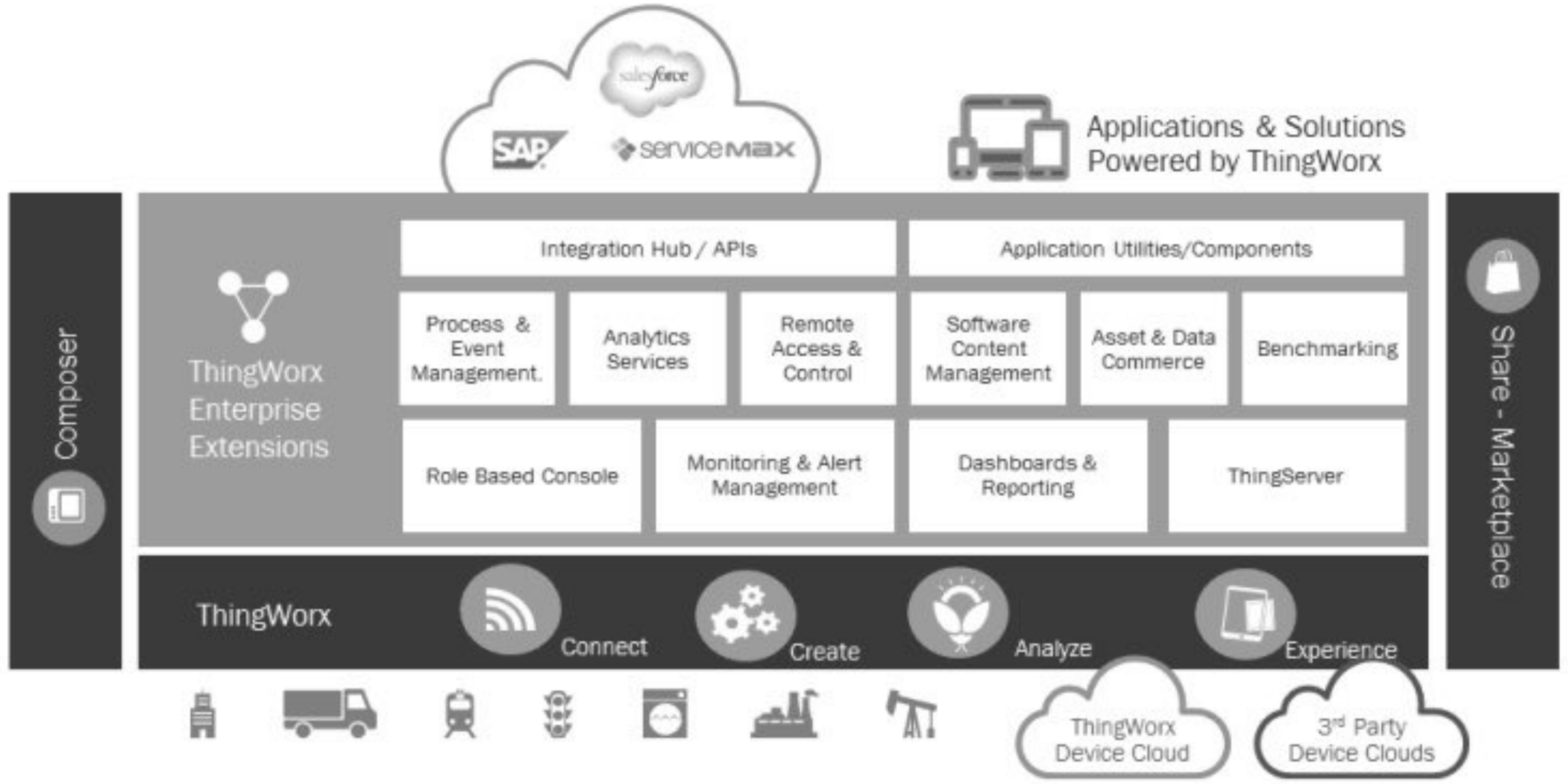
Cloud & Internet of Things

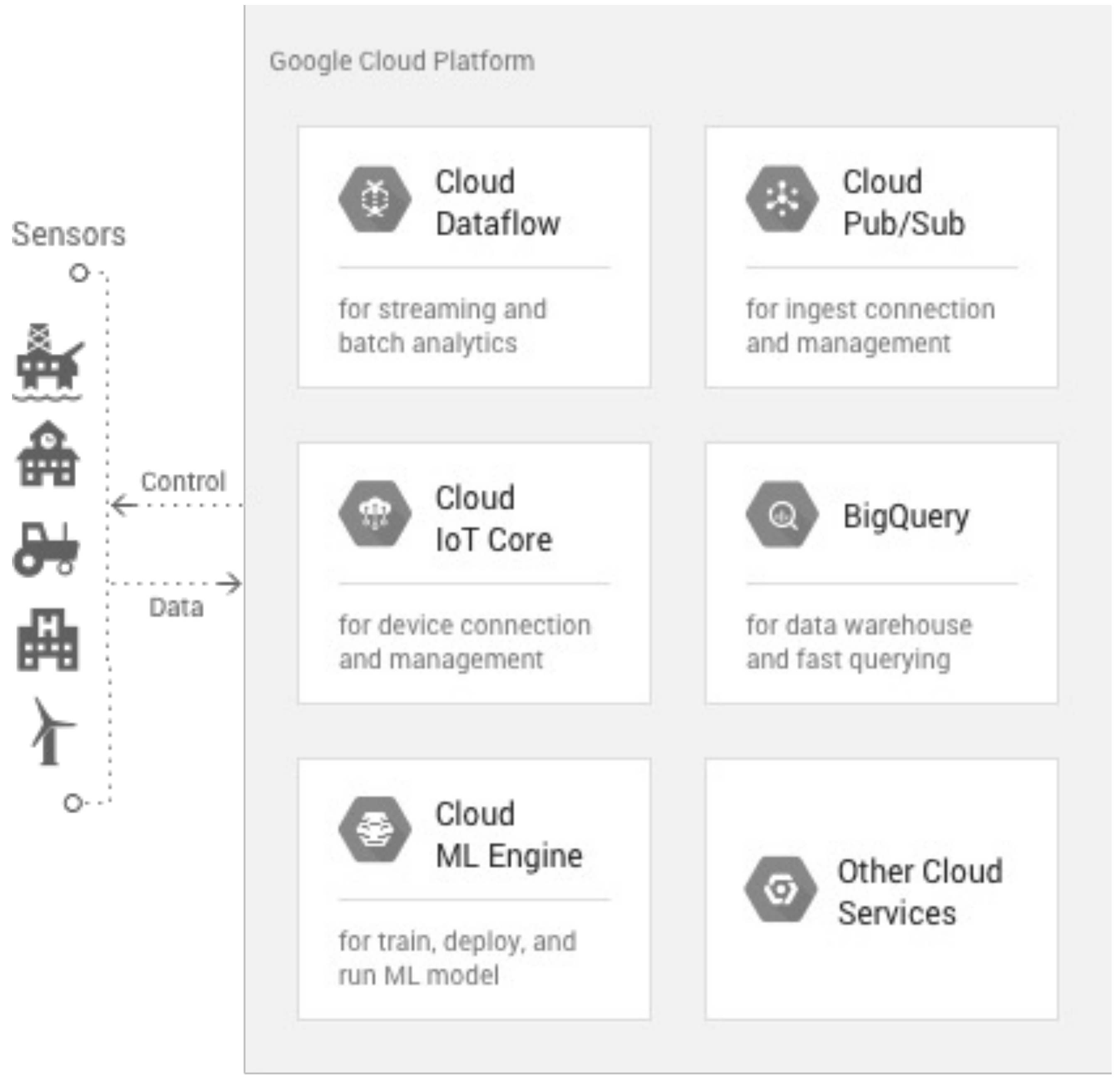
- Internet of Things (IoT) envisages overall merging of several “things” while utilizing internet as the backbone of the communication system to establish a smart interaction between people and surrounding objects
- Cloud, being the crucial component of IoT, provides valuable application specific services in many application domains
- A number of IoT cloud providers are currently emerging into the market to leverage suitable and specific IoT based services



Cloud & Internet of Things

- Main common IoT Cloud characteristics and features are:
  - Real-time event processing
  - Device management, inventory & registry
  - Secure communication
  - IoT Gateway & Hub for devices
  - Authentication and encryption
  - Device Clones / Replicas / Digital Twins
  - Intelligence & Cognitive Services
  - Huge Storage
  - Data Visualisation and Reporting
  - Multiple protocols support
  - Integration and third party services

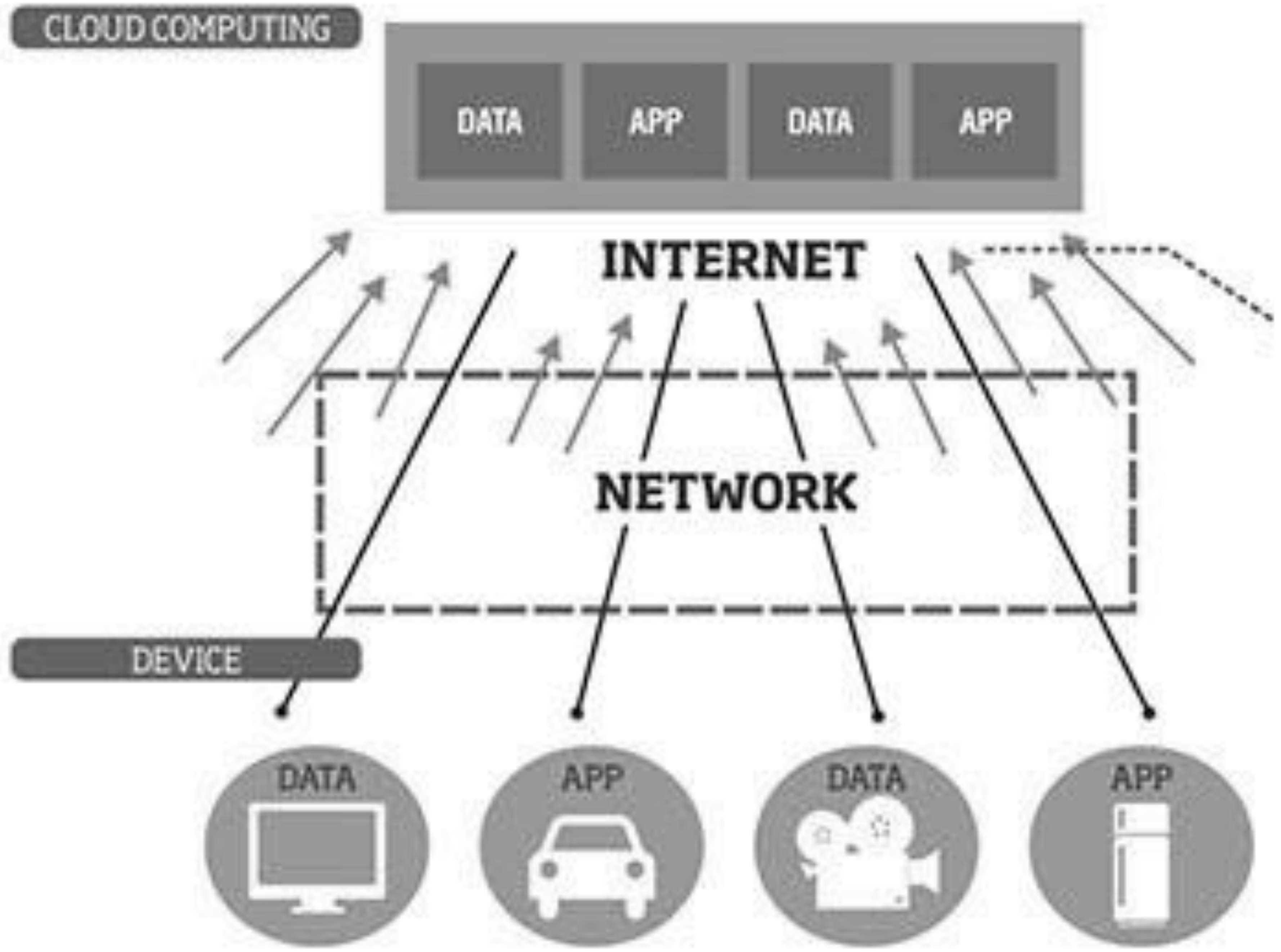




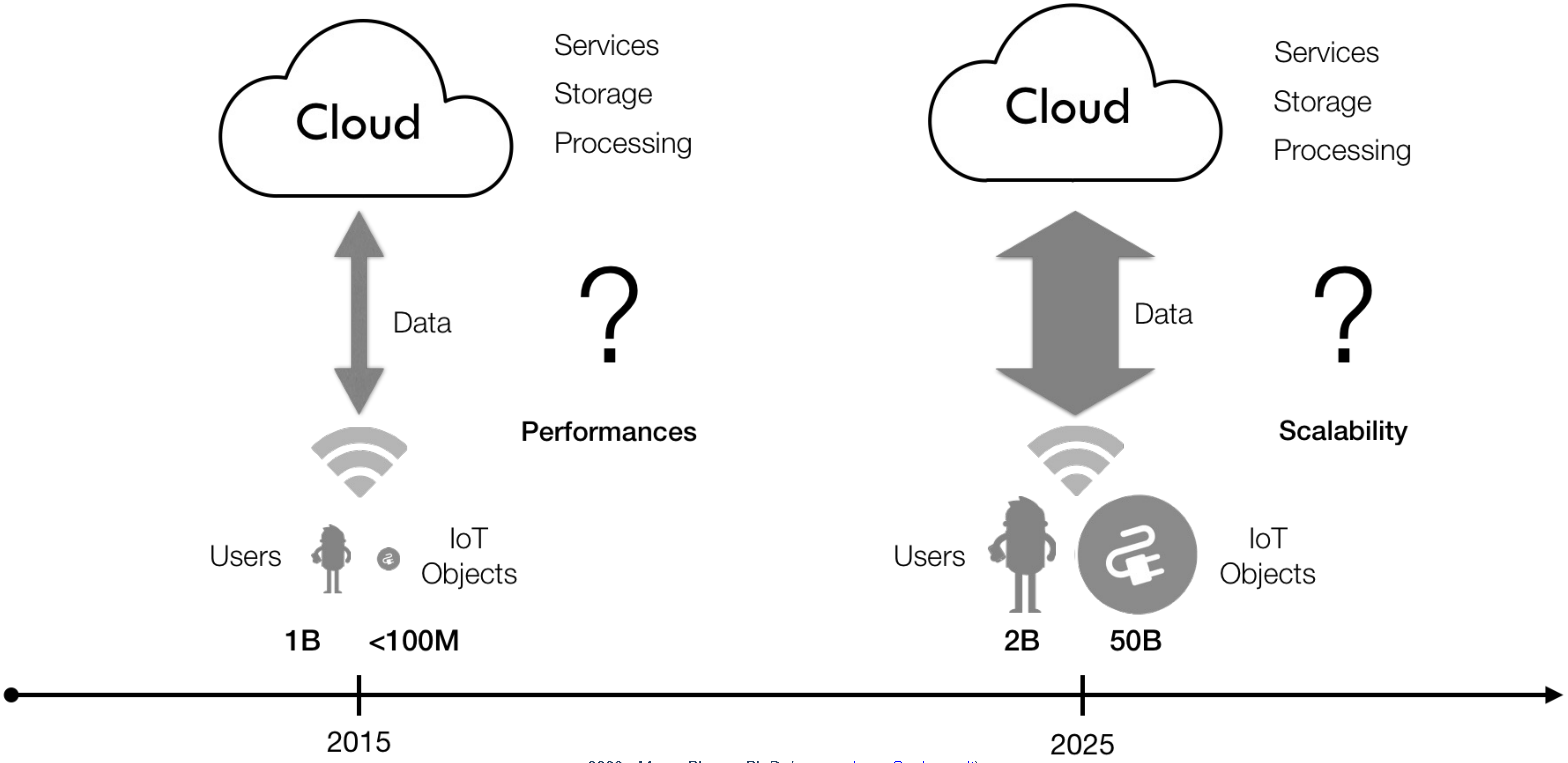
Build & train ML models in the cloud

Cloud & Internet of Things

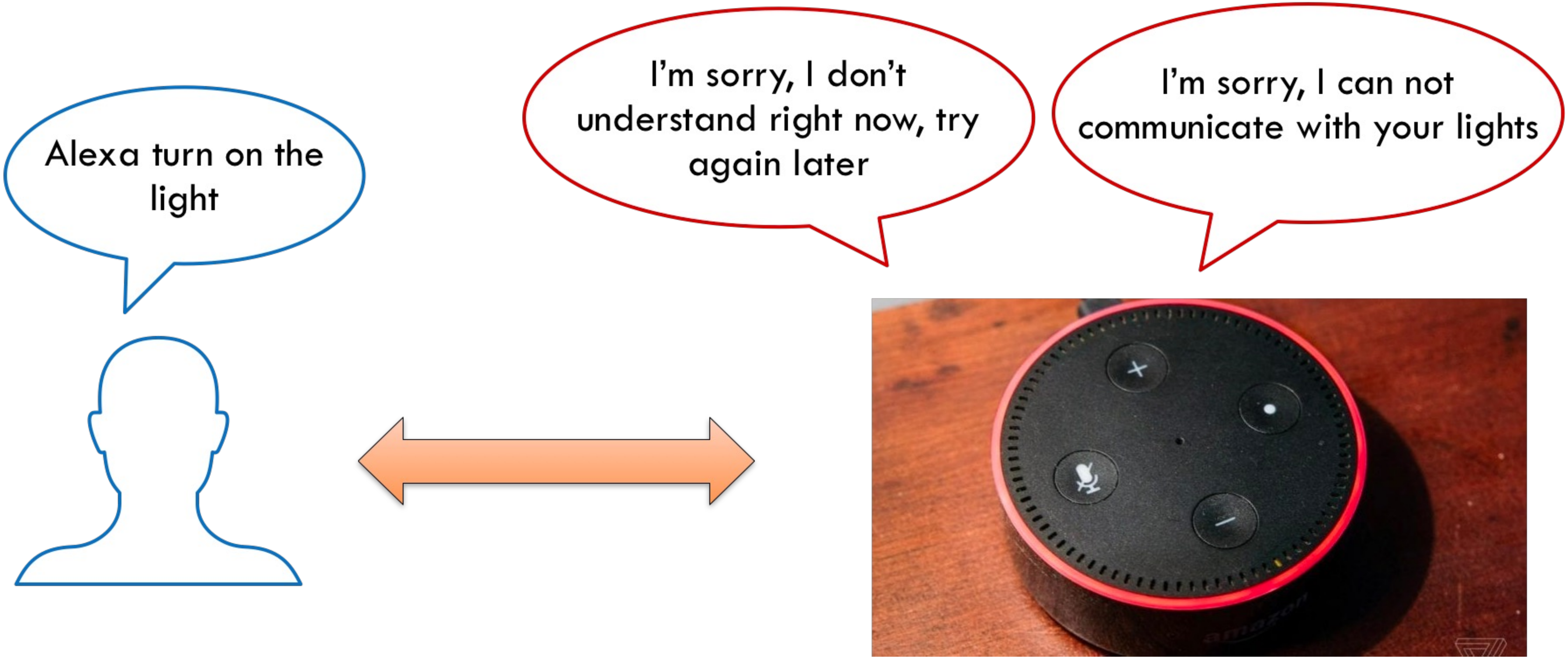
- Main common potential platform “open” issues are:
  - Hard to manage complex systems.
  - The limitation to install edge program on a custom platform.
  - Cost management & Control
  - **Strong bound with the vendor/provider**
  - Learning curve
- Is the Cloud always the best option to handle an IoT Project ? Is a Cloud-centric vision the only approach to design an IoT application ?



The Internet Architecture As We Know It



Cloud "Addition"



Connectivity Challenge


Got a tip? Let us know.

Follow Us

News Video Events CrunchBase

Message Us Search

STARTUP BATTLEFIELD Now Accepting Applications For Startup Battlefield At Disrupt SF Apply Today

Nest

Popular Posts

The Future of VR Storytelling
3 days ago

Atari is embracing the Internet of Things with new smart home devices
2 days ago

Artificial intelligence is changing SEO faster than you think
4 days ago

iContainers, Expedia for transporting freight internationally, raises $6.7 million
a day ago

Pebble's new Core pocket computer will ship with Alexa voice integration
5 days ago

Uber launches UberBike in

Nest's Smart Home Apps Are Back Online Following Outages

Posted Jan 9, 2016 by Lucas Matney (@lucas_matney)

858 SHARES

Nest users who depend on their connected products to monitor their homes experienced a hiccup in connectivity Saturday that left users scrambling to figure out what was happening.

In the absence of any initial details from Nest, many users flocked to Twitter to vent their frustrations with the company's products, and also reflect on the risks inherent to trusting connected devices with sensitive tasks like monitoring children.

CrunchBase

Nest Labs

FOUNDED
2010

OVERVIEW
Nest Labs is a home automation producer of programmable, self-learning, sensor-driven, Wi-Fi-enabled thermostats, smoke detectors, and other security systems. It introduced the Nest Learning Thermostat in 2011 as its first product. The Nest Protect smoke and carbon monoxide detector was then introduced in October 2013. After acquiring Dropcam, the rebranded Nest Cam was introduced in June 2015.

LOCATION
Palo Alto, CA

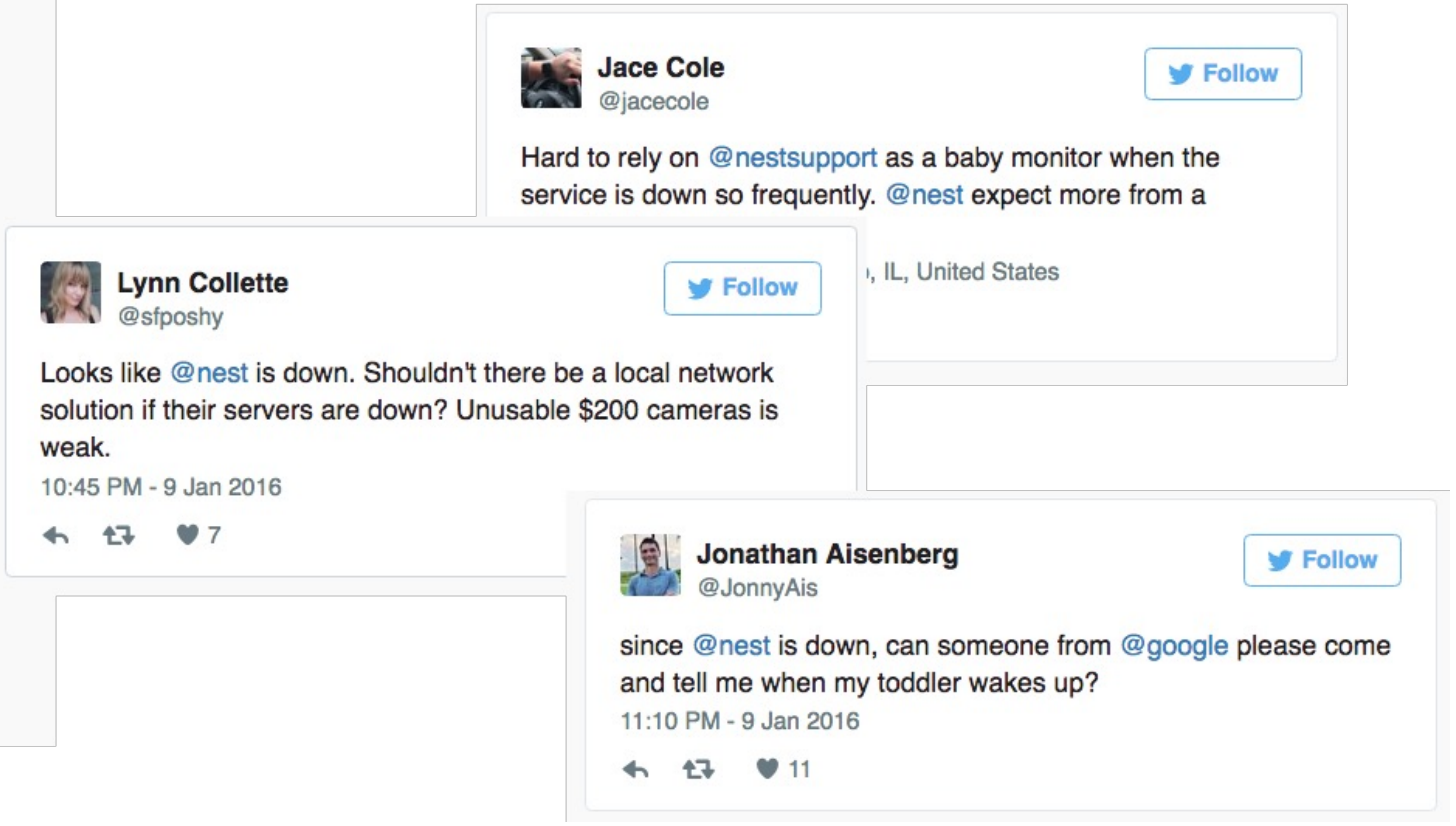
CATEGORIES
Home Automation

FOUNDERS
Tony Fadell, Matt Rogers

WEBSITE
http://www.nest.com

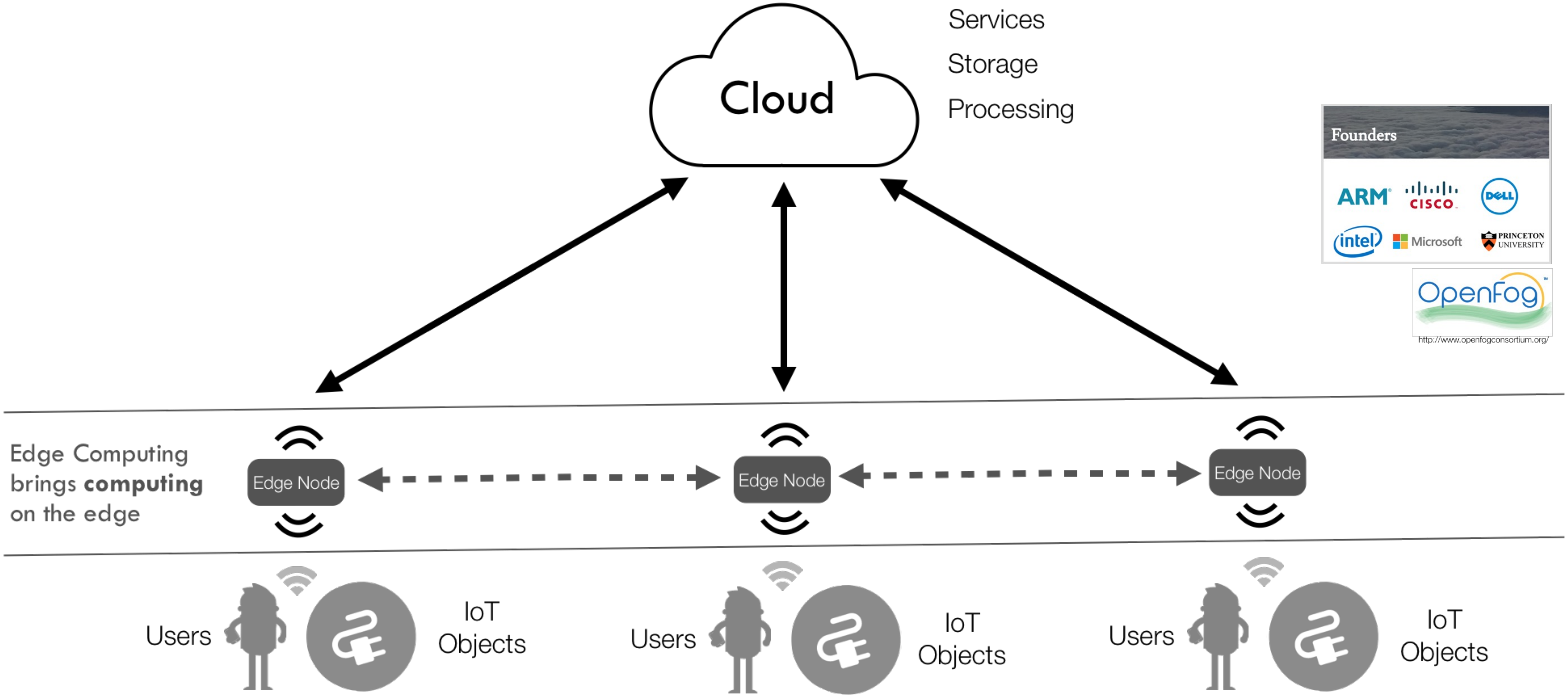
Full profile for Nest Labs

"Nest users who depend on their connected products to monitor their homes experienced a hiccup in connectivity Saturday that left users scrambling to figure out what was happening."



http://techcrunch.com/2016/01/09/nests-smart-home-apps-are-back-online-following-outages/

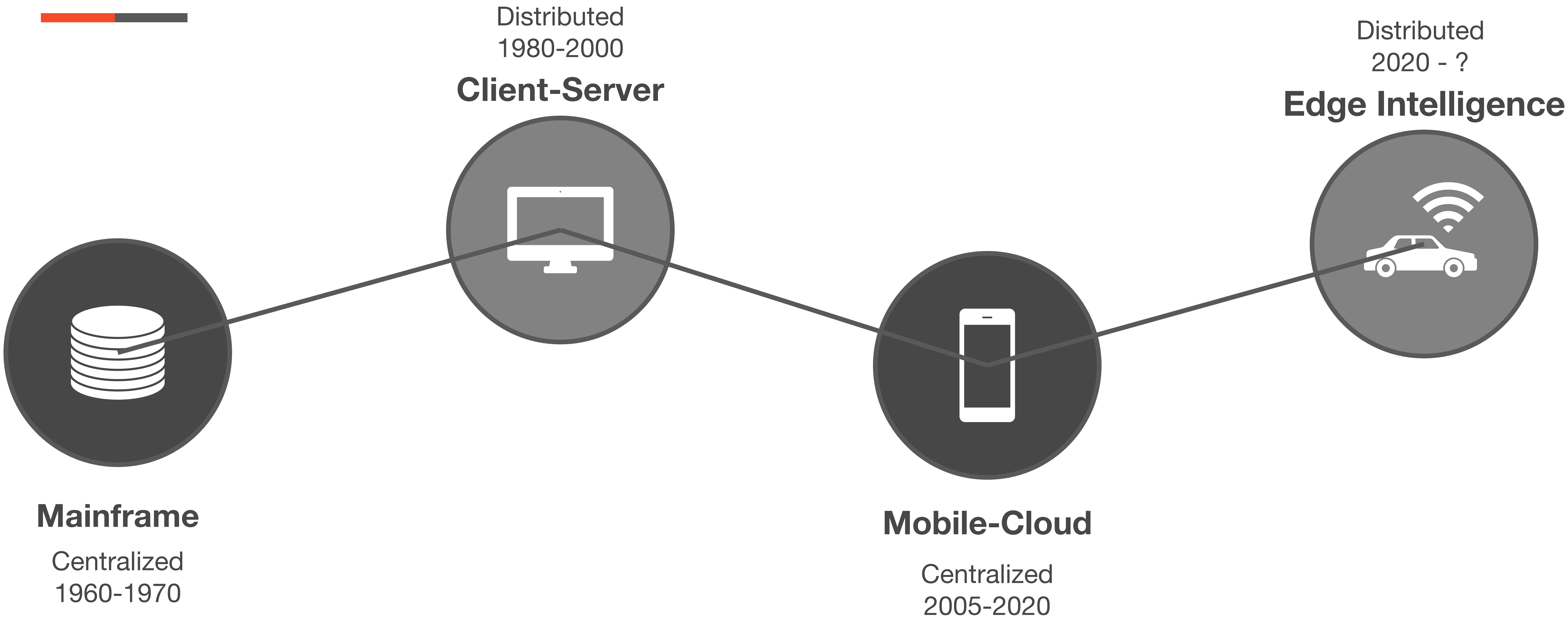
Envisioned Edge/Fog Architecture



Founders

http://www.openfogconsortium.org/

Cycles of Computing



45% of IoT created data will be stored and processed, analyzed and acted at the edge by 2019

5.6 Billion of IoT devices will be connected at the Edge by 2020

IDC FutureScape WorldWide Internet of Things Predictions, 2016

Shared Central vs Local Computation

- The computing models of the last seven decades are oscillating **between the use of shared and central resources or exclusive and local compute power**. Key factors in deciding the direction of the curve are advances in computing and communication.
  - Cheap and powerful compute power pushes towards the use of local resources.
  - Cheap and fast communication technologies enable the use of shared, central resources.
- Networked PCs and client-server models created balanced distributed computing model between the two extremes of Mainframes and standalone computer enabled by high-speed local area networks.

The Cloud Centric Vision

- The technology trend moved towards **central resources with the early web model** enabled by **cheap and fast wide area networking, compared to the level of data being transferred**.
- Shared and **centralized resources are highly efficient**, as they maximize the utilization of compute resources and provide elasticity and **scalability**.
- Given their central location, typically in data centres, they can be more easily **secured**, and their **lifecycle management is less complex** than in distributed Systems.
- However, **they need highly available communication channels of sufficient bandwidth** and speed to reach end users, which may incur significant cost.
- The cloud computing model seems to have found a happy compromise in the distributed computing spectrum, balancing the pros and cons between exclusive/local versus shared/central, but this solution will not last.

IoT Disrupts the Cloud

- The IoT disrupts the cloud compute model by introducing new usage scenarios resulting in the following key requirements:
  - **Real-time:** often decisions need to be made within tens of milliseconds. Today's communication infrastructure and the laws of physics require local decision-making, as a roundtrip to the cloud would take an excessive amount of time.
  - **Connectivity:** today's mobile networks are often spotty and cannot guarantee connectivity to the cloud. Hence decision making must occur locally
  - **Data volume:** the amount of data generated by sensors can be huge, for example hundreds of high-resolution cameras creating video streams at 30 frames per second, which could block wide-area communication channels.
- **The disruption of the cloud model is not the replacement of the Cloud but rather its extension to the edge**

A new Computational Model - Edge Computing

- The Cloud will continue to exist. For example, certain functions are best performed in the cloud, such as the training of predictive analytics algorithms, as typically only the cloud holds the necessary data
- **Edge Computing** will provide compute power and storage in the space between the device and the cloud. Edge compute devices include:
  - IoT Gateways
  - Routers
  - Micro data centres on the Edge and on mobile network base stations
- The new model will be a **fully distributed** computing model. It will support a wide range of interaction and communication paradigms

A Distributed Point of View

- **Autonomous, local decision-making** based on incoming IoT data and cached enterprise information
- **Peer-to-peer networking**, for example security cameras communicating amongst themselves about an object within their scope
- **Edge networking**, for example platoon driving, i.e. vehicles self-organizing into groups which travel together, orchestrated and controlled by a micro data centre in the base station of a mobile network
- **Distributed queries** across data that is stored in devices, in the cloud and anywhere in between
- **Distributed data management**, for example data aging: which data to store, where and for how long
- Self-learning algorithms that **learn and execute on the edge**, or **learn in the cloud and execute on the edge**, or **learn and execute in the cloud**
- Isolation, involving devices which are disconnected for a long time, operating on minimal energy consumption to maximize lifespan

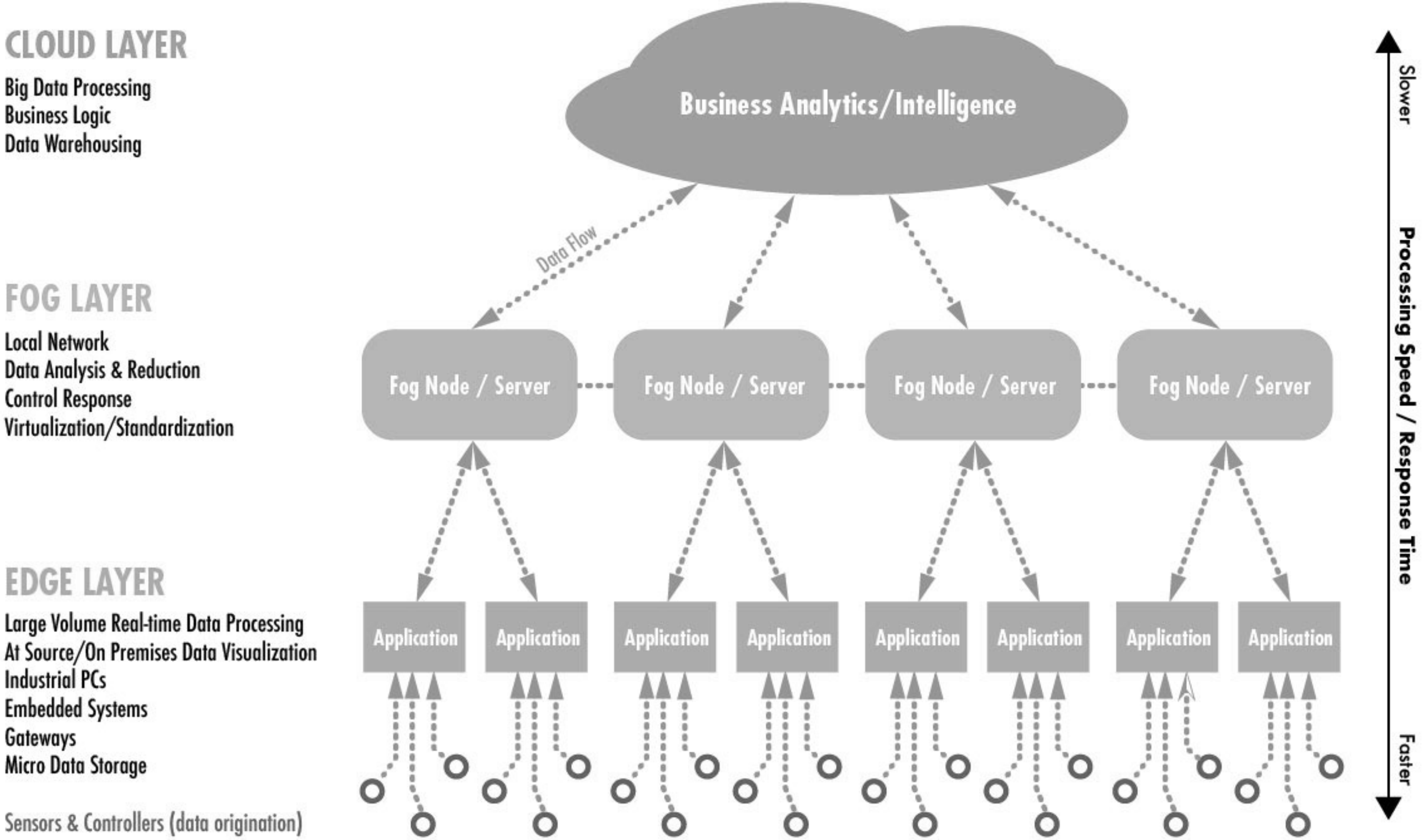
Intelligence at the Edge

- Through the introduction of intelligence at the edge nodes, systems can:
  - **take decisions more quickly and efficiently**, as the round trip delay in contacting the cloud is removed;
  - reach decisions according to local identity management and access control policies, **securing the data close to its source**;
  - **reduce communication costs** by limiting communication over public wide area networks
- This huge new opportunities come from technology evolutions in several fields:
  - Manufactured devices and 5G networks,
  - Algorithms and new standards in software-defined networking,
  - Mobile Edge Computing - Multi-access Edge Computing (MEC Architecture)
  - IoT evolution and standard approaches
  - Analytics and device and data ownership

Edge/Fog Computing

- Edge/Fog Computing brings computation, storage, and networking resources at the edge of the access network in order to enable:
  - real-time data processing
  - minimal latency
  - best quality of service/experience
  - independence from Internet connectivity for operations
  - flexible networking, security, and privacy
- The novel Edge/Fog Computing paradigm is the reference architecture for the IoT and will involve several actors and technologies for its innovation, implementation, and deployments

Edge vs Fog Computing

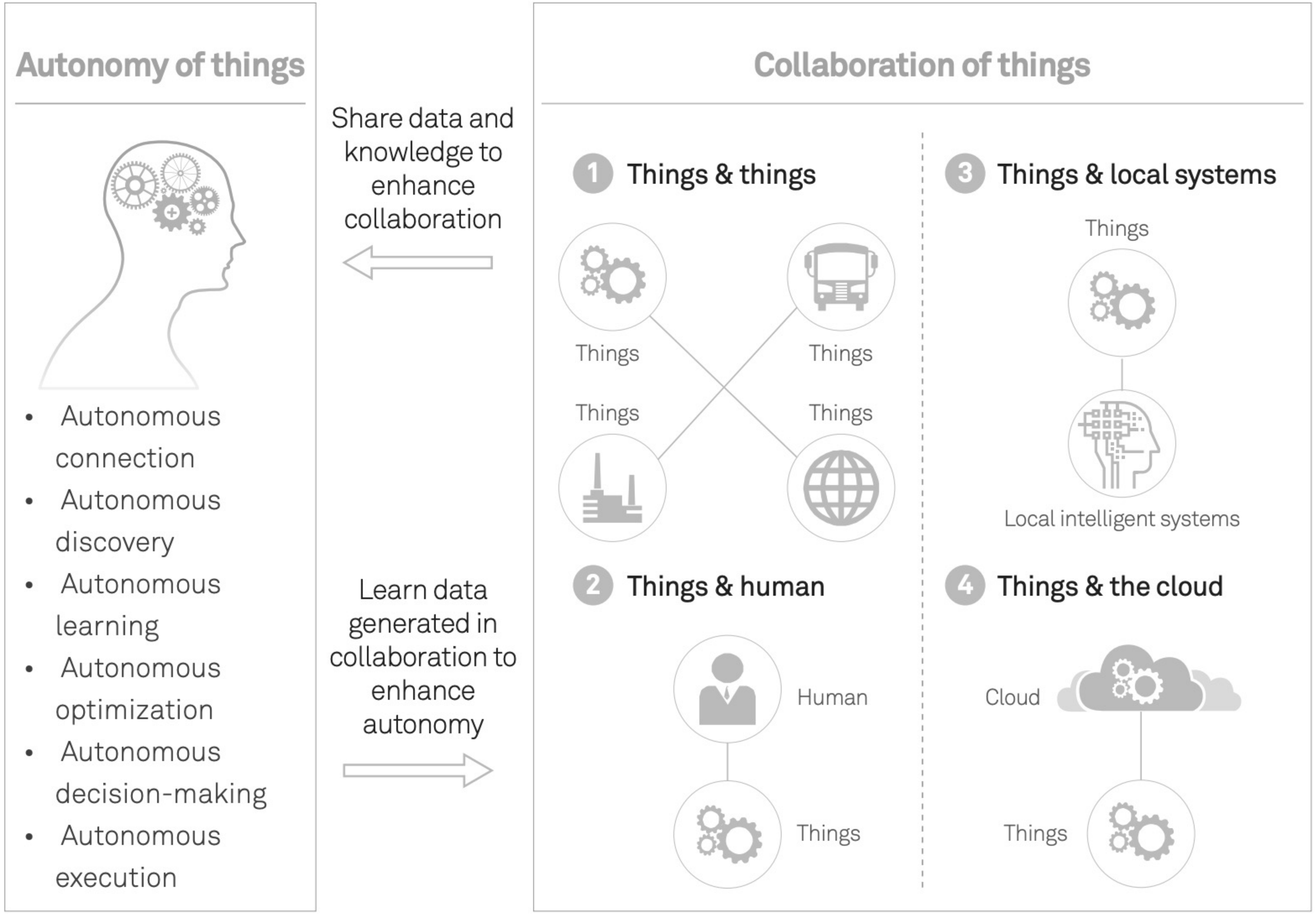


https://www.winsystems.com/cloud-fog-and-edge-computing-whats-the-difference/

Edge vs Fog Computing

- **Fog Computing** and **Edge Computing** look and are really similar for several aspects since they both involve bringing intelligence, storage and processing resources closer to the creation of data
- The key difference between the two lies in **where the location of intelligence and compute power is placed**
- **Fog Computing** environment places intelligence at the Local Area Network (LAN). This architecture transmits data from endpoints to a gateway, where it is then transmitted to sources for processing and return transmission
- **Edge Computing** places intelligence and processing power **in devices** such as embedded automation controllers and Smart Objects
- Edge computing offers many advantages over traditional architectures such as optimizing resource usage in a cloud-computing system. **Performing computations at the edge of the network reduces network traffic, which reduces the risk of a data bottleneck**. Edge computing also **improves security** by encrypting data closer to the network core, while optimizing data that's further from the core for performance
- *In this lecture we will use the term Edge Computing associated to the idea of "Working on the Edge" and considering both layers associated to Edge and Fog Computing*

Intelligent Distributed Architecture



- An intelligent distributed Edge Architecture drives intelligence capabilities to enable autonomy and collaboration of things, people and services:
  - **Autonomy of things**: Things can perform autonomous connection, discovery, learning, optimization, decision-making, and execution.
  - **Collaboration of things**: including collaboration between things, between things and human, between things and local smart systems, and between things and the cloud

Edge Computing - Characteristics & Attributes

- **Connectivity**
    - Connectivity is the basis of Edge Computing. The diversity of connected physical objects and application scenarios requires that the Edge layer provides connection functionalities through various network interfaces, protocols, topologies, network deployment and configuration, and network management and maintenance.
    - Connectivity needs to fully draw on the advanced research achievements in the network field, such as TSN, SDN, NFV, NaaS, WLAN, NB-IoT, and 5G. Additionally, connectivity needs to consider interoperability with a variety of existing industrial buses (in particular in IIoT Scenarios)
- **First Entry of Data**
    - As a bridge between the physical and digital worlds, **Edge Computing is the first entry point for data**. With mass, real-time, and complete data, edge computing implements data management and creates values based on the data lifecycle, supporting innovative applications such as **predictive maintenance**, **asset efficiency** and **management**. In addition, as the first entry of data, the Edge layer also faces the challenges caused by **real-time, determinacy, and heterogeneity and diversity**

Edge Computing - Characteristics & Attributes

- **Constraint**
    - Edge computing products need to adapt to work in several and heterogenous conditions and operating environments such as industrial sites (with interferences, dust, vibration, current/voltage fluctuations)
    - Specific challenging scenarios will require constrained power consumption, cost, and space limitation for edge devices.
    - In critical environment Edge computing products need to be integrated and optimized through hardware and software to adapt to various conditions and constraints and support diverse scenarios of industry digitalization
- **Distribution and Scalability**
    - In actual deployment, edge computing needs to support **distributed computing and storage**, achieve dynamic scheduling and unified management of distributed resources, support distributed intelligence, autonomous scalability and deliver distributed security capabilities.
    - Software innovations and Cloud Computing technologies such as **Virtualization**, **Containers** and **Lambda Functions are migrating from the Cloud to the Edge in order to support scalability and reliability of Edge Software Architectures**

Edge Computing - Characteristics & Attributes

- **Collaboration and Interoperability**
  - **Collaboration** and **Interoperability** are important foundation for the digital transformation of industries and applications on the Edge. Edge computing must support collaboration in connection, data, management, control, application, and security.
  - Edge Applications must be based on **standard communication protocols and mechanisms**. All the system components provide a highly modular design, which defines a set of uniform interfaces that can be used to be easily integrated into new or existing business applications. In particular:
    - thanks to the standardization module, RESTful CoAP and HTTP interfaces are defined for interactions with smart objects and legacy components;
    - RESTful HTTP API and MQTT/AMQP communications can be used to provide access to computed analysis and real- time metrics by external applications and Cloud Services

Values of Edge Computing

- **Mass and Heterogeneous Connection**
  - Networks are the cornerstone of system interconnection and data aggregation transmission. With the surge in the number of connected devices, networks face enormous challenges in terms of Operations and Maintenance, management, flexible expansion, and reliability.
  - The complexity of handle a large number of heterogeneous connections must be solved in an intelligent and efficient way at the Edge in order ensure real-time reliability of connections and interoperability of services both at the Edge and with the Cloud
- **Real-Time Services**
  - Industrial and Low-Latency Applications (e.g., Vehicular and Mobility service) have high real-time requirements, even within 10 milliseconds in some scenarios. If data analysis and control logic is implemented only on the Cloud, it is difficult to meet the real-time requirements of services

Values of Edge Computing

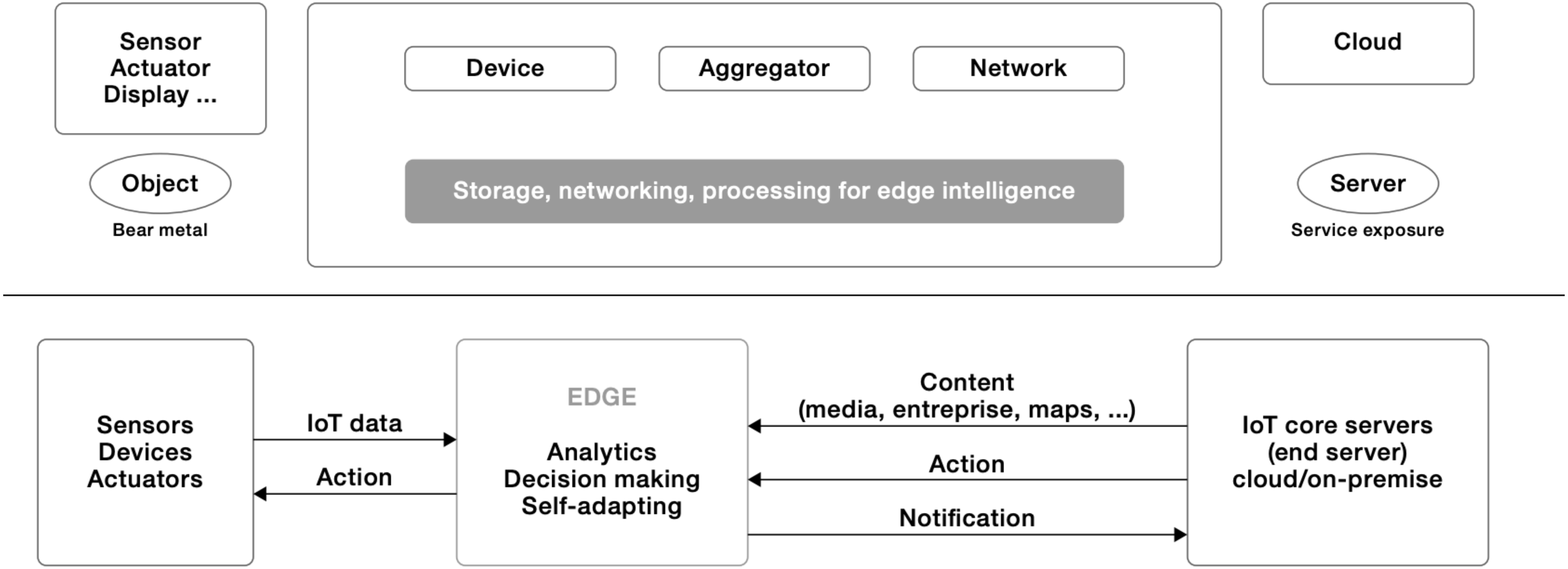
- **Data Optimization**
  - Today, almost every deployed application contain a huge amount of heterogeneous data from several sources. Data optimization must be implemented for data aggregation and unified presentation and openness, so that the data can serve intelligent Edge applications in a flexible and efficient manner.
- **Smart Application**
  - Business process optimization, Operation and Maintenance (O&M) automation, and service innovation drive applications to be smart. **Edge intelligence delivers significant efficiency and cost advantages**. Intelligent applications represented by predictive maintenance are driving industries to transition to new service models and business models.
- **Security and Privacy Protection**
  - Security is critical to cloud and edge computing, requiring end-to-end protection. The network edge is close to Internet of Things (IoT) devices, making access control and threat protection difficult. Edge security includes device security, network security, data security, and application security. The integrity and confidentiality of key data, as well as protection of mass production or personal data are also key areas of focus for security.

Collaboration between Edge and Cloud

- Cloud computing is suitable for non-real-time low latency and critical tasks, long-period data and business decision-making scenarios
- Edge computing plays an strategic role in scenarios such as real-time, low latency, short-period data and local decision-making potentially also in case of offline context.
- Edge and cloud computing are two important foundations for the digital transformation of industries. The collaboration between them in respect of network, service, application, and intelligence will represent a fundamental pillar to support more scenarios and unleash greater value in industry digitalization and IoT spreading

Point of Collaboration	Edge Computing	Cloud Computing
Network	Data aggregation (TSN + OPCUA)	Data analysis
Service	Agent	Service orchestration
Application	Micro applications	Lifecycle management of applications
Intelligence	Distributed reasoning	Centralized training

IoT & Edge Architecture



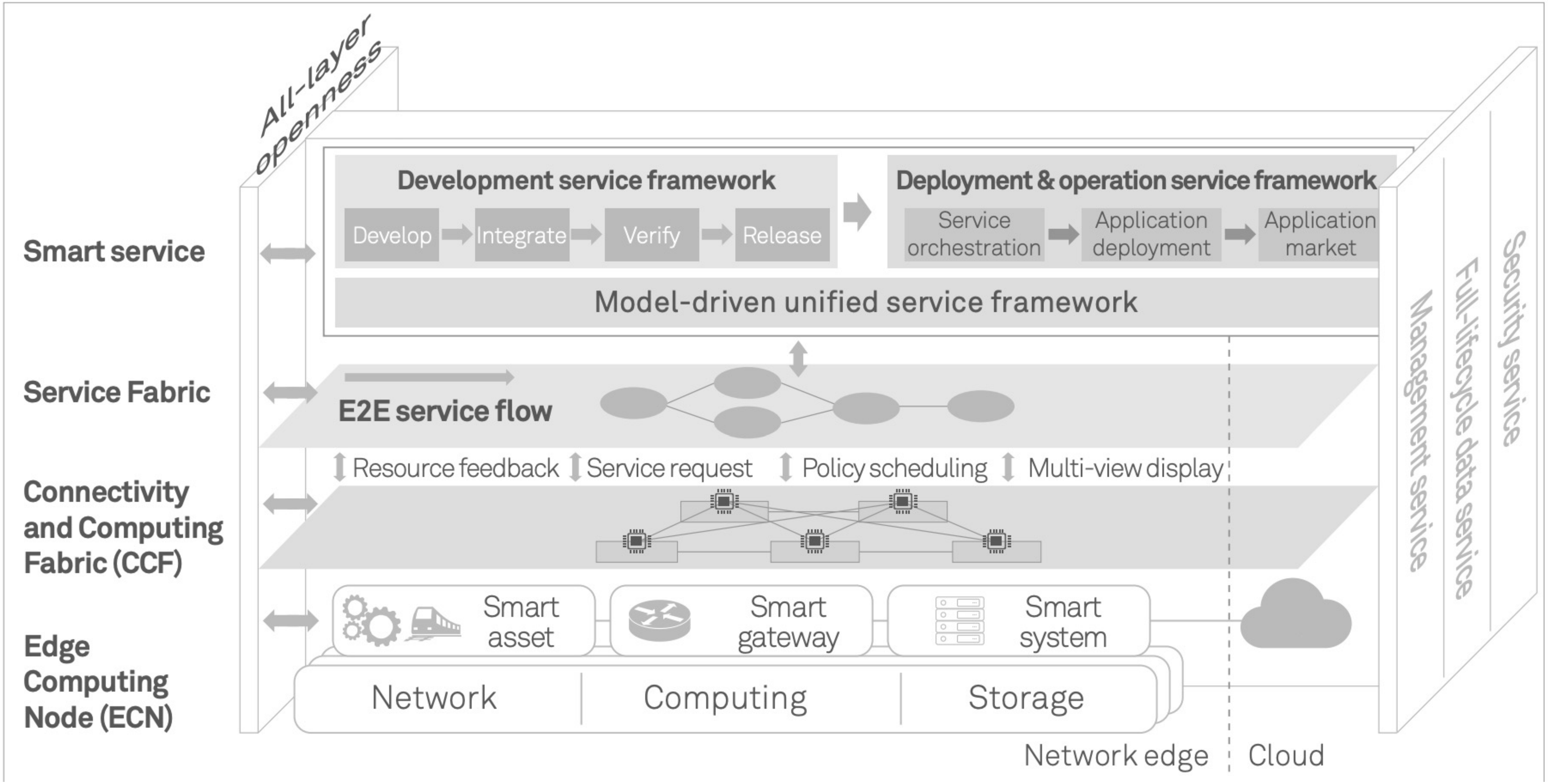
Working on the Edge

- The aim to define a general view and approach to Edge Computing Architectures in order to achieve:
    - **Coordination Between the Physical and Digital Worlds:** A real-time, systematic cognitive model of the physical world is established. The status of the physical world is predicted, and the running of the physical world emulated in the digital world. This approach simplifies reconstruction of the physical world and drives the physical world to optimize the operation of physical systems.
    - **Cross-Industry Collaboration:** Information and Communications Technology (ICT) industry and vertical industries can build and reuse knowledge modelling systems in their own realms. The ICT industry shields the complexity of ICT technologies using the horizontal edge computing model and reference architecture.

Working on the Edge

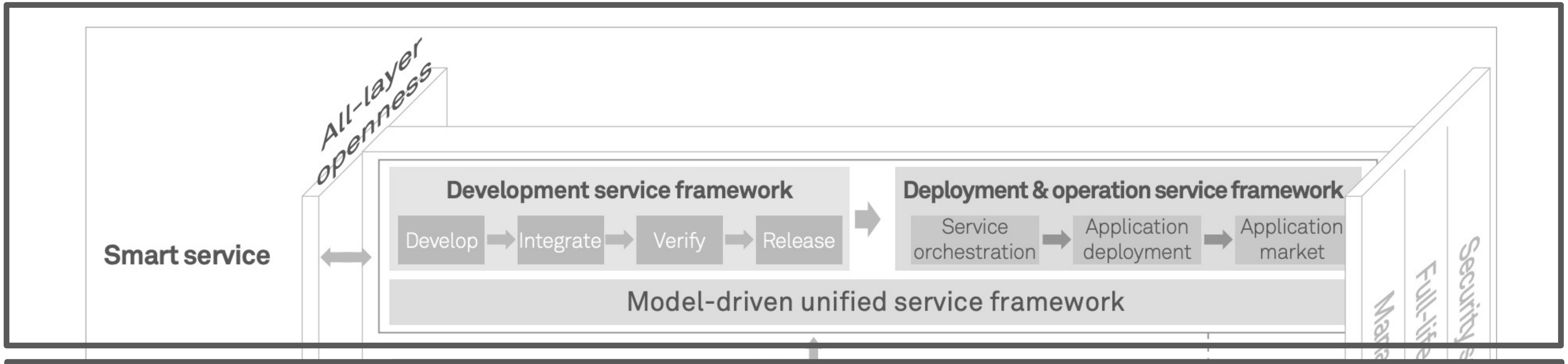
- and to achieve also:.
    - **Reduced System Heterogeneity and Simplified Cross-Platform Migration:** Model-based interfaces between systems, subsystems, services, and new and legacy systems enable interaction, simplifying integration of these systems. Using the model, software interfaces can be decoupled from development languages, platforms, tools, and protocols, which reduces the complexity of cross-platform migration
    - **Effective Support for System Lifecycle Activities:** System lifecycle activities include full lifecycle activities of application development services, deployment and operation services, data processing services, and security services. To address these challenges, the ICT industry is adopting technological innovation such as virtualization, Software-Defined Networking (SDN), model-driven Service Orchestrator (SO), and microservices.

Working on the Edge



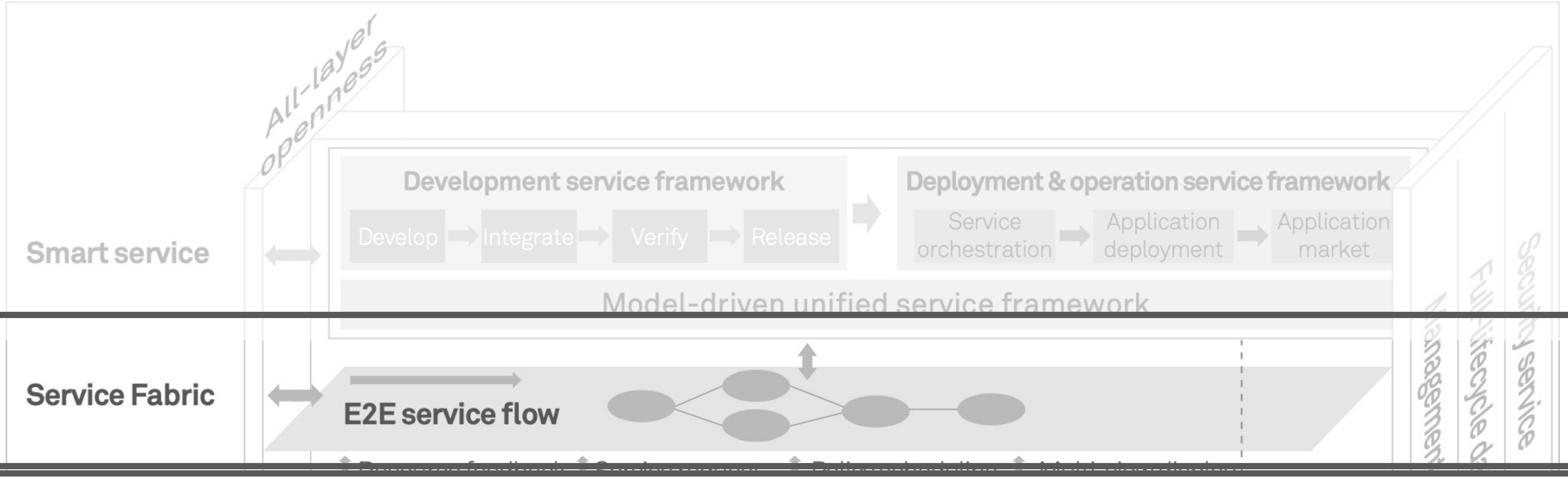
Edge Computing Reference Architecture 2.0 - http://en.ecconsortium.net/Uploads/file/20180328/1522232376480704.pdf

Working on the Edge



- **Smart Service** based on:
  - A common model and development and Deployment Framework
  - Intelligent coordination between service development and deployment
  - Consistent software development interfaces and automatic deployment and operations thorugh Continuous Integration and Continuous Delivery (CI/CD) technologies

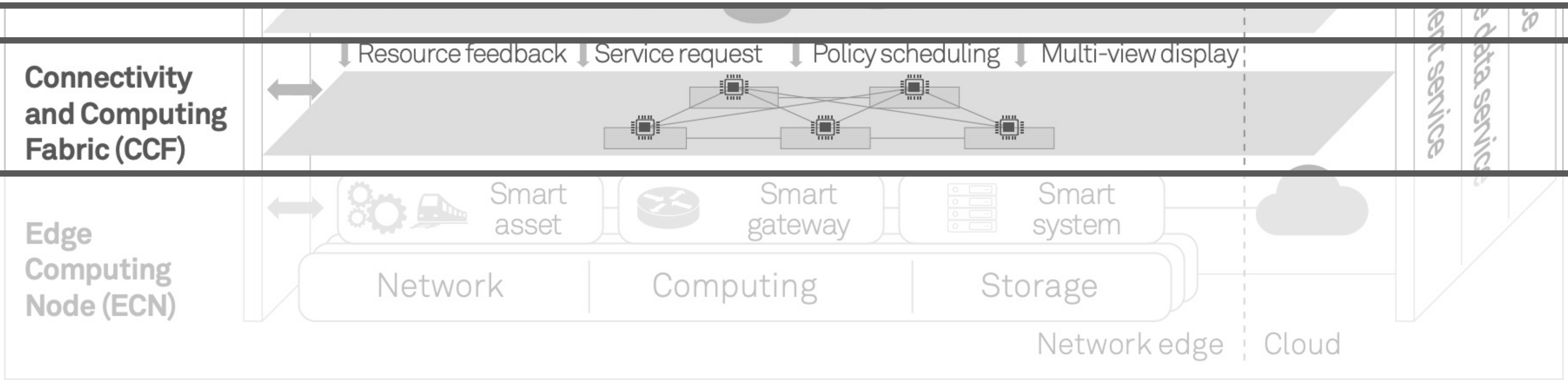
Working on the Edge



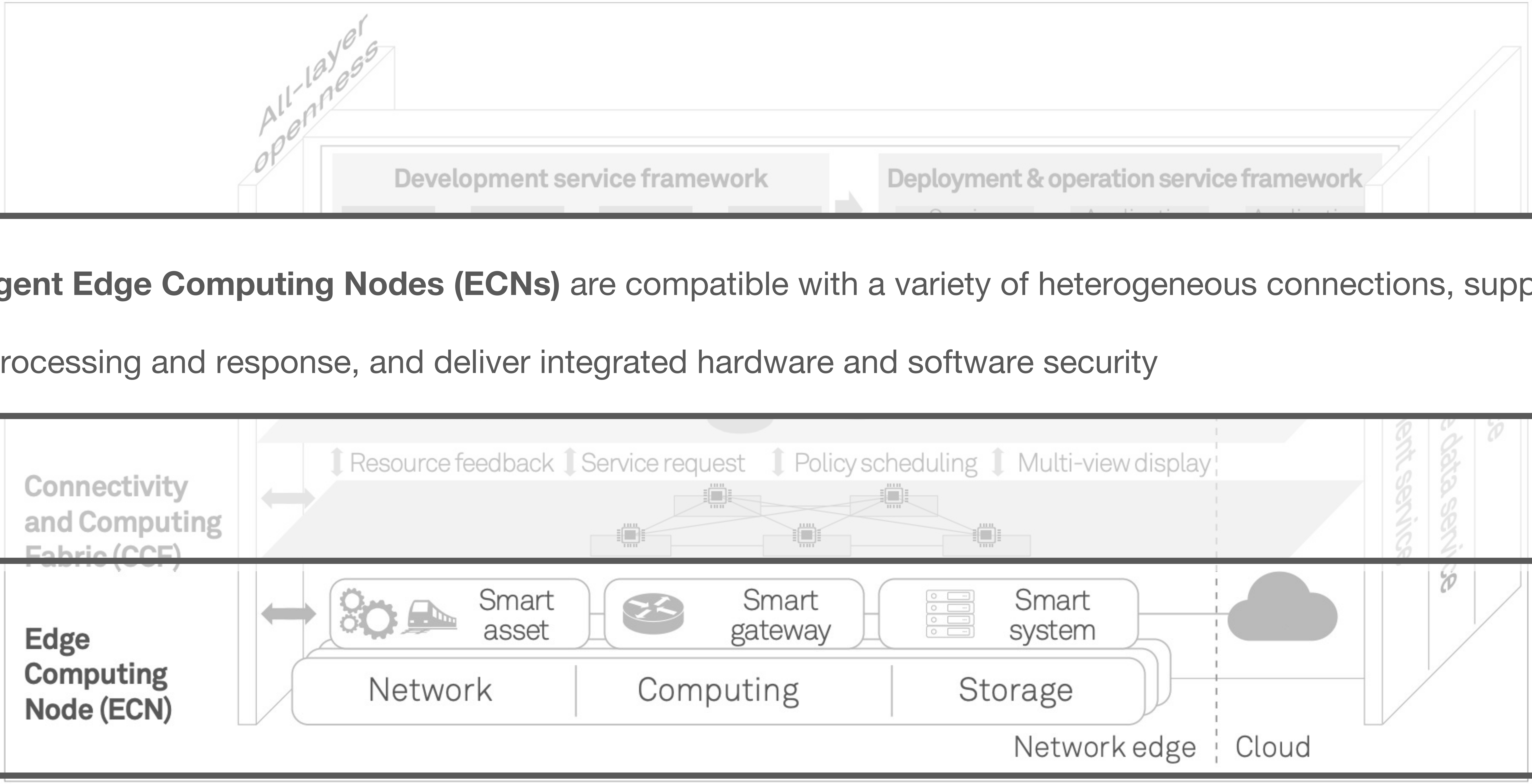
- **Service Orchestration** to enable Edge to Edge (E2E) communication and interoperability.
- The aim in to allow existing service to easily communicate and interact considering both applications from the same or different producers/developers. This solution will allow the creation of new applications, services and business models and the enhancement and enrichment of collected and analyzed data

Working on the Edge

- **Connectivity and Computing Layer/Fabric** enables a simplified architecture and simplifies the distributed edge intelligence architecture for services.
- It allows the connection and communication of involved elements in a shared and secure network that may be also controlled by software and potentially fully virtualized



Working on the Edge



- **Intelligent Edge Computing Nodes (ECNs)** are compatible with a variety of heterogeneous connections, support real-time processing and response, and deliver integrated hardware and software security

Network Provider & Edge Computing

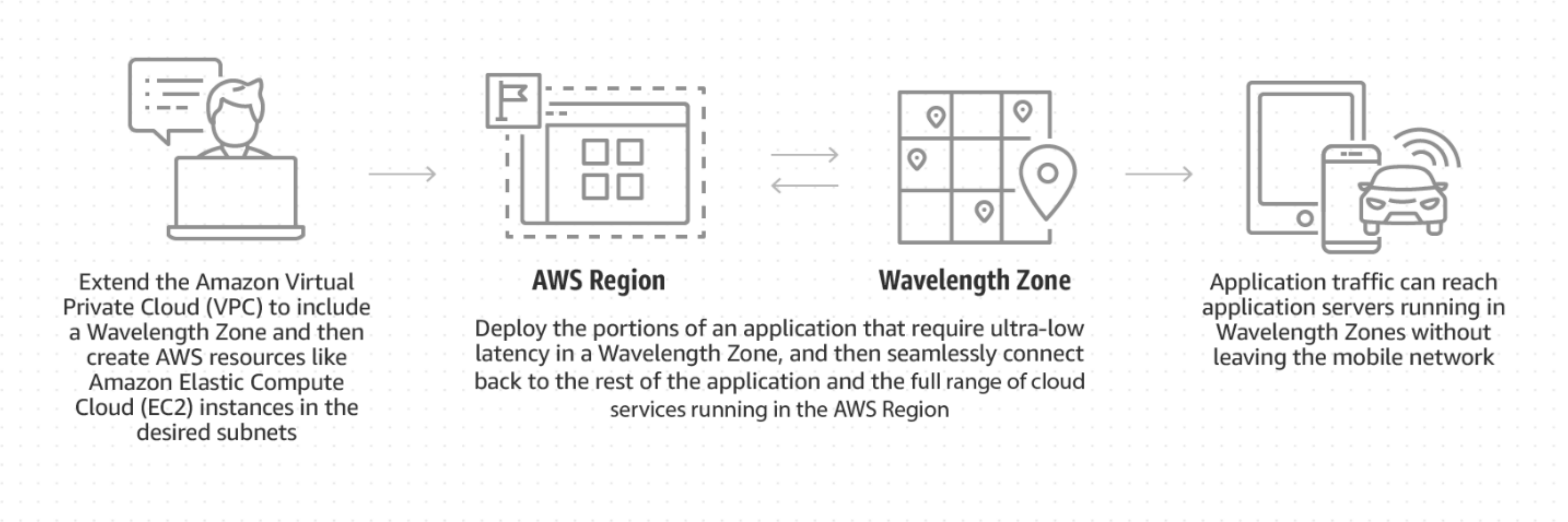
aws

Contact Sales | Support | English | My Account | Create an AWS Account

re:Invent | Products | Solutions | Pricing | Documentation | Learn | Partner Network | AWS Mar

AWS Wavelength | Overview | Features | FAQs

How it works



Technical use cases

Ultra-low latency compute

Deliver high-performance compute at the edge for faster application performance and emerging interactive applications like game streaming, virtual reality, and real-time rendering that require latencies of single-digit milliseconds to end-users.

Edge data processing

Enable and offload data processing tasks to take place at the mobile provider network edge to conserve 5G device resources like power and bandwidth for use cases like industrial automation, smart cities, IoT, and autonomous vehicles.

https://www.vodafone.com/business/news-and-insights/company-news/....

Vodafone Business and Amazon Web Services to bring edge computing closer to customers

Company news

Subscribe | Share

04 Dec 2019

AWS Wavelength supports ultra-low latency by enabling compute and storage services at the edge of Vodafone's network

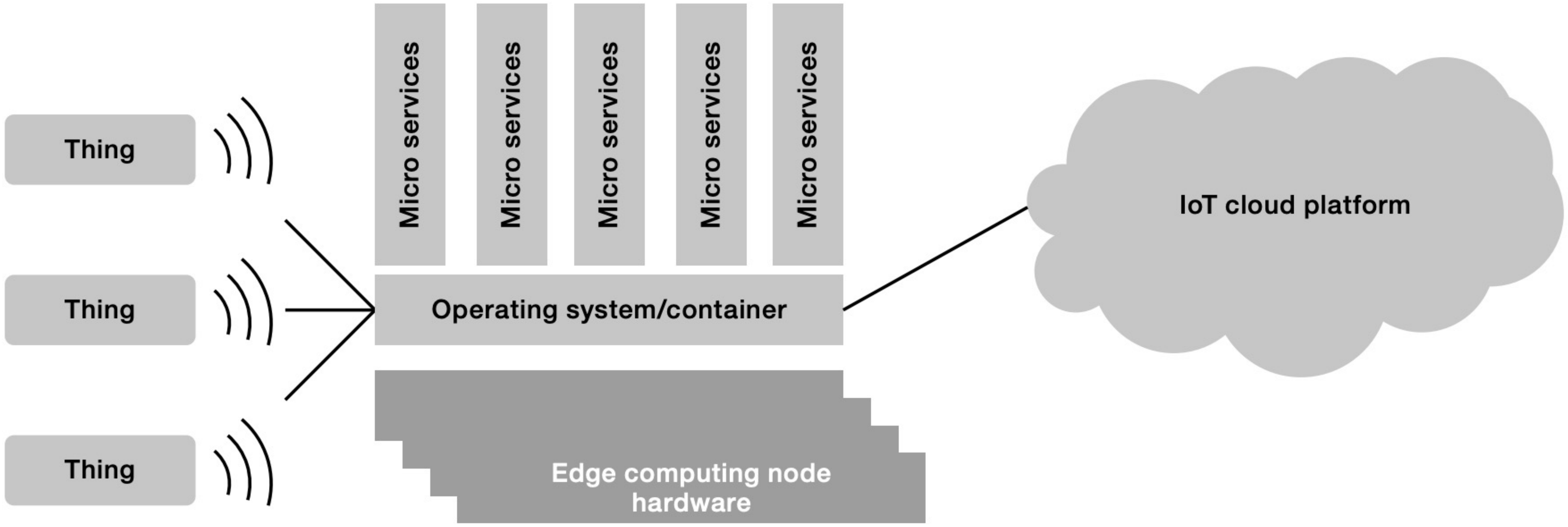
Vodafone Business is collaborating with Amazon Web Services (AWS) to make AWS Wavelength available in Europe. AWS Wavelength provides developers with the ability to build applications that serve end users with single-digit millisecond latencies over the 5G network. AWS Wavelength embeds AWS compute and storage services at the edge of telecommunications providers' 5G networks, enabling developers to serve use cases that require ultra-low latency.

Enabling AWS Wavelength at the edge of Vodafone's 5G network is designed to provide multi-access edge computing capabilities to developers, Internet of Things (IoT), devices and end users. Responsiveness matters when it comes to artificial intelligence, augmented and virtual reality, video analytics, autonomous vehicles, robotics and drone control. These applications require latencies that are 5-10 times lower to deliver business impact. Vodafone Business and AWS are bringing the AWS cloud closer to the devices that need it by running AWS Wavelength in strategic locations within Vodafone's 5G network.

https://aws.amazon.com/it/wavelength/

Edge Computing Software Evolution

- Most of the IoT ECN technology, i.e. the technology which is the host for edge processing capabilities, runs on flavours of the Linux OS while using different kinds of processor architectures.
- An industry-wide trend is emerging to package edge computing capabilities into microservices and deploy them within containers on IoT ECNs
- Containers provide security through isolation; they also serve as deployment units that simplify lifecycle management through less interdependency and complexity



Edge Operating System - IoT Gateway

- In Edge Computing, data will be processed, analyzed and aggregated at the network edge near things or data sources.
- The IoT gateway is the perfect host of all these capabilities. The main of IoT gateway include cloud offloading, private data filtering, data aggregation, protocol translation, caching etc.
- The OS running on IoT gateways is usually a general purpose OS such as Linux. Horizontal decoupling brings openness to gateways.
- Third party applications can be deployed on gateway OS via a Host OS, a container (LXC, Docker) or a virtual machine. The basic function of a gateway is always networking.
- The IoT gateway should support:
    - rich network protocols including Layer 2 (L2) and Layer 3 (L3)
    - virtual local area network (VLAN)
    - routing protocols, multicast protocols and reliability protocols, so that the gateway can have flexible networking capabilities and can interconnect with equipments of third-party manufacturers.

IoT & Edge Means Interoperability

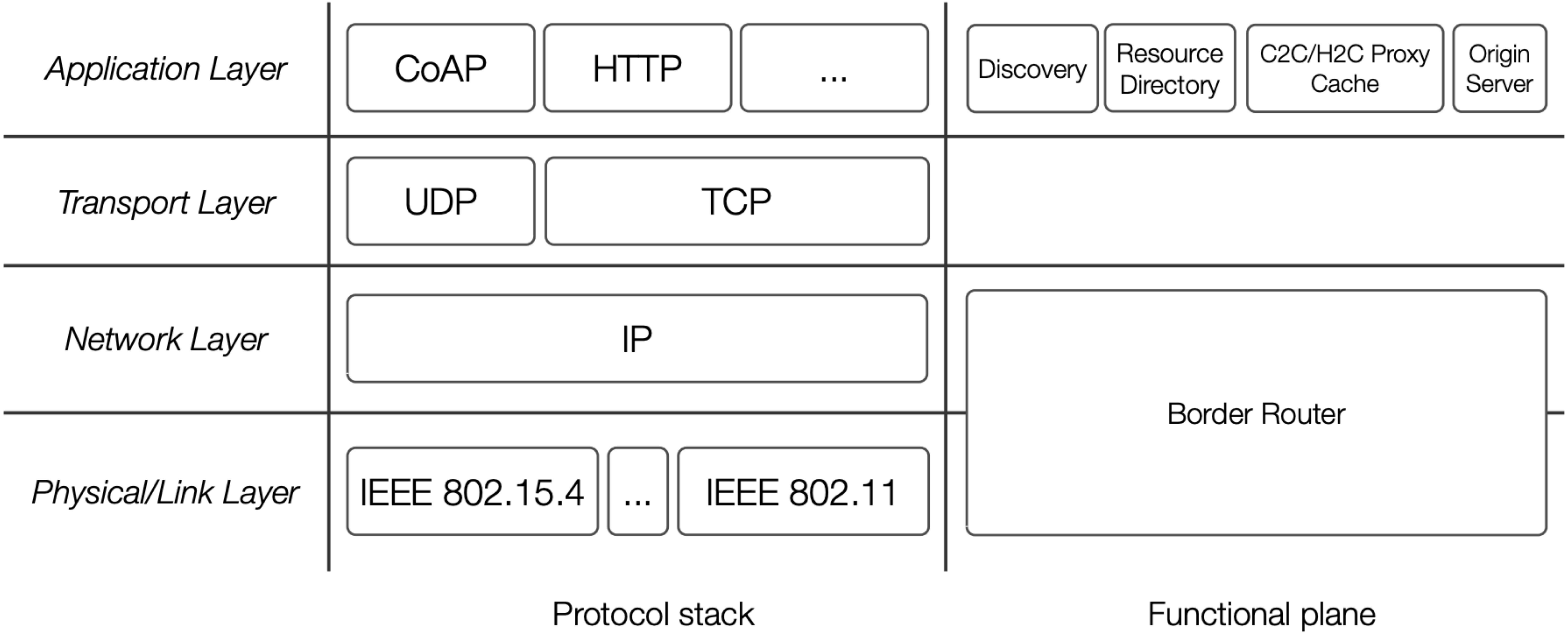
- **Today's IoT largely exists in isolation**, and it has been **impossible to realize a truly interconnected world where devices are interoperable**.
- Consistent APIs over the connectivity, security, together with the adoption of standard protocols and data formats are strategic elements to solve the interoperability issues
- Interoperability is one of the main pillars of the IoT and in particular for IoT applications on the Edge.
- The of existing standards and widespread protocols and approaches coming from the existing Internet (IP-based stack) and the World Wide Web. Main advantages are:
  - Local & remote interaction with objects and services (device-to-device and device-to-cloud approaches)
  - Share of well-known standard protocols and data/resource representations
  - Easy management, integration, and composition of services
  - Management of heterogeneity and dynamism
  - Efficient data collection from multiple sources
  - Self Discovery and Autonomous Configurability

The IoT Hub

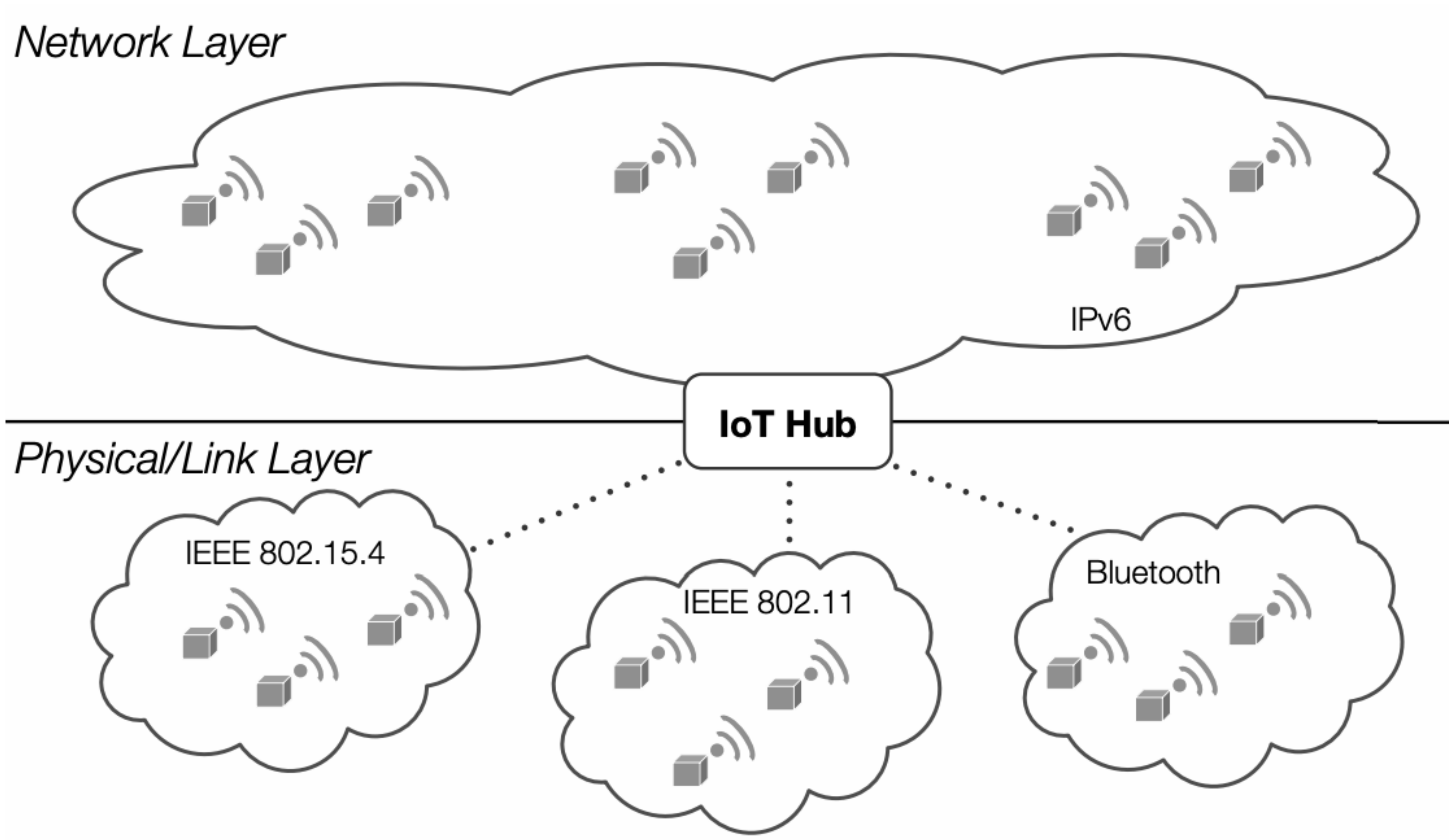
- The proposed IoT Hub is a Fog node placed at the edge of multiple physical networks with the goal of creating an IP-based IoT network to be used as an infrastructure for the deployment of WoT applications. As opposite to Smart Objects (SOs), the IoT Hub is a node that does not present strong limitations on the computational power and available memory, nor strict requirements on energy consumption.
- The IoT Hub plays a fundamental role by implementing the following functions at different layers of the protocol stack:
  - **Border Router**
  - **Service and Resource Discovery**
  - **Resource Directory (RD)**
  - **Origin Server (OS)**
  - **CoAP-to-CoAP (C2C) Proxy**
  - **HTTP-to-CoAP (H2C) Proxy**
  - **Cache**

5. S. Cirani, G. Ferrari, N. Iotti and M. Picone, "The IoT hub: a fog node for seamless management of heterogeneous connected smart objects," 2015 12th Annual IEEE International Conference on Sensing, Communication, and Networking - Workshops (SECON Workshops), Seattle, WA, 2015, pp. 1-6. doi: 10.1109/SECONW.2015.7328145. https://ieeexplore.ieee.org/document/7328145

The IoT Hub

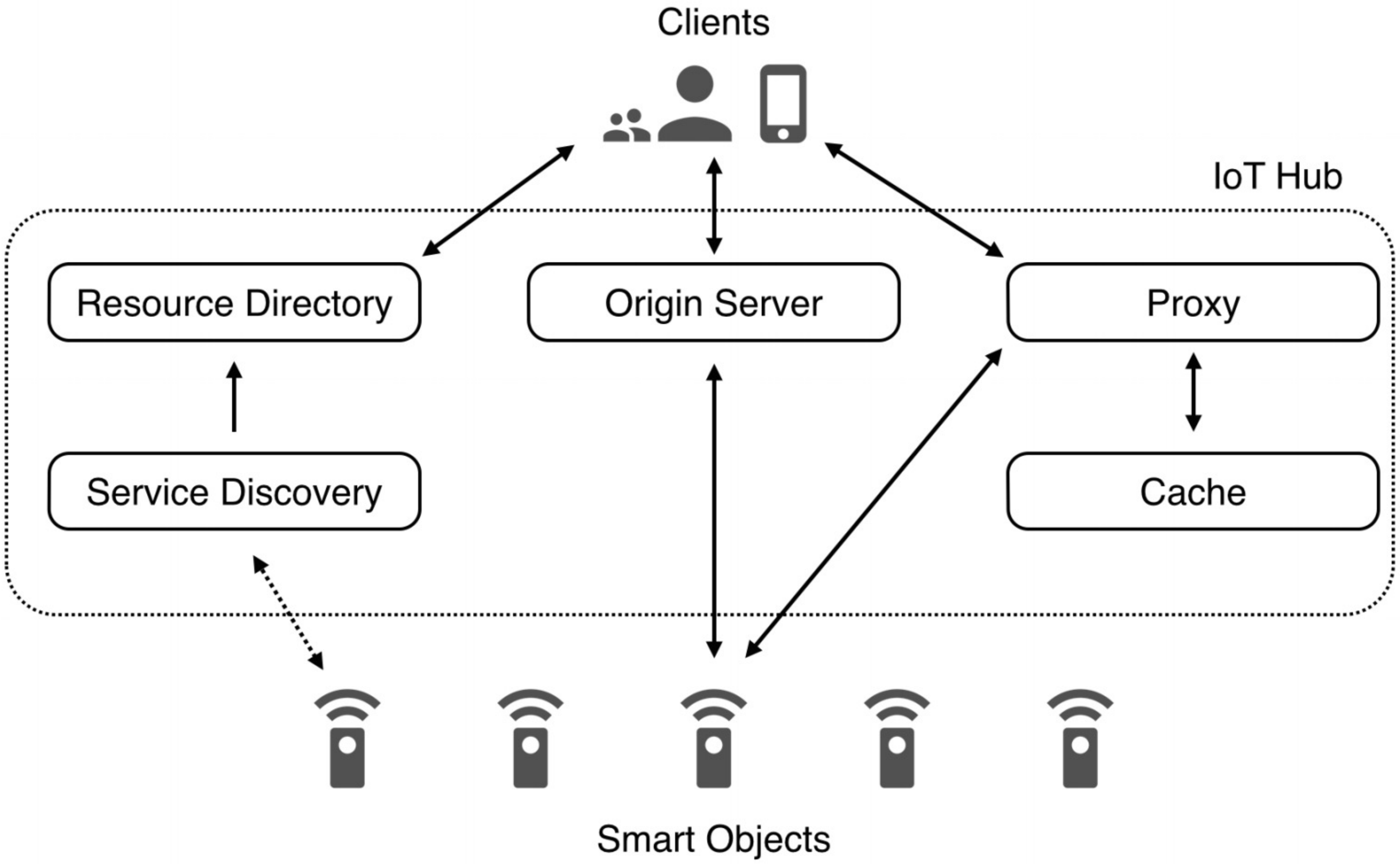

Application Layer
CoAP
HTTP
...
Discovery
Resource Directory
C2C/H2C Proxy Cache
Origin Server
Transport Layer
UDP
TCP
Network Layer
IP
Border Router
Physical/Link Layer
IEEE 802.15.4
...
IEEE 802.11
Protocol stack
Functional plane

The IoT Hub - Link Layer Functions



- **Border Router:** the IoT Hub is the gateway/bridge between one or more constrained networks (e.g., IEEE 802.15.4);
- The border router function implemented by the IoT Hub are two-fold. On the one hand, the IoT Hub provides intra-network IP routing, for instance by acting as coordinator of a RPL-based IEEE 802.15.4 multihop network.
- On the other hand, it provides a single addressing space for multiple networks it is attached

The IoT Hub - Application Layer Functions



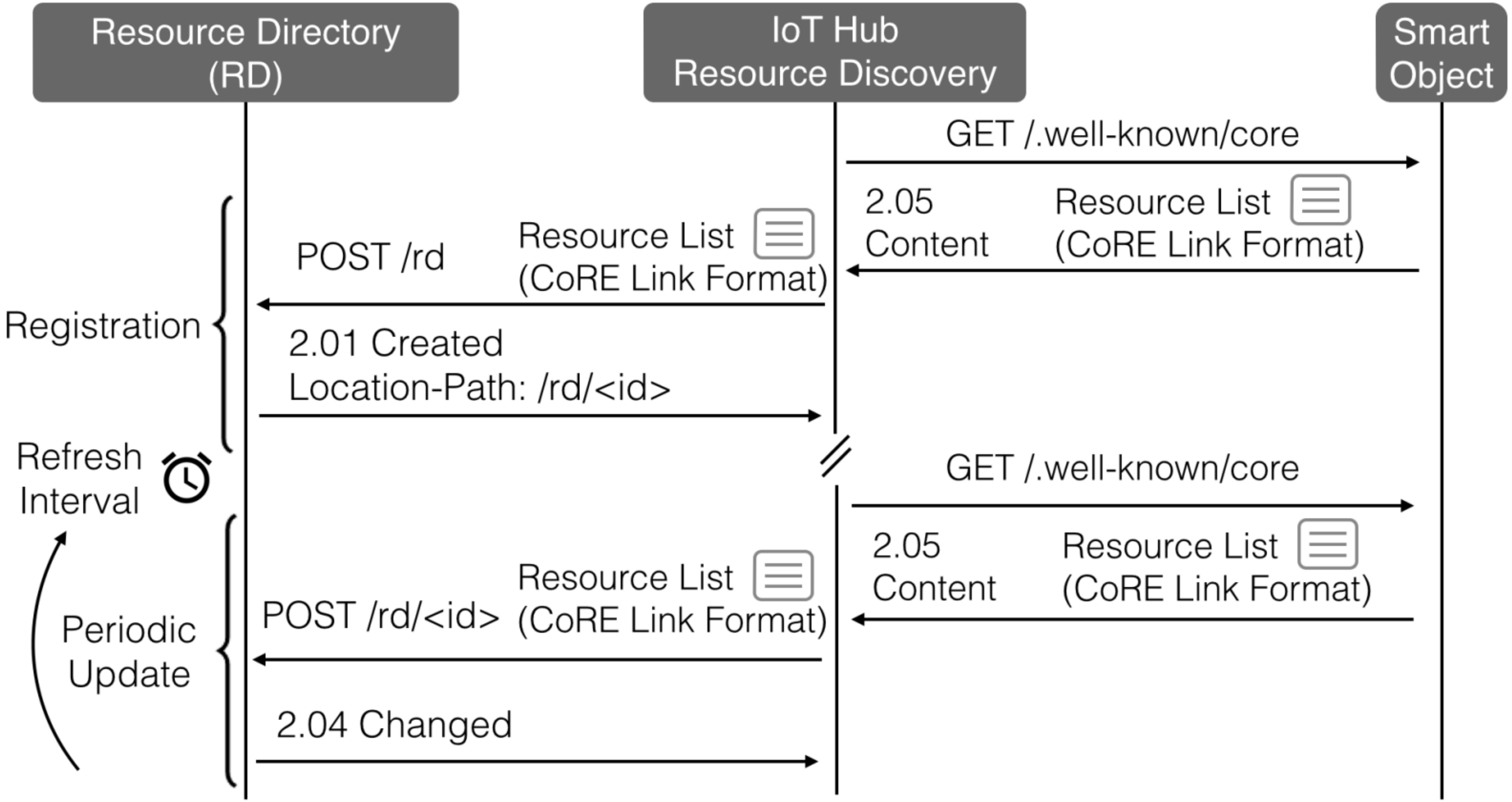
The IoT Hub - Service and Resource Discovery

- The IoT Hub implements Service and Resource Discovery in order to constantly have knowledge of the SOs that it is supposed to manage. Service Discovery aims at retrieving the scheme, IP address, and port at which a SO endpoint is reachable. This function is implemented according to multiple mechanisms:
  - Zero configuration mechanisms based on Multicast DNS (mDNS – RFC 6762) and DNS-Based Service Discovery (DNS-SD – RFC 6763)
  - The use of NMap for network scanning;
  - Other mechanisms, such as Lightweight multicast forwarding and Distributed P2P Service Discovery
- Once endpoints have been discovered, they can be requested to retrieve the list of resources that they host.
  - For CoAP-based SOs, the **/.well-known/core** URI can be used for this purpose. The list of resources is returned in CoRE Link Format (RFC 6690)
  - For HTTP-based SOs, the list might be returned using Web Linking (RFC 5988). Support for proprietary SOs that are not standard-compliant, other mechanisms can be integrated.

The IoT Hub - Resource Directory

- The Resource Directory (RD) is an specialized element that serves as a registry for all the resources in the overall IP network managed by the IoT Hub. Resources are characterized by a link (i.e., a URI) and a set of attributes that can be used to describe them, as defined in RFC 6690
- The RD provides
  - a function set for the registration, update;
  - deletion of resources
  - a function set for the lookup of resources
- The RD can be queried by clients to discover links to resources. Requests can contain query parameters to filter the results. The RD can be accessed using CoAP.
- The Resource Discovery module interacts with the RD so when resources are discovered, they are registered to the RD and then periodically updated

The IoT Hub - Resource Directory & Discovery



The IoT Hub - Resource Directory & Discovery

- Resource Directory Registration Steps:
  a. The Resource Discovery module registers resources on the RD by sending a CoAP POST request with the links to the discovered resources.
  b. The RD replies with the URI of the newly created resource, which is then used for any update or deletion operation
  c. The registration is also repeated by reporting links that allow access through the CoAP-to-CoAP proxy and HTTP-to-CoAP proxy, which can be used by clients that need a proxy (with optional protocol translation)

The IoT Hub - Origin Server

- In some cases, SOs are characterized by very limited capabilities and cannot act as a server and host resources.
- When this occurs, SOs act as clients and need an external node to maintain resources on their behalf. In order to support this kind of situation, the IoT Hub provides an Origin Server (OS) function.
- Clients can request the IoT Hub to create and maintain resources on their behalf. In this case, clients also need to take care of the registration of the newly created resource on the RD, since its attributes are unknown to the OS.
- The IoT Hub is then going to be used to access resources.

The IoT Hub - C2C/H2C Proxy and Cache

- The IoT Hub can act as an intermediary for communication with Smart Objects.
- The IoT Hub can act either as a:
  - **CoAP-to-CoAP (C2C)** proxy (to allow external CoAP clients to access constrained resources that cannot be reached directly)
  - **HTTP-to-CoAP** proxy (to allow HTTP clients to allow access to CoAP resources).
- There are several reasons that motivate the presence of a proxy in IoT scenarios:
  - to shield the constrained network from the outside, e.g. for security reasons (DoS attacks);
  - to support integration with the existing web through legacy HTTP clients;
  - to reduce network load
  - to support data formats that might not be suitable for constrained applications, e.g. XML.

The IoT Hub - C2C/H2C Proxy and Cache

- An additional reason to have a Proxy and Intelligent intermediate node is to: ***ensure high availability on resources through caching***
- For this reason the IoT Hub also implements a **cache**, which is particularly useful in presence of proxies. In fact, the IoT Hub keeps a “fresh” representation of resources that can be used to serve requests without requiring to forward the request to the Smart Object
- This is expedient to avoid unnecessary computation on constrained devices and also to make it possible for clients to get information also in case of duty-cycled devices, which may not be reached at all times.

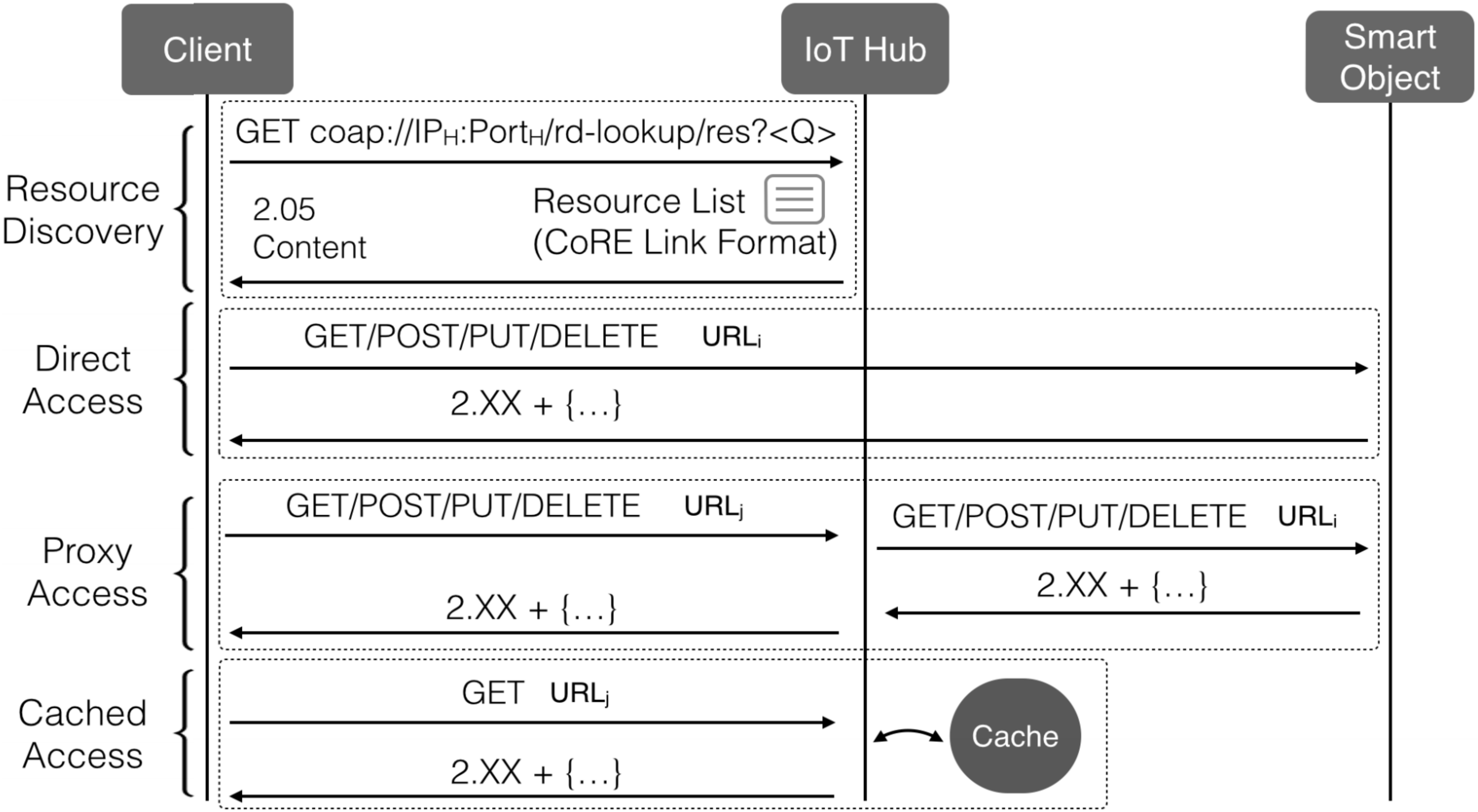
The IoT Hub - Resource Access

- The presence of the IoT Hub enables clients to effectively discover and access resources hosted by heterogeneous SOs.
- The added value of the IoT Hub is its capability to hide completely the diverse nature of SOs, in terms of hardware, wired/wireless communication protocol, and application-layer protocol used, to clients, which can interact with them using uniform interfaces and without requiring any prior configuration.
- Interaction with SOs through the IoT Hub occurs according the following steps (next Slides) …

The IoT Hub - Resource Access - Steps

1. The client discovers the IoT Hub through a suitable mechanism, such as DNS-SD;
2. The client queries the RD of the IoT Hub to get a list of resources that match its interests (resource discovery); the RD returns the list of links in CoRE Link Format;
3. the client selects the resource to access; since each resource has been registered with multiple links (direct, through CoAP-to-CoAP proxy, or through HTTP-to-CoAP proxy), the client can also select to use the URI that best fits its characteristics;
4. The client sends requests to the SO according to the URI selected at the previous step (resource access) and according to the type of interfaces exposed by the resource (e.g., a sensor resource might support only read – GET – operations, while actuators might also support write – POST/PUT – operations).

The IoT Hub - Resource Access - Steps



IoT Hub & Virtual Replication (E2C & E2E)

- Large-scale IoT platforms for interconnected devices are required to effectively manage resources provided by smart objects.
- In order to cope with a large number of resources in a scalable, seamless, and secure way it is possible to dynamically handling IoT Hub Virtual Instances both at Cloud and Edge Level
- The basic concept is the Virtual IoT Hub: a Cloud-based “entity” replicating all the functions of a physical IoT Hub, which external clients will query to access resources
- IoT Hub virtual replicas are constantly synchronized (through a dedicated protocol) with the physical IoT Hub through a low-overhead protocol based on Message Queue Telemetry Transport (MQTT).
- <u>The paper associated to this research has been focused on the Cloud side Virtualization to support Edge to Cloud scenarios but it can be also applied in Edge to Edge contexts</u>

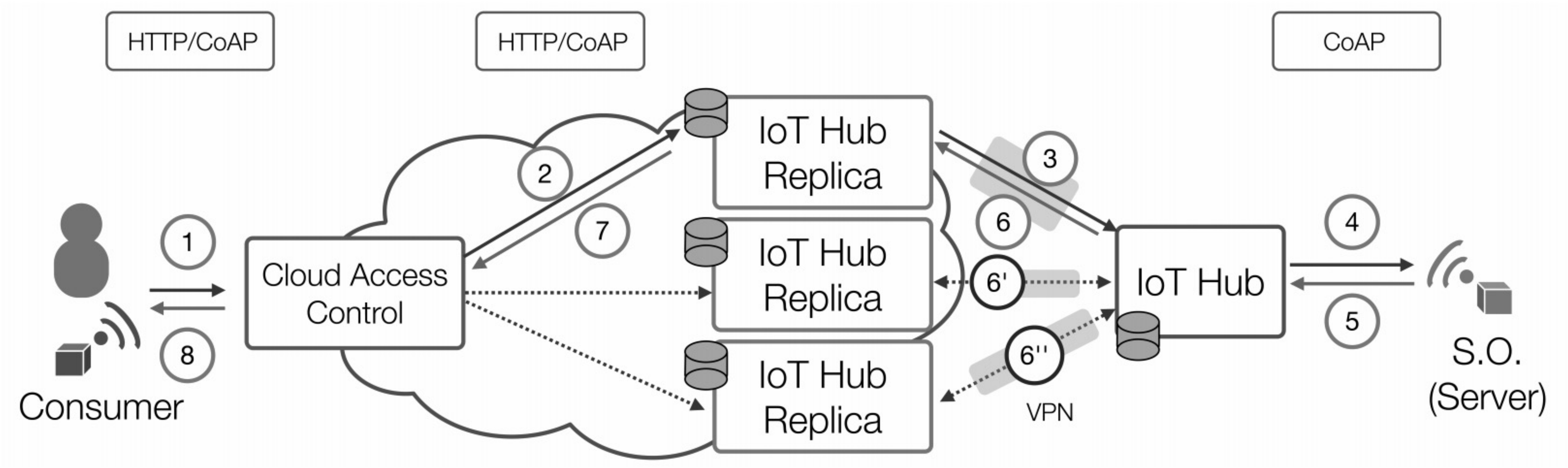
5. *Cirani, S.; Ferrari, G.; Mancin, M.; Picone, M. Virtual Replication of IoT Hubs in the Cloud: A Flexible Approach to Smart Object Management. J. Sens. Actuator Netw. 2018, 7, 16.* *https://www.mdpi.com/2224-2708/7/2/16*

IoT Virtual Hub - Operational Scenarios

- Resource access through the proposed platform can occur according to the following three operational models implemented by a smart object:
  - CoAP server;
  - Observable CoAP server;
  - CoAP client.
- Operational Methodologies available are:
  - Polling Resources
  - Observing Resources
  - Pushing Resources

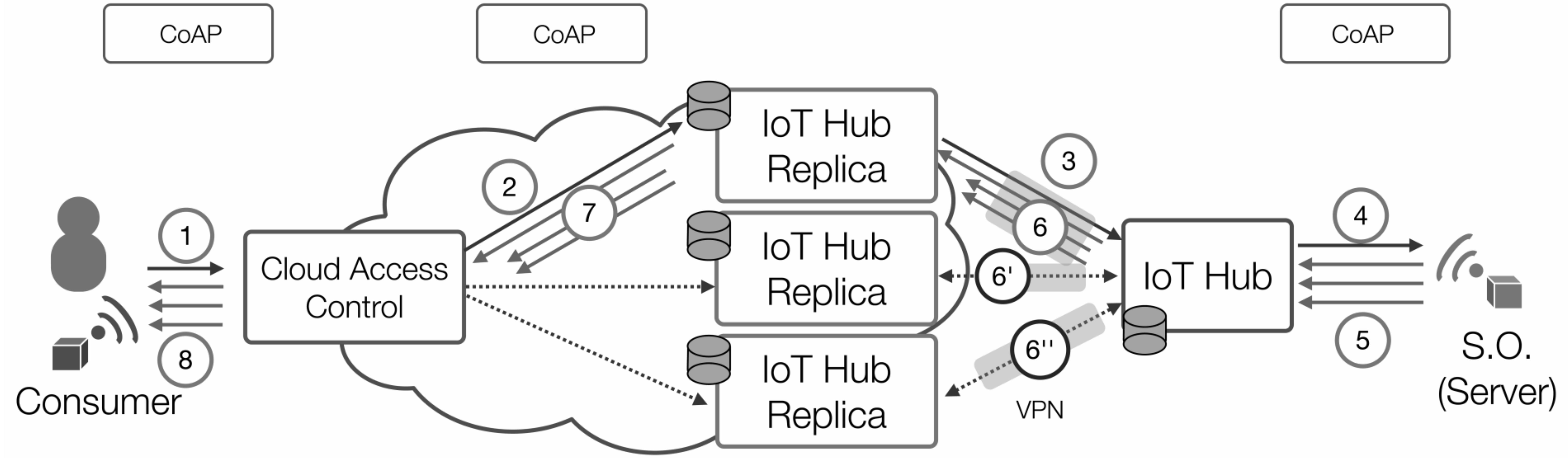
IoT Virtual Hub - Polling Resources

- A HTTP/Constrained Application Protocol (CoAP) client requests a resource to the Cloud platform, which internally selects a suitable IoT Hub replica and forwards the request. The request reaches the smart object only if neither the replica nor the IoT Hub have stored a fresh cached representation of the resource. Other replicas (which are not in the path of the request) are kept in sync using the synchronization protocol.



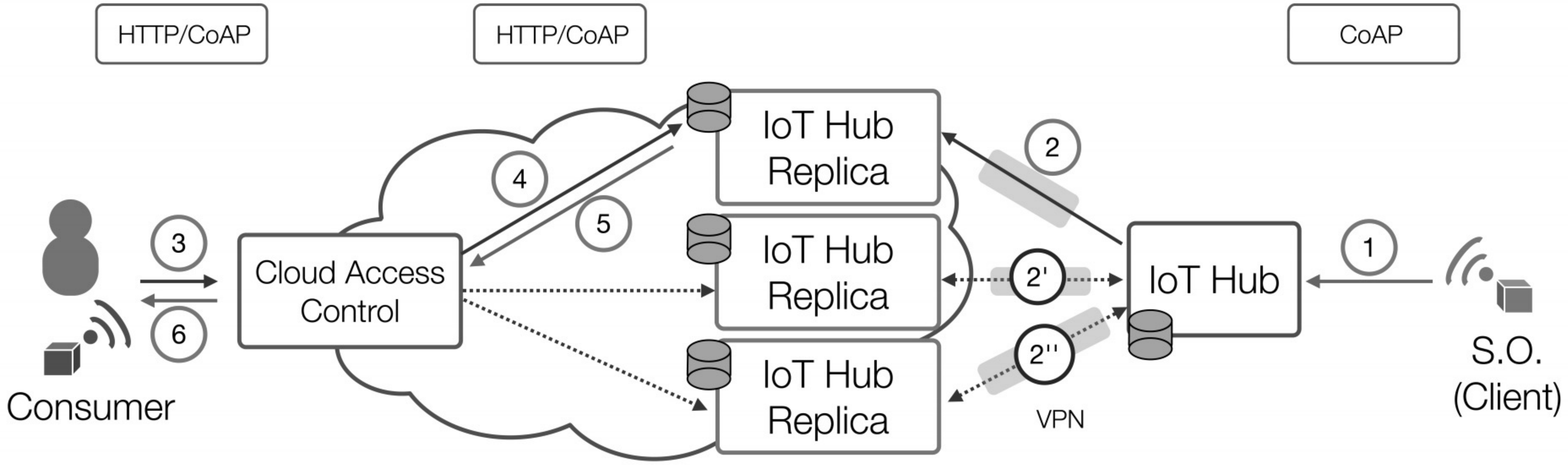
IoT Virtual Hub - Observing Resources

- A CoAP client requests a resource to the Cloud platform using the CoAP Observe option. The Cloud platform internally selects a suitable IoT Hub replica and forwards the request. The observe request is then forwarded to the IoT Hub and to the smart object thus creating an “observe chain”. Resource updates are then sent from the smart object back to the IoT Hub, then to the replica, and finally to the CoAP client. Other replicas (which are not in the path of the request) are kept in sync using the synchronization protocol.

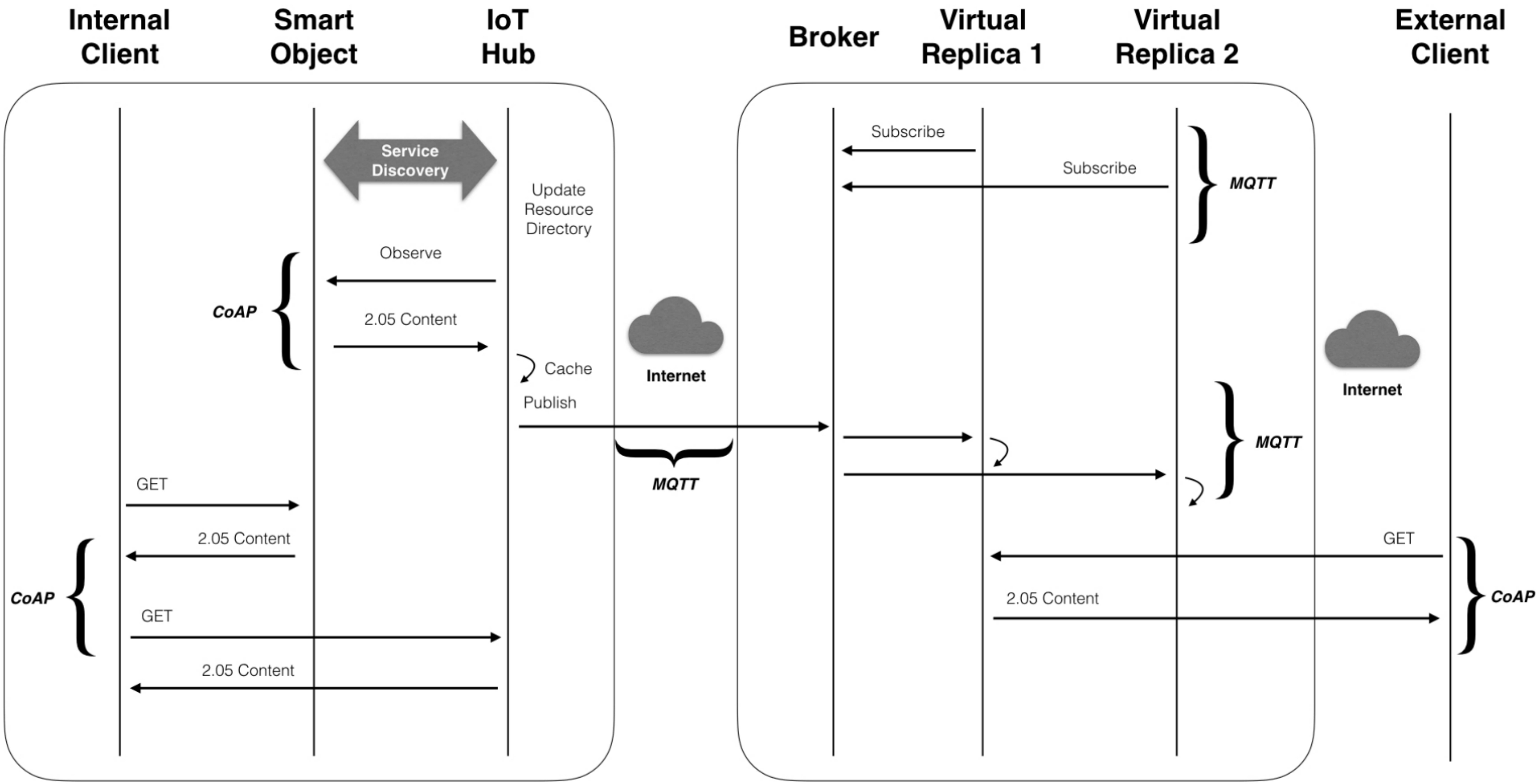


IoT Virtual Hub - Pushing Resources

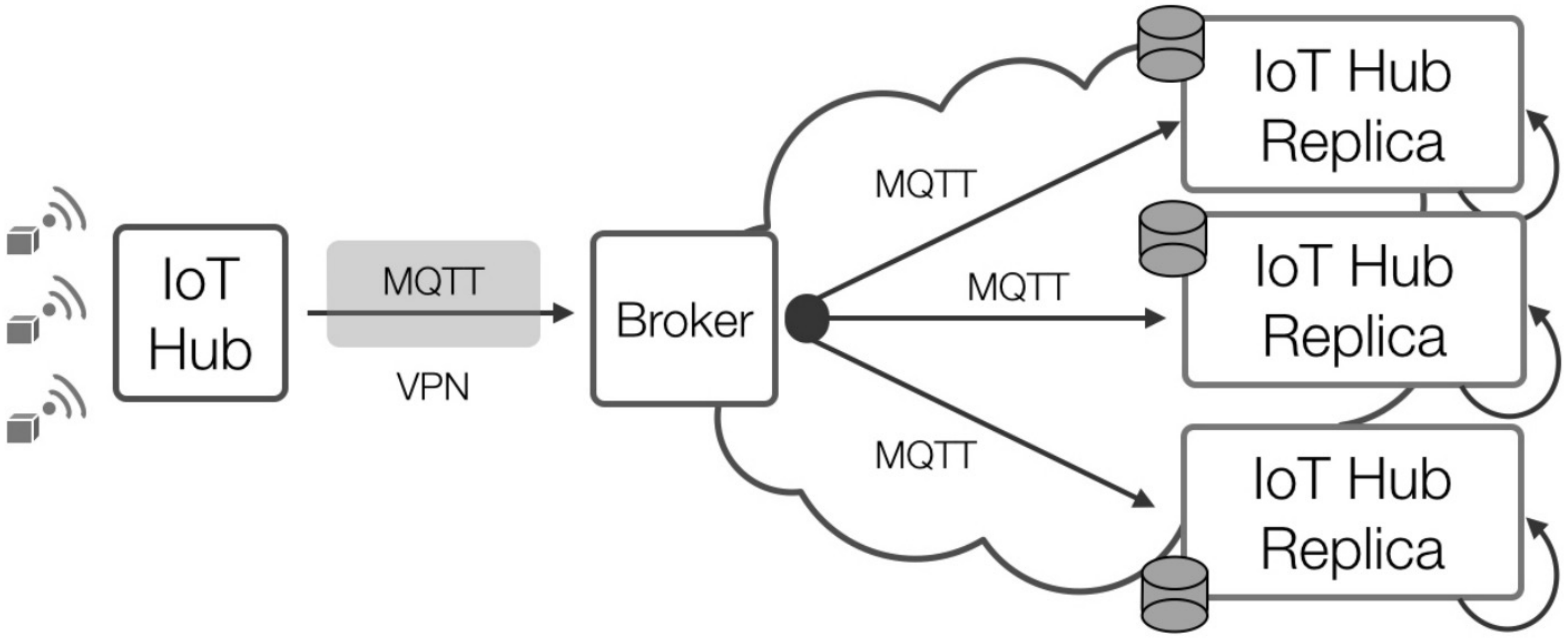
- A smart object acting as a CoAP client posts and updates resources on the IoT Hub, which acts as origin server. All replicas are kept in sync using the synchronization protocol. External client can request resources, which will served by a replica



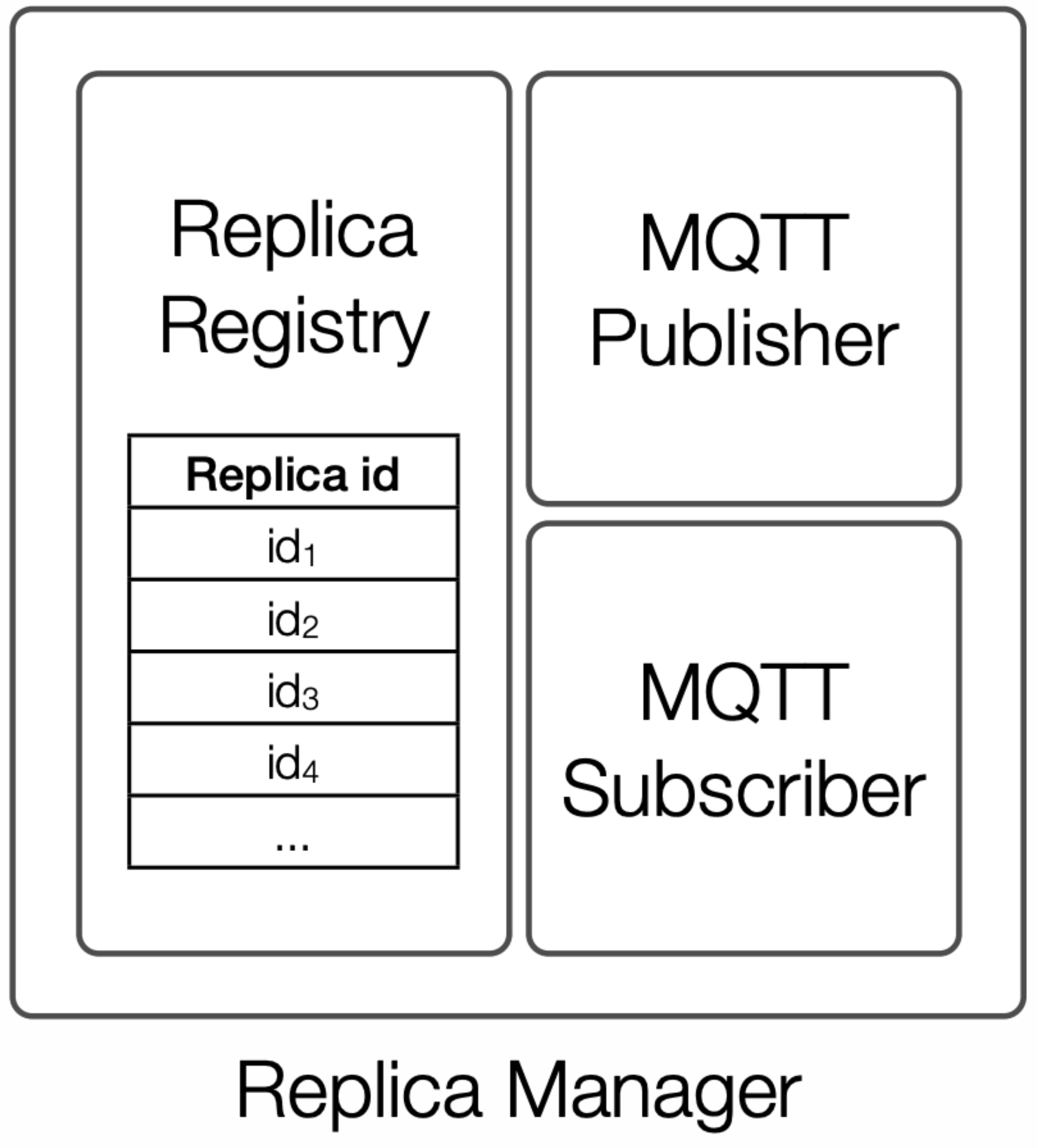
Synchronization Protocol



Replica Communication & Management



The broker-based message flow between the IoT Hub and its replicas

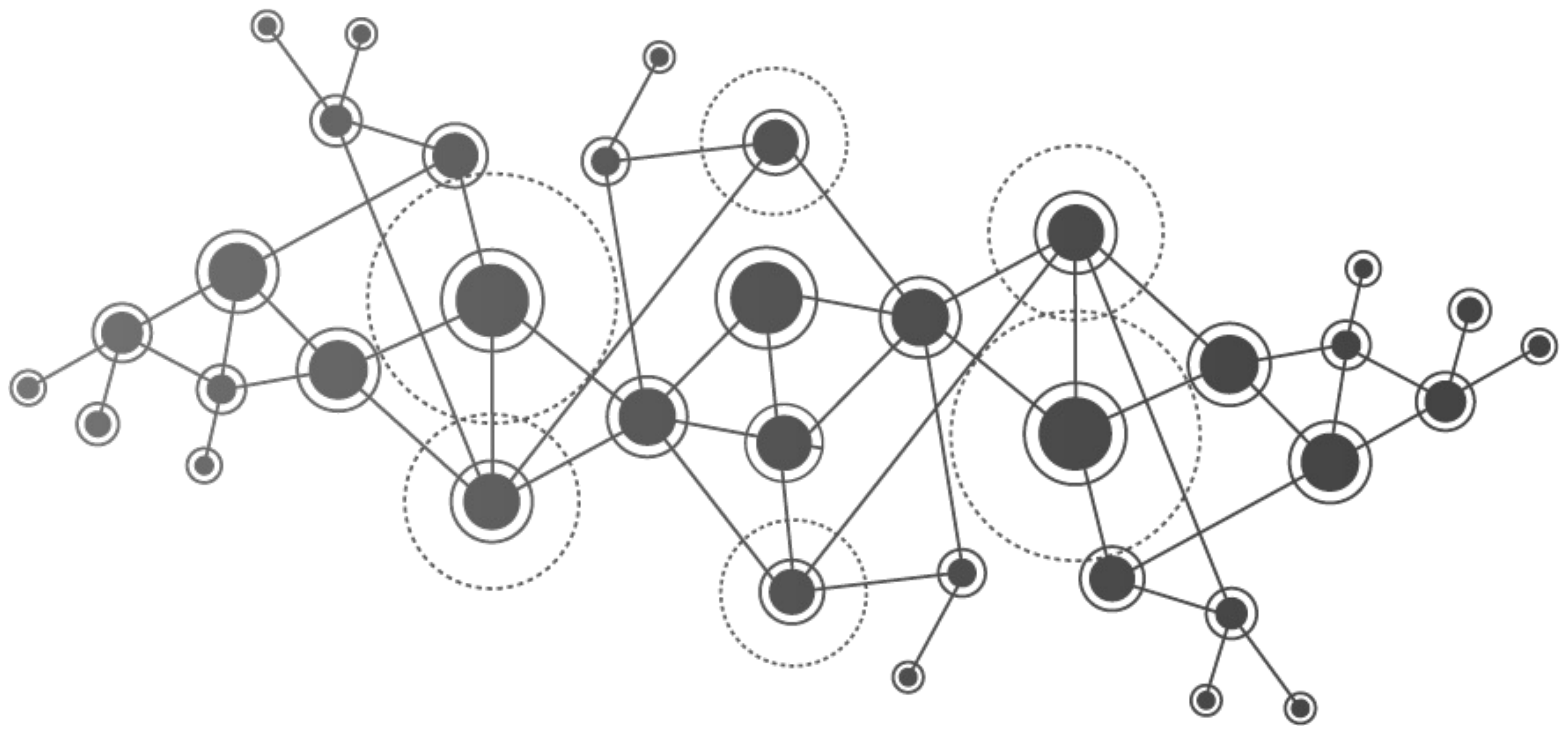


Internal structure of the Replica Registry module of the IoT Hub

References

1. “Internet of Things Applications - From Research and Innovation to Market Deployment”, edited by O. Vermesan and P. Friess, pp. 287-313, The River Publishers Series in Communications, Alborg, Denmark, 2014. ISBN: 9788793102941 [http://riverpublishers.com/view_details.php?book_id=245]
2. Edge Computing Reference Architecture 2.0 - http://en.ecconsortium.net/Uploads/file/20180328/1522232376480704.pdf
3. Developing Software for Multi-Access Edge Computing - https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp20ed2_MEC_SoftwareDevelopment.pdf
4. “Edge Computing: challenges, solutions and architectures arising from the integration of Cloud Computing with Internet of Things”, PhD Dissertation of Ermanno Guardo - Università degli Studi di Catania. http://dspace.unict.it/bitstream/10761/3908/1/GRDRNN84L09C351V-Tesi%20Dottorato%20-%20Ermanno%20Guardo.pdf
5. CLOUD, FOG AND EDGE COMPUTING – WHAT’S THE DIFFERENCE? - https://www.winsystems.com/cloud-fog-and-edge-computing-whats-the-difference/
6. Edge Intelligence White Paper: https://www.iec.ch/whitepaper/pdf/IEC_WP_Edge_Intelligence.pdf
7. Azure IoT a Practical Entry to New Retail https://www.slideshare.net/DanielLi72/azure-iot-a-practical-entry-to-new-retail
8. Welcome to the Edge - https://qwilt.com/welcome-edge-evolution-cloud-computing/
9. S. Cirani, G. Ferrari, N. Iotti and M. Picone, "The IoT hub: a fog node for seamless management of heterogeneous connected smart objects," 2015 12th Annual IEEE International Conference on Sensing, Communication, and Networking - Workshops (SECON Workshops), Seattle, WA, 2015, pp. 1-6. doi: 10.1109/SECONW.2015.7328145. https://ieeexplore.ieee.org/document/7328145
10. Cirani, S.; Ferrari, G.; Mancin, M.; Picone, M. Virtual Replication of IoT Hubs in the Cloud: A Flexible Approach to Smart Object Management. J. Sens. Actuator Netw. 2018, 7, 16. https://www.mdpi.com/2224-2708/7/2/16
11. SIMSEK, M. et al., 5G-Enabled Tactile Internet, IEEE Journal on Selected Areas in Communications, Vol. 34 (No. 3), March 2016
12. A survey of IoT cloud platforms
13. - https://www.sciencedirect.com/science/article/pii/S2314728816300149

UNIMORE
UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA



Intelligent Internet of Things

Edge Computing for the IoT

Prof. Marco Picone

A.A 2023/2024