Alliance for
Internet of Things
Innovation

Challenges in building edge clouds in conjunction with mobile networks

Omar Elloumi, AIOTI and Nokia
Bell-Labs Distinguished Member of Technical Staff

Brussels September 2020

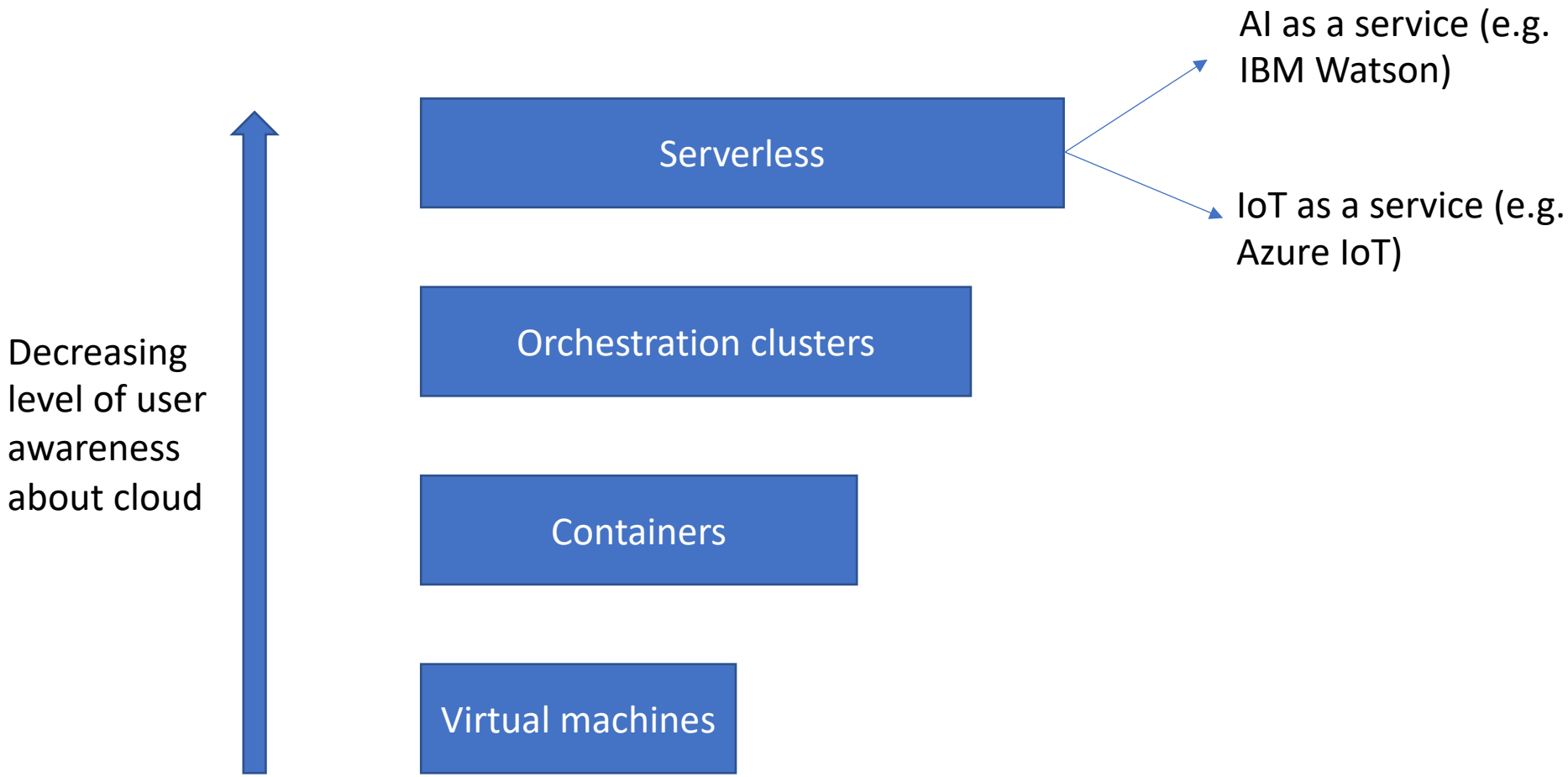
Motivations

- **Demand is here to stay:** IIoT, V2X, Gaming, AR/VR.
- **MNOs:**
  - have geographically dense points of presence that makes edge computing feasible
  - 5G just around the corner: Ultra-Reliable Low Latency Communications (URLLC) and massive Machine Type Communication (mMTC) call for complementary edge computing capabilities to realize the full market potential of 5G.

Let's start with what I consider a legitimate question…

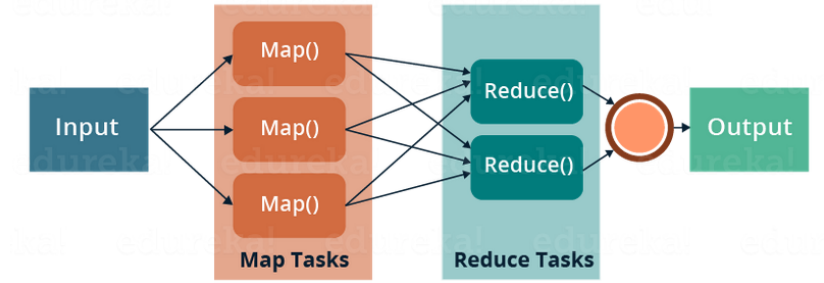
Cloud computing has existed for some time,
why deploying cloud at the edge makes things any different?

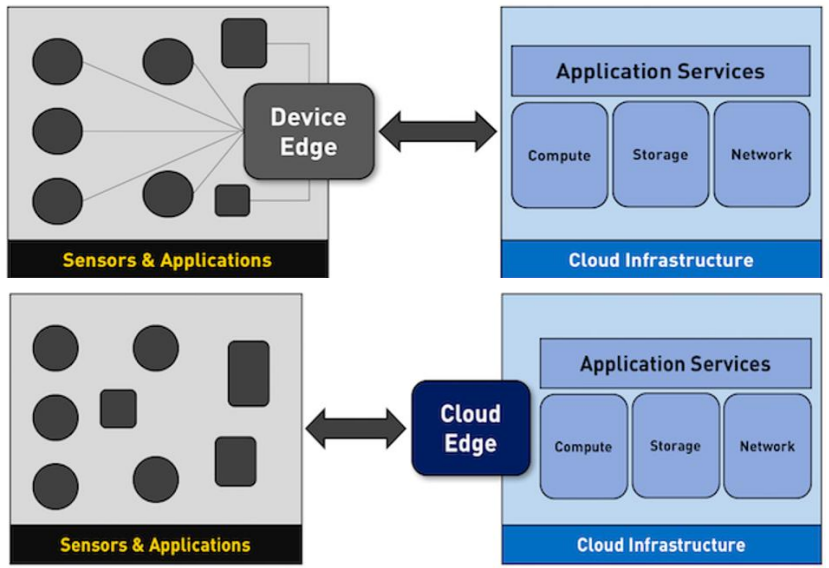
Understanding cloud evolutions, constantly moving up the layers



Cloud Computing Features for IoT Virtualisation

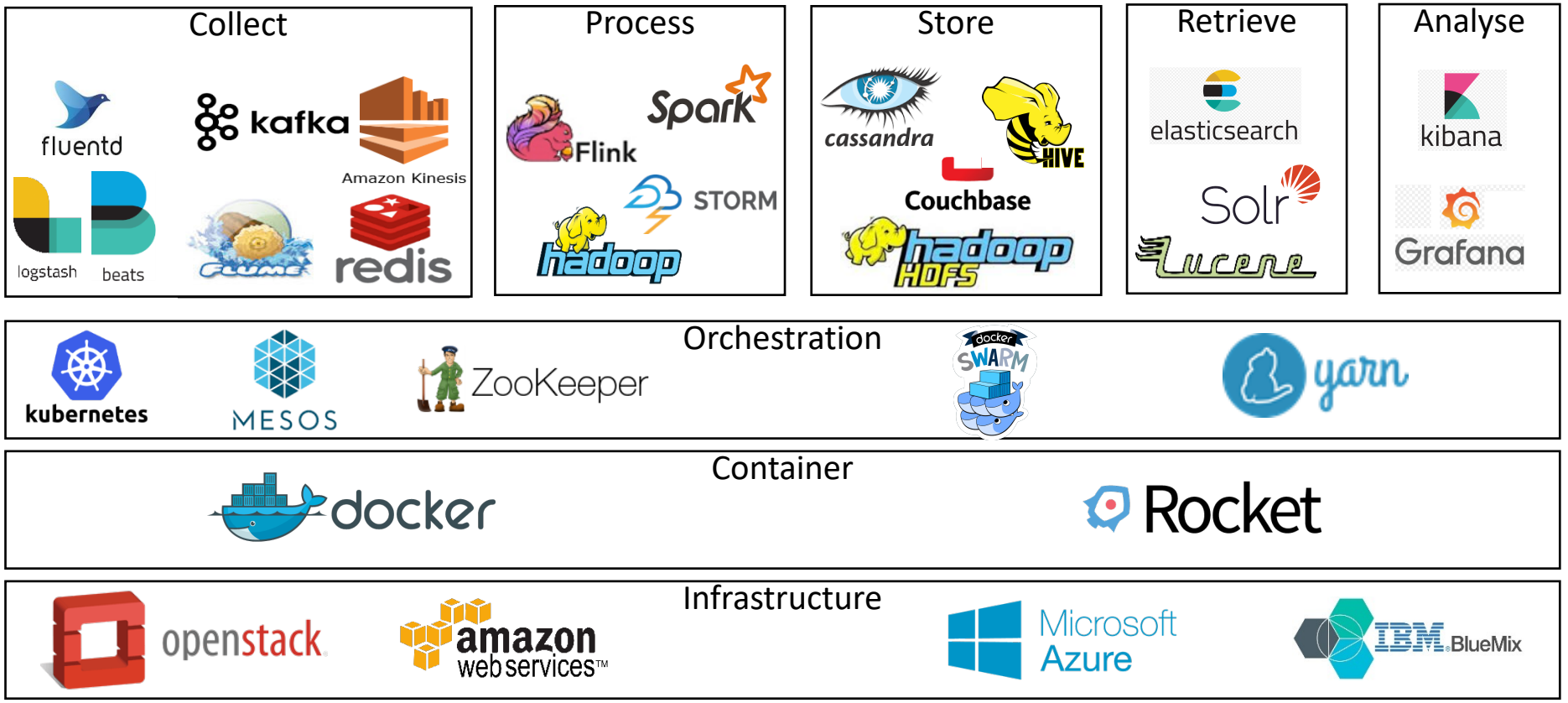
- Functional Requirements
  - Multi-tenancy
  - Massive Data Processing
- Non-functional requirements
  - High-throughput
  - High Availability
  - No single point of failure
  - Horizontal scalability
  - Low latency
    - In-memory MapReduce
    - In-memory database
    - Edge Computing
  - Security



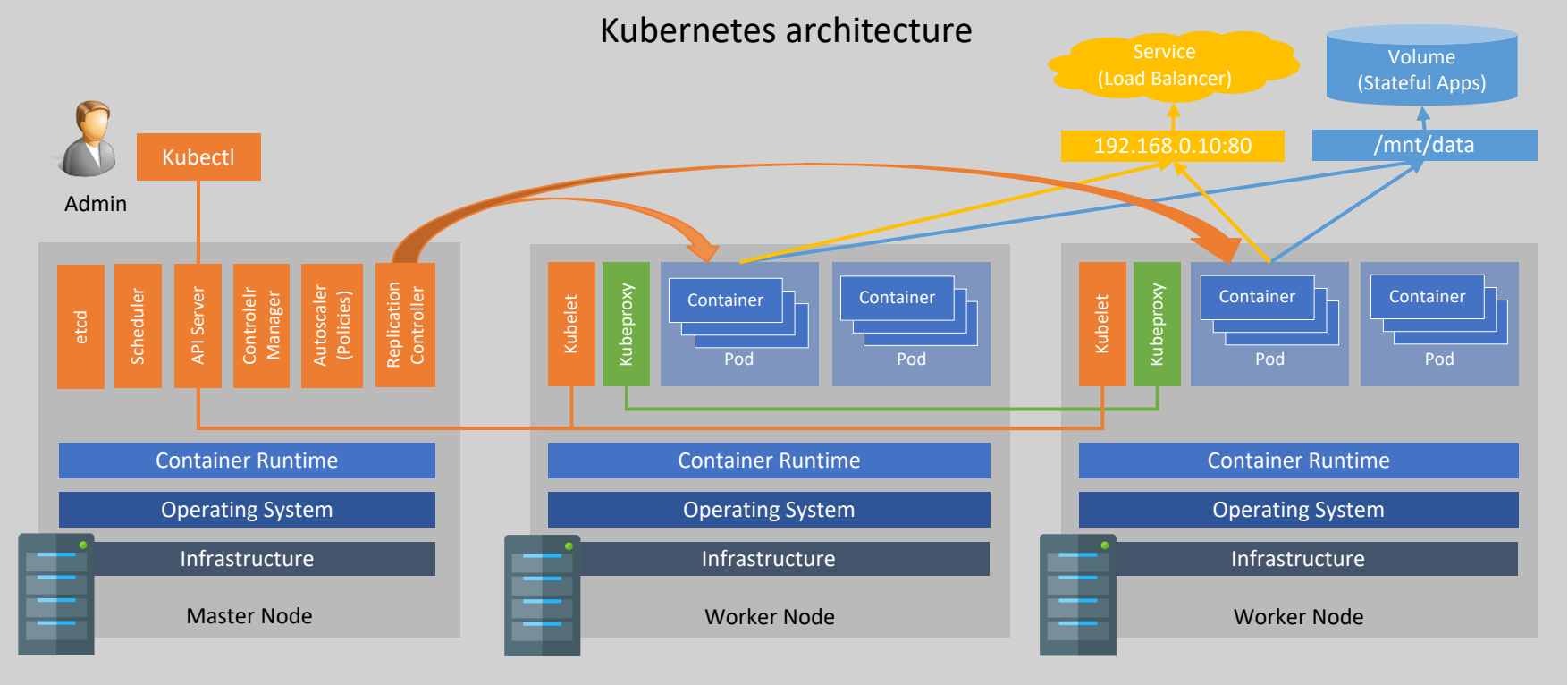


Source: STF 535 at EC-AIOTI WG03 Workshop

Relevant open source initiatives



Example, using Kubernetes



Value proposition: Run anywhere (Cloud infrastructure neutral platform), up and down autoscaling based on policies, load balancing, fault tolerance, replication control, automating deployment, management of containerized applications.

Edge cloud

- Is all of the above, but there are more specific challenges
  - **Discovery of, and optimal (re)-routing** to an edge cloud capable ofserving application clients (running on devices). When an application client wants to reach a server application, there's a need to discover the optimal edge cloud, one which runs instances of the server application, has the necessary resources (CPU, GPU, etc.) and provides the lowest network latency
  - **Service continuity:** when mobility occurs, it will be beneficial to seamlessly transfer the context from a stateful application instance in an edge cloud to an instance of the same application in a target edge cloud, e.g. one that provides a lower latency
  - **Smart application placement** to allow the optimized deployment of applications at the edge infrastructure based on criteria such as available resources, geographical areas, cost and latency requirements
  - **Cloud applications would enhance the user experience if they could leverage services offered by the network:** accessing information and services provided by the edge services such as the device location or QoS are key to enhance the user experience.
  - **Edge federation across multiple MNOs:** allowing MNOs to offer to the developers/customers the possibility to deploy their software across multiple domains and ensure service continuity when roaming on alternative networks.

Source: ETSI White Paper 36:

https://www.etsi.org/images/files/ETSIWhitePapers/ETSI_wp36_Harmonizing-standards-for-edge-computing.pdf

There's a role for standards bodies to play

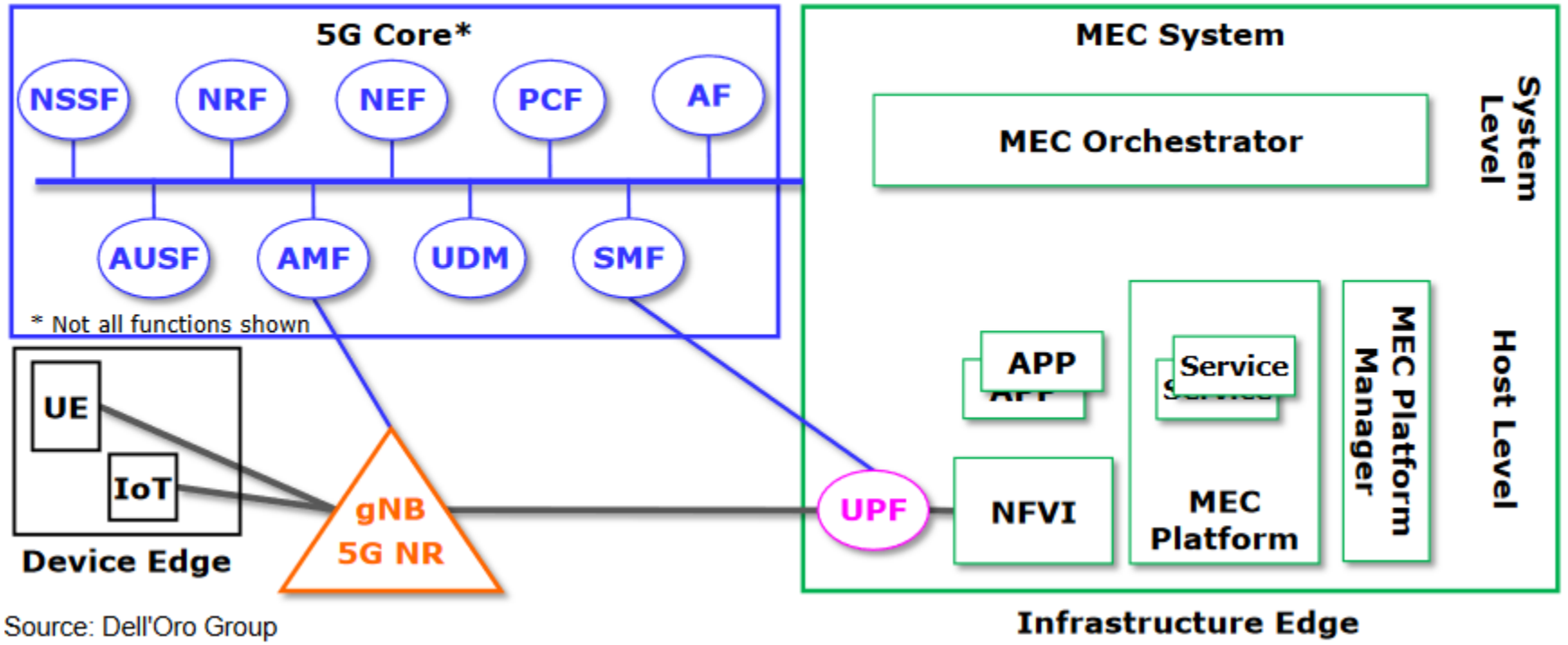
- **Multi stake holders:** user, MNO, cloud provider, application provider, etc.

- **Multi vendors:** device, apps, infrastructure, etc.

- Foundation standards in ETSI MEC and 3GPP
- Setting requirements: 5GAA, GSMA, etc.

Perfect fit for 5G

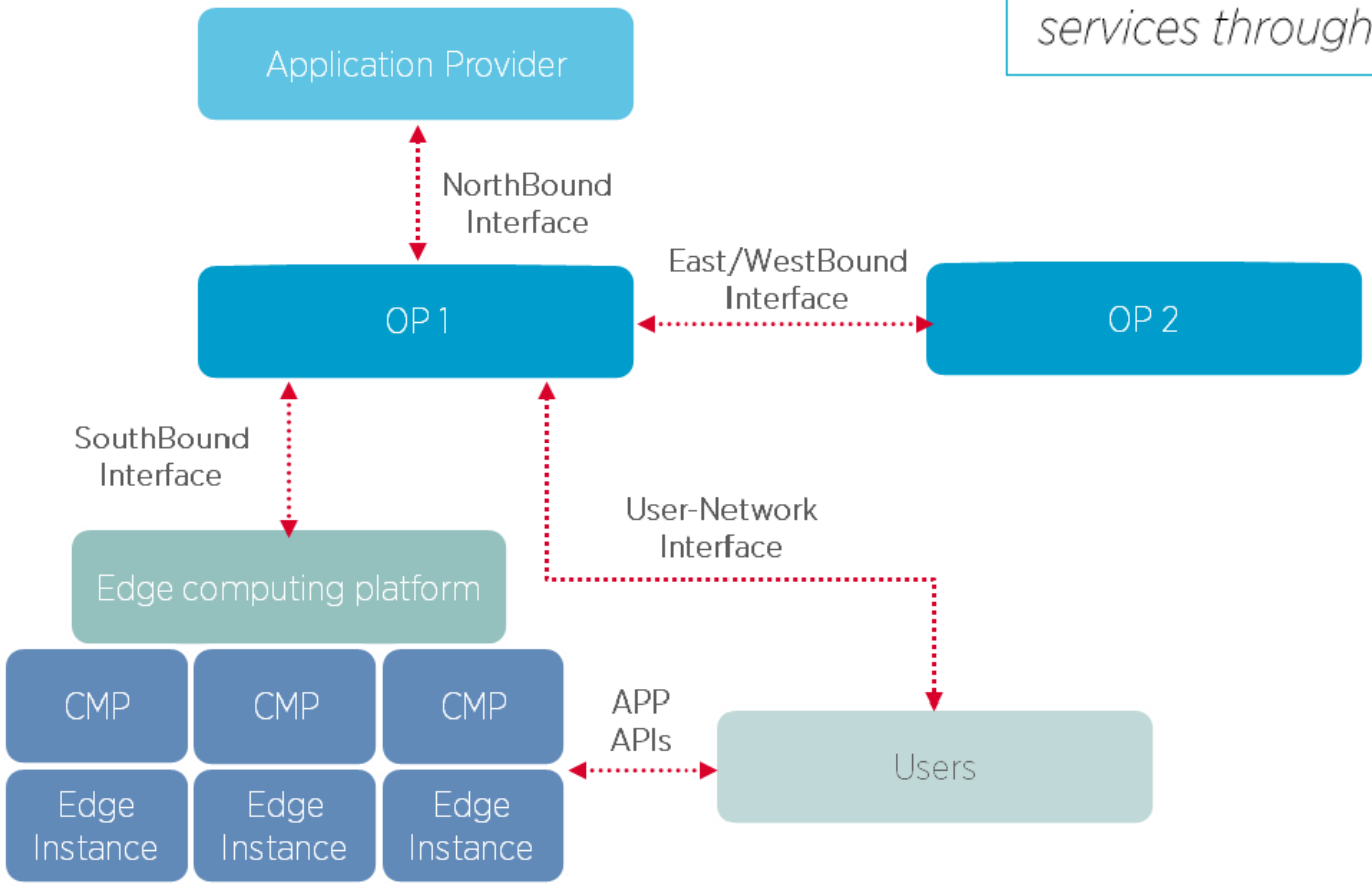
Figure 1: 5G MEC System Architecture Integrated into a 5G Network



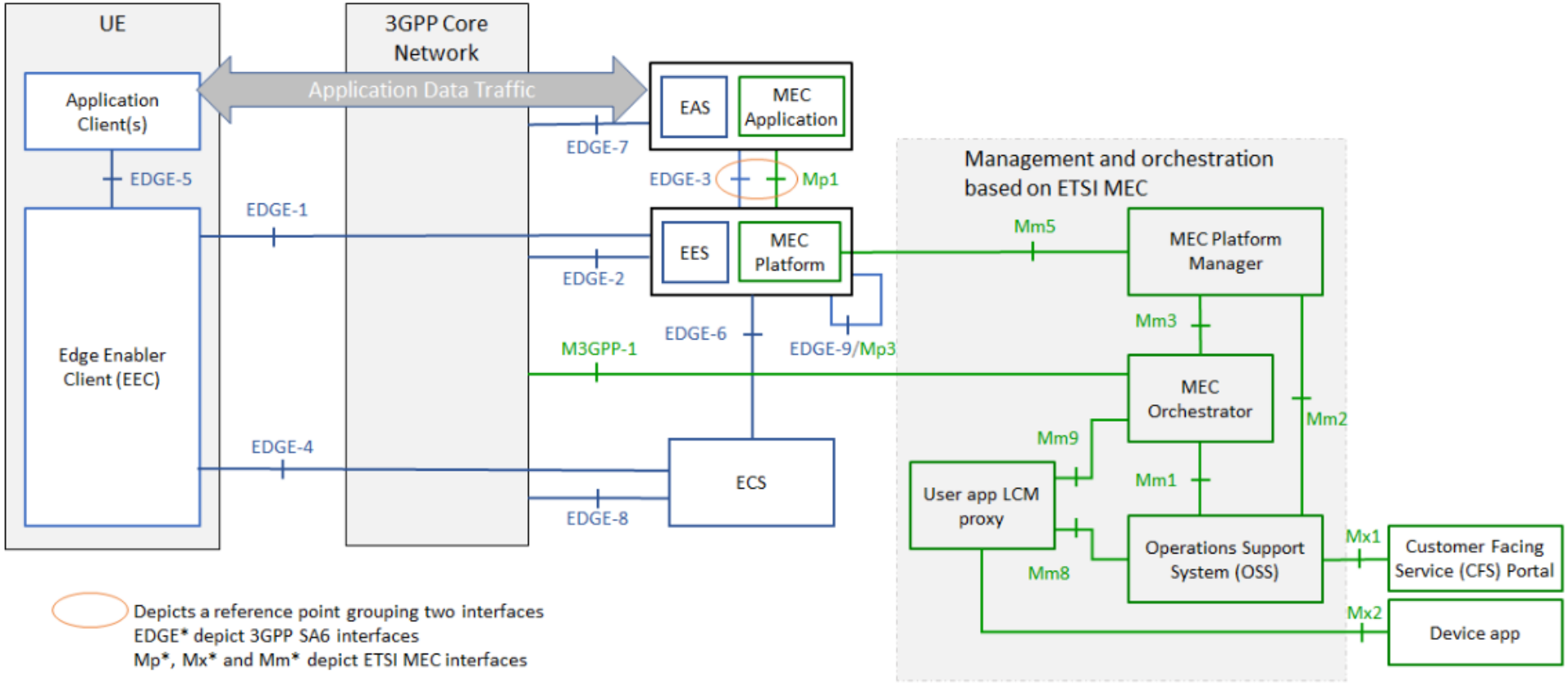
Source: Dell'Oro Group

Operator edge platform @GSMA

In Phase 1, the Operator Platform will federate multiple Operators' edge computing infrastructure to give application providers access to a global edge cloud to run innovative, distributed and low latency services through a set of common APIs.



Converge architecture for edge computing for MNOs



Value proposition: Run anywhere (Cloud infrastructure neutral platform), up and down autoscaling based on policies, load balancing, fault tolerance, replication control, automating deployment, management of containerized applications.

Concluding remarks

- Edge computing inherits the attributes of cloud computing
  - Certain constraints on resources
  - There are new challenges that have or being solved by standards:
    - Mobility, frequent mobility
    - Optimal routing
    - Context transfer (which might happen often because of resource constraints)
    - Federated edge platform

- Speaking about moving up the layers, data spaces, data sharing, federation of data marketplaces is a possible dimension.