Fog and edge computing in telecom networks

24 Feb 2017

Simone Mangiante

simone.mangiante@vodafone.com

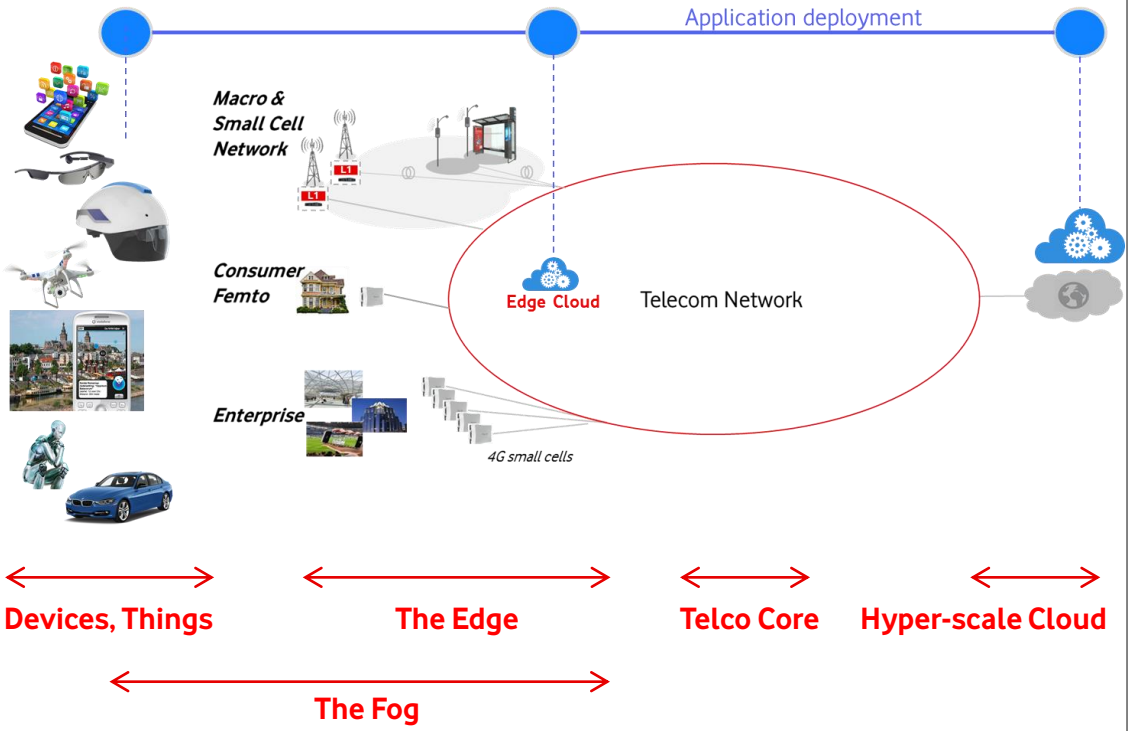
Group R&D

Agenda

1. Edge computing – What is it?
2. Use cases and drivers
3. Business rationale
4. Status, challenges and opportunities in edge computing
5. Conclusion and plans

Edge computing – What is it?

Edge Cloud Computing in a nutshell



Motivation

- Some mobile network functions cannot be fully centralised for a country or region.
- Some 3rd party applications don't work well when centrally deployed purely in a hyper-scale Cloud or on a device (e.g. poor responsiveness, not sufficiently real-time, poor battery lifetime of devices).

Solution

- Introduction of distributed cloud computing.
- Software deployment closer to devices, at the "Edge" of telecoms networks or in the "Fog".

Role of operator

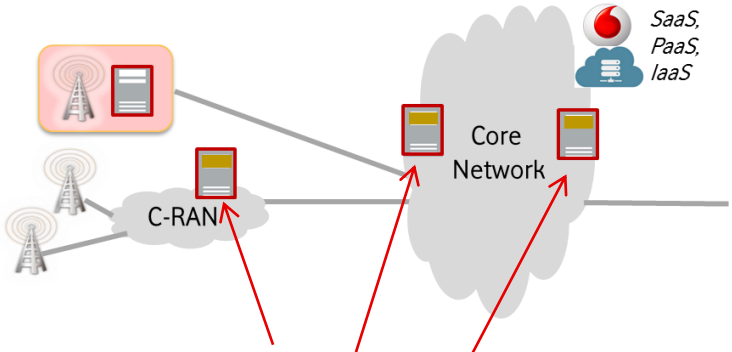
- Provides edge cloud computing infrastructure
- Hosts own software + 3rd party applications there
- Provides value-added services to hosted applications at the Edge through APIs
- Enables completely new applications and improves QoE for existing applications

The Move to the Edge – What is the Edge?

The edge differs for Tier 1 Cloud providers and network operators like Vodafone

Network Operator's Edge

Edge Locations: Potentially: Gi-LAN, BBU-hotel, access network hub, base station, fixed access network component, cable network component



Cloud computing servers at the edge

Gi-LAN: Operator's service LAN, BBU: base-band unit

Amazon's Edge

AWS Edge Locations: Amsterdam, The Netherlands (2), Dublin, Ireland, Frankfurt, Germany (3), London, England (3), Madrid, Spain, Marseille, France, Milan, Italy, Paris, France (2), Stockholm, Sweden, and Warsaw, Poland

Edge computing in more than 190 countries

Move to the Edge – What's the Motivation?

QoE, latency, bandwidth

Network operator's view

... to help our customers achieve

- better QoE,
- lower latency,
- higher throughput for delay-sensitive, CPU-hungry, real-time services,
- longer battery lifetime,
- more privacy,
- new 4G/5G services (for enterprises and consumers)

Amazon's view

"... to help our customers achieve

- lower latency and
- higher throughput, and
- to ensure that their data resides only in the Region they specify" [1)]

1) Src: http://aws.amazon.com/about-aws/global-infrastructure/

Structures of Edge Computing – 2 Tiers

Devices









Happening today

- ✓ AR applications (HP Aurasma, ...)
- ✓ 360 degree video (Nokia, ...)
- ✓ Live VR streaming (NextVR, ...)
- ✓ "Cloudlet" in locomotive (GE, ...)

On-premise
- Enterprise
- Residential

Access
- Fixed
- Mobile

Ramp to the Cloud
- Co-location providers
- CDNs

Cloud
- Public
- Private
- Hybrid
- Hyperscale

Structures of Edge Computing – Enterprise Edge

Devices

IoT Gateway
Enterprise Edge

Happening today

- ✓ IoT Gateways (HPE, ...)
- ✓ Enterprise edge (Cisco Edge Analytics Fabric + IBM Watson IoT, GE Digital Predix Machine + Predix, ...)

Cloud

- Public
- Private
- Hybrid
- Hyperscale

On-premise

- Enterprise
- Residential

Access

- Fixed
- Mobile

Ramp to the Cloud

- Co-location providers
- CDNs

Structures of Edge Computing – CDN Style

Devices

Happening since years

- ✓ CDN edge servers (Akamai, AWS, Azure, Google, ...)
- ✓ Edge computing of the finest

CDN

Cloud

- Public
- Private
- Hybrid
- Hyperscale

On-premise

- Enterprise
- Residential

Access

- Fixed
- Mobile

Ramp to the Cloud

- Co-location providers
- CDNs

Structures of Edge Computing – Cloud Hosting Hierarchy

Devices









IoT Gateway
Enterprise Edge

Emerging
- ✓ Real-time communication with AR/mixed reality, virtual avatars
- ✓ Equinix, EdgeConneX, ...

Co-location provider

Cloud
- Public
- Private
- Hybrid
- Hyperscale

On-premise
- Enterprise
- Residential

Access
- Fixed
- Mobile

Ramp to the Cloud
- Co-location providers
- CDNs

Structures of Edge Computing – MEC and Telco Style

Devices

Mobile Edge Computing

CORD Fixed Access

Being worked on

- ✓ ETSI MEC in progress
- ✓ Distributed cloud in 5G
- ✓ CORD for distributed cloud
- ✓ Fixed Access NFV

Cloud

- Public
- Private
- Hybrid
- Hyperscale

On-premise

- Enterprise
- Residential

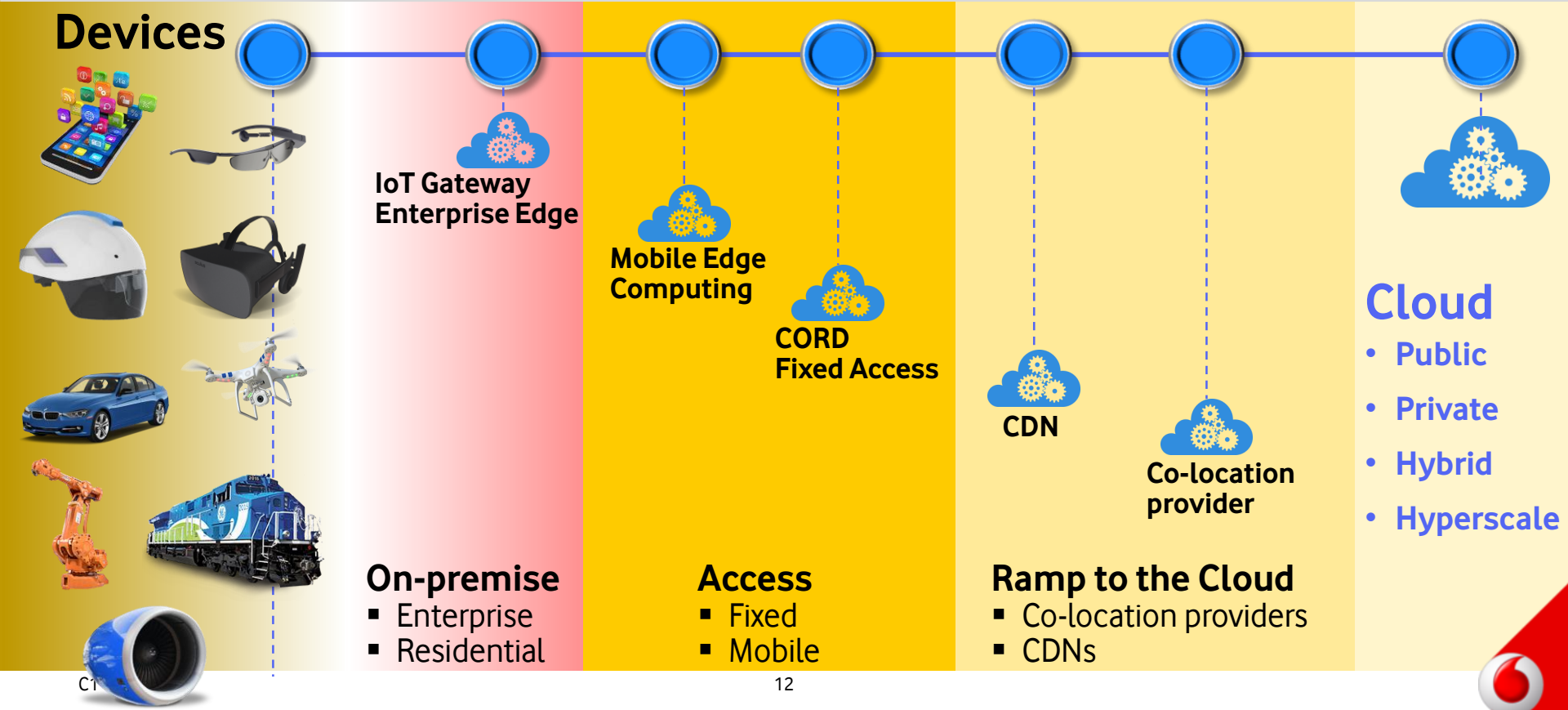
Access

- Fixed
- Mobile

Ramp to the Cloud

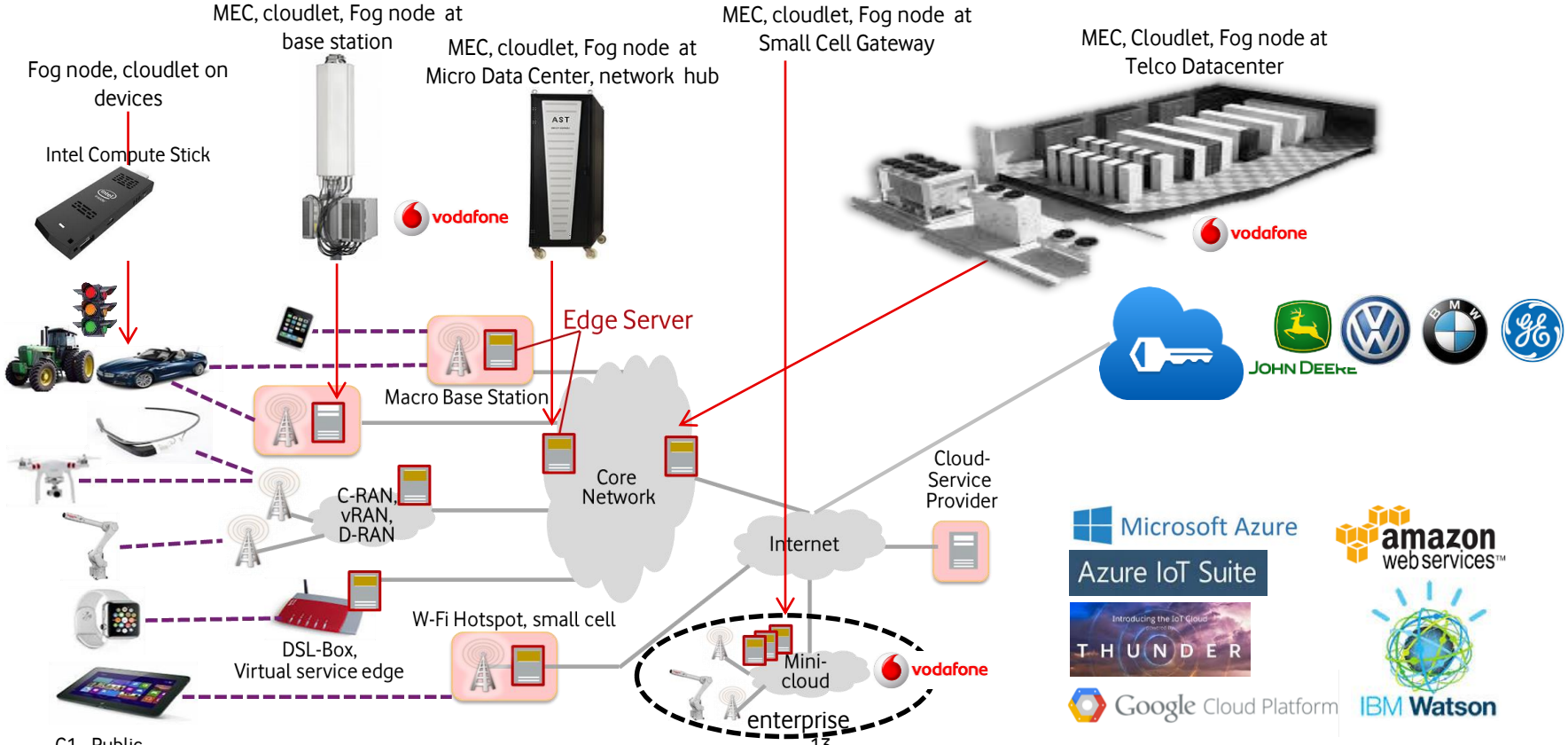
- Co-location providers
- CDNs

Structures of Edge Computing – Multiple Variants Exist



Future Possible Deployment of Edge Cloud Computing

Below physical view is relatively independent of flavour (Fog, Cloudlets, MEC servers)



Main drivers and use cases

Trends that Underpin the Emergence of Edge Computing

- CDNs
- IoT Gateways
- Enterprise NFV products
- Small cell controllers with open compute platforms
- Routers and switches with open compute platforms
- Mobile Edge Computing as defined by standards body ETSI ISG MEC
- Fog Computing as per OpenFog Consortium
- Cloudlets (micro datacenter)
- Industrial edge computing
- Moves to extend the hyperscale cloud
- Distributed telco cloud in 5G

Drivers for Edge Computing and Sample Use Cases

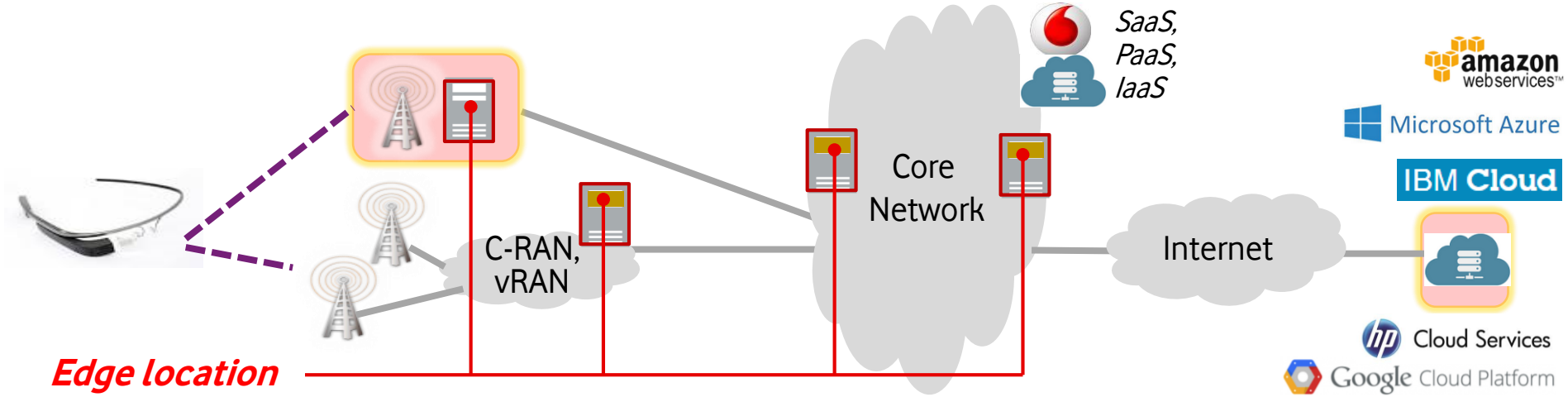
Main drivers in the next 5 years

- Development of low-latency 5G radio interfaces
- 5G use cases that require low latency (e.g. in automotive, industry 4.0, health: cognitive assistance, smart grid)
- Convergence Mobile/Cloud (see Telco datacenters)
- Wearable hardware (MS HoloLens, smart glasses)
- Progress with implementations of cognitive algorithms (face recognition, translation etc.)
- Progress with implementations of real-time analytics and big data

Use cases

- Security, safety data analytics
- Intelligent video analytics
- Operational analytics for IoT
- Local content caching
- Device location tracking
- Gaming
- Cloud robotics
- Augmented reality, assisted reality, cognitive assistance, mixed reality
- RAN-aware content optimisation
- Mobile video delivery optimisation using throughput guidance from the radio access

Move to the Edge – Which Software May Sit at an Operator's Edge?



Edge location

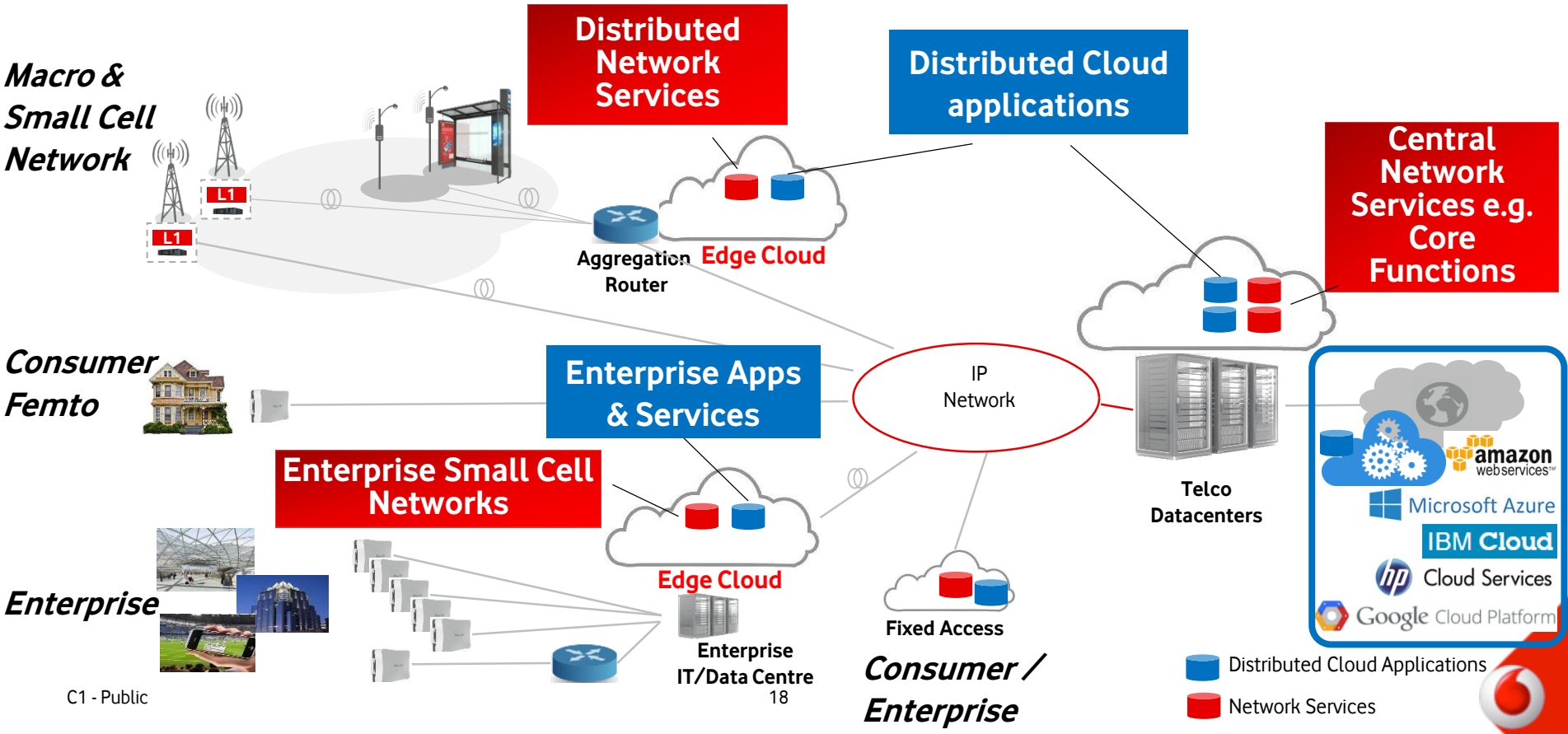
Application software at the edge

- Face recognition
- Object recognition
- Motion classifier
- Activity inference
- Augmented reality (object and data caches)
- Video real-time analytics, video mgmt (e.g. in stadium)
- RAN-aware applications
- Content optimisation and others

- Content servers
- Video storage
- Databases
- Enterprise integration
- …

A mix of different applications opens up compelling service opportunities

Distributed Cloud Computing With the Help of Telecoms Operators



Latency Requirements Vary Across a Broad Range

What counts is response time on application level

≤ 1ms

- **Remote control / telepresence** with real-time, synchronous **haptic** feedback
- **Industrial moving robots**
- **Industrial closed loop** control systems (e.g. 1ms cycles of polling data from sensors + actuators)
- **Negotiated automatic cooperative-driving** manoeuvres
- **Smart grid**: synchronous co-phasing of power suppliers (< 1ms)

≤ 10ms

- **Shared Haptic Virtual Environments**: several users perform tasks that require fine-motor skills
- **Tele-medical** applications (e.g. tele-diagnosis, tele-rehabilitation)
- **Augmented reality**
- **Education: Haptic overlay** trainer / learner for fine motor skills (e.g. for medical)
- **Smart grid** (3ms)
- **Process automation** (5ms)

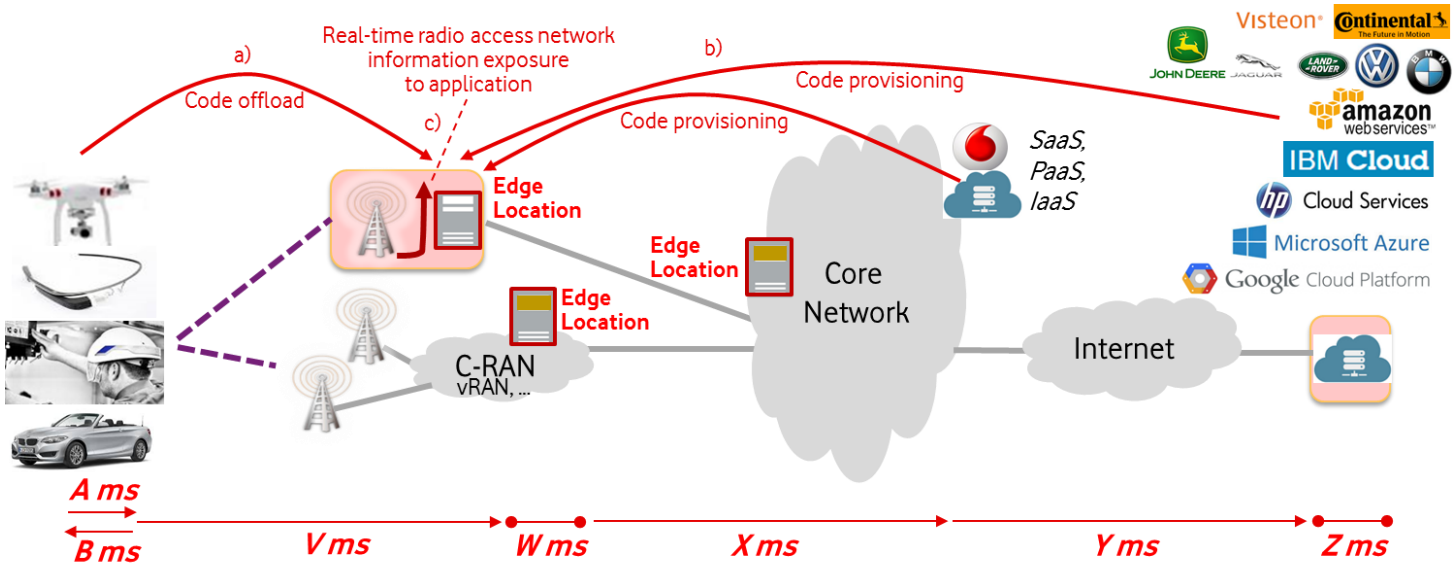
≤ 50ms

- **Serious gaming** (20ms)
- **Cognitive assistance** (20-40ms)
- **Virtual reality**
- **Cooperative driving** (20ms)
- **UAV control** (10 - 50ms)
- **Remote robot control** with haptic feedback (25ms)

≤ 100ms

- **Vehicle safety apps** (mutual awareness of vehicles for warning/alerting)
- **Assisted driving** (cars make cooperative decisions, but driver stays in control)

… And It's Latency on Application Level That Matters



In direction to the Cloud

A	sensor + embedded computing on device
V	radio interface
W	base band processing at edge
X	transmission through Core network
Y	transmission across Internet

Z processing in the Cloud

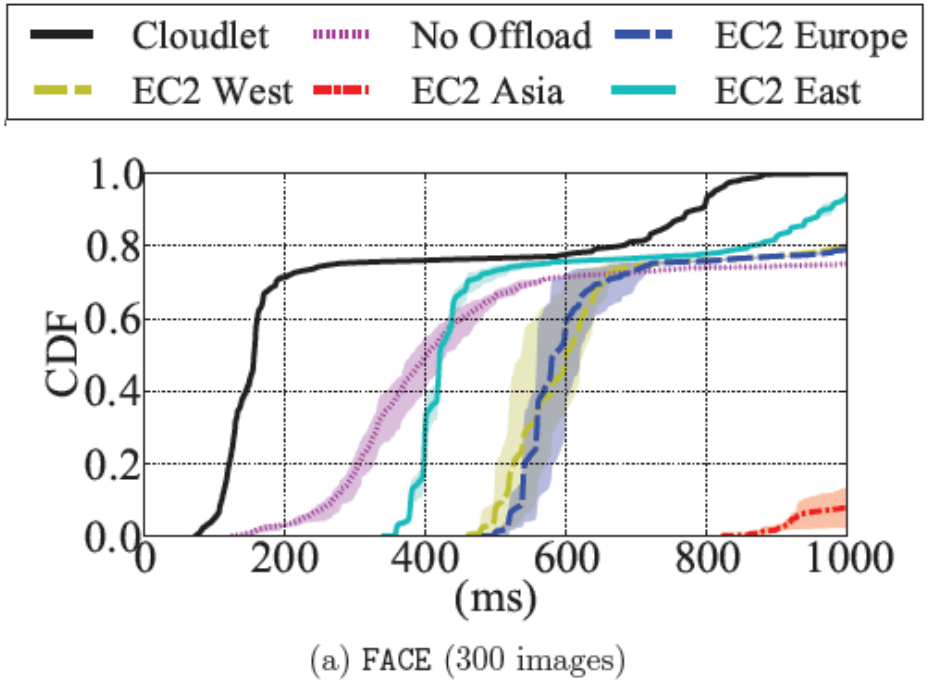
Back from the Cloud

B	embedded computing, actuator, UI presentation, human reaction time
V	radio interface
W	base band processing at edge
X	transmission through Core network
Y	transmission across Internet

Latency Reduction via Edge Computing Makes a Difference

Measurements for Face Recognition Application → A deployment of server code at an edge cloud in closer proximity to the end user devices clearly wins with lower latency.

CDF: cumulative distribution function. Better curves are to the left



(a) FACE (300 images)

Experiment set-up:

- Client server application
- Location of device: Pittsburgh, closest to AWS EC2 East (USA east coast)
- EC2: Amazon Web Services locations
- Cloudlet: is the local micro datacenter
- No offload: server component is run on the device itself (but lack of computing power)
- Radio access: LTE

Cumulative Distribution Function of application response time (milliseconds)
CDF = 0.5: is the median value

Source: Carnegie Mellon University, Prof. Satya

Vodafone Areas of Research and Proof of Concepts

- **ETSI:**
  - ETSI MEC PoC #8: Video Analytics
  - ETSI MEC PoC #9: Low latency industrial IoT
- **Vodafone R&D:**
  - Automotive
  - YouTube traffic offload at the edge
  - Edge platform for IoT
  - 360 video and VR streaming
- **Academia:**
  - Carnegie Mellon University: cloudlets and cognitive → Open Edge Computing

Business rationale

Edge Computing – What's in for Whom

Example

Tier 1 cloud providers: extend the Cloud

- Extend cloud computing to the 'last mile' (no presence there today)
- Thereby can offer better SLAs and QoE for their customers

Enterprises: new services & cost reduction

- Higher automation, real-time analytics
- More advanced autonomous, sensory systems, cognitive assistance,...

CSPs, Network Operators: expand the business

- Expand the scope of the smart pipe. Leverage Cloud, NFV/SDN, 5G
- Add Compute, Store to what is already in place: Networking

Infrastructure vendors (telecoms and IT)

- Leverage NFV/SDN and orchestration for edge cloud
- Provide the edge cloud software infrastructure

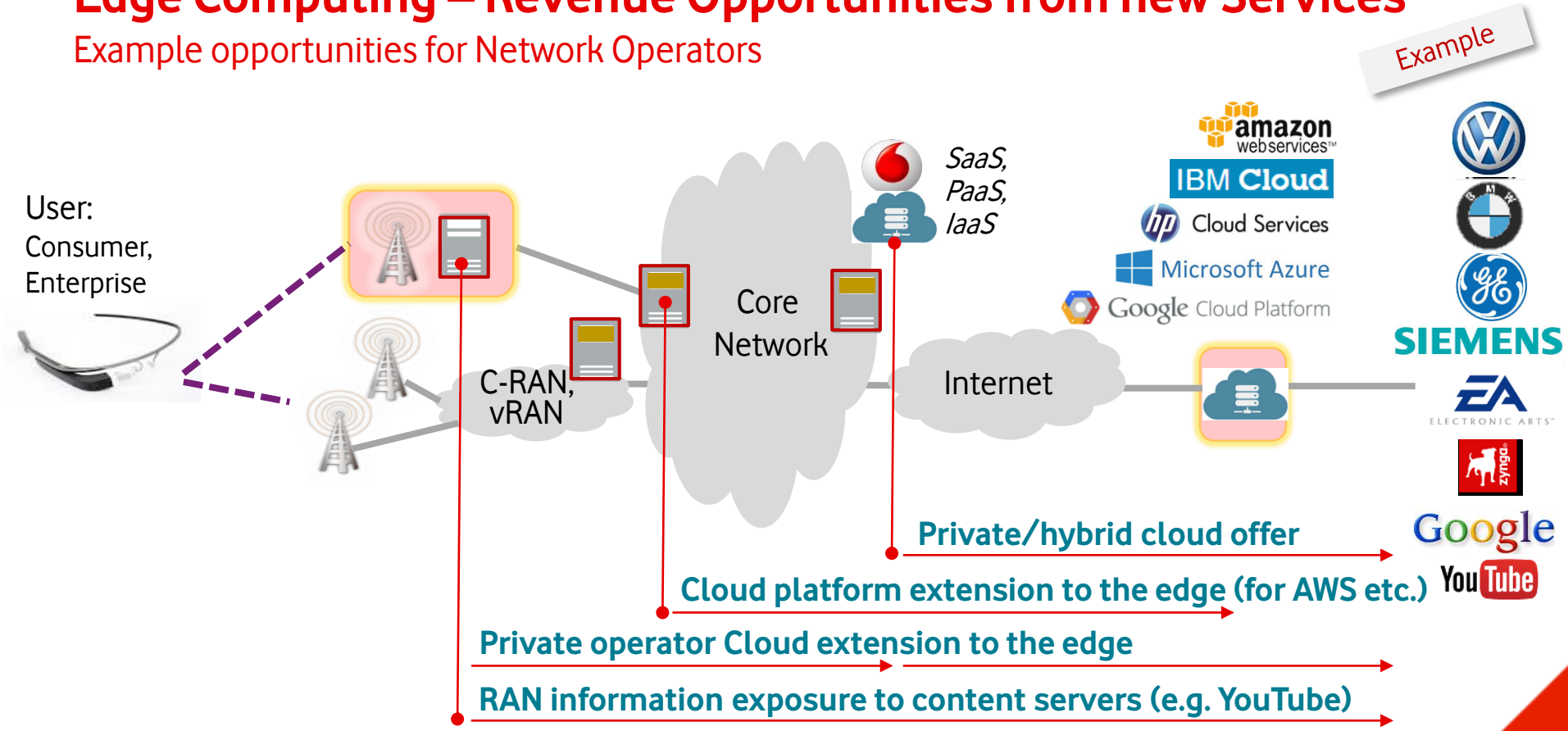
etc.

Why Edge Computing Supported by Telecoms Networks?

1. A new business opportunity for telecoms service providers
2. A natural way to support mobile devices, mobile "things" of the IoT
3. Cost synergies (e.g. cloudification of the Radio Access Network, Core, Central Office)
4. Unique value added from context awareness (e.g. radio access, location)
5. An enabler for various 5G use cases

Edge Computing – Revenue Opportunities from new Services

Example opportunities for Network Operators



Status, challenges and opportunities in edge computing

Technologies for Edge Computing

- **Fog computing** by Cisco, Princeton and publicised in 2012, see OpenFog Consortium

- **Cloudlets**: Carnegie Mellon University (CMU) has developed Cloudlets starting in 2009.
  - A cloudlet represents the middle tier of a 3-tier hierarchy: 'mobile device – cloudlet – cloud.
  - A cloudlet is a "data center in a box" or micro-data center .
  - In some situations, a cloudlet can also be co-located with a device itself (e.g. car, aeroplane).
  - Microsoft Research use both the notion of cloudlet and micro-data center interchangeably.
  - Vodafone Group R&D run open source cloudlet software in the R&D Lab.

- **MEC**: ETSI has created the Mobile Edge Computing architecture through the industry specification group MEC, which was co-founded by Vodafone and others in Sept 2014 .

- **Micro datacentres**: Microsoft Research published about the concept of micro datacentre . microDC is intended as an extension of today's hyperscale cloud datacentres (Microsoft Azure). Refers to Cloudlets. Purpose: to meet new application demands like lower latency and lower device battery consumption.

Challenges in Edge Computing

- Some use cases are only emerging
- Application developers lack awareness of emerging distributed cloud and stakeholders
- Collaboration of cloud hosting providers and CSPs for ‘edge to main cloud integration’ is lagging behind
- Standards for network API exposure at edge clouds not finished yet
- Edge Cloud APIs not widely available in SDKs, language libraries, developer tools, cloud management software
- Deployment mechanisms of code into edge clouds not well defined yet
- Integration with SDN/NFV/OSS

Open Issues - Generic

Follow-up questions to address

- Is there today already a link between **OSS/MANO for NFV and management of Cloud platforms** in a data centre? What are the implications for Cloud computing at the edge?
- Which are relevant **legal requirements** for cloud computing applications if any? E.g. need to consider Lawful Interception requirements from telecoms regulators.
- What are the **security and privacy requirements** for edge computing?
- Are there any Cloud-native applications available which could be used to **test** the edge computing concepts?
- How could some **open labs** be set up that enable developers/start-ups to experiment with edge cloud computing in a flexible way?
- The Cloudlet concept is based on Virtual Machines e.g. for security reasons. What would be the equivalent concept using **containers**, what would be the equivalent of real-time VM migration with Docker or other containers?
- Are the **top Cloud providers** working on new, innovative workload placement mechanisms too?

Open Issues – Architecture and technology

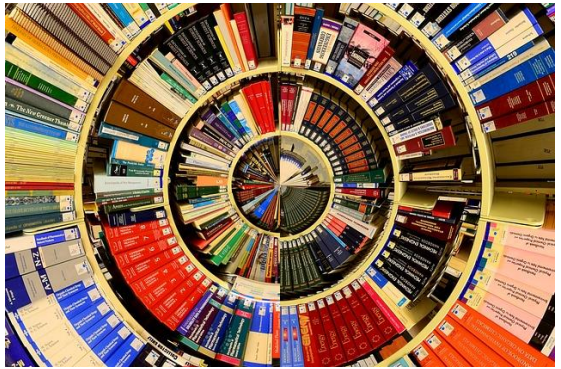
Follow-up questions to address

- How **robust and reliable** is the edge software architecture? **Resilience** solutions?
- How to **scale** at the edge of the network?
- How could edges be **packaged** up (sizing and automatic deployment) for consumer/enterprise markets?
- How to **orchestrate** cloud software at the edge together with the rest of the network?
- Which **other architectural approaches** exist related to edge-cloud?
- Best approaches to **integrate edge clouds with private and public clouds** for the benefit of developers?
- Is **OpenStack** 'the' opportunity to achieve portability of applications across (tiered) cloud computing locations including the Edge? Fit for purpose?
- What technology shall we use for **live VM migration across locations** ?
  We believe current datacentre live VM migration is not fit for purpose for edge computing and mobility.

Conclusion and plans

Summary and Conclusions

- Edge Computing celebrates **first successes**
- Edge Computing can be **enabled by telecoms** service providers
- **Application developers** to look beyond vertical silos
- Edge Computing **platforms + APIs** need **to be introduced to developer communities**: simple, easy
- Need **more experimentation** for commercial use cases: PoC, tests, trials, data points, ...
- **Technical challenges** need to be addressed → research, open source



Synergies with OpenFog

- Cloud integration
- Deployment and scalability
- New use cases and trials
- Horizontal testbeds
- Security and privacy
- Multi-tenancy
- Harmonization with other initiatives (IIC, ETSI, Open Edge Computing, ...)

Thank You

simone.mangiante@vodafone.com