IoT Edge Computing
Survey and Gap Analysis

Xavier de Foy (InterDigital), Dirk Kutscher (Huawei), Eve Schooler (Intel),
Debashish Purkayastha & Akbar Rahman (InterDigital)

T2TRG @ IETF 100 – Singapore

Introduction

- IoT today means handling a massive amount of data generated at the edge, reversing today's dominant data flow
  - Solutions involve edge computing, either at small scale data centers or further towards the edge
- This presentation covers
  - Motivations for IoT Edge Computing
  - An overview of major industry and research directions in this domain
  - Potential areas for future work based on the direction of those projects
- Its goal...
  - Gather input from the community, in particular to identify challenges that are relevant to IRTF

Motivations for IoT Edge Computing (see ref: [1])

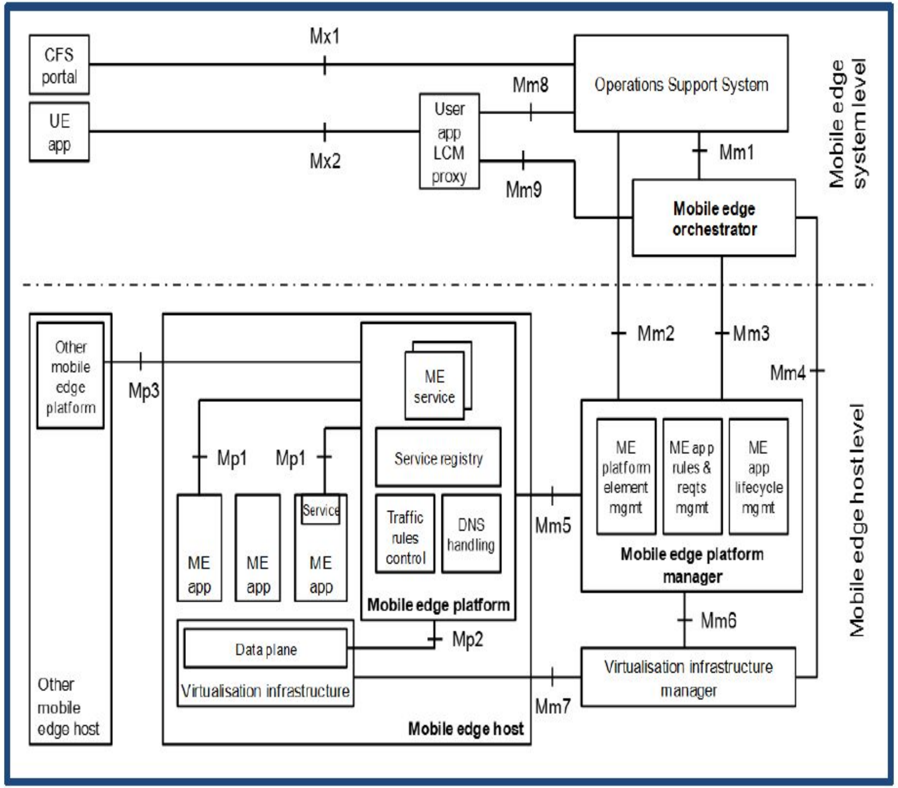
- Support high data volume at the edge
- Support highly time- and trust-sensitive applications
- Exploit opportunities for energy efficiency and cost reduction
- Adapt to intermittent connectivity
- … and “Open” the edge...
  - By enabling multiple providers to offer competing edge computing services
  - By enabling open (and secure) access to data
  - By enabling open (and secure) access to computing resources, e.g. multi-tenancy

Survey 1/6

Telecom Industry-Related Initiatives

Projects:

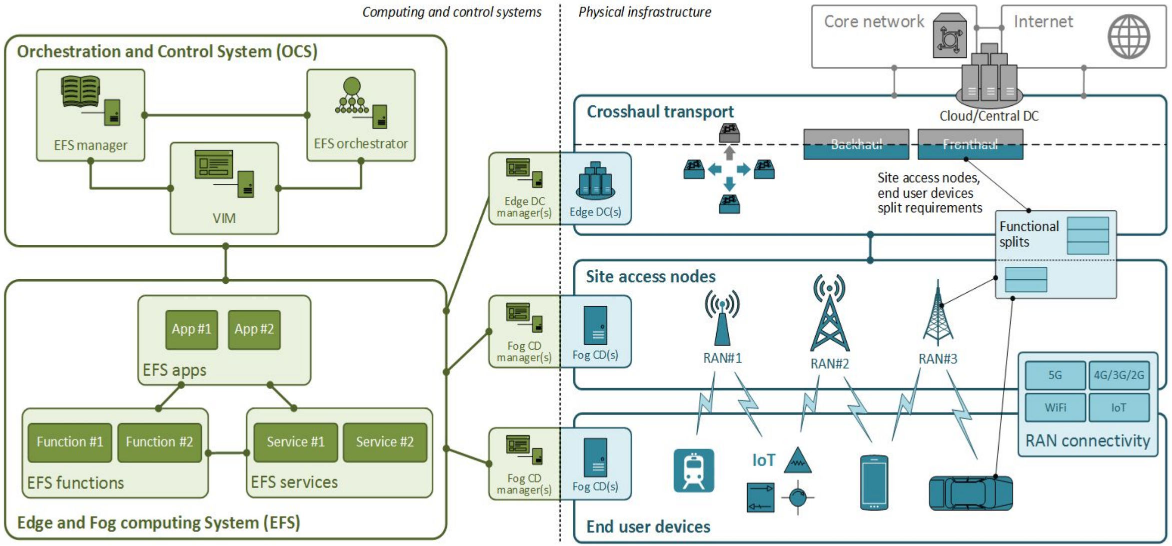
- ETSI Multi-Access Edge Computing, MEC* (architecture, services…)
- OpenEdgeComputing (cloudlet-based architecture for offloading)
- Telecom Infra Project (TIP) has an Edge Computing project, focuses on use cases
- 3GPP builds in some support for Edge Computing in 5G
- M-CORD, Mobile (networks) Central Office Re-architected as a Data center, integrates edge computing as a part of the 5G architecture
- ...



Ref: [2]

Survey 2/6
Telecom Industry-Related Initiatives

- …
- 5G Convergent Virtualized Radio Access Network, 5G-CORAL (research, combines telecom edge computing and fog)



Ref: [3]

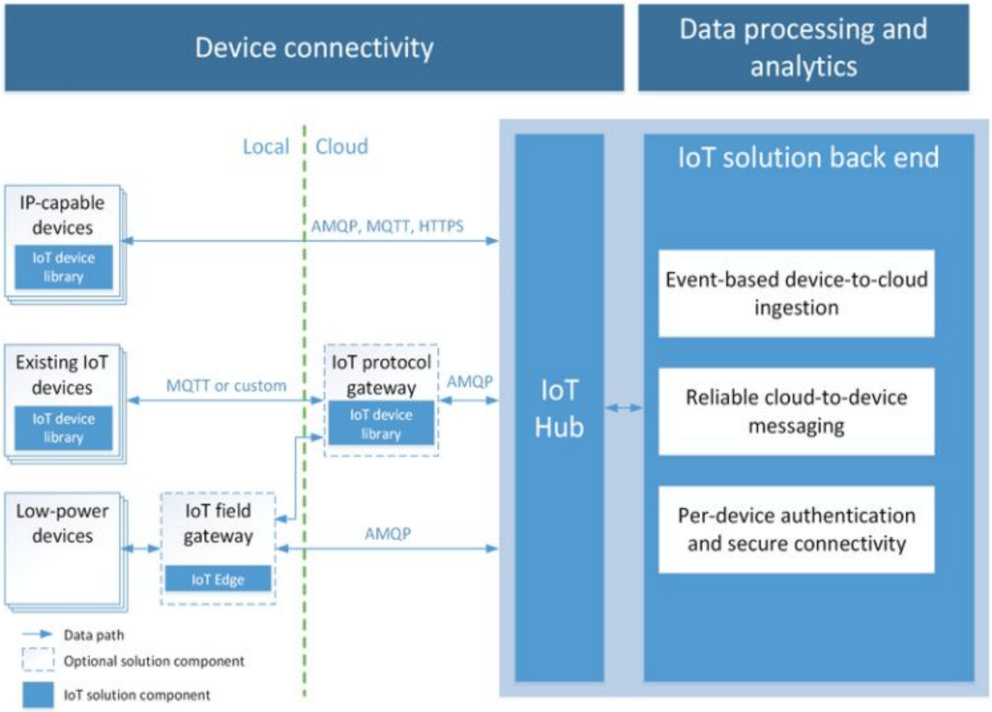
- Driven by the Telecom industry
- Integrate with Telcos' networks
- Technically: often NFV-based
- Evolving towards distributed computing, lightweight/fog computing

Survey 3/6
Intelligent IoT Gateway Model

Products & Projects

- Bosch (Prosyst gateway software),
- Siemens (IoT 2000 serie gateways),
- Microsoft (Azure IoT Edge)*,
- Amazon (Greengrass and Snowball Edge),
- EdgeX Foundry (IoT gateway open source)
- ...

- Originally: simple model, typically single tenant gateway running pre-provisioned code for data processing at the edge
- Typical protocols HTTPS, MQTT, AMQP, COAP, Modbus, OPC UA, DDS, etc.
- Different levels of control (and complexity) for application developers (from full control of embedded Linux gateways to high level programming model with Greengrass)

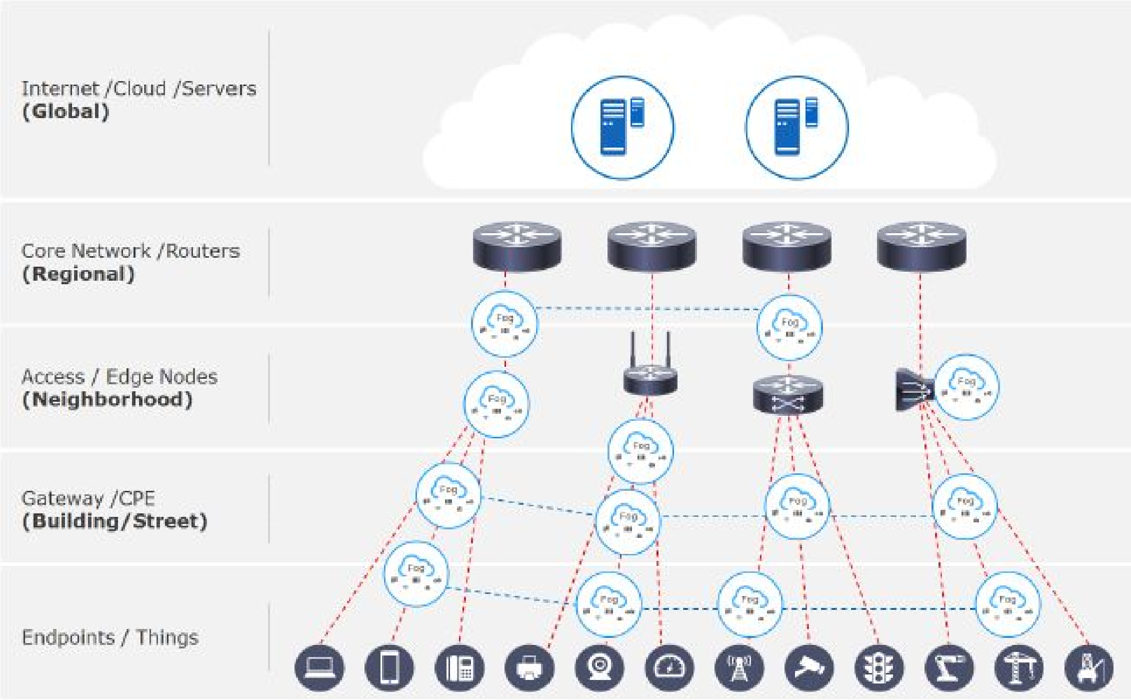


Ref: [4]

Survey 4/6

Intelligent IoT Gateway Model

- ...
- OpenFog (architecture): integrates computation, networking (including time sensitive networking), storage, control and acceleration
    - Linked to new IEEE P1934 WG on Fog Computing and Networking Architecture Framework
- Typically coupled with a distant cloud service through a client-server model and evolving towards clouds of clouds:
    - Gateways can be connected with each other in a tiered fashion,
    - East-west connectivity is envisioned in OpenFog



Ref: [5]

Survey 5/6
Emerging Trends
(Lightweight in-Network Computing)

Products & Projects

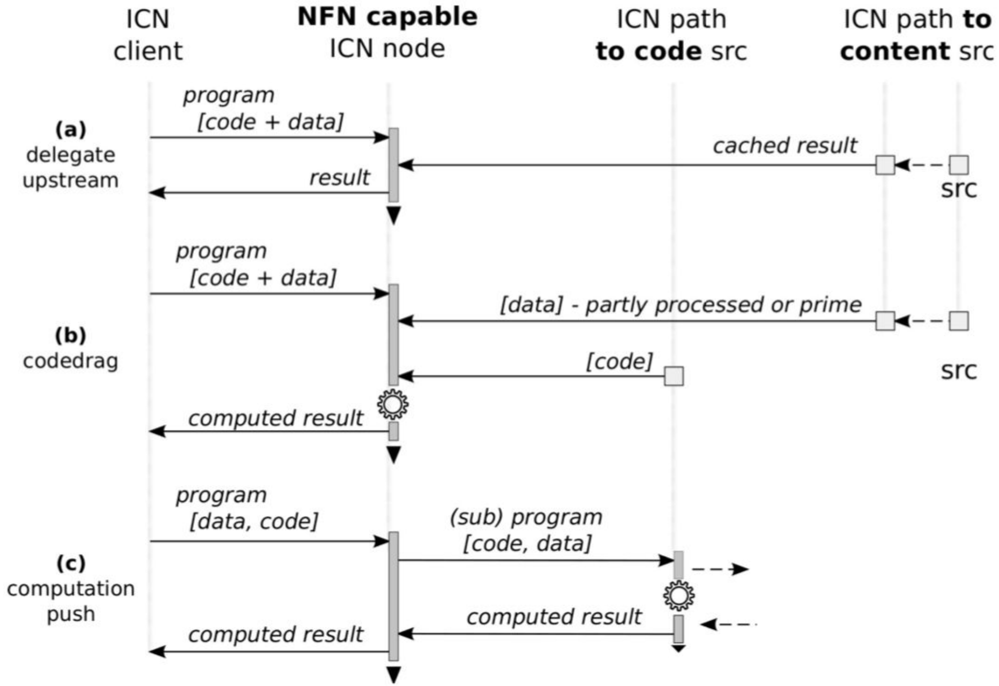
- Serverless computing products: Cloudflare Worker, Amazon Greengrass
  (exploit statelessness to decorrelate service from server location)
- Named Function Networking (NFN) ⇒ Next Slide
  (ICN-derived, more granular approach)
- Facility for Large-Scale Adaptive Media Experimentation, FLAME
  (transports IP over ICN, service routing in ICN layer to optimize traffic at the edge)

- Stateless functions are easier to dispatch to any server, local or remote (no state to carry over), enabling advanced data/service routing to be developed
- ICN technologies support intermittent connectivity, dissemination of local data in a manner less tied to applications.

Survey 6/6

ICN Named Function Networking (NFN)

- Accessing static data and dynamic computation results in one data-oriented framework
  - Benefitting from usual ICN features: data authenticity, caching etc.
  - Enabling network to perform various optimizations: move data to code etc.
- Enabling (secure) access to individual elements within Named Data Objects
  - Selective access to possibly pre-processed data elements
  - Difficult to do securely in a circuit-based edge computing model



Refs: [6], [7]

Initial Analysis

- **Computing devices**: mostly gateways, mini-data centers - not yet end devices
- **Computing models**: stateful (VMs or containers), stateless (serverless, NFN)
- **Communication models**: publish-subscribe (message-based, NFN/ICN)
- **Network traffic patterns**: mostly high volume upstream with throttling by edge computing devices (or deferred to off-peak hours or using physical shipping) - downlink for control, sw updates…
- **Storage models**: local storage on gateways, external DB in cloud or local IT cloud, distributed data (ICN)
- **IoT edge computing services**: protocol translation, analytics, transcoding, etc.
- **Management of EC**: often cloud-based/NFV Management and Orchestration

Gaps 1/3

- IoT EC evolves towards a **distributed computing model** leading to new challenges due to the dynamic and constrained environment at the edge:
  - On creating local cloud federations: protocol(s) to declare availability (e.g. m-DNS beyond local area, interaction with CoAP), status, capabilities, associations (with federations). Among other questions, where does the control live?
  - On creating clouds of clouds, which can find and help each other, share resources, make themselves available for federation , …
  - On operating edge clouds, e.g. determining an optimal placement for caching and computing considering service acceleration and other resource capabilities, routing of data and service requests, ...

Gaps 2/3

- IoT EC also evolves towards a more **open model**, with new challenges related to:
  - Open access to (managed, unmanaged or self-managed) compute/storage resources, e.g. using generic APIs for developers to ask for resources meeting specific requirements.
  - Open access to data to liberate data from “silos”, e.g. with APIs for data & meta-data lifecycle management, access control, auditing, managing impact on privacy.
  - Multi-tenancy, e.g. providing fair and secure allocation of resources to tenants.

Gaps 3/3

- IoT EC evolves towards some **lower end devices**, as they become more powerful, in part due to hardware acceleration (see Mist Computing definition from NIST [8])
    - Requiring dynamic and lightweight cooperation of things, edge devices, compute platforms
- This evolution can increase the challenge of **QoS at the edge**
    - Requiring dynamic network slicing (including multi-domain aspects)

Next Steps

After having reviewed technologies related to IoT EC, and looked at potential gaps:

(1) We are seeking help from the community to gather more input...

(2) ...and to brainstorm about what the IRTF should do…

- Is T2TRG the right research group?
  - Possibly with regards to the distributed and lightweight aspects of some challenges
- Which challenges are most related to IRTF?
  - Internet-related protocols are not within the scope of other organizations looking at IoT EC

References

[1] D. Kutscher, E. Schooler, T2TRG IETF-99 presentation, IoT Edge Computing Discussion @ IETF-98;

[2] ETSI MEC, Framework and Reference Architecture document; GS MEC003

[3] 5G-CORAL portal http://5g-coral.eu/

[4] Microsoft Azure IoT Edge, portal https://azure.microsoft.com/en-us/campaigns/iot-edge/

[5] OpenFog Consortium, OpenFog Reference Architecture for Fog Computing

[6] C. Marxer, C. Tschudin; Improved Content Addressability Through Relational Data Modeling and In-Network Processing Elements; ACM ICN 2016

[7] M. Sifalakis et al.; An Information-Centric Network for Computing the Distribution of Computations; ACM ICN 2014

[8] The NIST Definition of Fog Computing, Aug 2017, https://csrc.nist.gov/publications/drafts/800-191/sp800-191-draft.pdf