Edge Computing in 5G & Beyond

Abhay Karandikar

Director, Indian Institute of Technology Kanpur, Kanpur, India

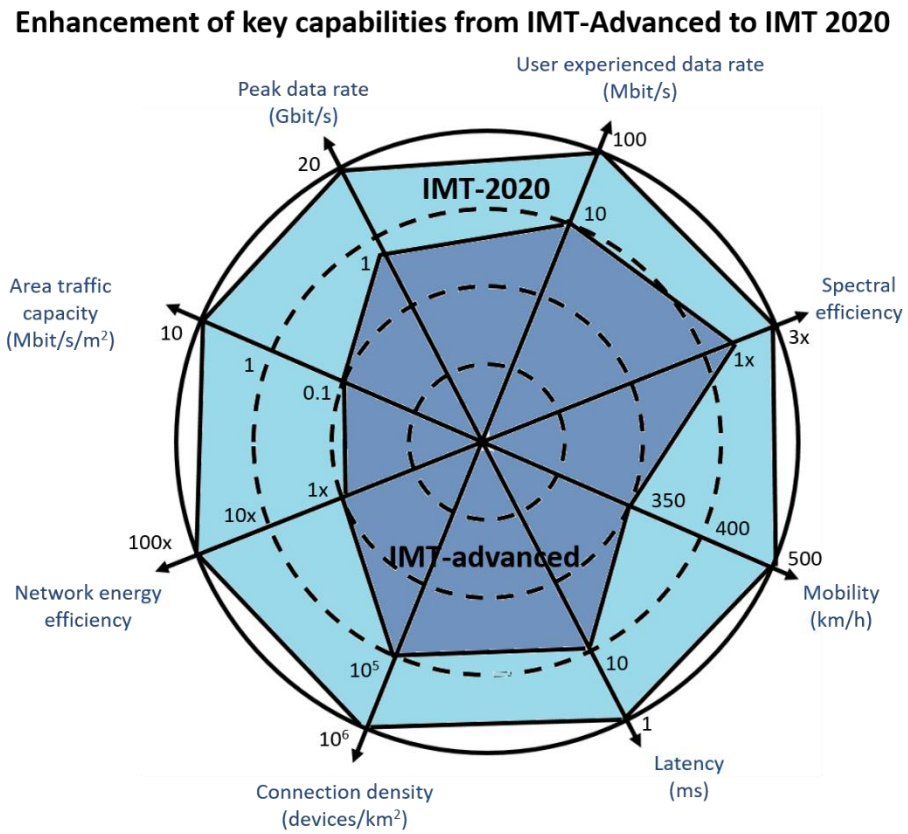
(On leave from Indian Institute of Technology Bombay, Mumbai, India)

director@iitk.ac.in

karandi@iitk.ac.in

Introduction to 5G

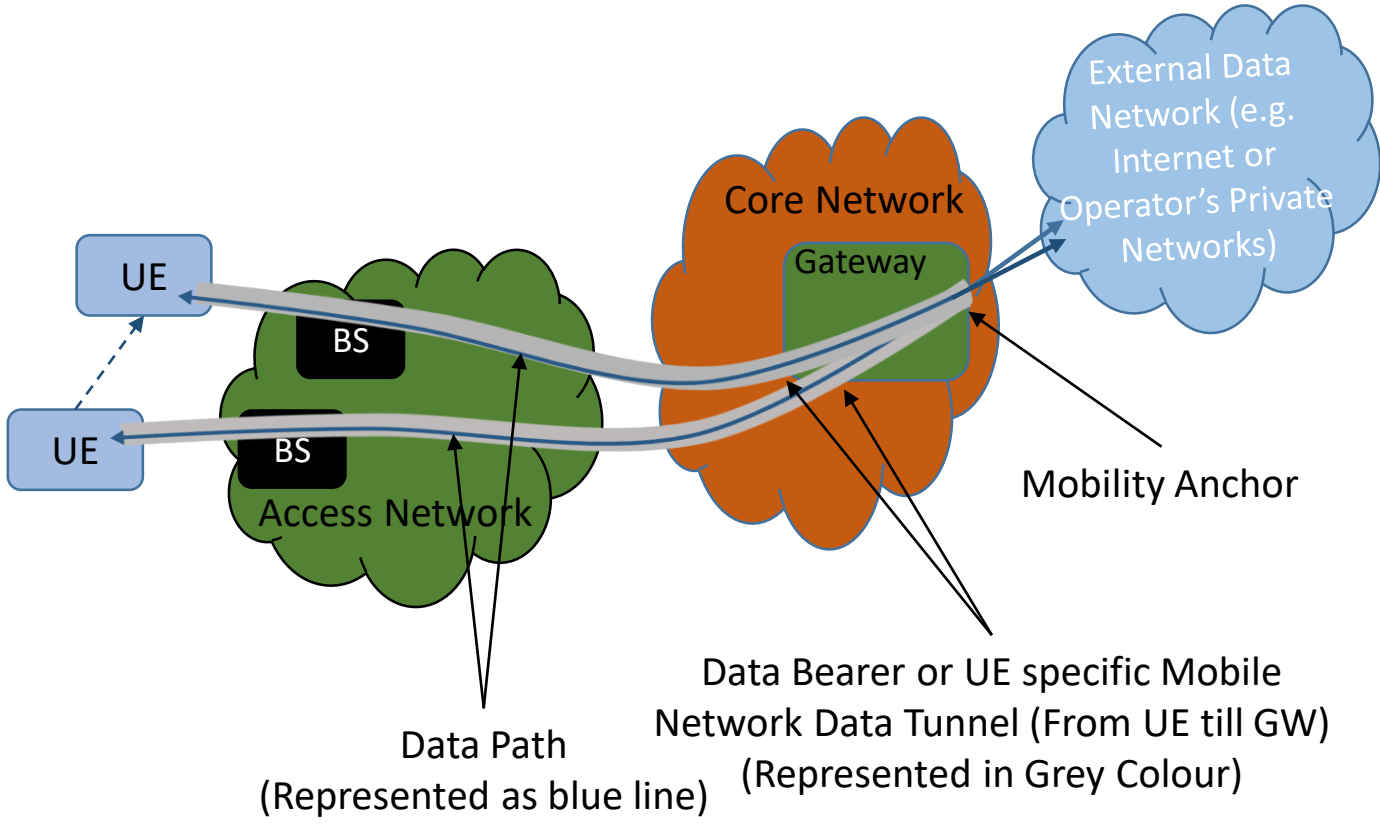
- Next generation of Mobile Communication System- IMT 2020
  - Enhanced capabilities over 4G
- Enhanced Mobile Broadband (eMBB)
  - Enhanced Mobility Support – 500 Km/h
  - Very High Peak Data Rate - 20 Gbps
  - High Spectral Efficiency – 30 bps/Hz
- Massive Machine-to-Machine Communication (MMTC)
  - Large no of devices in a small area – $10^6/km^2$
- Ultra Reliable Low Latency Communication (URLLC)
  - Extremely low Latency - ~ 1 ms latency over the air



Courtesy – International Telecommunication Union

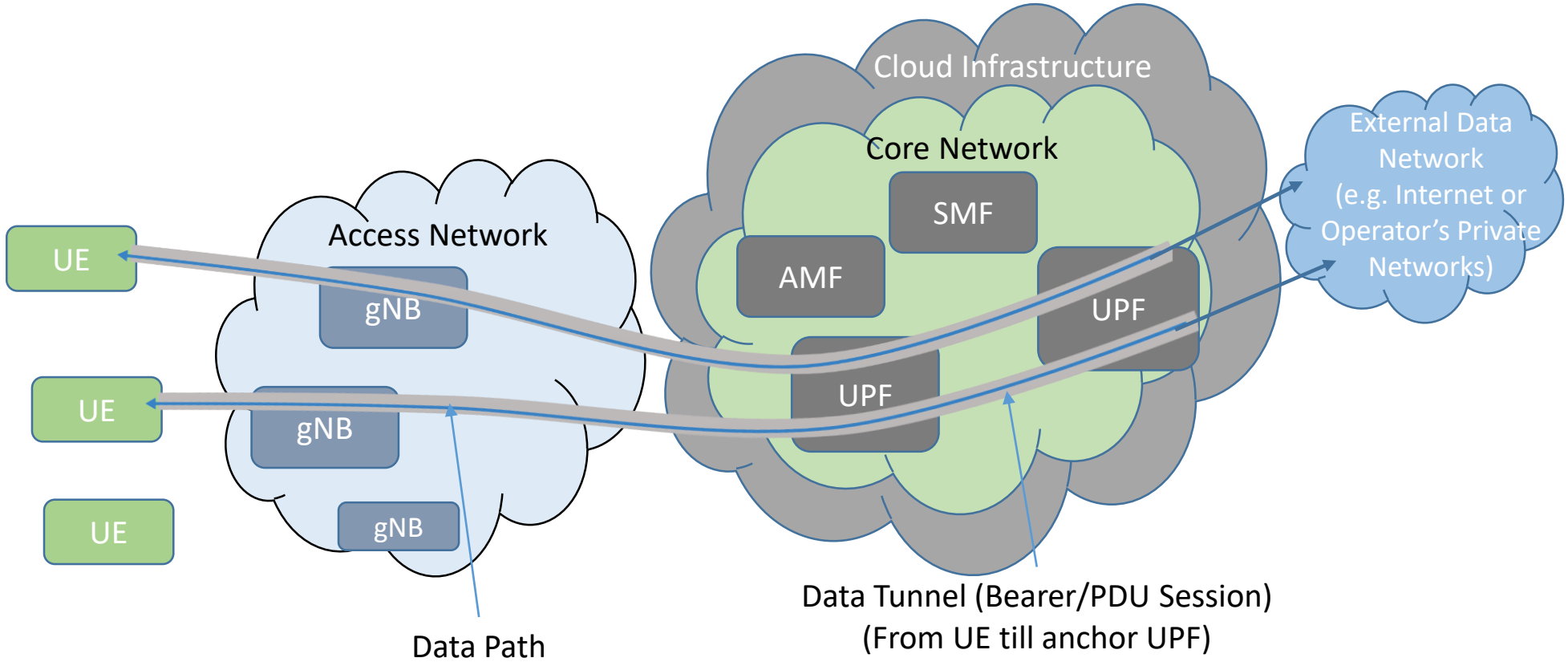
Typical Mobile Network Architecture

- Two Major Parts
  - Distributed Access Network
  - A Centralized Core Network
- Mobility- a key aspect
  - How to handle Mobility?
    - Data Tunnel between UE & Gateway in CN
    - CN Gateway - “Mobility Anchor” for UE
  - When a UE moves
    - Its point of attachment (BS) in the network may change
    - But its “Mobility Anchor” does not change
  - Packets from/to external network routed through the same Gateway for a UE even while it moves



5G Mobile Network Architecture

- Similar to the one Shown in the Previous slide
- Comprising of a Distributed Access & a Centralized Core Network
- Usage of Data Tunnel to Manage Mobility
- UPF - Mobility Anchor in the Core



Mobile Core typically deployed using a centralized cloud based infrastructure

Characteristics of Emerging Mobile Communication (1/2)

- Massive IoT
  - A key 5G use case
- Use Cases fuelling the growth of IoT
  - Smart Homes
  - Smart Cities & Villages
  - Smart Workplaces/Factories
  - Increased Automation Everywhere
- Most IoT Devices use Wireless Connectivity
  - Wi-Fi, Bluetooth, 4G and (5G in near future)
  - But Mobility not Important
  - Most IoT Devices - Stationary

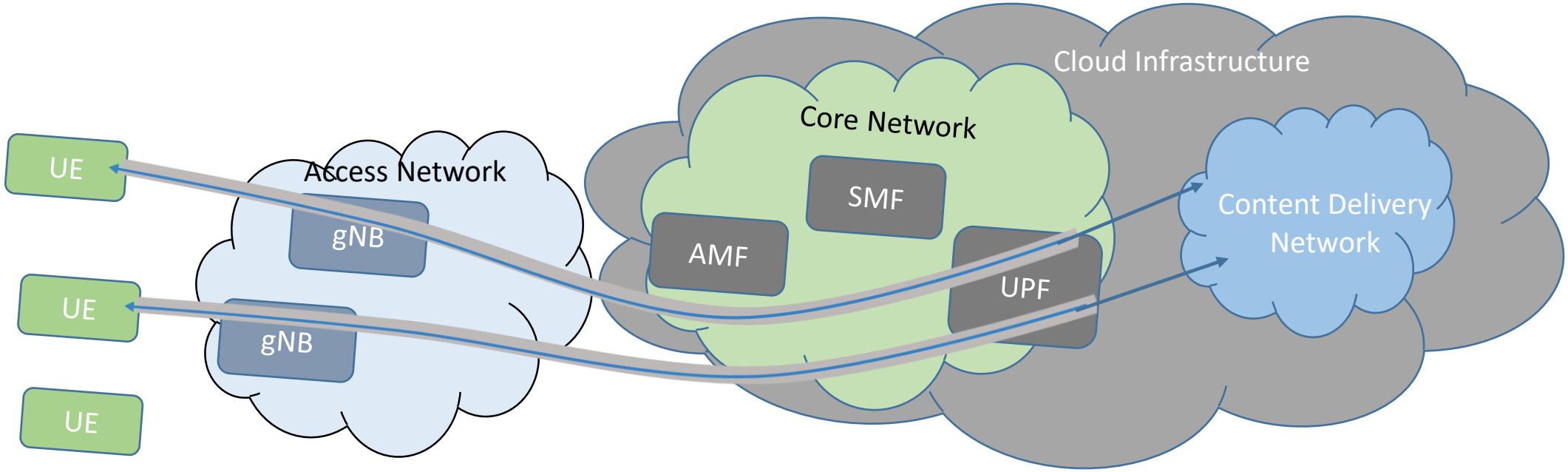
Characteristics of Emerging Mobile Communication (2/2)

- Increased Importance of Low Latency Communication
  - Mission Critical Communication
  - Industrial Control Systems
  - Mobile Health Care
  - V2X Communication etc.
- Need to Reduce Cost & Resource Consumption

Cloud Computing

- Remote Delivery of Compute/Storage Services (say over Internet)
  - Computation
    - Program Execution, Analytics, Intelligence etc.
  - Storage
    - Database etc.
- Virtually Unlimited Storage Capacity and Processing Power
- Scalability
- Business Continuity
  - Location Independence - Work from Anywhere/Anytime
- Economies of Scale & Cost Efficiency

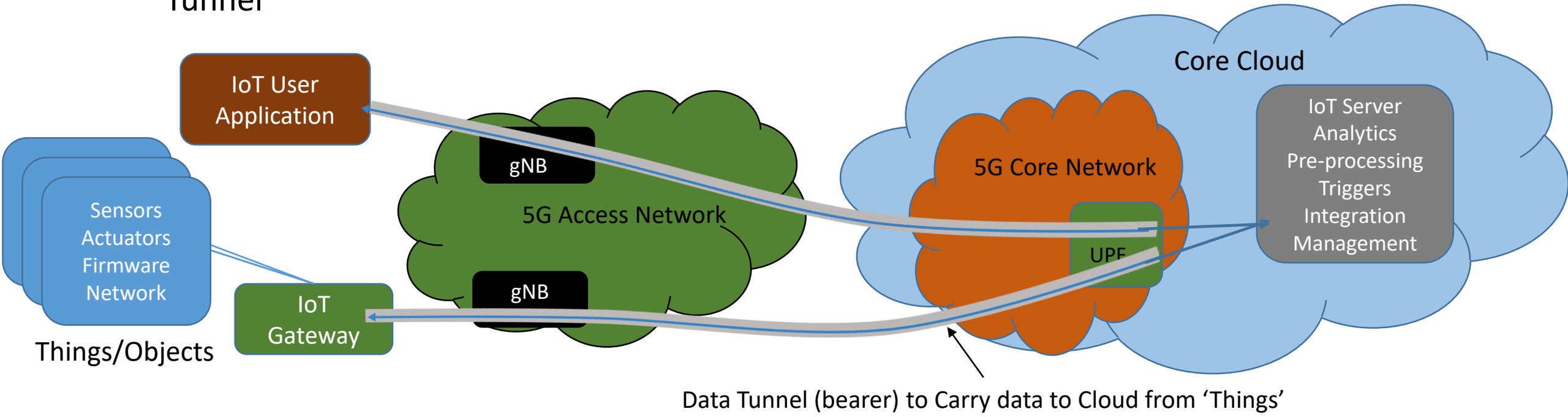
Cloud & 5G Use case – Content (Video) Delivery



- Content Delivery Network (CDN)
  - Delivery of Content to Users via Mobile Network
- CDN along with Mobile Core typically a part of Cloud based Infrastructure
- UE specific Data tunnel- CDN can only exist beyond the CN, i.e., beyond the UPFs (Gateways/Mobility Anchors)

IoT Integration with 5G – An Architecture

- IoT Server placed in Cloud along with 5G Core
- IoT Gateway acts as a UE
  - Data Tunnel via 5G Network to 'IoT Server…' in Cloud
  - Exchanges Information with IoT Server via Data Tunnel
- IoT User Application also acts as a UE
  - Separate Data Tunnel via 5G Network to 'IoT Server…'
  - Controls 'Things' via Cloud (via IoT Server)

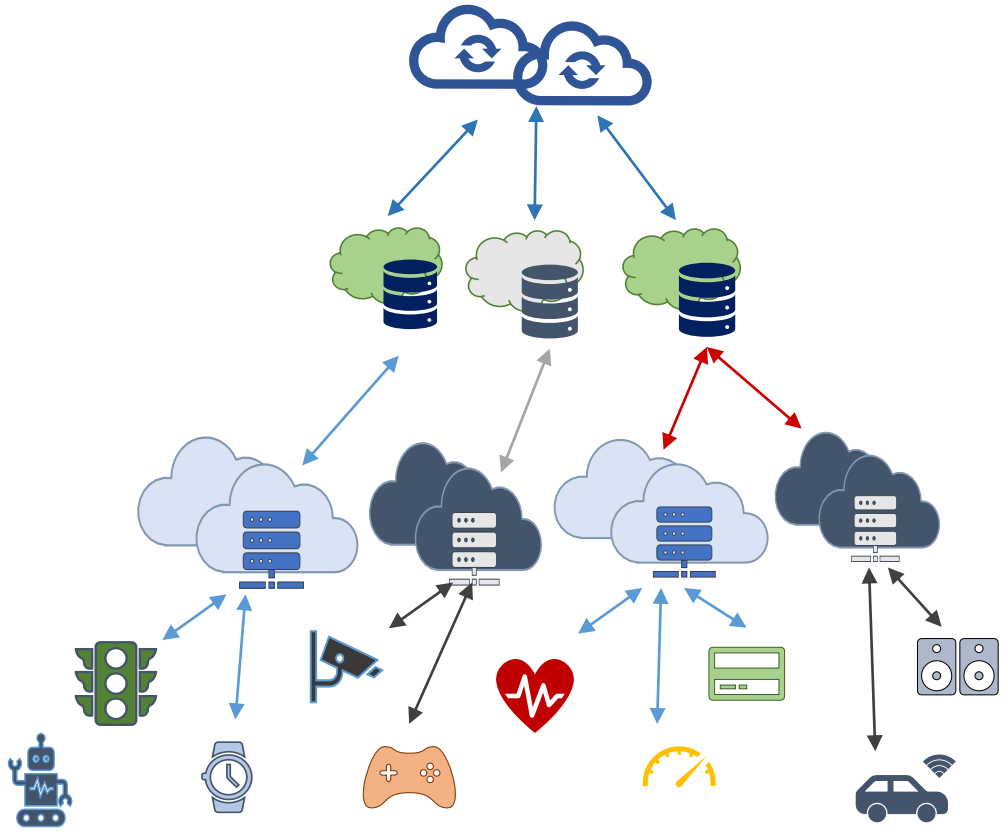


IoT with a Centralized Cloud via 5G - Issues

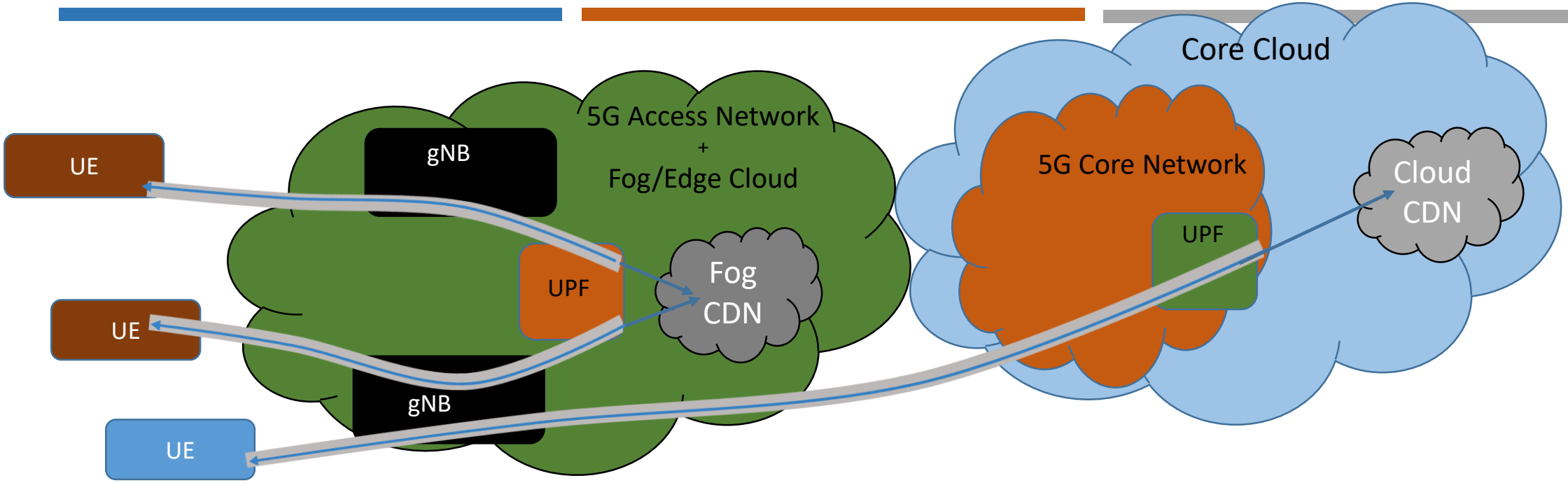
- Limitations of Cloud Computing
  - ‘Low Latency Communication’ Difficult to Achieve
  - Increased Resource Usage
- Push the Computing near ‘Things’ (Edge)
  - Computation/Storage near ‘Things’
  - Shorter Communication Paths
    - Improved Time Responsiveness
    - Reduced Resource Utilization

What is Fog Computing?

- Fog - **F**rom c**O**re to ed**G**e
- Support for Lower Latency Applications
- Location Awareness
- Reduced Network Bandwidth Usage
- However, there is an Issue
  - Can't handle Mobility of Devices
  - Need to form Continuum with Cloud
    - Fog for Stationary Devices, Cloud for Mobile ones
  - Fog and Cloud Complement each other



Fog/Edge Computing and 5G - CDN



- Existence of a Content Delivery Network (CDN) in Fog as well as Cloud
- Mobile UEs served from Cloud CDN
- Stationary UEs served from Fog CDN

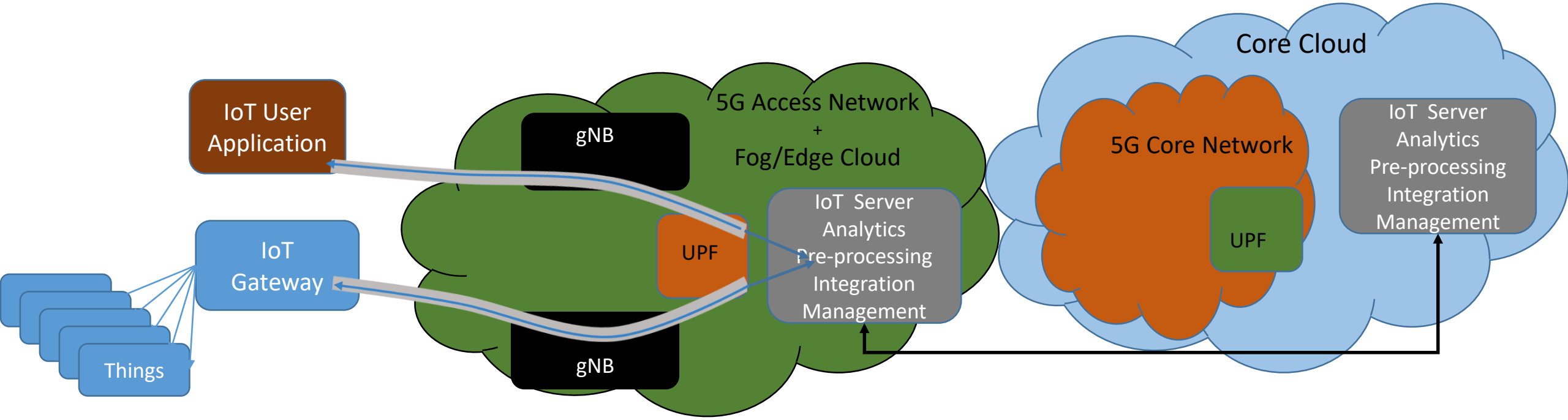
UE Stationary UE

UE Mobile UE

Fog/Edge Computing and 5G - IoT Use case

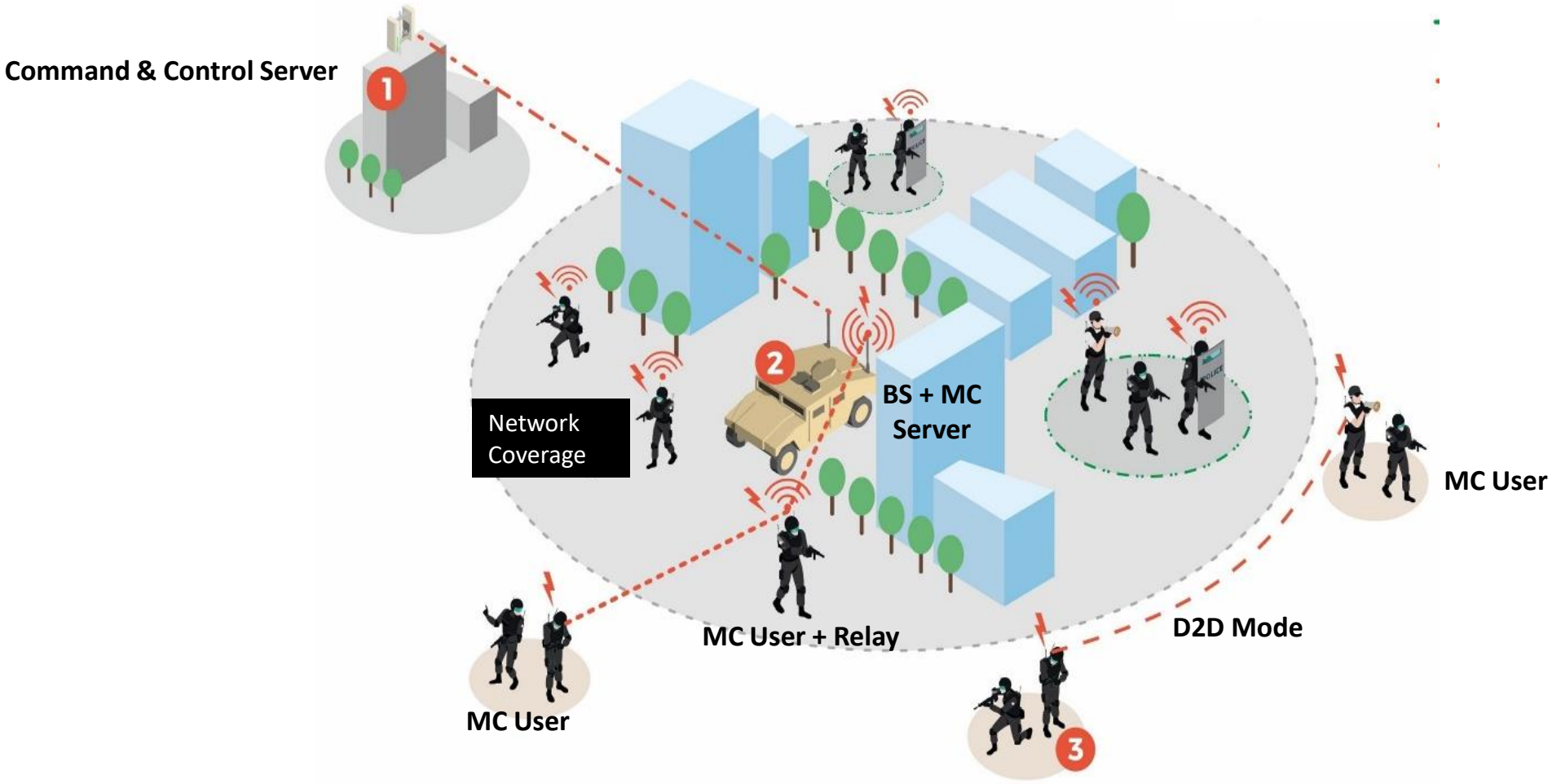
- Fog Based UPF Placement
- Placed in Access Network (Edge) Cloud

- IoT Server Placed in Fog as well as Cloud
- For Stationary Devices, Use Fog based Server
- For Mobile Devices use Cloud based Server – *Not shown in the Figure*

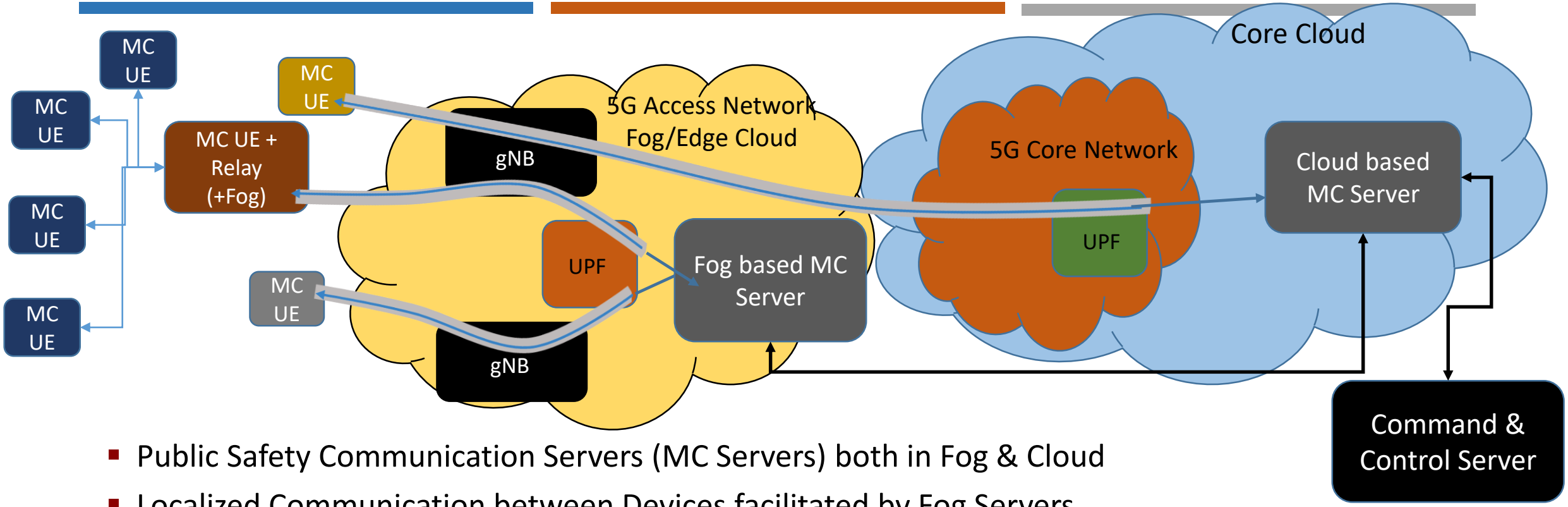


Public Safety Communication

- A Rapidly Deployable Mission Critical System



Fog Computing & 5G – Public Safety Communication



- Public Safety Communication Servers (MC Servers) both in Fog & Cloud
- Localized Communication between Devices facilitated by Fog Servers
  - Relay based Fog Element between a set of UEs
  - Fog Element in the vicinity of gNB too (in Access Fog)
- UE can communicate via Core Cloud also, if needed
- Command & Control Server Communication directly to Cloud (can use 5G network also for comm)

Frugal 5G Networks (IEEE P2061)

Frugal 5G Networks (IEEE P2061)

Refers to the vision of providing broadband access to rural areas by addressing these requirements and challenges

Source: Khaturia M, Jha P and Karandikar A, "Connecting the Unconnected: Towards Frugal 5G Network Architecture and Standardization", IEEE Communication Standards Magazine, June 2020.

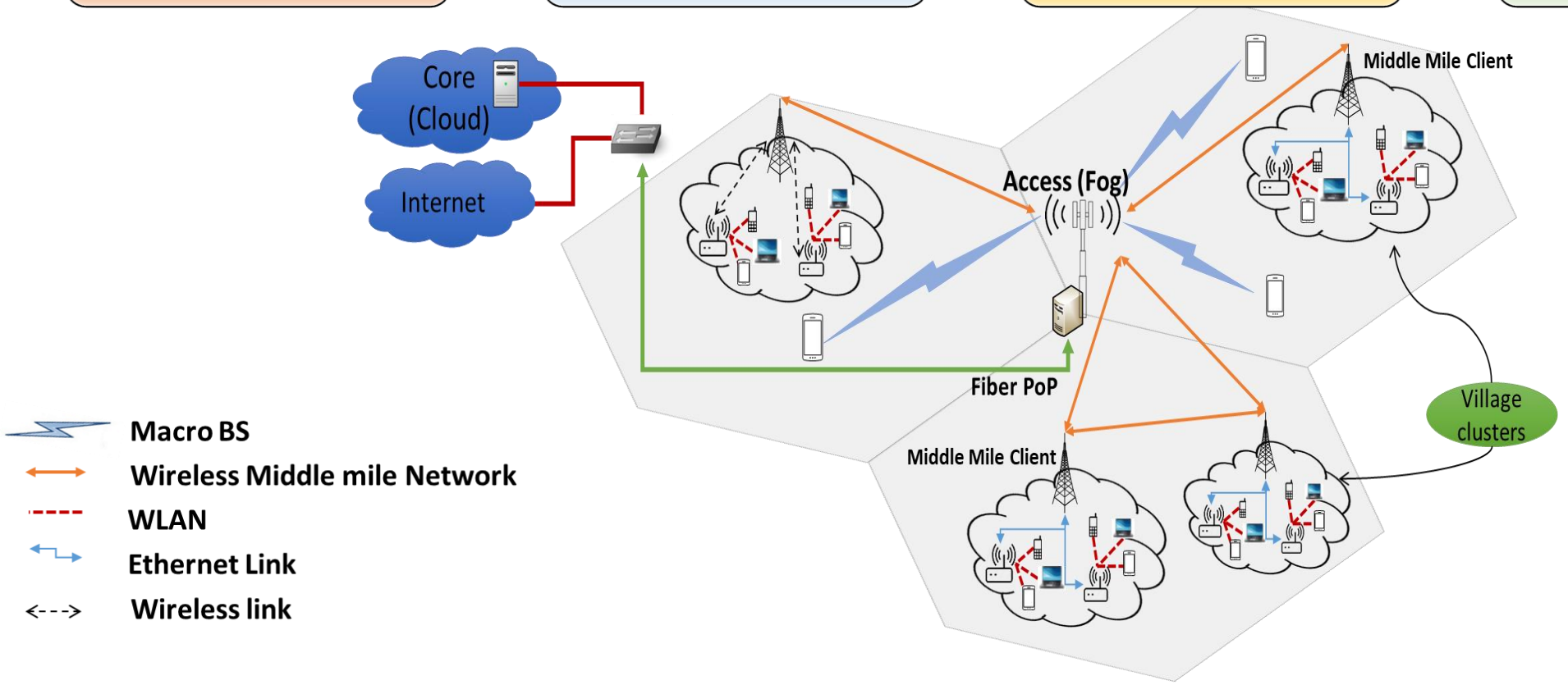
Network Architecture - Features

Large Coverage Area Cells to provide ubiquitous connectivity

Small Cells (WiFi Hotspots) as high speed access points

Wireless Middle Mile Network to backhaul data

Point to point wireless links to connect the nodes in villages



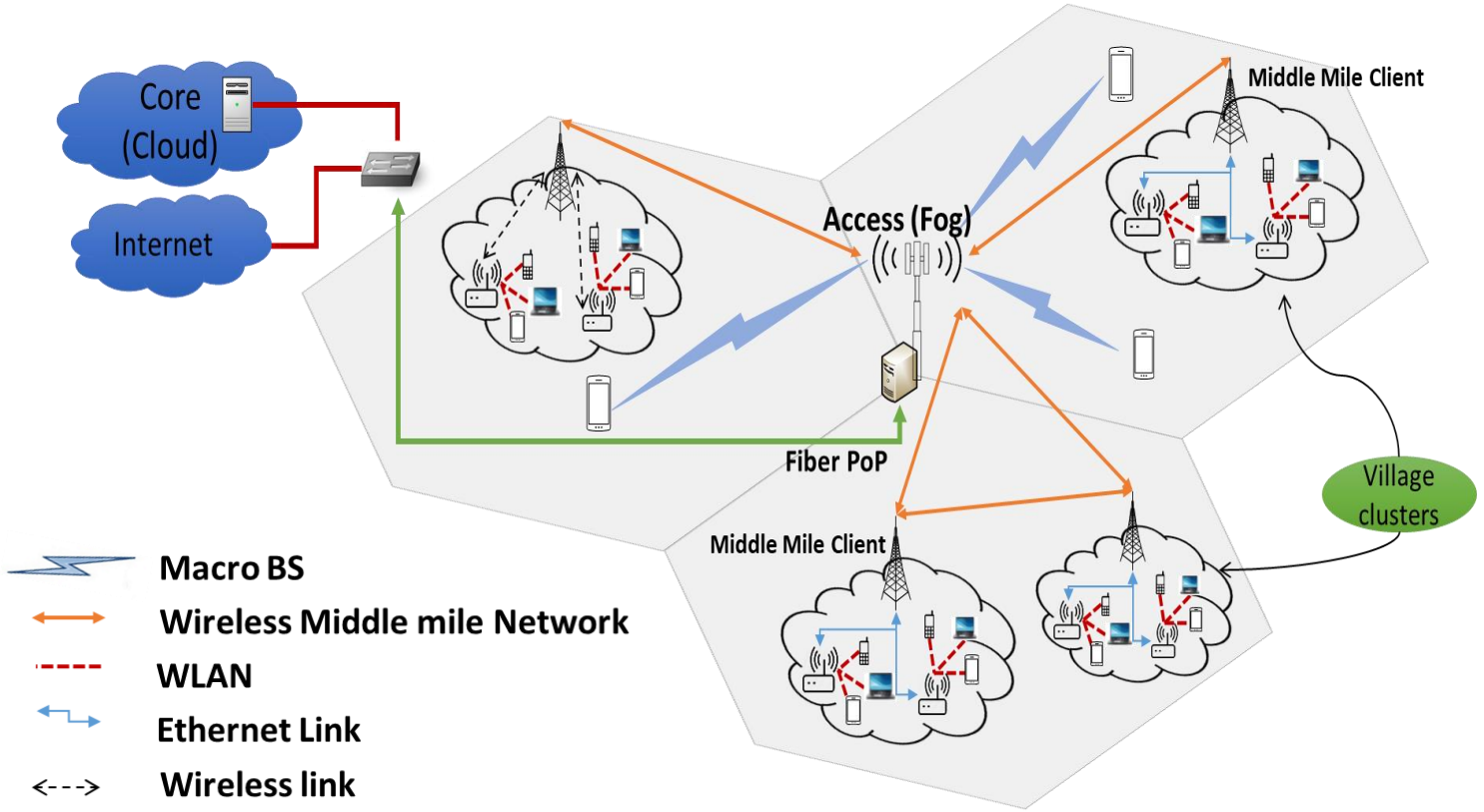
Source: Khaturia M, Jha P and Karandikar A , "Connecting the Unconnected: Towards Frugal 5G Network Architecture and Standardization", IEEE Communication Standards Magazine, June 2020.

IEEE Network Architecture- Features (contd..)

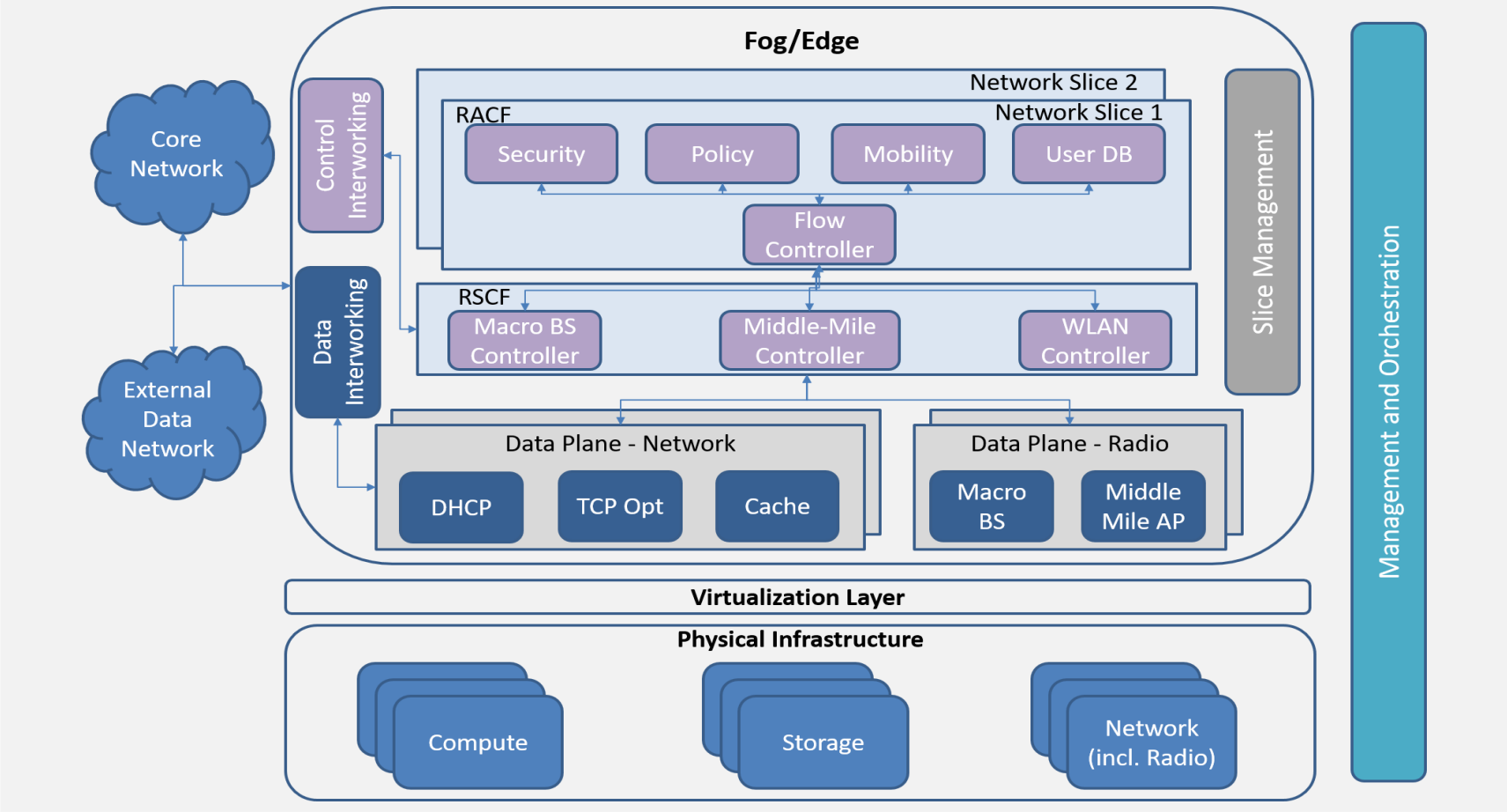
SDN based unified control - Efficient service delivery, Independent Evolution and Development of control/data plane entities

Usage of Virtual Network Functions - makes the system cost-effective

Intelligence at the edge - Enables local communication & reduces resource usage



Architecture – Fog/Edge Components



Hierarchical SDN based control of the multi-RAT network

RAT Specific Control Functions:

- Provides an abstract view of underlying RAN to higher level control entities;
- Enables unified control of multi-RAT network;
- Enables RAN virtualization and Network Slicing

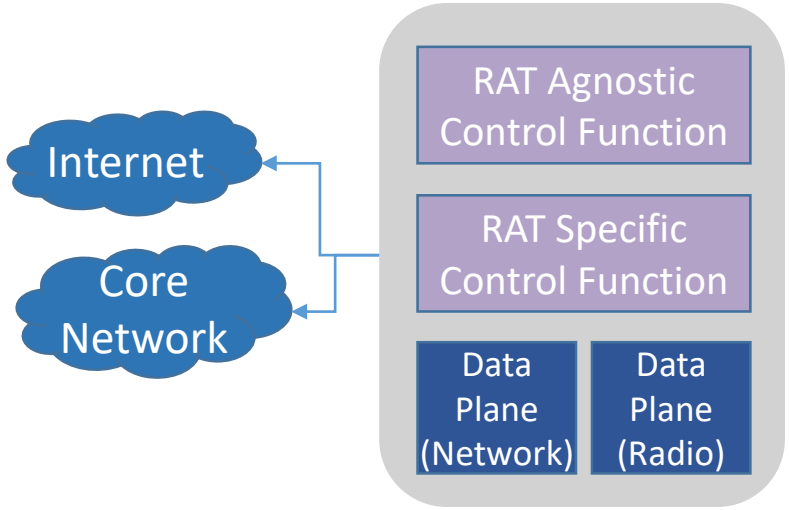
RAT Agnostic Control Functions:

- Operates over abstract resources provided by RSCFs;
- Analyses individual traffic flows and acts on it with help from other RACFs;
- Enables localized communication under fog element

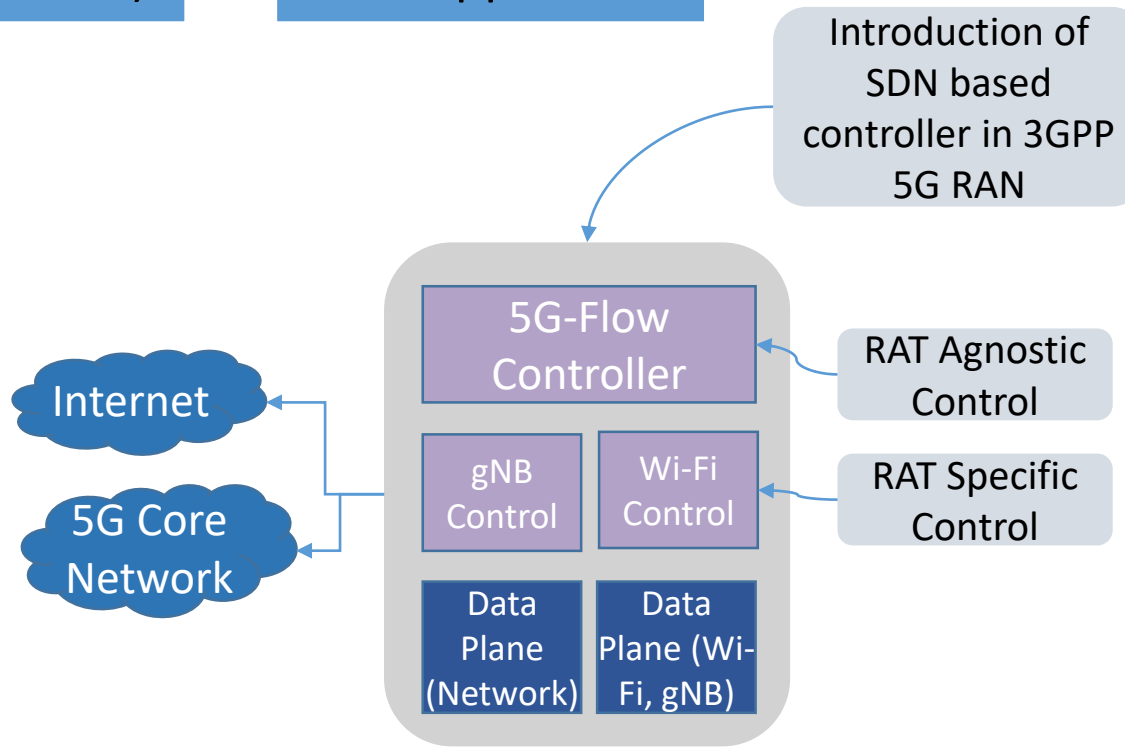
5G-Flow: Realizing Frugal-5G Architecture using 3GPP 5G

Desired Network Characteristics

- SDN Based Unified Access Control at Edge
- Direct Internet Connectivity from RAN (w/o Core)
- Local Communication Support



Frugal 5G Network (Edge Component)

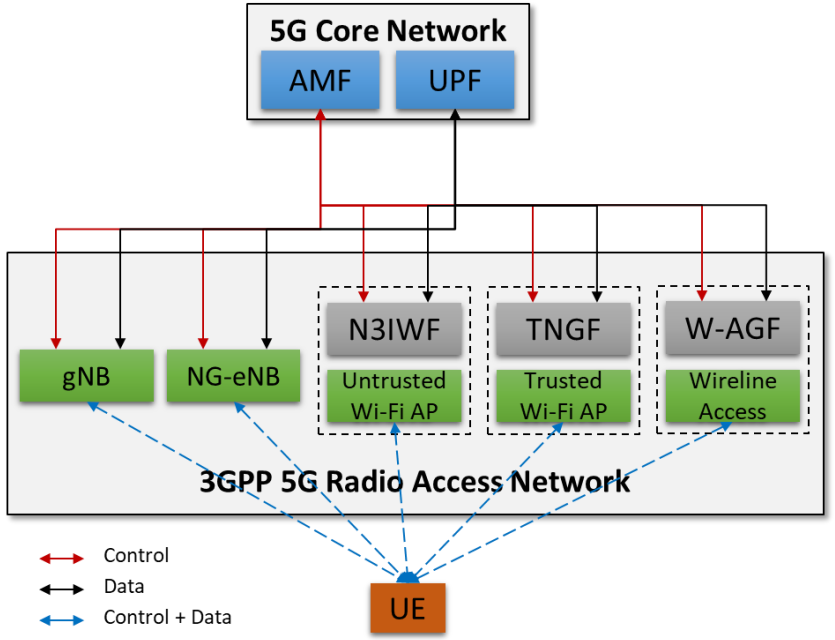


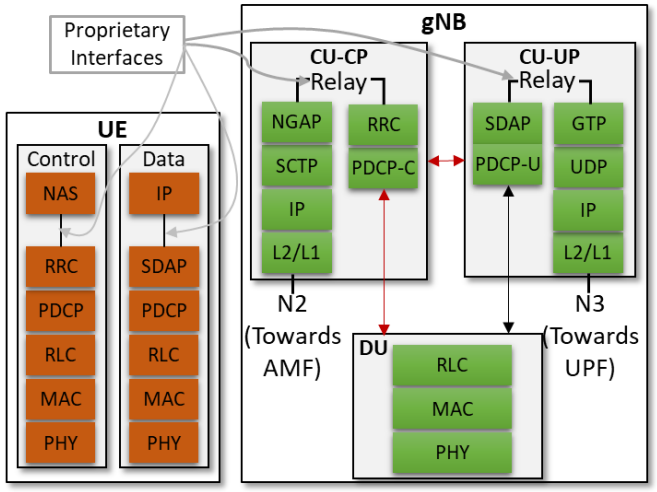
5G-Flow: Implementing Frugal 5G concepts at 3GPP 5G RAN

Source: Khaturia M, Jha P and Karandikar A, "5G-Flow: Flexible and Efficient 5G RAN Architecture Using OpenFlow." arXiv preprint arXiv:2010.07528 (2020).

Existing 3GPP 5G Architecture - Limitations

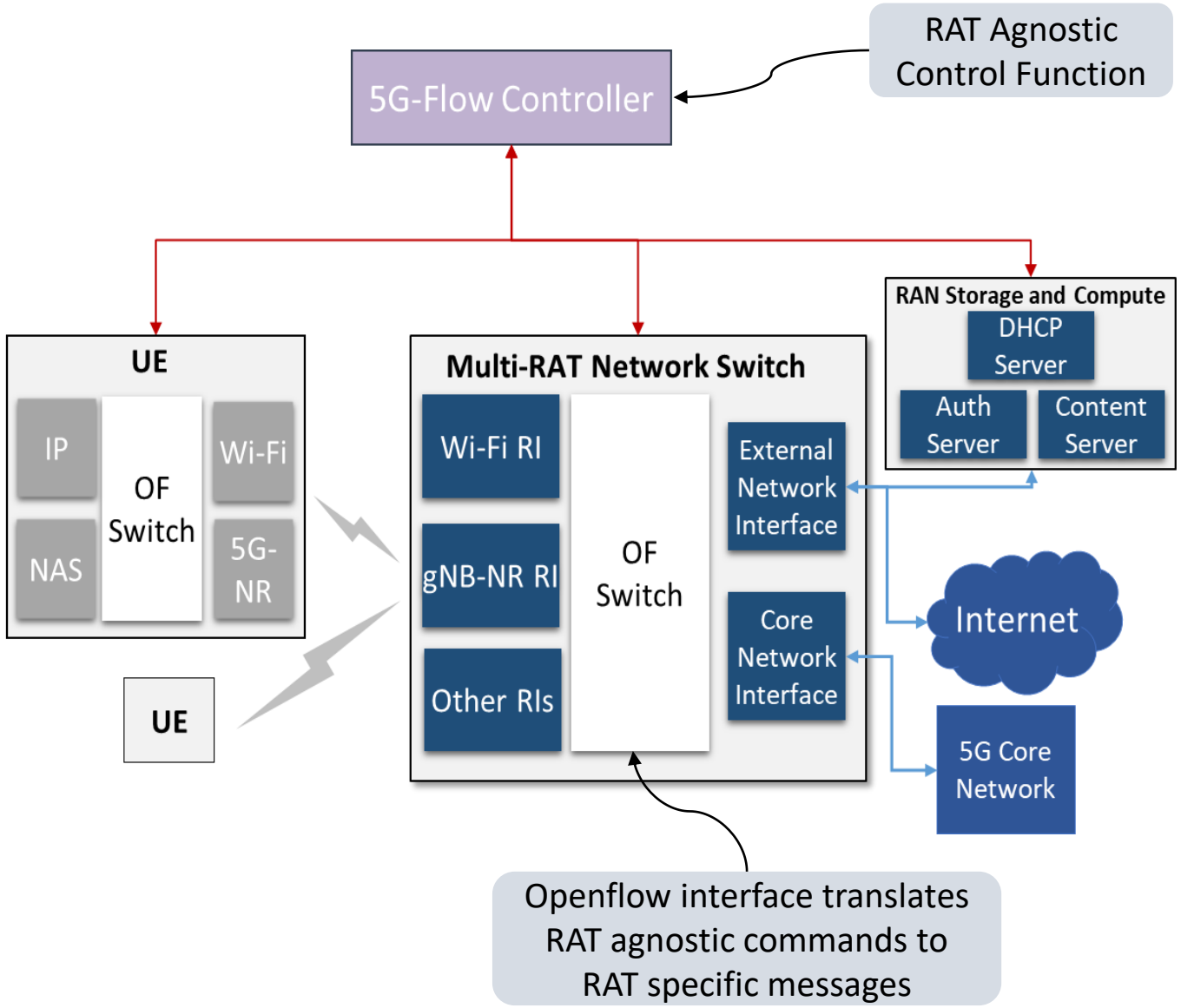
- Fragmented Decision Making in RAN
  - gNB, eNB, Wi-Fi APs …
  - Controlled and Managed Separately
- Unified Core but RAT Specific Inter-working functions
  - gNB, eNB, N3IWF, TNGF, W-AGF
  - Management Overhead
  - Non Optimal Multi-RAT Access
- Tight & Proprietary Coupling between Radio & CN Protocol Stacks in RAN
  - Leads to RAT Specific CN Interworking Function
  - Loss of Flexibility - Not possible to Connect 5G RAN to 4G Core
- Concurrent Multi-RAT Access for UE
  - Managed @Core - Access Traffic Steering, Switching & Splitting
  - Optimal Management of Multi-RAT Access Not Possible
    - RAN level information absent at Core





5G-Flow – Unified Multi-RAT RAN

- Logically Centralized Multi-RAT RAN Control
  - Light-weight OF (5G-Flow) Controller for Unified Control
- Decoupled Protocol Stacks at RAN Nodes and UE
  - CN and Radio Interface Stacks Decoupled
- OF-Switch based Unified Multi-RAT RAN Data Plane
  - Protocol Stacks used as Interfaces of an OF Switch
  - Even NAS Signaling Exchange treated as data passing through an OF-Switch



Summary

- Fog/Edge Computing together with 5G
  - Enables Many Use Cases
- Fog/Edge Computing
  - Easy to Integrate with the 5G Network
  - Lower Latency – Better Performance
  - Improved Resource Utilization
  - Addresses Certain Limitations of Cloud
- User/Device Mobility a key factor in the usage of Fog in 5G Network

THANK YOU